Essays on Labor Market Policy

Inaugural-Dissertation

zur Erlangung des Grades eines Doktors
der Wirtschafts- und Gesellschaftswissenschaften

durch

die Rechts- und Staatswissenschaftliche Fakultät der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Gero Stiepelmann

aus Bonn

Bonn

2025

Dekan:	Prof. Dr. Martin Böse
Erstreferent:	Prof. Dr. Moritz Kuhn
Zweitreferent:	Prof. Dr. Keith Küster
Tag der mündlichen Prüfung:	01. September 2025

Acknowledgements

I would like to express my sincere gratitude to my supervisors, Moritz Kuhn, Keith Kuester, and Christian Bayer. I am especially thankful to Moritz Kuhn, who continuously encouraged me to develop my own research ideas, present my work at international conferences, pursue a research stay abroad, and actively engage with the academic community. Our research discussions were always deeply insightful and motivating. Special thanks also go to Keith Kuester, who has supported my academic journey since my bachelor's studies. His guidance and advice have been invaluable throughout the years. Finally, I want to thank Christian Bayer for his insightful comments

I am also very grateful to the faculty and students at Yale University, who were always open and generous in discussing research. I am particularly thankful to Ilse Lindenlaub for hosting me at Yale and for her thoughtful advice on my job market paper. I also wish to thank Giuseppe Moscarini for the in-depth discussions on my research, and Martin Souchier and Michael Peters for their invaluable input in shaping and developing the third chapter of this dissertation.

I would like to extend my thanks to the Bonn Graduate School of Economics (BGSE), and especially to Silke Kinzig, for supporting my research stay abroad. I am also grateful to Kim Schneider for her outstanding assistance during the job market process. Further, I gratefully acknowledge financial support from the BGSE, the Research Training Group 2281, the Institute of Macroeconomics and Econometrics, and the Yale Graduate School of Arts and Sciences.

A heartfelt thank you goes to my fellow students and friends at BGSE for the many stimulating research discussions and their unwavering support, especially during the more challenging phases of the dissertation. This project would not have been possible without you. I am deeply thankful to my co-author Johannes Weber, as well as to Christina Brinkman and Leanne Nam, for their exceptional support throughout the job market period. Special thanks also go to Sophia Friedrich for her wholehearted support during the final months of my PhD.

Last but not least, I want to thank my parents, Heiko Stiepelmann and Doris Witteler-Stiepelmann, for their unconditional support throughout my entire academic journey and life in general. Their unwavering encouragement was a vital source of strength during the challenges of my PhD.

Contents

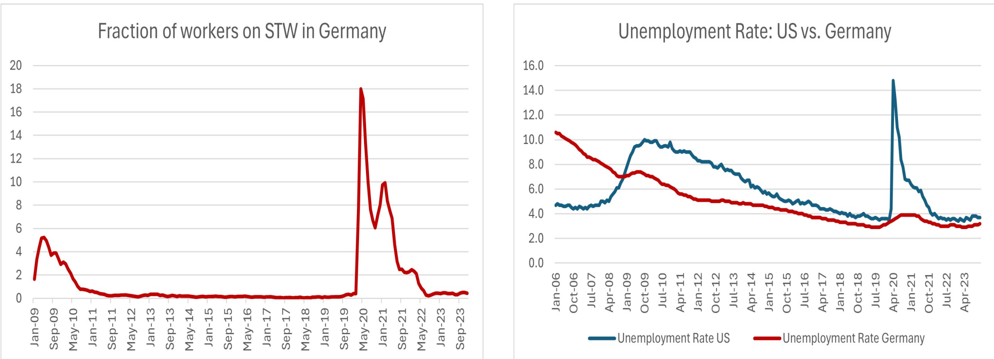
List of Figures

List of Tables

Introduction

The Great Recession and the COVID-19 pandemic have renewed interest in labor market policy instruments designed to protect workers from job and income losses. Among these, short-time work (STW) has emerged as the key policy tool to prevent separations in Germany and across Europe. It functions as a subsidy scheme that provides partial wage compensation to employees when firms sufficiently reduce working hours.

Figure 0.0.1. STW Take-Up and Unemployment Rates



(a) (b)

Notes: Plot (a) depicts the fraction of workers that have been on the STW system. The data for participation in the Short-Time Work System is taken from the Bundesagentur fuer Arbeit. Plot (b) depicts unemployment rates in Germany and the US over time. The unemployment rate for Germany is taken from the Federal Statistical Office of Germany. The unemployment rate for the US is taken from the U.S. Bureau of Labor Statistics.

The scale and economic significance of the STW system is most clearly illustrated by Germany's reaction to the Great Recession and the COVID-19 pandemic. In response to both crises, Germany significantly expanded the generosity and accessibility of its STW program, resulting in a substantial increase in worker participation. Figure 0.0.1 (a) depicts the extensive use of STW: at its peak, approximately 5.5% of the workforce was covered during the Great Recession, and around 18%

during the COVID-19 crisis. This widespread uptake likely contributed to stabilizing employment levels and preventing large-scale layoffs. In contrast, the United States, which did not employ comparable short-time work schemes, experienced sharp spikes in unemployment. As shown in Figure 0.0.1 (b), U.S. unemployment rates rose dramatically during both crises, while Germany maintained relatively stable levels.

However, despite the prevalence and significance of the STW system, there is limited theoretical work on the foundations and optimal design of STW policies. This dissertation seeks to fill that gap by developing a theoretical framework for evaluating and designing optimal STW policy. The analysis is grounded in a search-and-matching model of the labor market featuring Diamond-Mortensen-Pissarides-type search frictions, endogenous separations, flexible hours, and generalized Nash bargaining. The model is calibrated to the US economy.

The dissertation consists of three chapters. The first chapter addresses the fundamental question of why STW should be used in the first place. It views STW as a policy tool to mitigate socially inefficient separations that arise under a standard unemployment insurance (UI) system. Observing that European countries frequently adjusted their STW systems during recessions, the chapter proceeds to examine how STW should be optimally adjusted during a recession.

The second chapter focuses on a central critique of STW: by retaining workers in distressed firms, it may hinder labor reallocation and cause misallocation. This concern was notably raised by Cooper, Meyer, and Schott (2017), who argue that STW contributed to Germany's sluggish post-recession recovery relative to the U.S. . Therefore, the chapter investigates whether an optimally calibrated STW system induces labor misallocation and, if so, how government policy should respond to mitigate its effects.

The final chapter (joint with Johannes Weber) addresses the critique that there might be more effective instruments than STW to mitigate the fiscal externalities created by the UI system. While STW is widely used in Europe, the United States relies on a different approach. It has experience-rated its UI system. Firms that layoff more workers face higher UI contribution rates. Naturally, the question arises which of the two systems is the superior policy tool.

Reason to use STW. Unemployment insurance plays a vital role in protecting workers from income loss during unemployment. However, it also introduces distortions that depress job-finding rates and increase separation rates. Chapter 1 shows that STW can help to reduce inefficient separations. The key intuition is simple: while UI offers insurance after a separation, STW allows firms to avoid separation in the first place. By supporting firms in reducing working hours, STW helps lower their wage bills and thus reduces incentives for separation. At the same time, it keeps workers attached to the firm by partially compensating them for the lost income. Optimal STW benefits are set to target the cost a worker would impose on the government if they became unemployed and entered the UI system. Compared to broad-based employment subsidies, STW offers one notable advantage and one important drawback. Its key strength lies in its built-in screening mechanism: by conditioning support on a reduction in working hours below a certain threshold, STW naturally targets firms facing temporary productivity shortfalls—those most at risk of layoffs. This stands in contrast to a broad-based subsidy program (like the Paycheck Protection Program (PPP) in the U.S.), which has difficulty distinguishing between firms that genuinely required support and those that did not.

The major disadvantage, however, lies also in the mechanism's reliance on hours reduction, which may distort firm-level decisions. Firms and workers inefficiently reduce working hours to draw in more support from the government, introducing a new layer of inefficiency into the economy. Additionally, STW is no instrument that can be used to stabilize the job-finding rate.

The preceding analysis provides the foundation for evaluating the role of STW under varying economic conditions.

STW in Recessions. During both the Great Recession and COVID-19 pandemic, governments expanded STW systems—both in terms of generosity and eligibility. This raises the natural question: How should STW be optimally adjusted during recessions? The second part of chapter 1 finds that increasing STW benefit generosity in downturns is indeed optimal, aligning with recent policy practice. However, contrary to policy practice, the eligibility criteria should be tightened, not relaxed. Quantitatively, I find that in response to a 1% negative productivity shock, optimal STW benefits should rise by 12.5%, while firms must reduce working hours by an additional 6% to qualify for STW.

The rationale is as follows. Recessions typically involve sharp declines in job-finding rates, which increase the expected duration of unemployment spells. This makes separations more costly from a social perspective. As a result, the government wants to raise STW benefits to strengthen incentives for firms to retain workers. Yet, more generous benefits also increase the risk that firms that can survive without STW support enter the system. To mitigate these windfall effects, stricter eligibility rules (e.g., higher required reduction in working hours) are necessary.

STW and Labor Misallocation. A frequent concern with STW is that it impedes efficient worker reallocation by locking workers into unproductive firms. To study this, the second chapter introduces firm heterogeneity and information asymmetries into the model. Specifically, it distinguishes between firms with temporary and permanent productivity shocks. Temporary unproductive firms are likely to recover, while permanently unproductive ones are less.

The government, however, cannot observe a firm's type. This informational friction forces the government to balance two goals: preventing inefficient separations from temporarily distressed firms and promoting reallocation away from permanently unproductive ones. The model shows that the optimal policy would like to provide high subsidies to firms likely to recover and low subsidies to firms unlikely to do so. When firm types are unobservable, the government faces a trade-off: it can provide high STW benefits to support temporarily unproductive firms but at the cost of retaining workers in permanently inefficient firms — or vice versa. The optimal benefits rely on the distribution of firm types in the economy.

Surprisingly, I find that setting STW optimally almost completely eliminates the cost of misallocation. Furthermore, if the government can observe separation intentions (e.g., the government usually requires access to the financial records of the firm), it can improve targeting by letting STW benefits decline over time. As time passes, the pool of firms remaining on STW shifts toward those with more persistent negative shocks which motivates the government to reduce benefits to foster reallocation. If direct observation of separation intentions is not possible, eligibility based on hours reduction can screen out firms with permanent productivity shocks and makes a degressive STW system superfluous.

The chapter also offers first insights into how STW should respond to structural change. Structural shifts—such as sectoral decline or technological displacement—lead to an increase in firms with persistent productivity shocks. In such cases, the optimal policy response is the opposite of a recession response: If the shift in structural change is sufficiently large, STW benefits should be reduced to encourage labor reallocation. At the same time, eligibility criteria should be relaxed to extend support to the now larger mass of firms facing permanent shocks. Therefore, it is crucial to distinguish between recessionary shocks and structural change when designing STW policy.

STW or Lay-off Taxes. The final chapter shifts focus and examines whether there exist more effective systems than STW to combat the fiscal externality of the UI system. In Europe, STW and UI are typically combined. In contrast, the U.S. relies on experience-rated UI, where firms contribution to the UI system rise when laying off a worker which I model as lay-off tax. This raises a key policy question: Which system is superior?

The answer depends crucially on the prevalence of financial constraints. When financial constraints are rare, layoff taxes are preferable. They can efficiently deter inefficient separations without distorting firms' hours decisions, achieving the socially optimal separation rate.

However, in economies with many financially constrained firms, STW performs better. Layoff taxes are less effective at deterring separations when firms lack liquidity, while STW still enables firms to reduce labor costs without firing workers. Additionally, STW provides partial income insurance even when separation is avoided—something UI and lay-off taxes do not do.

The quantitative analysis shows that when more than 40% of firms are financially constrained, STW outperforms experience-rated UI.

Conclusion. This dissertation demonstrates that STW can serve as an effective policy instrument for mitigating inefficient separations caused by a UI system, particularly in the presence of firm-level financial constraints. Conditioning eligibility on working-hour reductions enables screening for firms in genuine need of support—an issue often encountered in conventional subsidy programs. When optimally designed, STW can significantly reduce inefficient separations without inducing notable labor misallocation. Nonetheless, minimizing distortions to hours worked remains a persistent challenge. The effectiveness of STW ultimately hinges on the government's ability to enforce and calibrate eligibility criteria to limit windfall effects. Importantly, optimal policy adjustments may differ markedly between recessions and periods of structural change.

Chapter 1

Optimal Short-Time Work Policy in Recessions

1.1 Introduction

The Great Recession and the COVID-19 pandemic have reignited interest in fundamental questions of labor economics: How can we prevent unemployment, and how can we protect workers from income loss due to unemployment? To address these challenges, policymakers in Europe have primarily relied on a combination of two policy tools: Unemployment Insurance (UI) and Short-Time Work (STW). UI replaces part of a worker's wage when workers become unemployed, while STW offers partial wage compensation to employees when employers temporarily reduce working hours below a certain eligibility threshold. In recent crises, policymakers expanded the STW system, increasing both its generosity and accessibility, while largely leaving the UI system unchanged. This raises the question of how STW and UI should be used together, and why we should adjust STW over the business cycle.

Although extensive literature exists on optimal UI,[1] the optimal use of STW and its interaction with UI are less well understood. Burdett and Wright (1989) and Braun and Brügemann (2017) are the first to analyze the interaction of STW and UI in a static implicit contract model using numerical simulations. However, Cahuc (2024) argues that the interplay between UI and STW remains conceptually vague and that clarifying the optimal interplay is pivotal for formulating effective policy

1. Several studies, including Baily (1978), Shavell and Weiss (1979), Hopenhayn and Nicolini (1997), and Chetty (2006) explore the optimal design of UI systems with a focus on job search incentives. Landais, Michaillat, and Saez (2018) extend this by incorporating vacancy posting incentives. Blanchard and Tirole (2008) and Cahuc and Zylberberg (2008) discuss layoff taxes to reduce UI-induced separations. Jung and Kuester (2015) and Michau (2015) examine the optimal UI design, addressing these issues with vacancy subsidies and layoff taxes. This paper investigates optimal UI design where STW could target vacancy postings and separations.

recommendations. Further, despite its extensive use in recessions, there exists no theory yet on how to use STW optimally over the business cycle.

To address this gap in the literature, I develop a real business cycle model incorporating Mortensen and Pissarides (1994) type matching frictions in the labor market. The model accounts for risk-averse workers who cannot save, flexible working hours, and endogenous separations. Workers' contracts about income, working hours, and separations are determined within a generalized Nash-Bargaining framework. The UI and STW systems are chosen optimally and are financed by lump sum income taxes.

The paper makes two key contributions. First, the paper is the first to provide closed-form expressions for the optimal policy mix between UI and STW. These show that the UI system is responsible for providing income insurance to workers while STW internalizes the fiscal externality of the UI system on separations, a result supported by the implicit contract literature. Further, the expressions show that the government trades off employment stabilization against distorting working hours with STW. Second, I allow STW to adjust optimally over the business cycle. My findings indicate that, in line with actual policy, optimal STW policy requires increasing benefits in recessions. However, contrary to current practices, the eligibility condition should be tightened. In a quantitative exercise calibrated to the US economy, I find that optimal adjustment of STW is sizable. In response to a 1% negative productivity shock, optimal STW benefits should rise by 12.5%, while firms must reduce working hours by an additional 6% to qualify for STW.

Search frictions, an element not accounted for in the implicit contract literature, prove to be crucial to understand why STW should be adjusted over the business cycle. The expressions show that a decline in the job-finding rate increases the social costs of separations due to prolonged spells on the UI system, prompting optimal STW benefits to rise. Simultaneously, larger STW benefits increase the incentives for firms to enter the STW system. To keep the number of firms and thus the distortion of working hours in check, the eligibility condition must become stricter. Surprisingly, integrating the STW system optimally with a UI system proves to be less fiscally expensive than relying solely on the UI system.

In more detail, the model entails two potential reasons for government intervention. First, risk-averse workers cannot insure themselves via savings or on the financial market. Second, the Hosios (1990) condition might not be fulfilled, causing inefficiencies in vacancy posting. To counter these inefficiencies, the Ramsey planner can choose the UI and STW system.

The UI system pays unemployed workers UI benefits as their only source of income while they are unemployed. STW consists of two instruments: the eligibility condition and STW benefits. Firms and workers choose working hours freely and qualify for STW when the hours worked fall below a specific eligibility threshold. Hijzen and Martin (2013) show that most STW systems in practice use this type of hours reduction as eligibility criterion. Under STW, the government compensates workers for every hour they work less than usual.

The modeling differs from the recent business cycle literature on STW. Balleer, Gehrke, Lechthaler, and Merkl (2016), Gehrke, Lechthaler, and Merkl (2017), Dengler and Gehrke (2021), and Cooper, Meyer, and Schott (2017) argue that working hours outside STW are inflexible and that STW's role is to flexibilize the intensive margin. Instead, I follow the spirit of the implicit contract literature of Burdett and Wright (1989), Van Audenrode (1994), and Braun and Brügemann (2017), where working hours are flexible, STW acts as a subsidy, and subsidizing hours reduction leads to a suboptimal low choice of working hours under STW. Cahuc, Kramarz, and Nevoux (2021) emphasize the empirical and quantitative relevance of these hours' distortions in a partial equilibrium search model. Following the approach of the implicit contract literature has the advantage that STW benefits reduce separations directly.

Based on a search and matching model with flexible intensive margin, I derive the optimal policy mix between the UI system and the STW system in steady state. Thereby, we have to weigh the advantages against the costs of each system carefully.

The UI system reacts to the inability of workers to insure themselves against income losses. Optimal UI benefits have to balance the classical trade-off. On the one hand, higher UI benefits offer risk-averse workers more income insurance in case of job loss. On the other hand, it distorts vacancy posting and separations (cf. Pissarides, 2000). When firms and workers negotiate contracts, workers face a trade-off: stay employed during low productivity times but accept a lower salary, or opt for a higher salary with greater unemployment risk. The UI system increases the outside option of the worker and skews this decision towards higher unemployment risk and salaries, inefficiently raising overall separation rates. Furthermore, higher UI benefits drive up workers' salary demands, reducing firms' revenues and, consequently, their hiring efforts. Both effects lead to inefficiently high unemployment levels.

As a subsidy, STW could potentially increase inefficient low vacancy posting and reduce inefficiently high separation rates. First, let us focus on the job-finding rate. Papers like Balleer et al. (2016), Giupponi and Landais (2018), or Cahuc, Kramarz, and Nevoux (2021) support the idea that STW increases vacancy postings. By considering the financing of the STW system, I find the opposite. In my model, STW and the UI system are financed balanced budget by lump sum income taxes. When posting vacancies, firms do not know yet whether they will produce

regularly in the future and finance the STW system, or whether they enter the STW system and gain from it. Therefore, a more generous STW system will not increase vacancy posting efforts, as it has to be financed by higher contribution when producing regularly.
Surprisingly, even when ignoring the fiscal costs of financing the STW system, I can demonstrate that the Ramsey planner would still opt not to stabilize job-finding rates. This is because the distortionary effects of the STW system are too costly.

On the contrary, STW effectively reduces separations as it reduces firms' salary costs during downturns. Optimal STW benefits respond to two opposing forces. First, they are designed to counterbalance the fiscal externality of the UI system. Essentially, STW should reduce firms' salary costs by the amount an additional unemployed worker increases the fiscal costs of the UI system to the government. In spirit, this is the same rule as for optimal layoff taxes derived by Blanchard and Tirole (2008) and Cahuc and Zylberberg (2008).
Second, optimal STW benefits have to take their own distortionary effects stressed by Burdett and Wright (1989) into account. When benefits are increased, firms and workers will opt for even lower hours to draw in more support from the government. Therefore, the Ramsey planner does not prevent all inefficient separations caused by the UI system. Instead, he trades off stabilizing employment against distorting average hours worked and thus adjusts STW benefits downwards. In the end, the planner will not prevent all inefficient separations but will also accept some degree of working hours distortion.

It is important to note that while STW secures jobs, it does not provide direct income insurance. Income insurance for unemployed workers is offered by the UI system. Firms, on the other hand, provide income protection against idiosyncratic productivity shocks as long as the worker remains employed, a finding supported by the implicit contract literature (cf. Rosen (1985) or Braun and Brügemann (2017)). In my search and matching model, firms and workers negotiate income, working hours, and separations within a generalized Nash-Bargaining framework before the idiosyncratic productivity of the match is known. They establish a contract contingent on the realization of productivity shocks. As a result, the risk-neutral firm offers the risk-averse worker income insurance in exchange for lower expected salaries. This entails stock-up during times on STW as we have seen in the COVID-19 pandemic.[2]

The optimal STW benefits show that STW acts like a subsidy system. A key chal-

2. In Germany, large companies like Volkswagen, Telekom, and Deutsche Bahn, along with major unions such as IG Metall and Verdi, supplemented their employees' income on STW to 78%–95% of their regular earnings, while the STW system alone only provided 60% (cf. Münchner, 2020).

lenge of regular subsidy systems is, however, to effectively target firms that genuinely need support, as productivity levels are often not directly observable. STW addresses this issue with its eligibility condition. STW requires firms to reduce working hours below a certain threshold to qualify for the program. Since it is costly for more productive firms to significantly cut hours, they tend to reduce working hours less than temporarily less productive firms. This creates a natural screening mechanism for productivity. The minimum hours reduction requirement can be adjusted so that only firms below a certain target productivity level qualify for STW. Teichgräber, Žužek, and Hensel (2022) discuss hours reduction as a screening instrument in a mechanism design model.

I can show that the optimal eligibility condition implements the separation threshold firms and workers would choose if they didn't have access to STW as eligibility threshold. A looser eligibility condition would reinforce the distortionary effects of the STW system without saving additional jobs, resulting in pure windfall effects as in Cahuc, Kramarz, and Nevoux (2021). A tighter eligibility condition would risk losing firms that could have been saved with STW.

Quantitatively, firms and workers are eligible to go on STW in steady state if hours worked fall roughly by 66% below their normal value. This is significantly stricter than the 30% required in the German system. Further, STW replaces roughly 75% of a worker's wage in steady state. Further, the distortionary effects of STW reduce the optimal net-transfers by roughly 30%.

In my model, recessions are caused by a negative aggregate productivity shock. Salaries are assumed to be rigid to solve the Shimer (2005) puzzle. Reflecting European practices, I allow STW to adjust over the business cycle while the UI system stays unaltered.
When analyzing STW during recessions, I focus on two potential cases. First, I calibrate the model to the US economy, where the current UI system replaces 45% of wages, and look at how STW should optimally adjust over the business cycle. In a second policy experiment, I implement the optimal UI rate in steady state and repeat the exercise. In the model, the optimal replacement rate is at roughly 35% of income.

Over both cases, the reaction of the optimal STW system remains consistent. A decline in the job-finding rate increases unemployment durations. Workers do not fully internalize the social costs of separation, as they can stay on the unemployment system longer. To internalize the additional social costs from separations, STW benefits need to be increased. However, higher STW benefits incentivize more firms and workers to enter the STW system. Consequently, the eligibility criteria must be tightened to prevent windfall effects and limit distortions in working hours.

Depending on the UI setup in steady state, STW contributes to welfare through different channels. Broadly speaking, the Ramsey planner has three main objectives: First, he wants to reduce efficiency losses from business cycle fluctuations. These can be examined by the amount of inefficient fluctuations in aggregate consumption. Second, he wants to smooth individual consumption across households. STW can achieve this through the extensive margin by stabilizing employment. Therefore, employment stabilization serves both an efficiency and an insurance role. Finally, the planner aims to keep working hour distortions minimal, ensuring that the disutility of work is used effectively in the production process.

Under the current UI system, both consumption and employment fluctuate inefficiently with the business cycle. Optimal STW policy stabilizes around 30% of inefficient aggregate consumption fluctuations and 50% of inefficient employment fluctuations. However, by raising STW benefits and increasing the number of workers in the STW system, it destabilizes average working hours by approximately 20%, which accounts for much of the remaining gap in consumption stabilization.

Surprisingly, with optimal UI in steady state, the planner's STW policy no longer needs to address aggregate consumption fluctuations. Optimal UI nearly maximizes aggregate consumption in the model. Excessively high UI benefits lead to high unemployment rates and low output, while overly low benefits discourage workers from leaving unproductive jobs, thus keeping inefficient firms afloat. Consequently, marginal job matches contribute little to output, and their loss or preservation has minimal impact on aggregate consumption stabilization. Nonetheless, employment rates still fluctuate inefficiently, allowing STW to smooth individual consumption indirectly via the extensive margin.

The paper contributes to two additional topics discussed in the literature. First, Balleer et al. (2016) argue that STW acts as an automatic stabilizer. By stabilizing separations, STW helps to stabilize employment, output, and consumption without requiring adjustments throughout the business cycle. In my model, similar results are obtained for employment but not for consumption. Since the eligibility condition is not adjusted, even firms that could survive without STW enter the program. This exacerbates the distortionary effects on hours worked, negating any positive effects of employment stabilization on consumption. This experiment emphasizes the necessity of adjusting STW over the business cycle.

Second, a major concern of STW systems is that subsidizing low productive matches causes allocative inefficiencies. Cooper, Meyer, and Schott (2017) argue in a search and matching model that STW keeps workers in unproductive matches and hinders their reallocation to more productive firms, effectively reducing overall productivity and output. Within my model, I find that the social planner must balance the costs of reallocating a worker via the labor market against the costs of keeping a worker in an unproductive occupation. If the STW system is set too generously, the concern of Cooper, Meyer, and Schott (2017) is valid. Conversely, if STW benefits are set too low, the opposite occurs and we lose output due to high unemployment rates. By setting STW benefits optimally, it is possible to re-align private and social incentives, thereby avoiding misallocation effects.

The remaining structure of the paper is as follows. Section 2 introduces the model, explores the decentralized economy, and solves for the social planner economy. Section 3 derives analytical expressions for optimal STW policy and explores its theoretical implications. Section 4 describes the calibration of the model. Section 5 applies the optimal STW policy to a supply-side recession. Section 6 analyzes optimal STW policy with optimal UI in steady state. Finally, section 7 concludes the paper.

1.2 Model

The economy is populated by a continuum of workers of measure one, infinitely many one-worker firms, and a continuum ν_t^f of firm owners. Each firm produces a homogeneous and non-storable good. The economy is closed. Each period, firms and workers are subject to aggregate and idiosyncratic shocks. Nonetheless, firms are ex-ante homogeneous in their match-efficiency.

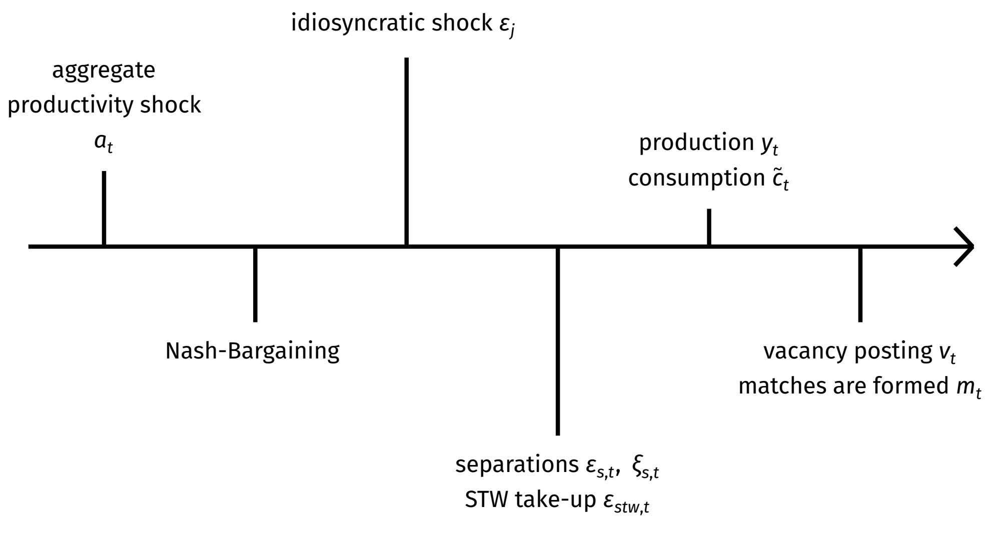


Figure 1.2.1. Period Timeline

The timeline of the period is structured as follows: At the start, firms experience an aggregate productivity shock. Before the idiosyncratic productivity shocks occur, generalized Nash-Bargaining takes place. Firms and workers write a contract specifying income, separations, hours of work, and short-time work (STW) take-up, all contingent on the realization of the idiosyncratic productivity shock. Following this, the idiosyncratic productivity is drawn. Separations and STW take-up take place. Then, output is produced based on working hours, and households consume. At the end of the period, vacancies are posted and new matches are formed. New matches don't produce until next period.

1.2.1 Decentralized Economy

In the decentralized economy, separations, vacancy postings, and working time are determined by firms and workers.

Firm Side. Each firm that enters a match with a worker can either produce or separate from the worker. There is an aggregate component a_t that is common to all matches and an idiosyncratic component ϵ_j that is, for analytical tractability, i.i.d. across time and matches with the distribution function $G(\epsilon)$.[3]
Firm-specific output $y_t(\epsilon, h_t(\epsilon))$ depends on the firm-specific productivity $a_t \cdot \epsilon$ which is divided in an aggregate productivity part a_t and the idiosyncratic part ϵ, the number of hours worked $h_t(\epsilon)$ and the resource costs of the firm $(\mu_\epsilon - \epsilon_j) \cdot c_f$:[4]

$$y_t(\epsilon, h_t(\epsilon)) \;=\; a_t \cdot \epsilon \cdot h_t(\epsilon)^\alpha - (\mu_\epsilon - \epsilon) \cdot c_f \quad \text{with} \quad E[(\mu_\epsilon - \epsilon) \cdot c_f] \;=\; 0$$

In line with Krause and Lubik (2007), I assume that the idiosyncratic shock $log(\epsilon_j)$ follows a normal distribution $\epsilon_j \sim \mathcal{N}(\mu, \sigma^2)$ with $\mu_\epsilon = E[\epsilon_j] = \exp(\mu + \frac{1}{2} \cdot \sigma^2)$. Furthermore, I assume that aggregate productivity follows an AR(1) process:

$$a_t \;=\; \mu_a + \rho_a \cdot (a_{t-1} - \mu_a) + \iota_t, \quad \rho_a \;\in\; [0, 1), \quad \iota_t \;\sim\; \mathcal{N}(0, \sigma_a^2)$$

Firms are assumed to be owned by firm owners. As a result, future cash flows are discounted using a stochastic discount factor, reflecting how firm owners weigh

3. Having persistent idiosyncratic shocks, we would need a state vector to keep track of the productivity distribution of the firms. This would make computing the Ramsey policy very difficult. c_f can also be interpreted as a measure for the persistence of the idiosyncratic productivity shocks.

4. Note that the cost shock of the firm is important, if we want to have a quantitatively realistic impact of the UI system on unemployment, endogenous separations, and time-independent idiosyncratic shocks in an otherwise analytically tractable model. It is a well-known problem that search and matching models overstate the importance of the UI system (see Costain and Reiter (2008)). To have a sensible impact of the UI system, we need a large surplus calibration. The bigger the surplus, the smaller the relative impact of a change in UI benefits. However, large surpluses lead to small separation incentives. Since the cost shock has an expectation value of zero, it allows for a large surplus calibration. At the same time, it affects the marginal firms the most, allowing for endogenous separations.

future marginal utility of consumption against today's:

$$Q^f_{t,t+1} \;=\; \beta \cdot \frac{u'(c^f_{t+1})}{u'(c^f_t)}$$

The value of a worker for a firm, that is not on STW, and whose idiosyncratic shock has realized to ϵ, is:

$$J_t(\epsilon) \;=\; y_t(\epsilon, h_t(\epsilon)) - w_t(h_t(\epsilon)) + E_t\left[Q^f_{t,t+1}\mathcal{J}_{t+1}\right]$$

The firm gets the production value of the match $y_t(\epsilon, h_t(\epsilon))$ but pays the wage-sum $w_t(h_t(\epsilon))$ dependent on the total working hours to the worker. $E_t\left[Q^f_{t,t+1}\mathcal{J}_{t+1}\right]$ denotes the expected discounted continuation value of the value of the worker for the firm.

The value of a worker for a firm, who is on STW, and whose idiosyncratic productivity has the value ϵ, can be written as:

$$J_{stw,t}(\epsilon) \;=\; y_t(\epsilon, h_{stw,t}(\epsilon)) - w_t(h_{stw,t}(\epsilon)) + E_t\left[Q^f_{t,t+1}\mathcal{J}_{t+1}\right]$$

Since working hours fall on STW, firms have to pay a smaller salary.

The access to the STW system is restricted by the government via the eligibility condition. As in Cahuc, Kramarz, and Nevoux (2021), firms and workers have the option to transition to STW when the number of hours worked falls below a specific threshold set by the government, denoted as D_t. This threshold serves as a criterion for determining eligibility for STW:

$$h_{stw,t}(\epsilon) \;\leq\; h_{stw,t}(\epsilon_{stw,t}) \;=\; D_t$$

It essentially means that firms and workers are eligible to participate in the STW system if they reduce their hours worked by a certain percentage below their normal level, $\frac{D_t-\bar{h}}{\bar{h}} \cdot 100\%$, where $\bar{h}$ represents the mean hours worked in the steady state. This eligibility condition is consistent with findings by Hijzen and Martin (2013), who identify that 15 out of 24 OECD countries with STW programs in place employ this minimum hours' reduction as an eligibility criterion. In subsequent sections, we say that the eligibility condition becomes looser when D_t increases, indicating that it becomes easier to enter into STW. The eligibility threshold, denoted as $\epsilon_{stw,t}$, is defined based on temporary productivity ϵ and is implicitly determined by the equation $D_t = h_{stw,t}(\epsilon_{stw,t})$. In the spirit of Teichgräber, Žužek, and Hensel (2022), the hours reduction criterion can be used as an instrument to screen for productivity and jobs at risk.

Depending on the values of D_t, respectively $\epsilon_{stw,t}$, the eligibility threshold may or may not be binding and can have various impacts on the economy. We need to consider four distinct cases, as illustrated in Figure 1.2.2.

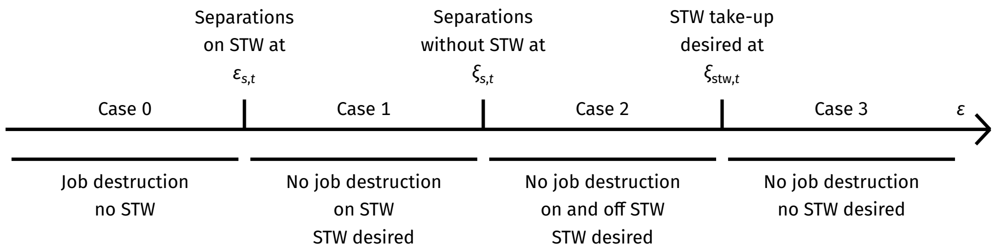


Figure 1.2.2. Thresholds and employment outcomes across productivity levels

Case 0: $\epsilon_{stw,t} < \epsilon_{s,t}$. In case 0, the eligibility threshold is stricter than the separation threshold for firms and workers on the STW system. Under these conditions, no firm or worker will ever access the STW system, rendering it obsolete. For the sake of notational brevity, we will exclude this case from further consideration in subsequent sections, as it does not limit the planner's choice set. We require $\epsilon_{stw,t} \geq \epsilon_{s,t}$. By setting $\epsilon_{stw,t} = \epsilon_{s,t}$, the Ramsey planner can still make the STW system obsolete.

Case 1: $\epsilon_{s,t} \leq \epsilon_{stw,t} < \xi_{s,t}$. Case 1 describes a situation where matches with lower productivity $\epsilon \in [\epsilon_{s,t}, \epsilon_{stw,t}]$ are allowed on the STW system and are rescued, while matches with higher productivity $\epsilon \in (\epsilon_{stw,t}, \xi_{s,t})$ are not allowed and dissolve. Here, $\xi_{s,t}$ denotes the separation threshold of matches without access to STW, determined within the generalized Nash-Bargaining framework.

Case 2: $\xi_{s,t} \leq \epsilon_{stw,t} < \xi_{stw,t}$. In case 2, all firms and workers that would dissolve without STW can enter the STW system. At the same time, the eligibility threshold denies matches with productivity $\epsilon \in (\epsilon_{stw,t}, \xi_{stw,t}]$ access to STW. Note that these matches want to take up STW but are not at risk of breaking up. Following Cahuc, Kramarz, and Nevoux (2021), I will refer to them as windfall effects. Here, $\xi_{stw,t}$ denotes the STW take-up threshold of firms and workers. This threshold determines the idiosyncratic productivity level at which firms and workers want to enter the STW system. It is also determined within the generalized Nash-Bargaining framework.

Case 3: $\xi_{stw,t} \leq \epsilon_{stw,t}$. In case 3, the eligibility condition becomes so loose that it no longer binds. Firms and workers do not want to take up STW. Without loss of generality, we can assume that the planner wants to set $\epsilon_{stw,t} \leq \xi_{stw,t}$ and excludes the case from further considerations. Setting $\epsilon_{stw,t} = \xi_{stw,t}$ has the same effect as setting $\xi_{stw,t} < \epsilon_{stw,t}$.

To wrap up, we have seen that only cases 1 and 2 are relevant for the subsequent analysis. Without loss of generality and for notational brevity, we require the eligibility threshold to be at least as large as the separation threshold of firms and workers with access to STW but not larger than the STW take-up threshold for firms and workers:

$$\epsilon_{s,t} \;\leq\; \epsilon_{stw,t} \;\leq\; \xi_{stw,t}$$

From this, we can derive the separation rate. The separation rate depends on the probability that firms in the STW system dissolve, plus the probability that firms and workers experience a productivity shock strong enough to cause dissolution but not strong enough to warrant entry into the STW system (case 1):

$$\rho_t \;=\; G(\epsilon_{s,t}) + \max\{G(\xi_{s,t}) - G(\epsilon_{stw,t}), 0\}$$

The expected value of a worker for a firm right before the idiosyncratic shock has realized can be denoted as:

$$\mathscr{J}_t \;=\; \int_{max\{\epsilon_{stw,t},\xi_{s,t}\}}^{\infty} J_t(\epsilon)dG(\epsilon) + \int_{\epsilon_{s,t}}^{\epsilon_{stw,t}} J_{stw,t}(\epsilon)dG(\epsilon) - \rho_t \cdot \left(w_{eu,t} + F\right) \quad (1.2.1)$$

When the idiosyncratic productivity exceeds both the eligibility threshold and the separation threshold for firms without access to STW, $\epsilon \geq \max\{\epsilon_{stw,t}, \xi_{s,t}\}$, the firm continues to operate regularly. If productivity falls within the interval $\epsilon \in [\epsilon_{s,t}, \epsilon_{stw,t}]$, firms shift to production under STW. Further, when firms decide to separate from a worker, they incur two types of costs. First, they must pay severance payments, denoted as $w_{eu,t}$. Severance payments compensate the worker for the loss of employment and are part of the contract that firms and workers bargain over at the beginning of the period. Second, firms face fixed costs of job destruction, represented by F. These costs include administrative and legal expenses associated with removing the worker from the payroll, as well as efficiency losses due to the need to restructure the production process.[5]

Firms post vacancies v_t until the expected costs of recruiting a worker equal the discounted expected value of a worker for the firm.

$$\frac{k_v}{q_t} \;=\; E_t\left[Q^f_{t+1}\mathscr{J}_{t+1}\right] \quad (1.2.2)$$

Here, q_t denotes the probability of filling a vacancy and k_v the costs of posting a vacancy.

5. Research by Kuhn, Manovskii, Bellmann, Gleiser, Hensgen, et al. (2021) indicates that firms often operate with coordinated teams and work processes. Separation from a worker disrupts this coordination, resulting in output losses. According to the study, firms view these costs as one of the main reasons for the use of STW.

Firm Owners. There exists a continuum ν_t^f of firm owners in the economy that have positive but decreasing utility of consumption $u()$. Firm owners consume the profits Π_t produced in the firm sector. Profits are spread equally among all firm owners. Besides that, they do not make any decisions. The value of a firm owner is denoted as:

$$V_t^f = u\left(c_t^f\right) + \beta \cdot E_t\left[V_{t+1}^f\right] \quad \text{with} \quad c_t^f = \frac{\Pi_t}{\nu_t^f}$$

Profits equal total output minus the wage bill, separation, and vacancy posting costs:

$$\begin{aligned}\Pi_t = n_t \cdot \Bigg(&\int_{\max\{\epsilon_{stw,t},\xi_{s,t}\}}^{\infty} \big(y_t(\epsilon, h_t(\epsilon)) - w_t(h_t(\epsilon))\big)\, dG(\epsilon) \\ &+ \int_{\epsilon_{s,t}}^{\epsilon_{stw,t}} \big(y_t(\epsilon, h_{stw,t}(\epsilon)) - w_t(h_{stw,t}(\epsilon))\big)\, dG(\epsilon)\Bigg) \\ &- \rho_t \cdot n_t \cdot (w_{eu,t} + F) - k_v \cdot v_t\end{aligned}$$

Worker Side. The value of an employed worker with idiosyncratic productivity ϵ can be written as:

$$\begin{aligned}V_t^w(\epsilon) &= u\big(w_t(h_t(\epsilon)) - \tau_{J,t} - \nu(h_t(\epsilon))\big) + \beta \cdot E_t\left[\mathcal{V}_{t+1}^w\right] \\ &\text{with} \quad u'(\cdot) > 0, \quad u''(\cdot) < 0, \quad \nu(0) = 0\end{aligned}$$

Workers derive utility from consumption and disutility from working $\nu(h)$. Each period, workers consume their after tax salary $w_t(h_t(\epsilon)) - \tau_{J,t}$. Further, workers are risk-averse. The use of the quasi-linear utility function excludes income effects and makes the theoretical results cleaner.[6] The expected value of entering next period's employment is denoted by $E_t\left[\mathcal{V}_{t+1}^W\right]$.

The value of an employed worker on STW can be denoted as:

$$\begin{aligned}V_{stw,t}^w(\epsilon) = u\Big(&\underbrace{w_t(h_{stw,t}(\epsilon))}_{\text{Reduced income by firm}} + \underbrace{\tau_{stw,t} \cdot (\bar{h} - h_{stw,t}(\epsilon))}_{\text{net transfer STW}} - \tau_{J,t} - \nu(h_{stw,t}(\epsilon))\Big) \\ &+ \beta \cdot E_t[\mathcal{V}_{t+1}^w]\end{aligned}$$

During STW, firms and workers will agree on reducing working hours. Consequently, the income of workers falls. The government now steps in and compensates the worker for every hour he works less than he would normally do. $\bar{h}$

6. Mathematically, it allows to map the model's flexible intensive margin into a standard search and matching model with risk aversion.

denotes the average hours worked in steady state. Note that only the least productive firms will reduce working hours sufficiently to enter the STW system. As a result, STW is a subsidy to the least productive matches.

The expected value of a worker at the beginning of the period is:

$$\begin{aligned}\mathcal{V}_t^w &= \int_{\max\{\epsilon_{stw,t},\xi_{s,t}\}}^{\infty} V_t^w(\epsilon)\, dG(\epsilon) + \int_{\epsilon_{s,t}}^{\epsilon_{stw,t}} V_{stw,t}^w(\epsilon)\, dG(\epsilon) \\ &\quad + \rho_t \cdot \big(u(w_{eu,t} - \tau_{J,t}) - u(b_t) + U_t\big)\end{aligned}$$

As in the equation 1.2.1 for the expected value of the firm, households work normally if the idiosyncratic productivity is large $\epsilon > \max\{\epsilon_{stw,t}, \epsilon_{s,t}\}$, go on STW if $\epsilon \in [\epsilon_{s,t}, \epsilon_{stw,t}]$ and get unemployed for $\epsilon < \epsilon_{s,t}$, respectively $\epsilon \in (\epsilon_{stw,t}, \xi_{s,t})$. When workers get unemployed, they receive severance payments $w_{eu,t}$. Workers still have to pay taxes $\tau_{J,t}$ on the severance payment. As in Jung and Kuester (2015), workers get no unemployment insurance in the period when they receive the severance payment.[7]

The value of an unemployed worker at the beginning of the period can be written as:

$$U_t = u(b_t) + \beta \cdot E_t\big[f_t \cdot \mathcal{V}_{t+1}^w + (1 - f_t) \cdot U_{t+1}\big]$$

Being unemployed, a worker receives unemployment benefits b_t. With probability f_t, the worker finds a job and gets the value of being employed at the beginning of the next period. Otherwise, the worker stays unemployed.

Nash-Bargaining. Firms and workers bargain over salaries $w_t(h)$, severance payments $w_{eu,t}$, the hours worked on STW $h_{stw,t}(\epsilon)$ and off STW $h_t(\epsilon)$, the voluntary STW take-up threshold $\xi_{stw,t}$ and the separation decisions with STW $\epsilon_{s,t}$ and without STW $\xi_{s,t}$ before the idiosyncratic productivity is known in a generalized Nash-Bargaining set-up. They can write a contract based on the realization of each idiosyncratic productivity state ϵ. η_{t-1} denotes the bargaining power of the worker. The bargaining solves:[8]

$$\max_{w_t(h), w_{eu,t}, h_t(\epsilon), h_{stw,t}(\epsilon), \xi_{s,t}, \xi_{stw,t}, \epsilon_{s,t}} \mathcal{J}_t^{1-\eta_{t-1}} \cdot (\mathcal{V}_t^w - U_t)^{\eta_{t-1}}$$

7. This reduces the elasticity of the separation rate on movements in the UI benefits, helping to solve the puzzle of Costain and Reiter (2008).

8. Derivations can be found in Appendix 1.F

The risk-neutral firm decides to offer the risk-averse worker a contract that insures the worker against any idiosyncratic productivity shock within that period: [9]

$$
\begin{aligned}
u'\Big(\underbrace{w_t(h_t(\epsilon)) - \tau_{J,t} - v(h_t(\epsilon))}_{\tilde{c}_t(\epsilon)}\Big) &= u'\Big(\underbrace{w_t(h_{stw,t}(\epsilon)) - \tau_{J,t} - v(h_{stw,t}(\epsilon))}_{\tilde{c}_{stw,t}(\epsilon)}\Big) \\
&= u'\Big(\underbrace{w_{eu,t} - \tau_{J,t}}_{\tilde{c}_{eu,t}}\Big)
\end{aligned}
$$

It guarantees the same consumption equivalent and therefore utility regardless of whether workers work regularly, are on STW, or get laid off. When workers get laid off, they receive severance payments for that period:

$$
\tilde{c}_t = \tilde{c}_t(\epsilon) = \tilde{c}_{stw,t}(\epsilon) = \tilde{c}_{eu,t}
$$

Once workers get laid off and do not find new employment within the same period, they transition to the UI system. To achieve the same utility on and off STW, firms must stock up part of the worker's salary on STW, a behavior we have actually seen in the COVID-19 pandemic. Note that utility is equalized, which gives limited room for the salaries to adjust according to the working hours.

The effect of risk aversion on the wage can also be seen in the wage-formula, respectively the formula for the consumption equivalents:

$$
\begin{aligned}
\tilde{c}_t = {} & \underbrace{\eta_{t-1} \cdot \left(z_t - \frac{1-n_t}{n_t} \cdot b_t - \rho_t \cdot F + (1-\rho_t) \cdot \theta_t \cdot k_v\right) + (1-\eta_{t-1}) \cdot b_t}_{\text{Resource Share of Joint Surplus}} \\
& \underbrace{- (1-\eta_{t-1}) \cdot \int_{b_t}^{\tilde{c}_t} \left[\frac{u'(x)}{u'(\tilde{c}_t)} - 1\right] dx}_{\text{Penalty for Risk Aversion of Workers} > 0}
\end{aligned}
\tag{1.2.3}
$$

The formula for consumption equivalents consists of two parts. The first part describes how the additional resources that the match generates, beyond what the UI system provides, are shared between firms and workers. The outcome is the same as under risk-neutrality.

9. Derivations of the optimality conditions implied by the Nash-Bargaining can be found in the appendix in section 1.F.

In the formula, z_t represents the expected consumption equivalents produced by the match:

$$z_t = \int_{\max\{\epsilon_{stw,t},\xi_{s,t}\}}^{\infty} y_t(\epsilon, h_t(\epsilon)) - v(h_t(\epsilon))dG(\epsilon) + \int_{\epsilon_{s,t}}^{\epsilon_{stw,t}} y(\epsilon, h_{stw,t}(\epsilon)) - v(h_{stw,t}(\epsilon))dG(\epsilon)$$

The second part of the formula describes how risk aversion reduces workers' income. On the one hand, risk-averse workers prefer a steady income stream. The risk-neutral firm knows that and provides income insurance against any idiosyncratic shocks to the worker. In exchange for the income insurance, workers are willing to give up part of their salary. On the other hand, incomplete unemployment insurance erodes workers' outside options, as the marginal utility of consumption is high when workers depend on UI benefits. Firms exploit this by threatening to walk away from negotiations, securing lower salary levels. The threat becomes more potent as the gap between the consumption equivalents and UI benefits grows.

Note that firms only offer income insurance against idiosyncratic shocks but not against aggregate shocks. Aggregate shocks have full pass through to the salary of the worker. This flexibility in salary causes the so-called Shimer Puzzle, which states that search and matching models struggle to generate sufficient cyclical fluctuations. The puzzle is commonly resolved by introducing wage-rigidity (see Hall, 2005) or, in my case, rigid salaries. In the implementation of rigid salaries, I follow Jung and Kuester (2015) and assume procyclical bargaining power of the firms.

$$(1 - \eta_t) = \exp(\gamma_w \cdot a_t), \quad \gamma_w > 0$$

We can relate the expression to rigid salaries as follows: If productivity falls in recessions, but salaries are rigid, a larger share of the joint surplus is claimed by the workers. In a model with Nash-Bargaining, this is equivalent to reducing the firms' or respectively increasing the workers' bargaining power. Fahr and Abbritti (2011), for instance, show that the existence of wage adjustment costs leads to the procyclical bargaining power of the firm.

As in the efficient bargaining setup of Trigari (2006), hours are chosen to maximize the joint surplus. As a result, outside STW, the marginal product of hours worked needs to equal its marginal disutility. This is the solution the social planner would choose as well (see equation 1.2.7):

$$\underbrace{\frac{\partial y_t(\epsilon, h_t(\epsilon))}{\partial h_t(\epsilon)}}_{\text{Marginal Product of Labor}} = \underbrace{v'(h_t(\epsilon))}_{\text{Marginal Disutility of Work}} \tag{1.2.4}$$

Workers with low idiosyncratic productivity work less to save disutility of hours worked, while those with high idiosyncratic productivity will work more to make use of the extra productivity boost. This result can be interpreted as some kind of perfect working time account. Working time accounts let workers do overtime in good times while reducing working time in bad times. Such flexible working times gain importance, for example, in Germany (see Ellguth, Gerner, and Zapf, 2018). A reduction in aggregate productivity will reduce the working hours of every worker in the economy.

Firms and workers want to access the STW system when the surplus gain from the STW subsidy exceeds the loss from working hours reduction:

$$\underbrace{\tau_{stw,t} \cdot \left(\bar{h} - h_{stw,t}(\xi_{stw,t})\right)}_{\text{Surplus Gain from STW Subsidy}}$$
$$= \underbrace{y_t(\xi_{stw,t}, h_{stw,t}(\xi_{stw,t})) - v(h_{stw,t}(\xi_{stw,t})) - y_t(\xi_{stw,t}, h_t(\xi_{stw,t})) + v(h_t(\xi_{stw,t}))}_{\text{Surplus Loss from suboptimal low Working Hours}}$$

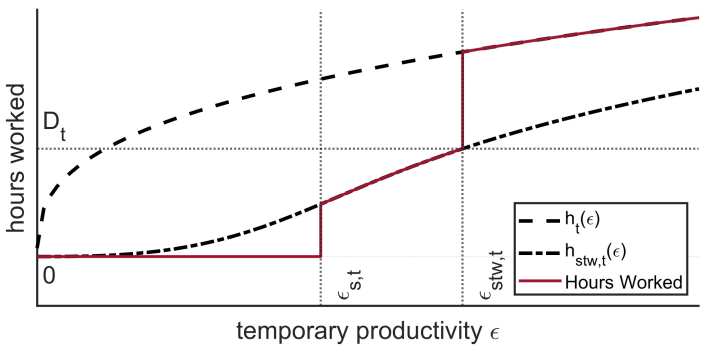
If firms and workers are highly productive, they will decide not to enter the STW system, as they can only attract benefits if they reduce the working hours below the usual level $\bar{h}$. Instead, they want to work more than usual to exploit the benefits of the extra productivity. All matches with an hours choice of $h_t(\epsilon) < \bar{h}$ will want to enter the STW system to exploit benefits. Naturally, the government would want to set a stricter eligibility condition, as otherwise more than 50% of the workforce would want to enter the STW system.

Working hours on STW are chosen sub-optimally low. By reducing the number of hours worked, firms and workers can not only reduce disutility from work but can also attract more STW benefits (see Cahuc, Kramarz, and Nevoux, 2021):

$$\underbrace{\frac{\partial y_t(\epsilon, h_{stw,t}(\epsilon))}{\partial h_{stw,t}(\epsilon)}}_{\text{Marginal Product of Labor}} = \underbrace{v'(h_{stw,t}(\epsilon))}_{\text{Marginal Disutility of Work}} + \underbrace{\tau_{stw,t}}_{\text{STW Benefits}} \tag{1.2.5}$$

To provide a visual representation of the hours distortion effect of STW for the case $\xi_{s,t} \leq \epsilon_{stw,t}$, figure 1.2.3 displays the relationship between hours worked and idiosyncratic productivity. When productivity is high, workers tend to work their normal hours. However, as productivity decreases, both firms and workers have the option to utilize the STW program. Under STW, working hours are reduced below the optimal level. We refer to this reduction of hours as the hours distortion effect of STW. The loss in working hours is represented as the area between the number of hours worked without STW and the actual hours worked on STW. Its impact on the optimal provision of STW will be discussed extensively in subsequent sections. If productivity declines even further, separations occur, and working hours fall to zero.

Figure 1.2.3. Hours Distortion Effect of STW



Notes: The figure illustrates how STW influences the hours choice decision. $h_t(\varepsilon)$ denotes the hours that firms and workers would choose if they were not on STW. $h_{stw,t}(\varepsilon)$ denotes the hours choice if they were on STW. The red line shows the actual hours choice dependent on being on regular production $\varepsilon > \varepsilon_{stw,t}$, on STW $\varepsilon \in [\varepsilon_{s,t}, \varepsilon_{stw,t}]$ or separated $\varepsilon < \varepsilon_{s,t}$. The differences between the red line and the dashed line on STW show the hours distortion effect of STW.

Separations occur if the joint surplus, after the idiosyncratic shock has been realized, becomes negative. The separation threshold without access to STW can thus be determined as:

$$y_t(\xi_{s,t}, h_t(\xi_{s,t})) - v(h_t(\xi_{s,t})) + F + \frac{1-\eta_t \cdot f_t}{1-\eta_t} \cdot \frac{k_v}{q_t} = 0$$

The separation threshold with STW can be determined as:

$$y_t(\epsilon_{s,t}, h_{stw,t}(\epsilon_{s,t})) - v(h_{stw,t}(\epsilon_{s,t})) + (\bar{h} - h_{stw,t}(\epsilon_{s,t})) \cdot \tau_{stw,t} + F + \frac{1-\eta_t \cdot f_t}{1-\eta_t} \cdot \frac{k_v}{q_t} = 0 \tag{1.2.6}$$

Firms and workers want to separate if period output minus disutility from work becomes negative, but are disincentivized by potential separation costs. Furthermore, firms want to hoard workers to save search costs for a new worker, while the worker would lose its expected value of being employed by the separation. This value is reduced by the opportunity of the worker to find a new job, which is represented by $1-\eta_t \cdot f_t < 1$. Notably, STW increases the joint surplus and disincentivizes separations. Firms are committed to insure workers against income fluctuations, even in bad times. Higher STW benefits reduce the salary-commitment of firms in bad times and thus decrease separation incentives.

Budget Constraint Government. I assume that the government must balance its budget every period. Income taxes finance the UI system and the STW system.

$$n_t \cdot \tau_{J,t} = \left(1 - n_t\right) \cdot b_t + n_t \cdot \int_{\epsilon_{s,t}}^{\epsilon_{stw,t}} \tau_{stw,t} \cdot \left(\bar{h} - h_{stw,t}(\epsilon)\right) \cdot dG(\epsilon)$$

The government will determine the UI and the STW system endogenously. The tax is adjusted accordingly.

Labor Market Flows. Based on the timing of the economy, we can formulate the law of motion of employment n_t:

$$n_t = (1 - \rho_t) \cdot n_{t-1} + m_{t-1}$$

Here, n_t denotes the number of employed workers at the beginning of the period. m_t denotes the number of newly formed matches. $1 - n_t + \rho_t \cdot n_t$ denotes the number of unemployed workers after separations took place. Unemployed workers are matched with vacancies v_t according to a Cobb-Douglas matching function:

$$m_t = \bar{m} \cdot v_t^{1-\gamma} \cdot (1 - n_t + \rho_t \cdot n_t)^{\gamma}$$

The parameter $\bar{m}$ determines the matching efficiency, and $\gamma \in (0, 1)$ denotes the elasticity of the matching function for unemployment. The labor market tightness is defined as the ratio of vacancies to unemployed $\theta_t = \frac{v_t}{1-n_t+\rho_t \cdot n_t}$. Based on the matching function and the labor market tightness, we can derive the probability to find a job f_t and the probability to fill a vacancy q_t:

$$f_t = \bar{m} \cdot \theta_t^{1-\gamma}, \qquad q_t = \bar{m} \cdot \theta_t^{-\gamma}$$

The number of separations s_t can be determined by:

$$s_t = \rho_t \cdot n_t$$

Market Clearing. The market clearing is defined via consumption equivalents. These can be used to pay for aggregate vacancy posting costs $v_t \cdot k_v$, separation costs $s_t \cdot F$, and consumption equivalents of employed $\tilde{c}_t^w$ and unemployed c_t^u workers as well as firm owners c_t^f:

$$n_t \cdot z_t = \underbrace{v_t \cdot k_v + s_t \cdot F}_{\text{Reallocation Costs}} + n_t \cdot \tilde{c}_t^w + (1 - n_t) \cdot c_t^u + v_t^f \cdot c_t^f$$

1.2.2 Social Planner

I assume that the social planner equally weights the utility of every household. The planner can freely allocate consumption, working hours, separations, and the job-finding rate, given the production technology of the economy (I) and the matching technology, respectively, the law of motion of employment (II).

$$\begin{aligned} W_t^P \;=\; & \max_{\theta_t, \epsilon_{s,t}, h_t(\epsilon)} n_t \cdot \int_0^\infty u(\tilde{c}_t^w(\epsilon))\, dG(\epsilon) + (1-n_t)\cdot u(c_t^u) + \nu_t^f \cdot u(c_t^f) \\ & + \beta \cdot E_t\left[W_{t+1}^P\right] \end{aligned}$$

subject to

$$\begin{aligned} \text{(I)} \quad n_t \cdot \int_0^\infty & \tilde{c}_t^w(\epsilon)\, dG(\epsilon) + (1-n_t)\cdot c_t^u + \nu_t^f \cdot c_t^f \\ & = n_t \cdot \int_{\epsilon_{s,t}}^\infty \left[y(\epsilon, h_t(\epsilon)) - v_t(h_t(\epsilon))\right] dG(\epsilon) \\ & - \theta_t \cdot \left(1 - n_t + G(\epsilon_{s,t})\cdot n_t\right)\cdot k_v - n_t \cdot G(\epsilon_{s,t}) \cdot F \\ \text{(II)} \quad n_{t+1} \;=\; & (1 - G(\epsilon_{s,t}))\cdot n_t + f(\theta_t)\cdot\left(1 - n_t + G(\epsilon_{s,t})\cdot n_t\right) \end{aligned}$$

Since workers and firm owners are risk-averse, the social planner wants to offer the same consumption equivalents and thus utility, regardless of whether a worker is employed or unemployed.

$$\tilde{c}_t \;=\; \tilde{c}_t^w(\epsilon) \;=\; c_t^u \;=\; c_t^f$$

Note that in the decentralized economy, firms can insure employed workers against idiosyncratic productivity shocks but cannot insure unemployed workers. Unemployed workers need to resort to the UI system. As in the decentralized economy, the social planner cannot insure households against aggregate shocks.

Since the planner can allocate resources freely, he tries to maximize output minus disutility from work and reallocation costs of a worker via the labor market. Just like firms and workers outside STW, the planner selects working hours such that the marginal productivity of hours worked is equal to the marginal disutility derived from work:[10]

$$\underbrace{\frac{\partial y_t(\epsilon, h_t(\epsilon))}{\partial h_t(\epsilon)}}_{\text{Marginal Product of Labor}} \;=\; \underbrace{v'(h_t(\epsilon))}_{\text{Marginal Disutility of Work}} \tag{1.2.7}$$

10. The derivations of the optimality conditions of the planner can be found in appendix 1.G.

As a result, working hours are optimally determined in the absence of STW intervention. The social planner discounts future welfare using a stochastic discount factor reflecting how households weigh future marginal utility of consumption against today's:

$$Q_{t,t+1} = \beta \cdot \frac{u'(\tilde{c}_{t+1})}{u'(\tilde{c}_t)}$$

The optimal hiring condition can be written as:

$$\underbrace{\frac{k_v}{q_t}}_{\text{Recruitment Costs}} = \underbrace{(1-\gamma)}_{\substack{\text{Static}\\ \text{Congestion Externality}}} \cdot \underbrace{E_t\Big[Q_{t,t+1}\Big(\int_{\epsilon_{s,t+1}}^{\infty}[y_{t+1}(\epsilon, h_{t+1}(\epsilon)) - v\big(h_{t+1}(\epsilon)\big)]dG(\epsilon) - G(\epsilon_{s,t+1}) \cdot F\Big)\Big]}_{\text{Expected Increase in Welfare}}$$
$$+ E_t\left[Q_{t,t+1} \underbrace{\big(1-\gamma \cdot f_{t+1}\big)}_{\substack{\text{Dynamic}\\ \text{Congestion Externality}}} \cdot \underbrace{(1-G(\epsilon_{s,t+1})) \cdot \frac{k_v}{q_{t+1}}}_{\text{Saved Recruitment Costs}}\right]$$

By creating and filling a new vacancy, the planner increases output and saves recruitment costs in the subsequent period. However, an increase in hiring also leads to a congestion externality, as firms compete for the available pool of unemployed workers. This externality has both a static and intertemporal component.

First, when firms post vacancies, they reduce the probability of other firms filling their vacancies, leading to higher recruitment costs. This relationship is reflected in the term $(1-\gamma) < 1$.

Second, keeping workers employed reduces unemployment and thus increases the labor market tightness. A larger labor market tightness reduces the probability of filling a vacancy and increases recruitment costs for other firms. The effect is captured in the term $(1-\gamma \cdot f_{t+1}) < 1$, which discounts the potential future recruitment cost savings.

The optimal labor market tightness is determined such that the expected costs of filling a vacancy equal its social benefits.

From the perspective of a social planner, separations should occur if the costs of keeping an unproductive match alive surpass the social costs of reallocating a worker via the labor market:

$$\underbrace{a_t \cdot \epsilon_{s,t} \cdot h_t(\epsilon_{s,t})^\alpha - (\mu_\epsilon - \epsilon_{s,t}) \cdot c_f - v\big(h_t(\epsilon_{s,t})\big)}_{\text{Social costs from keeping unproductive matches alive}} = \underbrace{-F - \frac{1-\gamma \cdot f_t}{1-\gamma} \cdot \frac{k_v}{q_t}}_{\text{Social costs from reallocating a worker via the labor market}}$$

The costs of reallocating a worker via the labor market entail the costs of employee turnover for the firm. These are the costs of separating from an old and recruiting a new worker, and the opportunity costs of leaving a worker outside production. The opportunity costs rise with a fall in the job-finding rate, as the worker stays longer unemployed.

$$F + \frac{1-\gamma \cdot f_t}{1-\gamma} \cdot \frac{k_v}{q_t} = \underbrace{F + \frac{k_v}{q_t}}_{\text{Expected Costs of Employee Turnover}} + \underbrace{(1-f_t) \cdot \frac{\gamma}{1-\gamma} \cdot \frac{k_v}{q_t}}_{\text{Opportunity Costs of having a Worker Outside Production}}$$

Note that having a worker outside production also entails a positive search externality as unemployment rises, which increases the probability for all firms to recruit a worker. This is represented by $\gamma \in (0, 1)$ which reduces the costs of lost production.

1.3 Optimal STW Policy

1.3.1 Ramsey Problem

The Ramsey planner weights the utility of every worker equally. In order to bring the decentralized economy as close as possible to the social planner economy, the Ramsey planner can choose the UI benefits b_t and the parameters of the STW system: the eligibility condition D_t and the STW benefits $\tau_{stw,t}$. The Ramsey planner's decisions are subject to the decentralized labor market equilibrium. With UI benefits, the planner can adjust the income insurance provided to unemployed workers. With STW, the Ramsey planner can influence the separation rate. However, the use of STW also introduces a distortion in the choice of working hours, which is summarized in the welfare cost term $n_t \cdot \Omega_t$. The Ramsey problem can be denoted as:

$$
\begin{aligned}
W_t^G = \max_{D_t,\tau_{stw,t},b_t} & (1-n_t)\cdot u(b_t) + n_t \cdot u(\tilde{c}_t^w) \\
& + \nu_t^f \cdot u\Bigg(\Big[n_t \cdot \int_{\mathscr{B}_t} \big(y_t(\epsilon, h_t(\epsilon)) - v(h_t(\epsilon))\big)\, dG(\epsilon) \\
& \qquad - n_t \cdot \Omega_t - n_t \cdot \tilde{c}_t^w - (1-n_t)\cdot b_t - n_t \cdot \rho_t \cdot F - v_t \cdot k_v\Big] / \nu_t^f\Bigg) \\
& + \beta \cdot E_t\left[W_{t+1}^G\right]
\end{aligned}
$$

s.t. decentralized Equilibrium

The full problem is displayed in Appendix 1.I. Integration takes place over all productivity states with employment $\mathscr{B}_t = [\epsilon_{s,t}, \xi_{s,t}] \cup [\xi_{s,t}, \epsilon_{stw,t}]$. Since the primary focus of the paper is not on distributional conflicts between firm owners and employed workers, I set the number of firm owners ν_t^f such that firms and workers have the same amount of consumption units $c_t^f = \tilde{c}_t^w$. Inclusion of distributional conflicts between firm owners and workers would not fundamentally change the subsequent analysis, but would increase the complexity of the expressions.

The welfare costs of STW are defined as the difference between output minus disutility of work with and without the hours distortion effect of the STW system.

Definition 2, Welfare Costs of STW
The aggregate difference between output minus disutility of work with and without hours distortions of STW is defined as the welfare costs of STW $n_t \cdot \Omega_t$. Here, Ω_t can be denoted as:

$$
\Omega_t = \int_{\epsilon_{s,t}}^{\epsilon_{stw,t}} [\underbrace{y_t(\epsilon, h_t(\epsilon)) - v\big(h_t(\epsilon)\big)}_{\text{No Hours Distortion}} - \underbrace{y_t(\epsilon, h_{stw,t}(\epsilon)) + v\big(h_{stw,t}(\epsilon)\big)}_{\text{With Hours Distortion}}] dG(\epsilon)
$$

The optimal implementation of STW is significantly influenced by how the STW threshold and benefits impact the hours distortion problem associated with STW. The results are summarized in Lemma 1.

Lemma 1, Welfare Costs of STW
The welfare costs of STW $n_t \cdot \Omega_t$ are positive, and increase, if the eligibility condition gets looser or the STW benefits get more generous or both as:

$$\Omega_t \geq 0, \quad \frac{\partial \Omega_t}{\partial \tau_{stw,t}} > 0, \quad \frac{\partial \Omega_t}{\partial D_t} > 0, \quad \frac{\partial^2 \Omega_t}{\partial \tau_{stw,t} \partial D_t} > 0$$

PROOF: *Appendix 1.J*

First, the welfare costs must be positive. In the absence of STW, firms and workers would naturally choose the optimal number of hours worked. However, under STW, working hours are distorted downward, resulting in inefficiently low production levels.
Second, more generous STW benefits create stronger incentives for workers on STW to reduce their hours, thus severing the hours distortion effect.
Third, the welfare costs increase as the STW condition becomes looser. A looser eligibility condition allows more firms and workers to enter STW. Consequently, a greater number of them will choose sub-optimal low working hours, leading to a larger loss in output.
Finally, if the eligibility condition becomes looser and the STW benefits get more generous, the hours distortion effects are further exacerbated.

1.3.2 Optimal UI given STW

Before examining the joint determination of optimal UI and STW, it is helpful to explore the optimal UI system given a STW system that may not be optimally set, as described in Proposition 1. This analysis provides insight into the role of the STW system.

Similar to the social planner, the Ramsey planner seeks to insure workers against income losses due to job loss. Ideally, as shown by the social planner, the Ramsey planner would like to fully insure workers against income shocks $\tilde{c}^w = b$. However, it is well known that UI systems create fiscal externalities, leading to a wedge between $\tilde{c}^w$ and b, such that $\tilde{c}^w > b$. The wedge arises because UI systems enhance workers' outside option. As a result, workers and firms negotiate contracts with higher salaries and higher unemployment risk, leading to inefficiently low vacancy posting and inefficiently high separation rates. Strikingly, STW cannot help to eliminate inefficiently low hiring efforts. However, STW can theoretically eliminate all inefficient separations.

Proposition 1, Optimal UI given STW in Steady State
Suppose the economy is in its non-stochastic steady state and a non-optimized STW system exists where its STW threshold is set so that $\epsilon_{stw} > \xi$. Then, the optimal UI benefits can be determined by

$$\underbrace{(1-n)\cdot\frac{u'(b)-u'(\tilde{c}^w)}{u'(\tilde{c}^w)}}_{\text{Provide Income Insurance}} = L_V \cdot \underbrace{\left(-\frac{\partial f^{ge}}{\partial b}\cdot u\right)}_{\text{Reduction Hiring by UI}} + L_S \cdot \underbrace{\left(n\cdot g(\epsilon_s)\cdot\frac{\partial \epsilon_s^{ge}}{\partial b}\right)}_{\text{Increased Separations by UI}}$$

where L_V determines the social value from hiring one additional worker while L_S denotes the social value of the marginal match.

$$L_V = \underbrace{\frac{\eta-\gamma}{(1-\gamma)\cdot(1-\eta)}\cdot\beta\cdot\mathcal{J}}_{\text{Deviation from Hosios Condition}} + \underbrace{FE}_{\text{Fiscal Externality UI on Hiring}}$$

$$L_S = \underbrace{(1-f)\cdot FE}_{\text{Fiscal Externality UI on Separations}} \underbrace{-\tau_{stw}\cdot(\bar{h}-h_{stw}(\epsilon_s))}_{\text{STW subsidy}}$$

The Fiscal Externality FE of the UI system is defined as

$$FE = \frac{\beta}{1-\beta\cdot\big(1-G(\epsilon_s)\big)\cdot(1-f)}\cdot\left[b+\frac{1-n}{n}\cdot b\right]$$

PROOF: *Appendix 1.I*

To understand this outcome, we examine the efficiency costs of UI in detail. The term $\left(-\frac{\partial f^{ge}}{\partial b}\cdot u\right)$ captures the reduction in hiring efforts and thus job-finding rates in response to higher UI benefits, with the superscript *ge* indicating that all general equilibrium effects are considered. The term L_V represents the social value of hiring an additional worker. Together, these terms reflect the welfare loss from reduced hiring due to increased UI benefits.

The social value of hiring an additional worker depends on the deviation from the Hosios (1990) condition and the fiscal externality (FE) of UI on hiring. When workers' bargaining power is too high ($\eta > \gamma$), firms face elevated wages, reducing the private value of hiring below the social value. The UI system reinforces this effect, as unemployment benefits, b, and the avoidance of income tax in unemployment, $\frac{1-n}{n}\cdot b$, increase the worker's outside option. Both effects together indicate inefficiently low hiring rates, making reductions in hiring socially costly. Notably, the absence of STW in this expression shows that STW cannot realign private and social incentives for hiring. Any increase in STW's generosity is offset by a tax increase required to fund it, leaving the firm's value and hiring incentives unchanged.

The term $\left(n\cdot g(\epsilon_s)\cdot\frac{\partial \epsilon_s^{ge}}{\partial b}\right)$ represents the number of additional separations resulting from larger UI benefits. L_S denotes the social value of the marginal match to the economy. Together, the terms represent the welfare loss from inflated separations due to UI benefits.

The social value of the marginal match depends on the fiscal externality of the UI system and the STW benefits. The option to rely on the UI system enables workers to negotiate higher salaries in exchange for accepting greater unemployment risk. Consequently, matches may dissolve despite having a positive social value, making it costly to increase separations. STW can reduce the number of inefficient separations by reducing salary commitments for firms in bad times. The expression indicates that sufficiently high STW benefits can eliminate all inefficient separations, setting the social value of the marginal match equal to zero.

This result offers a new perspective on Cooper, Meyer, and Schott (2017)'s concern that STW reduces allocative efficiency by impeding the reallocation of workers between unproductive and productive firms. If STW benefits are set too generously, Cooper, Meyer, and Schott (2017)'s argument is valid, suggesting that the social value of maintaining a marginal match is negative $L_S < 0$; in this scenario, it would be socially efficient for the worker to leave an unproductive firm and seek new employment. However, if STW benefits are sufficiently low, this concern becomes less relevant, as the social value of the match remains positive $L_S > 0$.

1.3.3 Optimal STW Benefits, Eligibility Condition and UI Benefits

We are now well-prepared to analyze the joint determination of the STW system and UI system. We start by analyzing the optimal provision of STW benefits described in Proposition 2:

Proposition 2, Optimal STW benefits in Steady State
Consider the economy as previously described. Assume that it has converged to its non-stochastic steady state. Then, the optimal STW benefits τ_{stw} are determined by:

$$\underbrace{\tau_{stw} \cdot \left(\bar{h} - h_{stw}(\epsilon_s)\right)}_{\text{Net-Transfer STW}} = \underbrace{(1-f) \cdot FE}_{\substack{\text{A: Fiscal Externality UI on Separations} \\ > 0}} - \underbrace{L_S^\star}_{\substack{\text{B: Welfare Costs Rescuing Worker with STW} \\ > 0}} - \underbrace{\tilde{BE}}_{C\text{: Bargaining Effect}}$$

PROOF: *Appendix 1.I*

The determination of optimal STW benefits, denoted as τ_{stw}, follows a two-step procedure, as outlined in the left-hand side of the formula in Proposition 2. First, the Ramsey planner calculates the optimal net-transfer to the least productive matches. This reflects the amount of resources the planner intends to transfer to the marginal match.
Second, the planner must consider how the STW system influences the reduction in working hours by firms and workers, represented by $\bar{h} - h_{stw}(\epsilon_s)$. This reduction determines the amount of resources allocated to the match for given STW benefits τ_{stw}. Higher STW benefits result in a greater reduction in working hours, which in turn leads to a larger transfer of resources to the match. Using this information, the Ramsey planner adjusts the STW benefits to achieve the optimal net-transfer of resources.

The optimal net-transfer to the least productive matches comprises three components. Part A explains the rationale for STW's existence. Parts B and C discuss trade-offs of using STW. First, let us look at the reason to use STW:

Part A claims that the optimal transfer with STW has to be equal to the fiscal externality of UI benefits on separations $(1-f) \cdot FE$. From Proposition 1, we can infer that using this rule will eliminate all inefficient separations. But what exactly is the fiscal externality? The formula can be interpreted as the net present value of transfers a worker receives from the government from becoming unemployed (UI benefits + tax savings). Conversely, we can interpret the expression as the expected fiscal costs imposed on the government upon unemployment of the worker. In spirit, this is the rule Blanchard and Tirole (2008) and Cahuc and Zylberberg (2008) found in their analysis of optimal layoff taxes. To see this interpretation more clearly, Corollary 1 shows a simplified version of the fiscal externality under the assumption of no discounting:

Corollary 1, Fiscal externality UI on Separations under no Discounting
Assume no discounting $\beta = 1$. Then, the fiscal externality of UI on separations is

$$(1-f) \cdot FE = T \cdot b \quad \text{with} \quad T = \frac{1-f}{f}$$

where T denotes the duration a worker spends on the UI system.

PROOF: *Appendix 1.K*

Under no discounting, the fiscal externality of the UI system equals the expected duration a worker spends on the UI system when laid off, multiplied by the UI benefits. This is nothing else than the total costs an unemployed worker imposes on the UI system and has one important implication: larger UI benefits and a longer unemployment spell increase the social and fiscal costs of separations and thus optimal STW benefits.

Corollary 2, Higher Job-Finding rates decrease optimal STW Benefits
Assume risk neutrality and that the STW benefits are set optimally according to Proposition 2. Then, the welfare costs of the UI system and, thus, the STW benefits increase if the job-finding rate decreases:

$$\frac{\partial \tau_{stw}}{\partial f} < 0$$

PROOF: *Appendix 1.K*

The connection to the duration of the unemployment spell is particularly relevant for understanding how STW should optimally react to the business cycle. Section 4 demonstrates that recessions in the data are characterized by significant declines in the job-finding rate. In other words, unemployment spells significantly lengthen during recessions. These extended unemployment periods render separations costly from the perspective of the social planner, as workers do not contribute to production for an extended duration. However, workers do not fully internalize the additional social costs of separations, as they can stay on the UI system for longer. Consequently, recessions lead to an increase in inefficient separations. Optimal STW benefits must be raised to counteract this effect. This concept is formalized in Corollary 2.

Part B of the formula addresses Burdett and Wright (1989)'s objection that STW distorts working hours. The planner recognizes that the use of STW is costly as it downward distorts the hours choice in the economy. Consequently, the planner faces a trade-off between preventing socially undesirable separations and minimizing the distortions introduced by the STW system. To minimize distortions, STW benefits are reduced by the welfare costs of the hours distortion that an additional STW-supported rescue would impose.

Definition 3, Welfare Costs of Rescuing Worker with optimal STW
The welfare costs of rescuing a worker with STW are defined as:

$$L_S^\star = \frac{MWL_{\tau_{stw}}}{MMS_{\tau_{stw}}} = \frac{n \cdot \frac{\partial \Omega}{\partial \tau_{stw}}}{n \cdot g(\epsilon_s) \cdot \frac{\partial \epsilon_s}{\partial \tau_{stw}}}$$

Here $MWL_{\tau_{stw}}$ denotes the total welfare loss from marginally increasing STW benefits and $MML_{\tau_{stw}}$ denotes the total number of additional matches saved by marginally loosening STW benefits.

The welfare costs of rescuing an additional worker with STW, $L_S^\star$, depend on two factors. First, the total marginal welfare loss from increasing STW benefits, $MWL_{\tau_{stw}}$, which primarily depends on the size of the STW system and the number of workers it supports. Raising STW benefits distorts not only the working hours of newly rescued workers but also those already in the system. Consequently, rescuing a worker with STW is more costly in a larger STW system. Second, the cost depends on how many additional matches a marginal increase in STW benefits can rescue, $MMS_{\tau_{stw}}$. If STW is effective in saving many matches, the distortion per rescued match is low. However, if only a few matches can be saved, the social planner is discouraged from using STW to avoid excessive distortions. In the calibrated model, the welfare cost penalty is substantial, with STW reducing net transfers by approximately 30%.

An interesting effect related to the welfare costs of rescuing an additional worker with STW arises in the case of zero-STW (i.e., no work on STW). In this scenario, the additional distortions from STW benefits are minimal ($\frac{\partial \Omega}{\partial \tau_{stw}} \to 0$). As a result, the net transfer closely resembles the benefits provided without any hours distortion effects $L_S^\star = 0$.

Corollary 3, Looser Eligibility decreases optimal STW Benefits
Assume risk neutrality and that the eligibility condition D is exogenous such that $\epsilon_{stw} \geq \xi$. Otherwise, STW benefits are chosen according to Proposition 2. Then, a looser eligibility condition increases the welfare cost penalty of STW on the optimal net-transfer and, ceteris paribus, reduces the optimal STW benefits.

$$\frac{\partial \tau_{stw}}{\partial D} < 0$$

PROOF: *Appendix 1.K*

Corollary 3 clarifies the relationship between the eligibility condition and optimal STW benefits. It states that looser eligibility reduces the optimal level of STW benefits. This is because a looser eligibility condition distributes the distortionary effects of STW across a larger fraction of firms, increasing the welfare cost penalty $\frac{\partial^2 \Omega}{\partial \tau_{stw} \partial D} > 0$. The result has two important implications. First, it shows that we want to keep the eligibility condition strict to keep the distortionary effects low. Second, since the fraction of workers on STW typically rises during recessions, the STW penalty increases, potentially weakening STW's ability to stabilize the business cycle.

Part C addresses the issue that, under risk aversion in a Nash bargaining framework, salaries cannot be fully adjusted downwards. This becomes important as STW shifts contracts towards less unemployment risks but also lower salaries. Firms rely on salary cuts because the STW subsidy provides no net benefit due to the taxes required to fund it, while retaining unproductive workers remains costly. Then, the incomplete reduction of salaries dampens vacancy postings and amplifies separation incentives. The Ramsey planner accounts for these unwanted side effect by adjusting STW benefits downwards. Appendix 1.B offers a detailed discussion of the formula.

But why is it difficult to adjust salaries downwards? Under risk aversion, a key value of employment for workers is the protection it offers from falling back on UI benefits, where marginal utility of consumption is high. The larger the gap between the consumption equivalents provided by the firm and UI compensation, the greater the cost to the worker of falling back on the UI system and the larger the value of being employed. Equation 1.2.3 shows that firms exploit this effect during salary negotiations to reduce salaries. However, this fact also makes it difficult to cut salaries when external conditions demand it. When salaries are adjusted downward, the worker's match surplus declines more than proportionally because the value of being shielded from unemployment diminishes.

In the calibrated model, the effect reduces net-transfer by roughly 8% to get a sense of the size of the effect. The size itself depends on the degree of risk aversion and the difference between the consumption equivalent and UI benefits.

Corollary 4, STW has no direct Insurance Role
Suppose that $b = 0$ and the Hosios condition is met $\eta = \gamma$. Then, STW benefits are zero. This implies that STW itself is not used to provide direct income insurance.

PROOF: *Appendix 1.K*

One surprising effect of the analysis is that STW itself does not provide any direct income insurance. The sole reason for STW to exist is to counter the fiscal externality of the UI system. Thereby, STW indeed takes the role of an optimal layoff tax in the sense of Blanchard and Tirole (2008) and Cahuc and Zylberberg (2008). We can see this more clearly by setting unemployment benefits to zero and imposing the Hosios condition. In this case, optimal STW benefits are equal to zero. STW itself provides no direct income insurance as firms write a contract that insures workers against any idiosyncratic productivity shocks on the firm. Appendix 1.D derives the optimal layoff tax within the model and shows its similarity to the STW system.

Optimal Eligibility Condition. Proposition 2 highlights a significant drawback of STW: it distorts working hours. The eligibility condition shows how to manage the distortions in working hours.

The primary goal of STW is to provide support to firm-worker matches that would otherwise dissolve. This could be done with a simple subsidy scheme. However, a key challenge for governments with regular subsidy programs is the difficulty of identifying firms that truly need assistance. Therefore, the government would need to know their productivity, which is hard to observe in practice. STW addresses this issue by using reductions in working hours as a screening tool to reveal a firm's productivity level. The more unproductive a firm is at its current level, the more inclined it will be to reduce working hours. This allows the government to check productivity with the minimum hours reduction requirement as eligibility condition.[11]

To determine the optimal eligibility condition, we must therefore consider which productivity level the Ramsey planner wants to set as the threshold, denoted by $\epsilon_{stw,t}$, and adjust the minimum hours reduction threshold D_t accordingly. The logic behind the eligibility condition is to minimize the number of firms and workers accessing STW to keep distortions in working hours low. Proposition 3 states that this is achieved when the STW threshold equals the separation threshold of firms and workers without STW ($\epsilon_{stw} = \xi$, equation 1.3.1).

11. For a detailed discussion of this screening mechanism from a mechanism design perspective and its trade-off with working hours distortions, see Teichgräber, Žužek, and Hensel (2022).

Proposition 3, Optimal Eligibility Condition in Steady State

Consider the economy as previously described and assume that it has converged to its non-stochastic steady state. Then, the optimal eligibility condition $D = h_{stw}(\epsilon_{stw})$ is implicitly defined by the separation threshold of a firm without STW ($\epsilon_{stw} = \xi$).

$$\underbrace{S(\epsilon_{stw}) \; = \; y(\epsilon_{stw}, h(\epsilon_{stw})) - v\big(h(\epsilon_{stw})\big) + F + \frac{1-\eta \cdot f}{1-\eta} \cdot \frac{k_v}{q} \; = \; 0}_{\text{Joint Surplus without STW is zero}} \tag{1.3.1}$$

as long as the welfare costs of a looser eligibility condition are positive:

$$\underbrace{n \cdot \frac{\partial \Omega}{\partial \epsilon_{stw}}}_{\text{More hours distortion}} + L_V \cdot \underbrace{\left(-\frac{\partial f^{ge}}{\partial \epsilon_{stw}} \cdot u\right)}_{\text{Less hiring}} + L_S^\star \cdot \underbrace{\left(n \cdot g(\epsilon_s) \cdot \frac{\partial \epsilon_s^{ge}}{\partial \epsilon_{stw}}\right)}_{\text{More Separations}} \; \geq \; 0 \tag{1.3.2}$$

PROOF: *Appendix 1.I*

To better understand why we should set the eligibility condition equal to the separation threshold of firms and workers without access to STW ($\epsilon_{stw} = \xi$), we must examine the consequences of choosing a too loose ($\epsilon_{stw} > \xi$) or too strict ($\epsilon_{stw} < \xi$) STW threshold.

Let us begin with the case of a too strict eligibility condition ($\epsilon_{stw} < \xi$). In this case, there exist unproductive matches that are allowed onto the STW system while more productive matches are not, causing the latter to dissolve. Rescuing less productive matches while allowing more productive matches to dissolve would clearly be inefficient. To avoid such inefficiencies, the STW threshold needs to be set at least as loose as the separation threshold without STW ($\epsilon_{stw} \geq \xi$).

The costs of choosing the eligibility condition too loose $\epsilon_{stw} > \xi$ are described in equation 1.3.2. The main costs to consider are the additional welfare costs associated with the STW system ($\frac{\partial \Omega}{\partial \epsilon_{stw,t}} > 0$). The hours distortion effect spreads among more firms without saving additional workers. Furthermore, a looser eligibility condition reduces hiring $\left(-\frac{\partial f^{ge}}{\partial \epsilon_{stw}} \cdot u\right)$ and increases separation incentives $\left(n \cdot g(\epsilon_s) \cdot \frac{\partial \epsilon_s^{ge}}{\partial \epsilon_{stw}}\right)$ through distortion of working hours. STW combats the additional inefficient separations by raising STW benefits. Remember that $L_S^\star$ denotes the welfare costs of rescuing one additional worker with STW.
At first sight, easier access to the STW system should increase the surplus of firms and workers. However, the effect is nullified by an equivalent increase in the income tax. Even worse, the spread of the hours distortion effect reduces the expected output of firms, leading to a fall in the joint surplus. Hence, it must be optimal to set the eligibility condition at least as strict as the separation threshold

of firms and workers without STW ($\epsilon_{stw} \leq \xi$). This property is later referred to as the "no-windfall effect" condition.

Equation 1.3.2 also describes the sufficient condition for the eligibility condition. For $\eta \geq \gamma$, the condition is unambiguously positive and fulfilled. However, even if $\eta < \gamma$, it is hard to imagine that the condition would not be fulfilled. It would mean that firms post so many vacancies that it would be optimal to sacrifice output to reduce the profits of firms and thus decrease vacancy postings. Using STW in that case seems unreasonable.

Combining both statements, we can conclude that if the eligibility condition is too strict, the planner will be unable to save some firms worth saving. If it is too loose, it will exacerbate the distortionary effects of STW. Therefore, it is optimal to set the separation threshold of firms and workers without access to STW equal to the STW threshold, denoted as $\epsilon_{stw} = \xi$. In the end, only matches that would dissolve otherwise should be allowed on the STW system.

In contrast to the optimal STW benefits, it is not directly obvious from the expression of the optimal eligibility condition how the eligibility condition interacts with the unemployment insurance system and the STW benefits. Its connection is highlighted in Corollary 5:

Corollary 5, Relationship Eligibility Condition to UI and STW benefits
Suppose that the sufficient condition for the eligibility condition holds and that $S'(\epsilon_{stw}) > 0$. Then, ceteris paribus, the optimal eligibility condition becomes looser when UI benefits increase:

$$\frac{\partial D}{\partial b} > 0$$

Further assume that the government wants to implement the eligibility threshold ϵ_{stw}. Then the government has to react with a stricter eligibility to implement larger STW benefits.

$$\frac{\partial D}{\partial \tau_{stw}} < 0$$

PROOF: *Appendix 1.K*

When UI benefits increase, the optimal eligibility condition has to be loosened. Any increase in UI benefits raises the worker's outside option, leading to higher separation rates in firms without access to STW. If the eligibility condition isn't adjusted, these matches are destroyed even though they could have been saved with STW. To prevent unnecessary separations, the eligibility condition must be loosened. The expression becomes important when considering the interaction of optimal UI benefits with the STW system.[12]

Any increase in STW benefits will necessitate tightening the optimal eligibility condition. Any rise in STW benefits encourages firms to reduce working hours to qualify for STW and enter the system. To prevent matches that could survive without STW from accessing the program, the eligibility condition must be made stricter. This ensures that only the firms truly in need of support receive it, minimizing working hours distortions. The channel becomes important when considering how to react to the adjustment of STW benefits over the business cycle.

Corollary 6, STW saves fiscal costs
Suppose $\beta = 1$ and $\eta \geq \gamma$. Then, ceteris paribus, combining the UI system with an optimal STW system is less fiscally expensive than having a UI system only.

PROOF: *Appendix 1.K*

Corollary 6 claims that ceteris paribus, combining the UI system with an optimal STW system is less fiscally expensive than having a UI system only. To see this, let τ_J^{stw} denote the taxes needed to support a UI system with a hypothetical optimal STW system and τ_J^b the taxes needed with a UI system only. Inserting the optimal eligibility condition and the optimal STW benefits into the budget constraint gives:

$$\tau_J^{stw} \; < \; G(\epsilon_s) \cdot T \cdot b + \big(G(\xi_s) - G(\epsilon_s)\big) \cdot \big(T \cdot b - \tilde{\Omega} - \tilde{BE}\big) \; < \; G(\xi_s) \cdot T \cdot b \; = \; \tau_J^b$$

Note that $G(\xi)$ denotes the probability of separations without access to an STW system, and $G(\epsilon_s)$ represents the probability with access. The STW system lowers income taxes for two reasons. First, the optimal eligibility threshold $\epsilon_{stw} = \xi$ ensures that only firms facing layoffs can enter the STW system. Thus, the government supports, ceteris paribus, the same fraction of workers under both a UI-only system and a UI system with optimal STW system. The difference is that the costs of keeping a worker on the STW system are lower than on the UI system, as the

12. As necessary condition, $S'(\epsilon_{stw}) > 0$ states that the joint surplus of a match at the eligibility condition must increase. Theoretically, a loosening of the eligibility condition can increase the distortionary effects of the STW system and decrease the continuation value of the match. However, quantitatively, this is very unlikely to exceed the direct effect of higher productivity on the joint surplus.

welfare cost penalty of STW and the bargaining effect reduce optimal STW benefits below the fiscal costs of separations. Consequently, it is fiscally cheaper to retain workers through an optimal STW system than to allow them to enter the UI system.

Optimal UI benefits. Finally, we need to characterize optimal UI benefits, taking into account their impact on the optimal STW benefits. Proposition 4 offers an expression.

Unaltered, the Ramsey planner faces a trade-off between providing additional income insurance for workers and the economic distortions introduced by higher UI benefits. The negative impact of UI benefits on vacancy posting has not changed.

Compared to the case with a given STW system, the interpretation shifts slightly. An increase in UI benefits still inefficiently raises separations and is captured by $\left(n \cdot g(\epsilon_s) \cdot \frac{\partial \epsilon_s^{ge}}{\partial b}\right)$. However, under an optimal STW policy, the government can offset this by rescuing matches through raising STW benefits. This approach is costly, as increasing STW benefits also amplifies the system's distortionary effects. These costs are captured by L_S^*, which represents both the expense of rescuing an additional worker with STW and the social value of the marginal match.[13] Since L_S^* is positive, the social planner chooses not to prevent all inefficient separations.

Proposition 4, Optimal UI with optimal STW in Steady State
Suppose the economy is in its non-stochastic steady state and the STW system is set optimally. Then, the optimal UI benefits can be determined by

$$\underbrace{(1-n) \cdot \frac{u'(b) - u'(\tilde{c}^w)}{u'(\tilde{c}^w)}}_{\text{Provide Income Insurance}} = L_V \cdot \underbrace{\left(-\frac{\partial f^{ge}}{\partial b} \cdot u\right)}_{\text{Less Hiring}} + L_S^\star \cdot \underbrace{\left(n \cdot g(\epsilon_s) \cdot \frac{\partial \epsilon_s^{ge}}{\partial b}\right)}_{\text{More Separations}} + L_{STW}^\star \cdot \underbrace{\left(n \cdot g(\epsilon_{stw}) \cdot \frac{\partial \epsilon_{stw}^{ge}}{\partial b}\right)}_{\text{More Workers on STW}}$$

PROOF: *Appendix 1.K*

13. The costs of rescuing a worker with STW must be equivalent to the social value of the marginal match as the Ramsey planner will want to rescue workers on STW until the additional social benefit of rescuing them is equivalent to the additional social costs of the STW system.

Finally, the government must account for its impact on the optimal eligibility condition. When UI benefits rise, firms without access to STW will separate at higher productivity levels. In response, the government loosens the eligibility condition, increasing the share of workers in the STW system. The resulting welfare costs per additional worker are captured by $L^{\star}_{STW}$ and are expressed in definition 4:

Definition 4, Welfare Costs of Adding a Worker to STW
The welfare costs of adding a worker to the STW system by loosening the eligibility condition can be defined as follows:

$$L^{\star}_{STW} = \frac{MWL_{\epsilon_{stw}}}{MMS_{\epsilon_{stw}}} = \frac{n \cdot \frac{\partial \Omega}{\partial \epsilon_{stw}}}{n \cdot g(\epsilon_{stw})}$$

Here $MWL_{\epsilon_{stw}}$ denotes the total welfare loss from marginally loosening the eligibility threshold and $MMS_{\epsilon_{stw}}$ denotes the total number of additional matches on STW by marginally loosening the eligibility threshold.

The welfare costs of a looser eligibility condition per additional worker on STW, $L^{\star}_{STW}$, depend on two factors. First, $MWL_{\tau_{stw}}$ represents the total welfare loss from a marginal loosening of the STW threshold. A more generous STW system results in greater welfare losses from relaxing eligibility. This loss is divided by $MMS_{\epsilon_{stw}}$, the number of additional workers entering the STW system when the eligibility threshold is loosened.

1.4 Calibration and Solution Procedure

This section calibrates the model to the US economy, using a period length of one month. For the business cycle analysis, I allow only the STW system to adjust, while keeping the unemployment insurance (UI) system exogenously set. The baseline model used here incorporates wage rigidity and an exogenously determined UI system but excludes any STW mechanism. The choice of US data is particularly advantageous because, historically, the US has not implemented a nationwide STW system, thereby ensuring that the data is unaffected by such a system's influence. This allows for a clearer analysis of the model's implications.

Data used for Calibration. I calibrate the model to data from 1952:I to 2020:I. The unemployment rate is taken from the U.S. Bureau of Labor Statistics. Following Shimer (2005), the job-finding rate and separation rate are calculated using data on the absolute number of unemployed u^a, newly unemployed

[14] u^s and employed e^a workers from the U.S. Bureau of Labor Statistics: $f_t = 1 - \frac{u^a_{t+1} - u^s_{t+1}}{u^a_t}, s_t = \frac{u^s_{t+1}}{e_t \cdot (1 - \frac{1}{2} \cdot f_t)}$. For vacancies, I use the composite help-wanted index from Barnichon (2010). Average weekly hours $\bar{h}_t/4 = E_t[h_t(\epsilon)|\epsilon \geq \epsilon_{s,t}]/4$ and average labor productivity $p_t = E[y_t(\epsilon)|\epsilon \geq \epsilon_{s,t}]$ are retrieved for the non-farm business sector from the U.S. Bureau of Labor Statistics.

The business cycle properties are reported in Table 1.4.1. Following Shimer (2005), the table reports log-deviations from an HP-trend with smoothing parameter 10^5. The properties of the business cycle data are well known. Vacancies, unemployment, and labor market tightness are very volatile. The job-finding rate and the average hours worked are pro-cyclical, while separations are counter-cyclical. Separations are less volatile than the job-finding rate.

Table 1.4.1. Business Cycle Properties US Data

		v	f	ρ	u	θ	$\bar{h}$	p
Standard Deviation		20.13	14.31	8.2	20.49	39.67	0.81	1.91
Autocorrelation		0.95	0.95	0.77	0.95	0.96	0.92	0.9
	v	1	0.85	-0.55	-0.92	0.98	0.55	0.19
	f	-	1	-0.29	-0.93	0.91	0.38	0.09
	ρ	-	-	1	0.6	-0.59	-0.63	-0.4
Correlation	u	-	-	-	1	-0.98	-0.55	-0.23
	θ	-	-	-	-	1	0.57	0.22
	$\bar{h}$	-	-	-	-	-	1	0.46
	p	-	-	-	-	-	-	1

Notes: The table lists the second moments of the data. u, v, f, $\bar{h}$ and $G(\varepsilon_s)$ are expressed as quarterly averages of monthly series. p is the seasonally adjusted average labor productivity in the non-farm business sector. All variables are reported as log-deviations from an HP trend with smoothing parameter 10^5.

Calibrated Parameters. Table 1.4.3 summarizes the chosen parameter values, and Table 1.4.4 the respective business cycle properties of the model.

Following Jung and Kuester (2015), I set the discount factor to $\beta = 0.996$. As target steady states, I choose the monthly steady state job-finding rate of $f = 0.41$ and separation rate $\rho = 0.03$ from the data. To implement the job-finding rate, I set vacancy posting costs to $k_v = 0.209$. To implement the separation rate, the strength of the resource cost shock is set to $c_f = 14.214$. The matching efficiency parameter $\bar{m} = 0.383$ is determined by targeting a monthly vacancy filling rate of $q = 0.338$. This is the monthly equivalent of the quarterly job-filling rate of

14. Unemployed for less than 5 weeks

0.71 reported in Haan, Ramey, and Watson (2000). I set the bargaining power of the worker to $\eta = 0.65$, which is, according to Petrongolo and Pissarides (2001), within the reasonable set of parameter estimates. In order to ensure that inefficiencies in the steady state are only driven by the UI system, the Hosios-Condition (see Hosios, 1990) is implemented by setting the elasticity of the matching function with respect to unemployment equal to the bargaining power of the firm: $\gamma = \eta$. The unemployment benefits are set to $b = 0.45$, which ensures a replacement rate of 45% of the wage, which is the empirical value reported by Engen and Gruber (2001).

Table 1.4.2. Parameters

Parameter	Description	Value
Preferences		
β	Discount rate	0.996
ψ	Inverse Frisch-elasticity	1.5
ϕ	Coefficient relative risk aversion	2
Vacancies, Matching, Bargaining		
k_v	Vacancy posting costs	0.209
$\bar{m}$	Matching parameter	0.383
γ	Elasticity matching function w.r.t. unemployment	0.65
η	Bargaining power worker in steady state	0.65
γ_w	Degree of cyclicality of bargaining power of worker	15.779
Production and Separations		
α	Labor elasticity production function	0.65
μ_a	Mean aggregate productivity	1.0
$\sigma_a \cdot 100$	s.d. aggregate productivity	0.25
ρ_a	Autocorr. productivity shock	0.985
μ	Stears mean of lognormal distribution	0.094
σ	Stears steering variance of lognormal distribution	0.12
c_f	Strength resource cost shock	14.214
F	Separation costs	1.441
Labor Market Policy		
b	UI benefits	0.45
$\bar{h}$	"Normal" hours worked	0.839

Table 1.4.3. Parameters

The parameter $\bar{h}$ represents the mean hours worked in a firm and is set to its steady state value in the baseline economy: $\bar{h} = 0.839$. Similar to Christoffel and Linzert (2010), I set the labor elasticity of the production function to $\alpha = 0.65$. The disutility of work has the common functional form of v(h) $= \frac{h^{1+\psi}}{1+\psi}$, $\psi > 0$. The utility function is assumed to be CRRA $u(\tilde{c}_t) = \frac{\tilde{c}_t^{1-\phi}-1}{1-\phi}$ with constant relative

risk aversion parameter $\phi = 2$. Following Domeij and Floden (2006), I set the Frish-elasticity to 0.66, which implies $\psi = 1.5$. As Krause and Lubik (2007), I set the parameter for the variance of the log-normal distribution of the idiosyncratic shock to $\sigma = 0.12$. In order to normalize the wage to 1, the parameter that steers the mean of the log-normal distribution is set to $\mu = 0.094$.
In order to reach a standard deviation (s.d.) of 0.02 of labor productivity over the business cycle, I set the standard deviation of the aggregate productivity shock to $\sigma_a = 0.0025$ and follow Jung and Kuester (2015) in setting the autocorrelation to $\rho_a = 0.985$. To match the standard deviation of the job-finding rate of 0.0143, I set the degree of cyclicality of bargaining power of the firm to $\gamma_w = 15.5$. Further, I target a standard deviation of the separation rate of 0.0082, by setting the separation costs to F = 1.441.[15]

Table 1.4.4. Business Cycle Properties Baseline Model

		v	f	ρ	u	θ	$\bar{h}$	p
Standard Deviation		19.94	14.31	8.2	21.37	40.88	0.76	1.91
Autocorrelation		0.95	0.97	0.97	0.98	0.97	0.97	0.97
	v	1	1	-0.99	-0.98	1	1	1
	f	-	1	-1	-1	1	1	1
	ρ	-	-	1	1	-1	-1	-1
Correlation	u	-	-	-	1	-1	-1	-1
	θ	-	-	-	-	1	1	1
	$\bar{h}$	-	-	-	-	-	1	1
	p	-	-	-	-	-	-	1

Notes: The table reports the second moments of the model. As in the data of Shimer (2005), all variables are quarterly averages of monthly series and reported as log-deviations. p denotes the average output per person, that is $p = E[y_t(\varepsilon)|\varepsilon \geq \varepsilon_{s,t}]$.

Compare the business cycle facts from the baseline economy from table 1.4.4 to the US business cycle facts in table 1.4.1. With the calibration chosen above, we can closely replicate the business cycle properties from the data. Note that a large chunk of the fluctuations is driven by our assumption of the procyclical bargaining power of the firms. Therefore, a lot of these fluctuations must be inefficient, which gives room for the policymaker to intervene.

15. This is consistent with the value used in Silva and Toledo (2009) for the US economy as severance payments plus the wasteful separation costs account for roughly 8 weeks of the annual salary of a worker. Silva and Toledo (2009), respectivley Ahr and Ahr (2000) report that turnover costs vary between 25% and 200% of the annual salary. In this model, turnover costs would be at the lower end, with roughly 25% accounting for recruitment and wasteful separation costs as well as severance payments.

To solve the model, I rely on first-order perturbation using the code of Schmitt-Grohe and Uribe (2004) based on the symbolic toolbox of Matlab.

1.5 Optimal STW Policy in Recessions - Current UI

In the following sections, I will explore two possible scenarios for the implementation of STW policy. This section will examine how STW reacts within an economy under the current calibrated UI insurance system. The analysis will shed light on how optimal STW policy would function when introduced into the present economic system. In the subsequent section, I will analyze the implications of optimal STW policy when paired with an optimal UI system in steady state. The section discusses only the results for the recession. A table with the steady state value can be found in Appendix 1.E.

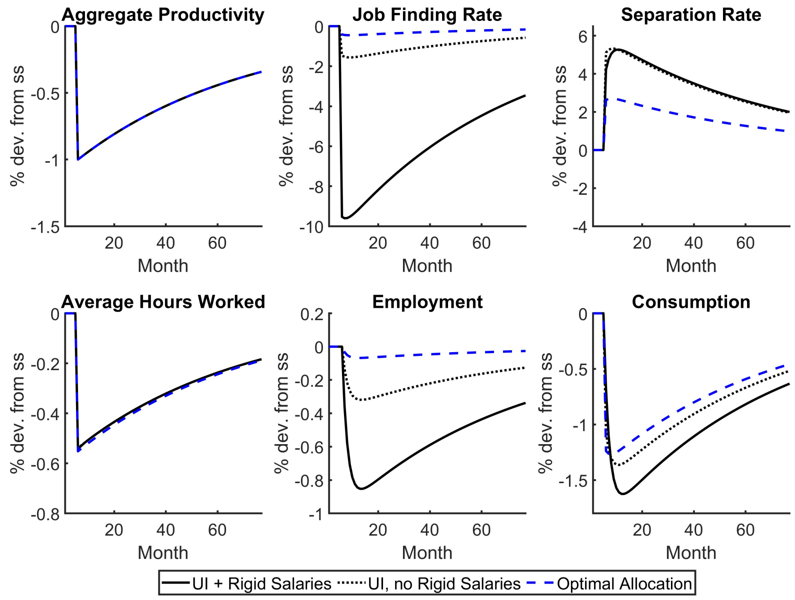
1.5.1 Inefficiencies in the Business Cycle

The business cycle is driven by real productivity shocks. Figure 1.5.1 shows the response of the planner economy, an economy with a UI system, and an economy with both a UI system and rigid salaries to a 1% negative aggregate productivity shock. We refer to the economy with UI system and rigid salaries as the baseline economy. Comparing the baseline economy and the economy with the UI system only to the planner allocation will give us a sense of the business cycle's inefficiencies.

Generally speaking, a reduction in aggregate productivity due to a negative productivity shock reduces the joint surplus of firm-worker matches. As a result, firms will be less willing to pay the vacancy posting costs. The number of vacancies and the job-finding rate fall. Furthermore, the reduced productivity implies that a larger fraction of firms and workers generate a negative surplus, leading to a larger separation rate. A reduction in the job-finding rate combined with an increase in the separation rate drives down employment. Output and consumption fall mainly due to the reduction of aggregate productivity.
Note that these fluctuations can be efficient to some extent (see Figure 1.5.1, blue line). The social planner would also increase separations to get rid of unproductive matches (cleansing effect) or reduce vacancy posting efforts if new workers add less to the output.
However, these fluctuations are inefficiently amplified by the existence of the UI system and rigid salaries.

The fall in the job-finding rate increases the distortionary effects of the UI system (see Corollary 2). Workers need more time to find a new job, which drives down the workers' outside option. The UI system can partially offset the effect, since

Figure 1.5.1. Inefficiencies in the Business Cycle



Notes: The figure shows impulse response functions for a 1% negative productivity shock. The black line shows the response of the baseline economy, that is, the economy with the distortionary effects of the UI system and rigid salaries as inefficiencies, but without STW system. The black dashed line shows the reaction of an economy without rigid salaries but with the distortionary effects of the UI system. The blue dashed line shows the response of the planner economy.

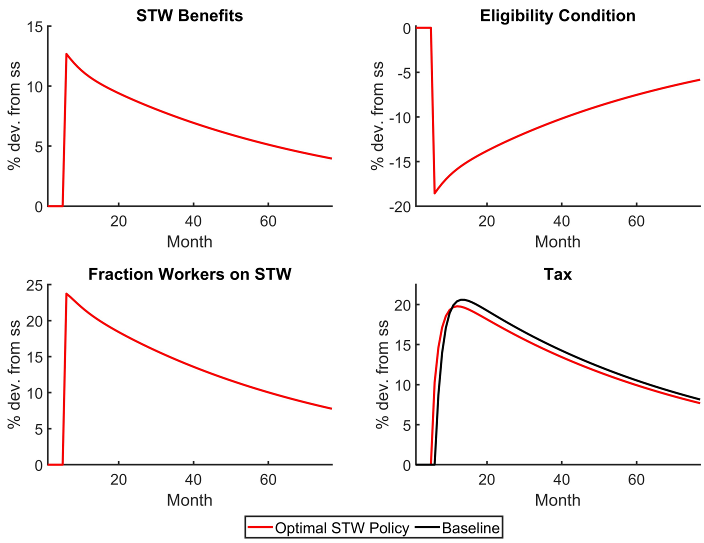
increasing the worker's unemployment spell also means prolonged payments from the UI system. This keeps the outside option of the workers and thus wages up, shrinks job postings, and inflates separations (see Figure 1.5.1, black vs. black dashed line). Furthermore, increased unemployment drives up the fiscal costs of the UI, forcing the government to increase taxes, amplifying the effect.

Rigid salaries further exacerbate inefficiencies in the business cycle (see Figure 1.5.1, black vs. dashed black line). In a recession, rigid salaries lead to a deviation from the Hosios-condition. Firms' share of the joint surplus falls, which makes them cut vacancies to save on vacancy posting costs. As a result, the job-finding rate plummets, leading to a large increase in undesirable unemployment and an aggravation of the distortionary effects of the UI system.

1.5.2 Optimal STW Policy

The Ramsey planner has three key objectives. First, he aims to minimize the efficiency costs of the business cycle, ensuring that the consumption response aligns as closely as possible with that of the social planner. Second, the Ramsey planner wants to smooth the consumption of households in the economy, which can be

Figure 1.5.2. Optimal STW Policy - Instruments



Notes: The figure shows the response of the economy with optimal STW system (red line) to a 1% negative productivity shock, and compares it to the baseline economy (black line).

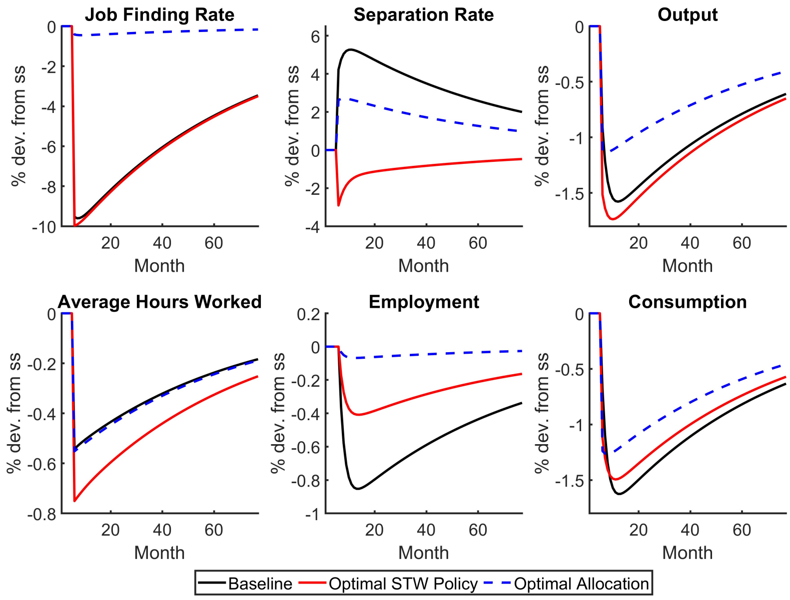
achieved through two channels: the intensive margin, via UI benefits, and the extensive margin, by reducing unemployment itself.

$$(1 - n) \cdot u(b) + n \cdot u(\tilde{c}) \tag{1.5.1}$$

Firms provide income insurance to workers as long as they are employed. Since the Ramsey planner can only control the STW system in the business cycle to provide equity, he aims to stabilize employment beyond sole efficiency considerations. Finally, the Ramsey planner aims to keep working hours as close to the social planner's optimal level as possible, ensuring that disutility of work is used efficiently in production. It is important to note that STW is the only reason why working hours can deviate from the socially optimal ones. In the following, we can judge the effectiveness of STW in stabilizing the business cycle by looking at the responses of average hours worked, employment, and consumption in the economy.

Figures 1.5.2 and 1.5.3 show the optimal response of the STW system and the economy to a negative 1% productivity shock. A striking result is that STW cannot stabilize the job-finding rate, despite the fact that the job-finding rate deviates significantly from its optimal level. This outcome is primarily due to the budget balance assumption and aligns with Proposition 1. Firms do not know at the beginning of the period whether they will receive STW benefits or be required to contribute to the system. This uncertainty nullifies any positive impact that a rise

Figure 1.5.3. Optimal STW Policy - Allocation



Notes: The figure shows the optimal response of the STW system (red line) to a 1% negative productivity shock and compares it to the baseline economy (black line) and the social planner (blue dashed line).

of STW benefits might have on vacancy postings.
The inability to stabilize the job-finding rate in recessions might seem surprising, particularly given that studies like Balleer et al. (2016) view STW as a tool that could potentially stabilize the job-finding rate. But the result is robust beyond the balanced budget assumption. Even if the Ramsey planner had the ability to adjust STW to influence the job-finding rate, he would not do so because the distortionary effects on working hours would outweigh any potential benefits. For a detailed discussion, look at Appendix 1.C.

Since STW cannot prevent the decline in the job-finding rate, it becomes increasingly difficult for workers to find new employment during recessions. As unemployment spells lengthen, the social costs of separations rise. In response, the Ramsey planner increases STW benefits, as suggested in Corollary 2, mirroring actual STW policy. This adjustment reduces separation rates. In fact, the Ramsey planner even lets separations fall during recessions to counterbalance the sharp drop in the job-finding rate, which would otherwise have a severe impact on employment.[16]

16. These results correspond surprisingly well to what actually happened in the Covid-19 crisis in Germany. Germany significantly increased the generosity of its STW system during the

In contrast to commonly applied STW policies, the eligibility condition does not need to be loosened during a recession; in fact, it needs to be tightened (see Figure 1.5.2). The increase in STW benefits incentivizes firms to reduce working hours, making more of them eligible for the system. To prevent these newly eligible firms from entering STW and profiting from windfall gains, the eligibility condition must be tightened, as outlined in Corollary 5. This ensures that only firms in genuine need of support benefit from the system.

Nonetheless, stabilizing the economy through oversteering the separation rate, increasing STW benefits, and expanding the fraction of workers on STW comes with two costs. First, by keeping unproductive workers employed, STW hinders the cleansing effect of recessions and leads to a further decline in average firm productivity. Second, larger STW benefits and a higher fraction of workers on STW amplify the distortionary effects of the system, destabilizing average hours worked. These factors explain why the optimal STW policy does not stabilize output as shown in Figure 1.5.3.

Fortunately, the impact on consumption is closer to the planner economy. By hoarding labor, STW reduces the need for reallocation of workers via the labor market, resulting in lower costs associated with firing and recruiting workers. Almost 30% of inefficient fluctuations in consumption can be eliminated by STW, despite distorting working hours.

One important note for policymakers is that using STW optimally over the business cycle is fiscally not more expensive than a system without STW (see figure 1.5.2). Since STW keeps employment stable, it prevents workers from entering the UI system, keeping its costs down in recessions. After the recession, the STW system should revert to its baseline values.

In conclusion, STW can be used to stabilize employment and consumption, but not output in recessions. However, its inability to influence the job-finding rate and its destabilization of average hours worked prevent it from reaching the planner allocation.

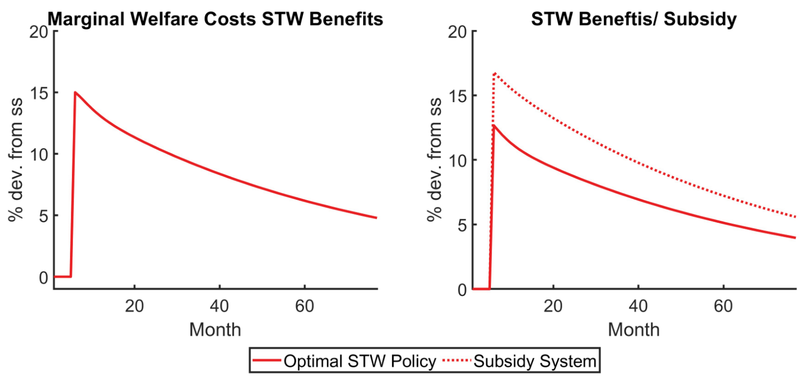
1.5.3 Welfare Costs of STW and Optimal STW Policy Adjustment

Proposition 2 states that the optimal STW benefits depend on two main effects. First, the reason for STW to exist is to offset the distortionary effects of the UI system on separations. The last section discussed its implications for the business cycle extensively. The second part of the formula looks at how the distortionary effects of the STW system influence the optimal provision of benefits. This section

pandemic. Weber and Röttger (2022) find that the separation rate fell even below the level before the crisis. Furthermore, new hires decreased.

investigates its impact on the business cycle. Figures 1.5.4 and 1.5.5 compare the response of the optimal STW system to a hypothetical subsidy system that pays benefits exactly when STW does, but without distorting working hours. This will give us a measure of how important working hours distortions are for the stabilization of the business cycle. Note that such a subsidy system is not implementable in practice as it would require knowledge of the productivity states in the economy. The use of STW is to elicit productivity via connecting the subsidy to a reduction of working hours, as exploited in Proposition 3.

Figure 1.5.4. Optimal STW Policy, Influence of Hours Distortion - Instruments

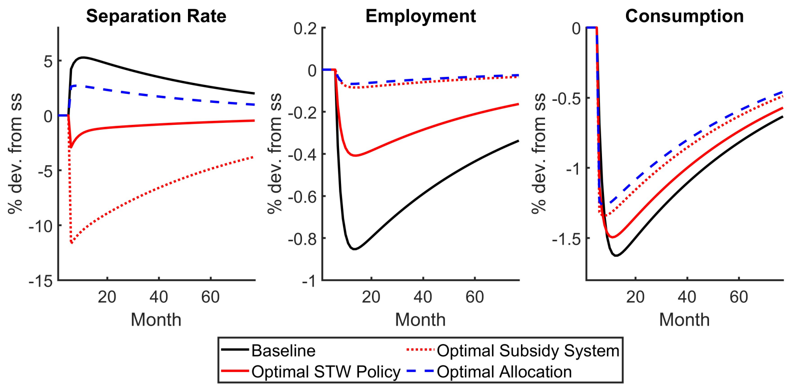


Notes: The figure compares the optimal adjustment of the STW system (solid line) to a hypothetical subsidy system without hours distortion (dashed line) to a 1% negative productivity shock.

Corollary 3 suggests that the distortion of working hours — and thus the welfare cost penalty of STW — must increase during a recession as the fraction of workers on STW expands (see Figure 5). Figure 1.5.4 shows that the marginal welfare costs of STW benefits $\frac{\partial \Omega_t}{\partial \tau_{stw,t}}$ rise indeed significantly in a recession. As more workers take advantage of the STW system during economic downturns, any increase in the STW benefits affects a larger number of firms, which significantly raises the overall distortionary effects and costs of STW benefits. Faced with higher marginal welfare costs of STW benefits, the planner responds by limiting the increase in STW benefits during recessions. The red dashed line shows how much the optimal transfer would increase if STW benefits did not have distortionary effects.

Consequently, the planner is unable to implement the optimal separation rate. Figure 1.5.5 shows that in the absence of distortions, the planner could reduce separations by more than 10 percent, achieving the optimal employment level. This, in turn, brings the consumption response close to the optimal level. However, full stabilization of consumption remains unattainable, as stabilizing employment solely by reducing separations leads to a decrease in the mean productivity of the economy. In contrast, when distortions are present, the planner allows separations to drop only by 3%, resulting in more pronounced fluctuations in employment. This, coupled with greater fluctuations in working hours, limits the planner's ability to implement the optimal consumption response.

Figure 1.5.5. Optimal STW Policy, Influence of Hours Distortion - Allocation



Notes: The figure shows the impulse response function of an economy with optimal STW system (solid red line) to a hypothetical subsidy system without hours distortion (dashed red line) for a 1% negative productivity shock. Further, it shows the response of the social planner economy (blue dashed line) and the baseline economy (black line).

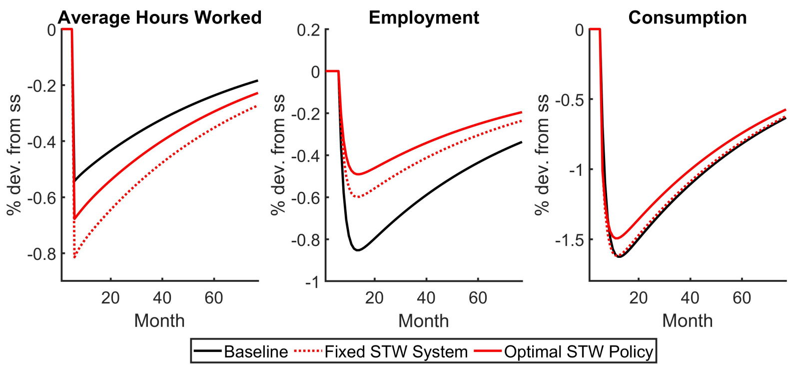
From this section, we can conclude that the distortion of working hours significantly reduces STW's capacity to stabilize both employment and consumption during recessions. In case of employment, it is responsible for almost all inefficient fluctuations, while in case of consumption, it is responsible for almost 80%, while roughly 20% stems from a reduction in average productivity.

1.5.4 Importance of a dynamic STW system

The last sections have shown that optimal STW policy requires the eligibility condition and STW benefits to be adjusted in the business cycle. Balleer et al. (2016) argue that STW acts as an automatic stabilizer. The system can stabilize employment and consumption without the need to be adjusted. Therefore, the question can be raised, how important a dynamic STW system is for business cycle stabilization. To answer the question, I implement the optimal STW system in steady state, but keep the eligibility condition and STW benefits constant in a recession. Figure 1.5.6 shows the results.

Employment is still stabilized compared to the baseline economy. During recessions, firms and workers opt to reduce working hours, allowing them to qualify for additional STW benefits, which increases the net transfer to the least productive matches. This mechanism mitigates the decline in the joint surplus of

Figure 1.5.6. Fixed STW System



Notes: The figure shows the impulse response function of an economy with fixed STW system (dashed line, red) to the optimal STW system (red line) and the baseline economy (black line) for a 1% negative productivity shock. The fixed STW system sets STW optimally in steady state but does not let the system adjust over the business cycle.

matches, leading to fewer separations. [17] While employment is stabilized, it is less effectively so than under optimal STW policy, as benefits cannot be increased to reduce separations during downturns.

However, despite not adjusting STW benefits, average hours worked decline more in the fixed STW system. This is due to the fact that the government does not tighten the eligibility conditions. Firms and workers that could continue working without STW opt to enter the system.[18] These windfall effects intensify the distortionary impact of STW and show how important the optimal implementation of the eligibility condition is. The combination of less stabilized employment and greater fluctuation in average working hours undermines the STW program's ability to stabilize consumption during recessions.

In conclusion, the analysis supports the findings of Balleer et al. (2016) that STW automatically stabilizes employment. It does not find stabilization of consumption. The main reason is that the eligibility condition of Balleer et al. (2016) does not allow for windfall effects.

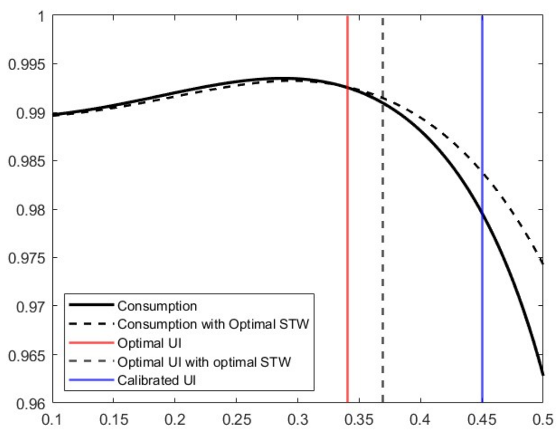
17. Note that this mechanism differs from that in Balleer et al. (2016). Their model features an inflexible intensive margin. STW allows for a reduction in hours and, thus, the wage bill. In recessions, firms can lower hours worked in response to a negative productivity shock. Thanks to the STW system, they consolidate wage expenditures and stabilize separations.

18. After a negative productivity shock, the job-finding rate decreases, and workers' unemployment prospects worsen, making it more challenging to find new employment. Consequently, workers are more inclined to stay with their current employer, working fewer hours for a reduced salary. However, due to the reduced working hours, they become eligible for the STW system despite not being in need of support, making their entrance into STW inefficient (see also Corollary 5).

1.6 Optimal STW Policy in Recessions - Optimal UI in SS

The previous section assumed a 45% replacement rate of the UI system, as in the data. However, when calculating the optimal UI system in the model, we see that the optimal replacement rate is at roughly 34% as shown in Figure 1.6.1.

Figure 1.6.1. Consumption and UI



Notes: The figure shows the steady state consumption levels in response to different UI replacement rates when there is only a UI system in place and when STW optimally reacts to the UI system.

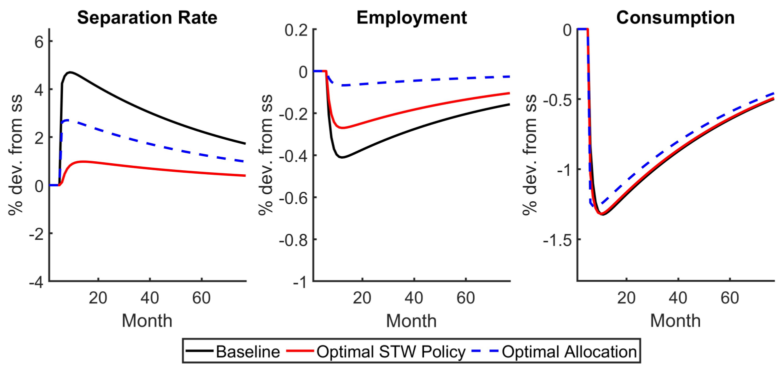
Interestingly, the optimal replacement rate is close to maximizing the economy's consumption levels. High UI benefits traditionally result in excessive unemployment, reducing overall production. Surprisingly, very low UI benefits can also lead to reduced consumption.

The reasoning is as follows: when UI benefits fall, unemployment becomes extremely costly for risk-averse workers. As a result, the value of maintaining employment rises for workers. In the Nash-Bargaining, workers prioritize job security, opting for contracts with low unemployment risk but also lower salaries. Equation 1.2.3 shows that firms gain a significant premium when workers are poorly insured against unemployment.

The issue is that unemployment can be productive in the economy. It allows workers to leave unproductive firms and reallocate to more productive ones. When UI benefits are too low, workers tend to remain in unproductive firms, which drags down overall productivity. If separation rates fall sufficiently, the drop in productivity can outweigh the benefits of higher employment, resulting in reduced output and, consequently, lower consumption.

Further, Figure 1.6.1 shows that when UI benefits exceed their optimal level, the

Figure 1.6.2. Optimal STW Policy with Optimal UI in Steady State - Allocation



Notes: The figure shows the optimal response of the STW system with optimal UI in steady state (red line) to a 1% negative productivity shock and compares it to the baseline economy with corresponding UI benefits in steady state (black line) and the social planner economy (blue dashed line).

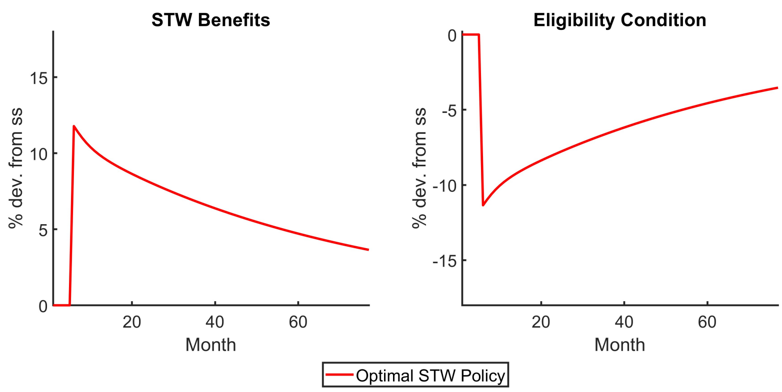
optimal STW policy reduces the number of inefficient separations from an efficiency point of view, resulting in a slower decline in consumption. This allows for a slightly higher optimal UI rate. However, this rate still remains close to the level that maximizes consumption. This has significant consequences for the business cycle.

Next to keeping working hours efficient, the planner has two primary objectives: First, insure workers against income losses by reducing employment fluctuations, and second, lower the efficiency costs of the business cycle by stabilizing consumption. Figure 1.6.3 shows that when UI benefits are set optimally in steady state, they barely contribute to inefficient consumption fluctuations over the business cycle. The relatively low UI level implies that unproductive firms can persist in steady state. Losing these firms during downturns is not particularly costly because their contribution to total production is minimal.

The costs of the UI system do not lie in higher unemployment levels that lead to a loss of consumption, but in the more risky contracts that workers agree to with firms and the subsequent loss of insurance. Remember, consumption smoothing can either be done at the intensive margin via the UI benefits or at the extensive margin by reducing unemployment (see equation 1.5.1).

Workers fail to recognize that the extra income from riskier contracts is offset by higher taxes, which are needed to finance the increased cost of the UI system. By subsidizing temporarily low productive matches, STW steps in to encourage firms and workers to negotiate less risky contracts, reducing employment fluctuations during recessions. Figure 1.6.3 shows that the optimal response of the STW system with the optimal steady state is very similar to the one with the too large benefit.

Figure 1.6.3. Optimal STW Policy with Optimal UI in Steady State - Instruments



Notes: The figure shows the optimal response of the STW system with optimal UI in steady state (red line) to a 1% negative productivity shock

In conclusion, when the UI system is set close to its optimal level, the primary role of STW is not to stabilize employment for the sake of consumption stabilization, but rather to provide better income insurance via the extensive margin to workers, helping them avoid overly risky labor contracts.

1.7 Discussion and Conclusion

In conclusion, the model presented in this paper demonstrates that STW can be a valuable complement to the UI system. While the UI system offers income insurance to workers, the STW system helps mitigate its distortionary effects. STW itself does not provide direct income insurance, as firms insure workers against idiosyncratic productivity shocks while employed. In fact, eliminating the UI system and implementing the Hosios condition removes any role for STW in the economy. To reduce the distortionary impacts of the UI system during recessions, STW is adjusted by offering more generous benefits and tightening the eligibility condition. When UI benefits are set too generously, this helps to stabilize separations, employment, total working hours, and consumption. When UI benefits are set optimally in steady state, business cycles do not show efficiency losses in production. However, they still face problems to insure workers against unemployment. While not offering insurance via the intensive margin, STW helps by stabilizing employment to offer insurance and smooth consumption between households via the extensive margin.

Despite its benefits, STW has two main shortcomings that prevent it from fully implementing the planner's solution. First, STW distorts working hours. When setting optimal STW benefits, the planner faces a trade-off between implementing the optimal separation rate and minimizing the distortion of working hours. This makes it crucial to adjust STW over the business cycle. If STW is not adjusted, the distortion of working hours can exacerbate fluctuations, undermining STW's ability to stabilize total working hours.

Second, unlike alluded to in papers like Balleer et al. (2016), Giupponi and Landais (2018), or Cahuc, Kramarz, and Nevoux (2021), STW cannot stabilize the job-finding rate. If STW benefits are financed through a tax on salaries, any increase in STW benefits will be offset by a corresponding increase in the tax rate, nullifying the impact on the joint surplus of firms and workers. Even when considering lump sum taxes, allowing STW to influence vacancy creation directly, the planner would refrain from stabilizing the job-finding rate, as the additional distortions to working hours would be too costly.

The paper suggests three avenues for further research. First, an intriguing result from the theoretical section is that STW functions similarly to a layoff tax, in the sense of Blanchard and Tirole (2008) and Cahuc and Zylberberg (2008). The key difference, however, is that STW distorts working hours, making it less effective than layoff taxes. Despite this, layoff taxes and STW are fundamentally different instruments: STW is a subsidy, while layoff taxes are a penalty. In the model, their similarity arises because firms are assumed to be never financially constrained, enabling them to offer insurance to workers and pay layoff costs regardless of circumstances. But what if firms do face financial constraints? In such a scenario, STW might provide an insurance component for workers that layoff taxes cannot. Additionally, paying penalties could become unfeasible for financially constrained firms. A comparison of STW and layoff taxes under financial constraints would be a valuable area of exploration. Chapter 3 deals with this topic.

Second, Cooper, Meyer, and Schott (2017) argue that a major drawback of STW is its potential to reduce allocative efficiency by incentivizing workers to remain in less productive occupations, thereby hindering their reallocation to more productive firms. However, I find that STW can strike a balance between reducing allocative inefficiency and minimizing the costs of reallocating workers through the labor market, effectively leaving no room for allocative inefficiencies. This outcome is based on the assumption that the shock duration is uniform for all workers. But what happens if firms and workers experience shocks of varying durations? Investigating how optimal STW policy would respond to such differences in shock duration could provide new insights into how STW should deal with allocative inefficiencies. Chapter 2 deals with this topic.

Finally, the question remains of whether the performance of STW could be enhanced by combining it with other labor market instruments. One significant limitation of STW is its inability to stabilize job-finding rates. It may be advantageous to explore the potential benefits of combining STW with a vacancy subsidy.

References

Ahr, Paul R, and Thomas B Ahr. 2000. *Overturn Turnover: Why some employees leave, why some employees stay, and ways to keep the ones who you want to stay*. Causeway Publishing Company. [44]

Baily, Martin Neil. 1978. "Some aspects of optimal unemployment insurance." *Journal of Public Economics* 10 (3): 379–402. [7]

Balleer, Almut, Britta Gehrke, Wolfgang Lechthaler, and Christian Merkl. 2016. "Does short-time work save jobs? A business cycle analysis." *European Economic Review* 84 (C): 99–122. [9, 12, 48, 51, 52, 56]

Barnichon, Regis. 2010. "Building a composite Help-Wanted Index." *Economics Letters* 109 (3): 175–78. [42]

Blanchard, Olivier J., and Jean Tirole. 2008. "The Joint Design of Unemployment Insurance and Employment Protection: A First Pass." *Journal of the European Economic Association* 6 (1): 45–77. [7, 10, 32, 36, 56]

Braun, Helge, and Björn Brügemann. 2017. "Welfare Effects of Short-Time Compensation." Tinbergen Institute Discussion Papers 17-010/VI. Tinbergen Institute. [7, 9, 10]

Burdett, Kenneth, and Randall Wright. 1989. "Unemployment Insurance and Short-Time Compensation: The Effects on Layoffs, Hours per Worker, and Wages." *Journal of Political Economy* 97 (6): 1479–96. [7, 9, 10, 33]

Cahuc, Pierre. 2024. "The Micro and Macro Economics of Short-Time Work." IZA Discussion Papers 17111. Institute of Labor Economics (IZA). [7]

Cahuc, Pierre, Francis Kramarz, and Sandra Nevoux. 2021. "The Heterogeneous Impact of Short-Time Work: From Saved Jobs to Windfall Effects." IZA Discussion Papers 14381. Institute of Labor Economics (IZA). [9, 11, 15, 16, 22, 56]

Cahuc, Pierre, and André Zylberberg. 2008. "Optimum income taxation and layoff taxes." *Journal of Public Economics* 92 (10-11): 2003–19. [7, 10, 32, 36, 56]

Chetty, Raj. 2006. "A general formula for the optimal level of social insurance." *Journal of Public Economics* 90 (10-11): 1879–901. [7]

Christoffel, Kai, and Tobias Linzert. 2010. "The Role of Real Wage Rigidity and Labor Market Frictions for Inflation Persistence." *Journal of Money, Credit and Banking* 42 (7): 1435–46. [43]

Cooper, Russell, Moritz Meyer, and Immo Schott. 2017. "The Employment and Output Effects of Short-Time Work in Germany." NBER Working Papers 23688. National Bureau of Economic Research, Inc. [9, 12, 13, 31, 56]

Costain, James S., and Michael Reiter. 2008. "Business cycles, unemployment insurance, and the calibration of matching models." *Journal of Economic Dynamics and Control* 32 (4): 1120–55. [14, 19]

Dengler, Thomas, and Britta Gehrke. 2021. "Short-Time Work and Precautionary Savings." IZA Discussion Papers 14329. Institute of Labor Economics (IZA). [9]

Domeij, David, and Martin Floden. 2006. "The Labor-Supply Elasticity and Borrowing Constraints: Why Estimates are Biased." *Review of Economic Dynamics* 9 (2): 242–62. [44]

Ellguth, Peter, Hans-Dieter Gerner, and Ines Zapf. 2018. "Flexible working times become more and more important." *IAB Kurzbericht*, (15): 8. [22]

Engen, Eric M., and Jonathan Gruber. 2001. "Unemployment insurance and precautionary saving." *Journal of Monetary Economics* 47 (3): 545–79. [43]

Fahr, Stephan, and Mirko Abbritti. 2011. "Macroeconomic implications of downward wage rigidities." Working Paper Series 1321. European Central Bank. [21]

Gehrke, Britta, Wolfgang Lechthaler, and Christian Merkl. 2017. "The German labor market in the Great Recession: Shocks and institutions." IAB Discussion Paper 201714. Institut für Arbeitsmarkt- und Berufsforschung (IAB), Nürnberg [Institute for Employment Research, Nuremberg, Germany]. [9]

Giupponi, Giulia, and Camille Landais. 2018. "Subsidizing labor hoarding in recessions: the employment and welfare effects of short time work." CEP Discussion Papers dp1585. Centre for Economic Performance, LSE. [9, 56]

Haan, Wouter J. den, Garey Ramey, and Joel Watson. 2000. "Job Destruction and Propagation of Shocks." *American Economic Review* 90 (3): 482–98. [43]

Hall, Robert E. 2005. "Employment Fluctuations with Equilibrium Wage Stickiness." *American Economic Review* 95 (1): 50–65. [21]

Hijzen, Alexander, and Sébastien Martin. 2013. "The Role of Short-Time Work Schemes during the Global Financial Crisis and Early Recovery: A Cross-Country Analysis." IZA Discussion Papers 7291. Institute of Labor Economics (IZA). [9, 15]

Hopenhayn, Hugo, and Juan Pablo Nicolini. 1997. "Optimal Unemployment Insurance." *Journal of Political Economy* 105 (2): 412–38. [7]

Hosios, Arthur J. 1990. "On the Efficiency of Matching and Related Models of Search and Unemployment." *Review of Economic Studies* 57 (2): 279–98. [8, 30, 43]

Jung, Philip, and Keith Kuester. 2015. "Optimal Labor-Market Policy in Recessions." *American Economic Journal: Macroeconomics* 7 (2): 124–56. [7, 19, 21, 42, 44]

Krause, Michael U., and Thomas A. Lubik. 2007. "The (ir)relevance of real wage rigidity in the New Keynesian model with search frictions." *Journal of Monetary Economics* 54 (3): 706–27. [14, 44]

Kuhn, Moritz, Iourii Manovskii, Lutz Bellmann, Patrick Gleiser, Sophie Hensgen, Christian Kagerl, Eva Kleifgen, Ute Leber, Michael Moritz, Laura Pohlan, Duncan Roth, Malte Schierholz, Jens Stegmaier, and Matthias Umkehrer. 2021. "Warum nutzen Arbeitgeber Kurzarbeit?" [17]

Landais, Camille, Pascal Michaillat, and Emmanuel Saez. 2018. "A Macroeconomic Approach to Optimal Unemployment Insurance: Theory." *American Economic Journal: Economic Policy* 10 (2): 152–81. [7]

Michau, Jean-Baptiste. 2015. "Optimal labor market policy with search frictions and risk-averse workers." *Labour Economics* 35 (C): 93–107. [7]

Mortensen, Dale T., and Christopher A. Pissarides. 1994. "Job Creation and Job Destruction in the Theory of Unemployment." *Review of Economic Studies* 61 (3): 397–415. [8]

Münchner, Thomas. 2020. "Kurzarbeitergeld aufstocken: 7 Wege, wie Sie wieder auf 100 Prozent Nettolohn kommen." [Online; accessed 21-August-2024]. [10]

Petrongolo, Barbara, and Christopher A. Pissarides. 2001. "Looking into the Black Box: A Survey of the Matching Function." *Journal of Economic Literature* 39 (2): 390–431. [43]

Pissarides, Christopher A. 2000. *Equilibrium Unemployment Theory, 2nd Edition*. Vol. 1, MIT Press Books 0262161877. The MIT Press. [9]

Rosen, Sherwin. 1985. "Implicit Contracts: A Survey." NBER Working Papers 1635. National Bureau of Economic Research, Inc. [10]

Schmitt-Grohe, Stephanie, and Martin Uribe. 2004. "Solving dynamic general equilibrium models using a second-order approximation to the policy function." *Journal of Economic Dynamics and Control* 28 (4): 755–75. [45]

Shavell, Steven, and Laurence Weiss. 1979. “The Optimal Payment of Unemployment Insurance Benefits over Time.” *Journal of Political Economy* 87 (6): 1347–62. [7]

Shimer, Robert. 2005. “The Cyclical Behavior of Equilibrium Unemployment and Vacancies.” *American Economic Review* 95 (1): 25–49. [11, 41, 42, 44]

Silva, José Ignacio, and Manuel Toledo. 2009. “Labor Turnover Costs And The Cyclical Behavior Of Vacancies And Unemployment.” *Macroeconomic Dynamics* 13 (S1): 76–96. [44]

Teichgräber, Julian, Simon Žužek, and Jannik Hensel. 2022. “Optimal short-time work: screening for jobs at risk.” ECON - Working Papers 402. Department of Economics - University of Zurich. [11, 15, 36]

Trigari, Antonella. 2006. “The Role of Search Frictions and Bargaining for Inflation Dynamics.” Working Papers 304. IGIER (Innocenzo Gasparini Institute for Economic Research), Bocconi University. [21]

Van Audenrode, Marc A. 1994. “Short-Time Compensation: Job Security, and Employment Contracts: Evidence from Selected OECD Countries.” *Journal of Political Economy* 102 (1): 76–102. [9]

Weber, Enzo, and Christof Röttger. 2022. “Es gab keinen Big Quit in Deutschland.” [49]

Appendix 1.A List of important variables and functions

Symbol	Description
n_t	Mass of employed workers at beginning of period
$u_t = 1 - n_t$	Mass of unemployed workers at beginning of period
v_t^f	Mass of firm owners
v_t	Mass of posted vacancies
$\theta_t = \frac{v_t}{1-n_t}$	Labor market tightness
f_t	Job-finding rate
q_t	Vacancy-filling rate
ρ_t	Separation rate
s_t	mass of separations
$\varepsilon_{s,t}$	Separation threshold (on STW)
$\xi_{s,t}$	Separation threshold (no STW)
$G(\varepsilon), g(\varepsilon)$	CDF and PDF of idiosyncratic productivity shocks
a_t	Aggregate productivity
α	Output elasticity (hours)
ψ	Inverse Frisch elasticity
D_t	Eligibility threshold for STW (on working hours)
$\varepsilon_{\text{stw},t}$	Eligibility threshold for STW (on productivity)
$\tau_{\text{stw},t}$	STW benefits per hour gap
b_t	Unemployment benefits
T	Expected time a worker spends on the UI system
$\tau_{J,t}$	Lump-sum production tax

Table 1.A.1. Important Variables and Functions

Symbol	Description
m_t	mass of matched workers
$v(h)$	Disutility of labor: $\frac{h^{1+\psi}}{1+\psi}$
$u(c)$	Worker utility from consumption net of disutility
$y_t(\varepsilon, h)$	Output
$y_t(\varepsilon)$	Shorthand for $y(\varepsilon, h(\varepsilon))$
$y_{\text{stw}.t}(\varepsilon)$	Output under STW hours
$z_t(\varepsilon, h)$	Output net of disutility (cons.-eq. units)
$z_t(\varepsilon)$	Shorthand for $z(\varepsilon, h(\varepsilon))$
$z_{\text{stw},t}(\varepsilon)$	Output net of disutility under STW hours
Ω_t	Welfare Costs of STW ε
$h_t(\varepsilon)$	Hours worked (non-STW)
$h_{\text{stw},t}(\varepsilon)$	Hours worked under STW
$\tilde{c}_t^w$	Consumption equivalent worker
$w_t(h)$	Total Wage functions
$w_{eu,t}$	Severance payments
c_t^f	Consumption firm owners
c_t^u	Consumption unemployed workers
Π_t	Profits
Q_t	Stochastic discount factor
$V_t^W(\varepsilon)$	Worker value (no STW), after idiosyncratic productivity shock
$V_{\text{stw},t}^W(\varepsilon)$	Worker value on STW, after idiosyncratic productivity shock
$\mathcal{V}_t^W$	Expected worker value, before idiosyncratic productivity shock
U_t	Value of unemployment
$J_t(\varepsilon)$	Firm value (no STW), after idiosyncratic productivity shock
$J_{\text{stw},t}(\varepsilon)$	Firm value on STW, after idiosyncratic productivity shock
$\mathcal{J}_t$	Expected firm value, before idiosyncratic productivity shock
FE	Fiscal Externality
BE	Bargaining Effect
L_v	Social value of a new hire
L_s	Social value of the marginal worker
$L_s^\star$	Social Value of marginal worker with optimal STW
$L_{stw}^\star$	Social Value of marginal worker that enters STW with optimal STW
$\text{MMS}_{\tau_{\text{stw}}}$	Marginal number of matches preserved with STW
$\text{MWL}_{\tau_{\text{stw}}}$	Marginal welfare loss caused by STW benefits
$\text{MMS}_{\varepsilon_{\text{stw}}}$	Marginal number of matches preserved with looser eligibility
$\text{MWL}_{\varepsilon_{\text{stw}}}$	Marginal welfare loss caused by looser eligibility

Table 1.A.2. Important Variables and Functions

Appendix 1.B The bargaining effect

As described in the main text, the bargaining effect illustrates how STW influences vacancy posting and separation behavior of firms via the wage channel. Diminishing returns to consumption make it harder for firms to reduce wages. However, the introduction of STW necessitates wage reductions to offset firms' losses during downturns. These adjustments are imperfect, resulting in relatively high wages, which, in turn, leads to fewer vacancies being posted and an increase in separations. Mathematically, this effect can be expressed as:

$$\tilde{BE} = \frac{\frac{BE \cdot \lambda_\theta}{n \cdot u'(\tilde{c}^w)}}{1 + \frac{\beta \cdot (1+\eta \cdot BE) \cdot \lambda_\theta}{n \cdot u'(\tilde{c}^w)}} \cdot \frac{1 - f(\theta)}{1 - \eta} \cdot \frac{k_v}{q(\theta)}$$

with

$$BE = \frac{\left(-\frac{u''(\tilde{c}^w)}{u'(\tilde{c}^w)}\right) \cdot \frac{u(\tilde{c}^w) - u(b)}{u'(\tilde{c}^w)}}{1 + (1-\eta) \cdot \left(-\frac{u''(\tilde{c}^w)}{u'(\tilde{c}^w)}\right) \cdot \frac{u(\tilde{c}^w) - u(b)}{u'(\tilde{c}^w)}} < \frac{1}{1-\eta}$$

Note that λ_θ is the Lagrange multiplier of the value of the firm, respectively, the job-creation condition. It captures welfare losses by posting insufficient vacancies and experiencing excessive separations. The imperfect reduction in wages leads to more separations, not only in firms utilizing STW but also in those without access to STW. To mitigate the additional loss of these matches, a looser eligibility condition must be implemented, increasing the hours distortion effects of STW.

BE captures the effect of risk aversion. Under risk-aversion, *BE* must be positive as:

$$\text{risk-aversion} \quad \Rightarrow \quad -u''(\tilde{c}_t^w) > 0$$

Note that the bargaining effect is zero under risk-neutrality:

$$\text{risk-neutrality} \quad \Rightarrow \quad u''(\tilde{c}_t^w) = 0 \quad \Rightarrow \quad BE = 0 \quad \Rightarrow \quad \tilde{BE} = 0$$

Appendix 1.C What if STW can directly influence Vacancy Posting?

The paper identifies the inability to stabilize the job-finding rate as one of the core problems of STW. In contrast, other authors such as Balleer, Gehrke, Lechthaler, and Merkl (2016), Giupponi and Landais (2018), or Cahuc, Kramarz, and Nevoux (2021) highlight the ability of STW to increase vacancy postings as a strength of STW. The difference lies in the financing of the STW system. While my model finances the STW system by income taxes, their models rely on lump sum taxes or do not consider financing of STW at all. This makes a big difference: any increase in the joint surplus of firms and workers from higher expected STW benefits is in my model offset by a corresponding rise in the production tax, making it impossible for the government to steer job-finding rates with STW.

Nonetheless, while the balanced budget assumption certainly is a good assumption for the long run, it might not need to hold in the short run. In fact, governments may choose to borrow in the short run to avoid raising taxes during recessions. Therefore, an expansion of the STW system financed by a deficit might still help stabilize the job-finding rates in recessions.

To explore this conjecture, I replace the production tax with a lump sum tax on all households. For simplicity, I assume risk neutrality. Lump-sum taxes in a model with risk-aversion suffer from distributional consequences that I do not want to discuss here. Under risk neutrality, lump sum taxes do not influence vacancy posting or separation decisions, allowing STW to directly influence vacancy posting. By increasing the generosity of the STW system, the government can now stimulate the expected joint surplus of firms and workers, encouraging vacancy creation.

Lemma 2
Under lump-sum taxation, the social value of hiring an additional worker changes to L_V'

$$L_V' = \underbrace{\frac{\eta-\gamma}{(1-\eta)\cdot(1-\gamma)}}_{\text{Congestion Externality}} + \underbrace{\frac{\beta}{1-\beta\cdot(1-G(\epsilon_s))\cdot(1-f)}\cdot b}_{\text{Fiscal Externality UI on Hiring (FE')}}$$
$$\underbrace{-\frac{\beta}{1-\beta\cdot(1-G(\epsilon_s))\cdot(1-f)}\cdot\int_{\epsilon_s}^{\epsilon_{stw}}(\bar{h}-h_{stw}(\epsilon))\cdot\tau_{stw}\cdot dG(\epsilon)}_{\text{STW increases vacancy posting}}$$

while $L_S^\star$ and $L_{STW}^\star$ stay the same.

PROOF: Appendix 1.L

This becomes clear analytically in Lemma 2, which derives the social value of hiring an additional worker. A more generous STW system, whether through a looser eligibility threshold or higher benefits, encourages vacancy posting, thereby lowering the social value of hiring an additional worker. Under an optimal STW policy, the social value of the marginal match, $L_S^\star$, remains unchanged because the social planner hires until the cost of an additional worker with STW equals its welfare gain.

Proposition 5 outlines the optimal eligibility condition in an economy with a lump sum tax. In determining this condition, the government faces a new trade-off, as described in equation 1.C.2. On one side, a looser eligibility condition raises the likelihood that a firm can use the STW program. This increases the expected benefits from the system, enhancing the joint surplus of firms and workers, which in turn raises job-finding rates and lowers separation rates. On the other side, relaxing the eligibility condition also spreads the distortionary effects on working hours across more firms, heightening the welfare costs associated with the STW system.

Proposition 5, Optimal Eligibility Condition in Steady-State - Lump Sum Tax *Consider the economy described in section 2.1. and replace the production tax by a lump sum tax. Further, assume that the economy has converged to its non-stochastic steady state. Then, the optimal eligibility condition $D = h_{stw}(\epsilon_{stw})$ is implicitly defined by the separation threshold of a firm without STW*

$$\underbrace{S(\epsilon_{stw}) \;=\; y(\epsilon_{stw}) - v\big(h(\epsilon_{stw})\big) + F + \frac{1-\eta\cdot f}{1-\eta}\cdot\frac{k_v}{q} \;=\; 0}_{\text{Joint Surplus without STW is zero}} \tag{1.C.1}$$

as long as the welfare costs of a looser eligibility outweigh its welfare gains:

$$\underbrace{n\cdot\frac{\partial\Omega}{\partial\epsilon_{stw}}}_{\text{Welfare Costs}} \;\geq\; \left[L_V'\cdot\underbrace{\frac{\partial f^{ge}}{\partial\epsilon_{stw}}}_{\text{More jobs}}\cdot u + L_S^\star\cdot\underbrace{\left(-n\cdot g(\epsilon_s)\cdot\frac{\partial\epsilon_s^{ge}}{\partial\epsilon_{stw}}\right)}_{\text{Less separations}}\right] \tag{1.C.2}$$

Otherwise it is determined by setting the welfare gains of a looser eligibility condition equal to its welfare costs.

PROOF: Appendix 1.L

There are two possible outcomes. In the first, the welfare gains from posting more vacancies by loosening the eligibility criteria outweigh the distortionary effects of the STW system. Here, the optimal eligibility condition is achieved by balancing the welfare gains of a looser eligibility condition against its welfare losses. In the second outcome, the welfare costs from increased working-hour distortions outweigh the benefits of additional vacancy posting. In this case, the eligibility condition should be set as strictly as possible. The optimal eligibility threshold in this case would match the separation threshold of a firm without access to STW in this period ($\epsilon_{stw,t} = \xi_{s,t}$). Determining which effect dominates is a quantitative question.

Proposition 6 establishes the optimal STW benefits under lump sum taxes. In contrast to the previous section, the Ramsey planner's objective is to implement the optimal expected net-transfer of the STW system, as indicated by equation 1.C.3. Potential future STW benefits are priced in the expected benefits. In the absence of income taxes, the STW system not only operates by increasing the period surplus, but also by raising the expected surplus of firms and workers.

Proposition 6, Optimal STW Subsidy in Steady State - Lump Sum Tax
Consider the economy described in section 2.1 and replace the production tax by a lump sum tax. Further, assume that the economy has converged to its non-stochastic steady state and the eligibility condition is set according to Proposition 5. Then, the optimal STW subsidy τ_{stw} is implicitly determined by the optimal expected net-transfer τ_{stw}^{net}:

$$\begin{aligned} \tau_{stw}^{net} &= \tau_{stw} \cdot (\bar{h} - h_{stw}(\epsilon_s)) \\ &+ \frac{\beta \cdot \big(1 - f\big)}{1 - \beta \cdot \big(1 - G(\epsilon_s)\big) \cdot (1 - f)} \cdot \int_{\epsilon_s}^{\epsilon_{stw}} \tau_{stw} \cdot (\bar{h} - h_{stw}(\epsilon)) \cdot dG(\epsilon) \end{aligned} \tag{1.C.3}$$

The optimal expected net-transfer τ_{stw}^{net} is determined by:

$$\tau^{net}_{stw} = \underbrace{(1-f)\cdot FE'}_{\substack{\text{A: Influence Distortionary Effect UI on Separations} \\ >0}} - \underbrace{L^{\star}_{S}}_{\substack{\text{B: Welfare Costs rescuing Worker with STW} \\ >0}} + \underbrace{G^{\star}_{S}}_{\text{C: Welfare Gain increasing joint Surplus with STW}} \tag{1.C.4}$$

The welfare Gain from increasing joint Surplus is defined as:

$$G^{\star}_{S} = \Bigg[\underbrace{L'_{V}\cdot\frac{\partial f^{ge}}{\partial \tau_{stw}}}_{\text{More jobs}}\cdot u + L^{\star}_{S}\cdot\underbrace{\left(-n\cdot g(\epsilon_s)\cdot\frac{\partial \epsilon^{ge}_{s}}{\partial \tau_{stw}}\right)}_{\text{Fewer separations}} + L^{\star}_{STW}\cdot\underbrace{\left(-n\cdot g(\epsilon_{stw})\cdot\frac{\partial \epsilon^{ge}_{stw}}{\partial \tau_{stw}}\right)}_{\text{Fewer workers on STW}}\Bigg]\Bigg/\left[g(\epsilon_s)\cdot\frac{\partial \epsilon_s}{\partial \tau_{stw}}\right]$$

PROOF: Appendix 1.L

The optimal expected net-transfer depends on three factors (see equation 1.C.4). Similar to Proposition 2, the expected net-tansfer should reduce the pay of the firm by the financial burden the worker would impose on the UI system (A) minus a penalty for the welfare costs of using the STW system (B). Different to Proposition 2, larger STW benefits can now increase the expected value of firms, stimulating vacancy posting and lessening separations even more. The incentive of smaller separations also reduces the number of workers that need to be rescued with STW. Therefore, the planner adjusts the STW benefits upwards (C). The Ramsey planner weighs the benefits of reducing inefficient separations and increasing suboptimal low vacancy posting efforts against distorting working hours in the economy. Whether a larger transfer of STW benefits indeed helps to stabilize the job-finding rate remains a quantitative question.

Propositions 5 and 6 highlight that STW indeed has the potential to stabilize job-finding rates as suggested by Balleer et al. (2016), Giupponi and Landais (2018), or Cahuc, Kramarz, and Nevoux (2021). Loosening the eligibility condition and increasing STW benefits both increase the joint surplus of firms and workers and thus incentivize vacancy posting. The prerequisite is that we do not have to worry about financing the system, respectively, the budget balance assumption. However, both propositions also stress that increasing vacancy posting incentives have always been traded off against the additional distortions that STW introduces into the system.

Table 1.C.1. Parameters for Model with Lump-Sum Tax

Parameter	Description	Value	Reason
ρ	Target ss separation rate	0.03	Data
f	Target ss job-finding rate	0.41	Data
q	Target ss vacancy filling rate	0.338	Haan, Ramey, and Watson (2000)
β	Discount rate	0.996	Jung and Kuester (2015)
ψ	Inverse Frisch-elasticity	1.5	Domeij and Floden (2006)
γ	Elasticity matching function w.r.t. unemployment	0.65	Shimer (2005)
η	Bargaining power worker	0.65	Implements Hosios condition
γ_w	Coeff. reaction bargaining power to productivity shock	15.5	s.d. job-finding rate 14.31 in data
F	Separation costs	1.01	s.d. separation rate of 8.2 in data
b	UI benefits	0.4	40% replacement rate of wage
α	Labor elasticity in production	0.65	Christoffel and Linzert (2010)
$\bar{h}$	"Normal" hours worked	0.834	Mean hours in baseline
ρ_a	Autocorr. productivity shock	0.985	Jung and Kuester (2015)
μ_a	Mean aggregate productivity	1.0	Normalization
$\sigma_a \cdot 100$	s.d. aggregate productivity	0.259	s.d. labor prod. of 1.91 in data
μ	Mean of lognormal productivity distribution	0.082	Normalizes wage to 1
σ	Variance of lognormal distribution	0.12	Krause and Lubik (2007)
$\bar{m}$	Matching parameter	0.383	Calculated by target ss
k_v	Vacancy posting costs	0.139	Calculated by target ss
c_f	Strength of resource cost shock	10.441	Calculated by target ss

Table 1.C.2. Business Cycle Properties Baseline Model with Lump-Sum Tax

		v	f	ρ	u	θ	$\bar{h}$	p
Standard Deviation		19.8	14.31	8.2	21.26	40.88	0.76	1.91
Autocorrelation		0.95	0.97	0.97	0.98	0.97	0.97	0.97
	v	1	1	-0.99	-0.98	1	1	1
	f	-	1	-1	-1	1	1	1
	ρ	-	-	1	1	-1	-1	-1
Correlation	u	-	-	-	1	-1	-1	-1
	θ	-	-	-	-	1	1	1
	$\bar{h}$	-	-	-	-	-	1	1
	p	-	-	-	-	-	-	1

Notes: The table reports the second moments of the model. As in the data of Shimer (2005), all variables are quarterly averages of monthly series and reported as log-deviations. p denotes the average output per person, that is $p = E[y_t(\varepsilon)|\varepsilon \geq \varepsilon_{s,t}]$.

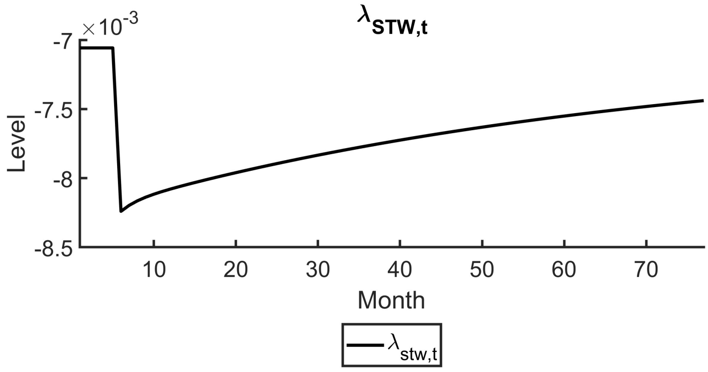
To evaluate which effect dominates, I use the same calibration strategy as described in section 4. However, due to the risk-neutrality assumption, I need to re-calibrate the model. For brevity, Table 1.C.1 lists all calibrated parameters with calibration strategy. Table 1.C.2 shows that the model closely replicates US business cycle facts.

The calibration reveals that the distortionary effects of the STW system clearly outweigh the benefits of additional vacancy posting. First of all, to evaluate how to set the eligibility condition, we can look at its Lagrange multiplier. Is the Lagrange multiplier negative, then the distortionary effects of STW dominate, and we want to set the eligibility condition as strict as possible:

$$\begin{aligned}\lambda_{stw,t} = & \left[L'_V \cdot \underbrace{\frac{\partial f^{ge}}{\partial \epsilon_{stw}}}_{\text{More jobs}} \cdot u + L^\star_S \cdot \underbrace{\left(-n \cdot g(\epsilon_s) \cdot \frac{\partial \epsilon_s^{ge}}{\partial \epsilon_{stw}} \right)}_{\text{Less separations}} + L^\star_S \cdot \underbrace{\left(-n \cdot g(\epsilon_{stw}) \cdot \frac{\partial \epsilon_{stw}^{ge}}{\partial \epsilon_{stw}} \right)}_{\text{Less separations}} \right] \\ & - \underbrace{n \cdot \frac{\partial \Omega}{\partial \epsilon_{stw}}}_{\text{Welfare Costs}} \\ \geq & \left[L'_V \cdot \underbrace{\frac{\partial f^{ge}}{\partial \epsilon_{stw}}}_{\text{More jobs}} \cdot u + L^\star_S \cdot \underbrace{\left(-n \cdot g(\epsilon_s) \cdot \frac{\partial \epsilon_s^{ge}}{\partial \epsilon_{stw}} \right)}_{\text{Less separations}} \right] - \underbrace{n \cdot \frac{\partial \Omega}{\partial \epsilon_{stw}}}_{\text{Welfare Costs}}\end{aligned}$$

Figure 1.C.1 shows that the Lagrange multiplier is indeed negative. In fact, it becomes even more negative in recessions. As benefits increase in recessions, it makes adding workers to the system increasingly costly.

Figure 1.C.1. STW must be set as strict as possible



Notes: The Figure shows the impulse response functions of a 1% negative productivity shock. It shows the reaction of the Lagrange multiplier for the eligibility condition.

Further, Figure 1.C.2 shows that it is too costly to increase STW benefits sufficiently to stabilize the job-finding rate. In fact, the impulse response function of the job-finding rate under a STW system with lump sum tax or distortionary taxation looks almost the same, despite the ability of the Ramsey Planner to stimulate vacancy postings with STW.

We can conclude that even when the Ramsey planner has the ability to directly influence vacancy postings with STW, he refrains from stabilizing the job-finding rate as the distortionary effects of the STW system are too costly.

Figure 1.C.2. Optimal STW Policy, Lump Sum vs Distortionary Taxation

Notes: The Figure shows the impulse response functions of a 1% negative productivity shock. It compares the response of the economy with optimal STW policy, financed by a lump sum income tax (red line) against a lump sum tax on households (black line).

Appendix 1.D Optimal Layoff Tax vs Optimal STW

The paper explores the optimal design of short-time work (STW) policy and concludes that STW addresses the fiscal externalities of the UI system. In the U.S., a system with similar purpose already exists. The UI system is experience-rated. This means that firms' contributions to the UI system increase when workers are laid off, effectively functioning as a layoff tax that firms must pay when they separate from employees.

In this policy experiment, I exchange the STW system for a layoff tax. Firms are required to pay a tax to the government when they lay off a worker. The revenue from the tax can be used to finance the UI system. Proposition 7 derives the optimal layoff tax, showing that it serves the same purpose as the STW system within the model.

Proposition 7, Optimal STW benefits in steady state
Consider the economy as previously described. Assume that it has converged to its non-stochastic steady state. Then, the optimal layoff tax τ_S is determined by:

$$\tau_S = \underbrace{(1-f)\cdot FE}_{\text{A: Fiscal Externality UI} > 0} - \underbrace{\tilde{BE}}_{C\text{: Bargaining Effect}}$$

PROOF: Appendix 1.M

The key distinction between STW and layoff taxes in the model is that layoff taxes do not distort working hours. This difference is reflected in the comparison between the optimal layoff tax and the optimal net transfer of STW benefits (as discussed in Proposition 2 versus Proposition 7). The planner reduces the optimal STW benefits to minimize the distortionary effects associated with the STW system. As a result, we can conclude that, within the model, layoff taxes are superior to STW benefits due to their ability to avoid these distortions.

This result might be surprising, as STW and layoff taxes should be two fundamentally different instruments. STW functions as a subsidy, while layoff taxes operate as a penalty. Given this, one might wonder how they can lead to similar outcomes within the model. There are two main reasons for this.

First, neither STW nor layoff taxes directly influences the job-finding rate. The budget constraint for the economy with STW system is:

$$n_t \cdot \tau_{J,t} = \left(1-n_t\right)\cdot b_t + n_t \cdot \int_{\epsilon_{s,t}}^{\epsilon_{stw,t}} \tau_{stw,t}\cdot\left(\bar{h} - h_{stw,t}(\epsilon)\right)\cdot dG(\epsilon)$$

The budget constraint for the economy with layoff tax is:

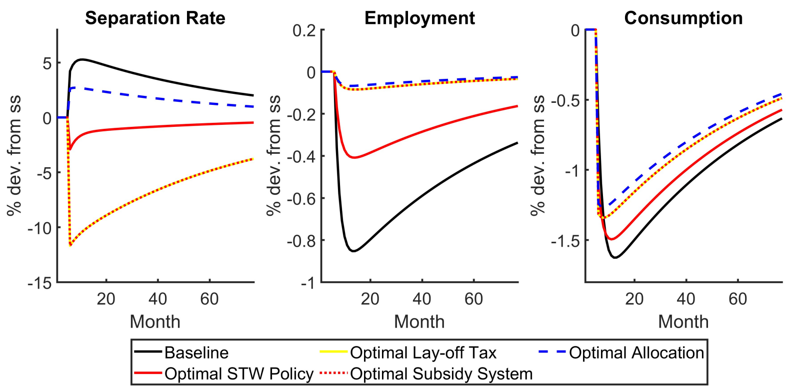
$$n_t \cdot \tau_{J,t} = \left(1-n_t\right)\cdot b_t - n_t \cdot \rho_t \cdot \tau_{S,t}$$

Increasing the generosity of STW can enhance the joint surplus of firms and workers; however, the subsequent rise in income taxes required to finance the STW system counteracts this benefit. Conversely, raising the layoff tax reduces the joint surplus of firms and workers. Nevertheless, an increase in the layoff tax also lowers the amount of income taxes needed to fund the UI system, which offsets the negative impact of the layoff tax on the joint surplus. If the surplus is not altered, job-finding rates do not alter.

Second, firms in the model are never financially constrained, which allows them to offer insurance to workers and pay layoff costs regardless of their financial situation. Under these circumstances, layoff taxes are clearly superior to STW. With financial constraints, things might be different. STW could be a tool to provide imperfect income insurance for the firm. Further, layoff taxes might become ineffective under financial constraints. A comparison of STW and layoff taxes in such circumstances would be interesting.

To see the quantitative difference between the optimal STW policy and optimal layoff taxes in the model, I calculated the impulse response functions to a 1% negative productivity shock. Figure 1.D.1 shows that, due to the absence of working hours distortions, layoff taxes are much better suited to stabilize the business cycle. In fact, it replicates the results of the hypothetical subsidy system from Section 5.4. Almost all inefficient fluctuations in employment and consumption can be solved.

Figure 1.D.1. Optimal Lay-off Taxes - Allocation



Notes: The figure shows the impulse response function of an economy with optimal STW system (red solid line) to a hypothetical subsidy system without hours distortion (red dashed line) and optimal layoff taxes (yellow line) for a 1% negative productivity shock and compares it to the response of the social planner economy (blue dashed line) and the baseline economy (black line).

Appendix 1.E Steady State

Table 1.E.1. Steady State Results

	baseline	optimal STW	optimal UI	optimal UI, STW
Job-finding rate	0.41	0.41	0.507	0.479
Separation rate	0.03	0.021	0.015	0.014
Fraction of workers on STW	-	0.009	-	0.004
Unemployment	0.07	0.049	0.027	0.03
Average hours worked	0.839	0.834	0.837	0.836
Consumption	0.979	0.984	0.992	0.991
UI benefits	0.45	0.45	0.34	0.37
STW benefits	-	0.735	-	0.543
Minimum hours reduction	-	-66%	-	-55%
Lump-Sum income tax	0.019	0.017	0.005	0.007

Notes: The table compares the steady state results between the baseline model, the model of optimal STW benefits with given UI benefits, the model with optimal UI benefits, and the model where STW benefits and UI benefits are chosen optimally. Further, note that minimum hours' reduction denotes how many working hours have to fall below their normal level to enter the STW system: $\frac{D-\bar{h}}{\bar{h}}$.

Table 1.E.1 presents the steady-state results of the model. Columns 1 and 2 compare the baseline economy to an economy that includes an optimal STW system. As predicted by the theory, STW has virtually no effect on the job-finding rate and operates almost entirely by reducing separation rates. The separation rate falls from 3% to 2.1%, which leads to a substantial decline in the unemployment rate. Consistent with the theoretical predictions, the introduction of STW also reduces government fiscal expenditures, which decline by approximately 10%.

Quantitatively, the model suggests that the optimal unemployment rate is lower than the current U.S. unemployment rate. Reducing UI benefits would significantly raise employment and expand consumption possibilities. When STW is introduced as part of the optimal policy mix, it directly addresses the fiscal externality that UI creates through its effect on separations. Proposition 4 shows that this enables the government to offer more generous UI benefits — a result that is also confirmed quantitatively. Interestingly, the reduced fiscal externality of the UI system is used entirely to increase the generosity of the UI system, rather than to further reduce unemployment.

Finally, note the difference in the eligibility condition between the economy with only optimized STW and the economy with both optimal UI and STW benefits. In the latter, a smaller reduction in working hours is required for a match to become eligible for STW. The reason is straightforward: with lower UI benefits, smaller STW benefits are needed to offset the fiscal externality of the UI benefits. This weakens the incentive for matches to reduce hours, making fewer matches eligible for STW.

Appendix 1.F Nash-Bargaining

Wage and Severance Payment.
The FOC for salary outside STW:

$$\eta_{t-1} \cdot u'(\tilde{c}_t(\epsilon)) \cdot g(\epsilon) \cdot \mathscr{J}_t \ = \ (1-\eta_{t-1}) \cdot g(\epsilon) \cdot \left(\mathscr{V}_t^W - U_t\right)$$

The FOC for salary on STW:

$$\eta_{t-1} \cdot u'(\tilde{c}_{stw,t}(\epsilon)) \cdot g(\epsilon) \cdot \mathscr{J}_t \ = \ (1-\eta_{t-1}) \cdot g(\epsilon) \cdot \left(\mathscr{V}_t^W - U_t\right)$$

The FOC for severance pay:

$$\eta_{t-1} \cdot u'(\tilde{c}_{eu,t}) \cdot g(\epsilon) \cdot \mathscr{J}_t \ = \ (1-\eta_{t-1}) \cdot g(\epsilon) \cdot \left(\mathscr{V}_t^W - U_t\right)$$

Rearranging gives:

$$u'(\tilde{c}_t(\epsilon)) \ = \ u'(\tilde{c}_{stw,t}(\epsilon)) \ = \ u'(\tilde{c}_{eu,t})$$

Workers are perfectly insured against idiosyncratic shocks in the period:

$$\tilde{c}_t^w \ = \ \tilde{c}_t(\epsilon) \ = \ \tilde{c}_{stw,t}(\epsilon) \ = \ \tilde{c}_{eu,t}$$

Using the full insurance result, we get the optimal wage equation:

$$\eta_{t-1} \cdot \mathscr{J}_t \ = \ (1-\eta_{t-1}) \cdot \frac{\mathscr{V}_t^W - U_t}{u'(\tilde{c}_t^w)}$$

Using the optimal wage equation, we gain expressions for the value of the firm and the surplus of employed workers, dependent on the joint surplus weighted by the marginal utility for firms and workers:

$$\begin{aligned}
\mathscr{J}_t \ &= \ (1-\eta_{t-1}) \cdot \left(\mathscr{J}_t + \frac{\mathscr{V}_t^W - U_t}{u'(\tilde{c}_t^w)}\right) \\
\mathscr{V}_t^W - U_t \ &= \ \eta_{t-1} \cdot \left(\mathscr{J}_t + \frac{\mathscr{V}_t^W - U_t}{u'(\tilde{c}_t^w)}\right)
\end{aligned}$$

The surplus of workers is denoted as:

$$
\begin{aligned}
\mathcal{V}_t^W &= \int_{\max\{\epsilon_{stw,t},\xi_{s,t}\}}^{\infty} V_t^W(\epsilon)dG(\epsilon) + \int_{\epsilon_{s,t}}^{\epsilon_{stw,t}} V_{stw,t}^W(\epsilon)dG(\epsilon) \\
&\quad + \rho_t \cdot (u(c_{eu,t}) - u(b_t) + U_t) \\
&= \int_{\max\{\epsilon_{stw,t},\xi_{s,t}\}}^{\infty} u(\tilde{c}_t(\epsilon))dG(\epsilon) + \int_{\epsilon_{s,t}}^{\epsilon_{stw,t}} u(\tilde{c}_{stw,t}(\epsilon))dG(\epsilon) + \rho_t \cdot u(c_{eu,t}) \\
&\quad + \rho_t \cdot \big(U_t - u(b_t)\big) + (1-\rho_t) \cdot E_t\left[\beta \cdot \frac{\mathcal{V}_{t+1}^W - U_{t+1}}{u'(\tilde{c}_t^w)}\right]
\end{aligned}
$$

Insert value of unemployed worker:

$$
\begin{aligned}
\mathcal{V}_t^W &= \int_{\max\{\epsilon_{stw,t},\xi_{s,t}\}}^{\infty} [u(\tilde{c}_t(\epsilon)) - u(b_t)]\,dG(\epsilon) + \int_{\epsilon_{s,t}}^{\epsilon_{stw,t}} [u(\tilde{c}_t(\epsilon)) - u(b_t)]\,dG(\epsilon) \\
&\quad + \rho_t \cdot \big[u(c_{eu,t}) - u(b_t)\big] \\
&\quad + (1-\rho_t) \cdot (1-f_t) \cdot \beta \cdot E_t\big[\mathcal{V}_{t+1}^W - U_{t+1}\big]
\end{aligned}
$$

Using the insurance result, we get:

$$
\mathcal{V}_t^W = u(\tilde{c}_t^w) - u(b_t) + (1-\rho_t) \cdot (1-f_t) \cdot \beta \cdot E_t\big[\mathcal{V}_{t+1}^W - U_{t+1}\big]
$$

Using the insurance result and dividing the expression by the marginal utility of consumption gives:

$$
\begin{aligned}
\frac{\mathcal{V}_t^W - U_t}{u'(\tilde{c}_t^w)} &= \frac{u(\tilde{c}_t^w) - u(b_t)}{u'(\tilde{c}_t^w)} + (1-\rho_t) \cdot (1-f_t) \cdot E_t\left[\beta \cdot \frac{\mathcal{V}_{t+1}^W - U_{t+1}}{u'(\tilde{c}_t^w)}\right] \\
&= \frac{u(\tilde{c}_t^w) - u(b_t)}{u'(\tilde{c}_t^w)} + (1-\rho_t) \cdot (1-f_t) \cdot E_t\left[\beta \cdot \frac{u'(\tilde{c}_{t+1}^w)}{u'(\tilde{c}_t^w)} \cdot \frac{\mathcal{V}_{t+1}^W - U_{t+1}}{u'(\tilde{c}_{t+1}^w)}\right] \\
&= \frac{u(\tilde{c}_t^w) - u(b_t)}{u'(\tilde{c}_t^w)} + (1-\rho_t) \cdot (1-f_t) \cdot E_t\left[Q_{t,t+1}^W \cdot \frac{\mathcal{V}_{t+1}^W - U_{t+1}}{u'(\tilde{c}_{t+1}^w)}\right]
\end{aligned}
$$

The value of a firm can be denoted as:

$$
\begin{aligned}
\mathcal{J}_t &= \int_{S_t} y_t(\epsilon) - v(h_t(\epsilon))dG(\epsilon) + \int_{\epsilon_{s,t}}^{\epsilon_{stw,t}} \tau_{stw,t} \cdot \big(\bar{h} - h_{stw,t}(\epsilon)\big)dG(\epsilon) - \Omega_t - \tilde{c}_t^w \\
&\quad - \tau_J + (1-\rho_t) \cdot E_t\big[Q_{t,t+1}^F \cdot \mathcal{J}_{t+1}\big]
\end{aligned}
$$

Insert government budget constraint:

$$
\begin{aligned}
\mathcal{J}_t \;=\; & \int_{S_t} y_t(\epsilon) - v(h_t(\epsilon))dG(\epsilon) + \int_{\epsilon_{s,t}}^{\epsilon_{stw,t}} \tau_{stw,t} \cdot \left(\bar{h} - h_{stw,t}(\epsilon)\right) dG(\epsilon) - \Omega_t - \tilde{c}_t^w \\
& - \int_{\epsilon_{s,t}}^{\epsilon_{stw,t}} \tau_{stw,t} \cdot \left(\bar{h} - h_{stw,t}(\epsilon)\right) dG(\epsilon) - \frac{1-n_t}{n_t} \cdot b_t \\
& + (1-\rho_t) \cdot E_t\left[Q_{t,t+1}^F \cdot \mathcal{J}_{t+1}\right]
\end{aligned}
$$

STW cancels out:

$$
\begin{aligned}
\mathcal{J}_t \;=\; & \int_{S_t} y_t(\epsilon) - v(h_t(\epsilon))dG(\epsilon) - \Omega_t - \tilde{c}_t^w \\
& - \frac{1-n_t}{n_t} \cdot b_t + (1-\rho_t) \cdot E_t\left[Q_{t,t+1}^F \cdot \mathcal{J}_{t+1}\right]
\end{aligned}
$$

Insert value of the firm and surplus of workers into wage equation:

$$
\begin{aligned}
\eta_{t-1} \cdot \Bigg(& \int_{S_t} y_t(\epsilon) - v(h_t(\epsilon))dG(\epsilon) + \int_{\epsilon_{s,t}}^{\epsilon_{stw,t}} \tau_{stw,t} \cdot \left(\bar{h} - h_{stw,t}(\epsilon)\right) dG(\epsilon) - \Omega_t - \tilde{c}_t^w \\
& - \frac{1-n_t}{n_t} \cdot b_t + (1-\rho_t) \cdot E_t\left[Q_{t,t+1}^F \cdot \mathcal{J}_{t+1}\right]\Bigg) \\
= \; & (1-\eta_{t-1}) \cdot \left(\frac{u(\tilde{c}_t^w) - u(b_t)}{u'(\tilde{c}_t^w)} + (1-\rho_t) \cdot E_t\left[Q_{t,t+1}^W \cdot \frac{\mathcal{V}_{t+1} - U_{t+1}}{u'(\tilde{c}_{t+1}^w)} \right]\right)
\end{aligned}
$$

Rearrange for income:

$$
\begin{aligned}
\eta_{t-1} \cdot \tilde{c}_t^w & + (1-\eta_{t-1}) \cdot \frac{u(\tilde{c}_t^w)}{u'(\tilde{c}_t^w)} \\
= \; & \eta_{t-1} \cdot \left(\int_{S_t} \left(y_t(\epsilon) - v(h_t(\epsilon))\right) dG(\epsilon) - \Omega_t - \tilde{c}_t^w - \frac{1-n_t}{n_t} \cdot b_t \right. \\
& \left. + (1-\rho_t) \cdot E_t\left[Q_{t,t+1}^F \cdot \mathcal{J}_{t+1}\right] \right) \\
& + (1-\eta_{t-1}) \cdot \left(\frac{u(b_t)}{u'(\tilde{c}_t^w)} - (1-\rho_t)(1-f_t) \cdot E_t\left[Q_{t,t+1}^W \cdot \frac{\mathcal{V}_{t+1} - U_{t+1}}{u'(\tilde{c}_{t+1}^w)} \right]\right)
\end{aligned}
$$

Insert optimal wage equation again:

$$\begin{aligned}
&\eta_{t-1} \cdot \tilde{c}_t^w + (1-\eta_{t-1}) \cdot \frac{u(\tilde{c}_t^w)}{u'(\tilde{c}_t^w)} \\
&\quad = \eta_{t-1} \cdot \Bigg(\int_{S_t} \big(y_t(\epsilon) - v(h_t(\epsilon))\big)\, dG(\epsilon) - \Omega_t - \tilde{c}_t^w \\
&\qquad - \frac{1-n_t}{n_t} \cdot b_t + (1-\rho_t) \cdot E_t\Big[Q_{t,t+1}^F \cdot \mathscr{J}_{t+1}\Big] \Bigg) \\
&\qquad + (1-\eta_{t-1}) \cdot \left(\frac{u(\tilde{c}_t^w)}{u'(\tilde{c}_t^w)} - (1-\rho_t)(1-f_t) \cdot E_t\left[Q_{t,t+1}^W \cdot \frac{\eta_t}{1-\eta_t} \cdot \mathscr{J}_{t+1} \right] \right)
\end{aligned}$$

Rearrange:

$$\begin{aligned}
&\eta_{t-1} \cdot \tilde{c}_t^w + (1-\eta_{t-1}) \cdot \frac{u(\tilde{c}_t^w)}{u'(\tilde{c}_t^w)} \\
&\quad = \eta_{t-1} \cdot \Bigg(\int_{S_t} \big(y_t(\epsilon) - v(h_t(\epsilon))\big)\, dG(\epsilon) - \Omega_t - \tilde{c}_t^w \\
&\qquad - \frac{1-n_t}{n_t} \cdot b_t + (1-\rho_t) f_t \cdot E_t\Big[Q_{t,t+1}^F \cdot \mathscr{J}_{t+1}\Big] \Bigg) \\
&\qquad + (1-\eta_{t-1}) \cdot \frac{u(b_t)}{u'(\tilde{c}_t^w)} \\
&\qquad - \frac{\eta_{t-1}-\eta_t}{1-\eta_t} \cdot (1-\rho_t)(1-f_t) \cdot E_t\Big[Q_{t,t+1}^W \cdot \mathscr{J}_{t+1}\Big]
\end{aligned}$$

Insert free entry condition:

$$\begin{aligned}
&\eta_{t-1} \cdot \tilde{c}_t^w + (1-\eta_{t-1}) \cdot \frac{u(\tilde{c}_t^w)}{u'(\tilde{c}_t^w)} \\
&\quad = \eta_{t-1} \cdot \Bigg(\int_{S_t} \big(y_t(\epsilon) - v(h_t(\epsilon))\big)\, dG(\epsilon) - \Omega_t - \tilde{c}_t^w \\
&\qquad - \frac{1-n_t}{n_t} \cdot b_t + (1-\rho_t) \cdot \theta_t \cdot k_v \Bigg) \\
&\qquad + (1-\eta_{t-1}) \cdot \frac{u(b_t)}{u'(\tilde{c}_t^w)} \\
&\qquad - \frac{\eta_{t-1}-\eta_t}{1-\eta_t} \cdot (1-\rho_t)(1-f_t) \cdot \frac{k_v}{q_t}
\end{aligned}$$

Hours Worked.
FOC of hours worked outside STW:

$$\begin{aligned}
&\alpha \cdot a_t \cdot \epsilon \cdot (h_t(\epsilon))^{\alpha-1} \cdot g(\epsilon) \cdot (1-\eta_{t-1}) \cdot \big(\mathscr{V}_t - U_t\big) \\
&\quad = v'(h_t(\epsilon)) \cdot u'(\tilde{c}_t^w) \cdot g(\epsilon) \cdot \eta_{t-1} \cdot \mathscr{J}_t
\end{aligned}$$

Inserting optimality condition of the wage gives:

$$\alpha \cdot a_t \cdot \epsilon \cdot (h_t(\epsilon))^{\alpha-1} \;=\; v'\big(h_t(\epsilon)\big)$$

FOC of hours worked on STW:

$$\begin{aligned}\alpha \cdot a_t \cdot \epsilon \cdot (h_{stw,t}(\epsilon))^{\alpha-1} \cdot g(\epsilon) \cdot (1-\eta_{t-1}) \cdot \big(\mathcal{V}_t^W - U_t\big) \\ = \big(v'(h_{stw,t}(\epsilon)) + \tau_{stw,t}\big) \cdot u'(\tilde{c}_t^w) \cdot g(\epsilon) \cdot \eta_{t-1} \cdot \mathcal{J}_t\end{aligned}$$

Inserting the optimality condition for the wage gives:

$$\alpha \cdot a_t \cdot \epsilon \cdot h_{stw,t}(\epsilon)^{\alpha-1} \;=\; v'(h_{stw,t}(\epsilon)) + \tau_{stw,t}$$

This is the condition for the optimal hours' choice.

Separations on STW.
The FOC of the separation threshold is:

$$\begin{aligned}&\eta_{t-1} \cdot g(\epsilon_{s,t}) \cdot (1-f_t) \cdot \beta \cdot E_t\big[\mathcal{V}_{t+1}^W - U_{t+1}\big] \cdot \mathcal{J}_t \\ &+ (1-\eta_{t-1}) \cdot g(\epsilon_{s,t}) \cdot \Big(y_{stw,t}(\epsilon_{s,t}) - v(h_{stw,t}(\epsilon_{s,t})) + \tau_{stw,t} \cdot \big(\bar{h} - h_{stw,t}(\epsilon_{s,t})\big) \\ &+ F + E_t\Big[Q_{t,t+1}^W \cdot \mathcal{J}_{t+1}\Big]\Big) \cdot \big(\mathcal{V}_t - U_t\big) \;=\; 0\end{aligned}$$

Insert optimal wage equation:

$$\begin{aligned}&(1-f_t) \cdot \beta \cdot E_t\left[\frac{\mathcal{V}_{t+1}^W - U_{t+1}}{u'(\tilde{c}_t^w)}\right] + F \\ &+ y_{stw,t}(\epsilon_{s,t}) - v(h_{stw,t}(\epsilon_{s,t})) + \tau_{stw,t} \cdot \big(\bar{h} - h_{stw,t}(\epsilon_{s,t})\big) + E_t\Big[Q_{t,t+1}^W \cdot \mathcal{J}_{t+1}\Big] \;=\; 0\end{aligned}$$

Account for stochastic multiplier of the workers:

$$\begin{aligned}&(1-f_t) \cdot E_t\left[Q_{t,t+1}^W \cdot \frac{\mathcal{V}_{t+1}^W - U_{t+1}}{u'(\tilde{c}_t^w)}\right] + F \\ &+ y_{stw,t}(\epsilon_{s,t}) - v(h_{stw,t}(\epsilon_{s,t})) + \tau_{stw,t} \cdot \big(\bar{h} - h_{stw,t}(\epsilon_{s,t})\big) + E_t\Big[Q_{t,t+1}^W \cdot \mathcal{J}_{t+1}\Big] \;=\; 0\end{aligned}$$

Use optimal wage equation again:

$$\begin{aligned}&(1-f_t) \cdot \frac{\eta_t}{1-\eta_t} \cdot E_t\Big[Q_{t,t+1}^W \cdot \mathcal{J}_{t+1}\Big] + F \\ &+ y_{stw,t}(\epsilon_{s,t}) - v(h_{stw,t}(\epsilon_{s,t})) + \tau_{stw,t} \cdot \big(\bar{h} - h_{stw,t}(\epsilon_{s,t})\big) + E_t\Big[Q_{t,t+1}^W \cdot \mathcal{J}_{t+1}\Big] \;=\; 0\end{aligned}$$

Rearrange:

$$y_{stw,t}(\epsilon_{s,t}) - v(h_{stw,t}(\epsilon_{s,t})) + \tau_{stw,t} \cdot \big(\bar{h} - h_{stw,t}(\epsilon_{s,t})\big) + F \\ + (1 - f_t) \cdot \frac{1 - \eta_t \cdot f_t}{1 - \eta_t} \cdot E_t\Big[Q^W_{t,t+1} \cdot \mathcal{J}_{t+1}\Big] \;=\; 0$$

Insert free-entry condition:

$$y_{stw,t}(\epsilon_{s,t}) - v(h_{stw,t}(\epsilon_{s,t})) + \tau_{stw,t} \cdot \big(\bar{h} - h_{stw,t}(\epsilon_{s,t})\big) + F \\ + (1 - f_t) \cdot \frac{1 - \eta_t \cdot f_t}{1 - \eta_t} \cdot \frac{k_v}{q_t} \;=\; 0$$

Separations outside STW.
The FOC of the separation threshold is:

$$\eta_{t-1} \cdot g(\xi_{s,t}) \cdot (1 - f_t) \cdot \beta \cdot E_t\big[\mathcal{V}^W_{t+1} - U_{t+1}\big] \cdot \mathcal{J}_t \\ + (1 - \eta_{t-1}) \cdot g(\xi_{s,t}) \cdot \Big(y_t(\xi_{s,t}) - v(h_t(\xi_{s,t})) + F + E_t\Big[Q^W_{t,t+1} \cdot \mathcal{J}_{t+1}\Big]\Big) \cdot \big(\mathcal{V}_t - U_t\big) \\ = \; 0$$

Insert optimal wage equation:

$$y_t(\xi_{s,t}) - v(h_t(\xi_{s,t})) + F + E_t\Big[Q^W_{t,t+1} \cdot \mathcal{J}_{t+1}\Big] \\ + (1 - f_t) \cdot \beta \cdot E_t\left[\frac{\mathcal{V}^W_{t+1} - U_{t+1}}{u'(\tilde{c}^W_t)}\right] \;=\; 0$$

Account for stochastic multiplier of the workers:

$$y_t(\xi_{s,t}) - v(h_t(\xi_{s,t})) + F + E_t\Big[Q^W_{t,t+1} \cdot \mathcal{J}_{t+1}\Big] \\ + (1 - f_t) \cdot E_t\left[Q^W_{t,t+1} \cdot \frac{\mathcal{V}^W_{t+1} - U_{t+1}}{u'(\tilde{c}^W_{t+1})}\right] \;=\; 0$$

Use optimal wage equation again:

$$y_t(\xi_{s,t}) - v(h_t(\xi_{s,t})) + F + E_t\Big[Q^W_{t,t+1} \cdot \mathcal{J}_{t+1}\Big] \\ + (1 - f_t) \cdot \frac{\eta_t}{1 - \eta_t} \cdot E_t\Big[Q^W_{t,t+1} \cdot \mathcal{J}_{t+1}\Big] \;=\; 0$$

Rearrange:

$$y_t(\xi_{s,t}) - v(h_t(\xi_{s,t})) + F \\ + (1 - f_t) \cdot \frac{1 - \eta_t \cdot f_t}{1 - \eta_t} \cdot E_t\Big[Q^W_{t,t+1} \cdot \mathcal{J}_{t+1}\Big] \;=\; 0$$

Insert free-entry condition:

$$y_t(\xi_{s,t}) - v(h_t(\xi_{s,t})) + F + (1 - f_t) \cdot \frac{1 - \eta_t \cdot f_t}{1 - \eta_t} \cdot \frac{k_v}{q_t} = 0$$

Appendix 1.G Derivation Social Planner

$$W_t^P = \max_{\theta_t, \epsilon_{s,t}, h_t(\epsilon)} n_t \cdot \int_0^\infty u(\tilde{c}_t^w(\epsilon))dG(\epsilon) + (1 - n_t) \cdot u(c_t^u) + \nu_t^f u(c_t^f) + \beta \cdot E_t\left[W_{t+1}^P\right]$$

subject to

$$\begin{aligned}
(I) \quad n_t \cdot \int_0^\infty \tilde{c}_t^w(\epsilon)dG(\epsilon) + (1 - n_t) \cdot c_t^u + \nu_t^f \cdot c_t^f &= \\
n_t \cdot \int_{\epsilon_{s,t}}^\infty y(\epsilon, h_t(\epsilon)) - v_t(h_t(\epsilon))dG(\epsilon) & \\
- \theta \cdot (1 - n_t + G(\epsilon_{s,t}) \cdot n_t) \cdot k_v - n_t \cdot G(\epsilon_{s,t}) \cdot F & \\
(II) \quad n_{t+1} = (1 - G(\epsilon_{s,t})) \cdot n_t + f(\theta_t) \cdot (1 - n_t + G(\epsilon_{s,t}) \cdot n_t) &
\end{aligned}$$

First Order Conditions (FOC) of the Planner .
FOC for consumption of employed workers:

$$\begin{aligned}
\frac{\partial}{\partial \tilde{c}_t^w(\epsilon)} &= \beta^t \cdot n_t \cdot g(\epsilon) \cdot u'(c_t(\epsilon)) - \beta^t \cdot n_t \cdot g(\epsilon) \cdot \lambda_{m,t} = 0 \\
\Leftrightarrow \lambda_{m,t} &= u'(\tilde{c}_t(\epsilon))
\end{aligned}$$

FOC for consumption of unemployed workers:

$$\begin{aligned}
\frac{\partial}{\partial \tilde{c}_t^w(\epsilon)} &= \beta^t \cdot (1 - n_t) \cdot u'(c_t^u) - \beta^t \cdot (1 - n_t) \cdot \lambda_{m,t} = 0 \\
\Leftrightarrow \lambda_{m,t} &= u'(c_t^u)
\end{aligned}$$

FOC for consumption of firm owners:

$$\begin{aligned}
\frac{\partial}{\partial \tilde{c}_t^w(\epsilon)} &= \beta^t \cdot \nu_t^f \cdot u'(c_t^f) - \beta^t \cdot \nu_t^f \cdot \lambda_{m,t} = 0 \\
\Leftrightarrow \lambda_{m,t} &= u'(c_t^f)
\end{aligned}$$

FOC for hours worked:

$$\frac{\partial \mathscr{L}}{\partial h_t(\epsilon)} = \beta^t \cdot \lambda_{m,t} \cdot n_t \cdot \left(\frac{\partial y_t(\epsilon, h_t(\epsilon))}{\partial h_t(\epsilon)} - v'(h_t(\epsilon)) \right) \cdot g(\epsilon) = 0$$
$$\Leftrightarrow \frac{\partial y_t(\epsilon, h_t(\epsilon))}{\partial h_t(\epsilon)} = v'(h_t(\epsilon))$$

This is equation 1.2.7.

FOC for employment:

$$\frac{\partial \mathscr{L}}{\partial n_t} = \beta^t \cdot \lambda_{m,t} \cdot \left(\int_{\epsilon_{s,t}}^{\infty} [y_t(\epsilon) - v(h_t(\epsilon))] dG(\epsilon) - G(\epsilon_{s,t}) \cdot F + \theta_t \cdot k_v \right)$$
$$- \beta^{t-1} \cdot \lambda_{n,t-1} + \beta^t \cdot \left(1 - \theta_t \cdot q(\theta_t)\right) \cdot (1 - G(\epsilon_{s,t})) \cdot \lambda_{n,t} = 0$$

$$\Leftrightarrow \quad \lambda_{n,t-1} = \beta \cdot \lambda_{m,t} \cdot \left(\int_{\epsilon_{s,t}}^{\infty} [y_t(\epsilon) - v(h_t(\epsilon))] dG(\epsilon) - G(\epsilon_{s,t}) \cdot F + \theta_t \cdot k_v \right)$$
$$+ \beta \cdot \left(1 - \theta_t \cdot q(\theta_t)\right) \cdot (1 - G(\epsilon_{s,t})) \cdot \lambda_{n,t}$$

$$\Leftrightarrow \quad \frac{\lambda_{n,t-1}}{\lambda_{m,t-1}} = \beta \cdot \frac{\lambda_{m,t}}{\lambda_{m,t-1}} \cdot \left(\int_{\epsilon_{s,t}}^{\infty} [y_t(\epsilon) - v(h_t(\epsilon))] dG(\epsilon) - G(\epsilon_{s,t}) \cdot F + \theta_t \cdot k_v \right)$$
$$+ \beta \cdot \frac{\lambda_{m,t}}{\lambda_{m,t-1}} \cdot \left(1 - \theta_t \cdot q(\theta_t)\right) \cdot (1 - G(\epsilon_{s,t})) \cdot \frac{\lambda_{n,t}}{\lambda_{m,t}}$$

FOC for the labor market tightness:

$$\frac{\partial \mathscr{L}}{\partial \theta_t} = -\beta^t \cdot \lambda_{m,t} \cdot k_v \cdot \left(1 - n_t + G(\epsilon_{s,t}) \cdot n_t\right)$$
$$+ \beta^t \cdot \lambda_{n,t} \cdot \left(q(\theta_t) + \theta_t \cdot q'(\theta_t)\right) \cdot \left(1 - n_t + G(\epsilon_{s,t}) \cdot n_t\right) = 0$$

$$\Leftrightarrow \quad \lambda_{n,t} = \frac{\lambda_{m,t}}{1 + \theta_t \cdot \frac{q'(\theta_t)}{q(\theta_t)}} \cdot \frac{k_v}{q(\theta_t)}$$

Note that we can express the elasticity of the matching function with respect to unemployment as:

$$\theta_t \cdot \frac{q'(\theta_t)}{q(\theta_t)} = -\gamma \cdot \theta_t \cdot \frac{\chi \cdot \theta_t^{-\gamma-1}}{\chi \cdot \theta_t^{-\gamma}} = -\gamma$$

Using this expression gives:

$$\frac{\lambda_{n,t}}{\lambda_{m,t}} = \frac{1}{1-\gamma} \cdot \frac{k_v}{q(\theta_t)}$$

FOC separation threshold:

$$-\frac{\partial \mathscr{L}}{\partial \epsilon_{s,t}} = \beta^t \cdot \lambda_{m,t} \cdot \big(y_t(\epsilon_{s,t}, h_t(\epsilon_{s,t})) - v(h_t(\epsilon_{s,t})) + F + \theta \cdot k_v\big) \cdot g(\epsilon_{s,t})$$
$$+ \beta^t \cdot \lambda_{n,t} \cdot (1 - f(\theta_t)) \cdot g(\epsilon_{s,t}) = 0$$

This is equivalent to:

$$\frac{\partial \mathscr{L}}{\partial \epsilon_{s,t}} = y_t(\epsilon_{s,t}, h_t(\epsilon_{s,t})) - v(h_t(\epsilon_{s,t})) + F + \theta_t \cdot k_v + (1 - f(\theta_t)) \cdot \frac{\lambda_{n,t}}{\lambda_{m,t}} \cdot = 0$$

Distribution of Income. It must hold:

$$\lambda_{m,t} = \lambda_{m,t} = \lambda_{m,t} \quad \Leftrightarrow \quad u'(\tilde{c}_t(\epsilon)) = u'(\tilde{c})^u_t = u'(c^f_t)$$
$$\Leftrightarrow \quad \tilde{c}_t := \tilde{c}_t(\epsilon) = \tilde{c}^u_t = c^f_t$$

Planner's Job-Creation Condition.
Insert Lagrange multiplier:

$$\frac{1}{1-\gamma} \cdot \frac{k_v}{q(\theta_{t-1})} = \beta \cdot \frac{u'(\tilde{c}_t)}{u'(\tilde{c}_{t-1})} \cdot \Big(\int_{\epsilon_{s,t}}^{\infty} [y_t(\epsilon) - v(h_t(\epsilon))]\, dG(\epsilon) - G(\epsilon_{s,t}) \cdot F$$
$$+ (1 - G(\epsilon_{s,t})) \cdot \theta_t \cdot k_v \Big)$$
$$+ \beta \cdot \frac{u'(\tilde{c}_t)}{u'(\tilde{c}_{t-1})} \cdot (1 - \theta_t \cdot q(\theta_t)) \cdot (1 - G(\epsilon_{s,t})) \cdot \frac{1}{1-\gamma} \cdot \frac{k_v}{q(\theta_t)}$$

Use that $\theta_t = \frac{\theta_t \cdot q(\theta_t)}{q(\theta_t)}$ gives:

$$\frac{1}{1-\gamma} \cdot \frac{k_v}{q(\theta_{t-1})} = \beta \cdot \frac{u'(\tilde{c}_t)}{u'(\tilde{c}_{t-1})} \cdot \Big(\int_{\epsilon_{s,t}}^{\infty} [y_t(\epsilon) - v(h_t(\epsilon))]\, dG(\epsilon) - G(\epsilon_{s,t}) \cdot F$$
$$+ (1 - G(\epsilon_{s,t})) \cdot \theta_t \cdot k_v \Big)$$
$$+ \beta \cdot \frac{u'(\tilde{c}_t)}{u'(\tilde{c}_{t-1})} \cdot (1 - G(\epsilon_{s,t})) \cdot \left[\frac{1 - \theta_t \cdot q(\theta_t)}{1-\gamma} + \theta_t \cdot q(\theta_t) \right] \cdot \frac{k_v}{q(\theta_t)}$$

Rearranging:

$$\frac{1}{1-\gamma}\cdot\frac{k_v}{q(\theta_{t-1})} = \beta\cdot\frac{u'(\tilde{c}_t)}{u'(\tilde{c}_{t-1})}\cdot\left(\int_{\epsilon_{s,t}}^{\infty}[y_t(\epsilon)-v(h_t(\epsilon))]\,dG(\epsilon)-G(\epsilon_{s,t})\cdot F\right)$$
$$+\beta\cdot\frac{u'(\tilde{c}_t)}{u'(\tilde{c}_{t-1})}\cdot(1-G(\epsilon_{s,t}))\cdot\frac{1-\gamma\cdot\theta_t\cdot q(\theta_t)}{1-\gamma}\cdot\frac{k_v}{q(\theta_t)}$$

Planner's Separation Decision.
Insert FOC for employment into FOC for separation threshold:

$$\frac{\partial\mathcal{L}}{\partial\epsilon_{s,t}} = y_t(\epsilon_{s,t},h_t(\epsilon_{s,t}))-v(h_t(\epsilon_{s,t}))+F+\theta_t\cdot k_v+(1-f(\theta_t))\cdot\frac{k_v}{q(\theta_t)} = 0$$

Using that $\theta_t=\frac{\theta_t\cdot q(\theta_t)}{q(\theta_t)}$ again gives:

$$\left[y_t(\epsilon_{s,t})-v(h_t(\epsilon_{s,t}))\right]+F+\frac{1-\gamma\cdot\theta_t\cdot q(\theta_t)}{1-\gamma}\cdot\frac{k_v}{q(\theta_t)} = 0$$

With $f_t=\theta_t\cdot q(\theta_t)$ and $q_t=q(\theta_t)$, this is equivalent to the separation decision of the planner expressed in the paper.

Appendix 1.H Derivation Decentralized Economy

Derivation Job-Creation Condition.

Joint surplus:

$$\mathcal{J}_t+\frac{\mathcal{V}_t-U_t}{u'(\tilde{c}_t^w)} = \int_{S_t}y_t(\epsilon)-v(h_t(\epsilon))dG(\epsilon)-\Omega_t-\frac{b_t}{n_t}-\tilde{c}_t^w+\frac{u(\tilde{c}_t^w)-u(b_t)}{u'(\tilde{c}_t^w)}$$
$$+(1-G(\epsilon_{s,t}))\cdot E_t\left[Q_{t+1}^F\cdot\mathcal{J}_{t+1}+\beta\cdot\frac{\mathcal{V}_{t+1}-U_{t+1}}{u'(\tilde{c}_t^w)}\right]$$

Insert optimal wage equation:

$$\frac{J_t}{1-\eta_{t-1}} = \int_{S_t}y_t(\epsilon)-v(h_t(\epsilon))dG(\epsilon)-\Omega_t-\frac{b_t}{n_t}-\tilde{c}_t^w+\frac{u(\tilde{c}_t^w)-u(b_t)}{u'(\tilde{c}_t^w)}$$
$$+(1-G(\epsilon_{s,t}))\cdot E_t\left[Q_{t+1}^F\cdot\mathcal{J}_{t+1}+(1-f_t)\cdot Q_{t+1}^W\cdot\frac{\eta_t}{1-\eta_t}\cdot J_{t+1}\right]$$

Assumption: $Q^F_{t+1} = Q^W_{t+1}$

$$\begin{aligned} \frac{J_t}{1-\eta_{t-1}} &= \int_{S_t} y_t(\epsilon) - v(h_t(\epsilon))dG(\epsilon) - \Omega_t - \frac{b_t}{n_t} - \tilde{c}^w_t + \frac{u(\tilde{c}^w_t) - u(b_t)}{u'(\tilde{c}^w_t)} \\ &\quad + (1 - G(\epsilon_{s,t})) \cdot \frac{\left(1 - \eta_t \cdot f_t\right)}{1-\eta_t} \cdot E_t\left[Q^W_{t+1} \cdot J_{t+1}\right] \end{aligned}$$

Insert free-entry condition:

$$\begin{aligned} \frac{1}{1-\eta_{t-1}} \cdot \frac{1}{Q^W_t} \cdot \frac{k_v}{q_{t-1}} &= \int_{S_t} y_t(\epsilon) - v(h_t(\epsilon))dG(\epsilon) - \Omega_t - \frac{b_t}{n_t} - \tilde{c}^w_t + \frac{u(\tilde{c}^w_t) - u(b_t)}{u'(\tilde{c}^w_t)} \\ &\quad + (1 - G(\epsilon_{s,t})) \cdot \frac{\left(1 - \eta_t \cdot f_t\right)}{1-\eta_t} \cdot \frac{k_v}{q_t} \end{aligned}$$

Rearrange:

$$\begin{aligned} \frac{1}{1-\eta_{t-1}} \cdot \frac{k_v}{q_{t-1}} &= Q^W_t \cdot \left(\int_{S_t} [y_t(\epsilon) - v(h_t(\epsilon))]\,dG(\epsilon) - \Omega_t - \frac{b_t}{n_t} - \tilde{c}^w_t + \frac{u(\tilde{c}^w_t) - u(b_t)}{u'(\tilde{c}^w_t)} \right) \\ &\quad + Q^W_t \cdot \left((1 - G(\epsilon_{s,t})) \cdot \frac{1 - \eta_t \cdot f_t}{1-\eta_t} \cdot \frac{k_v}{q_t} \right) \end{aligned}$$

Thus:

$$\begin{aligned} \frac{1}{1-\eta_t} \cdot \frac{k_v}{q_t} &= E_t\Bigg[Q^W_{t+1} \cdot \Bigg(\int_{S_{t+1}} [y_{t+1}(\epsilon) - v(h_{t+1}(\epsilon))]\,dG(\epsilon) - \Omega_{t+1} - \frac{b_{t+1}}{n_{t+1}} - \tilde{c}^w_{t+1} \\ &\quad + \frac{u(\tilde{c}^w_{t+1}) - u(b_{t+1})}{u'(\tilde{c}^w_{t+1})} \Bigg) \Bigg] \\ &\quad + E_t\left[Q^W_{t+1} \cdot \left((1 - G(\epsilon_{s,t+1})) \cdot \frac{1 - \eta_{t+1} \cdot f_{t+1}}{1-\eta_{t+1}} \cdot \frac{k_v}{q_{t+1}} \right) \right] \end{aligned}$$

Appendix 1.I Derivation Optimal Policy

The Ramsey planner chooses the minimum hours reduction condition D_t, the STW benefits $\tau_{stw,t}$, and UI benefits b_t subject to the labor market equilibrium. Check Appendix 1.F and 1.H for a derivation of core labor market equilibrium conditions.

$$\begin{aligned}W_t^G \;=\; \max_{D_t,\tau_{stw,t},b_t} & (1-n_t)\cdot u(b_t) + n_t \cdot u(\tilde{c}_t^w) \\ & + \nu_t^f \cdot u\Bigg(\Big[n_t \cdot \int_{\mathscr{B}_t} \big(y_t(\epsilon) - v(h_t(\epsilon))\big) dG(\epsilon) - n_t \cdot \Omega_t - n_t \cdot \tilde{c}_t^w \\ & \qquad - (1-n_t)\cdot b_t - \theta_t \cdot \big(1 - n_t + \rho_t \cdot n_t\big)\cdot k_v\Big] / \nu_t^f\Bigg) \\ & + \beta \cdot E_t W_{t+1}^G\end{aligned}$$

Set:

$$\mathscr{B}_{\sqcup} \;=\; [\epsilon_{s,t}, \epsilon_{stw,t}] \cup [\xi_{s,t}, \infty)$$

Job-Creation Condition:

$$\begin{aligned}\frac{1}{1-\eta_t}\cdot\frac{k_v}{q(\theta_t)} \;=\; & E_t\Big[Q_{t,t+1}^W\Big(\int_{S_t} y_{t+1}(\epsilon) - v(h_{t+1}(\epsilon))dG(\epsilon) - \Omega_{t+1} - \frac{b_{t+1}}{n_{t+1}} \\ & \qquad - \tilde{c}_{t+1} + \frac{u(\tilde{c}_{t+1}) - u(b_{t+1})}{u'(\tilde{c}_{t+1})}\Big)\Big] \\ & + E_t\Bigg[Q_{t,t+1}^W\Bigg((1-\rho_{t+1})\cdot\frac{1-\eta_{t+1}\cdot f(\theta_{t+1})}{1-\eta_{t+1}}\cdot\frac{k_v}{q(\theta_{t+1})}\Bigg)\Bigg]\end{aligned}$$

Income Insurance contract:

$$\begin{aligned}\eta_{t-1}\cdot\tilde{c}_t^w + (1-\eta_{t-1})\cdot\frac{u(\tilde{c}_t^w)}{u'(\tilde{c}_t^w)} & \\ = \; \eta_{t-1}\cdot\Bigg(\int_{S_t} & [y_t(\epsilon) - v(h_t(\epsilon))]dG(\epsilon) - \Omega_t - \frac{b_t}{n_t} + (1-\rho_t)\cdot\theta_t\cdot k_v\Bigg) \\ + (1-\eta_{t-1})\cdot & \frac{u(b_t)}{u'(\tilde{c}_t^w)} \\ - \frac{\eta_{t-1}-\eta_t}{1-\eta_t} & \cdot (1-\rho_t)\cdot(1-f_t)\cdot\frac{k_v}{q_t}\end{aligned}$$

Separation conditions without access to STW:

$$y_t(\xi_{s,t}, h_t(\xi_{s,t})) - v(h_t(\xi_{s,t})) + F + \frac{1-\eta_t\cdot f_t}{1-\eta_t}\cdot\frac{k_v}{q_t} \;=\; 0$$

Separation conditions on STW:

$$\begin{aligned}y_t(\epsilon_{s,t}, h_{stw,t}(\epsilon_{s,t})) - v(h_{stw,t}(\epsilon_{s,t})) + (\bar{h} - h_{stw,t}(\epsilon_{s,t}))\cdot\tau_{stw,t} & \\ + F + \frac{1-\eta_t\cdot f_t}{1-\eta_t}\cdot\frac{k_v}{q_t} \;=\; 0 &\end{aligned}$$

Working hours outside STW:

$$\alpha \cdot a_t \cdot \epsilon \cdot h_t(\epsilon)^{\alpha-1} \;=\; h_t(\epsilon)^{\psi}$$

Working hours on STW:

$$\alpha \cdot a_t \cdot \epsilon \cdot h_{stw,t}(\epsilon)^{\alpha-1} \;=\; h_{stw,t}(\epsilon)^{\psi} + \tau_{stw,t}$$

Welfare costs of STW:

$$\Omega_t \;=\; \int_{\epsilon_{s,t}}^{\epsilon_{stw,t}} [y_t(\epsilon, h_t(\epsilon)) - v\big(h_t(\epsilon)\big) - y_t(\epsilon, h_{stw,t}(\epsilon)) + v\big(h_{stw,t}(\epsilon)\big)]dG(\epsilon)$$

Stochastic discount factor:

$$Q^W_{t,t+1} \;=\; \beta \cdot \frac{u'(\tilde{c}^w_{t+1})}{u'(\tilde{c}^w_t)}$$

Law of motion of employment:

$$n_{t+1} \;=\; (1-\rho_t) \cdot n_t + f(\theta_t) \cdot (1 - n_t + \rho_t \cdot n_t)$$

Separation rate:

$$\rho_t \;=\; [\max\{G(\xi_{s,t}) - G(\epsilon_{stw,t}), 0\} + G(\epsilon_{s,t})]$$

1.I.1 Constrained Efficient Separation and Job-Creation Conditions

FOC for the consumption equivalent:

$$\begin{aligned}
\frac{\partial}{\partial \tilde{c}^w_t} \;=\; & \underbrace{n_t \cdot \left[u'(\tilde{c}^w_t) - \frac{\nu^f_t}{\nu^f_t} \cdot u'(\tilde{c}^w_t)\right]}_{=\,0} - \lambda_{\theta,t-1} \cdot Q_t \cdot \frac{u''(\tilde{c}^w_t)}{u'(\tilde{c}^w_t)} \cdot \frac{u(\tilde{c}^w_t) - u(b_t)}{u'(\tilde{c}^w_t)} \\
& - \lambda_{c,t} \cdot \left[1 - (1-\eta_{t-1}) \cdot \frac{u''(\tilde{c}^w_t)}{u'(\tilde{c}^w_t)} \cdot \frac{u(\tilde{c}^w_t) - u(b_t)}{u'(\tilde{c}^w_t)}\right] \\
& + \lambda_{Q,t} \cdot \beta \cdot \frac{u''(\tilde{c}^w_t)}{u'(\tilde{c}^w_{t-1})} - E_t\left[\lambda_{Q,t+1} \cdot \beta \cdot u''(\tilde{c}^w_t) \cdot \frac{u'(\tilde{c}^w_{t+1})}{u'(\tilde{c}^w_t)^2}\right] \;=\; 0
\end{aligned}$$

Changes in consumption influence the value of the joint surplus and thus hiring incentives:

$$\Leftrightarrow \lambda_{c,t} = \lambda_{\theta,t-1} \cdot Q_t \cdot BE_t + \eta \cdot u''(\tilde{c}_t^w) \cdot \frac{u'(\tilde{c}_{t-1}^w) \cdot u'(\tilde{c}_{t+1}^w) \cdot \lambda_{Q,t+1} - u'(\tilde{c}_t^w)^2 \cdot \lambda_{Q,t}}{u'(\tilde{c}_{t-1}^w) \cdot u'(\tilde{c}_t^w)^2}$$

$$\text{with} \quad BE_t = \frac{\left(-\frac{u''(\tilde{c}_t^w)}{u'(\tilde{c}_t^w)}\right) \cdot \frac{u(\tilde{c}_t^w)-u(b_t)}{u'(\tilde{c}_t^w)}}{1 + (1-\eta_{t-1}) \cdot \left(-\frac{u''(\tilde{c}_t^w)}{u'(\tilde{c}_t^w)}\right) \cdot \frac{u(\tilde{c}_t^w)-u(b_t)}{u'(\tilde{c}_t^w)}} < \frac{1}{1-\eta_{t-1}}$$

FOC for labor market tightness:

$$\begin{aligned}
\frac{\partial}{\partial \theta_t} = & -\underbrace{(1-n_t + G(\epsilon_{s,t}) \cdot n_t)}_{u_t} \cdot k_v \cdot \underbrace{\frac{\nu_t^f}{\nu_t^f} \cdot u'(c_t^f)}_{u'(\tilde{c}_t^w)} \\
& + \underbrace{(1-n_t + G(\epsilon_{s,t}) \cdot n_t)}_{u_t} \cdot (1-\gamma) \cdot q(\theta_t) \cdot E_t\left[\lambda_{n,t+1}\right] \\
& - \frac{1}{1-\eta} \cdot k_v \cdot \gamma \cdot \frac{\lambda_{\theta,t}}{f(\theta_t)} \\
& + \frac{1}{1-\eta} \cdot \left(\frac{\gamma}{f(\theta_t)} - \eta\right) \cdot \left(\beta \cdot \left(1 - G(\epsilon_{s,t})\right) \cdot \lambda_{\theta,t-1} - \lambda_{\epsilon,t} - \lambda_{stw,t}\right) \cdot k_v \\
& + \lambda_{c,t} \cdot \left(\eta \cdot (1 - G(\epsilon_{s,t})) \cdot k_v + \frac{\eta_t - \eta_{t-1}}{1-\eta_t} \cdot (1-\rho_t) \cdot \left(\frac{\gamma}{f(\theta_t)} - 1\right) \cdot k_v\right) = 0
\end{aligned}$$

Rearranging gives:

$$\begin{aligned}
\Leftrightarrow & \frac{1}{1-\gamma} \cdot \frac{k_v}{q(\theta_t)} + \frac{1}{1-\gamma} \cdot \frac{k_v}{q(\theta_t)} \\
& \cdot \frac{1}{1-\eta} \cdot \frac{1}{m(\theta_t)} \cdot \frac{\gamma \cdot \lambda_{\theta,t} - \left(\gamma - f(\theta_t) \cdot \eta\right) \cdot \left(\beta \cdot (1 - G(\epsilon_{s,t})) \cdot \lambda_{\theta,t-1} - \lambda_{\epsilon,t} - \lambda_{stw,t}\right)}{u'(\tilde{c}_t^w)} \\
& - \frac{\lambda_{c,t}}{u'(\tilde{c}_t^w)} \cdot \left(\eta \cdot (1 - G(\epsilon_{s,t})) \cdot k_v + \frac{\eta_t - \eta_{t-1}}{1-\eta_t} \cdot (1-\rho_t) \cdot \left(\frac{\gamma}{f(\theta_t)} - 1\right) \cdot k_v\right) \\
& \cdot \frac{\eta \cdot (1 - G(\epsilon_{s,t}))}{u_t} \cdot \frac{1}{1-\gamma} \cdot \frac{k_v}{q(\theta_t)} \\
= & \ \beta \cdot \frac{E_t\left[\lambda_{n,t+1}\right]}{u'(\tilde{c}_t^w)}
\end{aligned}$$

Define χ_t as:

$$
\begin{aligned}
\chi_t &= \\
&\frac{1}{1-\eta} \cdot \frac{1}{m(\theta_t)} \cdot \frac{\gamma \cdot \lambda_{\theta,t} - \left(\gamma - f(\theta_t) \cdot \eta\right) \cdot \left(\beta \cdot (1 - G(\epsilon_{s,t})) \cdot \lambda_{\theta,t-1} - \lambda_{\epsilon,t} - \lambda_{stw,t}\right)}{u'(\tilde{c}_t^w)} \\
&- \frac{\lambda_{c,t}}{u'(\tilde{c}_t^w)} \cdot \left(\eta \cdot (1 - G(\epsilon_{s,t})) \cdot k_v + \frac{\eta_t - \eta_{t-1}}{1 - \eta_t} \cdot (1 - \rho_t) \cdot \left(\frac{\gamma}{f(\theta_t)} - 1\right)\right. \\
&\left. \cdot k_v \cdot \frac{\eta \cdot (1 - G(\epsilon_{s,t}))}{u_t}\right.
\end{aligned}
$$

This simplifies the expression and gives the first building block for the constrained optimal job-creation and separation condition:

$$
\beta \cdot \frac{E_t\left[\lambda_{n,t+1}\right]}{u'(\tilde{c}_t^w)} = \frac{1 + \chi_t}{1 - \gamma} \cdot \frac{k_v}{q(\theta_t)}
$$

FOC for employment:

$$
\begin{aligned}
\frac{\partial}{\partial n_t} &= u(\tilde{c}_t^w) - u(b_t) \\
&+ \underbrace{\frac{\nu_t^f}{\nu_t^f} \cdot u'(c_t^f)}_{u'(\tilde{c}_t^w)} \cdot \left(\int_{\epsilon_{s,t}}^{\infty} y_t(\epsilon, h_t(\epsilon)) - v(h_t(\epsilon)) dG(\epsilon) - \Omega_t + b_t + (1 - G(\epsilon_{s,t})) \cdot \theta_t \cdot k_v\right) \\
&+ (\beta \cdot \lambda_{\theta,t-1} + \eta \cdot \lambda_{c,t}) \cdot \frac{b_t}{n_t^2} \\
&- \lambda_{n,t} + \beta \cdot (1 - f(\theta_t)) \cdot (1 - G(\epsilon_{s,t})) \cdot E_t\left[\lambda_{n,t+1}\right]
\end{aligned}
$$

Rearranging gives the second building block of the constrained efficient job-creation condition:

$$
\begin{aligned}
\Leftrightarrow & \frac{u'(\tilde{c}_{t-1}^w)}{u'(\tilde{c}_t^w)} \cdot \frac{\lambda_{n,t}}{u'(\tilde{c}_{t-1}^w)} \\
&= \int_{\epsilon_{s,t}}^{\infty} y_t(\epsilon, h_t(\epsilon)) - v(h_t(\epsilon)) dG(\epsilon) - \Omega_t + b_t + (1 - G(\epsilon_{s,t})) \cdot \theta_t \cdot k_v \\
&+ \frac{u(\tilde{c}_t^w) - u(b_t)}{u'(\tilde{c}_t^w)} + \frac{(\beta \cdot \lambda_{\theta,t-1} + \lambda_{c,t})}{n_t \cdot u'(\tilde{c}_t^w)} \cdot \frac{b_t}{n_t} \\
&+ \beta \cdot (1 - f(\theta_t)) \cdot (1 - G(\epsilon_{s,t})) \cdot \frac{E_t[\lambda_{n,t+1}]}{u'(\tilde{c}_t^w)}
\end{aligned}
$$

Inserting the first building block of the constrained efficient job-creation condition:

$$
\begin{aligned}
&\frac{1+\chi_{t-1}}{1-\gamma}\cdot\frac{k_v}{q(\theta_{t-1})} \;=\\
&\beta\cdot\frac{u'(\tilde{c}^w_t)}{u'(\tilde{c}^w_{t-1})}\cdot\left(\int_{\epsilon_{s,t}}^{\infty} y_t(\epsilon,h_t(\epsilon))-v(h_t(\epsilon))dG(\epsilon)-\Omega_t+b_t+(1-G(\epsilon_{s,t}))\cdot\theta_t\cdot k_v\right)\\
&+\beta\cdot\frac{u'(\tilde{c}^w_t)}{u'(\tilde{c}^w_{t-1})}\cdot\left(\frac{u(\tilde{c}^w_t)-u(b_t)}{u'(\tilde{c}^w_t)}+\frac{(\beta\cdot\lambda_{\theta,t-1}+\lambda_{c,t})}{n_t\cdot u'(\tilde{c}^w_t)}\cdot\frac{b_t}{n_t}\right)\\
&+\beta\cdot\frac{u'(\tilde{c}^w_t)}{u'(\tilde{c}^w_{t-1})}\cdot(1-G(\epsilon_{s,t}))\cdot(1-f(\theta_t))\cdot\frac{1+\chi_t}{1-\gamma}\cdot\frac{k_v}{q(\theta_t)}
\end{aligned}
$$

Rearranging gives the **constrained efficient job-creation condition:**

$$
\begin{aligned}
\Leftrightarrow&\frac{1+\chi_{t-1}}{1-\gamma}\cdot\frac{k_v}{q(\theta_{t-1})} \;=\\
&\beta\cdot\frac{u'(\tilde{c}^w_t)}{u'(\tilde{c}^w_{t-1})}\cdot\left(\int_{\epsilon_{s,t}}^{\infty} y_t(\epsilon,h_t(\epsilon))-v(h_t(\epsilon))dG(\epsilon)-\Omega_t+b_t\right)\\
&+\beta\cdot\frac{u'(\tilde{c}^w_t)}{u'(\tilde{c}^w_{t-1})}\cdot\left(\frac{u(\tilde{c}^w_t)-u(b_t)}{u'(\tilde{c}^w_t)}+\frac{(\beta\cdot\lambda_{\theta,t-1}+\lambda_{c,t})}{n_t\cdot u'(\tilde{c}^w_t)}\cdot\frac{b_t}{n_t}\right)\\
&+\beta\cdot\frac{u'(\tilde{c}^w_t)}{u'(\tilde{c}^w_{t-1})}\cdot(1-G(\epsilon_{s,t}))\cdot\frac{(1-\gamma\cdot f(\theta_t))+(1-f(\theta_t))\cdot\chi_t}{1-\gamma}\cdot\frac{k_v}{q(\theta_t)}
\end{aligned}
$$

The FOC for the optimal separation threshold is:

$$
\begin{aligned}
-\frac{\partial}{\partial\epsilon_{s,t}} \;=&\; n\cdot g(\epsilon_{s,t})\cdot\left[y_t(\epsilon_{s,t},h_{stw,t}(\epsilon_{s,t}))-v(h_{stw,t}(\epsilon_{s,t}))+F+k_v\cdot\theta_t\right]\cdot\underbrace{\frac{v^f_t}{v^f_t}\cdot\tilde{u}'(c^F_t)}_{u'(\tilde{c}^w_t)}\\
&+n_t\cdot g(\epsilon_{s,t})\cdot(1-f(\theta_t))\cdot E_t\left[\lambda_{n,t+1}\right]\\
&+\lambda_{\theta,t-1}\cdot g(\epsilon_{s,t})\\
&\quad\cdot\beta\cdot\left[y_t(\epsilon_{s,t},h_{stw,t}(\epsilon_{s,t}))-v(h_{stw,t}(\epsilon_{s,t}))+F+\frac{1-f(\theta_t)\cdot\eta}{1-\eta}\cdot\frac{k_v}{q(\theta_t)}\right]\\
&+\lambda_{c,t}\cdot g(\epsilon_{s,t})\cdot\left[y_t(\epsilon_{s,t},h_{stw,t}(\epsilon_{s,t}))-v(h_{stw,t}(\epsilon_{s,t}))+F+\theta_t\cdot k_v\right]\\
&-\lambda_{c,t}\cdot\frac{\eta_{t-1}-\eta_t}{1-\eta_t}\cdot g(\epsilon_{s,t})\cdot(1-f_t)\cdot\frac{k_v}{q_t}\\
&+\lambda_{\epsilon,t}\cdot\Big[a_t\cdot h_{stw,t}(\epsilon_{s,t})^{\alpha}+c_f\\
&\quad+\frac{\partial h_{stw,t}(\epsilon_{s,t})}{\partial\epsilon_{s,t}}\cdot\underbrace{\left(\alpha\cdot a_t\cdot\epsilon_{s,t}\cdot h_{stw,t}(\epsilon_{s,t})^{\alpha-1}-v'(h_{stw,t}(\epsilon_{s,t}))-\tau_{stw,t}\right)}_{=0}\Big]\\
=&\; 0
\end{aligned}
$$

Rearranging gives the second building block for the constrained efficient separation condition:

$$\begin{aligned}\Leftrightarrow\ & y_t(\epsilon_{s,t}, h_{stw,t}(\epsilon_{s,t})) - v(h_{stw,t}(\epsilon_{s,t})) + F + k_v \cdot \theta_t + (1 - f(\theta_t)) \cdot \frac{E_t\left[\lambda_{n,t+1}\right]}{u'(\tilde{c}_t^w)} \\ & - \frac{\eta_{t-1} - \eta_t}{1 - \eta_t} \cdot (1 - \rho_t) \cdot (1 - f_t) \cdot \frac{k_v}{q_t} \\ & - \frac{\beta \cdot \lambda_{\theta,t-1} + \lambda_{c,t}}{n_t \cdot u'(\tilde{c}_t^w)} \cdot \tau_{stw,t} - \frac{\eta \cdot \lambda_{c,t}}{n_t \cdot u'(\tilde{c}_t^w)} \cdot \frac{1 - f(\theta_t)}{1 - \eta} \cdot \frac{k_v}{q(\theta_t)} \\ & + \frac{1}{g(\epsilon_{s,t})} \cdot \frac{\lambda_{\epsilon,t}}{n_t \cdot u'(\tilde{c}_t^w)} \cdot \left(a_t \cdot h_{stw,t}(\epsilon_{s,t})^\alpha + c_f\right) = 0\end{aligned}$$

Inserting the first building block of the constrained efficient separation condition gives:

$$\begin{aligned}\Leftrightarrow\ & y_t(\epsilon_{s,t}, h_{stw,t}(\epsilon_{s,t})) - v(h_{stw,t}(\epsilon_{s,t})) + F + \theta_t \cdot k_v + \left(1 - f(\theta_t)\right) \cdot \frac{1 + \chi_t}{1 - \gamma} \cdot \frac{k_v}{q(\theta_t)} \\ quad & - \frac{\lambda_{\beta \cdot \theta,t-1} + \eta \cdot \lambda_{c,t}}{n_t \cdot u'(\tilde{c}_t^w)} \cdot \tau_{stw,t} - \frac{\eta \cdot \lambda_{c,t}}{n_t \cdot u'(\tilde{c}_t^w)} \cdot \frac{1 - f(\theta_t)}{1 - \eta} \cdot \frac{k_v}{q(\theta_t)} \\ & + \frac{1}{g(\epsilon_{s,t})} \cdot \frac{\lambda_{\epsilon,t}}{n_t \cdot u'(\tilde{c}_t^w)} \cdot \left(a_t \cdot h_{stw,t}(\epsilon_{s,t})^\alpha + c_f\right) = 0\end{aligned}$$

Rearranging gives the **constrained efficient optimal separation condition:**

$$\begin{aligned}\Leftrightarrow\ & y_t(\epsilon_{s,t}, h_{stw,t}(\epsilon_{s,t})) - v(h_{stw,t}(\epsilon_{s,t})) + F \\ & + \frac{1 - \gamma \cdot f(\theta_t) + \left(1 - f(\theta_t)\right) \cdot \chi_t}{1 - \gamma} \cdot \frac{k_v}{q(\theta_t)} \\ & - \frac{(1 + \eta \cdot BE_t) \cdot \lambda_{\theta,t}}{n_t \cdot u'(\tilde{c}_t^w)} \cdot \tau_{stw,t} - \frac{BE_t \cdot \lambda_{\theta,t}}{n_t \cdot u'(\tilde{c}_t^w)} \cdot \frac{1 - f(\theta_t)}{1 - \eta} \cdot \frac{k_v}{q(\theta_t)} \\ = & -\frac{1}{g(\epsilon_{s,t})} \cdot \frac{\lambda_{\epsilon,t}}{n_t \cdot u'(\tilde{c}_t^w)} \cdot \left(a_t \cdot h_{stw,t}(\epsilon_{s,t})^\alpha + c_f\right)\end{aligned}$$

1.I.2 Constrained Efficient vs. Decentralized Conditions Steady State

Inserting the Lagrange multiplier for the consumption equivalent lets the constrained efficient job-creation condition simplify to:

$$\begin{aligned}\frac{1 + \chi}{1 - \gamma} \cdot \frac{k_v}{q(\theta)} = & \ \beta \cdot \left(\int_{\epsilon_s}^{\infty} y(\epsilon, h(\epsilon)) - v(h(\epsilon)) dG(\epsilon) - \Omega + b \right) \\ & + \beta \cdot \left(\frac{u(\tilde{c}^w) - u(b)}{u'(\tilde{c}^w)} + \frac{\beta \cdot (1 + \eta \cdot BE) \cdot \lambda_\theta}{n \cdot u'(\tilde{c}^w)} \cdot \frac{b}{n} \right) \\ & + \beta \cdot (1 - G(\epsilon_s)) \cdot \frac{(1 - \gamma \cdot f(\theta)) + (1 - f(\theta)) \cdot \chi}{1 - \gamma} \cdot \frac{k_v}{q(\theta)}\end{aligned}$$

Next, we subtract the decentralized separation condition from the constrained efficient separation condition. To do this, recognize two auxiliary calculations. First, subtract the joint surplus of firms and workers from the social value of creating a new match:

$$\frac{1+\chi}{1-\gamma}\cdot\frac{k_v}{q(\theta)}-\frac{1}{1-\eta}\cdot\frac{k_v}{q(\theta)} = \frac{(1-\eta)\cdot(1+\chi)-(1-\gamma)}{(1-\gamma)\cdot(1-\eta)}\cdot\frac{k_v}{q(\theta)}$$
$$= \left(\chi-\frac{\eta-\gamma}{1-\eta}\right)\cdot\frac{1}{1-\gamma}\cdot\frac{k_v}{q(\theta)}$$

Second, do this for the continuation value:

$$\frac{(1-\gamma\cdot f(\theta))+(1-f(\theta))\cdot\chi}{1-\gamma}\cdot\frac{k_v}{q(\theta)}-\frac{1-\eta\cdot f(\theta)}{1-\eta}\cdot\frac{k_v}{q(\theta)}$$
$$= \frac{(1-\eta)\cdot\big((1-\gamma\cdot f(\theta))+(1-f(\theta))\cdot\chi\big)-(1-\gamma)\cdot(1-\eta\cdot f(\theta))}{(1-\eta)\cdot(1-\gamma)}\cdot\frac{k_v}{q(\theta)}$$
$$= \frac{(1-\eta)\cdot\big((1-\gamma\cdot f(\theta))+(1-f(\theta))\cdot\chi\big)-(1-\gamma)\cdot(1-\eta\cdot f(\theta))}{(1-\eta)\cdot(1-\gamma)}\cdot\frac{k_v}{q(\theta)}$$
$$= (1-f(\theta))\cdot\left(\chi-\frac{\eta-\gamma}{1-\eta}\right)\cdot\frac{1}{1-\gamma}\cdot\frac{k_v}{q(\theta)}$$

Putting both together gives the deviation of the constrained efficient separation condition from the decentralized:

$$\left(\chi-\frac{\eta-\gamma}{1-\eta}\right)\cdot\frac{1}{1-\gamma}\cdot\frac{k_v}{q(\theta)} = \beta\cdot\left(1+\frac{\beta\cdot(1+\eta\cdot BE)\cdot\lambda_\theta}{n\cdot u'(\tilde{c}^w)}\right)\cdot\frac{b}{n}$$
$$+\beta\cdot(1-G(\epsilon_s))\cdot(1-f(\theta))\cdot\left(\chi-\frac{\eta-\gamma}{1-\eta}\right)\cdot\frac{1}{1-\gamma}\cdot\frac{k_v}{q(\theta)}$$

Rearranging the **deviation from optimal job-creation condition** can be expressed as:

$$\left(\chi-\frac{\eta-\gamma}{1-\eta}\right)\cdot\frac{1}{1-\gamma}\cdot\frac{k_v}{q(\theta)}$$
$$= \frac{\beta}{1-\beta\cdot(1-f(\theta))\cdot(1-G(\epsilon_s))}\cdot\left(1+\frac{\beta\cdot(1+\eta\cdot BE)\cdot\lambda_\theta}{n\cdot u'(\tilde{c}^w)}\right)\cdot\frac{b}{n}$$

We can also interpret this as how many firms and workers undervalue the social value of the match.

The constrained optimal separation condition in steady state can be expressed as:

$$\Leftrightarrow \; y(\epsilon_s, h_{stw}(\epsilon_s)) - v(h_{stw}(\epsilon_s)) + F + \big(1 - \gamma \cdot f(\theta)\big) \cdot \frac{1+\chi}{1-\gamma} \cdot \frac{k_v}{q(\theta)}$$
$$- \frac{(1+\eta \cdot BE) \cdot \lambda_\theta}{n \cdot u'(\tilde{c}^w)} \cdot \tau_{stw} - \frac{BE \cdot \lambda_\theta}{n \cdot u'(\tilde{c}^w)} \cdot \frac{1-f(\theta)}{1-\eta} \cdot \frac{k_v}{q(\theta)}$$
$$+ \frac{1}{g(\epsilon_s)} \cdot \frac{\lambda_\epsilon}{n \cdot u'(\tilde{c}^w)} \cdot \big(a \cdot h(\epsilon_s)^\alpha + c_f\big) \; = \; 0$$

Next, subtract the decentralized separation condition:

$$- \tau_{stw,t} \cdot \big(\bar{h} - h_{stw,t}(\epsilon_{s,t})\big) + (1 - f(\theta)) \cdot \left(\chi - \frac{\eta - \gamma}{1-\eta}\right) \cdot \frac{1}{1-\gamma} \cdot \frac{k_v}{q(\theta)}$$
$$- \frac{(1+\eta \cdot BE) \cdot \lambda_{\theta,t}}{n \cdot u'(\tilde{c}^w)} \cdot \tau_{stw,t} \cdot \big(\bar{h} - h_{stw,t}(\epsilon_{s,t})\big) - \frac{BE \cdot \lambda_\theta}{n \cdot u'(\tilde{c}^w)} \cdot \frac{1-f(\theta)}{1-\eta} \cdot \frac{k_v}{q(\theta)}$$
$$+ \frac{1}{g(\epsilon_s)} \cdot \frac{\lambda_\epsilon}{n \cdot u'(\tilde{c}^w)} \cdot \big(a \cdot h(\epsilon_s)^\alpha + c_f\big) \; = \; 0$$

Rearrange for the **optimal STW subsidy.**

$$\Leftrightarrow \quad \tau_{stw,t} \cdot \big(\bar{h} - h_{stw,t}(\epsilon_s)\big) \; = \; (1 - f(\theta)) \cdot \left(\chi - \frac{\eta - \gamma}{1-\eta}\right) \cdot \frac{1}{1-\gamma} \cdot \frac{k_v}{q(\theta)}$$
$$- \frac{(1+\eta \cdot BE) \cdot \lambda_\theta}{n \cdot u'(\tilde{c}^w)} \cdot \tau_{stw} \cdot \big(\bar{h} - h_{stw,t}(\epsilon_s)\big)$$
$$+ \frac{1}{g(\epsilon_s)} \cdot \frac{\lambda_\epsilon}{n \cdot u'(\tilde{c}^w)} \cdot \big(a \cdot h(\epsilon_s)^\alpha + c_f\big)$$
$$- \frac{BE \cdot \lambda_\theta}{n \cdot u'(\tilde{c}^w)} \cdot \frac{1-f(\theta)}{1-\eta} \cdot \frac{k_v}{q(\theta)}$$

The subsidy is there to realign the optimal and decentralized separation condition. We can see this when we look at the **deviation from optimal separation condition**:

$$\left(1 + \frac{\beta \cdot (1+\eta \cdot BE) \cdot \lambda_\theta}{n \cdot u'(\tilde{c}^w)}\right)$$
$$\cdot \left(\frac{\beta \cdot (1 - f(\theta))}{1 - \beta \cdot (1 - f(\theta)) \cdot (1 - G(\epsilon_s))} \cdot \frac{b}{n} - \big(\bar{h} - h_{stw}(\epsilon_s)\big) \cdot \tau_{stw}\right)$$
$$+ \frac{1}{g(\epsilon_s)} \cdot \frac{\lambda_\epsilon}{n \cdot u'(\tilde{c}^w)} \cdot \big(a \cdot h(\epsilon_s)^\alpha + c_f\big) - \frac{BE \cdot \lambda_\theta}{n \cdot u'(\tilde{c}^w)} \cdot \frac{1-f(\theta)}{1-\eta} \cdot \frac{k_v}{q(\theta)} \; = \; 0$$

1.I.3 Optimal UI given STW Steady State

To calculate the optimal UI system given a non-optimized STW system in steady state, assume that the eligibility threshold is looser than the separation threshold without access to STW: $\epsilon_{stw,t} > \xi_{s,t}$. First, we have to form the FOC for the optimal UI benefits:

$$\begin{aligned}\frac{\partial}{\partial b_t} &= (1-n_t)\cdot\left(u'(b_t)-u'(\tilde{c}_t^w)\right)\\ &\quad -\lambda_{\theta,t}\cdot\beta\cdot\left[\left(\frac{u'(b_t)}{u'(\tilde{c}_t^w)}+\frac{n_t}{1-n_t}\right)-BE_t\cdot\left(\left(1-\eta\right)\cdot\frac{u'(b_t)}{u'(\tilde{c}_t^w)}-\eta\cdot\frac{n_t}{1-n_t}\right)\right]\end{aligned}$$

Note that this depends on the Lagrange multiplier of the job-creation condition/ private value of the match. To calculate it, we first have to calculate the FOC to the separation threshold of firms without access to STW. Note that its Lagrange multiplier must be zero:

$$\begin{aligned}\frac{\partial}{\partial \xi_{s,t}} &=\\ -\lambda_{\xi,t}\cdot&\left(a_t\cdot h_t(\xi_{s,t})^\alpha+c_f+\frac{\partial h_t(\xi_{s,t})}{\partial \xi_{s,t}}\cdot\underbrace{\left(\alpha\cdot a_t\cdot h_t(\xi_{s,t})^{\alpha-1}-v'(h_t(\xi_{s,t}))\right)}_{=0}\right)\cdot u'(\tilde{c}^w) = 0\\ \Leftrightarrow\ \lambda_{\xi,t} &= 0\end{aligned}$$

Next, note that the deviation of the private value of the match from the social value of the match can be denoted by:

$$\begin{aligned}&\left(\chi-\frac{\eta-\gamma}{1-\eta}\right)\cdot\frac{1}{1-\gamma}\cdot\frac{k_v}{q(\theta)} =\\ &\qquad\frac{\beta}{1-\beta\cdot(1-f(\theta))\cdot(1-G(\epsilon_s))}\cdot\left(1+\frac{\beta\cdot(1+\eta\cdot BE)\cdot\lambda_\theta}{n\cdot u'(\tilde{c}^w)}\right)\cdot\frac{b}{n}\end{aligned}$$

Rearranging for χ gives the right-hand side of the formula for calculating the Lagrange multiplier for the job-creation condition/ private value of a match:

$$\begin{aligned}\chi &= \frac{\eta-\gamma}{1-\eta}\\ &\quad+\frac{\beta}{1-\beta\cdot(1-f(\theta))\cdot(1-G(\epsilon_s))}\\ &\quad\cdot\left(1+\frac{\beta\cdot(1+\eta\cdot BE)\cdot\lambda_\theta}{n\cdot u'(\tilde{c}^w)}\right)\cdot\frac{b}{n}/\left(\frac{1}{1-\gamma}\cdot\frac{k_v}{q(\theta)}\right)\end{aligned}$$

The left-hand side of the equation can be expressed by the steady state value of χ.

$$\chi = \frac{1}{1-\eta} \cdot \frac{\gamma - \beta \cdot \left(\gamma - \left(1-(1-\eta)\cdot BE\right) \cdot \eta \cdot f(\theta)\right) \cdot (1-G(\epsilon_s))}{m(\theta)\cdot u'(\tilde{c}^w)} \cdot \lambda_\theta \\ - \frac{1}{1-\eta} \cdot \frac{(\gamma - \eta \cdot f(\theta)) \cdot \left(-\lambda_\epsilon - \lambda_\xi\right)}{m(\theta)\cdot u'(\tilde{c}^w)}$$

Note that the Lagrange multiplier for the separation condition without access to STW benefits is zero, which simplifies the equation to:

$$\chi = \frac{1}{1-\eta} \cdot \frac{\gamma - \beta \cdot \left(\gamma - \left(1-(1-\eta)\cdot BE\right) \cdot \eta \cdot f(\theta)\right) \cdot (1-G(\epsilon_s))}{m(\theta)\cdot u'(\tilde{c}^w)} \cdot \lambda_\theta \\ - \frac{1}{1-\eta} \cdot \frac{(\gamma - \eta \cdot f(\theta)) \cdot \left(-\lambda_\epsilon\right)}{m(\theta)\cdot u'(\tilde{c}^w)}$$

Now bring the left-hand side and the right-hand side of the equation together:

$$\frac{1}{1-\eta} \cdot \frac{\gamma - \beta \cdot \left(\gamma - \left(1-(1-\eta)\cdot BE\right) \cdot \eta \cdot f(\theta)\right) \cdot (1-G(\epsilon_s))}{m(\theta)\cdot u'(\tilde{c}^w)} \cdot \lambda_\theta = \\ + \frac{\eta - \gamma}{1-\eta} \\ + \frac{\beta}{1-\beta\cdot(1-f(\theta))\cdot(1-G(\epsilon_s))} \\ \cdot \left(1 + \frac{\beta\cdot(1+\eta\cdot BE)\cdot\lambda_\theta}{n\cdot u'(\tilde{c}^w)}\right) \cdot \frac{b}{n} / \left(\frac{1}{1-\gamma} \cdot \frac{k_v}{q(\theta)}\right) \\ + \frac{1}{1-\eta} \cdot \frac{(\gamma - \eta \cdot f(\theta)) \cdot \left(-\lambda_\epsilon\right)}{m(\theta)\cdot u'(\tilde{c}^w)}$$

Rearranging the equation gives:

$$\frac{1}{1-\eta} \cdot \frac{\gamma - \beta \cdot \left(\gamma - \left(1-(1-\eta)\cdot BE\right) \cdot \eta \cdot f(\theta)\right) \cdot (1-G(\epsilon_s))}{f(\theta)\cdot u'(\tilde{c}^w)} \cdot k_v \cdot \lambda_\theta = \\ + (1-\gamma)\cdot u \cdot q(\theta) \cdot \frac{\eta - \gamma}{(1-\gamma)\cdot(1-\eta)} \cdot \frac{k_v}{q(\theta)} \\ + (1-\gamma)\cdot u \cdot q(\theta) \cdot \frac{\beta}{1-\beta\cdot(1-f(\theta))\cdot(1-G(\epsilon_s))} \\ \cdot \left(1 + \frac{\beta\cdot(1+\eta\cdot BE)\cdot\lambda_\theta}{n\cdot u'(\tilde{c}^w)}\right) \cdot \frac{b}{n} \\ + \frac{1}{1-\eta} \cdot \frac{(\gamma - \eta \cdot f(\theta)) \cdot k_v}{f(\theta)} \cdot \frac{-\lambda_\epsilon}{u'(\tilde{c}^w)}$$

We still have to calculate the Lagrange multiplier for the separation condition λ_ϵ:

$$
\begin{aligned}
-\lambda_\epsilon \;=\; & \left(1+\frac{\beta\cdot(1+\eta\cdot BE)}{n\cdot u'(\tilde{c}^w)}\cdot\lambda_\theta\right)\cdot\Big(\frac{\beta\cdot(1-f(\theta))}{1-\beta\cdot(1-f(\theta))\cdot(1-G(\epsilon_s))}\cdot\frac{b}{n} \\
& -\tau_{stw}\cdot(\bar{h}-h_{stw}(\epsilon_s))\Big)\cdot\frac{g(\epsilon_s)\cdot n\cdot u'(\tilde{c}^w)}{a\cdot h_{stw}(\epsilon_s)^\alpha+c_f} \\
& -BE\cdot\lambda_\theta\cdot\frac{1-f(\theta)}{1-\eta}\cdot\frac{k_v}{q(\theta)}\cdot\frac{g(\epsilon_s)}{a\cdot h_{stw}(\epsilon_s)^\alpha+c_f}
\end{aligned}
$$

Inserting the Lagrange multiplier for separations into the formula above gives:

$$
\begin{aligned}
& \frac{1}{1-\eta}\cdot\frac{\gamma-\beta\cdot\Big(\gamma-\big(1-(1-\eta)\cdot BE\big)\cdot\eta\cdot f(\theta)\Big)\cdot(1-G(\epsilon_s))}{f(\theta)\cdot u'(\tilde{c}^w)}\cdot k_v\cdot\lambda_\theta \\
&= f'(\theta)\cdot u\cdot\Big(\frac{\eta-\gamma}{(1-\gamma)\cdot(1-\eta)}\cdot\frac{k_v}{q(\theta)}+\frac{\beta}{1-\beta\cdot(1-f(\theta))\cdot(1-G(\epsilon_s))} \\
&\quad\cdot\left(1+\frac{\beta\cdot(1+\eta\cdot BE)\cdot\lambda_\theta}{n\cdot u'(\tilde{c}^w)}\right)\cdot\frac{b}{n}\Big) \\
&\quad+\cdot\frac{\partial\epsilon_s}{\partial\theta}\cdot g(\epsilon_s)\cdot n\cdot\left(1+\frac{\beta\cdot(1+\eta\cdot BE)}{n\cdot u'(\tilde{c}^w)}\cdot\lambda_\theta\right) \\
&\quad\cdot\Big(\frac{\beta\cdot(1-f(\theta))}{1-\beta\cdot(1-f(\theta))\cdot(1-G(\epsilon_s))}\cdot\frac{b}{n}-\tau_{stw}\cdot\big(\bar{h}-h_{stw}(\epsilon_s)\big)\Big) \\
&\quad-BE\cdot\lambda_\theta\cdot\frac{1-f(\theta)}{1-\eta}\cdot\frac{k_v}{q(\theta)}\cdot\frac{g(\epsilon_s)}{a\cdot h_{stw}(\epsilon_s)^\alpha+c_f}\cdot\frac{1}{1-\eta}\cdot\frac{\gamma-\eta\cdot f(\theta)}{f(\theta)}\cdot k_v
\end{aligned}
$$

with

$$
\begin{aligned}
f'(\theta) &= (1-\gamma)\cdot q(\theta) \\
\frac{\partial\epsilon_s}{\partial\theta} &= \frac{1}{1-\eta}\cdot\frac{(\gamma-\eta\cdot f(\theta))\cdot k_v}{f(\theta)}\cdot\frac{1}{a\cdot h_{stw}(\epsilon_s)^\alpha+c_f}
\end{aligned}
$$

Next, solve for the Lagrange multiplier of the social value of the match:

$$
\begin{aligned}
\lambda_\theta \;=\; & \frac{u'(\tilde{c}^w)}{M}\cdot f'(\theta)\cdot u \\
& \cdot\left[\frac{\eta-\gamma}{(1-\eta)(1-\gamma)}\cdot\frac{k_v}{q(\theta)}+\frac{\beta}{1-\beta\cdot(1-f(\theta))\cdot(1-G(\epsilon_s))}\cdot\frac{b}{n}\right] \\
& +\frac{u'(\tilde{c}^w)}{M}\cdot\frac{\partial\epsilon_s}{\partial\theta}\cdot g(\epsilon_s)\cdot n \\
& \cdot\left[\frac{\beta\cdot(1-f(\theta))}{1-\beta\cdot(1-f(\theta))\cdot(1-G(\epsilon_s))}\cdot\frac{b}{n}-\tau_{stw}\cdot(\bar{h}-h_{stw}(\epsilon_s))\right]
\end{aligned}
$$

I call M the inverse multiplier of the match:

$$M = \underbrace{\frac{\gamma - \beta \cdot \big(\gamma - \big(1 - (1-\eta) \cdot BE\big) \cdot \eta \cdot f(\theta)\big) \cdot (1 - G(\epsilon_s)) + \big(\gamma - \eta \cdot f(\theta)\big) \cdot BE \cdot \frac{1-f(\theta)}{1-\eta} \cdot \frac{k_v}{q(\theta)} \cdot \frac{g(\epsilon_s)}{a \cdot h_{stw}(\epsilon_s)^{\alpha} + c_f}}{(1-\eta) \cdot f(\theta)}}_{\text{regular multiplier effect}}$$
$$- \underbrace{\frac{\beta \cdot (1 - G(\epsilon_s))}{1 - \beta \cdot (1 - f(\theta)) \cdot (1 - G(\epsilon_s))} \cdot \frac{\beta \cdot (1 + \eta \cdot BE)}{n} \cdot \frac{b}{n} \cdot f'(\theta) \cdot u}_{\text{tax effect via vacancy posting}}$$
$$- \underbrace{\frac{\beta \cdot (1 - G(\epsilon_s))}{1 - \beta \cdot (1 - f(\theta)) \cdot (1 - G(\epsilon_s))} \cdot \frac{\beta \cdot (1 + \eta \cdot BE)}{n} \cdot \left(\frac{b}{n} - \tau_{stw} \cdot (\bar{h} - h_{stw}(\epsilon_s))\right) \cdot \frac{\partial \epsilon_s}{\partial \theta} \cdot g(\epsilon_s) \cdot n}_{\text{tax effect via separations}}$$

M captures the feedback effects that altering the value of the firm in one period has on the labor market tightness. The feedback effects can be divided into three channels, which can be illustrated by considering an increase in UI benefits. The first channel is the regular channel, where an increase in UI benefits depresses the joint surplus of firms and workers, not only in the present but also in the future. The second channel emerges through the free entry condition. As the joint surplus declines, firms reduce vacancy postings. This reduction leads to higher unemployment, which, in turn, increases the fiscal burden of the UI system. To cover these additional costs, higher income taxes are required, which further depresses the joint surplus, thereby amplifying the initial impact.

Finally, the third channel arises through the separation condition. A decrease in the joint surplus increases the separation rate, as the continuation value of the match between firms and workers diminishes. More separations result in higher unemployment, which once again raises the costs of the UI system, leading to the need for higher taxes. This creates a reinforcing loop that further reduces the joint surplus of firms and workers.

It's important to note that the last effect is mitigated by the STW system. Larger STW benefits reduce separation incentives, thereby lowering the number of workers entering the UI system. As a result, the increase in unemployment is smaller, which in turn lessens the distortionary tax effect caused by the UI system. STW itself has a neutral impact on the joint surplus of firms and workers. On one hand, an increase in STW benefits raises the joint surplus through the subsidy effect. On the other hand, this increase is offset by the need to raise taxes to finance the system, which decreases the joint surplus. Ultimately, these two effects cancel each other out.

In the following, I require the inverse multiplier to be positive, $M > 0$. Otherwise, the model would not converge to its steady state. We can abbreviate the inverse multiplier by recognizing that it captures the general equilibrium effect that a change in the surplus or value of the firm has on the labor market tightness:

$$\frac{\partial \theta^{ge}}{\partial S} = \frac{1}{M} \quad \text{or} \quad \frac{\partial \theta^{ge}}{\partial J} = \frac{1}{1-\eta} \cdot \frac{1}{M}$$

We can rewrite the formula for the Lagrange multiplier as:

$$\begin{aligned}
\lambda_\theta \;=\; & \frac{u'(\tilde{c}^w)}{1-\eta} \cdot \frac{\partial \theta^{ge}}{\partial J} \cdot f'(\theta) \cdot u \\
& \cdot \left[\frac{\eta - \gamma}{\big(1-\eta\big) \cdot (1-\gamma)} \cdot \frac{k_v}{q(\theta)} + \frac{\beta}{1 - \beta \cdot (1 - f(\theta)) \cdot (1 - G(\epsilon_s))} \cdot \frac{b}{n} \right] \\
& + \frac{u'(\tilde{c}^w)}{1-\eta} \cdot \frac{\partial \theta^{ge}}{\partial J} \cdot \frac{\partial \epsilon_s}{\partial \theta} \cdot g(\epsilon_s) \cdot n \\
& \cdot \left[\frac{\beta \cdot (1 - f(\theta))}{1 - \beta \cdot (1 - f(\theta)) \cdot (1 - G(\epsilon_s))} \cdot \frac{b}{n} - \tau_{stw} \cdot (\bar{h} - h_{stw}(\epsilon_s)) \right]
\end{aligned}$$

So the optimal UI benefits can be characterized by:

$$\begin{aligned}
(1-n)\cdot\big(u'(b) - u'(\tilde{c}^w)\big) \;=\; & \\
& \lambda_\theta \cdot \beta \cdot \left[\left(\frac{u'(b)}{u'(\tilde{c}^w)} + \frac{n}{1-n} \right) - BE \cdot \left(\big(1-\eta\big) \cdot \frac{u'(b)}{u'(\tilde{c}^w)} - \eta \cdot \frac{n}{1-n} \right) \right]
\end{aligned}$$

Let L_V and L_S denote the social value of an additional hire, respectively, from retrieving the marginal match:

$$\begin{aligned}
L_V \;=\; & \frac{\eta - \gamma}{\big(1-\eta\big) \cdot (1-\gamma)} \cdot \frac{k_v}{q(\theta)} + \frac{\beta}{1 - \beta \cdot (1 - f(\theta)) \cdot (1 - G(\epsilon_s))} \cdot \frac{b}{n} \\
L_S \;=\; & \frac{\beta \cdot (1 - f(\theta))}{1 - \beta \cdot (1 - f(\theta)) \cdot (1 - G(\epsilon_s))} \cdot \frac{b}{n} - \tau_{stw} \cdot (\bar{h} - h_{stw}(\epsilon_s))
\end{aligned}$$

Further, note that most of the right-hand side of the optimal UI benefits denote the first derivative of the value of the firm to an increase in UI benefits:

$$-\frac{\partial J}{\partial b} \;=\; \beta \cdot (1-\eta) \cdot \left[\left(\frac{u'(b)}{u'(\tilde{c}^w)} + \frac{n}{1-n} \right) - BE \cdot \left(\big(1-\eta\big) \cdot \frac{u'(b)}{u'(\tilde{c}^w)} - \eta \cdot \frac{n}{1-n} \right) \right]$$

Then, we can express the equation for the optimal UI benefits as:

$$\begin{aligned}
(1-n)\cdot\frac{u'(b) - u'(\tilde{c}^w)}{u'(\tilde{c}^w)} \;=\; & \\
& L_V \cdot \left(-\frac{\partial J}{\partial b} \right) \cdot \frac{\partial \theta^{ge}}{\partial J} \cdot f'(\theta) \cdot u + L_S \cdot \frac{\partial J}{\partial b} \cdot \frac{\partial \theta^{ge}}{\partial J} \cdot \left(-\frac{\partial \epsilon_s}{\partial \theta} \right) \cdot g(\epsilon_s) \cdot n
\end{aligned}$$

We can also express this as:

$$(1-n) \cdot \frac{u'(b) - u'(\tilde{c}^w)}{u'(\tilde{c}^w)} \;=\; L_V \cdot \left(-\frac{\partial f^{ge}}{\partial b} \right) \cdot u + L_S \cdot \frac{\partial \epsilon_s^{ge}}{\partial b} \cdot g(\epsilon_s) \cdot n$$

1.I.4 Optimal STW benefits Steady State

First, derive the FOC of working time on STW:

$$\frac{\partial}{\partial h_{stw,t}(\epsilon)} = -\left(n_t \cdot \underbrace{\frac{\nu_t^f}{\nu_t^f} \cdot \tilde{u}(c_t^F)}_{\tilde{u}'(\tilde{c}_t^w)} + \lambda_\theta + \eta \cdot \lambda_{c,t} \right) \cdot \frac{\partial \Omega_t}{\partial h_{stw,t(\epsilon)}}$$
$$- \lambda_{h_{stw},t}(\epsilon) \cdot \left(\alpha \cdot (\alpha - 1) \cdot a_t \cdot \epsilon \cdot h_{stw,t}(\epsilon)^{\alpha-2} - \nu''(h_{stw,t}(\epsilon))\right) = 0$$

We can use this to calculate the Lagrange multiplier for the working time on STW. Note that the multiplier depends on the idiosyncratic productivity ϵ:

$$\lambda_{h_{stw},t}(\epsilon) = -n_t \cdot u'(\tilde{c}_t^w) \cdot \frac{\left(1 + \frac{(1+\eta \cdot BE)\cdot \lambda_{\theta,t-1}}{n_t \cdot u'(\tilde{c}_t^w)}\right)}{(\alpha - 1) \cdot \alpha \cdot a_t \cdot \epsilon \cdot h_{stw}(\epsilon)^{\alpha-2} - \nu''(h_{stw,t}(\epsilon))} \cdot \frac{\partial \Omega_t}{\partial h_{stw,t}(\epsilon)}$$

The FOC for STW benefits is:

$$\frac{\partial}{\partial \tau_{stw,t}} = -\int_{\epsilon_{s,t}}^{\epsilon_{stw,t}} \lambda_{h_{stw},t}(\epsilon) dG(\epsilon)$$
$$- \lambda_{\epsilon,t} \cdot \Bigg[\left(\bar{h} - h_{stw,t}(\epsilon_{s,t})\right)$$
$$+ \frac{\partial h_{stw,t}(\epsilon_{s,t})}{\partial \tau_{stw,t}} \cdot \underbrace{\left(\alpha \cdot a_t \cdot \epsilon_{s,t} \cdot h_{stw,t}(\epsilon_{s,t})^{\alpha-1} - \nu'(h_{stw,t}(\epsilon_{s,t})) - \tau_{stw,t}\right)}_{=0} \Bigg]$$
$$= 0$$

Next, we can insert the Lagrange multiplier from the optimal hours choice on STW:

$$\Leftrightarrow -n_t \cdot u'(\tilde{c}_t^w) \cdot \left(1 + \frac{(1 + \eta \cdot BE) \cdot \lambda_{\theta,t-1}}{n_t \cdot u'(\tilde{c}_t^w)}\right) \cdot \int_{\epsilon_{s,t}}^{\epsilon_{stw,t}} \frac{\partial \Omega_t}{\partial h_{stw,t}(\epsilon)} \cdot \frac{h_{stw,t}(\epsilon)}{\tau_{stw,t}} dG(\epsilon)$$
$$- \lambda_{\epsilon,t} \cdot \left(\bar{h} - h_{stw,t}(\epsilon_{s,t})\right) = 0$$

Here, the FOC of working hours for STW benefits can be denoted as:

$$\frac{h_{stw,t}(\epsilon)}{\tau_{stw,t}} = \frac{1}{(\alpha - 1) \cdot \alpha \cdot a_t \cdot \epsilon \cdot h_{stw}(\epsilon)^{\alpha-2} - \nu''(h_{stw,t}(\epsilon))} < 0$$

Rearranging gives us the Lagrange multiplier for the separation condition.

$$\Leftrightarrow \frac{\lambda_{\epsilon,t}}{n_t \cdot u'(\tilde{c}_t^w)} = -\frac{1 + \frac{\beta \cdot (1+\eta \cdot BE_t)\cdot \lambda_{\theta,t-1}}{n_t \cdot u'(\tilde{c}_t^w)}}{\bar{h} - h_{stw,t}(\epsilon_{s,t})} \cdot \frac{\partial \Omega_t}{\partial \tau_{stw,t}}$$

The planner decides against preventing all inefficient separations to save on costs of hours distortion.

Next, insert the Lagrange multiplier into the constrained efficient separation condition and rearrange for the optimal STW subsidy:

$$\begin{aligned}
\Leftrightarrow \quad & \tau_{stw} \cdot \left(\bar{h} - h_{stw}(\epsilon_s)\right) \\
&= \frac{\beta \cdot (1 - f(\theta))}{1 - \beta \cdot (1 - f(\theta)) \cdot (1 - G(\epsilon_s))} \cdot \left(1 + \frac{\beta \cdot (1 + \eta \cdot BE) \cdot \lambda_\theta}{n \cdot u'(\tilde{c}^w)}\right) \cdot \frac{b}{n} \\
&- \frac{(1 + \eta \cdot BE) \cdot \lambda_{\theta,t}}{n \cdot u'(\tilde{c}^w)} \cdot \tau_{stw,t} \cdot \left(\bar{h} - h_{stw,t}(\epsilon_{s,t})\right) \\
&- \frac{1}{g(\epsilon_s)} \cdot \left(1 + \frac{\beta \cdot (1 + \eta \cdot BE) \cdot \lambda_\theta}{n \cdot u'(\tilde{c}^w)}\right) \cdot \frac{\partial \Omega}{\partial \tau_{stw}} \cdot \frac{a \cdot h_{stw}(\epsilon_s)^\alpha + c_f}{\bar{h} - h_{stw}(\epsilon_s)} \\
&- \frac{BE \cdot \lambda_\theta}{n \cdot u'(\tilde{c}^w)} \cdot \frac{1 - f(\theta)}{1 - \eta} \cdot \frac{k_v}{q(\theta)}
\end{aligned}$$

Note that STW benefits reduce separation incentives by keeping unproductive workers employed. At the same time, they reduce the costs of the UI system. We can see this in the following equation by the term $\left(1 + \frac{\beta \cdot (1+\eta \cdot BE) \cdot \lambda_\theta}{n \cdot u'(\tilde{c}^w)}\right)$.

$$\begin{aligned}
\Leftrightarrow \quad & \left(1 + \frac{\beta \cdot (1 + \eta \cdot BE) \cdot \lambda_\theta}{n \cdot u'(\tilde{c}^w)}\right) \cdot \tau_{stw,t} \cdot \left(\bar{h} - h_{stw,t}(\epsilon_{s,t})\right) \\
&= \left(1 + \frac{\beta \cdot (1 + \eta \cdot BE) \cdot \lambda_\theta}{n \cdot u'(\tilde{c}^w)}\right) \cdot \left(\frac{\beta \cdot (1 - f(\theta))}{1 - \beta \cdot (1 - f(\theta)) \cdot (1 - G(\epsilon_s))} \cdot \frac{b}{n} \right. \\
&\qquad \left. - \frac{1}{g(\epsilon_s)} \cdot \frac{\partial \Omega}{\partial \tau_{stw}} \cdot \frac{a \cdot h_{stw}(\epsilon_s)^\alpha + c_f}{\bar{h} - h_{stw}(\epsilon_s)}\right) \\
&- \frac{BE \cdot \lambda_\theta}{n \cdot u'(\tilde{c}^w)} \cdot \frac{1 - f(\theta)}{1 - \eta} \cdot \frac{k_v}{q(\theta)}
\end{aligned}$$

Interestingly, both effects are equally strong, and the term crosses out, leaving us only with the bargaining effect:

$$\begin{aligned}
\Leftrightarrow \quad \tau_{stw,t} \cdot \left(\bar{h} - h_{stw,t}(\epsilon_{s,t})\right) &= \frac{\beta \cdot (1 - f(\theta))}{1 - \beta \cdot (1 - f(\theta)) \cdot (1 - G(\epsilon_s))} \cdot \frac{b}{n} \\
&- \frac{1}{g(\epsilon_s)} \cdot \frac{\partial \Omega}{\partial \tau_{stw}} \cdot \frac{a \cdot h_{stw}(\epsilon_s)^\alpha + c_f}{\bar{h} - h_{stw}(\epsilon_s)} \\
&- \frac{\frac{BE \cdot \lambda_\theta}{n \cdot u'(\tilde{c}^w)}}{1 + \frac{\beta \cdot (1+\eta \cdot BE) \cdot \lambda_\theta}{n \cdot u'(\tilde{c}^w)}} \cdot \frac{1 - f(\theta)}{1 - \eta} \cdot \frac{k_v}{q(\theta)}
\end{aligned}$$

Note that

$$\frac{a_t \cdot h_{stw,t}(\epsilon_{s,t})^{\alpha} + c_f}{\bar{h} - h_{stw,t}(\epsilon)} = \frac{\frac{\partial S_t(\epsilon_{stw,t})}{\partial \epsilon_{s,t}}}{\frac{\partial S_{stw,t}(\epsilon_{s,t})}{\tau_{stw,t}}} = -\frac{1}{-\frac{\frac{\partial S_t(\epsilon_{stw,t})}{\partial \epsilon_{s,t}}}{\frac{\partial S_{stw,t}(\epsilon_{s,t})}{\tau_{stw,t}}}} = -\frac{1}{\frac{\partial \epsilon_{s,t}}{\partial \tau_{stw,t}}} > 0$$

and define the bargaining effect as:

$$\tilde{BE} = \frac{\frac{BE\cdot\lambda_\theta}{n\cdot u'(\bar{c}^w)}}{1 + \frac{\beta\cdot(1+\eta\cdot BE)\cdot\lambda_\theta}{n\cdot u'(\bar{c}^w)}} \cdot \frac{1-f(\theta)}{1-\eta} \cdot \frac{k_v}{q(\theta)}$$

Then we get the formula for the optimal STW benefits as in Proposition 2:

$$\begin{aligned} \tau_{stw} \cdot \left(\bar{h} - h_{stw}(\epsilon_{s,t})\right) = \; & \frac{\beta \cdot (1-f(\theta))}{1-\beta \cdot (1-f(\theta)) \cdot (1-G(\epsilon_s))} \cdot \frac{b}{n} \\ & - \frac{1}{g(\epsilon_s)} \cdot \frac{\partial \Omega}{\partial \tau_{stw}} \cdot \left[-\frac{1}{\frac{\partial \epsilon_s}{\partial \tau_{stw}}} \right] \\ & - \tilde{BE} \end{aligned}$$

1.I.5 Optimal Eligibility Condition Steady State

This section derives the optimal eligibility condition. It aims to show, using proof by contradiction, that $\epsilon_{stw,t} = \xi_{s,t}$ as long as the sufficient condition of Proposition 3 holds:

$$\underbrace{n \cdot \frac{\partial \Omega}{\partial \epsilon_{stw}}}_{\text{More hours distortion}} + L_V \cdot \underbrace{\left(-\frac{\partial f^{ge}}{\partial \epsilon_{stw}} \cdot u\right)}_{\text{Less hiring}} + L_S^{\star} \cdot \underbrace{\left(n \cdot g(\epsilon_s) \cdot \frac{\partial \epsilon_s^{ge}}{\partial \epsilon_{stw}}\right)}_{\text{More Separations}} \geq 0$$

The section will examine each possible value of $\epsilon_{stw,t}$, except for $\epsilon_{stw,t} = \xi_{s,t}$ itself, to demonstrate that they cannot constitute an optimum.

case 0: $\boldsymbol{\epsilon_{stw,t} < \epsilon_{s,t}}$.

Case 0 examines the case where the eligibility threshold is stricter than the separation threshold of a firm with access to STW. As a result, STW does not exist in this economy. By excluding any $\epsilon_{stw,t} < \epsilon_{s,t}$ from the optimization problem of the Ramsey planner, we do not restrict the choice set of the Ramsey planner. The Ramsey planner can always choose $\epsilon_{stw,t} = \epsilon_{s,t}$ or $\tau_{stw,t} = 0$ to eliminate the STW system from the economy.

case 1: $\epsilon_{s,t} \leq \epsilon_{stw,t} < \xi_{s,t}$;

Case 1 describes the situation where the eligibility condition is so strict that matches with $\epsilon_p \in (\epsilon_{stw,t}, \xi_{s,t})$ are not allowed to go onto the STW system and subsequently dissolve. At the same time, matches with $\epsilon_u \in [\epsilon_{s,t}, \epsilon_{stw,t}]$ are allowed to enter the STW system and are rescued. The paragraph wants to show that it cannot be optimal to rescue less productive matches while letting more productive matches dissolve $\epsilon_u < \epsilon_p$. To verify the claim, it has to be shown that the Ramsey planner would always want to loosen the eligibility threshold as long as we are in case 1. The Ramsey planner always wants to loosen the eligibility condition if the following condition is met:

$$\frac{\partial}{\partial \epsilon_{stw,t}} \overset{!}{>} 0$$

Showing this would mean that no $\epsilon_{stw,t} \in [\epsilon_{s,t}, \xi_{s,t})$ is optimal. So we start by calculating the first-order derivative of the Lagrange problem:

$$\begin{aligned}
\frac{\partial}{\partial \epsilon_{stw,t}} = {} & n_t \cdot \underbrace{\frac{\nu_t^f}{\nu_t^f} \cdot \tilde{u}'(c_t^F)}_{u'(\tilde{c}_t^w)} \cdot g(\epsilon_{stw,t}) \cdot \Big(y_t(\epsilon, h_{stw,t}(\epsilon_{stw,t})) - v(h_{stw,t}(\epsilon_{stw,t})) + F + \theta_t \cdot k_v\Big) \\
& + n_t \cdot g(\epsilon_{stw,t}) \cdot (1 - f(\theta_t)) \cdot E_t\left[\lambda_{n,t+1}\right] \\
& + \lambda_{\theta,t-1} \cdot g(\epsilon_{stw,t}) \cdot \beta \cdot \Big(y_t(\epsilon, h_{stw,t}(\epsilon_{stw,t})) - v(h_{stw,t}(\epsilon_{stw,t})) + F \\
& \quad + \frac{1 - f(\theta_t) \cdot \eta_t}{1 - \eta_t} \cdot \frac{k_v}{q(\theta_t)}\Big) \\
& + \lambda_{c,t} \cdot g(\epsilon_{stw,t}) \cdot \Big(y_t(\epsilon_{stw,t}, h_{stw,t}(\epsilon_{stw,t})) - v(h_{stw,t}(\epsilon_{stw,t})) + F + \theta_t \cdot k_v\Big) \overset{!}{>} 0
\end{aligned}$$

$$\begin{aligned}\Leftrightarrow\quad & y_t(\epsilon_{stw,t}, h_{stw,t}(\epsilon_{stw,t})) - v(h_{stw,t}(\epsilon_{stw,t})) + F + \theta_t \cdot k_v + (1 - f(\theta_t)) \cdot \frac{E_t\left[\lambda_{n,t+1}\right]}{u'(\tilde{c}_t)} \\ & + \frac{\lambda_{\theta,t-1}}{n_t \cdot u'(\tilde{c}_t^w)} \cdot \beta \cdot \Big(y_t(\epsilon, h_{stw,t}(\epsilon_{stw,t})) - v(h_{stw,t}(\epsilon_{stw,t})) + F \\ & \quad + \frac{1 - f(\theta_t) \cdot \eta_t}{1 - \eta_t} \cdot \frac{k_v}{q(\theta_t)}\Big) \\ & + \frac{\lambda_{c,t}}{n_t \cdot u'(\tilde{c}_t^w)} \cdot \Big(y_t(\epsilon, h_{stw,t}(\epsilon_{stw,t})) - v(h_{stw,t}(\epsilon_{stw,t})) + F + \theta_t \cdot k_v\Big) \overset{!}{>} 0\end{aligned}$$

$$\begin{aligned}\Leftrightarrow\quad & y_t(\epsilon_{stw,t}, h_{stw,t}(\epsilon_{stw,t})) - v(h_{stw,t}(\epsilon_{stw,t})) + F + \theta_t \cdot k_v + (1 - f(\theta_t)) \cdot \frac{E_t\left[\lambda_{n,t+1}\right]}{u'(\tilde{c}_t^w)} \\ & + \left(\frac{\beta \cdot (1 + \eta_{t-1} \cdot BE_t) \cdot \lambda_{\theta,t-1}}{n_t \cdot u'(\tilde{c}_t^w)}\right) \cdot \Big(y_t(\epsilon_{stw,t}, h_{stw,t}(\epsilon_{stw,t})) - v(h_{stw,t}(\epsilon_{stw,t})) + F \\ & \quad + \frac{1 - f(\theta_t) \cdot \eta_t}{1 - \eta_t} \cdot \frac{k_v}{q(\theta_t)}\Big) \\ & - \frac{\beta \cdot BE_t \cdot \lambda_{\theta,t-1}}{n_t \cdot u'(\tilde{c}_t^w)} \cdot \frac{1 - f(\theta_t)}{1 - \eta_t} \cdot \frac{k_v}{q(\theta_t)} \overset{!}{>} 0\end{aligned}$$

Next, we insert the Lagrange multiplier for the law of motion of employment $E_t[\lambda_{n,t+1}]$:

$$\begin{aligned}& y_t(\epsilon_{stw,t}, h_{stw,t}(\epsilon_{stw,t})) - v(h_{stw,t}(\epsilon_{stw,t})) + F + \theta_t \cdot k_v \\ & + (1 - f(\theta_t)) \cdot \frac{1 + \chi_t}{1 - \gamma} \cdot \frac{k_v}{q(\theta_t)} \\ & + \left(\frac{\beta \cdot (1 + \eta_{t-1} \cdot BE_t) \cdot \lambda_{\theta,t-1}}{n_t \cdot u'(\tilde{c}_t)}\right) \cdot \Big(y_t(\epsilon_{stw,t}, h_{stw,t}(\epsilon_{stw,t})) - v(h_{stw,t}(\epsilon_{stw,t})) + F \\ & \quad + \frac{1 - f(\theta_t) \cdot \eta_t}{1 - \eta_t} \cdot \frac{k_v}{q(\theta_t)}\Big) \\ & - \frac{\beta \cdot BE_t \cdot \lambda_{\theta,t-1}}{n_t \cdot u'(\tilde{c}_t)} \cdot \frac{1 - f(\theta_t)}{1 - \eta_t} \cdot \frac{k_v}{q(\theta_t)} = 0\end{aligned}$$

$$\begin{aligned}\Leftrightarrow\quad & y_t(\epsilon_{stw,t}, h_{stw,t}(\epsilon_{stw,t})) - v(h_{stw,t}(\epsilon_{stw,t})) + F \\ & + \frac{(1 - \gamma \cdot f(\theta_t)) + (1 - f(\theta_t)) \cdot \chi_t}{1 - \gamma} \cdot \frac{k_v}{q(\theta_t)} \\ & + \left(\frac{\beta \cdot (1 + \eta_{t-1} \cdot BE_t) \cdot \lambda_{\theta,t-1}}{n_t \cdot u'(\tilde{c}_t)}\right) \cdot \Big(y_t(\epsilon_{stw,t}, h_{stw,t}(\epsilon_{stw,t})) - v(h_{stw,t}(\epsilon_{stw,t})) + F \\ & \quad + \frac{1 - f(\theta_t) \cdot \eta_t}{1 - \eta_t} \cdot \frac{k_v}{q(\theta_t)}\Big) \\ & - \frac{\beta \cdot BE_t \cdot \lambda_{\theta,t-1}}{n_t \cdot u'(\tilde{c}_t)} \cdot \frac{1 - f(\theta_t)}{1 - \eta_t} \cdot \frac{k_v}{q(\theta_t)} = 0\end{aligned}$$

To reduce the notational burden, let us define the difference in the surplus between a match at productivity level $\epsilon_{stw,t}$ and a match at the separation threshold without STW as $\epsilon_{s,t}$ as:

$$y_t(\epsilon_{stw,t}, h_{stw,t}(\epsilon_{stw,t})) - \nu(h_{stw,t}(\epsilon_{stw,t})) - y_t(\epsilon_{s,t}, h_{stw,t}(\epsilon_{s,t})) + \nu(h_{stw,t}(\epsilon_{s,t})) \\ = z_{stw,t}(\epsilon_{stw,t}) - z_{stw,t}(\epsilon_{s,t}) \; > \; 0$$

The surplus of a match with higher productivity $\epsilon_{stw,t} > \epsilon_{s,t}$ is, of course, larger than a match with smaller productivity. Next, we subtract the socially optimal separation condition on STW from the expression above. We can interpret the operation as deducting the social value of rescuing a match with productivity level $\epsilon_{s,t}$ from the social value of rescuing a match with productivity level $\epsilon_{s,t}$:

$$\left(1 + \frac{\beta \cdot (1 + \eta_{t-1} \cdot BE_t) \cdot \lambda_{\theta,t-1}}{n_t \cdot u'(\tilde{c}_t)}\right) \cdot \left(z_{stw,t}(\epsilon_{stw,t}) - z_{stw,t}(\epsilon_{s,t})\right) \\ - \frac{1}{g(\epsilon_s)} \cdot \frac{\lambda_\epsilon}{n \cdot u'(\tilde{c}^w)} \cdot \left(a \cdot h(\epsilon_s)^\alpha + c_f\right) \overset{!}{>} 0$$

From the FOC of the eligibility condition, we get the Lagrange multiplier for the decentralized separation condition. Inserting into the equation above gives:

$$\left(1 + \frac{\beta \cdot (1 + \eta_{t-1} \cdot BE_t) \cdot \lambda_{\theta,t-1}}{n_t \cdot u'(\tilde{c}_t)}\right) \cdot \left(z_{stw,t}(\epsilon_{stw,t}) - z_{stw,t}(\epsilon_{s,t})\right) \\ + \frac{1}{g(\epsilon_{s,t})} \cdot \left(1 + \frac{\beta \cdot (1 + \eta \cdot BE_t) \cdot \lambda_{\theta,t-1}}{n_t \cdot u'(\tilde{c}^w_t)}\right) \cdot \frac{\partial \Omega_t}{\partial \tau_{stw,t}} \cdot \frac{a_t \cdot h_t(\epsilon_{s,t})^\alpha + c_f}{\bar{h} - h_{stw,t}(\epsilon_{s,t})} \overset{!}{>} 0$$

Rearranging, we see that the Ramsey planner always wants to choose a looser eligibility condition

$$\left(1 + \frac{\beta \cdot (1 + \eta_{t-1} \cdot BE_t) \cdot \lambda_{\theta,t-1}}{n_t \cdot u'(\tilde{c}_t)}\right) \cdot \left(z_{stw,t}(\epsilon_{stw,t}) - z_{stw,t}(\epsilon_{s,t}) \\ + \frac{1}{g(\epsilon_{s,t})} \cdot \frac{\partial \Omega_t}{\partial \tau_{stw,t}} \cdot \frac{a_t \cdot h_t(\epsilon_{s,t})^\alpha + c_f}{\bar{h} - h_{stw,t}(\epsilon_{s,t})}\right) > 0$$

as long as:

$$\left(1 + \frac{\beta \cdot (1 + \eta_{t-1} \cdot BE_t) \cdot \lambda_{\theta,t-1}}{n_t \cdot u'(\tilde{c}^w_t)}\right) > 0$$

Assuming steady state and rearranging, we see that the expression closely resembles the expression for the sufficient condition:

$$n \cdot u'(\tilde{c}^w) \cdot \frac{\partial \Omega}{\partial \epsilon_{stw}} + (1 + \eta \cdot BE) \cdot \lambda_\theta \cdot \frac{\partial \Omega}{\partial \epsilon_{stw}} \geq 0$$

However, we need to determine λ_θ. Remember the expression for χ :

$$\begin{aligned}\chi \;=\; & \frac{1}{1-\eta}\cdot\frac{\gamma-\beta\cdot\left(\gamma-\left(1-(1-\eta)\cdot BE\right)\cdot\eta\cdot f(\theta)\right)\cdot(1-G(\epsilon_s))}{m(\theta)\cdot u'(\tilde{c}^w)}\cdot\lambda_\theta\\ & -\frac{1}{1-\eta}\cdot\frac{(\gamma-\eta\cdot f(\theta))\cdot\left(-\lambda_\epsilon\right)}{m(\theta)\cdot u'(\tilde{c}^w)}\\ & -\frac{1}{1-\eta}\cdot\frac{(\gamma-\eta\cdot f(\theta))\cdot\left(-\lambda_\xi\right)}{m(\theta)\cdot u'(\tilde{c}^w)}\end{aligned}$$

We know λ_ϵ but we still need to determine λ_ξ to determine λ_θ.

$$\begin{aligned}-\frac{\partial}{\partial\xi_{s,t}} \;=\; & n_t\cdot g(\xi_{s,t})\cdot\left[y_t(\xi_{s,t},h_t(\xi_{s,t}))-v(h_t(\xi_{s,t}))+F+k_v\cdot\theta_t\right]\cdot\underbrace{\Lambda_t\cdot\tilde{u}'(c_t^F)}_{u'(\tilde{c}_t^w)}\\ & +n_t\cdot g(\xi_{s,t})\cdot(1-f(\theta_t))\cdot E_t\left[\lambda_{n,t+1}\right]\\ & +\lambda_{\theta,t-1}\cdot g(\xi_{s,t})\cdot\beta\cdot\left[y_t(\xi_{s,t},h_t(\xi_{s,t}))-v(h_t(\xi_{s,t}))+F+\frac{1-f(\theta_t)\cdot\eta}{1-\eta}\cdot\frac{k_v}{q(\theta_t)}\right]\\ & +\lambda_{c,t}\cdot g(\xi_{s,t})\cdot\left[y_t(\xi_{s,t},h_t(\xi_{s,t}))-v(h_t(\xi_{s,t}))+F+\theta_t\cdot k_v\right]\\ & +\lambda_{\xi,t}\cdot\Bigg[a_t\cdot h_t(\xi_{s,t})^\alpha+c_f\\ & \quad+\frac{\partial h_t(\xi_{s,t})}{\partial\xi_{s,t}}\cdot\underbrace{\left(\alpha\cdot a_t\cdot\xi_{s,t}\cdot h_{stw,t}(\xi_{s,t})^{\alpha-1}-v'(h_{stw,t}(\xi_{s,t}))-\tau_{stw,t}\right)}_{=0}\Bigg] \;=\; 0\end{aligned}$$

Rearranging gives the social value for rescuing matches with productivity $\xi_{s,t}$.

$$\begin{aligned}& y_t(\xi_{s,t},h_{stw,t}(\xi_{s,t}))-v(h_{stw,t}(\xi_{s,t}))+F+k_v\cdot\theta_t+(1-f(\theta_t))\cdot\frac{E_t\left[\lambda_{n,t+1}\right]}{u'(\tilde{c}_t^w)}\\ & \quad-\frac{\beta\cdot(1+\eta\cdot BE_t)\cdot\lambda_{\theta,t-1}}{n_t\cdot u'(\tilde{c}_t^w)}\cdot\Bigg[y_t(\xi_{s,t},h_t(\xi_{s,t}))-v(h_t(\xi_{s,t}))+F\\ & \qquad\qquad+\frac{1-f(\theta_t)\cdot\eta}{1-\eta}\cdot\frac{k_v}{q(\theta_t)}\Bigg]\\ & \quad-\frac{\beta\cdot\eta\cdot BE_t\cdot\lambda_{\theta,t-1}}{n_t\cdot u'(\tilde{c}_t^w)}\cdot\frac{1-f(\theta_t)}{1-\eta}\cdot\frac{k_v}{q(\theta_t)}\\ & \quad=-\frac{1}{g(\xi_{s,t})}\cdot\frac{\lambda_{\xi,t}}{n_t\cdot u'(\tilde{c}_t^w)}\cdot\left(a_t\cdot h_t(\xi_{s,t})^\alpha+c_f\right)\end{aligned}$$

We can express the social value of rescuing a worker with productivity $\xi_{s,t}$ in dependence of the social value of rescuing the marginal matches $\xi_{s,t}$:

$$
\begin{aligned}
&- \frac{1}{g(\xi_{s,t})} \cdot \frac{\lambda_{\xi,t}}{n_t \cdot u'(\tilde{c}_t^w)} \cdot \left(a_t \cdot h_t(\xi_{s,t})^\alpha + c_f\right) = \\
&\quad - \frac{1}{g(\epsilon_{s,t})} \cdot \frac{\lambda_{\epsilon,t}}{n_t \cdot u'(\tilde{c}_t^w)} \cdot \left(a_t \cdot h_{stw,t}(\epsilon_{s,t})^\alpha + c_f\right) \\
&\quad + \left(1 + \frac{\beta \cdot (1 + \eta \cdot BE_t) \cdot \lambda_{\theta,t-1}}{n_t \cdot u'(\tilde{c}_t^w)}\right) \cdot \left(z_t(\xi_{s,t}) - z_{stw,t}(\epsilon_{s,t})\right)
\end{aligned}
$$

Higher productive matches must be socially more valuable than less productive matches. We can see this by the fact: $\left(z_t(\xi_{s,t}) - z_{stw,t}(\epsilon_{s,t})\right) > 0$. Inserting the Lagrange multiplier for the separation threshold with STW gives:

$$
\begin{aligned}
&- \frac{1}{g(\xi_{s,t})} \cdot \frac{\lambda_{\xi,t}}{n_t \cdot u'(\tilde{c}_t^w)} \cdot \left(a_t \cdot h_t(\xi_{s,t})^\alpha + c_f\right) = \\
&+ \left(1 + \frac{\beta \cdot (1 + \eta \cdot BE_t) \cdot \lambda_{\theta,t-1}}{n_t \cdot u'(\tilde{c}_t^w)}\right) \cdot \left(\frac{1}{g(\epsilon_s)} \cdot \frac{\partial \Omega}{\partial \tau_{stw}} \cdot \frac{a \cdot h_{stw}(\epsilon_s)^\alpha + c_f}{\bar{h} - h_{stw}(\epsilon_s)} \right. \\
&\quad \left. + \left(z_t(\xi_{s,t}) - z_{stw,t}(\epsilon_{s,t})\right)\right)
\end{aligned}
$$

Using the same procedure as in section H.3, we can derive the Lagrange multiplier λ_θ and get the following statement:

$$
\begin{aligned}
&\underbrace{n \cdot \frac{\partial \Omega}{\partial \epsilon_{stw}}}_{\text{More hours distortion}} + L_V \cdot \underbrace{\left(-\frac{\partial f^{ge}}{\partial \epsilon_{stw}} \cdot u\right)}_{\text{Less hiring}} + L_S^\star \cdot \underbrace{\left(n \cdot g(\epsilon_s) \cdot \frac{\partial \epsilon_s^{ge}}{\partial \epsilon_{stw}}\right)}_{\text{More Separations}} \\
&+ \underbrace{\left(L_S^\star + z_t(\xi_{s,t}) - z_{stw,t}(\epsilon_{s,t})\right) \cdot n \cdot g(\xi_s) \cdot \frac{\partial \xi_s^{ge}}{\partial \epsilon_{stw}}}_{\text{Loose additional workers that could have been employed}} > 0
\end{aligned}
$$

The statement must be true by assumption, as

$$
\underbrace{n \cdot \frac{\partial \Omega}{\partial \epsilon_{stw}}}_{\text{More hours distortion}} + L_V \cdot \underbrace{\left(-\frac{\partial f^{ge}}{\partial \epsilon_{stw}} \cdot u\right)}_{\text{Less hiring}} + L_S^\star \cdot \underbrace{\left(n \cdot g(\epsilon_s) \cdot \frac{\partial \epsilon_s^{ge}}{\partial \epsilon_{stw}}\right)}_{\text{More Separations}} > 0
$$

proving the argument!

case 2: $\xi_{s,t} < \epsilon_{stw,t} < \xi_{stw,t}$ Assume now that we choose a looser eligibility condition. The first derivative for the eligibility condition is:

$$
\frac{\partial}{\partial \epsilon_{stw,t}} = - \underbrace{\frac{\nu_t^f}{\nu_t^f} \cdot \tilde{u}'(\tilde{c}_t^w)}_{u'(\tilde{c}_t^w)} \cdot n_t \cdot \frac{\partial \Omega_t}{\partial \epsilon_{stw,t}} - \beta \cdot \left(1 + \eta \cdot BE_t\right) \cdot \lambda_{\theta,t-1} \cdot \frac{\partial \Omega_t}{\partial \epsilon_{stw,t}}
$$

If we can show that

$$\frac{\partial}{\partial \epsilon_{stw}} = \tilde{u}'(\tilde{c}^w) \cdot n \cdot \frac{\partial \Omega}{\partial \epsilon_{stw}} + \beta \cdot \big(1 + \eta \cdot BE\big) \cdot \lambda_\theta \cdot \frac{\partial \Omega}{\partial \epsilon_{stw}} > 0$$

then it must always be better to set a stricter eligibility condition. This is practically the necessary condition of proposition 3. To show this, we have to calculate the Lagrange multiplier for the job-creation condition again. Note that we can show that the Lagrange multiplier for the separation threshold for STW is zero:

$$\begin{aligned}\frac{\partial}{\partial \xi_{s,t}} &= -\lambda_{\xi,t} \cdot \left(a_t \cdot h_t^\alpha(\xi_{s,t}) + c_f + \frac{\partial h_t(\xi_{s,t})}{\partial \xi_{s,t}} \cdot \underbrace{\big(\alpha \cdot a_t \cdot \xi_{s,t} \cdot h_t(\xi_{s,t})^{\alpha-1} - v'(h_t(\xi_{s,t}))\big)}_{=0} \right) \\ &= 0 \\ &\Leftrightarrow \lambda_{\xi,t} = 0\end{aligned}$$

Any increase in the separation condition without STW does not affect separations. $\lambda_\xi = 0$ simplifies the expression for χ to:

$$\begin{aligned}\chi &= \frac{1}{1-\eta} \cdot \frac{\gamma - \beta \cdot \big(\gamma - \big(1 - (1-\eta) \cdot BE\big) \cdot \eta \cdot f(\theta)\big) \cdot (1 - G(\epsilon_s))}{m(\theta) \cdot u'(\tilde{c}^w)} \cdot \lambda_\theta \\ &- \frac{1}{1-\eta} \cdot \frac{(\gamma - \eta \cdot f(\theta)) \cdot \big(-\lambda_\epsilon\big)}{m(\theta) \cdot u'(\tilde{c}^w)}\end{aligned}$$

Using the same steps as before, we can derive that if

$$\underbrace{n \cdot \frac{\partial \Omega}{\partial \epsilon_{stw}}}_{\text{More hours distortion}} + L_V \cdot \underbrace{\left(-\frac{\partial f^{ge}}{\partial \epsilon_{stw}} \cdot u \right)}_{\text{Less hiring}} + L_S^\star \cdot \underbrace{\left(n \cdot g(\epsilon_s) \cdot \frac{\partial \epsilon_s^{ge}}{\partial \epsilon_{stw}} \right)}_{\text{More Separations}} > 0$$

then the Ramsey planner always wants to set a stricter eligibility condition, which is the necessary condition from Proposition 3.

case 3: $\xi_{stw,t} \leq \epsilon_{stw,t}$.

Case 3 describes the situation in which the eligibility threshold is so loose that it is no longer binding. Firms and workers decide when to enter the STW system. A looser eligibility condition does not have an impact on the economy anymore. However, in case 2, we have seen that a stricter eligibility condition will reduce welfare losses from the use of STW. As a result, case 3 cannot be optimal.

1.I.6 Optimal UI considering optimal STW in Steady State

The FOC for the UI benefits for the UI system with optimal STW is the same as for the UI system given UI:

$$\begin{aligned}\frac{\partial}{\partial b_t} &= (1-n_t)\cdot\left(u'(b_t)-u'(\tilde{c}_t^w)\right)\\ &\quad -\lambda_{\theta,t}\cdot\beta\cdot\left[\left(\frac{u'(b_t)}{u'(\tilde{c}_t^w)}+\frac{n_t}{1-n_t}\right)-BE_t\cdot\left(\left(1-\eta\right)\cdot\frac{u'(b_t)}{u'(\tilde{c}_t^w)}-\eta\cdot\frac{n_t}{1-n_t}\right)\right]\end{aligned}$$

What changes again is the definition for λ_θ. To calculate λ_θ consider first how we calculate χ:

$$\begin{aligned}\chi\cdot\left(\frac{1}{1-\gamma}\cdot\frac{k_v}{q(\theta)}\right) &= \frac{\eta-\gamma}{1-\eta}\cdot\left(\frac{1}{1-\gamma}\cdot\frac{k_v}{q(\theta)}\right)\\ &\quad+\frac{\beta}{1-\beta\cdot(1-f(\theta))\cdot(1-G(\epsilon_s))}\cdot\left(1+\frac{\beta\cdot(1+\eta\cdot BE)\cdot\lambda_\theta}{n\cdot u'(\tilde{c}^w)}\right)\cdot\frac{b}{n}\end{aligned}$$

With an optimal set STW system, χ can be denoted as:

$$\begin{aligned}\chi &= \frac{1}{1-\eta}\cdot\frac{\gamma-\beta\cdot\left(\gamma-\left(1-(1-\eta)\cdot BE\right)\cdot\eta\cdot f(\theta)\right)\cdot(1-G(\epsilon_s))}{m(\theta)\cdot u'(\tilde{c}^w)}\cdot\lambda_\theta\\ &\quad-\frac{1}{1-\eta}\cdot\frac{(\gamma-\eta\cdot f(\theta))\cdot\left(-\lambda_\epsilon\right)}{m(\theta)\cdot u'(\tilde{c}^w)}\\ &\quad-\frac{1}{1-\eta}\cdot\frac{(\gamma-\eta\cdot f(\theta))\cdot\left(-\lambda_{stw}\right)}{m(\theta)\cdot u'(\tilde{c}^w)}\end{aligned}$$

The Lagrange multiplier for the decentralized separation condition in steady state is:

$$\frac{\lambda_\epsilon}{n\cdot u'(\tilde{c}^w)} = -\frac{1+\frac{\beta\cdot(1+\eta\cdot BE)\cdot\lambda_\theta}{n\cdot u'(\tilde{c}^w)}}{\bar{h}-h_{stw}(\epsilon_s)}\cdot\frac{\partial\Omega}{\partial\tau_{stw}}$$

Further, we need to calculate the Lagrange multiplier for the optimal separation condition $\xi_{s,t}=\epsilon_{stw,t}$. To do that, consider the FOC for the eligibility threshold:

$$\begin{aligned}\frac{\partial}{\partial\epsilon_{stw,t}} &= -\underbrace{\Lambda_t\cdot\tilde{u}(c_t^F)}_{u'(\tilde{c}_t^w)}\cdot n_t\cdot\frac{\partial\Omega_t}{\partial\epsilon_{stw,t}}-\beta\cdot\left(1+\eta\cdot BE_t\right)\cdot\lambda_{\theta,t-1}\cdot\frac{\partial\Omega_t}{\partial\epsilon_{stw,t}}\\ &\quad-\lambda_{stw,t}\cdot\left[a_t\cdot h_t(\epsilon_{s,t})^\alpha+c_f+\frac{\partial h_t(\epsilon_{s,t})}{\partial\epsilon_{s,t}}\cdot\underbrace{\left(\alpha\cdot a_t\cdot\epsilon_{s,t}\cdot h_t(\epsilon_{s,t})^{\alpha-1}-\nu'(h_t(\epsilon_{s,t}))\right)}_{=0}\right]\\ &= 0\end{aligned}$$

Rearranging gives an expression for the FOC of the Lagrange multiplier:

$$\begin{aligned}\lambda_{stw} &= -\frac{n \cdot u'(\tilde{c}^w) + \beta \cdot (1 + \eta \cdot BE) \cdot \lambda_\theta}{a \cdot h(\epsilon_{stw})^\alpha + c_f} \cdot \frac{\partial \Omega}{\partial \epsilon_{stw}} \\ &= -n \cdot u'(\tilde{c}^w) \cdot \left(1 + \frac{\beta \cdot (1 + \eta \cdot BE)}{n \cdot u'(\tilde{c}^w)} \cdot \lambda_\theta\right) \cdot \frac{\partial \Omega}{\partial \epsilon_{stw}} / (a \cdot h(\epsilon_{stw})^\alpha + c_f)\end{aligned}$$

Using our expression for ξ and inserting it into the expression for the deviation of the private and social value of the match gives:

$$\begin{aligned}&\frac{1}{(1-\eta)} \cdot \frac{\gamma - \beta \cdot \big(\gamma - \big(1 - (1-\eta) \cdot BE\big) \cdot \eta \cdot f(\theta)\big) \cdot (1 - G(\epsilon_s))}{f(\theta) \cdot u'(\tilde{c}^w)} \cdot \lambda_\theta = \\ &+ f'(\theta) \cdot u \cdot \Big(\frac{\eta - \gamma}{(1-\eta) \cdot (1-\gamma)} \cdot \mathscr{J} \\ &\quad + \frac{\beta}{1 - \beta \cdot (1 - f(\theta)) \cdot (1 - G(\epsilon_s))} \cdot \left(1 + \frac{\beta \cdot (1 + \eta \cdot BE) \cdot \lambda_\theta}{n \cdot u'(\tilde{c}^w)}\right) \cdot \frac{b}{n}\Big) \\ &+ \frac{1}{(1-\eta)} \cdot \left(1 + \frac{\beta \cdot (1 + \eta \cdot BE) \cdot \lambda_\theta}{n \cdot u'(\tilde{c}^w)}\right) \cdot \frac{\partial \Omega}{\partial \tau_{stw}} \cdot \frac{\partial \tau_{stw}}{\partial \epsilon_s} \cdot \frac{\partial \epsilon_s}{\partial \theta} \cdot n \\ &+ \frac{1}{(1-\eta)} \cdot \left(1 + \frac{\beta \cdot (1 + \eta \cdot BE) \cdot \lambda_\theta}{n \cdot u'(\tilde{c}^w)}\right) \cdot \frac{\partial \Omega}{\partial \epsilon_{stw}} \cdot \frac{\partial \epsilon_{stw}}{\partial \theta} \cdot n\end{aligned}$$

with

$$\begin{aligned}f'(\theta) &= (1-\gamma) \cdot q(\theta) \\ \frac{\partial \epsilon_{stw}}{\partial \theta} &= \frac{n \cdot (\gamma - \eta \cdot f(\theta))}{f(\theta)} / (a \cdot h_{stw}(\epsilon_{stw})^\alpha + c_f) \\ \frac{\partial \tau_{stw}}{\partial \epsilon_s} &= -\frac{a \cdot h_{stw}(\epsilon_s)^\alpha + c_f}{\bar{h} - h_{stw}(\epsilon_s)} \\ \frac{\partial \epsilon_s}{\partial \theta} &= \frac{1}{1-\eta} \cdot \frac{(\gamma - \eta \cdot f(\theta)) \cdot k_v}{f(\theta)} \cdot \frac{1}{a \cdot h_{stw}(\epsilon_s)^\alpha + c_f} \cdot k_v\end{aligned}$$

Next, solve for the Lagrange multiplier of the social value of the match:

$$\begin{aligned}\lambda_\theta = &\frac{u'(\tilde{c}^w)}{M'} \cdot \frac{\partial n}{\partial \theta} \cdot \Big[\frac{\eta - \gamma}{\big(1-\eta\big) \cdot (1-\gamma)} \cdot \frac{k_v}{q(\theta)} \\ &\quad + \frac{\beta}{1 - \beta \cdot (1 - f(\theta)) \cdot (1 - G(\epsilon_s))} \cdot \frac{b}{n}\Big] \\ &+ \frac{u'(\tilde{c}^w)}{M'} \cdot \frac{\partial \Omega}{\partial \tau_{stw}} \cdot \frac{\partial \tau_{stw}}{\partial \epsilon_s} \cdot \frac{\partial \epsilon_s}{\partial \theta} \cdot g(\epsilon_s) \cdot n \\ &+ \frac{u'(\tilde{c}^w)}{M'} \cdot \frac{\partial \Omega}{\partial \epsilon_{stw}} \cdot \frac{\partial \epsilon_{stw}}{\partial \theta} \cdot g(\epsilon_{stw}) \cdot n\end{aligned}$$

Define M' as the inverse multiplier in case of an optimal STW system.

$$M' = \underbrace{\frac{\gamma - \beta \cdot \left(\gamma - \left(1 - (1-\eta) \cdot BE\right) \cdot \eta \cdot f(\theta)\right) \cdot (1 - G(\epsilon_s))}{(1-\eta) \cdot f(\theta)}}_{\text{regular multiplier}}$$
$$- \underbrace{\frac{\beta}{1 - \beta \cdot (1 - f(\theta)) \cdot (1 - G(\epsilon_s))} \cdot \frac{(1 + \eta \cdot BE)}{n} \cdot \frac{b}{n} \cdot f'(\theta) \cdot u}_{\text{Amplification over tax}}$$
$$- \underbrace{\frac{\beta \cdot (1 + \eta \cdot BE)}{n} \cdot \frac{\partial \Omega}{\partial \tau_{stw}} \cdot \frac{\partial \tau_{stw}}{\partial \epsilon_s} \cdot \frac{\partial \epsilon_s}{\partial \theta} \cdot n}_{\text{Amplification over reaction of STW beneftis}}$$
$$- \underbrace{\frac{\beta \cdot (1 + \eta \cdot BE)}{n} \cdot \frac{\partial \Omega}{\partial \epsilon_{stw}} \cdot \frac{\partial \epsilon_{stw}}{\partial \theta} \cdot n}_{\text{Amplification over reaction of eligibility condition}}$$

Compared to the inverse multiplier of the exogenous STW system, the regular channel and the influence of vacancy posting on the costs of the UI system remain largely unchanged. However, the separation channel now accounts for the reaction of the STW system. When UI benefits increase, the joint surplus of firms and workers decreases, leading to a higher separation rate. In response, the STW system offers more generous benefits, but these benefits distort working hours, thereby reducing production. The subsidy effect of STW is effectively nullified by the corresponding increase in taxes, further decreasing the joint surplus of firms and workers.

Not only does the separation rate rise for firms with access to STW, but it also increases for firms without access to STW. As a result, the eligibility condition must be loosened, spreading the distortionary effects of reduced working hours across more firms. This, similar to an increase in STW benefits, leads to a decline in production, which in turn reduces the joint surplus of firms and workers.

Again, the paper assumes the inverse multiplier to be positive to guarantee convergence of the steady state. We can abbreviate the inverse multiplier by recognizing the fact that it captures the general equilibrium effect a change in the surplus or value of the firm has on the labor market tightness:

$$\frac{\partial \theta^{ge}}{\partial S} = \frac{1}{M'} \quad \text{or} \quad \frac{\partial \theta^{ge}}{\partial J} = \frac{1}{1-\eta} \cdot \frac{1}{M'}$$

Plugging this and the expression for λ_θ into the FOC for the optimal UI gives:

$$\underbrace{(1-n)\cdot\frac{u'(b)-u'(\tilde{c}^w)}{u'(\tilde{c}^w)}}_{\text{Provide additional Income Insurance}} =$$

$$+\,L_V\cdot\underbrace{\left(-\frac{\partial f^{ge}}{\partial b}\cdot u\right)}_{\text{Less Hiring}} + \underbrace{n\cdot\frac{\partial\Omega}{\partial\tau_{stw}}\cdot\frac{\partial\tau_{stw}}{\partial\epsilon_s}\cdot\frac{\partial\epsilon_s^{ge}}{\partial b}}_{\text{STW Benefits increase to counter Separations}} + \underbrace{n\cdot\frac{\partial\Omega}{\partial\epsilon_{stw}}\cdot\frac{\partial\epsilon_{stw}^{ge}}{\partial b}}_{\text{Looser eligibility condition}}$$

Let L_V denote the social value of an additional hire:

$$L_V = \left[\frac{\eta-\gamma}{\left(1-\eta\right)\cdot(1-\gamma)}\cdot\frac{k_v}{q(\theta)} + \frac{\beta}{1-\beta\cdot(1-f(\theta))\cdot(1-G(\epsilon_s))}\cdot\frac{b}{n}\right]$$

Further, note that most of the right-hand side of the optimal UI benefits denote the first derivative of the value of the firm to an increase in UI benefits:

$$-\frac{\partial J}{\partial b} = \beta\cdot(1-\eta)\cdot\left[\left(\frac{u'(b)}{u'(\tilde{c}^w)}+\frac{n}{1-n}\right)-BE\cdot\left(\left(1-\eta\right)\cdot\frac{u'(b)}{u'(\tilde{c}^w)}-\eta\cdot\frac{n}{1-n}\right)\right]$$

Further, denote:

$$\begin{aligned}
(-\frac{\partial f^{ge}}{\partial b})\cdot u &= f'(\theta)\cdot u\cdot\frac{\partial\theta^{ge}}{\partial J}\cdot\frac{\partial J}{\partial b}\\
\frac{\partial\epsilon_s^{ge}}{\partial b} &= \frac{\partial J}{\partial b}\cdot\frac{\partial\theta^{ge}}{\partial J}\cdot\left(-\frac{\partial\epsilon_s}{\partial\theta}\right)\\
\frac{\partial\epsilon_{stw}^{ge}}{\partial b} &= \frac{\partial J}{\partial b}\cdot\frac{\partial\theta^{ge}}{\partial J}\cdot\left(-\frac{\partial\epsilon_{stw}}{\partial\theta}\right)
\end{aligned}$$

We can also reformulate the optimality condition for UI benefits under an optimal STW policy to align it with the expression for the optimality condition of UI with a non-optimized STW system. To do this, note that

$$\begin{aligned}
n\cdot\frac{\partial\Omega}{\partial\tau_{stw}}\cdot\frac{\partial\tau_{stw}}{\partial\epsilon_s}\cdot\frac{\partial\epsilon_s^{ge}}{\partial b} &= \frac{n\cdot\frac{\partial\Omega}{\partial\tau_{stw}}}{n\cdot g(\epsilon_s)\cdot\frac{\partial\epsilon_s}{\partial\tau_{stw}}}\cdot n\cdot g(\epsilon_s)\cdot\frac{\partial\epsilon_s^{ge}}{\partial b}\\
&= L_S^{\star}\cdot n\cdot g(\epsilon_s)\cdot\frac{\partial\epsilon_s^{ge}}{\partial b}
\end{aligned}$$

and

$$\begin{aligned}
n\cdot\frac{\partial\Omega}{\partial\epsilon_{stw}}\cdot\frac{\partial\epsilon_{stw}^{ge}}{\partial b} &= \frac{n\cdot\frac{\partial\Omega}{\partial\tau_{stw}}}{n\cdot g(\epsilon_s)}\cdot n\cdot g(\epsilon_{stw})\cdot\frac{\partial\epsilon_s^{ge}}{\partial b}\\
&= L_{STW}^{\star}\cdot n\cdot g(\epsilon_{stw})\cdot\frac{\partial\epsilon_{stw}^{ge}}{\partial b}
\end{aligned}$$

Therefore, we can rewrite the condition as:

$$\underbrace{(1-n)\cdot\frac{u'(b)-u'(\tilde{c}^w)}{u'(\tilde{c}^w)}}_{\text{Provide Income Insurance}}$$
$$= L_V \cdot \underbrace{\left(-\frac{\partial f^{ge}}{\partial b}\cdot u\right)}_{\text{Less Hiring}} + L_S^\star \cdot \underbrace{\left(n\cdot g(\epsilon_s)\cdot\frac{\partial \epsilon_s^{ge}}{\partial b}\right)}_{\text{More Separations}} + L_{STW}^\star \cdot \underbrace{\left(n\cdot g(\epsilon_{stw})\cdot\frac{\partial \epsilon_{stw}^{ge}}{\partial b}\right)}_{\text{More Workers on STW}}$$

Appendix 1.J Lemma 1

(i). The welfare costs of STW can be denoted as:

$$\Omega_t = \int_{\epsilon_{s,t}}^{\epsilon_{stw,t}} [y_t(\epsilon) - v\big(h_t(\epsilon)\big) - y_{stw,t}(\epsilon) + v\big(h_{stw,t}(\epsilon)\big)]dG(\epsilon)$$

From Nash-Bargaining, we can infer:

$$y_t(\epsilon_{stw,t}) - v(h_t(\epsilon_{stw,t})) - \big[y_{stw,t}(\epsilon_{stw,t}) - v(h_{stw,t}(\epsilon_{stw,t}))\big] > 0$$

Note that in Nash-Bargaining $h_t(\epsilon_{stw,t})$ is chosen to maximize $y_t(\epsilon_{stw,t}) - v(h_t(\epsilon_{stw,t}))$. And $h_{stw,t}(\epsilon_{stw,t})$ is chosen to maximize $y_t(\epsilon_{stw,t}) - v(h_t(\epsilon_{stw,t})) + (\bar{h} - h_{stw,t}(\epsilon_{stw,t}))\cdot\tau_{stw,t}$. As long as $\tau_{stw,t} \neq 0$ we get $h_t(\epsilon_{stw,t}) \neq h_{stw,t}(\epsilon_{stw,t})$. Thus, the first part of the equation must be larger than the second, which fulfills the condition. This implies that $\Omega_t > 0$

(ii). From the derivative of the welfare costs for the STW benefits:

$$\frac{\partial\Omega_t}{\partial\tau_{stw,t}} = -\int_{\epsilon_{s,t}}^{\epsilon_{stw,t}} \frac{\partial h_{stw,t}(\epsilon)}{\partial\tau_{stw,t}}\cdot\big(\alpha\cdot a_t\cdot\epsilon\cdot h_{stw,t}(\epsilon)^{\alpha-1} - v'(h_{stw,t}(\epsilon))\big)dG(\epsilon) \overset{!}{>} 0$$

Using the optimality condition for the hours choice on STW, we can rewrite the equation to:

$$\frac{\partial\Omega_t}{\partial\tau_{stw,t}} = -\int_{\epsilon_{s,t}}^{\epsilon_{stw,t}} \frac{\partial h_{stw,t}(\epsilon)}{\partial\tau_{stw,t}}\cdot\tau_{stw,t}\cdot dG(\epsilon) \overset{!}{>} 0$$

Since the derivation of the number of hours worked for the STW benefits is negative for $\alpha < 1$

$$\frac{\partial h_{stw,t}(\epsilon)}{\partial \tau_{stw,t}} = \frac{1}{\alpha \cdot (\alpha - 1) \cdot a_t \cdot \epsilon \cdot h_{stw,t}(\epsilon)^{\alpha-2} - \psi \cdot h_{stw,t}^{\psi-1}} < 0$$
$$\text{for } \alpha \in (0,1) \text{ and } \psi, \epsilon, h_{stw,t}(\epsilon) > 0$$

and the STW benefits $\tau_{stw,t}$ need to be positive, we can conclude that the derivation for the welfare costs of STW must be positive.

(iii). First, note that if we loosen the eligibility condition, then we loosen the eligibility threshold. The relationship is implicitly defined as:

$$D = h_{stw}(\epsilon_{stw})$$

Using the implicit function theorem gives:

$$\frac{\partial \epsilon_{stw}}{\partial D} = -\frac{1}{h'_{stw}(\epsilon_{stw})} > 0$$

From the derivative of the welfare costs for the eligibility condition:

$$\frac{\partial \Omega_t}{\partial \epsilon_{stw,t}} = \Big(y_t(\epsilon_{stw,t}) - v(h_t(\epsilon_{stw,t})) - \big[y_{stw,t}(\epsilon_{stw,t}) - v(h_{stw,t}(\epsilon_{stw,t}))\big]\Big) \cdot g(\epsilon_{stw,t})$$
$$\overset{!}{>} 0$$
$$\Leftrightarrow y_t(\epsilon_{stw,t}) - v(h_t(\epsilon_{stw,t})) - \big[y_{stw,t}(\epsilon_{stw,t}) - v(h_{stw,t}(\epsilon_{stw,t}))\big] \overset{!}{>} 0$$

As argued in (i), this equation must hold due to the Nash-Bargaining set-up.

(iv). Using our result from (iii), it follows:

$$\frac{\partial^2 \Omega_t}{\partial \tau_{stw,t} \partial \epsilon_{stw,t}} = \underbrace{-\frac{\partial h_{stw,t}(\epsilon_{\epsilon_{stw,t}})}{\partial \tau_{stw,t}}}_{>0} \cdot \tau_{stw,t} \cdot g(\epsilon_{stw,t}) > 0$$

Appendix 1.K Corollaries

All corollaries are under the ceteris paribus assumption. For the corollaries, ceteris paribus implies that separation rates and job-finding rates are seen as exogenous. Therefore, all results have to be interpreted as partial equilibrium results.

Corollary 1. The fiscal externality on hiring is defined as:

$$FE = \frac{1}{1-\beta \cdot (1-G(\epsilon_s)) \cdot (1-f)} \cdot \left[b + \frac{1-n}{n} \cdot b \right]$$

Using the law of motion of employment, we can derive the steady state employment rate:

$$n = \frac{f}{f + G(\epsilon_s) \cdot (1-f)}$$

Rearranging the fiscal externality and setting the discount factor to 1 gives:

$$FE = \frac{1}{f + G(\epsilon_s) \cdot (1-f)} \cdot \frac{b}{n}$$

Multiplying and dividing by f recovers employment in the employment rate in the equation:

$$FE = \frac{1}{f} \cdot \frac{f}{f + G(\epsilon_s) \cdot (1-f)} \cdot \frac{b}{n}$$

Inserting the employment rate gives:

$$FE = \frac{b}{f}$$

This proves Corollary 1.

Corollary 2. The optimal STW benefits can be denoted as:

$$\tau_{stw} = \frac{1}{\bar{h} - h_{stw}(\epsilon_s)} \cdot \left(\frac{1-f}{1-\beta \cdot (1-G(\epsilon_s)) \cdot (1-f)} \cdot \frac{b}{n} - \frac{n \cdot \frac{\partial \Omega}{\partial \tau_{stw}}}{n \cdot g(\epsilon_s) \cdot \left(-\frac{\partial \epsilon_s}{\partial \tau_{stw}}\right)} - \tilde{BE} \right)$$

Without risk aversion, the bargaining effect is zero: $\tilde{BE} = 0$.

$$\tau_{stw} = \frac{1}{\bar{h} - h_{stw}(\epsilon_s)} \cdot \left(\frac{1-f}{1-\beta \cdot (1-G(\epsilon_s)) \cdot (1-f)} \cdot \frac{b}{n} - \frac{n \cdot \frac{\partial \Omega}{\partial \tau_{stw}}}{n \cdot g(\epsilon_s) \cdot \left(-\frac{\partial \epsilon_s}{\partial \tau_{stw}}\right)} \right)$$

Note that the working hours distortions do not depend on the job-finding rate. Thus, we can simplify notation to:

$$\tau_{stw} = \frac{1}{\bar{h} - h_{stw}(\epsilon_s)} \cdot \frac{\beta \cdot \left(1 - f\right)}{1 - \beta \cdot (1 - G(\epsilon_s)) \cdot (1 - f)} \cdot \frac{b}{n} - C$$

C denotes a constant that does not depend on f. Note that the steady state employment rate gives:

$$\tau_{stw} = \frac{1}{\bar{h} - h_{stw}(\epsilon_s)} \cdot \frac{1 - f}{1 - \beta \cdot (1 - G(\epsilon_s)) \cdot (1 - f)} \cdot \frac{b}{n(f)} - C$$

Taking the first derivatives gives:

$$\begin{aligned} -\frac{\partial \tau_{stw}}{\partial f} = & \frac{1}{\bar{h} - h_{stw}(\epsilon_s)} \cdot \frac{1 + \beta \cdot (1 - G(\epsilon_s)) \cdot f}{[1 - \beta \cdot (1 - G(\epsilon_s)) \cdot (1 - f)]^2} \cdot \frac{b}{n(f)} \\ & \frac{\beta \cdot (1 - f)}{1 - \beta \cdot (1 - G(\epsilon_s)) \cdot (1 - f)} \cdot \frac{b}{n(f)^2} \cdot n'(f) > 0 \end{aligned}$$

The first term reflects how much longer a worker laid off this period will remain unemployed due to the decline in the job-finding rate. The second term highlights that lower job-finding rates not only decrease employment but also increase the fiscal burden on the UI system, as more workers remain dependent on UI benefits for extended periods:

$$n'(f) = \frac{G(\epsilon_s)}{(f + G(\epsilon_s) \cdot (1 - f))^2}$$

We can conclude that UI benefits must rise after a fall in job-finding rates.

Corollary 3. Without risk aversion, the bargaining effect is zero: $\tilde{BE} = 0$. The optimal STW benefits can thus be expressed as:

$$\begin{aligned} \tau_{stw} = & \frac{1}{\bar{h} - h_{stw}(\epsilon_s)} \cdot \left(\frac{1 - f}{1 - \beta \cdot (1 - G(\epsilon_s)) \cdot (1 - f)} \cdot \frac{b}{n} \right. \\ & \left. - \frac{n \cdot \frac{\partial \Omega}{\partial \tau_{stw}}}{n \cdot g(\epsilon_s) \cdot \left(-\frac{\partial \epsilon_s}{\partial \tau_{stw}} \right)} \right) \end{aligned}$$

The first derivative is:

$$\frac{\partial \tau_{stw}}{\partial D} = -\frac{1}{\bar{h} - h_{stw}(\epsilon_s)} \underbrace{\frac{\partial^2 \Omega}{\partial \tau_{stw} \partial D}}_{> 0} / \underbrace{\left[-g(\epsilon_s) \cdot \frac{\partial \epsilon_s}{\partial \tau_{stw}} \right]}_{> 0} < 0$$

Increasing both STW benefits and eligibility condition increases the distortionary effects of the STW system. STW benefits have to be chosen stricter.

Corollary 4. T.b.s.: $b = 0$ and $\eta = \gamma \Rightarrow \tau_{stw} = 0$

We proceed by guess and verify. Guess that $\tau_{stw} = 0$. Then we know that the welfare costs of STW are zero, and so are its derivatives for the STW benefits

$$\frac{\partial \Omega}{\partial \tau_{stw}} = -\int_{\epsilon_s}^{\epsilon_{stw}} \frac{\partial h_{stw}(\epsilon)}{\partial \tau_{stw}} \cdot \tau_{stw} \cdot dG(\epsilon) = 0$$

and the eligibility condition:

$$\frac{\partial \Omega}{\partial D} = \frac{\partial \epsilon_{stw}}{\partial D} \cdot \big(y(\epsilon_{stw}) - v(h(\epsilon_{stw})) - [y_{stw}(\epsilon_{stw}) - v(h_{stw}(\epsilon_{stw}))]\big) \cdot g(\epsilon_{stw}) = 0$$

as $h_{stw}(\epsilon_s) = h_(\epsilon_s)$. Therefore, we can imply that the welfare cost penalty of STW must be equal to zero:

$$\left[n \cdot \frac{\partial \Omega}{\partial \tau_{stw}}\right] / \left[n \cdot g(\epsilon_s) \cdot \left(-\frac{\partial \epsilon_s}{\partial \tau_{stw}}\right)\right] = 0$$

Further, it is trivial that the fiscal externality of the UI system on separations with $b = 0$ is equal to zero:

$$(1 - f) \cdot FE = 0$$

Finally, we have to look at the bargaining effect. Appendix 1.B shows that the bargaining effect depends on the Lagrange multiplier of the value of the firm. We can write the Lagrange multiplier as

$$\lambda_\theta = L_V + L_S + L_{STW}$$

with

$$L_V = \frac{1}{M'} \cdot \underbrace{\frac{\eta - \gamma}{\big(1 - \eta\big) \cdot (1 - \gamma)}}_{\text{Deviation from Hosios Condition}} + \frac{1}{M'} \cdot \underbrace{FE}_{\text{Fiscal Externality}} / \frac{k_v}{q(\theta)}$$

$$LS = n \cdot g(\epsilon_s) \cdot u'(\tilde{c}^w) \cdot \frac{\partial \epsilon_s}{\partial J} \cdot \frac{\tau_{stw}}{\epsilon_s} \cdot \frac{\partial \Omega}{\partial \tau_{stw}} / M'$$

$$L_{STW} = n \cdot g(\epsilon_{stw}) \cdot u'(\tilde{c}^w) \cdot \frac{\partial D}{\partial J} \cdot \frac{\partial \Omega}{\partial D} / M'$$

When $\eta = \gamma$, then the Hosios condition is fulfilled. Further, $b = 0$ implies that $FE = 0$; thus, we can conclude that vacancy posting must be efficient, $L_V = 0$. Further, we already established that for $\tau_{stw} = 0$ STW does not distort the economy: $\frac{\partial \Omega}{\partial \tau_{stw}} = 0$ and $\frac{\partial \Omega}{\partial D} = 0$ implying that there are no inefficient separations $LS = 0$ and the eligibility condition does not cause any inefficiencies $L_{STW} = 0$. From this, we can infer that

$$\tilde{BE} = 0$$

Looking at the formula for the optimal STW benefits

$$\tau_{stw} = \frac{1}{\bar{h} - h_{stw}(\epsilon_s)} \cdot \left((1-f) \cdot FE - \left[n \cdot \frac{\partial \Omega}{\partial \tau_{stw}} \right] \Big/ \left[n \cdot g(\epsilon_s) \cdot \left(-\frac{\partial \epsilon_s}{\partial \tau_{stw}} \right) \right] - \tilde{BE} \right)$$
$$= 0$$

and given the fact that the fiscal externality on separations, the welfare cost penalty of STW and $\tilde{BE}$ are zero, we can infer that τ_{stw} is zero, verifying the initial guess.

Corollary 5. The optimal eligibility threshold is implicitly defined as:

$$S(\epsilon_{stw}) = y(\epsilon_{stw}, h(\epsilon_{stw})) - v\big(h(\epsilon_{stw})\big) + F + \frac{1 - \eta \cdot f}{1 - \eta} \cdot \frac{k_v}{q} = 0$$

Inserting the continuation value of the match gives:

$$\begin{aligned} & y(\epsilon_{stw}) - v\big(h(\epsilon_{stw})\big) + F \\ & + \beta \cdot (1 - \eta \cdot f) \\ & \cdot \frac{\int_{\epsilon_s}^{\infty} y(\epsilon) - v(h(\epsilon)) dG(\epsilon) - \Omega + G(\epsilon_s) \cdot F - \frac{1-n}{n} - \tilde{c}^w(b) + \frac{u(\tilde{c}^w(b)) - u(b)}{u'(\tilde{c}^w(b))}}{1 - \beta \cdot (1 - \eta \cdot f) \cdot (1 - G(\epsilon_s))} = 0 \end{aligned}$$

Note that also the consumption equivalent $\tilde{c}^w$ depends on the UI benefits b. The relationship between the consumption equivalents and the UI benefits is implicitly defined by:

$$\begin{aligned} \eta \cdot \tilde{c}^w + (1 - \eta) \cdot \frac{u(\tilde{c}^w)}{u'(\tilde{c}^w)} & \\ = \eta \cdot \left(\int_{\mathscr{B}} y(\epsilon) - v(h(\epsilon)) dG(\epsilon) - \Omega - \frac{1-n}{n} \cdot b + (1 - \rho) \cdot \theta \cdot k_v \right) & \\ + (1 - \eta) \cdot \frac{u(b)}{u'(\tilde{c}^w)} & \end{aligned}$$

Using the implicit function theorem, we can derive:

$$\frac{\partial \epsilon_{stw}}{\partial b} = \frac{\frac{\beta \cdot (1 - \eta \cdot f)}{1 - \beta \cdot (1 - \eta \cdot f) \cdot (1 - G(\epsilon_s))} \cdot \left(\frac{1-n}{n} \cdot b + \frac{u'(b)}{u'(\tilde{c}(b))} - c'(b) \cdot \left(-\frac{u''(\tilde{c}^w)}{u'(\tilde{c}^w)} \right) \cdot \frac{u(\tilde{c}^w) - u(b)}{u'(\tilde{c}^w)} \right)}{S'(\epsilon_{stw})}$$

By assumption, the denominator must be positive: $S'(\epsilon_{stw})$. An increase in productivity usually increases the joint surplus. In the particular case of a looser eligibility threshold, it can be that the extra distortionary effects dominate the increase in the joint surplus:

$$S'(\epsilon_{stw}) \;=\; a \cdot \epsilon_{stw} \cdot h(\epsilon_{stw})^{\alpha} + c_f - \frac{\beta \cdot (1-\eta \cdot f)}{1-\beta \cdot (1-\eta \cdot f) \cdot (1-G(\epsilon_s))} \cdot \frac{\partial \Omega}{\partial \epsilon_{stw}}$$

However, in a calibrated model, this seems to be unlikely. The nominator must be unambiguously positive. An increase in the UI benefits should always lead to a reduction in joint surplus and thus increase separation incentives, in this case, without access to the STW system:

$$\begin{aligned}
&\frac{1-n}{n} + \frac{u'(b)}{u'(\tilde{c}(b))} - c'(b) \cdot \left(-\frac{u''(\tilde{c}^w)}{u'(\tilde{c}^w)}\right) \cdot \frac{u(\tilde{c}^w) - u(b)}{u'(\tilde{c}^w)} \\
&= \frac{1-n}{n} + \frac{u'(b)}{u'(\tilde{c}(b))} - BE \cdot \left(-\eta \cdot \frac{1-n}{n} + (1-\eta) \cdot \frac{u'(b)}{u'(\tilde{c}(b))}\right) \\
&= (1+\eta \cdot BE) \cdot \frac{1-n}{n} + (1-(1-\eta) \cdot BE) \cdot \frac{u'(b)}{u'(\tilde{c}(b))} \;>\; 0
\end{aligned}$$

Note that we have seen in appendix 1.B that $\tilde{BE} < \frac{1}{1-\eta}$. Thus, the expression must be positive, proving that higher UI benefits will loosen the eligibility threshold and thus the minimum hours reduction condition:

$$\frac{\partial D}{\partial b} \;=\; \frac{D}{\epsilon_{stw}} \cdot \frac{\epsilon_{stw}}{\partial b} \;>\; 0$$

Next, we look at what happens when we offer more generous STW benefits. Suppose the government wants to impose the STW threshold ϵ_{stw}. Then we know that we need to set the following minimum hours reduction threshold:

$$D \;=\; h_{stw}(\epsilon_{stw})$$

When the government decides to reduce working, then we need to keep the eligibility condition stricter:

$$\frac{\partial D}{\partial \tau_{stw}} \;=\; \frac{\partial h_{stw}(\epsilon_{stw})}{\partial \tau_{stw}} \;<\; 0$$

We have already shown that in Lemma 1.

Corollary 6. The steady state budget constraint is:

$$n \cdot \tau_J \;=\; (1-n) \cdot b + n \cdot \int_{\epsilon_s}^{\epsilon_{stw}} \tau_{stw} \cdot \left(\bar{h} - h_{stw}(\epsilon)\right) \cdot dG(\epsilon)$$

Rearrange for the lump sum income tax:

$$\tau_J = \frac{(1-n)}{n} \cdot b + \int_{\epsilon_s}^{\epsilon_{stw}} \tau_{stw} \cdot \left(\bar{h} - h_{stw}(\epsilon)\right) \cdot dG(\epsilon)$$

Insert optimal eligibility condition and STW benefits:

$$\tau_J = \frac{(1-n)}{n} \cdot b + \int_{\epsilon_s}^{\xi_s} \frac{\bar{h} - h_{stw}(\epsilon)}{\bar{h} - h_{stw}(\epsilon_s)} \cdot \left((1-f) \cdot FE - \tilde{\Omega} - \tilde{BE}\right) dG(\epsilon)$$

with

$$\tilde{\Omega} = \left[n \cdot \frac{\partial \Omega}{\partial \tau_{stw}}\right] / \left[n \cdot g(\epsilon_s) \cdot \left(-\frac{\partial \epsilon_s}{\partial \tau_{stw}}\right)\right]$$

Set $\beta = 1$:

$$\tau_J = \frac{\left(1-f\right)}{f} \cdot G(\epsilon_s) \cdot b + \int_{\epsilon_s}^{\xi_s} \frac{\bar{h} - h_{stw}(\epsilon)}{\bar{h} - h_{stw}(\epsilon_s)} \cdot \left(\frac{1-f}{f} \cdot b - \tilde{\Omega} - \tilde{BE}\right) dG(\epsilon)$$

Note that

$$\frac{\bar{h} - h_{stw}(\epsilon)}{\bar{h} - h_{stw}(\epsilon_s)} < 1$$

Less hours reduction makes rescuing a job at risk less expensive than rescuing the marginal match. It must therefore apply:

$$\begin{aligned} \tau_J &< \frac{1-f}{f} \cdot G(\epsilon_s) \cdot b + \left(G(\xi_s) - G(\epsilon_s)\right) \cdot \left(\tilde{\Omega} + \tilde{BE}\right) \\ &= \underbrace{G(\xi_s) \cdot \frac{1-f}{f} \cdot b}_{\text{Costs of UI system without STW}} \quad \underbrace{-\left(G(\xi_s) - G(\epsilon_s)\right) \cdot \left(\tilde{\Omega} + \tilde{BE}\right)}_{\text{Costs saved with STW}} \end{aligned}$$

Ceteris paribus UI combined with optimal STW must be less expensive a UI system only.

Appendix 1.L Derivation Optimal STW Policy under Lump Sum Taxation

The argument for the minimum value condition for STW is equivalent to the one of the proof for Proposition 2 and is not repeated here. Under lump sum taxation and without risk aversion, the problem of the planner changes to:

$$\begin{aligned} W_t^P \; = \; & \max_{\theta_t, \epsilon_{s,t}, \epsilon_{stw,t}, h_t(\epsilon), \tau_{stw,t}} n_t \cdot \int_{\mathscr{B}_t}^{\infty} [y_t(\epsilon) - v\big(h_t(\epsilon)\big)] dG(\epsilon) - n_t \cdot \Omega_t \\ & - n_t \cdot \rho_t \cdot F - \theta_t \cdot (1 - n_t + \rho_t \cdot n_t) \cdot k_v + \beta \cdot E_t\big[W_{t+1}^P\big] \end{aligned}$$

Set:

$$\mathscr{B}_t \; = \; [\epsilon_{s,t}, \epsilon_{stw,t}] \cup [\xi_{s,t}, \infty)$$

Law of motion of employment:

$$n_{t+1} \; = \; (1 - \rho_t) \cdot n_t + f(\theta_t) \cdot (1 - n_t + \rho_t \cdot n_t)$$

Job-Creation Condition:

$$\begin{aligned} \frac{1}{1-\eta_t} \frac{k_v}{q(\theta_t)} \; = \; & \beta \cdot E_t\Big(\int_{\epsilon_{stw,t+1}}^{\infty} [y_{t+1}(\epsilon) - v\big(h_{t+1}(\epsilon)\big)] dG(\epsilon) \\ & + \int_{\epsilon_{s,t+1}}^{\epsilon_{stw,t+1}} [y_{stw,t+1}(\epsilon) - v\big(h_{stw,t+1}(\epsilon)\big) + \tau_{stw,t} \cdot \big(\bar{h} - h_{stw,t}(\epsilon)\big)] dG(\epsilon) \\ & - b_t + \rho_t \cdot (-F) + (1 - \rho_t) \cdot \frac{1 - \eta_{t+1} \cdot \theta_{t+1} \cdot q(\theta_{t+1})}{1 - \eta_{t+1}} \cdot \frac{k_v}{q(\theta_{t+1})} \Big) \end{aligned}$$

Separation Condition with STW:

$$\begin{aligned} & y_{stw,t}(\epsilon_{s,t}) - v\big(h_{stw,t}(\epsilon_{s,t})\big) + \tau_{stw,t} \cdot \big(\bar{h} - h_{stw,t}(\epsilon_{s,t})\big) \\ & + F + \frac{1 - \eta_t \cdot \theta_t \cdot q(\theta_t)}{1 - \eta_t} \cdot \frac{k_v}{q(\theta_t)} \; = \; 0 \end{aligned} \tag{1.L.1}$$

Separation Condition without access to STW:

$$y_t(\xi_{s,t}) - v\big(h_t(\xi_{s,t})\big) + F + \frac{1 - \eta_t \cdot \theta_t \cdot q(\theta_t)}{1 - \eta_t} \cdot \frac{k_v}{q(\theta_t)} \; = \; 0$$

Welfare costs of STW:

$$\Omega_t \; = \; \int_{\epsilon_{s,t}}^{\epsilon_{stw,t}} [y_t(\epsilon, h_t(\epsilon)) - v\big(h_t(\epsilon)\big) - y_t(\epsilon, h_{stw,t}(\epsilon)) + v\big(h_{stw,t}(\epsilon)\big)] dG(\epsilon)$$

Separation rate:

$$\rho_t \; = \; [\max\{G(\xi_{s,t}) - G(\epsilon_{stw,t}), 0\} + G(\epsilon_{s,t})]$$

For the same argument as in the main model, we make the assumption that the eligibility threshold must be at least as generous as the separation threshold of a firm without access to STW:

$$\epsilon_{stw,t} \geq \xi_{s,t}$$

1.L.1 Optimal STW benefits Steady State

First, we derive the optimality condition for the STW benefits:

$$\begin{aligned}
\frac{\partial}{\partial \tau_{stw,t}} &= -n_t \cdot \frac{\partial \Omega_t}{\partial \tau_{stw,t}} + \lambda_{\theta,t-1} \cdot \int_{\epsilon_{s,t}}^{\epsilon_{stw,t}} \left(\bar{h} - h_{stw,t}(\epsilon)\right) dG(\epsilon) \\
&\quad - \lambda_{\epsilon_s,t} \cdot \left(\bar{h} - h_{stw,t}(\epsilon_{s,t})\right) = 0 \\
\Leftrightarrow \lambda_{\epsilon_s,t} &= \frac{1}{\bar{h} - h_{stw,t}(\epsilon_{s,t})} \cdot \left(-n_t \cdot \frac{\partial \Omega_t}{\partial \tau_{stw,t}} + \lambda_{\theta,t-1} \cdot \int_{\epsilon_{s,t}}^{\epsilon_{stw,t}} \left(\bar{h} - h_{stw,t}(\epsilon)\right) dG(\epsilon)\right)
\end{aligned}$$

The optimality condition for employment can be written as:

$$\begin{aligned}
\frac{\partial}{\partial n_t} &= \int_{\mathscr{B}_t}^{\infty} \left[y_t(\epsilon) - v(h_t(\epsilon))\right] - \Omega_t - \rho_t \cdot F + (1-\rho_t) \cdot \theta_t \cdot k_v \\
&\quad - \lambda_{n,t} + \beta \cdot (1 - \theta_t \cdot q(\theta_t)) \cdot (1-\rho_t) \cdot E_t\left[\lambda_{n,t+1}\right] = 0
\end{aligned}$$

Correspondingly, the FOC for the labor market tightness is:

$$\begin{aligned}
\frac{\partial}{\partial \theta_t} &= -(1-n_t) \cdot k_v + \beta \cdot (1-n_t) \cdot (1-\gamma) \cdot q(\theta_t) \cdot E_t\left[\lambda_{n,t+1}\right] \\
&\quad - \frac{1}{1-\eta_t} \cdot \frac{k_v}{q(\theta_t)} \cdot \gamma \cdot \frac{\lambda_{\theta,t}}{\theta_t} \\
&\quad + \frac{1}{1-\eta_t} \cdot \left(\frac{\gamma}{\theta_t \cdot q(\theta_t)} - \eta_t\right) \cdot k_v \cdot \left((1-\rho_t) \cdot \lambda_{\theta,t-1} - \lambda_{\epsilon_s,t} - \lambda_{cs_1,t}\right) = 0
\end{aligned}$$

Rewriting the optimality condition for the labor market tightness gives:

$$\beta \cdot E_t\left[\lambda_{n,t+1}\right] = \left(\frac{1+\chi_t}{1-\gamma}\right) \cdot \frac{k_v}{q(\theta_t)}$$

With

$$\begin{aligned}
\chi_t &= \frac{1}{1-n_t} \cdot \frac{1}{1-\eta_t} \cdot \frac{1}{\theta_t \cdot q(\theta_t)} \cdot \Big(\gamma \cdot \lambda_{\theta,t} - \left(\gamma - \theta_t \cdot q(\theta_t) \cdot \eta_t\right) \\
&\quad \cdot \left(\left(1 - G(\epsilon_{s,t})\right) \cdot \lambda_{\theta,t-1} - \lambda_{\epsilon_s,t} - \lambda_{cs_1,t}\right)\Big)
\end{aligned}$$

Finally, we can derive the FOC for the separation threshold:

$$
\begin{aligned}
\frac{\partial}{\partial \epsilon_{s,t}} =& - \left(y_{stw,t}(\epsilon_{s,t}) - v(h_{stw,t}(\epsilon_{s,t})) - F\right) \cdot g(\epsilon_{s,t}) \cdot n_t \\
& - \lambda_{n,t} \cdot (1 - f(\theta_t)) \cdot g(\epsilon_{s,t}) \cdot n_t \\
& - \lambda_{\theta,t-1} \cdot \Bigg(y_{stw,t}(\epsilon_{s,t}) - v(h_{stw,t}(\epsilon_{s,t})) + (\bar{h} - h_{stw,t}(\epsilon_{s,t})) \cdot \tau_{stw,t} \\
& \quad + F + \frac{1 - \theta_t \cdot q(\theta_t) \cdot \eta_t}{1 - \eta_t} \cdot \frac{k_v}{q(\theta_t)} \Bigg) \cdot g(\epsilon_{s,t}) \\
& - \lambda_{\epsilon,t} \cdot \left[a_t \cdot h_{stw,t}(\epsilon_{s,t})^{\alpha} + c_f \right] \\
=& - y_{stw,t}(\epsilon_{s,t}) + v(h_{stw,t}(\epsilon_{s,t})) - F \\
& - \lambda_{n,t} \\
& - \lambda_{\epsilon,t} \cdot \frac{a_t \cdot h_{stw,t}(\epsilon_{s,t})^{\alpha} + c_f}{n_t \cdot g(\epsilon_{s,t})}
\end{aligned}
$$

Using the optimality condition for employment and the labor market density, we can derive the optimal vacancy posting condition:

$$
\begin{aligned}
\frac{1}{1-\gamma} \cdot \frac{\left(1 + \chi_t\right) \cdot k_v}{q_t} =& \beta \cdot E_t \left[\int_{\mathscr{B}_t}^{\infty} \left(y_t(\epsilon) - v(h_t(\epsilon))\right) dG(\epsilon) - \Omega_t + \rho_t \cdot (-F) \right] \\
& + \beta \cdot E_t \left[(1 - \rho_t) \cdot \frac{1 - \gamma \cdot f_{t+1} + \left(1 - f_{t+1}\right) \cdot \chi_{t+1}}{1 - \gamma} \cdot \frac{k_v}{q_{t+1}} \right]
\end{aligned}
$$

Likewise, we can derive the optimal separation decision by plugging the optimality condition for employment and the labor market density into the optimality condition for the separation threshold.

$$
\begin{aligned}
y_{stw,t}(\epsilon_{s,t}) - v(h_{stw,t}(\epsilon_{s,t})) + \lambda_{\epsilon_s,t} \cdot \left(a_t \cdot h_{stw,t}(\epsilon_{s,t})^{\alpha} + c_f\right) + F \\
+ \frac{1 - \gamma \cdot f_t + \left(1 - f_t\right) \cdot \chi_t}{1 - \gamma} \cdot \frac{k_v}{q_t} = 0
\end{aligned}
\tag{1.L.2}
$$

Next, we deduct the decentralized job-creation condition from the optimal vacancy posting condition:

$$
\begin{aligned}
\left(\chi_t - \frac{\eta_t - \gamma}{1 - \eta_t}\right) \cdot \frac{1}{1-\gamma} \cdot \frac{k_v}{q_t} =& \; b_t - \int_{\epsilon_{s,t}}^{\epsilon_{stw,t}} (\bar{h} - h_{stw,t}(\epsilon)) dG(\epsilon) \cdot \tau_{stw,t} \\
+ \beta \cdot E_t \left[\left(1 - G(\epsilon_{s,t+1})\right) \cdot (1 - f_{t+1}) \cdot \left(\chi_{t+1} - \frac{\eta_{t+1} - \gamma}{1 - \eta_{t+1}}\right) \cdot \frac{1}{1-\gamma} \cdot \frac{k_v}{q_{t+1}} \right]
\end{aligned}
$$

Likewise, we deduct the decentralized separation condition from the optimal separation condition and rearrange it for the STW subsidy:

$$\tau_{stw,t} \cdot \left(\bar{h} - h_{stw,t}(\epsilon)\right) = \lambda_{\epsilon,t} \cdot \left(a_t \cdot h_{stw,t}(\epsilon_{s,t})^{\alpha} + c_f\right) \cdot \frac{1 - G(\epsilon_{s,t})}{n_t \cdot g(\epsilon_{s,t})} + (1 - f_t) \cdot \left(\chi_t - \frac{\eta_t - \gamma}{1 - \eta_t}\right) \cdot \frac{1}{1 - \gamma} \cdot \frac{k_v}{q_t} \tag{1.L.3}$$

Plugging equation 1.L.2 into equation 1.L.3 and iterating forward gives:

$$\tau_{stw,t} \cdot \left(\bar{h} - h_{stw,t}(\epsilon_{s,t})\right) = \lambda_{\epsilon,t} \cdot \left(a_t \cdot h_{stw,t}(\epsilon_{s,t})^{\alpha} + c_f\right) \cdot \frac{1 - G(\epsilon_{s,t})}{n_t \cdot g(\epsilon_{s,t})} + \sum_{j=1}^{\infty} \left[\prod_{i=1}^{j} \beta \cdot (1 - f_{t+i-1}) \cdot (1 - G(\epsilon_{s,t+i}))\right] \cdot \left[b_{t+j} - \int_{\epsilon_{s,t}}^{\epsilon_{stw,t}} \left(\bar{h} - h_{stw,t}(\epsilon)\right) \cdot \tau_{stw,t}\right]$$

Inserting the Lagrange multiplier $\lambda_{\epsilon,t}$ and using the implicit function theorem so that:

$$\frac{a_t \cdot h_{stw,t}(\epsilon_{s,t})^{\alpha} + c_f}{\bar{h} - h_{stw,t}(\epsilon)} = \frac{\frac{\partial S_t(\epsilon_{stw,t})}{\partial \epsilon_{s,t}}}{\frac{\partial S_{stw,t}(\epsilon_{s,t})}{\tau_{stw,t}}} = -\frac{1}{\frac{\partial \epsilon_{s,t}}{\partial \tau_{stw,t}}} > 0$$

gives:

$$\begin{aligned} \tau_{stw,t} \cdot \left(\bar{h} - h_{stw,t}(\epsilon_{s,t})\right) = & \\ -\Bigg[\frac{n_t}{1 - G(\epsilon_{s,t})} \cdot \frac{\partial \Omega_t}{\partial \tau_{stw,t}} + \lambda_{\theta,t-1} \cdot \int_{\epsilon_{s,t}}^{\epsilon_{stw,t}} \left(\bar{h} - h_{stw,t}(\epsilon)\right) dG(\epsilon) \Bigg] & \\ \cdot \left(-\frac{1}{\frac{\partial \epsilon_{s,t}}{\partial \tau_{stw,t}}}\right) \cdot \frac{1 - G(\epsilon_{s,t})}{g(\epsilon_{s,t}) \cdot n_t} & \\ + \sum_{j=1}^{\infty} \left[\prod_{i=1}^{j} \beta \cdot (1 - f_{t+i-1}) \cdot (1 - G(\epsilon_{s,t+i}))\right] & \\ \cdot \left[b_{t+j} - \int_{\epsilon_{s,t+j}}^{\epsilon_{stw,t+j}} \left(\bar{h} - h_{stw,t+j}(\epsilon)\right) \cdot \tau_{stw,t+j} \, d\epsilon \right] & \end{aligned}$$

Assume that the system has converged to its non-stochastic steady state, then the STW subsidy can be expressed as:

$$\tau_{stw} \cdot (\bar{h} - h_{stw}(\epsilon_s)) =$$
$$\underbrace{\frac{\beta \cdot \left(1-f\right)}{1-\beta \cdot \left(1-G(\epsilon_s)\right) \cdot (1-f)} \cdot \left(b - \int_{\epsilon_s}^{\epsilon_{stw}} (\bar{h} - h_{stw}(\epsilon))dG(\epsilon) \cdot \tau_{stw}\right)}_{\text{FE on Separations}}$$
$$\underbrace{- \frac{1}{g(\epsilon_s)} \cdot \frac{\partial \Omega}{\partial \tau_{stw}} / \frac{\partial \epsilon_s}{\partial \tau_{stw}}}_{\text{Utility cost of using STW}}$$
$$\underbrace{+ \frac{\tilde{\lambda}_\theta}{g(\epsilon_s)} \cdot \int_{\epsilon_s}^{\epsilon_{stw}} \left(\bar{h} - h_{stw}(\epsilon)\right) dG(\epsilon) / \frac{\partial \epsilon_s}{\partial \tau_{stw}}}_{\text{STW increases depressed vacancy posting}}$$

This completes the proof of Proposition 6.

1.L.2 Optimal Eligibility Condition

First, let us form the FOC of the eligibility threshold:

$$\frac{\partial}{\partial \epsilon_{stw,t}} = -n_t \cdot \frac{\partial \Omega_t}{\partial \epsilon_{stw,t}} + \lambda_{\theta,t-1} \cdot \left[J_{stw}(\epsilon_{stw}) - J(\epsilon_{stw})\right] \cdot g(\epsilon_{stw,t}) - \lambda_{cs_1,t} = 0$$

According to the Kuhn-Tucker conditions the Lagrange multiplier must be zero $\lambda_{cs_1,t} = 0$ if the condition is not binding $\epsilon_{stw,t} > \xi_{s,t}$ or negative $\lambda_{cs_1,t} < 0$ if the condition is binding $\epsilon_{stw,t} = \xi_{s,t}$:

$$\lambda_{cs_1,t} = \begin{cases} \lambda_{\theta,t-1} \cdot \left[J_{stw,t}(\epsilon_{stw,t}) - J_t(\epsilon_{stw,t})\right] - \frac{\partial \Omega_t}{\partial \epsilon_{stw,t}} < 0, & \text{for} \quad \epsilon_{stw,t} = \xi_{s,t} \\ 0, & \text{for} \quad \epsilon_{stw,t} > \xi_{s,t} \end{cases}$$

If the condition is binding, we get the same eligibility condition as in the model with lump sum income taxes. If it is not, then we get a trade-off between hours distortions and increasing joint surplus of firms and workers to stimulate vacancy posting and reduce separations. To see this in the formula, we need to derive the Lagrange multiplier for the job-creation condition λ_θ. We follow the procedure from Appendix 1.I.

The **deviation from optimal job-creation condition** in steady state can be denoted as:

$$\left(\chi - \frac{\eta-\gamma}{1-\eta}\right) \cdot \frac{1}{1-\gamma} \cdot \frac{k_v}{q} = b - \int_{\epsilon_s}^{\epsilon_{stw}} (\bar{h} - h_{stw}(\epsilon))dG(\epsilon) \cdot \tau_{stw} \qquad (1.L.4)$$
$$+ \beta \cdot \left(1 - G(\epsilon_s)\right) \cdot (1-f) \cdot \left(\chi - \frac{\eta-\gamma}{1-\eta}\right) \cdot \frac{1}{1-\gamma} \cdot \frac{k_v}{q}$$

Rearranging for χ gives:

$$\chi = \frac{\eta-\gamma}{1-\eta} + \frac{1}{1-\beta\cdot(1-\rho)\cdot(1-f)}\cdot\left(b - \int_{\epsilon_s}^{\epsilon_{stw}} (\bar{h} - h_{stw}(\epsilon))dG(\epsilon)\cdot\tau_{stw}\right) \times \frac{1}{1-\gamma}\cdot\frac{k_v}{q}$$

From the optimality condition of the labor market tightness, we can derive a second expression for χ :

$$\chi_t = \frac{1}{1-n_t}\cdot\frac{1}{1-\eta_t}\cdot\frac{1}{\theta_t\cdot q(\theta_t)} \cdot\left(\gamma\cdot\lambda_{\theta,t} - \left(\gamma-\theta_t\cdot q(\theta_t)\cdot\eta_t\right)\cdot\left(\left(1-G(\epsilon_{s,t})\right)\cdot\lambda_{\theta,t-1} - \lambda_{\epsilon_s,t} - \lambda_{\xi_s,t}\right)\right)$$

Since we want to isolate $\lambda_{\theta,t}$ we need to find expressions for the Lagrange multipliers of $\lambda_{\epsilon_s,t}$ and $\lambda_{\xi_{s,t}}$. The Lagrange multiplier for the separation condition of firms with access to STW can be denoted by:

$$-\lambda_{\epsilon_s,t} = \frac{1}{\bar{h}-h_{stw,t}(\epsilon_{s,t})}\cdot\left(n_t\cdot\frac{\partial\Omega_t}{\partial\tau_{stw,t}} - \lambda_{\theta,t-1}\cdot\int_{\epsilon_{s,t}}^{\epsilon_{stw,t}}\left(\bar{h}-h_{stw,t}(\epsilon)\right)dG(\epsilon)\right)$$

Next, we can form the FOC for the separation threshold without access to STW benefits:

$$\frac{\partial}{\partial\xi_{s,t}} = \lambda_{cs_1,t} - \lambda_{\xi_s,t}\cdot\left(a_t\cdot h_{stw,t}(\xi_{s,t})^\alpha + c_f\right) = 0$$
$$\Leftrightarrow \quad \lambda_{cs_1,t} = \lambda_{\xi_s,t}\cdot\left(a_t\cdot h_{stw,t}(\xi_{s,t})^\alpha + c_f\right)$$

This allows us to calculate the Lagrange multiplier for the separation threshold of firms without access to STW. Note that we have a Kuhn-Tucker problem, so that $\lambda_{cs_1,t}$ and this $\lambda_{\xi_s,t}$ require a case distinction:

$$\lambda_{\xi_s,t} = \begin{cases} \frac{\lambda_{\theta,t-1}\cdot\left[J_{stw,t}(\epsilon_{stw,t})-J_t(\epsilon_{stw,t})\right]-\frac{\partial\Omega_t}{\partial\epsilon_{stw,t}}}{\left(a_t\cdot h_{stw,t}(\epsilon_{stw,t})^\alpha+c_f\right)}, & \text{for} \quad \epsilon_{stw,t} = \xi_{s,t} \\ 0, & \text{for} \quad \epsilon_{stw,t} > \xi_{s,t} \end{cases}$$

This can be rewritten using the indicator function:

$$\lambda_{\xi_s,t} = \frac{\mathbb{1}\{\epsilon_{stw,t} = \xi_{s,t}\}}{\left(a_t\cdot h_{stw,t}(\epsilon_{stw,t})^\alpha + c_f\right)}\cdot\left(\lambda_{\theta,t-1}\cdot\left[J_{stw,t}(\epsilon_{stw,t}) - J_t(\epsilon_{stw,t})\right] - n_t\cdot\frac{\partial\Omega_t}{\partial\epsilon_{stw,t}}\right)$$

In a next step, we can set both expressions for χ equal and insert our new expressions for the Lagrange multipliers:

$$\begin{aligned}
&\frac{1}{1-\eta} \cdot \frac{\gamma - \beta \cdot \big(\gamma - \eta \cdot f(\theta)\big) \cdot (1 - G(\epsilon_s))}{f(\theta) \cdot u'(\tilde{c}^w)} \cdot k_v \cdot \lambda_\theta \\
&= (1-\gamma) \cdot u \cdot q(\theta) \cdot \left(\frac{\eta - \gamma}{(1-\gamma) \cdot (1-\eta)} \cdot \frac{k_v}{q(\theta)} \right. \\
&\quad + \frac{\beta}{1 - \beta \cdot (1-\rho) \cdot (1-f)} \cdot \left. \left(b - \int_{\epsilon_s}^{\epsilon_{stw}} (\bar{h} - h_{stw}(\epsilon)) dG(\epsilon) \cdot \tau_{stw} \right) \right) \\
&\quad + \frac{1}{1-\eta} \cdot \frac{(\gamma - \eta \cdot f(\theta)) \cdot k_v}{f(\theta) \cdot u'(\tilde{c}^w)} \cdot \frac{1}{\bar{h} - h_{stw}(\epsilon_s)} \\
&\quad \cdot \left(\lambda_\theta \cdot \int_{\epsilon_s}^{\epsilon_{stw}} \big(\bar{h} - h_{stw}(\epsilon)\big) dG(\epsilon) - n \cdot \frac{\partial \Omega}{\partial \tau_{stw}} \right) \\
&\quad + \frac{1}{1-\eta} \cdot \frac{(\gamma - \eta \cdot f(\theta)) \cdot k_v}{f(\theta) \cdot u'(\tilde{c}^w)} \cdot \frac{\mathbb{1}\{\epsilon_{stw} = \xi_s\}}{a \cdot h_{stw}(\epsilon_{stw})^\alpha + c_f} \\
&\quad \cdot \left(\lambda_\theta \cdot [J_{stw}(\epsilon_{stw}) - J(\epsilon_{stw})] - n \cdot \frac{\partial \Omega}{\partial \epsilon_{stw}} \right).
\end{aligned}$$

Using the implicit function theorem, we can denote:

$$\begin{aligned}
\frac{\partial \tau_{stw}}{\partial \theta} &= -\frac{1}{1-\eta} \cdot \frac{(\gamma - \eta \cdot f(\theta))}{f(\theta) \cdot u'(\tilde{c}^w)} \cdot k_v \cdot \frac{1}{\bar{h} - h_{stw}(\epsilon_s)} \\
\frac{\partial \epsilon_{stw}}{\partial \theta} &= -\frac{1}{1-\eta} \cdot \frac{(\gamma - \eta \cdot f(\theta))}{f(\theta) \cdot u'(\tilde{c}^w)} \cdot k_v \cdot \frac{1}{\big(a \cdot h_{stw}(\epsilon_{stw})^\alpha + c_f\big)}
\end{aligned}$$

Further, note that

$$(1-\gamma) \cdot q(\theta) = f'(\theta)$$

This allows us to solve for λ_θ:

$$\begin{aligned}
\lambda_\theta &= \frac{f'(\theta) \cdot u}{M} \cdot \left(\frac{\eta - \gamma}{(1-\gamma) \cdot (1-\eta)} \cdot \frac{k_v}{q(\theta)} \right. \\
&\quad + \frac{\beta}{1 - \beta \cdot (1-\rho) \cdot (1-f)} \cdot \left. \left(b - \int_{\epsilon_s}^{\epsilon_{stw}} (\bar{h} - h_{stw}(\epsilon))\, dG(\epsilon) \cdot \tau_{stw} \right) \right) \\
&\quad + \frac{n}{M} \cdot \frac{\partial \tau_{stw}}{\partial \theta} \cdot \frac{\partial \Omega}{\partial \tau_{stw}} \\
&\quad + \mathbb{1}\{\epsilon_{stw} = \xi_s\} \cdot \frac{n}{M} \cdot \frac{\partial \epsilon_{stw}}{\partial \theta} \cdot \frac{\partial \Omega}{\partial \epsilon_{stw}}.
\end{aligned}$$

Here, M denotes the inverse multiplier:

$$
\begin{aligned}
M \;=\; & \frac{1}{1-\eta} \cdot \frac{\gamma - \beta \cdot \big(\gamma - \eta \cdot f(\theta)\big) \cdot (1 - G(\epsilon_s))}{f(\theta) \cdot u'(\tilde{c}^w)} \cdot k_v \\
& + \frac{\partial \tau_{stw}}{\partial \theta} \cdot \int_{\epsilon_s}^{\epsilon_{stw}} \big(\bar{h} - h_{stw}(\epsilon)\big)\, dG(\epsilon) \\
& + \frac{\partial \epsilon_{stw}}{\partial \theta} \cdot [J_{stw}(\epsilon_{stw}) - J(\epsilon_{stw})]
\end{aligned}
$$

This is nothing else but the general equilibrium effect, that an increase in the value of the firm has on the labor market tightness:

$$
\frac{\partial \theta^{ge}}{\partial J} \;=\; \frac{1}{M}
$$

Inserting this simplifies the expression to:

$$
\begin{aligned}
\lambda_\theta \;=\; & \frac{\partial \theta^{ge}}{\partial J} \cdot f'(\theta) \cdot u \cdot \Bigg(\frac{\eta - \gamma}{(1-\gamma) \cdot (1-\eta)} \cdot \frac{k_v}{q(\theta)} \\
& + \frac{\beta}{1 - \beta \cdot (1-\rho) \cdot (1-f)} \cdot \left(b - \int_{\epsilon_s}^{\epsilon_{stw}} (\bar{h} - h_{stw}(\epsilon))\, dG(\epsilon) \cdot \tau_{stw} \right) \Bigg) \\
& - \frac{\partial \theta^{ge}}{\partial J} \cdot \Bigg[\frac{\partial \tau_{stw}}{\partial \theta} \cdot n \cdot \frac{\partial \Omega}{\partial \tau_{stw}} \\
& \quad + \mathbb{1}\{\epsilon_{stw} \;=\; \xi_s\} \cdot \frac{\partial \epsilon_{stw}}{\partial \theta} \cdot n \cdot \frac{\partial \Omega}{\partial \epsilon_{stw}} \Bigg].
\end{aligned}
$$

Note that

$$
\begin{aligned}
\frac{\partial \theta^{ge}}{\partial J} \cdot \frac{\partial \tau_{stw}}{\partial \theta} \;=\; & \frac{\partial \theta^{ge}}{\partial J} \cdot \frac{\partial \epsilon_s}{\partial \theta} \cdot \frac{n \cdot g(\epsilon_s)}{n \cdot g(\epsilon_s) \cdot \left(-\frac{\partial \epsilon_s}{\partial \tau_{stw}}\right))} \\
\;=\; & \frac{n \cdot g(\epsilon_s) \cdot \frac{\partial \epsilon_s^{ge}}{\partial J}}{n \cdot g(\epsilon_s) \cdot \left(-\frac{\partial \epsilon_s}{\partial \tau_{stw}}\right))}.
\end{aligned}
$$

and that

$$
\begin{aligned}
\frac{\partial \theta^{ge}}{\partial J} \cdot \frac{\partial \epsilon_{stw}}{\partial \theta} \;=\; & \frac{\partial \theta^{ge}}{\partial J} \cdot \frac{\partial \epsilon_{stw}}{\partial \theta} \cdot \frac{n \cdot g(\epsilon_{stw})}{n \cdot g(\epsilon_{stw})} \\
\;=\; & \frac{n \cdot g(\epsilon_{stw}) \cdot \frac{\partial \epsilon_s^{ge}}{\partial J}}{n \cdot g(\epsilon_{stw})}.
\end{aligned}
$$

This changes the expression to:

$$\begin{aligned}
\lambda_\theta \;=\; & \frac{\partial f^{ge}}{\partial J} \cdot u \cdot \Bigg(\frac{\eta-\gamma}{(1-\gamma)\cdot(1-\eta)} \cdot \frac{k_v}{q(\theta)} \\
& + \frac{\beta}{1-\beta\cdot(1-\rho)\cdot(1-f)} \cdot \left(b - \int_{\epsilon_s}^{\epsilon_{stw}} (\bar{h} - h_{stw}(\epsilon)) dG(\epsilon) \cdot \tau_{stw} \right) \Bigg) \\
& - \left[n \cdot g(\epsilon_s) \cdot \frac{\partial \epsilon_s^{ge}}{\partial J} \right] \cdot \frac{n \cdot \frac{\partial \Omega}{\partial \tau_{stw}}}{n \cdot g(\epsilon_s) \cdot \left(-\frac{\partial \epsilon_s}{\partial \tau_{stw}} \right))} \\
& - \mathbb{1}\{\epsilon_{stw} \;=\; \xi_s\} \cdot \left[n \cdot g(\epsilon_{stw}) \cdot \frac{\partial \epsilon_{stw}^{ge}}{\partial J} \right] \cdot \frac{n \cdot \frac{\partial \Omega}{\partial \epsilon_{stw}}}{n \cdot g(\epsilon_{stw})}.
\end{aligned}$$

Define:

$$\begin{aligned}
L_V \;=\; & \frac{\eta-\gamma}{(1-\gamma)\cdot(1-\eta)} \cdot \frac{k_v}{q(\theta)} \\
& + \frac{\beta}{1-\beta\cdot(1-\rho)\cdot(1-f)} \cdot \left(b - \int_{\epsilon_s}^{\epsilon_{stw}} (\bar{h} - h_{stw}(\epsilon)) dG(\epsilon) \cdot \tau_{stw} \right), \\
L_S^\star \;=\; & \frac{n \cdot \frac{\partial \Omega}{\partial \tau_{stw}}}{n \cdot g(\epsilon_s) \cdot \left(-\frac{\partial \epsilon_s}{\partial \tau_{stw}} \right))}, \\
L_{STW}^\star \;=\; & \frac{n \cdot \frac{\partial \Omega}{\partial \epsilon_{stw}}}{n \cdot g(\epsilon_{stw})}.
\end{aligned}$$

This simplifies the expression for the Lagrange multiplier of the job-creation condition to:

$$\begin{aligned}
\lambda_\theta \;=\; & \frac{\partial f^{ge}}{\partial J} \cdot u \cdot L_V - \left[n \cdot g(\epsilon_s) \cdot \frac{\partial \epsilon_s^{ge}}{\partial J} \right] \cdot L_S^\star \\
& - \mathbb{1}\{\epsilon_{stw} \;=\; \xi_s\} \cdot \left[n \cdot g(\epsilon_{stw}) \cdot \frac{\partial \epsilon_{stw}^{ge}}{\partial J} \right] \cdot L_{STW}^\star
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \Omega}{\partial \epsilon_{stw}} \;>\; & \Bigg(\frac{\partial f^{ge}}{\partial J} \cdot u \cdot L_V - \left[n \cdot g(\epsilon_s) \cdot \frac{\partial \epsilon_s^{ge}}{\partial J} \right] \cdot L_S^\star \\
& - \mathbb{1}\{\epsilon_{stw} \;=\; \xi_s\} \cdot \left[n \cdot g(\epsilon_{stw}) \cdot \frac{\partial \epsilon_{stw}^{ge}}{\partial J} \right] \cdot L_{STW}^\star \Bigg) \cdot [J_{stw}(\epsilon_{stw}) - J(\epsilon_{stw})].
\end{aligned}$$

Simplifies to:

$$\frac{\partial \Omega}{\partial \epsilon_{stw}} > \frac{\partial f^{ge}}{\partial \epsilon_{stw}} \cdot u \cdot L_V + \left[n \cdot g(\epsilon_s) \cdot \left(-\frac{\partial \epsilon_s^{ge}}{\partial \epsilon_{stw}} \right) \right] \cdot L_S^{\star}$$
$$+ \mathbb{1}\{\epsilon_{stw} = \xi_s\} \cdot \left[n \cdot g(\epsilon_{stw}) \cdot \left(-\frac{\partial \epsilon_{stw}^{ge}}{\partial \epsilon_{stw}} \right) \right] \cdot L_{STW}^{\star}$$

Note that if $\epsilon_{stw} > \epsilon_s$ and

$$\frac{\partial \Omega}{\partial \epsilon_{stw}} > \frac{\partial f^{ge}}{\partial \epsilon_{stw}} \cdot u \cdot L_V + \left[n \cdot g(\epsilon_s) \cdot \left(-\frac{\partial \epsilon_s^{ge}}{\partial \epsilon_{stw}} \right) \right] \cdot L_S^{\star}$$

then $\epsilon_{stw} > \epsilon_s$ cannot be optimal, and the social planner would want to choose $\epsilon_{stw} = \epsilon_s$. Otherwise, the planner can set ϵ_{stw} so that

$$\frac{\partial \Omega}{\partial \epsilon_{stw}} = \frac{\partial f^{ge}}{\partial \epsilon_{stw}} \cdot u \cdot L_V + \left[n \cdot g(\epsilon_s) \cdot \left(-\frac{\partial \epsilon_s^{ge}}{\partial \epsilon_{stw}} \right) \right] \cdot L_S^{\star}$$

This is the condition from Proposition 3.

Appendix 1.M Derivation Optimal Layoff Taxes

In this exercise, we swap the STW system with a layoff tax. The Ramsey planner chooses the layoff tax τ_S subject to the labor market equilibrium.

$$\begin{aligned} W_t^G = \max_{\tau_{S,t}} \Big\{ & (1-n_t) \cdot u(b_t) + n_t \cdot u(\tilde{c}_t^w) \\ & + \nu_t^f \cdot u\Big(\Big[n_t \cdot \int_{\epsilon_{s,t}}^{\epsilon_{stw,t}} \big(y_t(\epsilon) - v(h_t(\epsilon)) \big) dG(\epsilon) \\ & \quad - n_t \cdot \tilde{c}_t^w - (1-n_t) \cdot b_t - \theta_t \cdot \big(1 - n_t + \rho_t \cdot n_t \big) \cdot k_v \Big] \Big/ \nu_t^f \Big) \\ & + \beta \cdot E_t W_{t+1}^G \Big\} \end{aligned}$$

Job-Creation Condition:

$$\begin{aligned} \frac{1}{1-\eta_t} \cdot \frac{k_v}{q(\theta_t)} = E_t \Big[Q_{t,t+1}^W \Big(& \int_{\epsilon_{s,t}}^{\epsilon_{stw,t}} y_{t+1}(\epsilon) - v(h_{t+1}(\epsilon)) dG(\epsilon) - \frac{1-n_{t+1}}{n_{t+1}} \cdot b_{t+1} \\ & - \tilde{c}_{t+1} + \frac{u(\tilde{c}_{t+1}) - u(b_{t+1})}{u'(\tilde{c}_{t+1})} \Big) \Big] \\ & + E_t \left[Q_{t,t+1}^W \left((1-\rho_{t+1}) \cdot \frac{1-\eta_{t+1} \cdot f(\theta_{t+1})}{1-\eta_{t+1}} \cdot \frac{k_v}{q(\theta_{t+1})} \right) \right] \end{aligned}$$

Income Insurance contract:

$$\eta_{t-1} \cdot \tilde{c}_t^w + (1-\eta_{t-1}) \cdot \frac{u(\tilde{c}_t^w)}{u'(\tilde{c}_t^w)}$$
$$= \eta_{t-1} \cdot \left(\int_{\epsilon_{s,t}}^{\epsilon_{stw,t}} \big(y_t(\epsilon) - v(h_t(\epsilon))\big) dG(\epsilon) - \frac{1-n_t}{n_t} \cdot b_t + (1-\rho_t) \cdot \theta_t \cdot k_v \right)$$
$$+ (1-\eta_{t-1}) \cdot \frac{u(b_t)}{u'(\tilde{c}_t^w)}$$
$$- \frac{\eta_{t-1} - \eta_t}{1-\eta_t} \cdot (1-\rho_t) \cdot (1-f_t) \cdot \frac{k_v}{q_t}$$

Separation conditions on STW:

$$y_t(\epsilon_{s,t}, h_{stw,t}(\epsilon_{s,t})) - v(h_{stw,t}(\epsilon_{s,t})) + \tau_{S,t} + F + \frac{1-\eta_t \cdot f_t}{1-\eta_t} \cdot \frac{k_v}{q_t} = 0$$

Working hours:

$$\alpha \cdot a_t \cdot \epsilon \cdot h_t(\epsilon)^{\alpha-1} = h_t(\epsilon)^{\psi}$$

Stochastic discount factor:

$$Q_{t,t+1}^W = \beta \cdot \frac{u'(\tilde{c}_{t+1}^w)}{u'(\tilde{c}_t^w)}$$

Law of motion of employment:

$$n_{t+1} = (1-\rho_t) \cdot n_t + f(\theta_t) \cdot (1 - n_t + \rho_t \cdot n_t)$$

Similar to Appendix 1.I, we can derive the **constrained efficient job-creation condition:**

$$\Leftrightarrow \frac{1+\chi_{t-1}}{1-\gamma} \cdot \frac{k_v}{q(\theta_{t-1})}$$
$$= \beta \cdot \frac{u'(\tilde{c}_t^w)}{u'(\tilde{c}_{t-1}^w)} \cdot \left(\int_{\epsilon_{s,t}}^{\infty} \big(y_t(\epsilon, h_t(\epsilon)) - v(h_t(\epsilon))\big) dG(\epsilon) + b_t \right)$$
$$+ \beta \cdot \frac{u'(\tilde{c}_t^w)}{u'(\tilde{c}_{t-1}^w)} \cdot \left(\frac{u(\tilde{c}_t^w) - u(b_t)}{u'(\tilde{c}_t^w)} + \frac{(\beta \cdot \lambda_{\theta,t-1} + \lambda_{c,t})}{n_t \cdot u'(\tilde{c}_t^w)} \cdot \frac{b_t}{n_t} \right)$$
$$+ \beta \cdot \frac{u'(\tilde{c}_t^w)}{u'(\tilde{c}_{t-1}^w)} \cdot (1 - G(\epsilon_{s,t})) \cdot \frac{(1-\gamma \cdot f(\theta_t)) + (1-f(\theta_t)) \cdot \chi_t}{1-\gamma} \cdot \frac{k_v}{q(\theta_t)}$$

Following Appendix 1.I we can derive a condition that expresses the deviation of the decentralized job-creation condition from of the constrained efficient job-creation condition in steady state and call it **deviation from optimal job-creation**

condition.

$$\left(\chi - \frac{\eta - \gamma}{1 - \eta}\right) \cdot \frac{1}{1 - \gamma} \cdot \frac{k_v}{q(\theta)}$$
$$= \frac{\beta}{1 - \beta \cdot (1 - f(\theta)) \cdot (1 - G(\epsilon_s))} \cdot \left(1 + \frac{\beta \cdot (1 + \eta \cdot BE) \cdot \lambda_\theta}{n \cdot u'(\tilde{c}^w)}\right) \cdot \frac{b}{n}$$

Following Appendix 1.I, we can also derive the **constrained efficient optimal separation condition:**

$$\begin{aligned}
\Leftrightarrow \quad & y_t(\epsilon_{s,t}, h_{stw,t}(\epsilon_{s,t})) - v(h_{stw,t}(\epsilon_{s,t})) + F \\
& + \frac{1 - \gamma \cdot f(\theta_t) + \left(1 - f(\theta_t)\right) \cdot \chi_t}{1 - \gamma} \cdot \frac{k_v}{q(\theta_t)} \\
& - \frac{(1 + \eta \cdot BE_t) \cdot \lambda_{\theta,t}}{n_t \cdot u'(\tilde{c}_t^w)} \cdot \tau_{S,t} \\
& - \frac{BE_t \cdot \lambda_{\theta,t}}{n_t \cdot u'(\tilde{c}_t^w)} \cdot \frac{1 - f(\theta_t)}{1 - \eta} \cdot \frac{k_v}{q(\theta_t)} \\
= & -\frac{1}{g(\epsilon_{s,t})} \cdot \frac{\lambda_{\epsilon,t}}{n_t \cdot u'(\tilde{c}_t^w)} \cdot \left(a_t \cdot h_{stw,t}(\epsilon_{s,t})^\alpha + c_f\right)
\end{aligned}$$

Imposing steady state and subtracting the decentralized job-creation condition lets us rearrange for the **optimal STW subsidy.**

$$\begin{aligned}
\Leftrightarrow \quad \tau_S \; = \; & (1 - f(\theta)) \cdot \left(\chi - \frac{\eta - \gamma}{1 - \eta}\right) \cdot \frac{1}{1 - \gamma} \cdot \frac{k_v}{q(\theta)} \\
& - \frac{(1 + \eta \cdot BE) \cdot \lambda_\theta}{n \cdot u'(\tilde{c}^w)} \cdot \tau_S \\
& + \frac{1}{g(\epsilon_s)} \cdot \frac{\lambda_\epsilon}{n \cdot u'(\tilde{c}^w)} \cdot \left(a \cdot h(\epsilon_s)^\alpha + c_f\right) \\
& - \frac{BE \cdot \lambda_\theta}{n \cdot u'(\tilde{c}^w)} \cdot \frac{1 - f(\theta)}{1 - \eta} \cdot \frac{k_v}{q(\theta)}
\end{aligned}$$

Inserting the deviation from optimal job-creation condition gives:

$$\begin{aligned}
\tau_S \; = \; & \frac{\beta \cdot (1 - f(\theta))}{1 - \beta \cdot (1 - f(\theta)) \cdot (1 - G(\epsilon_s))} \cdot \frac{b}{n} + \frac{\frac{1}{g(\epsilon_s)} \cdot \frac{\lambda_\epsilon}{n \cdot u'(\tilde{c}^w)}}{\left(1 + \frac{\beta \cdot (1 + \eta \cdot BE) \cdot \lambda_\theta}{n \cdot u'(\tilde{c}^w)}\right)} \cdot \left(a \cdot h(\epsilon_s)^\alpha + c_f\right) \\
& - \frac{\frac{BE \cdot \lambda_\theta}{n \cdot u'(\tilde{c}^w)}}{\left(1 + \frac{\beta \cdot (1 + \eta \cdot BE) \cdot \lambda_\theta}{n \cdot u'(\tilde{c}^w)}\right)} \cdot \frac{1 - f(\theta)}{1 - \eta} \cdot \frac{k_v}{q(\theta)}
\end{aligned}$$

Forming the first derivative for the optimal layoff tax gives:

$$\frac{\partial}{\partial \tau_S} \; = \; -\lambda_\epsilon \; = \; 0$$

Note that it sets the Lagrange multiplier for the decentralized separation condition to zero. This means that the layoff tax can impose the optimal number of separations. Applying this gives the result from Proposition 7.

$$\tau_S = \frac{\beta \cdot (1 - f(\theta))}{1 - \beta \cdot (1 - f(\theta)) \cdot (1 - G(\epsilon_s))} \cdot \left[1 + \frac{1-n}{n}\right] \cdot b - \tilde{BE}$$

References

Balleer, Almut, Britta Gehrke, Wolfgang Lechthaler, and Christian Merkl. 2016. “Does short-time work save jobs? A business cycle analysis.” *European Economic Review* 84 (C): 99–122. [64, 67]

Cahuc, Pierre, Francis Kramarz, and Sandra Nevoux. 2021. “The Heterogeneous Impact of Short-Time Work: From Saved Jobs to Windfall Effects.” IZA Discussion Papers 14381. Institute of Labor Economics (IZA). [64, 67]

Christoffel, Kai, and Tobias Linzert. 2010. “The Role of Real Wage Rigidity and Labor Market Frictions for Inflation Persistence.” *Journal of Money, Credit and Banking* 42 (7): 1435–46. [68]

Domeij, David, and Martin Floden. 2006. “The Labor-Supply Elasticity and Borrowing Constraints: Why Estimates are Biased.” *Review of Economic Dynamics* 9 (2): 242–62. [68]

Giupponi, Giulia, and Camille Landais. 2018. “Subsidizing labor hoarding in recessions: the employment and welfare effects of short time work.” CEP Discussion Papers dp1585. Centre for Economic Performance, LSE. [64, 67]

Haan, Wouter J. den, Garey Ramey, and Joel Watson. 2000. “Job Destruction and Propagation of Shocks.” *American Economic Review* 90 (3): 482–98. [68]

Jung, Philip, and Keith Kuester. 2015. “Optimal Labor-Market Policy in Recessions.” *American Economic Journal: Macroeconomics* 7 (2): 124–56. [68]

Krause, Michael U., and Thomas A. Lubik. 2007. “The (ir)relevance of real wage rigidity in the New Keynesian model with search frictions.” *Journal of Monetary Economics* 54 (3): 706–27. [68]

Shimer, Robert. 2005. “The Cyclical Behavior of Equilibrium Unemployment and Vacancies.” *American Economic Review* 95 (1): 25–49. [68]

Chapter 2

Optimal STW Policy and Labor Misallocation

2.1 Introduction

Short-time work (STW), a subsidy scheme conditional on reduced working hours, has become the primary labor market instrument in Germany to prevent job losses during economic downturns. It was used extensively during both the COVID-19 crisis and the Great Recession, with peak participation rates reaching 18% and 5.5% of the workforce, respectively. During the COVID-19 crisis, the scheme was also adopted by other major European economies. For instance, in France and Italy, 31% and 35% of workers, respectively, entered STW schemes, with the OECD average reaching 20.1%.[1]

While STW is widely recognized for its effectiveness in mitigating job losses, concerns have risen about unintended side effects as the program's scope has expanded. A particularly prominent concern in both academic and policy debates is that STW may hinder the reallocation of workers from unproductive to more productive firms. Cooper, Meyer, and Schott (2017) argue that this misallocation problem was a key reason why Germany's recovery from the Great Recession lagged behind that of the United States. This paper investigates whether the labor misallocation problem persists when STW is set optimally and explores how the policy should react to it.

To address this question, I develop a search and matching model of the labor market with Diamond-Mortensen-Pissarides type matching frictions (see Mortensen and Pissarides (1994)) that incorporates endogenous separations, flexible working hours, and generalized Nash bargaining and calibrate it to US data. In the model, firms are subject to either transitory shocks, characterized by a high prob-

1. Fitzenberger and Walwei, 2023.

ability of recovery, or permanent shocks, which have a low probability of recovery. The unemployment insurance (UI) and STW system are funded through lump-sum taxes on firm production. In this setting, the government aims to provide generous per period subsidies to firms facing transitory shocks to prevent unnecessary separations, while limiting support to firms with permanent productivity shocks to encourage efficient worker reallocation. It is assumed that the government cannot directly observe the persistence of a firm's shock.

When looking at optimal STW policy, the Ramsey planner faces a trade-off: generous STW benefits may lead to excessive retention of workers in permanently unproductive matches, while insufficient benefits risk premature separations of temporarily unproductive matches. I find that the optimal STW policy aims to minimize the welfare losses associated with both excess retention and premature separation. It does so by weighting firm types according to how many matches a change in STW benefits can preserve for each type, and adjusts the level of benefits accordingly. Quantitatively, the results suggest that optimal STW policy can mitigate a large share of welfare losses due to misallocation of workers.

When the government has reliable information about which jobs are at risk, it can even improve the outcome by using the time a worker spends on STW as a screening device. By reducing benefits with time spent on STW, the planner encourages separations in firms with permanent shocks, promoting reallocation. In contrast, when information is scarce, the government can use the extent of working hours reductions as eligibility condition. This type of eligibility condition naturally works as screening mechanism that helps target low-productivity firms. At the same time, it screens out firms with permanent shocks, making a declining benefit schedule unnecessary.

In more detail: Following Braun and Brügemann (2017), Teichgräber, Žužek, and Hensel (2022), and Stiepelmann (2024), I justify inefficient separations through the fiscal externality of the unemployment insurance (UI) system.[2] In settings with risk-averse workers and incomplete markets, the government faces a well-known trade-off: On the one hand, it aims to insure workers against income loss through generous UI benefits. On the other hand, higher UI benefits increase the worker's outside option, thereby discouraging vacancy creation and resulting in excessive job separations. To mitigate these distortions, STW schemes offer subsidies tied to reductions in working hours. This allows firms to lower labor costs without severing employment relationships, especially during temporary downturns in productivity.

2. Other mechanism of STW discussed in the literature entail the flexibilisation of the intensive margin (see Balleer, Gehrke, Lechthaler, and Merkl (2016) and Cooper, Meyer, and Schott (2017)), or financial constraints and wage rigidity (see Giupponi and Landais (2018))

This paper examines how to address the fiscal externality of UI systems with STW when firms are subject to both transitory and permanent productivity shocks. To eliminate all inefficient separations induced by the UI system, I show that STW benefits must be set such that the expected total transfer a worker receives on the STW system equals the expected UI benefits the worker would receive upon separation. This ensures that UI benefits do not generate incentives for premature separations. Since workers are ex-ante identical, they all face the same expected UI benefit upon separation. However, firms differ in their probability of recovery. Workers in firms experiencing permanent productivity shocks tend to remain on STW longer. Consequently, optimal per-period STW benefits must be lower in firms with low probability of recovery than in firms with a high probability of recovery.

Governments face an asymmetric information problem: they cannot directly observe whether a firm has a high or low probability of recovery. Interestingly, short-time work (STW) systems contain an inherent mechanism that helps mitigate this issue. Firms experiencing transitory productivity shocks are generally more willing to reduce working hours than firms facing permanent productivity losses. Because STW subsidies are typically linked to the extent of hours reductions, the system implicitly channels more generous support to firms with temporary shocks. As a result, STW naturally targets firms with transitory shocks, offering them higher subsidies.

Quantitatively, I find that while STW systems alleviate the welfare costs associated with worker misallocation, they do not eliminate them. STW still under-subsidizes firms with permanent productivity shocks and over-subsidizes firms with permanent productivity shocks. Further, although linking subsidies to reductions in working hours helps target firms with transitory shocks and partially addresses the misallocation problem, it introduces a significant drawback: it distorts firms' incentives, encouraging inefficient reductions in hours worked. This distortionary effect has been extensively discussed by Burdett and Wright (1989), Cahuc, Kramarz, and Nevoux (2021), and Stiepelmann (2024).

When choosing the optimal STW benefits, the government thus faces two challenges. On the one hand, it must address the distortion of working hours. On the other hand, the government still faces a misallocation problem. Setting benefits too generously leads to excess retention in permanently unproductive firms, impeding efficient labor reallocation. Conversely, setting benefits too low results in premature separations from firms with transitory productivity shocks.

Analytically, I show that optimal STW benefits can be derived through a three-step procedure. First, the government computes the hypothetical optimal STW benefit for each firm type — specifically, firms facing either transitory or permanent productivity shocks. These type-specific benefits reflect the trade-off between enforcing an efficient separation threshold for the firm type and limiting distortions to working hours induced by STW. This is done by setting the STW benefits below the transfers needed to implement the efficient separation threshold. Second, the government assigns weights to each firm type based on the sensitivity of match survival to changes in STW benefits. Greater weight is given to firm types for which marginal increases in benefits lead to the preservation of more employment relationships. In the final step, the government sets a uniform STW benefit level such that the weighted average of the hypothetical type-specific transfers matches the transfer provided under the actual STW policy. This procedure ensures that the system minimizes welfare losses arising from both excess retention and premature separation.

Quantitatively, I find that STW is highly effective in mitigating the welfare costs associated with worker misallocation. In fact, the welfare losses arising from misallocation are an order of magnitude smaller than those resulting from distortions in working hours. This suggests that while STW succeeds in addressing the misallocation problem to a significant extent, the primary source of inefficiency in such systems stems from their impact on working hours.

Depending on how the government designs the eligibility criteria, the STW system can be tailored to further reduce the welfare costs associated with worker misallocation. I distinguish between two cases. In case 1, the government is able to credibly verify whether a firm intends to separate from a worker. In practice, governments require firms to submit financial records or other documentation to demonstrate economic distress. In case 2, the government cannot credibly infer a firm wants to separate from a worker. Instead, eligibility is determined by whether the firm reduces working hours below a specified threshold. Almost all STW systems employ this kind of eligibility condition.

In case 1, the government can rule out windfall effects. It restricts access to STW to only those firms that cannot survive without STW support. In this case, it is optimal to reduce STW benefits over the duration of time a firm spends in the program. The reasoning is straightforward: firms facing transitory shocks tend to recover and exit the STW system more quickly, while those with permanent shocks remain on it longer. As a result, the share of permanently unproductive firms on STW increases over time. To encourage the reallocation of workers from these firms to more productive ones, the government finds it optimal to reduce STW benefits gradually.

In case 2, the government lacks reliable information about separations. However, it can apply a minimum hours reduction threshold as eligibility condition. As shown by Teichgräber, Žužek, and Hensel (2022), reduced hours serve as a screening mechanism for low productivity. The intuition is straightforward. When firms are highly productive, reducing working hours is costly because it leads to a substantial loss in output. In contrast, when firms experience a temporary decline in productivity, it becomes more attractive for both firms and workers to reduce hours to save on disutility of labor.

In addition to this mechanism, I find that the minimum hours-reduction threshold helps differentiate firms experiencing transitory versus permanent productivity shocks. Firms with transitory shocks are more willing to sustain productivity losses due to their higher likelihood of recovery compared to firms facing permanent shocks. Consequently, they are also more willing to reduce working hours before resorting to layoffs. By imposing a sufficiently strict hours-reduction requirement, the government can effectively exclude firms with permanent productivity shocks from STW eligibility.

When setting the eligibility condition for STW, the government faces a trade-off. A loose condition includes firms with permanent productivity shocks, but leads to windfall effects among many temporarily unproductive firms that would have retained workers even without STW support. A strict condition, by contrast, screens out permanently unproductive firms, including those that the Ramsey planner would want to support with STW benefits. The optimal choice depends on the relative share of firms with transitory versus permanent shocks in the economy. Quantitatively, I find that if more than 30% of firms face transitory shocks, the government prefers a strict eligibility condition. In this case, a declining STW benefit schedule becomes unnecessary, as low-productivity firms have already been screened out.

Finally, the model offers first insights into how STW policies should be optimally adjusted during periods of structural change. Consider a situation in which certain sectors experience declining demand or face obsolete business models, while others are expanding. In the model, this corresponds to a rise in the share of firms experiencing persistent productivity shocks — both positive and negative. The optimal policy response depends critically on how the government has designed the STW eligibility criteria.

If the government can credibly restrict access to firms that would otherwise have separated from their workers, then optimal STW benefits should decline in response to structural change. A higher share of firms with permanent negative productivity shocks raises the costs of retaining workers in unproductive jobs. Reducing STW benefits in this context fosters the necessary reallocation of labor to more productive uses.

By contrast, if the government relies on a minimum hours reduction threshold to determine eligibility, the optimal response depends on the intensity of the structural change. Suppose that in normal times, the economy has a high share of temporarily unproductive firms, and the existing threshold already screens out most permanently unproductive firms. In this case, a modest increase in permanent shocks requires no policy change. However, if the shift is substantial — leading to a significant rise in the number of firms with permanent productivity losses — then the government may need to loosen the eligibility condition to accommodate these additional firms. In such a scenario, optimal STW benefits should fall sharply, as the system must now contend with a higher risk of inefficient worker retention and additional windfall effects among temporarily unproductive firms. Note that, according to this theory, the government should adjust STW policy during structural change in a manner that is the exact opposite of what Stiepelmann (2024) recommends for recessions. While a recession calls for more generous STW benefits combined with stricter eligibility condition, big structural change shift may require a reduction in benefits and a relaxation of the eligibility condition.

Literature. The literature highlights three primary side effects of STW.[3] First, STW distorts working hours. As the level of STW subsidies rises with reduced working hours, matches opt for lower working hours to gain additional benefits. Burdett and Wright (1989) and Van Audenrode (1994) discuss these effects in an implicit contract model. Cahuc, Kramarz, and Nevoux (2021) introduces them into a partial equilibrium search and matching model. Stiepelmann (2024) shows in a general equilibrium search and matching model that these effects reduce the ability of optimal STW policy to stabilize the business cycle with STW. This paper also takes hours distortion effects into account. In fact, later sections will show that these distortions reduce the responsiveness of optimal STW benefits to changes in the share of permanently versus temporarily unproductive firms. As a result, the optimal policy becomes less sensitive to structural change, and the optimal STW benefits need to decline less steeply with the duration a worker spends in the STW system.

Second, STW may give rise to an adverse selection problem, where workers who are not at risk of job loss opt to enter the STW system. Cahuc, Kramarz, and Nevoux (2021) show that such windfall effects are costly as they spread the hours distortion effect to a larger number of firms. Teichgräber, Žužek, and Hensel (2022) show that hours reduction is used as screening mechanism to reduce the selection of productive firms into STW. This paper adds a new dimension to the costs associated with windfall effects. It shows that temporarily unproductive firms with a high probability of recovery are willing to retain workers even without STW support, while firms with low recovery prospects require STW assistance to do so.

3. An additional concern is raised by Bossler, Osiander, Schmidtke, and Trappmann (2023). They argue that firms might misreport actual working hours, which they view as a drawback of using STW.

This makes it impossible for the government to eliminate windfall effects through a minimum hours reduction threshold as an eligibility condition, without simultaneously excluding firms that the government would want to support.

Third, Cooper, Meyer, and Schott (2017) claim that STW results in a loss of allocative efficiency, which is empirically supported by Giupponi and Landais (2022). The idea is that it sustains employment in unproductive firms, thereby impeding productive firms' ability to find workers and expand. This obstruction to worker reallocation leads to a reduction in overall output, despite maintaining low unemployment. While this mechanism is fundamental in the discussion of STW policy, the paper leaves two questions unanswered:

Both papers examine the provision of STW from a positive perspective and do not address how STW should optimally respond to its potential to cause labor misallocation. From a theoretical point of view, a suboptimally calibrated and overly generous STW system can trivially retain workers in unproductive jobs, thereby reducing allocative efficiency. This paper takes a more normative approach to the issue. It asks whether labor misallocation remains a problem when the STW system is optimally designed. The analysis shows that heterogeneity in shock duration across firms is key to the emergence of labor misallocation. If all firms experienced shocks of equal duration, there would be no scope for labor misallocation in a canonical DMP model under an optimal STW policy. The paper also outlines how the government should address labor misallocation when designing STW schemes.

2.2 Model

The economy is populated by a continuum of workers of measure one and infinitely many one-worker firms. The economy is closed. Each firm produces a homogeneous and non-storable good. Firms are subject to idiosyncratic productivity shocks that can differ in their duration. Firms open costly vacancies before the realization of the productivity shock is known. The productivity shock is drawn first when new matches are formed. Based on the productivity shock and its duration, firms decide to separate from a worker, and production takes place.

Firm Side. The value of a worker for a firm outside STW is defined by:

$$J(\epsilon,\lambda) \;=\; y(\epsilon,h(\epsilon)) - w + \beta \cdot \left[(1-\lambda)\cdot J(\epsilon,\lambda) + \lambda \cdot \bar{J}\right]$$

Firms discount the future with β. With probability λ they draw a new idiosyncratic productivity level $\epsilon \geq 0$ from the c.d.f $G_\lambda(\epsilon)$ and a new probability $\lambda \in [0,1]$ from the c.d.f. $P(\lambda)$. The corresponding p.d.f. is $g_\lambda(\epsilon)$, respectively $p(\lambda)$. Note that we allow the c.d.f. for the productivity level to depend on λ. λ can be seen as a measure for the persistence of the productivity level. $1/\lambda$ denotes the expected duration until a new productivity level is drawn. If λ is small, we say that a firm got a permanent productivity shock, which makes the firm either permanently productive or unproductive. If λ is large, we say that a firm has a transitory productivity shock, which makes the firm either temporarily productive or unproductive. The expected value of a firm right before a new productivity level and its persistence is drawn is denoted as $\bar{J}$.

A firm receives the output $y(\epsilon, h)$ of the match and pays a salary w to the worker. The output depends on the current productivity ϵ, the aggregate productivity level A, the number of hours worked h, and the cost shock $(\mu_\epsilon - \epsilon) \cdot c_f$. Output increases linearly in productivity and with positive and decreasing marginal returns to working hours, $\alpha \in (0, 1)$:

$$y(\epsilon, h) \;=\; A \cdot \epsilon \cdot h^\alpha - (\mu_\epsilon - \epsilon) \cdot c_f \quad \text{with} \quad E_{G_\lambda}[y(\epsilon)] \;=\; E_{G_\lambda}[A \cdot \epsilon \cdot h(\epsilon)^\alpha]$$

On STW, the government pays STW benefits $s(\lambda, h, k)$ to the worker. At this stage, those benefits can depend on three items: First, the shock duration $1/\lambda$, second, the working hours h, and, finally, the time k a worker spent on the STW system. We will vary the benefits depending on the information assumption, specified in a separate section. This also applies to the choice of the eligibility condition, that is, the rule under which workers can enter the STW system. In general, firms and workers can enter the STW once their productivity falls below the eligibility threshold

$$\epsilon \;\leq\; R_{stw}(\lambda)$$

The STW system is set up so that the government can vary the benefits within the first K periods. After that, the government sets constant benefits $s(\lambda, h, K)$. If K=1, we return to a system in which STW benefits are independent of the duration of time spent on STW. Note that the benefits can depend on the working hours h and thus will influence the hours choice decision. Thus, working hours and separation thresholds might differ over the time spent on the STW system: $h_{stw,k}(\epsilon) \neq h_{stw,k+1}(\epsilon)$, $R_{stw}(\lambda, k) \neq R_{stw}(\lambda, k+1)$

The value of a worker for a firm can thus be denoted as:

$$\begin{aligned} J_{stw}(\epsilon, \lambda, k) \;&=\; y(\epsilon, h_{stw,k}(\epsilon)) - w_{stw}(\epsilon, \lambda, k) \\ &\quad + \beta \cdot \Big[(1 - \lambda) \cdot \mathbb{1}(\epsilon \;\geq\; R(\lambda, k + 1)) \cdot J_{stw}(\epsilon, \lambda, k + 1) + \lambda \cdot \bar{J} \Big] \end{aligned}$$

A difference in the separation threshold becomes relevant when STW benefits decline over time and the corresponding separation thresholds rise. Suppose a firm does not draw a new productivity shock in the next period. Then, even if its productivity remains unchanged, it may now choose to separate from the worker due to the higher threshold. Only if productivity is sufficiently high — specifically, if $\epsilon > R(\lambda, k+1)$ — the worker will remain employed.

After K periods, the government chooses fixed STW benefits. By setting them to zero, the government could effectively end the STW system. The value of a worker for a firm that spent at least K periods on STW is thus:

$$\begin{aligned} J_{stw}(\epsilon, \lambda, K) &= y(\epsilon, h_{stw,K}(\epsilon)) - w_{stw}(\epsilon, \lambda, K) \\ &\quad + \beta \cdot \Big[(1-\lambda) \cdot J_{stw}(\epsilon, \lambda, K) + \lambda \cdot \bar{J} \Big] \end{aligned}$$

Finally, $\bar{J}$ describes the value of a worker for a firm before the idiosyncratic productivity ϵ and its persistence $1/\lambda$ are drawn:

$$\begin{aligned} \bar{J} &= -\tau + \int_0^1 \left(\int_{\max\{R_{stw}(\lambda), \xi(\lambda)\}}^{\infty} J(\epsilon, \lambda) dG_\lambda(\epsilon) \right) dP(\lambda) \\ &\quad + \int_0^1 \int_{R(\lambda,1)}^{\max\{R_{stw}(\lambda), R(\lambda,1)\}} J_{stw}(\epsilon, \lambda, 1) dG_\lambda(\epsilon) dP(\lambda) \end{aligned}$$

If the productivity level of the firms is larger than the STW threshold $\epsilon > R_{stw}(\lambda)$, then they produce normally. If the productivity falls into the interval $\epsilon \in [R(\lambda, 1), R_{stw}(\lambda)]$, then firms are unproductive enough to enter the STW system but is still so productive that it can survive with the help of the STW system. If the productivity falls further $\epsilon < R(\lambda, 1)$, then firms will separate from their workers even with STW benefits. There exist one special case, reflected by $\max\{R_{stw}(\lambda), \xi(\lambda)\}$. In fact, the government can set the eligibility threshold $\xi(\lambda) > R_{stw}(\lambda)$ so strict that there exist firms that cannot survive without STW but whose productivity is not low enough to enter the STW system. Here $\xi(\lambda)$ denotes the separation threshold without STW benefits. For simplicity, the tax τ is paid at shock-arrival.

Job-creation is determined by the free-entry condition:

$$\frac{k_v}{q} = \bar{J}$$

Firms post vacancies until the expected gain from filling those vacancies equals the expected costs. k_v is defined as the vacancy posting costs, while q is the probability of filling the vacancy. Vacancies are posted before the productivity level and the shock persistence of the match is known.

Due to search frictions, the labor market is not perfectly competitive. Firms make positive profits and pay dividends to the workers. Dividends equal the expected gain from production minus wage costs for the workers, taxes paid to the government,t and vacancy posting costs:

$$\begin{aligned} \Pi \;=\; & \int_0^1 \Big(\sum_{k=1}^{K} \frac{n(\lambda,k)}{1-\rho(\lambda,k)} \cdot \Big[\int_{\max\{R_{stw}(\lambda),\xi(\lambda)\}}^{\infty} [y(\epsilon,h(\epsilon)) - w]\, dG_\lambda(\epsilon) \\ & + \int_{\max\{R(\lambda,1),\dots,R(\lambda,k)\}}^{\max\{R_{stw}(\lambda),R(\lambda,1),\dots,R(\lambda,k)\}} \big[y(\epsilon,h_{stw,k}(\epsilon)) - w_{stw}(\epsilon,\lambda,k) \big] G_\lambda(\epsilon) \Big] dP(\lambda) \\ & - v \cdot k_v - n^s \cdot \tau \end{aligned}$$

Worker Side. The utility of a worker employed in a firm with productivity level ϵ and persistence λ can be denoted as:

$$V(\epsilon,\lambda) \;=\; w + \Pi - \phi(h(\epsilon)) + \beta \cdot \big[(1-\lambda) \cdot V(\epsilon,\lambda) + \lambda \cdot \bar{V} \big]$$

Workers are assumed to be risk-neutral.[4] They derive utility from consumption and disutility from working hours $h(\epsilon)$. They receive a salary w from the firms plus dividend payments. With probability λ their firm draws a new productivity level and persistence giving an expected value of $\bar{V}$ to the workers.

The utility of a worker after k periods on the STW system can be denoted as:

$$\begin{aligned} V_{stw}(\epsilon,\lambda,k) \;=\; & w_{stw}(\epsilon,\lambda,k) + s(\lambda,h_{stw,k}(\epsilon),k) + \Pi - \phi(h_{stw,k}(\epsilon)) \\ & + \beta \cdot \big[(1-\lambda) \cdot \mathbb{1}(\epsilon \;\geq\; R(\lambda,k+1)) \cdot V_{stw}(z,\lambda,k+1) + \lambda \cdot \bar{V} \big] \end{aligned}$$

On STW, workers get the reduced income from the firms plus STW benefits from the government according to the subsidy function $s(\lambda,h,k)$. After K periods, the government chooses fixed STW subsidies. In this case, the value of a worker is denoted as:

$$\begin{aligned} V_{stw}(\epsilon,\lambda,K) \;=\; & w_{stw}(\epsilon,\lambda,K) + s(\lambda,h_{stw,K}(\epsilon),K) + \Pi - \phi(h_{stw,K}(\epsilon)) \\ & + \beta \big[(1-\lambda) \cdot V_{stw}(z,\lambda,K) + \lambda \bar{V} \big] \end{aligned}$$

The utility of an unemployed worker is:

$$U \;=\; b + \beta \cdot \big[f \cdot \bar{V} + (1-f) \cdot U \big]$$

Workers get unemployment benefits b from the government. With probability f the worker finds a new job. With probability $1-f$ he remains unemployed.

4. Stiepelmann (2024) shows that in a setting with financially unconstrained firms and risk-averse workers who cannot save, STW serves no insurance purpose and is used solely to counteract the fiscal externality of UI. Since the focus of the present model is on worker misallocation, risk aversion is excluded for simplicity.

The expected utility workers derive from employment, before the productivity shock is realized and the shock duration is known, can be denoted as:

$$\begin{aligned}\bar{V} &= \int_0^1 \int_{\max\{R_{stw}(\lambda),\xi(\lambda)\}}^{\infty} V(\epsilon,\lambda) dG_\lambda(\epsilon) dP(\lambda) \\ &+ \int_0^1 \int_{R(\lambda,1)}^{\max\{R_{stw}(\lambda),R(\lambda,1)\}} V_{stw}(\epsilon,\lambda,1) dG_\lambda(\epsilon) dP(\lambda) \\ &+ \int_0^1 \rho(\lambda,1)\cdot U \cdot dP(\lambda)\end{aligned}$$

Nash-Bargaining. Firms and workers bargain over wages, working hours, and separation thresholds. Bargaining takes place before the realization of the productivity shock and its duration. The contract remains in effect for the entire duration of the employment relationship. Employment terminates when the worker separates from the firm.

$$\max_{w,w_{stw}(z,\lambda),h(\epsilon),h_{stw,k}(\epsilon),\xi(\lambda),R(\lambda,k)} \bar{J}^{1-\eta}(\bar{V}-U)^{\eta}$$

η denotes the bargaining power of the worker. The bargaining outcomes are derived in Appendix 2.B.3. Salaries are determined by a surplus splitting rule. Firms get a share of $1-\eta$ from the joint surplus $\bar{S}=\bar{J}+\bar{V}-U$. Workers receive η.

$$\bar{J} = 1-\eta\cdot\bar{S}, \qquad \bar{V}-U = \eta\cdot\bar{S}$$

Working hours outside STW are set so that the marginal product of labor equals the marginal disutility from working:

$$\frac{\partial y(\epsilon,h(\epsilon))}{\partial h} = \phi'(h(\epsilon)) \tag{2.2.1}$$

If productivity is large, then working hours are large. If productivity is small, then working hours are small.

If the government adjusts its subsidy function according to the hours choice on STW, then the slope of the hours choice function influences the working hours decision:

$$\frac{\partial y(\epsilon,h_{stw,k}(\epsilon))}{\partial h} = \phi'(h_{stw,k}(\epsilon)) - \frac{\partial s(\lambda,h_{stw,k}(\epsilon),k)}{\partial h} \tag{2.2.2}$$

Note that if benefits increase with a fall in working hours $\frac{\partial s(\lambda,h_{stw,k}(\epsilon))}{\partial h}<0$, then working hours are decreased below the socially optimal level. Firms and workers have an incentive to choose suboptimal low working hours to secure additional benefits from the government.

Since wages are determined by the surplus splitting rule, firms will decide to separate when the joint surplus of firms and workers becomes negative. For notational brevity, let us denote the per-period surplus — defined as output net of the disutility of work — for both regular work and STW as follows:

$$z(\epsilon) \;=\; y(\epsilon, h(\epsilon)) - \phi(h), \quad z_{stw}(\epsilon) \;=\; y(\epsilon, h_{stw}(\epsilon)) - \phi(h_{stw}(\epsilon))$$

Accordingly, we can determine the separation threshold of a firm without access to STW benefits $\xi(\lambda)$ as:

$$z(\xi(\lambda)) - b + \frac{\lambda - \eta \cdot f}{1 - \eta} \cdot \frac{k_v}{q} \;=\; 0$$

Note that the threshold $\xi(\lambda)$ depends on λ. When firms and workers are at the brink of separation, λ can be interpreted as the probability of recovery. A higher recovery probability — or, equivalently, a shorter expected duration of the negative productivity shock — leads firms to adopt a lower separation threshold. Firms are more willing to retain workers despite low current productivity, in anticipation of higher profitability in the near future.

The separation thresholds under the STW system are determined by firms' expectations about the future STW benefit schedule. In this case, however, it is crucial to account for whether the STW benefit schedule is increasing or decreasing over time. For a decreasing STW system, in which

$$s(\lambda, h_{\mathrm{stw},k}(R(\lambda, k)), k) \;\geq\; s(\lambda, h_{\mathrm{stw},k+1}(R(\lambda, k+1)), k+1),$$

The separation threshold must rise with the duration of time spent in the STW program. Firms at the margin of separation recognize that if they do not draw a new productivity shock, their continuation value in the next period is zero, prompting separation. The corresponding separation threshold $R(\lambda, k)$ is therefore characterized by:

$$z(R(\lambda, k)) + s(\lambda, h_{stw,k}(R(\lambda, k)), k) - b + \frac{\lambda - \eta \cdot f}{1 - \eta} \cdot \frac{k_v}{q} \;=\; 0$$

In the case of an increasing benefits schedule, the separation threshold in the first period effectively determines the separation thresholds in all subsequent periods. If separation was not optimal at lower benefit levels in the initial period, then it will also not be optimal at higher benefit levels in later periods. The corresponding separation threshold $R(\lambda, 1)$ is pinned down by:

$$\begin{aligned}
&\sum_{k=1}^{K-1} \lambda \cdot (1-\lambda)^{k-1} \cdot \left[z_{stw,k}(R(\lambda,1)) + \tau_{stw}(k) \cdot \left(\bar{h} - h_{stw,k}(R(\lambda,1))\right)\right] \\
&+ (1-\lambda)^{K-1} \cdot \left[z_{stw,K}(R(\lambda,1)) + \tau_{stw}(K) \cdot \left(\bar{h} - h_{stw,K}(R(\lambda,1))\right)\right] \\
&- b + \frac{\lambda - \eta \cdot f}{1-\eta} \cdot \frac{k_v}{q(\theta)} \;=\; 0
\end{aligned}$$

Note that STW benefits in all future periods play a crucial role in the firm's decision in period 1 whether to remain in the STW system.

Labor market flows. Unemployed u and vacancies v are matched according to a Cobb-Douglas matching function:

$$m \;=\; \psi \cdot v^{1-\gamma} \cdot u^{\gamma}$$

The parameter ψ determines the matching efficiency, and $\gamma \in (0,1)$ denotes the elasticity of the matching function for unemployment. The labor market tightness is defined as the ratio of vacancies to unemployed $\theta = \frac{v}{u}$. Based on the matching function and the labor market tightness, we can derive the probability to find a job f and the probability to fill a vacancy q:

$$\begin{aligned}
f(\theta) &= \psi \cdot \theta^{1-\gamma}, \\
q(\theta) &= \psi \cdot \theta^{\gamma}
\end{aligned}$$

The number of employed workers equals the total number of workers minus the number of unemployed workers $n = 1 - u$. Employment can also be expressed as the sum of all employed workers with shock duration $1/\lambda$ in period k after shock arrival:

$$n \;=\; \int_0^1 \sum_{k=1}^{K} n(\lambda, k) d\lambda$$

Likewise, the number of workers on STW n_{stw} is the sum of all workers with shock duration $1/\lambda$ that spend k periods on the STW system:

$$n_{stw} \;=\; \int_0^1 \sum_{k=1}^{K} n_{stw}(\lambda, k) d\lambda$$

To keep notation simple, it is constructive to denote the number of matches that are exposed to new productivity shocks. These concern the number of new formed matches $f \cdot u$ plus those matches that are hit by a new productivity shock $\lambda \cdot n(\lambda, k)$:

$$n^s = \theta \cdot q(\theta) \cdot u + \int_0^1 \sum_{k=1}^{K} \lambda \cdot n(\lambda, k) d\lambda$$

This allows us to define the number of firms with shock duration $1/\lambda$ that spent k periods on the STW system as:

$$n(\lambda, k) = (1-\lambda)^{k-1} \cdot \big(1 - \rho(\lambda, k)\big) \cdot p(\lambda) \cdot n^s \quad \text{for} \quad k = 1, 2, ..., K-1$$

Likewise, the number of firms that spent at least K periods on the STW system is defined by:

$$n(\lambda, K) = \frac{(1-\lambda)^{K-1}}{\lambda} \cdot \big(1 - \rho(\lambda, K)\big) \cdot p(\lambda) \cdot n^s$$

In both equations, $p(\lambda)$ can be interpreted as the probability of getting a shock with duration $1/\lambda$. In period k, $(1-\lambda)^k$ of the initial firms have not received a new shock yet. After at least K periods, a fraction $\frac{(1-\lambda)^{K-1}}{\lambda}$ of firms did not receive a new shock. With probability $1 - \rho(\lambda, k)$ the worker was not separated until period k. The separation probability is defined by:

$$\begin{aligned} \rho(\lambda, k) &= G_\lambda\big(\max\{R(\lambda, 1), ..., R(\lambda, k)\}\big) \\ &\quad + G_\lambda\big(\max\{\xi(\lambda), R_{stw}(\lambda)\}\big) - G_\lambda\big(\max\{R_{stw}(\lambda), R(\lambda, 1), ..., R(\lambda, k)\}\big) \end{aligned}$$

Note that this equation highlights two important features. First, the max operator in the first line reflects that depending on the use of the STW system, different separation thresholds become important: in a falling STW system, the separation threshold in period k determines the separation rate; likewise, in an increasing STW system, the threshold from period 1 determines the separation rate, as firms that chose not to separate at lower benefit levels will also not do so at higher levels. Second, the second line of the equation captures the possibility that some firms are denied access to STW, even though access could have prevented separations and preserved the match.

Accordingly, we can also define the number of workers in firms with shock duration $1/\lambda$ that spent k periods, respectively at least K periods on STW as:

$$\begin{aligned} n_{stw}(\lambda, k) &= (1-\lambda)^{k-1} \cdot \rho_{stw}(\lambda, k) \cdot p(\lambda) \cdot n^s \quad \text{for} \quad k = 1, 2, ..., K-1 \\ n_{stw}(\lambda, K) &= \frac{(1-\lambda)^{K-1}}{\lambda} \cdot \rho_{stw}(\lambda, K) \cdot p(\lambda) \cdot n^s \end{aligned}$$

Here, $\rho_{stw}(\lambda, k)$ defines the probability that a worker enters STW and is not separated after k periods:

$$\rho_{stw}(\lambda, k) \;=\; G\big(\max\{R_{stw}(\lambda), R(\lambda,1), ..., R(\lambda,k)\}\big) - G\big(\max\{R(\lambda,1), ..., R(\lambda,k)\}\big)$$

Government. The government needs to balance its budget at each time t. It bears the costs of the unemployment insurance system and the costs of the STW system. Its revenue stems from the lump sum tax. The budget constraint of the government can be written as:

$$\begin{aligned} n^s \cdot \tau \;=\;& (1-n) \cdot b \\ &+ \int_0^1 \sum_{k=1}^{K} n_{stw}(\lambda,k) \cdot \int_{\max\{R(\lambda,1),...,R(\lambda,k)\}}^{\max\{\xi(\lambda),R_{stw}(\lambda)\}} s(\lambda, h_{stw,k}(\epsilon), k) \cdot \frac{dG_\lambda(\epsilon)}{\rho_{stw}(\lambda,k)} dP(\lambda) \end{aligned}$$

2.2.1 Various Policy Regimes

In the following, I analyze three STW systems that differ in their information requirements. Each STW system is defined by two components: an eligibility condition and a subsidy function. The eligibility condition specifies a threshold $R_{\text{stw}}(\lambda)$, below which a match's idiosyncratic productivity ϵ must fall for the firm to qualify for STW support. The subsidy function compensates firms for reduced working hours by replacing part of the worker's lost income. Specifically, it provides a transfer of

$$\tau_{\text{stw}} \cdot \left(\bar{h} - h_{\text{stw}}\right)$$

where τ_{stw} denotes the per-hour STW benefit, $\bar{h}$ is the firm's normal working hours, and h_{stw} is the reduced hours under STW. To isolate the impact of linking transfers to working hours, I also define a benchmark case in which the government provides a simple lump-sum transfer or subsidy, independent of hours worked.

Jobs at Risk and Shock Persistence Observable. In the first case, I assume that the government can perfectly observe whether a firm is willing to separate from a worker or not and is able to distinguish between permanent and transitory shocks. Further, I assume that the government can observe the duration each worker has spent in the STW system.

The last assumption is relatively easy for governments to verify. In contrast, eliciting information about whether a job is genuinely at risk of termination and whether a productivity shock is temporary requires more effort. To acquire this information, governments typically require firms to provide evidence that their economic difficulties are indeed temporary and to provide access to their financial records and books.

Using the available information, the government is eager to implement an eligibility condition where workers can only enter STW if they would otherwise have been separated:

$$S(\epsilon, \lambda) \;\leq\; 0$$

Stiepelmann (2024) argues that a looser eligibility condition cannot be optimal as it would distort working hours of firms without rescuing additional workers. Furthermore, a stricter eligibility condition would separate the most productive workers on STW, which may also not be optimal. Note that the corresponding STW threshold depends on duration of the shock it has received: $R_{stw}(\lambda) = \xi(\lambda)$.

Further, the government will also want to depend the STW benefits on the duration of the productivity shock. That way, the government can distinguish between firms with transitory shocks and permanent shocks with the transfer it pays. Additionally, the government will want to exploit information about the time a worker spends on STW. The corresponding STW subsidies can thus be defined as:

$$s(\lambda, h, k) \;=\; \tau_{stw}(\lambda, k) \cdot \left(\bar{h} - h_{stw,k}(\epsilon)\right)$$

Jobs at Risk but not Shock Persistence Observable. In the following case, I assume that the government cannot distinguish between transitory and permanent shocks anymore. In practice, not even firms have certainty about how persistent a shock to their business model is, and it might be relatively hard for the government to infer from the financial records alone. Therefore, it might be hard to find credible information for the government about the actual duration of the shock so that the government cannot depend their STW benefits on it:

$$s(\lambda, h, k) \;=\; \tau_{stw} \cdot \left(\bar{h} - h_{stw,k}(\epsilon)\right), \qquad s(\lambda, h, k) \;=\; \tau_{stw}(k) \cdot \left(\bar{h} - h_{stw,k}(\epsilon)\right)$$

I will distinguish between a case where the government depends their working hours on the time a worker spent on STW and a case where this is not the case as reference point. Since it is still observable whether a worker would be separated, the same eligibility condition as in case 1 applies:

$$S(\epsilon, \lambda) \;\leq\; 0$$

Note that from the fact that a firm separates, the government cannot automatically learn its shock persistence.

Only working hours observable. Finally, I assume that the government can neither observe whether a firm is at risk nor distinguish between permanent and transitory shocks. Checking financial records would not provide any usable information. The only credibly observable variable is the number of hours worked on STW. In practice, governments not only screen firms by reviewing their financial records but also employ a second tool to identify jobs at risk. Specifically, they require workers to reduce their working hours below a certain threshold D in order to qualify for entry into the STW system.

$$h_{stw}(\epsilon) \;\leq\; D$$

In this case, the eligibility threshold does not depend on the probability of recovery $h_{stw}(R_{stw}) = D$. Regarding the STW benefits, I still consider a case that depends working hours on time spent on STW and a case without this feature as a benchmark:

$$s(\lambda, h, k) \;=\; \tau_{stw} \cdot \left(\bar{h} - h_{stw}(\epsilon)\right), \qquad s(\lambda, h, k) \;=\; \tau_{stw}(k) \cdot \left(\bar{h} - h_{stw,k}(\epsilon)\right)$$

Subsidy System. Finally, I define a subsidy system for case 1 and case 2 of the information assumptions. This will help us understand how the STW system handles information asymmetry about the shock duration. In case 1, the government can directly choose the subsidy dependent on the probability of recovery λ, working hours h, and time spent on the STW system λ: s(λ,k). In case 2, the government can only choose a uniform subsidy, respectively a subsidy that depends on time spent on the subsidy system.

$$s(\lambda, h, k) \;=\; s, \quad s(\lambda, h, k) \;=\; s(k)$$

Table 2.4.1 in section 2.4.4 gives a good overview of the different combinations of information available to the government and instruments. The comparison of the different policy regimes allows us to clearly identify the costs of labor misallocation and enables us to compare them to the costs of hours distortions and windfall effects.

2.2.2 Inefficiencies in the Economy

There are two reasons for the government to intervene in the labor market. First, it is well known that the UI system distorts the job-creation and job-destruction decisions in the economy (see Pissarides (2000)). The UI benefits provided to workers increase their outside options, allowing them to demand higher wages. Consequently, firms face higher labor costs, which in turn reduce their expected profits. As a result, firms post suboptimal few vacancies from the perspective of the social planner. Furthermore, smaller expected profits decrease the willingness of firms to hoard labor, leading to inflated separation rates. For a detailed discussion see also Stiepelmann (2024).

Second, the Hosios-Condition (see Hosios (1990)) might not be fulfilled. When posting vacancies, firms don't consider that they reduce the probability of other firms to find a worker (congestion externality) and increase the probability for workers to find a job (thick market externality). In consequence, firms post either too many vacancies, inflating vacancy posting costs, or too few vacancies, leading to too much unemployment. Under a Cobb-Douglas matching function, the Hosios-condition states that firms post the optimal number of vacancies if the bargaining power of a worker equals the elasticity of the matching function with respect to unemployment:

$$\eta = \gamma$$

The next section discusses how optimal STW policy can deal with these two inefficiencies.

2.2.3 Ramsey Problem

For simplicity, we consider an economy without discounting $\beta \to 1$. The Ramsey planner weighs the utility of every household equally. He chooses the policy systems described in section 2.1 such that they maximize output minus disutility of work and costs from vacancy creation, subject to the decentralized economy.

$$W = \int_0^1 \sum_{k=1}^{K} \Big[\underbrace{n(\lambda,k) \cdot \tilde{z}(\lambda,k)}_{\text{Undistorted Utility Created by Matches}} - \underbrace{n_{stw}(\lambda,k) \cdot \tilde{\Omega}(\lambda,k)}_{\text{Welfare Costs of STW}} \Big] dP(\lambda) - \underbrace{v \cdot k_v}_{\text{Costs of Vacancy Creation}}$$

Welfare can be expressed as the total undistorted utility generated by matches, minus the welfare costs associated with the use of short-time work (STW), and minus the costs of posting vacancies. Definition 1 introduces $\tilde{z}(\lambda,k)$ as the average undistorted welfare created by a match with shock duration $1/\lambda$, k periods after the shock occurred. This measure captures output net of the disutility of work, assuming that working hours are chosen optimally.

Definition 1, Average Undistorted Utility created by Matches
The average undistorted period utility created by a match with shock duration $1/\lambda$, *k periods after the shock has arrived is defined by:*

$$\bar{z}(\lambda,k) = \int_{A(\lambda,k)} z(\epsilon)\cdot\frac{dG_\lambda(\epsilon)}{\rho(\lambda,k)}$$

where $A(\lambda,k)$ is defined as:

$$\begin{aligned} A(\lambda,k) =& [\max\{R(\lambda,1),...,R(\lambda,k)\}, \max\{R_{stw}(\lambda), R(\lambda,1),...,R(\lambda,k)\}] \cup \\ & [\max\{R_{stw}(\lambda), \xi(\lambda)\}, \infty) \end{aligned}$$

Definition 2 defines $\tilde{\Omega}(\lambda,k)$ as the average utility loss caused by hours distortions induced by the STW system for a match with shock duration $1/\lambda$ that has spent k periods in the STW program. This loss is measured as the difference between output minus the disutility of work under optimal hours and output minus disutility when working hours are distorted by the STW subsidy.

Definition 2, Welfare Costs Hours Distortions
The average welfare cost of a match in STW with shock duration $1/\lambda$ that spent k periods on STW is defined as the average difference between the period utility created by the match under optimal working hours and working hours distorted by STW.

$$\tilde{\Omega}(\lambda,k) = \int_{\max\{R(\lambda,1),...,R(\lambda,K)\}}^{\max\{R_{stw}(\lambda),R(\lambda,1),...,R(\lambda,K)\}} \left[z(\epsilon) - z_{stw}(\epsilon)\right]\frac{dG_\lambda(\epsilon)}{\rho_{stw}(\lambda,k)}$$

Likewise, the expected welfare costs of a match with shock duration $1/\lambda$, k periods after shock arrival is defined as:

$$\Omega(\lambda,k) = \rho_{stw}(\lambda,k)\cdot\tilde{\Omega}(\lambda,k)$$

Lemma 1 summarizes some of its properties. Note that the welfare costs are non-negative. If $h_{stw}(\epsilon) > h(\epsilon)$, then workers work suboptimal many hours, leading to inflated disutility from work. If $h_{stw}(\epsilon) < h(\epsilon)$, then workers work too few hours, leading to an output loss. If $h_{stw}(\epsilon) = h(\epsilon)$, then workers work the optimal number of hours, and the welfare costs are zero.

Generally, the government wants to increase subsidies if the number of hours worked falls. Note, however, that if subsidies increase faster in the fall in working hours, firms and workers have a higher incentive to choose suboptimal low working hours to secure additional subsidies.

Lemma 1, Welfare Effect Hours Distortion
The welfare costs of STW are non-negative:

$$\tilde{\Omega}(\lambda, k) \;\geq\; 0$$

If the STW benefits rise or the eligibility thresholds get looser, then the welfare costs from STW rise as well:

$$\frac{\partial \tilde{\Omega}(\lambda, k)}{\partial \tau_{stw}} \;\geq\; 0, \qquad \frac{\partial \Omega(\lambda, k)}{\partial R_{stw}(\lambda)} \;\geq\; 0$$

PROOF: Appendix 2.B.4

2.3 Calibration

This section calibrates the model to the U.S. economy. In the following sections, I present the theoretical and quantitative results jointly, which helps to clarify the underlying mechanisms and to strengthen the interpretation of the policy trade-offs.

I will refer to the model without a STW system as the baseline model. For simplicity, I assume that the shock persistence follows a Bernoulli distribution. I assume that a firm gets hit by a transitory shock with probability p_T and a persistent shock with probability $1 - p_T$.

$$P(\lambda) \;=\; \begin{cases} p_T & \lambda \;=\; \lambda_T \\ 1 - p_T & \lambda \;=\; \lambda_P \\ 0 & \text{else} \end{cases}$$

Productivity shocks are log-normally distributed and conditional on shock persistence $\log(\tilde{e}_i) \sim \mathcal{N}(\mu, \sigma_i^2)$ for $i = T, P$. I will distinguish two cases. In the first, labeled equal variance, I will assume that the location parameter for the variance of the log-normal distribution is independent of the shock duration $\sigma^2 = \sigma_T^2 = \sigma_P^2$. In the second, labeled unequal variance, I will assume that the variance for productivity shocks of firms with permanent productivity shocks is smaller than the variance for firms with temporary productivity shocks: $\sigma_T > \sigma_P$. The motivation for considering the unequal variance case stems from the observation that firms with transitory shocks typically require much larger negative shocks to trigger separation, making separations rare. By assigning higher variance to transitory and lower to permanent shocks, the model aligns separation probabilities across both firm types. I follow Stiepelmann (2024) in setting the location parameter for the mean to $\mu = 0.094$ and the variance to $\sigma^2 = 0.12$. For the unequal variance case, I consider $\sigma_T = 0.15$ and $\sigma_P = 0.09$. I set the probability of receiving a transitory productivity shock to 0.5. Further, I set the aggregate productivity to A = 1.5. High aggregate productivity helps to resolve the Costain-Reiter Puzzle (see Costain and Reiter (2008)). In order to generate a realistic policy elasticity, the semi-elasticity of unemployment with respect to an increase in UI benefits must lie between 2 and 3.5. The model produces a semi-elasticity of 3.24, which falls well within this empirically plausible range.

Parameter	Description	Value
Preferences		
ψ	Inverse Frisch-elasticity	1.5
Vacancies, Matching, Bargaining		
k_v	Vacancy posting costs	0.2092
$\bar{m}$	Matching parameter	0.3832
γ	Elasticity matching function w.r.t. unemployment	0.65
η	Bargaining power worker in steady state	0.65
Production and Separations		
α	Labor elasticity production function	0.65
A	Aggregate Productivity	1.5
μ	Stears mean of lognormal distribution	0.094
σ	Stears steering variance of lognormal distribution	0.12
c_f	Strength resource cost shock	1.2620
- unequal variance		
σ_T, σ_P	Stears steering variance of lognormal distribution	0.09, 0.15
c_f	Strength resource cost shock	1.4344
Shock Duration		
p_T	Prob. to get transitory shock	0.5
λ_T	Prob. to receive shock, transitory	1/3
λ_P	Prob. to receive shock, permanent	1/6
Labor Market Policy		
b	UI benefits	0.8493
$\bar{h}$	"Normal" hours worked	1.055

Table 2.3.1. Parameters

Both model versions target job-finding rates of $f = 0.41$ and separation rates of $\rho = 0.03$ from the data in Stiepelmann (2024). The period length is a month. To implement the job-finding rate, I set vacancy posting costs to $k_v = 0.2092$. To implement the separation rate, the strength of the resource cost shock is set to $c_f = 1.2620$ for the model with equal variance and $c_f = 1.4344$ for the model with unequal variance. The matching efficiency parameter $\bar{m} = 0.383$ is determined by targeting a monthly vacancy filling rate of $q = 0.338$. This is the monthly equivalent of the quarterly job-filling rate of 0.71 reported in Haan, Ramey, and Watson

(2000). Firms experiencing transitory shocks are assumed to draw a new productivity shock, on average, every $1/\lambda_T = 3$ months. In contrast, firms with permanent shocks draw new shocks less frequently, on average every $1/\lambda_P = 6$ months. These parameters are calibrated to ensure that, on average, firms remain in the STW (short-time work) system for a duration of 4 to 5 months. This range aligns with empirical findings on average STW durations reported by Brinkmann, Jäger, Kuhn, Saidi, and Wolter (2024).

I set the bargaining power of the worker to $\eta = 0.65$, which is, according to Petrongolo and Pissarides (2001), within the reasonable set of parameter estimates. In order to ensure that inefficiencies in the steady state are only driven by the UI system, the Hosios-Condition (see Hosios, 1990) is implemented by setting the elasticity of the matching function with respect to unemployment equal to the bargaining power of the firm: $\gamma = \eta$. The unemployment benefits are set to $b = 0.8493$, which ensures a replacement rate of 45% of the wage, which is the empirical value reported by Engen and Gruber (2001). The parameter $\bar{h}$ represents the mean hours worked in a firm and is set to its steady state value in the baseline economy: $\bar{h} = 1.055$. Similar to Christoffel and Linzert (2010), I set the labor elasticity of the production function to $\alpha = 0.65$. The disutility of work has the common functional form of $\phi(h) = \frac{h^{1+\psi}}{1+\psi}$, $\psi > 0$. Following Domeij and Floden (2006), I set the Frish-elasticity to 0.66, which implies $\psi = 1.5$.

If not noted differently, the calibration variant with unequal variance is used as reference.

2.4 Results

2.4.1 Jobs at Risk and Shock Persistence Observable

This section examines case 1 in which the government has access to a rich set of information. The government can observe the duration of the productivity shock, job at risk, working hours and time k a worker has received benefits. As a benchmark, I calculate the optimal per-period subsidy $s^*(\lambda)$ that the planner would allocate to matches. This allows us to characterize the optimal separation rates and provides insights into the costs associated with the hours distortion introduced by the STW system. In turn, this helps to assess the relative importance of labor misallocation costs.

Lemma 2
Suppose the government can observe whether a job is at risk $S(\epsilon, \lambda) < 0$, the expected duration $1/\lambda$ of the idiosyncratic productivity shock, working hours h, and the time a worker has received benefits k. Then the optimal subsidy $s^*(\lambda)$ can be written as:

$$s^*(\lambda) \;=\; \underbrace{\frac{\lambda}{f} \cdot b}_{\text{Fiscal Externality UI}}$$

PROOF: Appendix 2.C.1.4

It is important to note that optimal transfers per period are independent of working hours and time spent on the subsidy system. Linking subsidies to working hours can lead to distortions in working hours. Additionally, the geometric distribution is memoryless, meaning that a firm that has spent k periods on the system has the same probability of becoming productive again as a firm that has just joined. This results in the same optimal subsidies per period, independent of the time a worker received benefits. The optimal flow transfers are designed to replace the amount of UI benefits that the marginal worker forfeits when choosing to remain with a firm. This ensures that the UI system does not provide an incentive for workers to enter unemployment, effectively eliminating any inefficient separations. Corollary 1 states that $s^*(\lambda)$ can implement the optimal separation thresholds, meaning that all separations and all worker-retention decisions are efficient.

Corollary 1
1. $s^*(\lambda)$ *implements the optimal separation thresholds*
2. *Firms with permanent productivity shocks get smaller transfers*
3. $s^*(\lambda) > b$ *if* $\lambda > f$; $\quad s^*(\lambda) = b$ *if* $\lambda = f$; $\quad s^*(\lambda) < b$ *if* $\lambda < f$

PROOF: Appendix 2.C.1.4

A key feature of the optimal flow transfer is that it decreases with respect to the expected shock duration $1/\lambda$ (as observed in Corollary 1). This implies that the Ramsey planner aims to allocate higher transfers per period to matches with transitory shocks compared to those with permanent shocks. The rationale behind this is straightforward: since workers are identical and face the same outside option in terms of UI benefits and job-finding rates, the government needs to implement the same expected transfer with the subsidy system, irrespective of the persistence of productivity shocks. However, matches with permanent productivity shocks are expected to stay on the subsidy system for longer compared to matches with transitory shocks. To achieve the same expected transfer amount, the government needs to select smaller transfers per period for matches with permanent shocks.

Note that we could also write the equation above as

$$T_{Sub} \cdot s^*(\lambda) = T_{UI} \cdot b \quad \text{with} \quad T_{Sub} = \frac{1}{\lambda},\ T_{UI} = \frac{1}{f}$$

so that the left-hand side of the equation denotes the expected benefits that the worker receives when staying employed, while the right-hand side of the equation denotes the expected UI benefits a worker would receive when becoming unemployed. Here, T_{Sub} denotes the time a worker spends on the subsidy system, while T_{UI} denotes the expected time a worker spends on the UI until he finds a new job.

Another noteworthy observation is that when $\lambda = f$, the optimal flow transfers should equal the UI benefits. In this scenario, it is equally likely for a worker to return to a productive state by remaining with the firm as it is by finding a new job through unemployment. If matches are characterized by a transitory shock such that $\lambda > f$, a worker is more likely to revert to a productive state by staying with the same firm rather than searching for a new job through unemployment. Consequently, the government focuses on protecting these matches by selecting a flow subsidy greater than UI benefits. Conversely, if workers experience a permanent productivity shock such that $\lambda < f$, the Ramsey planner knows that it is unlikely for the worker to return to a productive state. In this case, transitioning to unemployment enables a faster transition to a productive state. Consequently, the planner opts for smaller transfers than UI benefits to promote separations and facilitate the reallocation of workers in the labor market.

Next, we examine how optimal STW benefits should be characterized when the government can observe the persistence of a firm's productivity shock. Lemma 3 establishes the optimal design of STW benefits in this setting. As in the case with the simple subsidy, the government adjusts STW benefits so that the total resources transferred to a marginal match via STW equal the UI benefits the worker would have received upon separation.

Definition 3 Let τ_{stw}^{net} *denote the net-transfer of resources to a match on STW:*

$$\tau_{stw}^{net} = \tau_{stw} \cdot (\bar{h} - h_{stw})$$

Note that the net transfer per period — defined in Definition 3 — depends on the extent to which firms and workers are willing to reduce working hours. The greater the reduction in hours, the smaller h_{stw}, the more resources are transferred to the match for a given level of STW benefits. The government must consider this behavioral response when determining optimal STW benefits.

Lemma 3
Suppose the government can observe whether a job is at risk $S(\epsilon,\lambda)<0$, the expected duration $1/\lambda$ of the idiosyncratic productivity shock and the time a worker has received benefits k, then optimal STW benefits $\tau_{stw}(\lambda)$ can be written as:

$$\tau_{stw}(\lambda)\cdot(\bar{h}-h_{stw}(R(\lambda))) = \underbrace{\frac{\lambda}{f}\cdot b}_{\text{Fiscal Externality UI}} - \underbrace{\left[n_{stw}(\lambda)/\frac{\partial n(\lambda)}{\partial \tau_{stw}}\right]\cdot\frac{\partial\tilde{\Omega}(\lambda)}{\partial \tau_{stw}}}_{\text{Welfare Costs Rescuing an addtional Worker with STW in a Firm with Shock Duration } 1/\lambda}$$

where the marginal number of matches of type λ preserved by STW benefits is

$$\frac{\partial n(\lambda)}{\partial \tau_{stw}} = p(\lambda)\cdot\frac{g_\lambda(R_{stw}(\lambda))}{\lambda}\cdot\left(-\frac{\partial R_{stw}(\lambda)}{\partial \tau_{stw}}\right)\cdot n^s$$

PROOF: Appendix 2.C.1.5

The key difference between a simple subsidy or transfer and STW is that STW connects the transfer to hours reduction. The higher the STW benefits, the higher the incentives for firms and workers to reduce working hours to draw more government funding into the match. Therefore, the government needs to take into account its impact on working hours. To keep distortions low, it adjusts STW benefits downwards.

Note that the welfare costs of STW depend on two key factors. First, $\frac{\partial\tilde{\Omega}(\lambda)}{\partial \tau_{stw}}$ captures the marginal distortion to working hours caused by an increase in STW benefits. Second, the ratio $n_{stw}(\lambda)/\frac{\partial n(\lambda)}{\partial \tau_{stw}}$ measures how many firms choice get distorted on the STW system relative to the number of additional matches saved by increasing benefits. If this ratio is high — that is, if relatively few matches are rescued compared to the number of distorted firms — then increasing STW benefits is particularly costly.

2.4.2 Jobs at Risk but not Shock Persistence Observable

Governments often face difficulties in accurately assessing the persistence of shocks faced by individual firms. This section analyzes the optimal level of STW benefits under the assumption that the government cannot distinguish between transitory and permanent shocks.

2.4.2.1 Uniform Systems

For simplicity, this section analyzes the optimal subsidy and short-time work (STW) benefits under the assumption that the government does not condition policy instruments on the duration of a worker's participation in the system. Since the policy instruments are invariant to both shock persistence (due to information asymmetry) and the time spent in the system, I refer to these as *uniform instruments*. This assumption will be relaxed in later sections. As a reference point, I begin with the subsidy system. Lemma 4 presents the optimal uniform subsidy in the case where shock persistence is unobservable.

Lemma 4

Suppose the government can observe whether a job is at risk $S(\epsilon, \lambda) < 0$ but cannot observe the expected duration $1/\lambda$ of the idiosyncratic productivity shock, then the optimal subsidy s_ω can be written as

$$s_\omega = E_\omega[s^*(\lambda)]$$

with weights

$$\omega(\lambda) = \frac{\frac{\partial n(\lambda)}{\partial s_\omega}}{\int_0^1 \frac{\partial n(\lambda)}{\partial s_\omega} d\lambda}$$

where the marginal number of matches of type λ preserved by the subsidy is

$$\frac{\partial n(\lambda)}{\partial s_\omega} = n^s \cdot p(\lambda) \cdot \frac{g_\lambda(R(\lambda))}{\lambda} \cdot \frac{\partial R(\lambda)}{\partial s_\omega}$$

PROOF: Appendix 2.C.1.6

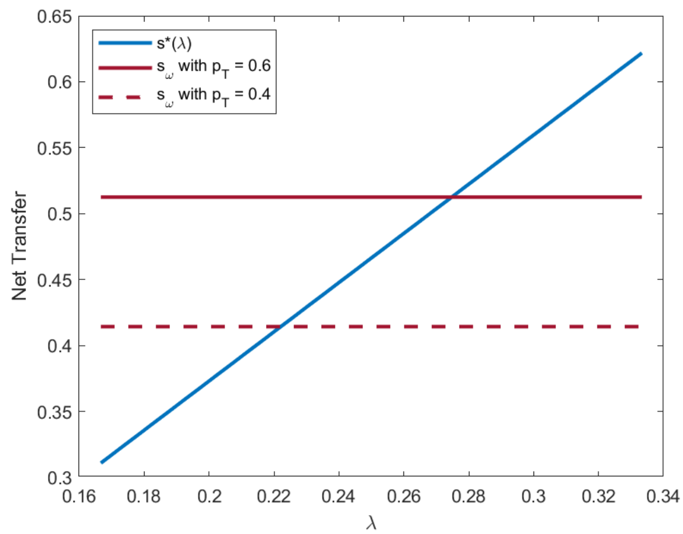
When the government cannot observe the persistence of a firm's productivity shock, it faces a fundamental trade-off in assigning subsidies: On the one hand, it would like to provide generous benefits to firms experiencing transitory shocks, as their probability of recovery is high. On the other hand, it aims to limit support to firms facing permanent shocks, where the likelihood of recovery is low.

For my calibration, this trade-off is illustrated clearly in Figure 2.4.1. The blue line shows the optimal transfer schedule $s^*(\lambda)$, which implements the optimal separation threshold for each probability of recovery. In contrast, a uniform subsidy s_w fails to tailor benefits to firm-specific conditions: it over-subsidizes firms with low chances of recovery ($s_\omega > s^*(\lambda)$), thereby keeping workers in unproductive matches that should be reallocated. At the same time, it under-subsidizes firms with high recovery potential ($s_\omega < s^*(\lambda)$), leading to the unnecessary destruction of matches that the social planner would prefer to preserve.

Both types of misallocation — excess retention and premature separation — result in an inefficient allocation of labor across the economy.

The government aims to minimize welfare losses stemming from both excessive retention and premature separation. To do so, it constructs a weighted average of type-specific optimal transfers. Greater weight is assigned to firms with shock persistence $1/\lambda$ if an increase in the uniform subsidy preserves more matches of that type. This is captured by the derivative $\frac{\partial n(\lambda)}{\partial s_\omega}$. The logic is straightforward: when more matches of a particular type are sensitive to the subsidy, the welfare cost of over- or under-subsidizing that group becomes larger, warranting greater consideration in the policy design.

Figure 2.4.1. Optimal Subsidy with and without Information Asymmetry



Notes: The figure compares the optimal subsidy function without information asymmetry regarding shock persistence ($s^*(\lambda)$, blue line) to the optimal uniform subsidy under asymmetric information (s_ω, red lines). The solid orange line represents an economy with a high share of transitory shocks; the dashed orange line reflects a shift toward an economy with a higher share of permanent shocks.

These weights are particularly sensitive to the distribution of firms across different levels of shock persistence. Figure 2.4.1 illustrates that an increase in the relative number of firms experiencing permanent productivity shocks leads the government to choose a lower optimal transfer. The intuition is straightforward: when a larger share of firms faces permanent productivity shocks, the cost of retaining workers in unproductive jobs becomes higher. As a result, the government reduces benefits to foster reallocation of workers via the labor market.

This mechanism is particularly relevant during periods of structural change. Suppose certain sectors of the economy face declining demand or obsolete business models, while others are expanding. From the model's perspective, this corresponds to an increase in the number of firms experiencing persistent shocks, both positive and negative. In such a scenario, the model suggests that reducing worker retention incentives can facilitate the reallocation of labor toward more productive uses.

Proposition 1 looks at the optimal STW benefits under the assumption that the government cannot distinguish between transitory and permanent shocks. In general, the government faces a similar problem to that of a uniform subsidy. It might over- or under-subsidize firms, leading to excess retention or premature separations. In contrast to the uniform subsidy, however, it is ex-ante not clear which firms STW over- or under-subsidizes: firms with transitory or firms with permanent shocks. This is for two reasons:

First, it is not ex-ante clear whether the optimal net transfer, $\tau^{\text{net}}_{\text{stw}}(\lambda)$, increases or decreases with the persistence of the shock. On the one hand, the government has an incentive to subsidize firms with a low probability of recovery less generously to foster reallocation over the labor market. On the other hand, rescuing firms with transitory shocks may be more costly, as suggested by the ratio of workers on STW to the marginal effect of the subsidy on firm rescue, $n_{\text{stw}}(\lambda) \Big/ \frac{\partial n(\lambda)}{\partial \tau_{\text{stw}}}$. In particular, when the probability of recovery is high, firms must be hit by relatively large negative productivity shocks to consider laying off workers. Since such shocks might be rare, it might be more costly to rescue firms facing transitory shocks than those facing more persistent ones. The blue line in Figure 2.4.2, however, shows that the misallocation effect of workers dominates quantitatively in our calibration.

Second, STW links subsidies to reductions in working hours. As a result, it effectively provides more generous support to firms experiencing transitory productivity shocks than to those facing permanent shocks. This makes it unclear whether the subsidy schedule implied by STW is steeper in the probability of recovery than the optimal net transfer. The underlying mechanism is as follows: firms facing transitory productivity shocks are willing to retain workers even at relatively low productivity levels, whereas firms facing permanent shocks will separate for higher productivity levels:

$$R(\lambda) \; < \; R(\lambda') \quad \text{iff} \quad \lambda \; > \; \lambda'$$

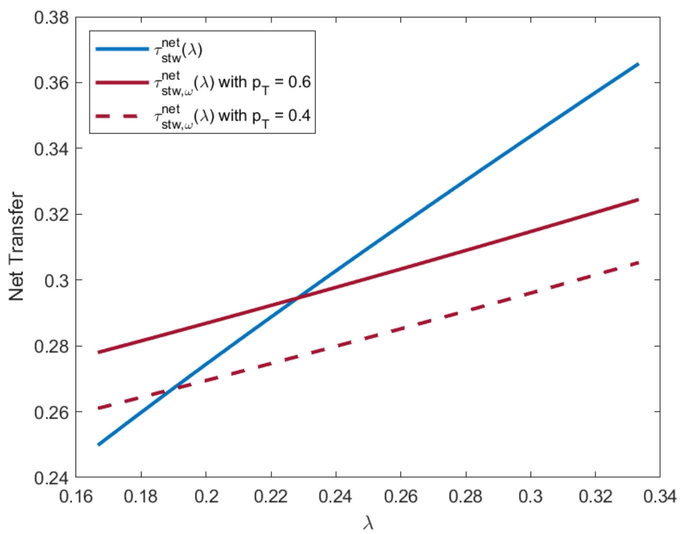
As a result, firms with transitory productivity shocks will be willing to reduce working hours more before separating from a worker than firms with permanent productivity shocks:

$$h_{stw}(R(\lambda)) < h_{stw}(R(\lambda')) \quad \text{iff} \quad \lambda > \lambda'$$

As a result, firms facing transitory productivity shocks automatically receive higher transfers through the STW system than firms facing permanent shocks. Figure 2.4.2 illustrates this effect. The orange line represents the net subsidy under the optimal STW scheme when the government cannot distinguish between transitory and permanent shocks. In this case, the STW system implicitly subsidizes firms with a higher probability of recovery more than those with a lower probability of recovery.

However, the resulting subsidy function is flatter than the optimal subsidy schedule that would be implemented if the government could observe shock persistence. Consequently, the STW system still under-subsidizes firms with transitory shocks and over-subsidizes firms with permanent shocks. Nonetheless, the distortion is less severe compared to a uniform subsidy scheme.

Figure 2.4.2. Optimal Net Transfer STW with and without Information Asymmetry



Notes: The figure compares the optimal net-transfers imposed by the optimal STW benefit function without information asymmetry regarding shock persistence ($\tau^{net}_{stw}(\lambda) = \tau_{stw}(\lambda) \cdot (\bar{h} - h_{stw}(R(\lambda)))$, blue line) to the optimal net-transfers imposed by the optimal uniform STW benefits under asymmetric information ($\tau^{net}_{stw,\omega}(\lambda) = \tau_{stw,\omega} \cdot (\bar{h} - h_{stw}(R(\lambda)))$, red lines). The solid orange line represents an economy with a high share of transitory shocks; the dashed orange line reflects a shift toward an economy with a higher share of permanent shocks.

Proposition 1
Suppose the government can observe whether a job is at risk but not the expected duration $1/\lambda$ of the idiosyncratic productivity shock, then optimal STW benefits $\tau_{stw,\omega}$ can be characterized by

$$\underbrace{\tau_{stw,\omega} \cdot E_{\omega^{stw}}\Big[\bar{h} - h_{stw}(R(\lambda))\Big]}_{\substack{\text{Weighted Net Transfer from optimal STW Benefits} \\ \text{with Asymmetric Information}}} = \underbrace{E_{\omega^{stw}}\Big[\tau_{stw}(\lambda) \cdot (\bar{h} - h_{stw}(R(\lambda)))\Big]}_{\substack{\text{Weighted Net Transfer from optimal STW Benefits} \\ \text{without Asymmetric Information}}}$$

with weights

$$\omega^{stw}(\lambda) = \frac{\frac{\partial n(\lambda)}{\partial \tau_{stw,\omega}}}{\int_0^1 \frac{\partial n(\lambda)}{\partial \tau_{stw,\omega}} d\lambda}$$

PROOF: Appendix 2.C.1.7

Proposition 1 characterizes how the government can minimize the welfare costs arising from both excess retention and premature separation. Specifically, the government assigns more weight to firms with a given shock persistence if an increase in STW benefits leads to a greater number of matches being preserved for that type. In such cases, over- or under-subsidizing these matches becomes more costly. To minimize the cost of worker misallocation, Proposition 1 suggests setting the expected transfer under a short-time work system, when subsidies cannot be conditioned on shock persistence, equal to that under a system in which such conditioning is possible.

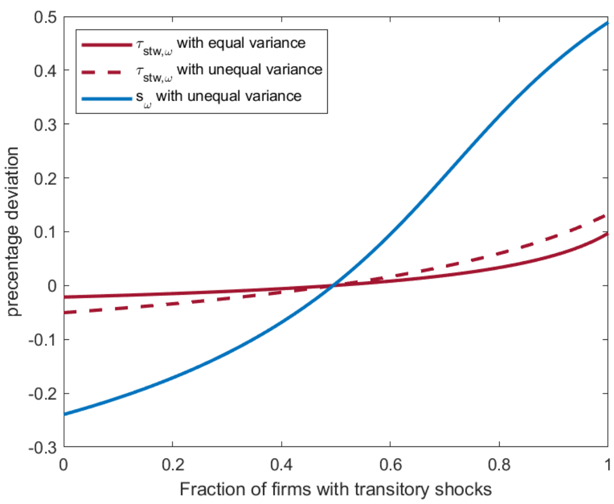
Next, we can examine how STW should respond when the distribution of firms with transitory to permanent shocks shifts. Figure 2.4.2 shows an increase in the relative number of firms with transitory shocks. Similarly to a system with uniform subsidies, STW benefits should be reduced as more weight is applied to the costs of keeping workers in unproductive occupations. Likewise, the model would suggest reducing STW benefits in times of structural change. Finally, we compare the sensitivity of the optimal uniform subsidy and STW benefits to shifts in the distribution of firms toward transitory or permanent shocks. Figure 2.4.3 illustrates this comparison: the orange lines represent the optimal STW benefits, while the blue line depicts the optimal uniform subsidy, both plotted against the share of firms experiencing transitory shocks.

A key observation is that STW benefits respond far less to changes in the share of transitory firms than the uniform subsidy does. This difference in responsiveness has important implications discussed in a later section and is driven by two main factors:

First, as previously discussed, the STW system inherently targets firms with transitory shocks with larger subsidies. Because of this, the misallocation costs associated with a fixed STW benefit are smaller, and less adjustment is required in response to shifts in the firm distribution.

Second, the distortionary costs of rescuing firms with transitory shocks tend to be higher. These firms typically require larger productivity shocks before they would separate from a worker. Such shocks may be relatively rare. As a result, the ratio of distorted firms to additional matches saved — expressed as $n_{\text{stw}}(\lambda) \Big/ \frac{\partial n(\lambda)}{\partial \tau_{\text{stw}}}$—can be large, making benefit increases less efficient.

Figure 2.4.3. Responsiveness of Subsidy and STW Benefits to Fraction of Transitory Firms



Notes: The blue line shows the optimal uniform subsidy under information asymmetry regarding shock persistence s_ω in response to the fraction of firms with transitory shocks in the economy. The red lines show the optimal uniform STW benefits under information asymmetry $\tau_{stw,\omega}$ with respect to the fraction of firms with transitory shocks. The solid red line assumes that the variance of the productivity shock is the same for firms with transitory and permanent productivity shocks. The dashed line assumes the variance is greater for firms with transitory productivity shocks. All three cases are expressed as percentage deviations from the optimal benefits in an economy where the fraction of firms with a transitory shock is $p_T = 0.5$.

This point is illustrated in Figure 3 by comparing the solid and dashed orange lines. In the dashed scenario, the variance of productivity shocks — conditional on shock persistence — is adjusted so that the separation probabilities of firms with transitory and permanent shocks become more similar. This adjustment reduces the cost associated with rescuing transitory matches, thereby increasing the responsiveness of optimal STW benefits to the distribution of shock persistence. As the figure shows, under this condition, STW benefits become more sensitive to changes in the share of transitory firms in the economy.

2.4.2.2 Time dependent Systems

The last section abstracted from the possibility of depending STW benefits on the time a worker has received them. This section adds the dimension.

The intuition for setting the optimal uniform subsidy — when the government cannot distinguish between transitory and permanent shocks but can condition benefits on the time a worker spends in the system — is closely related to Lemma 5. As before, the goal is to minimize misallocation arising from excess retention and premature separation. This is achieved by placing greater weight on shock persistence levels where subsidies save more matches.

Lemma 5
Suppose the government can observe whether a job is at risk $S(\epsilon,\lambda) < 0$ but cannot observe the expected duration $1/\lambda$ of the idiosyncratic productivity shock. Further, suppose it can base the benefits on the duration of the STW system k. Then the optimal subsidy s_{ω_k} dependent on time spent on the STW system k can be written as

$$s_{\omega_k} = E_{\omega_k}[s^*(\lambda)]$$

with weights

$$\omega_k(\lambda) = \frac{\frac{\partial n_k(\lambda,k)}{\partial s_{\omega_k}}}{\int_0^1 \frac{\partial n(\lambda,k)}{\partial s_{\omega_k}} d\lambda}$$

where the marginal number of matches of type (λ, k) preserved by the subsidy is

$$\frac{\partial n(\lambda,k)}{\partial s_{\omega_k}} = (1-\lambda)^{k-1} \cdot n^s \cdot p(\lambda) \cdot g_\lambda(R(\lambda,k)) \cdot \frac{\partial R(\lambda,k)}{\partial s_{\omega_k}}$$

PROOF: Appendix 2.C.1.8

The key difference in this setting is that the relative composition of transitory and permanent firms evolves over time. Unproductive firms with transitory shocks have a higher likelihood of becoming productive again than those with permanent shocks, and therefore tend to exit the subsidy system more quickly. As a result, the share of firms with transitory shocks declines with the duration a worker receives the subsidy. From the formula for the weights, this result can be seen by recognizing that

$$(1-\lambda)^{k-1} \ < \ (1-\lambda')^{k-1} \quad \textit{iff} \quad \lambda \ > \ \lambda'$$

The government anticipates this dynamic. Over time, it places increasing weight on firms with permanent shocks — those more likely to remain in the system. To limit the welfare costs of keeping workers in persistently unproductive matches, the government reduces the uniform subsidy as the duration of time spent on the subsidy system increases.

This result is formally established in Corollary 3. Figure 2.4.4 illustrates the quantitative response of optimal transfers to the evolving composition of firms over time on the subsidy system.

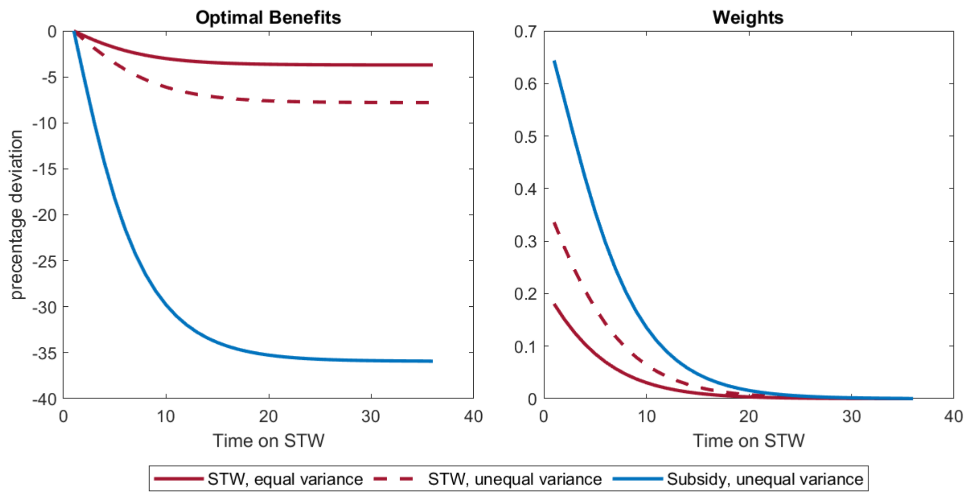
Corollary 2 s_{ω_k} falls with k

Note that, in contrast to many real-world subsidy programs such as STW or the Payment Protection Program in the US, my findings suggest that the subsidy should not be terminated after a fixed duration. The memoryless property of the geometric distribution drives this result. Although firms with permanent shocks increasingly dominate the pool of firms as time on the subsidy system progresses, those that remain are still just as likely to become productive again as newly entering firms with permanent shocks. This time-invariant probability of recovery justifies continuing to provide positive subsidies, even for firms that have been in the system for a long time.

Proposition 2 characterizes how optimal STW benefits should be adjusted when the government cannot observe whether a firm is experiencing a transitory or permanent productivity shock, but can condition STW benefits on the duration of time a firm has spent in the program. The logic behind setting optimal STW benefits in this context hinges on whether the optimal net-transfer function is increasing or decreasing with respect to time in the system. This, in turn, depends on the relative cost of rescuing transitory matches: if transitory matches are expensive to support due to low responsiveness to subsidies, the optimal benefit path may decline over time; if they are relatively cheap to preserve, benefits increase with time.

When the optimal net transfers with STW increase, the theory is straightforward. For every time k spent on the STW system, the government sets the expected optimal STW transfers without distinguishing between shock persistence equal to the optimal expected transfers of a system that can observe the shock persistence. Again, the government takes into account that the relative number of firms will shift towards firms with permanent productivity shocks. To foster reallocation of workers, the government reduces benefits.

Figure 2.4.4. Optimal Benefits and Time Spent on the System



Notes: The blue line represents the subsidy system when there exists information asymmetry over the shock duration. The red lines represent the STW system when there exists information asymmetry. The solid red line represents the system when firms with transitory and permanent shocks have the same variance in the productivity shocks. The dashed red line considers the case when the variance for productivity shocks in firms with transitory shocks is larger than in firms with permanent shocks. Optimal benefits are expressed as percentage deviations from period 1 benefit levels. Weights are reported as absolute weights.

When the distortionary costs of rescuing firms with transitory shocks dominate, the logic for setting STW benefits changes. If STW benefits are increasing over time, firms and workers will only consider separating in the first period. Once a firm remains on the system beyond the initial period, the rising value of future STW benefits provides an increasing incentive to stay matched.

In this setting, the government's objective shifts toward minimizing the distortionary impact of STW on firm behavior. Since the cost of distorting firms with permanent productivity shocks is lower — due to their better ratio of firms on STW to firms that can be rescued with STW — it becomes optimal to increase STW benefits over time. Doing so concentrates distortions into periods with a larger relative fraction of firms with permanent productivity shocks, thus making the distortions less costly.

Proposition 2

Suppose the government can observe whether a job is at risk $S(\epsilon, \lambda) < 0$ but cannot observe the expected duration $1/\lambda$ of the idiosyncratic productivity shock. Further, suppose it can base the benefits on the time k a worker has spent on STW. If the net-transfer of the system rises, then the optimal benefits can be expressed as

$$\underbrace{\tau_{stw}(k) \cdot E_{\omega_k^{stw}}\Big[\bar{h} - h_{stw,k}(R(\lambda,k))\Big]}_{\substack{\text{Weighted Net Transfer from optimal STW Benefits} \\ \text{in period k on STW under Asymmetric Information}}} = \underbrace{E_{\omega_k^{stw}}\Big[\tau_{stw}(\lambda,k) \cdot (\bar{h} - h_{stw,k}(R(\lambda,k)))\Big]}_{\substack{\text{Weighted Net Transfer from optimal STW Benefits} \\ \text{in period k on STW without Asymmetric Information}}}$$

with weights

$$\omega_k^{stw}(\lambda) = \frac{\frac{\partial n(\lambda,k)}{\partial \tau_{stw}(k)}}{\int_0^1 \frac{\partial n(\lambda,k)}{\partial \tau_{stw}(k)} d\lambda}$$

where the marginal number of matches of type (λ, k) preserved by the benefits is

$$\frac{\partial n(\lambda,k)}{\partial \tau_{stw}(k)} = (1-\lambda)^{k-1} \cdot n^s \cdot p(\lambda) \cdot \frac{g_\lambda(R(\lambda,k))}{\lambda} \cdot \frac{\partial R(\lambda,k)}{\partial \tau_{stw}(k)}$$

If the net subsidy of the system decreases with the duration of STW benefits, then the STW benefits are determined by:

$$\underbrace{E_{\omega_k^{stw}}\left[\left(n_{stw}(\lambda,k) / \frac{\partial n(\lambda,k)}{\partial \tau_{stw}}\right) \cdot \frac{\partial \tilde{\Omega}(\lambda,k)}{\partial \tau_{stw}}\right]}_{\substack{\text{Weighted Average Welfare Costs of rescuing Marginal Worker with STW} \\ \text{in period k on STW}}} = \underbrace{E_{\omega_k^{stw}}\Big[LS(\lambda)\Big]}_{\substack{\text{Weighted Average Social Value Marginal Match} \\ \text{in period k on STW}}}$$

where $LS(\lambda)$ denotes the social value of the marginal job with shock persistence $1/\lambda$

$$LS(\lambda) = \frac{\lambda}{f} \cdot b - \sum_{k=1}^{\infty} \lambda \cdot (1-\lambda)^{k-1} \cdot \tau(k) \cdot (\bar{h} - h_{stw,k}(R(\lambda,k)))$$

PROOF: Appendix 2.C.2.1

Figure 2.4.4 illustrates the optimal STW benefits over time, along with the corresponding weights the government assigns to different types of productivity shocks. The left panel shows that, in all cases, optimal STW benefits decline with time spent on the system — reflecting the dominant role of the worker misallocation motive. The right panel explains this decline: it is driven by the falling ratio of firms with transitory shocks to those with permanent shocks. Firms with transitory shocks are more likely to exit the system early due to their higher recovery rates.

Importantly, however, STW benefits should decline more gradually than under a uniform subsidy system. This is due to two countervailing factors. First, as discussed earlier, STW systems inherently allocate more support to transitory firms, reducing the need for a steep drop in benefits over time. Second, the number of transitory firms receiving productivity shocks large enough to trigger separation may be small. This increases the cost of rescuing matches via STW and lowers the planner's incentive to assign high weight to matches with transitory shocks. As a result, the relative share of transitory to permanent shocks changes only modestly over time.

Together, these factors imply a more moderate decline in STW benefits as time on the system increases. This is clearly illustrated in Figure 2.4.4: when shocks that lead to separation in temporarily unproductive firms become more likely, the planner places greater weight on transitory firms, leading to a sharper decline in benefits. In contrast, when separations remain rare among transitory firms, the decline in benefits is more gradual, reflecting the higher cost of distortions and reduced importance of transitory matches in the planner's objective.

2.4.3 Only working hours observable

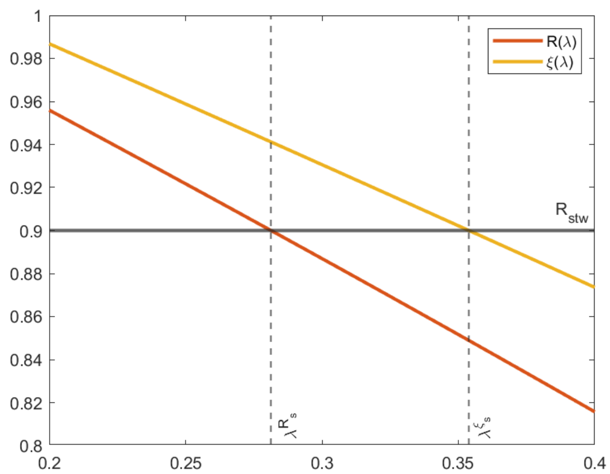
In the previous sections, we assumed that the government knows exactly which firms would separate from a worker and which would not. In practice, governments attempt to elicit this information by examining firms' financial records. But what if this information is not reliably extractable? This section explores how the government can still screen for jobs at risk under such uncertainty. Again, we abstract from the time spent on the STW system at first.

STW systems include a minimum working hours reduction threshold that firms must meet to qualify for support. By requiring a reduction in hours, the government can indirectly screen for a firm's current productivity. Since workers suffer disutility from work, firms and workers will always choose to reduce working hours more when productivity is lower. This allows the government to screen out a firm's productivity from observed reductions in hours. The productivity threshold for eligibility can therefore be characterized by:

$$h_{stw}(R_{stw}) = D$$

Effectively, firms and workers can enter the STW scheme once their idiosyncratic productivity, ϵ, falls below a threshold R_{stw}, regardless of the persistence of the shock the firm is experiencing. This design, however, makes it difficult to effectively screen for a job at risk. The issue is illustrated in Figure 2.4.5. The yellow line depicts the separation thresholds for firms without access to STW benefits, depending on their probability of recovery.

Figure 2.4.5. Separation Thresholds with and without STW



Notes: The orange line shows the separation threshold for matches receiving STW support, conditional on the match's recovery probability. The yellow line depicts the corresponding threshold for matches without access to STW. The area between the two curves represents all (ε, λ) combinations in which STW prevents separations that would otherwise occur. The gray line illustrates the threshold imposed by a minimum hours reduction requirement for eligibility, allowing only firms with sufficiently low productivity to enter the STW system.

Ideally, firms should be granted access to STW only if, in the absence of support, they would separate from their workers. However, firms with a high probability of recovery are often willing to retain workers even at lower productivity levels, while those with low probability of recovery want to separate for higher productivity levels. Since eligibility for STW is based on a single productivity threshold, the government faces a problem: firms with permanent productivity shocks that could have been saved with STW are excluded from the program, while firms with transitory shocks that would not have separated in the first place still qualify and receive support. Figure 2.4.5 highlights this problem.

Proposition 3
Suppose the government can only observe the working hours of a firm, and suppose for simplicity that $P(\lambda)$ is continuous. Then the optimal eligibility condition $D = h_{stw}(\epsilon_{stw})$ is implicitly defined by

$$p_{stw} \cdot \underbrace{E_{\omega^D}\left[SW(\epsilon,\lambda)\middle|\lambda^{\xi_s}(R_{stw}) \geq \lambda \geq \lambda^R(R_{stw})\right]}_{\substack{\text{Average Social Welfare Gain from Preserving a Match}\\ \text{with low Recovery Probability}}}$$

$$= p_f \cdot \underbrace{E_{\omega^D}\left[\left(n(\lambda)/\frac{\partial n(\lambda)}{\partial R_{stw}}\right) \cdot \frac{\partial \Omega(\lambda)}{\partial R_{stw}}\middle|\lambda \geq \lambda^{\xi}(R_{stw})\right]}_{\substack{\text{Average Social Costs from distorting Hours in a Match}\\ \text{with high Recovery Probability}}}$$

Where $SW(\epsilon,\lambda)$ denotes the social value of the rescued match

$$SW(\epsilon,\lambda) = \underbrace{\frac{\lambda}{f} \cdot b - \tau^{net}_{stw}(\lambda)}_{\text{Social Value of the Marginal Match } SW(R(\lambda),\lambda)} + \underbrace{z_{stw}(\epsilon) - z_{stw}(R(\lambda))}_{\text{Extra Welfare through Productivity}}$$

$\lambda^R(R_{stw})$ denotes the recovery probability below which firms are screened out from STW and $\lambda^{\xi}(R_{stw})$ denotes the recovery probability above which firms go onto STW without the need of STW support:

$$\lambda^R(R_{stw}) = R^{-1}(R_{stw}) \qquad \lambda^{\xi}(R_{stw}) = \xi^{-1}(R_{stw})$$

p_{stw} denotes the marginal fraction of firms that enter the STW system and are in need of STW support, and p_f denotes the marginal fraction of firms that enter the STW system and do not need STW support

$$p_{stw} = P(\lambda^{\xi}(R_{stw})) - P(\lambda^R(R_{stw})) \qquad p_f = 1 - P(\lambda^{\xi}(R_{stw}))$$

and weights are defined as

$$\omega^D(\lambda) = \frac{\frac{\partial n(\lambda)}{\partial R_{stw}}}{\int_0^1 \frac{\partial n(\lambda)}{\partial R_{stw}} d\lambda}$$

where the marginal matches affected by the STW threshold are

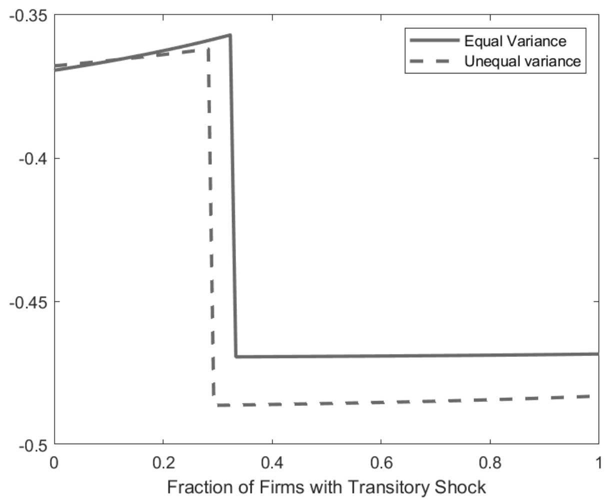
$$\frac{\partial n(\lambda)}{\partial R_{stw}} = n^s \cdot p(\lambda) \cdot \frac{g_{\lambda}(R_{stw})}{\lambda}$$

PROOF: Appendix 2.C.4.3

Proposition 3 characterizes the optimal setting of the eligibility threshold. For simplicity, it is assumed that the distribution, about $P(\lambda)$, is continuous. The government faces a trade-off. By loosening the eligibility criterion, it can prevent the dissolution of matches with productivity R_{stw} and recovery probabilities $\lambda \in [\lambda^R, \lambda^\xi]$. However, this also allows firms with the same productivity level but higher recovery probabilities $\lambda > \lambda^\xi$ to enter the program, even though they would have retained their workers without support, leading to an unnecessary distortion of working hours.

The optimal eligibility condition sets the social value of rescuing additional matches with a looser eligibility condition equal to the additional distortions it introduces into the economy.

Figure 2.4.6. Minimum Hours Reduction Threshold (Eligibility Condition)



Notes: The figure illustrates the optimal eligibility condition as a function of the share of workers on STW. Eligibility is here defined by the required reduction in working hours relative to normal hours, given by $\frac{D-\bar{h}}{\bar{h}}$, where D denotes the maximum working hours allowed under STW and $\bar{h}$ represents normal working hours. The solid line looks at an economy where the variance of the productivity shocks is the same for firms with transitory and permanent shocks while the dashed line looks at an economy where the variance of the productivity shock for firms with transitory shocks is higher than with permanent.

Figure 2.4.6 illustrates the optimal eligibility condition in an economy with two types of firms: those experiencing permanent productivity shocks and those facing transitory shocks. When the share of transitory firms is small, the government opts for a looser eligibility threshold. This allows it to rescue all firms with permanent productivity shocks, maximizing the welfare gains from preserving these matches. Although some firms with transitory shocks may enter the system without needing support, leading to windfall effects, the associated welfare loss remains limited as long as the share of transitory firms is low.

As the fraction of firms with transitory shocks increases, however, the welfare costs from these windfall effects rise, while the absolute gains from rescuing firms with permanent shocks diminish. Eventually, the welfare losses from inefficiently supporting transitory firms outweigh the benefits of preserving permanent matches. At this point, the government tightens the eligibility condition, effectively reducing windfall effects from transitory shocks, but at the cost of excluding many firms with permanent productivity shocks from receiving support.

Proposition 4 characterizes how the optimal benefit level should be set under the new eligibility condition. It mirrors the condition in Proposition 1, with one key difference: the planner now accounts for the fact that firms with a recovery probability below $\lambda < \lambda^R$ are effectively screened out by the eligibility threshold.

Proposition 4
Suppose the government can observe whether a job is at risk but not the expected duration $1/\lambda$ of the idiosyncratic productivity shock, then optimal STW benefits τ_{stw} can be written as:

$$\underbrace{\tau_{stw,\omega} \cdot E_{\omega^{stw}}\left[\bar{h} - h_{stw}(R(\lambda))\middle|\lambda \geq \lambda^R(R_{stw})\right]}_{\substack{\text{Weighted Net Transfer from optimal STW Benefits} \\ \text{on STW under Asymmetric Information}}} = \underbrace{E_{\omega^{stw}}\left[\tau_{stw}(\lambda) \cdot (\bar{h} - h_{stw}(R(\lambda)))\middle|\lambda \geq \lambda^R(R_{stw})\right]}_{\substack{\text{Weighted Net Transfer from optimal STW Benefits} \\ \text{on STW without Asymmetric Information}}}$$

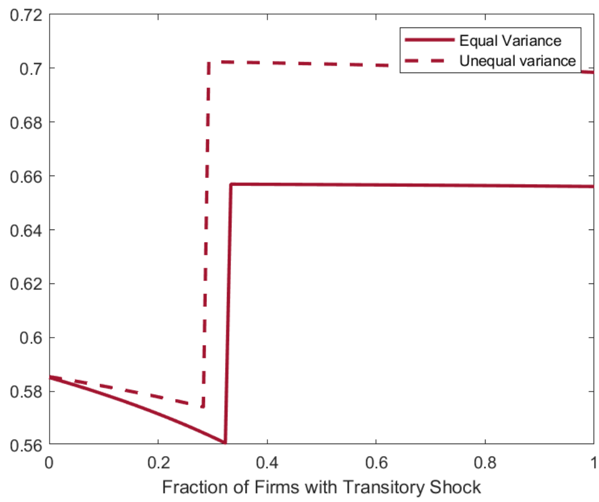
PROOF: Appendix 2.C.4.4

Figure 2.4.7 plots the optimal short-time work (STW) benefits as a function of the share of firms facing transitory productivity shocks in the economy. When the fraction of firms with permanent shocks is high, the government sets a loose eligibility condition that includes all such firms. In this case, optimal benefits are relatively low to encourage the reallocation of workers through the labor market.

Surprisingly, as the share of firms with transitory shocks increases, optimal STW benefits decline further. This is due to growing windfall effects: a rising number of transitory-shock firms enter the system despite not needing support to survive. As a result, the marginal effectiveness of additional STW benefits diminishes —

the number of firms that can be saved relative to those that get distorted becomes smaller. This is reflected in an increase in the ratio $\left[n_{\text{stw}}(\lambda)/\frac{\partial n(\lambda)}{\partial \tau_{\text{stw}}}\right]$. In response, the planner reduces STW benefits to limit the distortionary effects of the STW benefits.

Figure 2.4.7. Optimal STW Benefits



Notes: The figure depicts the optimal uniform STW benefits $\tau_{stw,\omega}$ when there exists information asymmetry about shock persistence and the government chooses the minimum hours reduction condition as eligibility criterion. The solid line looks at an economy where the variance of the productivity shocks is the same for firms with transitory and permanent shocks, while the dashed line looks at an economy where the variance of the productivity shock for firms with transitory shocks is higher than with permanent.

Once the share of firms with transitory shocks becomes sufficiently large, the planner chooses to tighten the eligibility condition in order to curb windfall effects. This policy shift leads to a sharp increase in optimal STW benefits, driven by two main factors. First, the tighter eligibility criterion significantly reduces the number of firms participating in the STW system, thereby lowering the distortionary cost of providing more generous benefits. Second, the stricter threshold effectively screens out permanently unproductive firms. With fewer such firms in the system, the need to incentivize reallocation through lower benefits diminishes, allowing the planner to raise support levels.

At this point, only firms with transitory productivity shocks remain eligible for STW. As a result, further increases in their share do not substantially affect the optimal benefit level.

When it comes to our interpretation of structural change, the model provides a different perspective than in earlier sections. The optimal response of the STW system to structural change depends critically on the initial share of firms with permanent productivity shocks. If this share is low, then permanent-shock firms are already being screened out by a tight eligibility condition, and the government does not need to adjust the system in response to structural change.

However, if a structural shift increases the number of firms with permanent productivity shocks a lot, the government may choose to include them in the STW system. This requires loosening the eligibility condition, which in turn calls for lower STW benefits to limit the distortionary effects of the system. By contrast, if the economy already had a large fraction of firms with permanent shocks, then the increase reinforces their presence in the system, and in this case, optimal benefits should rise to better support them.

This highlights the central role of the eligibility condition in shaping the optimal design of the STW system, particularly in the face of structural shifts.

2.4.4 Welfare Comparison

Finally, we examine how the introduction of STW affects overall welfare under different information settings. Broadly, STW can create two types of welfare costs. First, it may incentivize firms and workers to choose inefficiently low working hours, distorting labor input. Second, STW — or subsidy systems more generally — can lead to misallocation of labor. This includes both excess retention of workers in permanently unproductive firms and premature separations in firms experiencing only transitory productivity shocks. Table 1 presents the welfare gains under different regimes, allowing us to assess the relative importance of the hours distortion effect and the welfare costs associated with worker misallocation.

Table 2.4.1. Welfare Increase in different Models

	Jobs at risk + shock persistence	Jobs at risk, not shock persistence	Only Working hours
Subsidy - uniform	2.6730%	2.6139%	-
Subsidy - time dependent	2.6730%	2.6209%	-
STW - uniform	0.8359%	0.7868%	0.4653%
STW - time dependent	0.8359%	0.7901%	0.4653%

Notes: The table shows the percentage increase in consumption equivalents between the optimal subsidy, respectviely STW system with respect to the baseline economy. The baseline economy considers a fraction of firms with transitory shocks of $p_T = 0.5$, so the economy under the information setting "only working hours" sets the strict eligibility condition.

It is striking that an optimal subsidy or STW system can effectively keep the welfare costs from worker misallocation low. When we compare policy systems that have full information about the shock persistence of firms to those that suffer from this information problem, we observe only minor welfare losses due to misallocation. Consequently, allowing STW benefits to depend on the duration a worker spends in the system also yields only limited additional welfare gains.

In contrast, the welfare difference between the subsidy system and the STW system is substantial, underscoring the significant costs associated with working hours distortions. The situation deteriorates further when the government cannot reliably identify jobs at risk. In this case, the minimum hours reduction threshold introduces a misclassification problem: some firms with permanent productivity shocks may be inefficiently excluded from STW, while others with transitory shocks may enter the system despite not requiring support. In this calibration, the government decides to perfectly screen out permanent unproductive firms, making it superfluous to depend on STW benefits on duration.

2.5 Conclusion

One of the fundamental concerns surrounding job retention schemes such as STW is the potential for labor misallocation. While previous research based on suboptimally designed systems — such as in Cooper, Meyer, and Schott (2017)—concludes that STW exacerbates misallocation of workers, this paper investigates whether optimal STW policy also leads to misallocation and, if so, how it should be addressed.

The analysis shows that even under optimal design, STW policies result in some degree of worker misallocation. The Ramsey planner aims to provide generous subsidies to temporarily unproductive firms in order to avoid inefficient job destruction. Simultaneously, limited support is offered to permanently unproductive firms to promote necessary worker reallocation.

STW systems partially accommodate this objective. Firms facing transitory shocks are more willing to reduce working hours compared to those experiencing permanent productivity losses. Since STW subsidies are typically linked to the extent of hours reductions, the system naturally channels more generous support to firms with temporary shocks, thereby alleviating the labor misallocation problem.

However, quantitatively, the paper finds that while working hours provide a useful screening mechanism, it is insufficient to resolve the misallocation issue fully. In particular, the STW system tends to undersubsidize firms with transitory shocks and oversubsidize those with permanent shocks, resulting in both premature separations and excess retention.

Optimal STW benefits are designed to minimize the welfare losses associated with these two distortions. The optimal STW policy is calibrated such that the weighted average of transfers across firm types — those experiencing permanent versus transitory productivity shocks — equals the average weighted transfer level in an environment where the duration of shocks is unobservable. Greater weight is assigned to firm types for which marginal changes in STW generosity lead to greater match preservation. This targeting principle enables the STW system to substantially reduce the welfare costs associated with misallocation. Quantitatively, I find that the primary source of welfare loss under STW stems from distortions to working hours rather than from labor misallocation.

In practice, most STW systems employ a limited duration of the STW system. When firms can perfectly check their eligibility condition, the paper finds support for a tiered system. Reducing STW benefits over time helps to screen out firms with permanent productivity shocks. However, when STW systems rely on a minimum hours reduction threshold as the eligibility condition, it serves as a mechanism to screen out permanently unproductive firms, making it unnecessary to adjust STW benefits over the duration of participation.

The paper offers two promising avenues for future research. First, it provides a framework to quantify the welfare costs imposed by information asymmetries in the STW system, offering guidance on the extent to which governments should invest in firm-level monitoring to enhance policy targeting. A natural extension would be to empirically examine how effectively governments can identify jobs at risk and the administrative costs associated with doing so. For example, Fitzenberger and Walwei (2023) documents that the administrative burden of implementing STW in Germany was substantial during the COVID-19 pandemic.

Second, the analysis offers a foundation for studying the optimal design of STW in the context of structural change. The findings suggest that in the face of pronounced structural shifts, reducing STW benefit generosity may be warranted to facilitate labor reallocation, while simultaneously relaxing eligibility criteria to accommodate a larger share of firms experiencing persistent productivity shocks. A promising direction for future research would be to embed the framework in a richer, multi-sector model with explicit sectoral heterogeneity and reallocation frictions arising from sector-specific human capital.

References

Balleer, Almut, Britta Gehrke, Wolfgang Lechthaler, and Christian Merkl. 2016. "Does short-time work save jobs? A business cycle analysis." *European Economic Review* 84 (C): 99–122. [134]

Bossler, Mario, Christopher Osiander, Julia Schmidtke, and Mark Trappmann. 2023. "Free riding on short-time work allowances? Results from an experimental survey design." *Kyklos* 76 (4): 882–901. [138]

Braun, Helge, and Björn Brügemann. 2017. "Welfare Effects of Short-Time Compensation." Tinbergen Institute Discussion Papers 17-010/VI. Tinbergen Institute. [134]

Brinkmann, Christina, Simon Jäger, Moritz Kuhn, Farzad Saidi, and Stefanie Wolter. 2024. "Short-Time Work Extensions." ECONtribute Discussion Papers Series 340. University of Bonn, and University of Cologne, Germany. [154]

Burdett, Kenneth, and Randall Wright. 1989. "Unemployment Insurance and Short-Time Compensation: The Effects on Layoffs, Hours per Worker, and Wages." *Journal of Political Economy* 97 (6): 1479–96. [135, 138]

Cahuc, Pierre, Francis Kramarz, and Sandra Nevoux. 2021. "The Heterogeneous Impact of Short-Time Work: From Saved Jobs to Windfall Effects." IZA Discussion Papers 14381. Institute of Labor Economics (IZA). [135, 138]

Christoffel, Kai, and Tobias Linzert. 2010. "The Role of Real Wage Rigidity and Labor Market Frictions for Inflation Persistence." *Journal of Money, Credit and Banking* 42 (7): 1435–46. [154]

Cooper, Russell, Moritz Meyer, and Immo Schott. 2017. "The Employment and Output Effects of Short-Time Work in Germany." NBER Working Papers 23688. National Bureau of Economic Research, Inc. [133, 134, 139, 175]

Costain, James S., and Michael Reiter. 2008. "Business cycles, unemployment insurance, and the calibration of matching models." *Journal of Economic Dynamics and Control* 32 (4): 1120–55. [153]

Domeij, David, and Martin Floden. 2006. "The Labor-Supply Elasticity and Borrowing Constraints: Why Estimates are Biased." *Review of Economic Dynamics* 9 (2): 242–62. [154]

Engen, Eric M., and Jonathan Gruber. 2001. "Unemployment insurance and precautionary saving." *Journal of Monetary Economics* 47 (3): 545–79. [154]

Fitzenberger, Bernd, and Ulrich Walwei. 2023. "Kurzarbeitergeld in der Covid-19-Pandemie: Lessons learned." IAB-Forschungsbericht 202305. Institut für Arbeitsmarkt- und Berufsforschung (IAB), Nürnberg [Institute for Employment Research, Nuremberg, Germany]. [133, 176]

Giupponi, Giulia, and Camille Landais. 2018. "Subsidizing labor hoarding in recessions: the employment and welfare effects of short time work." CEP Discussion Papers dp1585. Centre for Economic Performance, LSE. [134]

Giupponi, Giulia, and Camille Landais. 2022. "Subsidizing Labour Hoarding in Recessions: The Employment and Welfare Effects of Short-time Work." *Review of Economic Studies* 90 (4): 1963–2005. eprint: https://academic.oup.com/restud/article-pdf/90/4/1963/51440201/rdac069.pdf. [139]

Haan, Wouter J. den, Garey Ramey, and Joel Watson. 2000. "Job Destruction and Propagation of Shocks." *American Economic Review* 90 (3): 482–98. [153]

Hosios, Arthur J. 1990. "On the Efficiency of Matching and Related Models of Search and Unemployment." *Review of Economic Studies* 57 (2): 279–98. [150, 154]

Mortensen, Dale T., and Christopher A. Pissarides. 1994. "Job Creation and Job Destruction in the Theory of Unemployment." *Review of Economic Studies* 61 (3): 397–415. [133]

Petrongolo, Barbara, and Christopher A. Pissarides. 2001. "Looking into the Black Box: A Survey of the Matching Function." *Journal of Economic Literature* 39 (2): 390–431. [154]

Pissarides, Christopher A. 2000. *Equilibrium Unemployment Theory, 2nd Edition*. Vol. 1, MIT Press Books 0262161877. The MIT Press. [149]

Stiepelmann, Gero. 2024. "Optimal Short-Time Work Policy in Recessions." Working Papers. University of Bonn. [134, 135, 138, 142, 148, 150, 152, 153]

Teichgräber, Julian, Simon Žužek, and Jannik Hensel. 2022. "Optimal short-time work: screening for jobs at risk." ECON - Working Papers 402. Department of Economics - University of Zurich. [134, 136, 138]

Van Audenrode, Marc A. 1994. "Short-Time Compensation: Job Security, and Employment Contracts: Evidence from Selected OECD Countries." *Journal of Political Economy* 102 (1): 76–102. [138]

Appendix 2.A List of important variables and functions

Symbol	Description
$G_\lambda(\varepsilon), g_\lambda(\varepsilon)$	C.d.f. and p.d.f. of idiosyncratic productivity shock
$P(\lambda), p(\lambda)$	C.d.f. and p.d.f. of shock duration distribution
n	Mass of employed workers
$u = 1 - n$	Mass of unemployed workers
$n(\lambda)$	Mass of matches with expected shock duration $1/\lambda$
$n(\lambda, k)$	Mass of matches with expected shock duration $1/\lambda$, k periods after shock
$n_{stw}(\lambda)$	Mass of matches with expected shock duration $1/\lambda$ on STW
$n_{stw}(\lambda, k)$	Mass of matches with expected shock duration $1/\lambda$, k periods on STW
v	Mass of posted vacancies
$\theta = \frac{v}{1-n}$	Labor market tightness
f	Job-finding rate
q	Vacancy-filling rate
$\rho(\lambda)$	Separation rate matches with recovery probability λ
$\rho(\lambda, k)$	Separation rate matches with recovery probability λ, after k periods on STW
$\rho_{stw}(\lambda)$	Probability to enter STW for matches with expected shock duration $1/\lambda$
$\rho_{stw}(\lambda, k)$	probability to enter STW k periods on the system
$R(\lambda)$	Separation threshold (on STW) in firms with recovery probability λ
$R(\lambda, k)$	Separation threshold (on STW) in firms with recovery probability λ, k periods on the system
$\xi(\lambda)$	Separation threshold (no STW) in firms with recovery probability λ
D	Eligibility threshold for STW (on working hours) - only in case 3
$R_{\mathrm{stw}}(\lambda)$	Eligibility threshold for STW (on productivity)
$s(\lambda, h, k)$	General subsidy function
s	Subsidy
$s^*(\lambda)$	Optimal subsidy without information asymmetry
s_ω	Optimal subsidy with information asymmetry
$s_\omega(k)$	Optimal subsidy with information asymmetry, k periods on the system
τ_{stw}	STW benefits
$\tau_{\mathrm{stw}}(\lambda)$	Optimal STW benefits without Information asymmetry
$\tau_{stw,\omega}$	Optimal STW benefits with information asymmetry
$\tau_{\mathrm{stw},\omega}(k)$	Optimal STW benefits with information asymmetry, k periods on the system
$\tau^{net}_{\mathrm{stw}}$	Net-transfer STW benefits
$\tau^{net}_{\mathrm{stw}}(\lambda)$	Net-transfer with optimal STW benefits without information asymmetry
$\tau^{net}_{\mathrm{stw},\omega}(\lambda)$	Net-transfer optimal STW benefits with information asymmetry
$\tau^{net}_{\mathrm{stw},\omega}(\lambda, k)$	Net-transfer optimal STW benefits with information asymmetry, k periods on the system
b	Unemployment benefits
τ	Lump sum tax per shock

Table 2.A.1. Important Variables and Functions

Symbol	**Description**
m	mass of matched workers
$\phi(h)$	Disutility of labor: $\frac{h^{1+\psi}}{1+\psi}$
$y(\varepsilon, h)$	Output
$z(\varepsilon, h)$	Output net of disutility (cons.-eq. units)
$z(\varepsilon)$	Shorthand for $z(\varepsilon, h(\varepsilon))$
$z_{\text{stw}}(\varepsilon)$	Output net of disutility under STW hours
$\tilde{\Omega}(\lambda)$	Average welfare cost of a match in STW with shock duration $1/\lambda$. Further, $\Omega(\lambda) = \rho_{stw}(\lambda) \cdot \tilde{\Omega}(\lambda)$
$\tilde{\Omega}(\lambda, k)$	Average welfare cost of a match in STW with shock duration $1/\lambda$ that spent k periods on STW. Further, $\Omega(\lambda, k) = \rho_{stw}(\lambda, k) \cdot \tilde{\Omega}(\lambda, k)$
$h(\varepsilon)$	Hours worked (non-STW)
$h_{\text{stw},k}(\varepsilon)$	Hours worked under STW in period k of the system
w	Total wage
$w_{stw}(\varepsilon, \lambda, k)$	Total wage function on STW
Π	Profits
$V(\varepsilon, \lambda)$	Worker value (no STW), in firm with expected shock duration $1/\lambda$
$V_{stw}(\varepsilon, \lambda, k)$	Worker value on STW, in firm with expected shock duration $1/\lambda$, k periods on STW
$\bar{V}$	Expected worker value, before idiosyncratic shocks realize
U	Value of unemployment
$J(\varepsilon, \lambda)$	Firm value (no STW) with expected shock duration $1/\lambda$
$J_{\text{stw}}(\varepsilon, \lambda, k)$	Firm value on STW with expected shock duration $1/\lambda$, k periods on STW
$\bar{J}$	Expected firm value, before idiosyncratic shocks realize

Table 2.A.2. Important Variables and Functions

Appendix 2.B Decentralized Economy

2.B.1 Equilibrium conditions employment

This section calculates various equilibrium expressions for employment. These are defined by:

1. The number of workers that draw a shock this period:

$$n^s = \theta \cdot q(\theta) \cdot (1-n) + \int_0^1 \sum_{k=1}^{K-1} \lambda \cdot n(\lambda, k) + \lambda \cdot n(\lambda, K) dP(\lambda)$$

2. The number of firms with shock duration $1/\lambda$, in period k after shock arrival $\forall \lambda \in [0,1]$ and $k \in \mathbb{N}_+,\ k \leq K$:

$$n(\lambda, k) = (1-\lambda)^{k-1} \cdot \big(1 - \rho(\lambda, k)\big) \cdot p(\lambda) \cdot n^s$$

3. The number of firms with shock duration $1/\lambda$, in period K or more periods after shock arrival $\forall \lambda \in [0,1]$:

$$n(\lambda, K) = \frac{(1-\lambda)^{K-1}}{\lambda} \cdot \big(1 - \rho(\lambda, K)\big) \cdot p(\lambda) \cdot n^s$$

4. Total employment

$$n = \int_0^1 \sum_{k=1}^{K-1} n(\lambda, k) + n(\lambda, K) d\lambda$$

5. Separation rate after shock:

$$\begin{aligned} \rho(\lambda, k) &= G\big(\max\{R(\lambda, 1), ..., R(\lambda, k)\}\big) \\ &\quad + G\big(\max\{\xi(\lambda), R_{stw}(\lambda)\}\big) - G\big(\max\{R_{stw}(\lambda), R(\lambda, 1), ..., R(\lambda, k)\}\big) \end{aligned}$$

First we need to calculate n^s. To do this, we need an expression for $\int_0^1 \sum_{k=1}^{K-1} \lambda \cdot n(\lambda, k) + \lambda \cdot n(\lambda, K) dP(\lambda)$. Therefore, note that

$$\begin{aligned} \lambda \cdot n(\lambda, k) &= \lambda \cdot (1-\lambda)^{k-1} \cdot \big(1 - \rho(\lambda, k)\big) \cdot p(\lambda) \cdot n^s \\ \Leftrightarrow \int_0^1 \lambda \cdot n(\lambda, k) d\lambda &= \int_0^1 \sum_{k=1}^{K-1} \lambda \cdot (1-\lambda)^{k-1} \cdot \big(1 - \rho(\lambda, K)\big) \cdot dP(\lambda) \cdot n^s \end{aligned}$$

and

$$\lambda \cdot n(\lambda, K) = (1-\lambda)^{K-1} \cdot \big(1-\rho(\lambda, K)\big) \cdot p(\lambda) \cdot n^s$$
$$\int_0^1 \lambda \cdot n(\lambda, K) \cdot d\lambda = \int_0^1 (1-\lambda)^{K-1} \cdot \big(1-\rho(\lambda, K)\big) \cdot dP(\lambda) \cdot n^s$$

Denote

$$E_K[\rho(R(\lambda, k))] = \sum_{k=1}^{K-1} \lambda \cdot (1-\lambda)^{k-1} \cdot \big(1-\rho(\lambda, k)\big) + (1-\lambda)^{K-1} \cdot \big(1-\rho(\lambda, K)\big)$$

This allows us to express the expression as:

$$\int_0^1 \sum_{k=1}^{K-1} \lambda \cdot n(\lambda, k) + \lambda \cdot n(\lambda, K) dP(\lambda) = \left(1 - \int_0^1 E_K[\rho(\lambda, k)] dP(\lambda)\right) \cdot n^s$$

Insert n^s

$$\begin{aligned} &\int_0^1 \sum_{k=1}^{K-1} \lambda \cdot n(\lambda, k) + \lambda \cdot n(\lambda, K)\, dP(\lambda) \\ &= \left(1 - \int_0^1 E_K[\rho(\lambda, k)]\, dP(\lambda)\right) \cdot \Bigg(\theta \cdot q(\theta) \cdot (1-n) \\ &\qquad + \int_0^1 \sum_{k=1}^{K-1} \lambda \cdot n(\lambda, k) + \lambda \cdot n(\lambda, K)\, dP(\lambda)\Bigg) \end{aligned}$$

Rearranging gives:

$$\begin{aligned} &\int_0^1 \sum_{k=1}^{K-1} \lambda \cdot n(\lambda, k) + \lambda \cdot n(\lambda, K)\, dP(\lambda) \\ &= \frac{1 - \int_0^1 E_K[\rho(\lambda, k)]\, dP(\lambda)}{\int_0^1 E_K[\rho(\lambda, k)]\, dP(\lambda)} \cdot \theta \cdot q(\theta) \cdot (1-n) \end{aligned}$$

This gives an expression for n^s:

$$n^s = \frac{\theta \cdot q(\theta)}{\int_0^1 E_K[\rho(\lambda, k)] dP(\lambda)} \cdot (1-n)$$

Inserting n^s into $n(\lambda,k)$, $n(\lambda,K)$ gives:

$$n(\lambda,k) = (1-\lambda)^{k-1} \cdot (1-\rho(\lambda,k)) \cdot p(\lambda) \cdot \frac{\theta \cdot q(\theta)}{\int_0^1 E_K[\rho(\lambda,k)]dP(\lambda)} \cdot (1-n)$$
$$n(\lambda,K) = \frac{(1-\lambda)^{K-1}}{\lambda} \cdot (1-\rho(\lambda,K)) \cdot p(\lambda) \cdot \frac{\theta \cdot q(\theta)}{\int_0^1 E_K[\rho(\lambda,K)]dP(\lambda)} \cdot (1-n)$$

Inserting it into the expression for employment pins down n:

$$n = \frac{\int_0^1 \frac{\theta \cdot q(\theta)}{\lambda} \cdot \left(1 - E_K[\rho(\lambda,k)]\right) dP(\lambda)}{\int_0^1 E_K[\rho(\lambda,k)]dP(\lambda)} \cdot (1-n)$$

Rearranging for n gives:

$$n = \frac{\int_0^1 \frac{\theta \cdot q(\theta)}{\lambda} \cdot \left(1 - E_K[\rho(\lambda,k)]\right) dP(\lambda)}{\int_0^1 \frac{\theta \cdot q(\theta)}{\lambda} \cdot \left(1 - E_K[\rho(\lambda,k)]\right) dP(\lambda) + \int_0^1 E_K[\rho(\lambda,k)]dP(\lambda)}$$

Consequently n^s can be written as:

$$n^s = \frac{\theta \cdot q(\theta)}{\int_0^1 \frac{\theta \cdot q(\theta)}{\lambda} \cdot \left(1 - E_K[\rho(\lambda,k)]\right) dP(\lambda) + \int_0^1 E_K[\rho(\lambda,k)]dP(\lambda)}$$

2.B.2 Job-Creation Condition

First, let us calculate the joint surplus of firms and workers after the idiosyncratic shock has realized to (ϵ, λ):

$$\begin{aligned} S(\epsilon,\lambda) &= J(\epsilon,\lambda) + V(\epsilon,\lambda) - U \\ &= z(\epsilon) + (1-\lambda) \cdot (J(\epsilon,\lambda) + V(\epsilon,\lambda)) + \lambda \cdot \left(\bar{J} + \bar{V}\right) \\ &- b - f \cdot V - (1-f) \cdot U \\ &= z(\epsilon) - b + (1-\lambda) \cdot (J(\epsilon,\lambda) + V(\epsilon,\lambda) - U) + \lambda \bar{J} + (\lambda - f) \cdot (\bar{V} - U) \end{aligned}$$

Insert the optimality condition for the wage-sum:

$$S(\epsilon,\lambda) = z(\epsilon) - b + (1-\lambda) \cdot S(\epsilon,\lambda) + \left(\lambda - f \cdot \eta\right) \cdot \bar{S}$$

Rearranging gives:

$$S(\epsilon,\lambda) = \frac{1}{\lambda} \cdot [z(\epsilon) - b + \left(\lambda - f \cdot \eta\right) \cdot \bar{S}]$$

The joint surplus of a worker on STW can be denoted as:

$$S(\epsilon,\lambda,0) \;=\; J_{stw}(\epsilon,\lambda,0) + V_{stw}(\epsilon,\lambda,0) - U$$

Using the same steps as above, we can rewrite the equation to:

$$\begin{aligned}
S_{stw}(\epsilon,\lambda,1) \;&=\; J_{stw}(\epsilon,\lambda,1) + V_{stw}(\epsilon,\lambda,1) - U \\
&=\; z_{stw,1}(\epsilon) - b + \tau_{stw}(1)\cdot\left(\bar{h} - h_{stw,1}(\epsilon)\right) \\
&\quad + (1-\lambda)\cdot \mathbb{1}(\epsilon \;>\; R(\lambda,2))\cdot S(\epsilon,\lambda,1) \\
&\quad + \left(\lambda - \eta\cdot f\right)\cdot \bar{S} \\
&=\; z_{stw,1}(\epsilon) - b + \tau_{stw}(1)\cdot\left(\bar{h} - h_{stw,1}(\epsilon)\right) \\
&\quad + (1-\lambda)\cdot \mathbb{1}(\epsilon \;>\; R(\lambda,2))\cdot \left(z_{stw,2}(\epsilon)\right. \\
&\quad \left. + \tau_{stw}(2)\cdot\left(\bar{h} - h_{stw,2}(\epsilon)\right) - b\right) \\
&\quad + (1-\lambda)^2\cdot \mathbb{1}(\epsilon \;>\; R(\lambda,2))\cdot \mathbb{1}(\epsilon \;>\; R(\lambda,3))\cdot S(\epsilon,\lambda,3) \\
&\quad + \left(\lambda - \eta\cdot f\right)\cdot \bar{S} \\
&\quad + (1-\lambda)\cdot\left(\lambda - \eta\cdot f\right)\cdot \bar{S} \\
&=\; \sum_{k=1}^{K-1}(1-\lambda)^{k-1}\cdot\prod_{i=2}^{k-1}\mathbb{1}(\epsilon \;>\; R(\lambda,i))\cdot\left(z_{stw,k}(\epsilon)\right. \\
&\quad \left. + \tau_{stw}(k)\cdot\left(\bar{h} - h_{stw,k}(\epsilon)\right) - b + \left(\lambda - \eta\cdot f\right)\cdot\bar{S}\right) \\
&\quad + (1-\lambda)^{K-1}\cdot\prod_{i=2}^{K-1}\mathbb{1}(\epsilon \;>\; R(\lambda,i))\cdot S(\epsilon,\lambda,K)
\end{aligned}$$

After K periods, the STW does not change anymore:

$$\begin{aligned}
S_{stw}(\epsilon,\lambda,K) \;&=\; z_{stw,K}(\epsilon) + \tau_{stw}(K)\cdot\left(\bar{h} - h_{stw,K}(\epsilon)\right) \\
&\quad + (1-\lambda)\cdot S_{stw}(\epsilon,\lambda,K) + (\lambda - \eta\cdot f)\cdot\bar{S}
\end{aligned}$$

Rearranging gives:

$$S_{stw}(\epsilon,\lambda,K) \;=\; \frac{1}{\lambda}\cdot[z_{stw,K}(\epsilon) + \tau_{stw}(K)\cdot\left(\bar{h} - h_{stw,K}(\epsilon)\right) - b + (\lambda - \eta\cdot f)\cdot\bar{S}]$$

The surplus on STW can thus be denoted as:

$$\begin{aligned}
S_{stw}(\epsilon,\lambda,1) \;&=\; \sum_{k=1}^{K-1}(1-\lambda)^{k-1}\cdot\prod_{i=2}^{k-1}\mathbb{1}(\epsilon \;>\; R(\lambda,i)) \\
&\quad \cdot\left(z_{stw,k}(\epsilon) + \tau_{stw}(k)\cdot\left(\bar{h} - h_{stw,k}(\epsilon)\right) - b + \left(\lambda - \eta\cdot f\right)\cdot\bar{S}\right) \\
&\quad + \frac{(1-\lambda)^{K-1}}{\lambda}\cdot\prod_{i=2}^{K-1}\mathbb{1}(\epsilon \;>\; R(\lambda,i)) \\
&\quad \cdot\left(z_{stw,K}(\epsilon) + \tau_{stw}(K)\cdot\left(\bar{h} - h_{stw,K}(\epsilon)\right) - b + (\lambda - \eta\cdot f)\cdot\bar{S}\right)
\end{aligned}$$

The expected surplus of a match before the idiosyncratic productivity shocks are known can be denoted as:

$$\bar{S} = -\tau + \int_0^1 \int_{\max\{\xi(\lambda),R_{stw}(\lambda)\}}^{\infty} S(\epsilon,\lambda)dG(\epsilon)dP(\lambda) + \int_0^1 \int_{R(\lambda,1)}^{\max\{R_{stw}(\lambda),R(\lambda,1)\}} S_{stw}(\epsilon,\lambda,1)dG(\epsilon)dP(\lambda)$$

Insert the expressions for the surplus after the idiosyncratic shocks have been realized:

$$\begin{aligned}\bar{S} = & -\tau + \int_0^1 \int_{\max\{\xi(\lambda),R_{stw}(\lambda)\}}^{\infty} \frac{1}{\lambda}\cdot\Big[z(\epsilon) - b + \big(\lambda - f\cdot\eta\big)\cdot\bar{S}\Big]dG(\epsilon) \\ & + \int_0^1 \sum_{k=1}^{K-1}(1-\lambda)^{k-1}\cdot\int_{\max\{R(\lambda,1),\ldots,R(\lambda,k)\}}^{\max\{R_{stw}(\lambda),R(\lambda,1),\ldots,R(\lambda,k)\}} \Big[z_{stw,k}(\epsilon) + \tau_{stw}(k)\cdot\big(\bar{h} - h_{stw,k}(\epsilon)\big) \\ & - b + \big(\lambda - \eta\cdot f\big)\cdot\bar{S}\Big]dP(\lambda) \\ & + \int_0^1 \frac{(1-\lambda)^{K-1}}{\lambda}\cdot\int_{\max\{R(\lambda,1),\ldots,R(\lambda,K)\}}^{\max\{R_{stw}(\lambda),R(\lambda,1),\ldots,R(\lambda,K)\}} \Big[z_{stw,K}(\epsilon) + \tau_{stw}(K)\cdot\big(\bar{h} - h_{stw,K}(\epsilon)\big) \\ & - b + (\lambda - \eta\cdot f)\cdot\bar{S}\Big]dP(\lambda)\end{aligned}$$

Note that the government budget constraint has the following form:

$$\begin{aligned}n^s\cdot\tau = & (1-n)\cdot b \\ & + \int_0^1 \sum_{k=1}^{K}\frac{n(\lambda,k)}{1-\rho(\lambda,k)} \\ & \cdot\int_{\max\{R(\lambda,1),\ldots,R(\lambda,k)\}}^{\max\{R_{stw}(\lambda),R(\lambda,1),\ldots,R(\lambda,k)\}} \tau_{stw}(k)\cdot\big(\bar{h} - h_{stw}(\epsilon)\big)\cdot dG_\lambda(\epsilon)\,dP(\lambda)\end{aligned}$$

Inserting the definitions for n and $n(\lambda,k)$ gives:

$$\begin{aligned}\tau = & \frac{1-n}{n^s}\cdot b \\ & + \int_0^1 \sum_{k=1}^{K}(1-\lambda)^{k-1}\cdot\int_{\max\{R(\lambda,1),\ldots,R(\lambda,k)\}}^{\max\{R_{stw}(\lambda),R(\lambda,1),\ldots,R(\lambda,k)\}} \tau_{stw}(k)\cdot\big(\bar{h} - h_{stw}(\epsilon)\big)dG_\lambda(\epsilon)dP(\lambda) \\ & + \int_0^1 \frac{(1-\lambda)^{K-1}}{\lambda}\cdot\int_{\max\{R(\lambda,1),\ldots,R(\lambda,K)\}}^{\max\{R_{stw}(\lambda),R(\lambda,1),\ldots,R(\lambda,K)\}} \tau_{stw}(K)\cdot\big(\bar{h} - h_{stw}(\epsilon)\big)dG_\lambda(\epsilon)dP(\lambda)\end{aligned}$$

Using the equilibrium outcomes for n and n^s lets us rewrite the budget constraint to:

$$
\begin{aligned}
\tau \;=\; & \frac{1}{\lambda}\cdot\big(1-E_K[\rho(R(\lambda,k))]\big)\cdot\frac{1-n}{n}\cdot b \\
& + \int_0^1 \sum_{k=1}^{K}(1-\lambda)^{k-1}\cdot\int_{\max\{R(\lambda,1),\dots,R(\lambda,k)\}}^{\max\{R_{stw}(\lambda),R(\lambda,1),\dots,R(\lambda,k)\}} \tau_{stw}\cdot\big(\bar{h}-h_{stw}(\epsilon)\big)\,dG_\lambda(\epsilon)dP(\lambda) \\
& + \int_0^1 \frac{(1-\lambda)^{K-1}}{\lambda}\cdot\int_{\max\{R(\lambda,1),\dots,R(\lambda,K)\}}^{\max\{R_{stw}(\lambda),R(\lambda,1),\dots,R(\lambda,K)\}} \tau_{stw}\cdot\big(\bar{h}-h_{stw}(\epsilon)\big)\,dG_\lambda(\epsilon)dP(\lambda)
\end{aligned}
$$

Inserting it into the surplus equation cancels out the direct impact of the joint surplus on job creation.

$$
\begin{aligned}
\bar{S} \;=\; & \int_0^1 \int_{\max\{\xi(\lambda),R_{stw}(\lambda)\}}^{\infty} \frac{1}{\lambda}\cdot[z(\epsilon)-b-\frac{1-n}{n}\cdot b+\big(\lambda-f\cdot\eta\big)\cdot\bar{S}]dG(\epsilon) \\
& + \int_0^1 \sum_{k=1}^{K-1}(1-\lambda)^{k-1}\cdot\int_{\max\{R(\lambda,1),\dots,R(\lambda,k)\}}^{\max\{R_{stw}(\lambda),R(\lambda,1),\dots,R(\lambda,k)\}} [z_{stw,k}(\epsilon)-b-\frac{1-n}{n}\cdot b \\
& \qquad + \big(\lambda-\eta\cdot f\big)\cdot\bar{S}]dG(\epsilon)\,dP(\lambda) \\
& + \int_0^1 \frac{(1-\lambda)^{K-1}}{\lambda}\cdot\int_{\max\{R(\lambda,1),\dots,R(\lambda,K)\}}^{\max\{R_{stw}(\lambda),R(\lambda,1),\dots,R(\lambda,K)\}} [z_{stw,K}(\epsilon)-b-\frac{1-n}{n}\cdot b \\
& \qquad + (\lambda-\eta\cdot f)\cdot\bar{S}]dG(\epsilon)\,dP(\lambda)
\end{aligned}
$$

We can simplify the expression to

$$
\begin{aligned}
\bar{S} \;=\; & \int_0^1 \Bigg(\sum_{k=1}^{K-1}(1-\lambda)^{k-1}\cdot\Bigg[\int_{A(\lambda,k)} z(\epsilon)dG_\lambda(\epsilon)-\Omega(\lambda,k)-\frac{b}{n} \\
& \qquad + \big(\lambda-\eta\cdot\theta\cdot q(\theta)\big)\cdot\bar{S}\Bigg]\Bigg)dP(\lambda) \\
& + \int_0^1 \Bigg(\frac{(1-\lambda)^{K-1}}{\lambda}\cdot\Bigg[\int_{A(\lambda,K)} z(\epsilon)dG_\lambda(\epsilon)-\Omega(\lambda,K)-\frac{b}{n} \\
& \qquad + \big(\lambda-\eta\cdot\theta\cdot q(\theta)\big)\cdot\bar{S}\Bigg]\Bigg)dP(\lambda)
\end{aligned}
$$

with

$$
\begin{aligned}
A(\lambda,k) \;=\; & [\max\{R(\lambda,1),\dots,R(\lambda,k)\},\max\{R_{stw}(\lambda),R(\lambda,1),\dots,R(\lambda,k)\}]\cup \\
& [\max\{R_{stw}(\lambda),\xi(\lambda)\},\infty)
\end{aligned}
$$

Insert the free entry condition and make use of the optimality condition for the joint surplus:

$$
\begin{aligned}
\frac{k_v}{q(\theta)} &= J = (1-\eta)\cdot\bar{S} \\
\Leftrightarrow \bar{S} &= \frac{1}{1-\eta}\cdot\frac{k_v}{q(\theta)}
\end{aligned}
$$

This gives:

$$
\begin{aligned}
\frac{1}{1-\eta}\cdot\frac{k_v}{q(\theta)} = &\int_0^1\Bigg(\sum_{k=1}^{K-1}(1-\lambda)^{k-1}\cdot\Bigg[\int_{A(\lambda,k)} z(\epsilon)dG_\lambda(\epsilon)-\Omega(\lambda,k)-\frac{b}{n} \\
&\qquad+\frac{\lambda-\eta\cdot\theta\cdot q(\theta)}{1-\eta}\cdot\frac{k_v}{q(\theta)}\Bigg]\Bigg)dP(\lambda) \\
&+\int_0^1\Bigg(\frac{(1-\lambda)^{K-1}}{\lambda}\cdot\Bigg[\int_{A(\lambda,K)} z(\epsilon)dG_\lambda(\epsilon)-\Omega(\lambda,K)-\frac{b}{n} \\
&\qquad+\frac{\lambda-\eta\cdot\theta\cdot q(\theta)}{1-\eta}\cdot\frac{k_v}{q(\theta)}\Bigg]\Bigg)dP(\lambda)
\end{aligned}
$$

In case of $s(\lambda, h_{stw,k}, k) \geq s(\lambda, h_{stw,k+1}, k+1)$ we can write the job-creation condition as:

$$
\begin{aligned}
\frac{1}{1-\eta}\cdot\frac{k_v}{q(\theta)} = &\int_0^1\Bigg(\sum_{k=1}^{K-1}(1-\lambda)^{k-1}\cdot\Bigg[\int_{R(\lambda,k)}^{\infty} z(\epsilon)dG_\lambda(\epsilon)-\Omega(\lambda,k)-\frac{b}{n} \\
&\qquad+\frac{\lambda-\eta\cdot\theta\cdot q(\theta)}{1-\eta}\cdot\frac{k_v}{q(\theta)}\Bigg]\Bigg)dP(\lambda) \\
&+\int_0^1\Bigg(\frac{(1-\lambda)^{K-1}}{\lambda}\cdot\Bigg[\int_{R(\lambda,K)}^{\infty} z(\epsilon)dG_\lambda(\epsilon)-\Omega(\lambda,K)-\frac{b}{n} \\
&\qquad+\frac{\lambda-\eta\cdot\theta\cdot q(\theta)}{1-\eta}\cdot\frac{k_v}{q(\theta)}\Bigg]\Bigg)dP(\lambda)
\end{aligned}
$$

If $s(\lambda, h_{stw,k}, k) \leq s(\lambda, h_{stw,k+1}, k+1)$ we can write the job-creation condition as:

$$
\begin{aligned}
\frac{1}{1-\eta}\cdot\frac{k_v}{q(\theta)} = &\int_0^1\Bigg(\sum_{k=1}^{K-1}(1-\lambda)^{k-1}\cdot\Bigg[\int_{R(\lambda,1)}^{\infty} z(\epsilon)dG_\lambda(\epsilon)-\Omega(\lambda,k)-\frac{b}{n} \\
&\qquad+\frac{\lambda-\eta\cdot\theta\cdot q(\theta)}{1-\eta}\cdot\frac{k_v}{q(\theta)}\Bigg]\Bigg)dP(\lambda) \\
&+\int_0^1\Bigg(\frac{(1-\lambda)^{K-1}}{\lambda}\cdot\Bigg[\int_{R(\lambda,1)}^{\infty} z(\epsilon)dG_\lambda(\epsilon)-\Omega(\lambda,K)-\frac{b}{n} \\
&\qquad+\frac{\lambda-\eta\cdot\theta\cdot q(\theta)}{1-\eta}\cdot\frac{k_v}{q(\theta)}\Bigg]\Bigg)dP(\lambda)
\end{aligned}
$$

2.B.3 Bargaining

$$\max_{w,h(\epsilon,\lambda),h_{stw}(\epsilon,\lambda)\xi(\lambda),R(\lambda)} \bar{J}^{1-\eta} \cdot \left(\bar{V} - U\right)^{\eta}$$

FOC for w:

$$-(1-\eta) \cdot \left(\frac{\bar{V} - U}{\bar{J}}\right)^{\eta} + \eta \cdot \left(\frac{\bar{J}}{\bar{V} - U}\right)^{1-\eta} = 0$$

Rearranging gives:

$$\begin{aligned} J &= \left(1-\eta\right) \cdot \bar{S} \\ \bar{V} - U &= \eta \cdot \bar{S} \end{aligned}$$

Insert into the bargaining problem:

$$\max_{h(\epsilon,\lambda),h_{stw,k}(\epsilon,\lambda),\xi(\lambda),R(\lambda)} \left(1-\eta\right)^{1-\eta} \cdot \eta^{\eta} \cdot \bar{S}$$

FOC for $h(\epsilon, \lambda)$

$$\begin{aligned} &(1-\eta)^{1-\eta} \cdot \eta^{\eta} \cdot \frac{1}{\lambda} \cdot E_K\left[\left(\alpha \cdot A \cdot \epsilon \cdot h(\epsilon,\lambda)^{-\alpha} - \phi'(h(\epsilon,\lambda))\right)\right] \cdot p(\lambda) \cdot g_{\lambda}(\epsilon) = 0 \\ \Leftrightarrow \quad & \alpha \cdot A \cdot \epsilon \cdot h(\epsilon,\lambda)^{-\alpha} = \phi'(h(\epsilon,\lambda)) \\ \Leftrightarrow \quad & \alpha \cdot A \cdot \epsilon \cdot h(\epsilon)^{-\alpha} = \phi'(h(\epsilon)) \end{aligned}$$

FOC for $h_{stw,k}(\epsilon, \lambda)$

$$\begin{aligned} &(1-\eta)^{1-\eta} \cdot \eta^{\eta} \cdot (1-\lambda)^{k} \cdot \Big(\alpha \cdot A \cdot \epsilon \cdot h_{stw,k}(\epsilon,\lambda)^{-\alpha} \\ &\qquad - \phi'(h_{stw,k}(\epsilon,\lambda)) - \tau_{stw}(k)\Big) \cdot p(\lambda) \cdot g_{\lambda}(\epsilon) = 0 \\ \Leftrightarrow \quad & \alpha \cdot A \cdot \epsilon \cdot h_{stw,k}(\epsilon,\lambda)^{-\alpha} = \phi'(h_{stw,k}(\epsilon,\lambda)) + \tau_{stw}(k) \\ \Leftrightarrow \quad & \alpha \cdot A \cdot \epsilon \cdot h_{stw,k}(\epsilon)^{-\alpha} = \phi'(h_{stw,k}(\epsilon)) + \tau_{stw}(k) \end{aligned}$$

Suppose that $\xi(\lambda) > R_{stw}(\lambda)$. Then the FOC for $\xi(\lambda)$ is:

$$\frac{p(\lambda)}{\lambda} \cdot g_{\lambda}(\xi(\lambda)) \cdot \left(z_{stw}(R(\lambda,k)) - b + \left(\lambda - \eta \cdot f\right) \cdot \bar{S}\right) = 0$$

Inserting the free entry condition and rearranging gives:

$$z_{stw}(R(\lambda,k)) - b + \frac{\left(\lambda - \eta \cdot f\right)}{1-\eta} \cdot \frac{k_v}{q} = 0$$

Suppose that $R(\lambda,k) < R_{stw}(\lambda)$ and suppose that the subsidy function is weakly decreasing over time spent on STW. Then the FOC for $R(\lambda,k)$ is:

$$(1-\lambda)^k \cdot \frac{p(\lambda)}{\lambda} \cdot g_\lambda(R_{stw}(\lambda)) \cdot \Big(z_{stw}(R(\lambda,k)) - b + \big(\lambda - \eta \cdot f\big) \cdot \bar{S}\Big) \;=\; 0$$

Inserting the free entry condition and rearranging gives:

$$z_{stw}(R(\lambda,k)) - b + \frac{\big(\lambda - \eta \cdot f\big)}{1-\eta} \cdot \frac{k_v}{q(\theta)} \;=\; 0$$

The derivations for $R(\lambda,K)$ follow the same logic and are not replicated here. Suppose that $R(\lambda,1) < R_{stw}(\lambda)$ and suppose that the subsidy function is weakly increasing over time spent on STW. Then the FOC for $R(\lambda,1)$ is:

$$\begin{aligned}\Big(\sum_{k=1}^{K-1}\lambda \cdot (1-\lambda)^{k-1} \cdot \big(z_{stw}(R(\lambda,1)) + \tau_{stw}(k) \cdot \big(\bar{h} - h_{stw}(R(\lambda,1))\big)\big)\\ +\frac{(1-\lambda)^{K-1}}{\lambda} \cdot \big(z_{stw}(R(\lambda,1)) + \tau_{stw}(K) \cdot \big(\bar{h} - h_{stw}(R(\lambda,1))\big)\big)\\ -b + \big(\lambda - \eta \cdot f\big) \cdot \bar{S}\Big) \cdot \frac{p(\lambda)}{\lambda} \cdot g_\lambda(R(\lambda,1)) \;=\; 0\end{aligned}$$

Rearranging gives:

$$\begin{aligned}&\sum_{k=1}^{K-1}\lambda \cdot (1-\lambda)^{k-1} \cdot \big[z_{stw,k}(R(\lambda,1)) + \tau_{stw}(k) \cdot \big(\bar{h} - h_{stw,k}(R(\lambda,1))\big)\big]\\ &+ (1-\lambda)^{K-1} \cdot \big[z_{stw,K}(R(\lambda,1)) + \tau_{stw}(K) \cdot \big(\bar{h} - h_{stw,K}(R(\lambda,1))\big)\big]\\ &- b + \frac{\lambda - \eta \cdot f}{1-\eta} \cdot \frac{k_v}{q(\theta)} \;=\; 0\end{aligned}$$

2.B.4 Hours Distortions

(i). The welfare costs of STW can be denoted as:

$$\tilde{\Omega}(\lambda,k) \;=\; \int_{\max\{R(\lambda,1),\ldots,R(\lambda,K)\}}^{\max\{R_{stw}(\lambda),R(\lambda,1),\ldots,R(\lambda,K)\}} \Big[z(\epsilon) - z_{stw}(\epsilon)\Big] \frac{dG_\lambda(\epsilon)}{\rho_{stw}(\lambda,k)}$$

From Nash-Bargaining, we can infer:

$$z(\epsilon) - z_{stw}(\epsilon) \;>\; 0$$

Note that in Nash-Bargaining, $h(\epsilon)$ is chosen to maximize $z(\epsilon,h(\epsilon))$. And $h_{stw}(\epsilon)$ is chosen to maximize $z(\epsilon_{stw},h_{stw}(\epsilon)) + \tau_{stw} \cdot (\bar{h} - h_{stw}(\epsilon))$. As long as $\tau_{stw} \neq 0$ we get $h(\epsilon) \neq h_{stw}(\epsilon)$. Thus, the first part of the equation must be larger than the second, which fulfills the condition. This implies that $\Omega \geq 0$.

(ii). Form the derivative of the welfare costs for the STW benefits:

$$\frac{\partial \Omega_t}{\partial \tau_{stw,t}} = -\int_{\max\{R(\lambda,1),\ldots,R(\lambda,K)\}}^{\max\{R_{stw}(\lambda),R(\lambda,1),\ldots,R(\lambda,K)\}} \frac{\partial h_{stw}(\epsilon)}{\partial \tau_{stw}} \cdot \Big(\alpha \cdot A \cdot \epsilon \cdot h_{stw}(\epsilon)^{\alpha-1} - \phi'(h_{stw}(\epsilon))\Big) dG(\epsilon) \overset{!}{>} 0$$

Using the optimality condition for the hours choice on STW, we can rewrite the equation to:

$$\frac{\partial \Omega}{\partial \tau_{stw}} = -\int_{\max\{R(\lambda,1),\ldots,R(\lambda,K)\}}^{\max\{R_{stw}(\lambda),R(\lambda,1),\ldots,R(\lambda,K)\}} \frac{\partial h_{stw}(\epsilon)}{\partial \tau_{stw}} \cdot \tau_{stw} \cdot dG(\epsilon) \overset{!}{\geq} 0$$

Since the derivation of the number of hours worked for the STW benefits is negative for $\alpha < 1$

$$\frac{\partial h_{stw}(\epsilon)}{\partial \tau_{stw}} = \frac{1}{\alpha \cdot (\alpha - 1) \cdot A \cdot \epsilon \cdot h_{stw}(\epsilon)^{\alpha-2} - \psi \cdot h_{stw}(\epsilon)^{\psi-1}} < 0$$
$$\text{for} \quad \alpha \in (0,1), \ \psi > 0, \ \epsilon > 0, \ h_{stw}(\epsilon) > 0$$

and the STW benefits τ_{stw} need to be positive, we can conclude that the derivation for the welfare costs of STW must be weakly positive.

(iii). Form the derivative of the welfare costs for the eligibility condition. When $R_{stw}(\lambda)$ is large enough then:

$$\frac{\partial \Omega}{\partial R_{stw}(\lambda)} = [z(R_{stw}(\lambda), h(R_{stw}(\lambda))) - z(R_{stw}(\lambda), h_{stw}(R_{stw}(\lambda)))] \cdot g_\lambda(R_{stw}(\lambda)) \overset{!}{\geq} 0$$
$$\Leftrightarrow z(R_{stw}(\lambda), h(R_{stw}(\lambda))) - z(R_{stw}(\lambda), h_{stw}(R_{stw}(\lambda))) \overset{!}{\geq} 0$$

As argued in (i), this equation must hold due to the Nash-Bargaining set-up. Otherwise, the derivative is zero trivially.

Appendix 2.C Ramsey Planner

2.C.1 Ramsey Problem — Job at risk observable, weakly decreasing benefits schedule

For simplicity, it is assumed that $\beta \to 1$. Further, I use a guess-and-verify approach to calculate the optimal benefits. I guess that the optimal subsidy schedule is weakly decreasing with time spent on the subsidy system: $s(\lambda, h_{stw,k}(R(\lambda,k)),k)(\lambda,k) \geq s(\lambda, h_{stw,k+1}(R(\lambda,k+1)),k+1)$. This implies that separation thresholds must be weakly increasing: $R(\lambda,k) \leq R(\lambda,k+1)$. To verify the guess, the optimal subsidy schedule calculated by solving the Ramsey problem must be weakly decreasing. Thus, the Ramsey problem becomes:

$$\begin{aligned}
\max_{s(\lambda,h,k)} \quad & n^s \cdot \int_0^1 \left(\sum_{k=1}^{K-1} (1-\lambda)^{k-1} \cdot \int_{R(\lambda,k)}^{\infty} z(\epsilon) - \Omega(\lambda,k) dG_\lambda(\epsilon) \right) dP(\lambda) d\lambda \\
& + n^s \cdot \int_0^1 \left(\frac{(1-\lambda)^{K-1}}{\lambda} \cdot \left[\left[\int_{R(\lambda,K)}^{\infty} z(\epsilon) dG_\lambda(\epsilon) - \Omega(\lambda,K) \right] \right] \right) dP(\lambda) d\lambda \\
& - (1-n) \cdot \theta \cdot k_v
\end{aligned}$$

subject to

1. The number of workers that draw a shock this period:

$$n^s = \theta \cdot q(\theta) \cdot (1-n) + \int_0^1 \sum_{k=1}^{K-1} \lambda \cdot n(\lambda,k) + \lambda \cdot n(\lambda,K) dP(\lambda)$$

2. The number of firms with shock duration $1/\lambda$, in period k after shock arrival $\forall \lambda \in [0,1]$ and $k \in \mathbb{N}_+$, $k \leq K$:

$$n(\lambda,k) = (1-\lambda)^{k-1} \cdot (1 - G(R(\lambda,k)) \cdot p(\lambda) \cdot n^s$$

3. The number of firms with shock duration $1/\lambda$, in period K or more periods after shock arrival $\forall \lambda \in [0,1]$:

$$n(\lambda,K) = \frac{(1-\lambda)^{K-1}}{\lambda} \cdot (1 - G_\lambda(R(\lambda,K)) \cdot p(\lambda) \cdot n^s$$

4. Total employment

$$n = \int_0^1 \sum_{k=1}^{K-1} n(\lambda,k) + n(\lambda,K) d\lambda$$

5. Job-creation condition:

$$\begin{aligned}\frac{1}{1-\eta}\cdot\frac{k_v}{q(\theta)} \;=\; & \int_0^1 \left(\sum_{k=1}^{K-1}(1-\lambda)^{k-1}\cdot\left[\int_{R(\lambda,k)}^{\infty} z(\epsilon)\,dG_\lambda(\epsilon)-\Omega(\lambda,k)-\frac{b}{n}\right.\right.\\ & \left.\left.+\frac{\lambda-\eta\cdot\theta\cdot q(\theta)}{1-\eta}\cdot\frac{k_v}{q(\theta)}\right]\right)dP(\lambda)\\ & +\int_0^1\left(\frac{(1-\lambda)^{K-1}}{\lambda}\cdot\left[\int_{R(\lambda,K)}^{\infty} z(\epsilon)\,dG_\lambda(\epsilon)-\Omega(\lambda,K)-\frac{b}{n}\right.\right.\\ & \left.\left.+\frac{\lambda-\eta\cdot\theta\cdot q(\theta)}{1-\eta}\cdot\frac{k_v}{q(\theta)}\right]\right)dP(\lambda)\end{aligned}$$

6. Separation Conditions for firms without access to a subsidy $\forall\lambda\in[0,1]$:

$$z(\xi(\lambda))-b+\frac{\lambda-\eta\cdot f}{1-\eta}\cdot\frac{k_v}{q(\theta)} \;=\; 0$$

7. Separation condition with access to a subsidy $\forall\lambda\in[0,1]$ and $k\in\mathbb{N}_+,\ k\le K$:

$$z_{stw,k}(R(\lambda,k))+s(\lambda,h_{stw,k}(R(\lambda,k)),k)-b+\frac{\lambda-\eta\cdot f}{1-\eta}\cdot\frac{k_v}{q(\theta)} \;=\; 0$$

2.C.1.1 FOCs

FOC for n^s:

$$\begin{aligned}\frac{\partial}{\partial n^s} =& \int_0^1\frac{1}{\lambda}\cdot E\left(\sum_{k=1}^{K-1}(1-\lambda)^{k-1}\cdot\int_{R(\lambda,k)}^{\infty} z(\epsilon)-\Omega(\lambda,k)dG_\lambda(\epsilon)\right)dP(\lambda)d\lambda\\ & +\int_0^1\left(\frac{(1-\lambda)^{K-1}}{\lambda}\cdot\int_{R(\lambda,K)}^{\infty} z(\epsilon)-\Omega(\lambda,k)dG_\lambda(\epsilon)\right)dP(\lambda)d\lambda\\ & -\int_0^1\sum_{k=1}^{K-1}(1-\lambda)^{k-1}\cdot(1-G(R(\lambda,k)))\cdot\mu(\lambda,k)\cdot dP(\lambda)d\lambda\\ & -\int_0^1\frac{(1-\lambda)^{K-1}}{\lambda}\cdot(1-G_\lambda(R(\lambda,K)))\cdot\mu(\lambda,K)\cdot dP(\lambda)d\lambda\\ & -\mu_n^s \;=\; 0\end{aligned}$$

FOC for $n(\lambda,k)$

$$\frac{\partial}{\partial n(\lambda,k)} \;=\; \lambda\cdot\mu_n^s-\mu_n(\lambda,k)+\lambda_n \;=\; 0$$

FOC for $n(\lambda,k)$

$$\frac{\partial}{\partial n(\lambda,k)} \;=\; \lambda\cdot\mu_n^s-\mu_n(\lambda,k)+\mu_n \;=\; 0$$

FOC for n

$$
\begin{aligned}
\frac{\partial}{\partial n} =\ & k_v \cdot \theta - \mu_n - \theta \cdot q(\theta) \cdot \mu_n^s \\
& + \mu_\theta \cdot \frac{b}{n^2} \cdot \int_0^1 \Bigg(\sum_{k=1}^{K-1} (1-\lambda)^{k-1} \cdot (1 - G(R(\lambda,k))) \\
& \quad + \frac{(1-\lambda)^{K-1}}{\lambda} \cdot (1 - G_\lambda(R(\lambda,K))) \Bigg) dP(\lambda) \\
=\ & 0
\end{aligned}
$$

FOC for θ

$$
\begin{aligned}
\frac{\partial}{\partial \theta} =\ & -k_v \cdot (1-n) + (1-\gamma) \cdot q(\theta) \cdot (1-n) \cdot \mu_n^s \\
& - \frac{1}{1-\eta} \cdot \gamma \cdot k_v \cdot \frac{\mu_\theta}{f} \\
& + \int_0^1 \frac{1}{1-\eta} \cdot \frac{\lambda \cdot \gamma - f \cdot \eta}{f} \\
& \cdot \left(\frac{p(\lambda)}{\lambda} \cdot (1 - E_K[G(\lambda,k)]) \cdot \mu_\theta - \mu_R(\lambda) - \mu_{R_{stw}}(\lambda) \right) d\lambda \\
=\ & 0
\end{aligned}
$$

FOC for $R(\lambda,k)$

$$
\begin{aligned}
\frac{\partial}{\partial R(\lambda,k)} =\ & -(1-\lambda)^{k-1} \cdot g_\lambda(R(\lambda,k)) \cdot p(\lambda) \cdot n^s \\
& \cdot \left(z_{stw,k}(R(\lambda,k)) + \lambda_n^s(\lambda,k) \right) \\
& - \mu_\theta \cdot (1-\lambda)^{k-1} \cdot g_\lambda(R(\lambda,k)) \cdot p(\lambda) \\
& \cdot \left(z_{stw,k}(R(\lambda,k)) - \frac{b}{n} + \frac{\lambda - \eta \cdot \theta \cdot q(\theta)}{1-\eta} \cdot \frac{k_v}{q(\theta)} \right) \\
& - \mu_R(\lambda,k) \cdot \left(A \cdot \epsilon \cdot h_{stw,k}(R(\lambda,k))^\alpha + c_f \right) \\
=\ & 0
\end{aligned}
$$

FOC for $R(\lambda,K)$

$$\begin{aligned}
\frac{\partial}{\partial R(\lambda,K)} = & -\frac{(1-\lambda)^{K-1}}{\lambda}\cdot g_\lambda(R(\lambda,K))\cdot p(\lambda)\cdot n^s \\
& \cdot\left(z_{stw,k}(R(\lambda,K))+\lambda_n^s(\lambda,K)\right) \\
& -\mu_\theta\cdot\frac{(1-\lambda)^{K-1}}{\lambda}\cdot g_\lambda(R(\lambda,K))\cdot p(\lambda) \\
& \cdot\left(z_{stw,k}(R(\lambda,K))-\frac{b}{n}+\frac{\lambda-\eta\cdot\theta\cdot q(\theta)}{1-\eta}\cdot\frac{k_v}{q(\theta)}\right) \\
& -\mu_R(\lambda,K)\cdot\left(A\cdot\epsilon\cdot h_{stw,K}(R(\lambda,K))^\alpha+c_f\right) \\
= & \ 0
\end{aligned}$$

2.C.1.2 Calculate optimal job-creation condition

Insert FOC for $n(\lambda,k)$ into FOC for n gives:

$$\begin{aligned}
\mu_n(\lambda,k) = & \ \mu_\theta\cdot\frac{b}{n^2}\cdot\int_0^1\Bigg(\sum_{k=1}^{K-1}(1-\lambda)^{k-1}\cdot(1-G(R(\lambda,k))) \\
& +\frac{(1-\lambda)^{K-1}}{\lambda}\cdot\left(1-G_\lambda(R(\lambda,K))\right)\Bigg)dP(\lambda) \\
& +k_v\cdot\theta+\left(\lambda-\theta\cdot q(\theta)\right)\cdot\mu_n^s \\
= & \ 0
\end{aligned}$$

Inserting FOC for $n(\lambda,K)$ into FOC for n gives:

$$\begin{aligned}
\mu_n(\lambda,K) = & \ \mu_\theta\cdot\frac{b}{n^2}\cdot\int_0^1\Bigg(\sum_{k=1}^{K-1}(1-\lambda)^{k-1}\cdot(1-G(R(\lambda,k))) \\
& +\frac{(1-\lambda)^{K-1}}{\lambda}\cdot\left(1-G_\lambda(R(\lambda,K))\right)\Bigg)dP(\lambda) \\
& +k_v\cdot\theta+\left(\lambda-\theta\cdot q(\theta)\right)\cdot\mu_n^s \\
= & \ 0
\end{aligned}$$

Rearrange FOC for n^s for μ_n^s:

$$\begin{aligned}
\mu_n^s = & \int_0^1\left(\sum_{k=1}^{K-1}(1-\lambda)^{k-1}\cdot\int_{R(\lambda,k)}^{\infty}(z(\epsilon)-\Omega(\lambda,k))\,dG_\lambda(\epsilon)\right)dP(\lambda)d\lambda \\
& +\int_0^1\left(\frac{(1-\lambda)^{K-1}}{\lambda}\cdot\int_{R(\lambda,K)}^{\infty}(z(\epsilon)-\Omega(\lambda,k))\,dG_\lambda(\epsilon)\right)dP(\lambda)d\lambda \\
& +\int_0^1\left(\sum_{k=1}^{K-1}(1-\lambda)^{k-1}\cdot\mu(\lambda,k)+\frac{(1-\lambda)^{K-1}}{\lambda}\cdot\mu(\lambda,K)\right)dP(\lambda)d\lambda
\end{aligned}$$

Insert $\mu_n(\lambda, k)$ and $\mu_n(\lambda, K)$:

$$\begin{aligned}
\mu_n^s \;=\; & \int_0^1 \left(\sum_{k=1}^{K-1} (1-\lambda)^{k-1} \cdot \int_{R(\lambda,k)}^{\infty} (z(\epsilon) - \Omega(\lambda,k))\, dG_\lambda(\epsilon) \right) dP(\lambda) d\lambda \\
& + \int_0^1 \left(\frac{(1-\lambda)^{K-1}}{\lambda} \cdot \int_{R(\lambda,K)}^{\infty} (z(\epsilon) - \Omega(\lambda,k))\, dG_\lambda(\epsilon) \right) dP(\lambda) d\lambda \\
& + \int_0^1 \sum_{k=1}^{K-1} (1-\lambda)^{k-1} \cdot (1 - G(R(\lambda,k))) \cdot \left(k_v \cdot \theta + (\lambda - \theta \cdot q(\theta)) \cdot \mu_n^s \right) dP(\lambda) d\lambda \\
& + \int_0^1 \frac{(1-\lambda)^{K-1}}{\lambda} \cdot \left(1 - G_\lambda(R(\lambda,K))\right) \cdot \left(k_v \cdot \theta + (\lambda - \theta \cdot q(\theta)) \cdot \mu_n^s \right) dP(\lambda) d\lambda \\
& + \mu_\theta \cdot \frac{b}{n^2} \\
& \cdot \left[\int_0^1 \left(\sum_{k=1}^{K-1} (1-\lambda)^{k-1} \cdot (1 - G(R(\lambda,k))) + \frac{(1-\lambda)^{K-1}}{\lambda} \cdot (1 - G_\lambda(R(\lambda,K))) \right) dP(\lambda) \right]^2
\end{aligned}$$

Insert definitions of n and n^s:

$$\begin{aligned}
\mu_n^s \;=\; & \int_0^1 \left(\sum_{k=1}^{K-1} (1-\lambda)^{k-1} \cdot \int_{R(\lambda,k)}^{\infty} z(\epsilon) - \Omega(\lambda,k) dG_\lambda(\epsilon) \right) dP(\lambda) d\lambda \\
& + \int_0^1 \left(\frac{(1-\lambda)^{K-1}}{\lambda} \cdot \int_{R(\lambda,K)}^{\infty} z(\epsilon) - \Omega(\lambda,k) dG_\lambda(\epsilon) \right) dP(\lambda) d\lambda \\
& + \int_0^1 \sum_{k=1}^{K-1} (1-\lambda)^{k-1} \cdot (1 - G(R(\lambda,k))) \cdot \left(k_v \cdot \theta + (\lambda - \theta \cdot q(\theta)) \cdot \mu_n^s \right) \cdot dP(\lambda) d\lambda \\
& + \int_0^1 \frac{(1-\lambda)^{K-1}}{\lambda} \cdot (1 - G_\lambda(R(\lambda,K))) \cdot \left(k_v \cdot \theta + (\lambda - \theta \cdot q(\theta)) \cdot \mu_n^s \right) \cdot dP(\lambda) d\lambda \\
& + \mu_\theta \cdot \frac{b}{(n^s)^2}
\end{aligned}$$

We can simplify the expression to:

$$\begin{aligned}
\mu_n^s \;=\; & \int_0^1 E_K \left[\int_{R(\lambda,k)}^{\infty} z(\epsilon) - \Omega(\lambda,k) dG_\lambda(\epsilon) \right] dP(\lambda) d\lambda \\
& + \int_0^1 \left(1 - E_K [G_\lambda(R(\lambda,K))] \right) \cdot \left(k_v \cdot \theta + (\lambda - \theta \cdot q(\theta)) \cdot \mu_n^s \right) \cdot dP(\lambda) d\lambda \\
& + \mu_\theta \cdot \frac{b}{(n^s)^2}
\end{aligned}$$

Rearranging the FOC for θ gives:

$$\mu_n^s \;=\; \frac{1+\chi}{1-\gamma} \cdot \frac{k_v}{q(\theta)}$$

with

$$\begin{aligned}\chi \;=\; & \frac{1}{1-\eta}\cdot\gamma\cdot k_v\cdot\frac{\mu_\theta}{f\cdot(1-n)}\\ & -\int_0^1 \frac{1}{1-\eta}\cdot\frac{\lambda\cdot\gamma-f\cdot\eta}{f\cdot(1-n)}\cdot\left(\frac{p(\lambda)}{\lambda}\cdot\big(1-E_K[G(\lambda,k)]\big)\cdot\mu_\theta\right.\\ & \left. -\mu_R(\lambda)-\mu_{R_{stw}}(\lambda)\right)d\lambda\end{aligned}$$

This allows us to derive the **constrained efficient job-creation condition**:

$$\begin{aligned}\frac{1+\chi}{1-\gamma}\cdot\frac{k_v}{q(\theta)} \;=\; & \int_0^1 E_K\left[\int_{R(\lambda,k)}^{\infty}(z(\epsilon)-\Omega(\lambda,k))\,dG_\lambda(\epsilon)\right]dP(\lambda)d\lambda\\ & +\int_0^1\Big(1-E_K[G_\lambda(R(\lambda,K))]\Big)\cdot\Bigg(\\ & \quad\frac{\lambda\cdot\gamma}{1-\gamma}\cdot\frac{k_v}{q(\theta)}-\frac{\theta\cdot q(\theta)}{1-\gamma}\cdot\frac{k_v}{q(\theta)}\\ & \quad+\frac{(\lambda-\theta\cdot q(\theta))\cdot\chi}{1-\gamma}\cdot\frac{k_v}{q(\theta)}\Bigg)dP(\lambda)d\lambda\\ & +\mu_\theta\cdot\frac{b}{(n^s)^2}\end{aligned}$$

Rearranging gives:

$$\begin{aligned}\frac{1+\chi}{1-\gamma}\cdot\frac{k_v}{q(\theta)} \;=\; & \int_0^1 E_K\left[\int_{R(\lambda,k)}^{\infty}(z(\epsilon)-\Omega(\lambda,k))\,dG_\lambda(\epsilon)\right]dP(\lambda)d\lambda\\ & +\int_0^1\big(1-E_K[G_\lambda(R(\lambda,K))]\big)\cdot\frac{1}{1-\gamma}\cdot\frac{k_v}{q(\theta)}\\ & \cdot\big(\lambda\cdot\gamma-\theta\cdot q(\theta)+(\lambda-\theta\cdot q(\theta))\cdot\chi\big)dP(\lambda)d\lambda\\ & +\mu_\theta\cdot\frac{b}{(n^s)^2}\end{aligned}$$

Next we can calculate the deviation of the constrained efficient job creation condition from the constrained efficient:

$$\begin{aligned}
&\frac{1+\chi}{1-\gamma}\cdot\frac{k_v}{q(\theta)}-\frac{1}{1-\eta}\cdot\frac{k_v}{q(\theta)}\\
&=\int_0^1 E_K\left[\int_{R(\lambda,k)}^{\infty}(z(\epsilon)-\Omega(\lambda,k))\,dG_\lambda(\epsilon)\right]dP(\lambda)d\lambda\\
&\quad+\int_0^1\big(1-E_K\left[G_\lambda(R(\lambda,K))\right]\big)\cdot\frac{1}{1-\gamma}\cdot\frac{k_v}{q(\theta)}\\
&\quad\cdot\big(\lambda-\theta\cdot q(\theta)\cdot\gamma+(\lambda-\theta\cdot q(\theta))\cdot\chi\big)\,dP(\lambda)d\lambda\\
&\quad+\mu_\theta\cdot\frac{b}{(n^s)^2}\\
&\quad-\int_0^1 E_K\left[\int_{R(\lambda,k)}^{\infty}\left(z(\epsilon)-\frac{b}{n}-\Omega(\lambda,k)\right)dG_\lambda(\epsilon)\right]dP(\lambda)d\lambda\\
&\quad-\int_0^1\big(1-E_K\left[G_\lambda(R(\lambda,K))\right]\big)\cdot\frac{1}{1-\eta}\cdot\frac{k_v}{q(\theta)}\\
&\quad\cdot\big(\lambda-\theta\cdot q(\theta)\cdot\eta\big)\,dP(\lambda)d\lambda
\end{aligned}$$

Rearranging gives:

$$\begin{aligned}
\left(\chi-\frac{\eta-\gamma}{1-\eta}\right)\cdot\frac{k_v}{q(\theta)} &= \int_0^1\frac{1}{\lambda}E_K[1-G(\lambda,k)]\cdot\frac{b}{n}+\mu_\theta\cdot\frac{b}{(n^s)^2}\\
&\quad+E_K\left[(1-G(\lambda,k))\cdot(\lambda-\theta\cdot q(\theta))\right]\cdot\left(\chi-\frac{\eta-\gamma}{1-\eta}\right)\cdot\frac{k_v}{q(\theta)}
\end{aligned}$$

This is equivalent to:

$$\begin{aligned}
\left(\chi-\frac{\eta-\gamma}{1-\eta}\right)\cdot\frac{k_v}{q(\theta)} &= \left(1+\frac{\mu_\theta}{n^s}\right)\int_0^1\frac{1}{\lambda}\cdot E_K\left[1-G(\lambda,k)\right]\cdot\frac{b}{n}\\
&\quad+\int_0^1\frac{1}{\lambda}\cdot E_K\bigg[(1-G(\lambda,k))\cdot\big(\lambda-\theta\cdot q(\theta)\big)\\
&\quad\cdot\left(\chi-\frac{\eta-\gamma}{1-\eta}\right)\cdot\frac{k_v}{q(\theta)}\bigg]dP(\lambda)
\end{aligned}$$

Finally, we can find an expression for $\left(\chi-\frac{\eta-\gamma}{1-\eta}\right)\cdot\frac{k_v}{q(\theta)}$:

$$\left(\chi - \frac{\eta - \gamma}{1 - \eta}\right) \cdot \frac{k_v}{q(\theta)}$$
$$= \frac{\displaystyle\int_0^1 \frac{1}{\lambda} \cdot \big(1 - E_K\left[G_\epsilon(\lambda, k)\right]\big)\, dP(\lambda)}{\displaystyle\int_0^1 \frac{\theta \cdot q(\theta)}{\lambda} \cdot \big(1 - E_K\left[G_\epsilon(\lambda, k)\right]\big)\, dP(\lambda) + \int_0^1 E_K\left[G_\epsilon(\lambda, k)\right] dP(\lambda)}$$
$$\cdot \left(1 + \frac{\mu_\theta}{n^s}\right) \cdot \frac{b}{n}$$

Inserting the equilibrium expression for employment n lets us simplify the equation to:

$$\left(\chi - \frac{\eta - \gamma}{1 - \eta}\right) \cdot \frac{k_v}{q(\theta)} = \left(1 + \frac{\mu_\theta}{n^s}\right) \cdot \frac{b}{f}$$

2.C.1.3 Calculate optimal separation conditions

Subtracting the decentralized separation condition simplifies the FOC for $R(\lambda, k)$:

$$\begin{aligned}
&(1 - \lambda)^{k-1} \cdot g_\lambda(R(\lambda, k)) \cdot p(\lambda) \cdot n^s \cdot \big(z_{stw,k}(R(\lambda, k)) + \mu_n^s(\lambda, k)\big) \\
&- \mu_\theta \cdot (1 - \lambda)^{k-1} \cdot g_\lambda(R(\lambda, k)) \cdot p(\lambda) \cdot \left[s(\lambda, h_{stw,k}, k) + \frac{1 - n}{n} \cdot b\right] \\
&+ \mu_R(\lambda, k) \cdot \big(A \cdot \epsilon \cdot h_{stw,k}(R(\lambda, k))^\alpha + c_f\big) = 0
\end{aligned}$$

Subtracting the decentralized separation condition simplifies the FOC for $R(\lambda, K)$:

$$\begin{aligned}
&\frac{(1 - \lambda)^{K-1}}{\lambda} \cdot g_\lambda(R(\lambda, K)) \cdot p(\lambda) \cdot n^s \cdot \big(z_{stw,k}(R(\lambda, K)) + \mu_n^s(\lambda, K)\big) \\
&- \mu_\theta \cdot \frac{(1 - \lambda)^{K-1}}{\lambda} \cdot g_\lambda(R(\lambda, K)) \cdot p(\lambda) \cdot \left[s(\lambda, h_{stw,k}, k) + \frac{1 - n}{n} \cdot b\right] \\
&+ \mu_R(\lambda, K) \cdot \big(A \cdot \epsilon \cdot h_{stw,K}(R(\lambda, K))^\alpha + c_f\big) = 0
\end{aligned}$$

Insert the Lagrange multiplier $\mu_n(\lambda, k)$ in FOC for $R(\lambda, k)$:

$$\begin{aligned}
&(1 - \lambda)^{k-1} \cdot g_\lambda(R(\lambda, k)) \cdot p(\lambda) \cdot n^s \\
&\quad \cdot \left(z_{stw,k}(R(\lambda, k)) + \frac{\mu_\theta}{n^s} \cdot \frac{b}{n} + k_v \cdot \theta + (\lambda - \theta \cdot q(\theta)) \cdot \mu_n^s\right) \\
&- \mu_\theta \cdot (1 - \lambda)^{k-1} \cdot g_\lambda(R(\lambda, k)) \cdot p(\lambda) \cdot \left[s(\lambda, h_{stw,k}, k) + \frac{1 - n}{n} \cdot b\right] \\
&+ \mu_R(\lambda, k) \cdot \big(A \cdot \epsilon \cdot h_{stw,k}(R(\lambda, k))^\alpha + c_f\big) = 0
\end{aligned}$$

Insert the Lagrange multiplier $\mu_n(\lambda, K)$ in FOC for $R(\lambda, K)$:

$$\begin{aligned}
&\frac{(1-\lambda)^{K-1}}{\lambda} \cdot g_\lambda(R(\lambda,K)) \cdot p(\lambda) \cdot n^s \\
&\quad \cdot \left(z_{stw,k}(R(\lambda,K)) + \frac{\mu_\theta}{n^s} \cdot \frac{b}{n} + k_v \cdot \theta + (\lambda - \theta \cdot q(\theta)) \cdot \mu_n^s \right) \\
&- \mu_\theta \cdot \frac{(1-\lambda)^{K-1}}{\lambda} \cdot g_\lambda(R(\lambda,K)) \cdot p(\lambda) \cdot \left[s(\lambda, h_{stw,k}, k) + \frac{1-n}{n} \cdot b \right] \\
&+ \mu_R(\lambda,K) \cdot \left(A \cdot \epsilon \cdot h_{stw,K}(R(\lambda,K))^\alpha + c_f \right) = 0
\end{aligned}$$

Insert the Lagrange multiplier μ_n^s in FOC for $R(\lambda, k)$:

$$\begin{aligned}
&(1-\lambda)^{k-1} \cdot g_\lambda(R(\lambda,k)) \cdot p(\lambda) \cdot n^s \\
&\quad \cdot \left(z_{stw,k}(R(\lambda,k)) + \frac{\mu_\theta}{n^s} \cdot b + \frac{\lambda \cdot \gamma - \theta \cdot q(\theta) + (\lambda - \theta \cdot q(\theta)) \cdot \chi}{1-\gamma} \cdot \frac{k_v}{q(\theta)} \right) \\
&- \mu_\theta \cdot (1-\lambda)^{k-1} \cdot g_\lambda(R(\lambda,k)) \cdot p(\lambda) \cdot s(\lambda, h_{stw,k}, k) \\
&+ \mu_R(\lambda,k) \cdot \left(A \cdot \epsilon \cdot h_{stw,k}(R(\lambda,k))^\alpha + c_f \right) = 0
\end{aligned}$$

Insert the Lagrange multiplier μ_n^s in FOC for $R(\lambda, K)$:

$$\begin{aligned}
&\frac{(1-\lambda)^{K-1}}{\lambda} \cdot g_\lambda(R(\lambda,K)) \cdot p(\lambda) \cdot n^s \\
&\quad \cdot \left(z_{stw,k}(R(\lambda,K)) + \frac{\mu_\theta}{n^s} \cdot b + \frac{\lambda \cdot \gamma - \theta \cdot q(\theta) + (\lambda - \theta \cdot q(\theta)) \cdot \chi}{1-\gamma} \cdot \frac{k_v}{q(\theta)} \right) \\
&- \mu_\theta \cdot \frac{(1-\lambda)^{K-1}}{\lambda} \cdot g_\lambda(R(\lambda,K)) \cdot p(\lambda) \cdot s(\lambda, h_{stw,k}, k) \\
&+ \mu_R(\lambda,K) \cdot \left(A \cdot \epsilon \cdot h_{stw,K}(R(\lambda,K))^\alpha + c_f \right) = 0
\end{aligned}$$

Subtracting the decentralized separation condition from FOC for $R(\lambda, k)$ gives:

$$\begin{aligned}
&(1-\lambda)^{k-1} \cdot g_\lambda(R(\lambda,k)) \cdot p(\lambda) \cdot n^s \\
&\quad \cdot \left(\left(1 + \frac{\mu_\theta}{n^s}\right) \cdot b - s(\lambda, h_{stw,k}, k) + (\lambda - \theta \cdot q(\theta)) \cdot \left(\chi - \frac{\eta - \gamma}{1-\eta} \right) \cdot \frac{k_v}{q(\theta)} \right) \\
&- \mu_\theta \cdot (1-\lambda)^{k-1} \cdot g_\lambda(R(\lambda,k)) \cdot p(\lambda) \cdot s(\lambda, h_{stw,k}, k) \\
&+ \mu_R(\lambda,k) \cdot \left(A \cdot \epsilon \cdot h_{stw,k}(R(\lambda,k))^\alpha + c_f \right) = 0
\end{aligned}$$

Subtracting the decentralized separation condition from FOC for $R(\lambda, K)$ gives:

$$\begin{aligned}
&\frac{(1-\lambda)^{K-1}}{\lambda} \cdot g_\lambda(R(\lambda,K)) \cdot p(\lambda) \cdot n^s \\
&\quad \cdot \left(\left(1 + \frac{\mu_\theta}{n^s}\right) \cdot b - s(\lambda, h_{stw,k}, k) + (\lambda - \theta \cdot q(\theta)) \cdot \left(\chi - \frac{\eta - \gamma}{1-\eta} \right) \cdot \frac{k_v}{q(\theta)} \right) \\
&- \mu_\theta \cdot \frac{(1-\lambda)^{K-1}}{\lambda} \cdot g_\lambda(R(\lambda,K)) \cdot p(\lambda) \cdot s(\lambda, h_{stw,k}, k) \\
&+ \mu_R(\lambda,K) \cdot \left(A \cdot \epsilon \cdot h_{stw,K}(R(\lambda,K))^\alpha + c_f \right) = 0
\end{aligned}$$

Subtracting the decentralized separation condition from FOC for $R(\lambda, k)$ gives:

$$\begin{aligned}&(1-\lambda)^{k-1}\cdot g_\lambda(R(\lambda,k))\cdot p(\lambda)\cdot n^s\cdot(\lambda-\theta\cdot q(\theta))\\&\quad\cdot\left(\left(1+\frac{\mu_\theta}{n^s}\right)\cdot b-s(\lambda,h_{stw,k},k)+\left(\chi-\frac{\eta-\gamma}{1-\eta}\right)\cdot\frac{k_v}{q(\theta)}\right)\\&-\mu_\theta\cdot(1-\lambda)^{k-1}\cdot g_\lambda(R(\lambda,k))\cdot p(\lambda)\cdot s(\lambda,h_{stw,k},k)\\&+\mu_R(\lambda,k)\cdot\left(A\cdot\epsilon\cdot h_{stw,k}(R(\lambda,k))^\alpha+c_f\right) \;=\; 0\end{aligned}$$

Subtracting the decentralized separation condition from FOC for $R(\lambda, K)$ gives:

$$\begin{aligned}&\frac{(1-\lambda)^{K-1}}{\lambda}\cdot g_\lambda(R(\lambda,K))\cdot p(\lambda)\cdot n^s\\&\quad\cdot\left(\left(1+\frac{\mu_\theta}{n^s}\right)\cdot b-s(\lambda,h_{stw,k},K)+(\lambda-\theta\cdot q(\theta))\cdot\left(\chi-\frac{\eta-\gamma}{1-\eta}\right)\cdot\frac{k_v}{q(\theta)}\right)\\&-\mu_\theta\cdot\frac{(1-\lambda)^{K-1}}{\lambda}\cdot g_\lambda(R(\lambda,K))\cdot p(\lambda)\cdot s(\lambda,h_{stw,k},K)\\&+\mu_R(\lambda,K)\cdot\left(A\cdot\epsilon\cdot h_{stw,K}(R(\lambda,K))^\alpha+c_f\right) \;=\; 0\end{aligned}$$

Rearranging the FOC for $R(\lambda, k)$ gives:

$$\mu_R(\lambda,k) \;=\; \frac{(1-\lambda)^{k-1}\cdot g_\lambda(R(\lambda,k))\cdot p(\lambda)}{\left(A\cdot\epsilon\cdot h_{stw,k}(R(\lambda,k))^\alpha+c_f\right)}\cdot n^s\cdot\left(1+\frac{\mu_\theta}{n^s}\right)\cdot\left(\frac{\lambda}{f}\cdot b-s(\lambda,h_{stw,k},k)\right)$$

Rearranging the FOC for $R(\lambda, K)$ gives:

$$\mu_R(\lambda,K) \;=\; \frac{\frac{(1-\lambda)^{K-1}}{\lambda}\cdot g_\lambda(R(\lambda,K))\cdot p(\lambda)}{\left(A\cdot\epsilon\cdot h_{stw,K}(R(\lambda,K))^\alpha+c_f\right)}\cdot n^s\cdot\left(1+\frac{\mu_\theta}{n^s}\right)\cdot\left(\frac{\lambda}{f}\cdot b-s(\lambda,h_{stw,K},K)\right)$$

2.C.1.4 Optimal Subsidy — Jobs at Risk and Shock Persistence Observable

FOC for subsidy under full information:

$$\frac{\partial}{\partial s(\lambda,h,k)} \;=\; -\mu_R(\lambda,k) \;=\; 0$$

The optimal subsidy thus is:

$$s(\lambda,h,k) \;=\; s^*(\lambda) \;=\; \frac{\lambda}{f}\cdot b$$

Note that the Lagrange multiplier for the separation condition is zero. Thus, we can infer that the full information subsidy can implement the optimal separation condition, proofing Corollary 1. The remainder of Corollary 1 is trivial.

2.C.1.5 Optimal STW benefits — Jobs at Risk and Shock Persistence Observable

FOC for STW benefits with horizon k under full information:

$$\frac{\partial}{\partial \tau_{stw}(\lambda,k)} = -\mu_R(\lambda,k)\cdot\left(\bar{h}-h_{stw,k}(R(\lambda,k))\right) - (1-\lambda)^{k-1}\cdot p(\lambda)\cdot\left(1+\frac{\mu_\theta}{n^s}\right)\cdot n^s\cdot\frac{\partial\Omega(\lambda,k)}{\partial\tau_{stw}(\lambda,k)} = 0$$

Insert the Lagrange multiplier for the FOC for $R(\lambda,k)$:

$$\begin{aligned}&(1-\lambda)^{k-1}\cdot g_\lambda(R(\lambda,k))\cdot\left(-\frac{\partial R(\lambda,k)}{\partial\tau_{stw}(\lambda,k)}\right)\cdot p(\lambda)\cdot n^s\cdot\left(1+\frac{\mu_\theta}{n^s}\right)\\&\cdot\left(\frac{\lambda}{f}\cdot b-\tau_{stw}(\lambda,k)\cdot\left(\bar{h}-h_{stw,k}(R(\lambda,k))\right)\right)\\&-(1-\lambda)^{k-1}\cdot p(\lambda)\cdot\left(1+\frac{\mu_\theta}{n^s}\right)\cdot n^s\cdot\frac{\partial\Omega(\lambda,k)}{\partial\tau_{stw}(\lambda,k)} = 0\end{aligned}$$

with

$$\frac{\partial R(\lambda,k)}{\partial\tau_{stw}(\lambda,k)} = -\frac{\bar{h}-h_{stw,k}(R(\lambda,k))}{A\cdot\epsilon\cdot h_{stw,k}(R(\lambda,k))^\alpha+c_f}$$

The optimal subsidy thus is:

$$\tau_{stw}(\lambda,k)\cdot\left(\bar{h}-h_{stw,k}(R(\lambda,k))\right) = \frac{\lambda}{f}\cdot b-\frac{n_{stw}(\lambda,k)}{\rho_{stw}(\lambda,k)}\cdot\frac{\partial\Omega(\lambda,k)}{\partial\tau_{stw}(\lambda,k)}\Big/\frac{\partial n(\lambda,k)}{\partial\tau_{stw}(\lambda,k)}$$

with

$$\frac{\partial n(\lambda,k)}{\partial\tau_{stw}(\lambda,k)} = (1-\lambda)^{k-1}\cdot g_\lambda(R(\lambda,k))\cdot\left(-\frac{\partial R(\lambda,k)}{\partial\tau_{stw}(\lambda,k)}\right)\cdot p(\lambda)\cdot n^s$$

Using Definition 1 we can rewrite the expression to:

$$\tau_{stw}(\lambda,k)\cdot\left(\bar{h}-h_{stw,k}(R(\lambda,k))\right) = \frac{\lambda}{f}\cdot b-\left(n_{stw}(\lambda,k)\Big/\frac{\partial n(\lambda,k)}{\partial\tau_{stw}(\lambda,k)}\right)\cdot\frac{\partial\tilde{\Omega}(\lambda,k)}{\partial\tau_{stw}(\lambda,k)}$$

Note that $\tau_{stw}(\lambda,k)$ is independent of k. To prove this, we can use the guess-and-verify approach. We guess that $\tau_{stw}(\lambda,k)=\tau_{stw}(\lambda,k)\ \ \forall\ \ k\in\mathbb{N}_+,\ k\le K$. This implies that $R(\lambda,k)=R(\lambda)$ and $h_{stw,k}(R(\lambda,k))=h_{stw}(R(\lambda))\ \ \forall\ \ k\in\mathbb{N}_+,\ k\le K$ as well. Trivially, this implies that

$$\frac{\partial\Omega(\lambda,k)}{\partial\tau_{stw}(\lambda,k)} = \frac{\partial\Omega(\lambda)}{\partial\tau_{stw}(\lambda)}\ \ \forall\ \ k\ \in\ \mathbb{N}_+,\ k\ \le\ K$$

Finally, we have to check whether the following quotient is independent of k:

$$n_{stw}(\lambda,k)/\frac{\partial n(\lambda,k)}{\partial \tau_{stw}(\lambda,k)} = \frac{(1-\lambda)^{k-1}\cdot \rho_{stw}(\lambda,k)\cdot p(\lambda)\cdot n^s}{(1-\lambda)^{k-1}\cdot g_\lambda(R(\lambda,k))\cdot\left(-\frac{\partial R(\lambda,k)}{\partial \tau_{stw}(\lambda,k)}\right)\cdot p(\lambda)\cdot n^s}$$
$$= \frac{\rho_{stw}(\lambda,k)}{g_\lambda(R(\lambda,k))\cdot\left(-\frac{\partial R(\lambda,k)}{\partial \tau_{stw}(\lambda,k)}\right)}$$

Knowing that the separation threshold and the hours choice are independent of k as long as τ_{stw} is independent of k implies trivially that we can write the quotient independent of k:

$$n_{stw}(\lambda,k)/\frac{\partial n(\lambda,k)}{\partial \tau_{stw}(\lambda,k)} = \frac{\rho_{stw}(\lambda)}{g_\lambda(R(\lambda))\cdot\left(-\frac{\partial R(\lambda)}{\partial \tau_{stw}(\lambda)}\right)}$$

We can rewrite the condition for optimal STW benefits as:

$$\tau_{stw}(\lambda,k)\cdot\left(\bar{h}-h_{stw,k}(R(\lambda,k))\right) = \tau_{stw}(\lambda)\cdot\left(\bar{h}-h_{stw}(R(\lambda))\right)$$
$$= \frac{\lambda}{f}\cdot b-\left(n_{stw}(\lambda)/\frac{\partial n(\lambda)}{\partial \tau_{stw}(\lambda)}\right)\cdot\frac{\partial\tilde{\Omega}(\lambda)}{\partial \tau_{stw}(\lambda)}$$

2.C.1.6 Optimal Subsidy — Jobs at Risk, not Shock Persistence Observable

FOC for subsidy when only jobs at risk are observable for firms that spend $k \leq K$ periods on the subsidy scheme:

$$\frac{\partial}{\partial s} = -\int_0^1\sum_{k=1}^{K}\mu_R(\lambda,k)d\lambda = 0$$

Insert for the Lagrange multiplier:

$$\int_0^1\sum_{k=1}^{K}(1-\lambda)^{k-1}\cdot g_\lambda(R(\lambda,k))\cdot p(\lambda)\cdot\frac{\partial R(\lambda,k)}{\partial s}\cdot n^s\cdot\left(1+\frac{\mu_\theta}{n^s}\right)\cdot\left(\frac{\lambda}{f}\cdot b-s\right)d\lambda$$
$$+\int_0^1\frac{(1-\lambda)^{K-1}}{\lambda}\cdot g_\lambda(R(\lambda,k))\cdot p(\lambda)\cdot\frac{\partial R(\lambda,k)}{\partial s}\cdot n^s\cdot\left(1+\frac{\mu_\theta}{n^s}\right)\cdot\left(\frac{\lambda}{f}\cdot b-s\right)d\lambda$$
$$= 0$$

with

$$\frac{\partial R(\lambda,k)}{\partial s} = -\frac{1}{A\cdot\epsilon\cdot h_{stw,k}(R(\lambda,k))^\alpha+c_f}$$

Note that a uniform subsidy implies that the separation threshold does not change with time spent on the subsidy system:

$$R(\lambda, k) = R(\lambda)$$

This lets us simplify the FOC for the subsidy to:

$$\int_0^1 \frac{1}{\lambda} \cdot g_\lambda(R(\lambda)) \cdot p(\lambda) \cdot \frac{\partial R(\lambda)}{\partial s} \cdot n^s \cdot \left(\frac{\lambda}{f} \cdot b - s\right) d\lambda = 0$$

Rearranging for s gives:

$$s = \int_0^1 \frac{\frac{1}{\lambda} \cdot g_\lambda(R(\lambda)) \cdot p(\lambda) \cdot \frac{\partial R(\lambda)}{\partial s} \cdot n^s}{\int_0^1 \frac{1}{\lambda} \cdot g_\lambda(R(\lambda)) \cdot p(\lambda) \cdot \frac{\partial R(\lambda)}{\partial s} \cdot n^s \cdot d\lambda} \cdot \left(\frac{\lambda}{f} \cdot b\right) d\lambda$$

Note that $s(\lambda) = \frac{\lambda}{f} \cdot b$. Denote $\omega(\lambda)$ as weight for the optimal subsidy:

$$\begin{aligned}
\omega(\lambda) &= \frac{\frac{1}{\lambda} \cdot g_\lambda(R(\lambda)) \cdot p(\lambda) \cdot \frac{\partial R(\lambda)}{\partial s} \cdot n^s}{\int_0^1 \frac{1}{\lambda} \cdot g_\lambda(R(\lambda)) \cdot p(\lambda) \cdot \frac{\partial R(\lambda)}{\partial s} \cdot n^s \cdot d\lambda} \\
&= \frac{\frac{\partial n(\lambda)}{\partial s}}{\int_0^1 \frac{\partial n(\lambda)}{\partial s} \cdot d\lambda}
\end{aligned}$$

This gives:

$$s = E_\omega[s(\lambda)] = \int_0^1 \omega(\lambda) s(\lambda) d\lambda$$

2.C.1.7 Optimal STW Benefits — Jobs at Risk, not Shock Persistence Observable

FOC for subsidy when only jobs at risk are observable:

$$\begin{aligned}
\frac{\partial}{\partial \tau_{stw}} = &- \int_0^1 \sum_{k=1}^{K} \mu_R(\lambda, k) \cdot \left(\bar{h} - h_{stw,k}(R(\lambda, k))\right) d\lambda \\
&- \int_0^1 \sum_{k=1}^{K-1} (1-\lambda)^{k-1} \cdot p(\lambda) \cdot \left(1 + \frac{\mu_\theta}{n^s}\right) \cdot n^s \cdot \frac{\partial \Omega(\lambda, k)}{\partial \tau_{stw}} \cdot d\lambda \\
&- \frac{(1-\lambda)^{K-1}}{\lambda} \cdot p(\lambda) \cdot \left(1 + \frac{\mu_\theta}{n^s}\right) \cdot n^s \cdot \frac{\partial \Omega(\lambda, K)}{\partial \tau_{stw}} = 0
\end{aligned}$$

Insert for the Lagrange multiplier:

$$\int_0^1 \sum_{k=1}^{K-1} (1-\lambda)^{k-1} \cdot g_\lambda(R(\lambda,k)) \cdot p(\lambda) \cdot \frac{\partial R(\lambda,k)}{\partial \tau_{stw}} \cdot n^s \cdot \left(1 + \frac{\mu_\theta}{n^s}\right)$$
$$\cdot \left(\frac{\lambda}{f} \cdot b - \tau_{stw} \cdot \left(\bar{h} - h_{stw}(R(\lambda,k))\right)\right) d\lambda$$
$$+ \int_0^1 \frac{(1-\lambda)^{K-1}}{\lambda} \cdot g_\lambda(R(\lambda,k)) \cdot p(\lambda) \cdot \frac{\partial R(\lambda,k)}{\partial \tau_{stw}} \cdot n^s \cdot \left(1 + \frac{\mu_\theta}{n^s}\right)$$
$$\cdot \left(\frac{\lambda}{f} \cdot b - \tau_{stw} \cdot \left(\bar{h} - h_{stw}(R(\lambda,k))\right)\right) d\lambda$$
$$- \int_0^1 \sum_{k=1}^{K-1} (1-\lambda)^{k-1} \cdot p(\lambda) \cdot \left(1 + \frac{\mu_\theta}{n^s}\right) \cdot n^s \cdot \frac{\partial \Omega(\lambda,k)}{\partial \tau_{stw}} \cdot d\lambda$$
$$- \int_0^1 \frac{(1-\lambda)^{K-1}}{\lambda} \cdot p(\lambda) \cdot \left(1 + \frac{\mu_\theta}{n^s}\right) \cdot n^s \cdot \frac{\partial \Omega(\lambda,K)}{\partial \tau_{stw}} \cdot d\lambda = 0$$

Note that uniform STW benefits imply that the separation threshold does not change with time spent on the subsidy system.

$$R(\lambda,k) = R(\lambda)$$

This also applies to the welfare costs of STW:

$$\frac{\partial \Omega(\lambda,k)}{\partial \tau_{stw}} = \frac{\partial \Omega(\lambda,k)}{\partial \tau_{stw}}$$

This simplifies the FOC for the STW benefits to:

$$\int_0^1 \frac{1}{\lambda} \cdot g_\lambda(R(\lambda)) \cdot p(\lambda) \cdot \frac{\partial R(\lambda)}{\partial \tau_{stw}} \cdot n^s \cdot \left(\frac{\lambda}{f} \cdot b - \tau_{stw} \cdot \left(\bar{h} - h_{stw}(R(\lambda))\right)\right) d\lambda$$
$$- \int_0^1 \frac{1}{\lambda} \cdot p(\lambda) \cdot n^s \cdot \frac{\partial \Omega(\lambda)}{\partial \tau_{stw}} \cdot d\lambda = 0$$

Rearranging for the net transfers gives:

$$\tau_{stw} \cdot \left(\bar{h} - h_{stw}(R(\lambda))\right)$$
$$= \int_0^1 \frac{\frac{1}{\lambda} \cdot g_\lambda(R(\lambda)) \cdot p(\lambda) \cdot \frac{\partial R(\lambda)}{\partial \tau_{stw}} \cdot n^s}{\int_0^1 \frac{1}{\lambda} \cdot g_\lambda(R(\lambda)) \cdot p(\lambda) \cdot \frac{\partial R(\lambda,k)}{\partial \tau_{stw}} \cdot n^s \cdot d\lambda}$$
$$\cdot \left(\frac{\lambda}{f} \cdot b - \left(\frac{n_{stw}(\lambda)}{\frac{\partial n(\lambda)}{\partial \tau_{stw}(\lambda)}}\right) \cdot \frac{\partial \tilde{\Omega}(\lambda)}{\partial \tau_{stw}(\lambda)} \cdot \frac{\partial \Omega(\lambda)}{\partial \tau_{stw}}\right) d\lambda$$

Note that $s(\lambda) = \frac{\lambda}{f} \cdot b$. Denote $\omega^{stw}(\lambda)$ as weight for the optimal subsidy:

$$\omega^{stw}(\lambda) = \frac{\frac{1}{\lambda} \cdot g_\lambda(R(\lambda)) \cdot p(\lambda) \cdot \frac{\partial R(\lambda)}{\partial \tau_{stw}} \cdot n^s}{\int_0^1 \frac{1}{\lambda} \cdot g_\lambda(R(\lambda)) \cdot p(\lambda) \cdot \frac{\partial R(\lambda)}{\partial \tau_{stw}} \cdot n^s \cdot d\lambda} = \frac{\frac{\partial n(\lambda)}{\partial \tau_{stw}}}{\int_0^1 \frac{\partial n(\lambda)}{\partial \tau_{stw}} \cdot d\lambda}$$

This gives:

$$\int_0^1 \omega^{stw}(\lambda) \cdot \tau_{stw} \cdot \left(\bar{h} - h_{stw}(R(\lambda))\right) d\lambda = \int_0^1 \omega^{stw}(\lambda) \left[\frac{\lambda}{f} \cdot b - \left(n_{stw}(\lambda) / \frac{\partial n(\lambda)}{\partial \tau_{stw}(\lambda)}\right) \cdot \frac{\partial \tilde{\Omega}(\lambda)}{\partial \tau_{stw}(\lambda)}\right] d\lambda$$

or equivalently:

$$E_{\omega^{stw}}\left[\tau_{stw,\omega} \cdot \left(\bar{h} - h_{stw}(R(\lambda))\right)\right] = E_{\omega^{stw}}\left[\frac{\lambda}{f} \cdot b - \left(n_{stw}(\lambda) / \frac{\partial n(\lambda)}{\partial \tau_{stw}(\lambda)}\right) \cdot \frac{\partial \tilde{\Omega}(\lambda)}{\partial \tau_{stw}(\lambda)}\right]$$

Using the notation of the main part, we can also express this as dependent on the optimal STW benefits without information asymmetry:

$$E_{\omega^{stw}}\left[\tau_{stw} \cdot \left(\bar{h} - h_{stw}(R(\lambda))\right)\right] = E_{\omega^{stw}}\left[\tau_{stw}(\lambda) \cdot \left(\bar{h} - h_{stw}(R(\lambda))\right)\right]$$

2.C.1.8 Optimal Subsidy — Time dependent System

FOC for subsidy when only jobs at risk are observable for firms that spend $k \leq K$ periods on the subsidy scheme:

$$\frac{\partial}{\partial s(k)} = -\int_0^1 \mu_R(\lambda, k) d\lambda = 0$$

Insert for the Lagrange multiplier:

$$\int_0^1 (1-\lambda)^{k-1} \cdot g_\lambda(R(\lambda,k)) \cdot p(\lambda) \cdot \frac{\partial R(\lambda,k)}{\partial s(k)} \cdot n^s \cdot \left(1 + \frac{\mu_\theta}{n^s}\right) \cdot \left(\frac{\lambda}{f} \cdot b - s(k)\right) d\lambda = 0$$

with

$$\frac{\partial R(\lambda,k)}{\partial s(k)} = -\frac{1}{A \cdot \epsilon \cdot h_{stw,k}(R(\lambda,k))^\alpha + c_f}$$

Rearranging for $s(k)$ gives:

$$s(k) = \int_0^1 \frac{(1-\lambda)^{k-1} \cdot g_\lambda(R(\lambda,k)) \cdot p(\lambda) \cdot \frac{\partial R(\lambda,k)}{\partial s(k)} \cdot n^s}{\int_0^1 (1-\lambda)^{k-1} \cdot g_\lambda(R(\lambda,k)) \cdot p(\lambda) \cdot \frac{\partial R(\lambda,k)}{\partial s(k)} \cdot n^s \cdot d\lambda} \cdot \left(\frac{\lambda}{f} \cdot b\right) d\lambda$$

Note that $s(\lambda) = \frac{\lambda}{f} \cdot b$. Denote $\omega(\lambda, k)$ as weight for the optimal subsidy:

$$\begin{aligned}
\omega_k(\lambda) &= \frac{(1-\lambda)^{k-1} \cdot g_\lambda(R(\lambda,k)) \cdot p(\lambda) \cdot \frac{\partial R(\lambda,k)}{\partial s(k)} \cdot n^s}{\int_0^1 (1-\lambda)^{k-1} \cdot g_\lambda(R(\lambda,k)) \cdot p(\lambda) \cdot \frac{\partial R(\lambda,k)}{\partial s(k)} \cdot n^s \cdot d\lambda} \\
&= \frac{\frac{\partial n(\lambda,k)}{\partial s(k)}}{\int_0^1 \frac{\partial n(\lambda,k)}{\partial s(k)} \cdot d\lambda}
\end{aligned}$$

This gives:

$$s(k) = E_{\omega_k}[s(\lambda,k)] = \int_0^1 \omega_k(\lambda) s(\lambda,k) d\lambda$$

FOC for subsidy when only jobs at risk are observable for firms that spend at least K periods on the subsidy scheme:

$$\frac{\partial}{\partial s(k)} = -\int_0^1 \mu_R(\lambda, K) d\lambda = 0$$

Following the steps above, we can define weights

$$\omega_K(\lambda) = \frac{\frac{(1-\lambda)^{K-1}}{\lambda} \cdot g_\lambda(R(\lambda,K)) \cdot p(\lambda) \cdot \frac{\partial R(\lambda,K)}{\partial \tau_{stw}(K)} \cdot n^s}{\int_0^1 \frac{(1-\lambda)^{K-1}}{\lambda} \cdot g_\lambda(R(\lambda,K)) \cdot p(\lambda) \cdot \frac{\partial R(\lambda,K)}{\partial s(K)} \cdot n^s \cdot d\lambda} = \frac{\frac{\partial n(\lambda,K)}{\partial s(K)}}{\int_0^1 \frac{\partial n(\lambda,K)}{\partial s(K)} \cdot d\lambda}$$

so that we can express $s(K)$ as:

$$s(K) = E_{\omega_K}[s(\lambda,K)] = \int_0^1 \omega_K(\lambda) s(\lambda,K) d\lambda$$

2.C.2 Corollary 2

First, we have to show that the weight distribution in k+1 is first order stochastically dominated by the weight distribution in k:

$$\int_0^x \omega_k(\lambda) d\lambda \leq \int_0^x \omega_{k+1}(\lambda) d\lambda \quad \forall x \in [0,1]$$

$$\int_0^x \frac{\frac{\partial n(\lambda,k)}{\partial \tau_{stw}}}{\int_0^1 \frac{\partial n(\lambda,k)}{\partial \tau_{stw}} d\lambda} d\lambda \;\leq\; \int_0^x \frac{\frac{\partial n(\lambda,k+1)}{\partial \tau_{stw}}}{\int_0^1 \frac{\partial n(\lambda,k+1)}{\partial \tau_{stw}} d\lambda} d\lambda \quad \forall x \in [0,1]$$

$$\int_0^x \frac{\frac{\partial n(\lambda,k)}{\partial \tau_{stw}}}{\int_0^1 \frac{\partial n(\lambda,k)}{\partial \tau_{stw}} d\lambda} d\lambda \;\leq\; \int_0^x \frac{(1-\lambda)\cdot \frac{\partial n(\lambda,k)}{\partial \tau_{stw}}}{\int_0^1 (1-\lambda)\cdot \frac{\partial n(\lambda,k)}{\partial \tau_{stw}} d\lambda} d\lambda \quad \forall x \in [0,1]$$

Define

$$\alpha(\lambda) = 1-\lambda$$

Note that $\alpha'(\lambda) = -1 < 0$. Further define

$$f(\lambda) = \frac{\frac{\partial n(\lambda,k)}{\partial \tau_{stw}}}{\int_0^1 \frac{\partial n(\lambda,k)}{\partial \tau_{stw}} d\lambda} d\lambda$$

Then we can rewrite the inequality as:

$$\int_0^x f(\lambda) d\lambda \;\leq\; \int_0^x \frac{\alpha(\lambda)\cdot f(\lambda)}{\int_0^1 \alpha(\lambda)\cdot f(\lambda) d\lambda} \cdot d\lambda \quad \forall x \in [0,1]$$

Define

$$B(x) = \int_0^x \left(\frac{\alpha(\lambda)}{\int_0^1 \alpha(\lambda)\cdot f(\lambda) d\lambda} - 1 \right) \cdot f(\lambda)\cdot d\lambda \;\geq\; 0 \quad \forall x \in [0,1]$$

Note that since $\alpha'(\lambda) < 0$ there must exist λ^* such that

$$\alpha(\lambda^*) = \int_0^1 \alpha(\lambda)\cdot f(\lambda) d\lambda$$

This implies that $B(x) \geq 0 \ \forall x \in [0,\lambda^*)$ as

$$\left(\frac{\alpha(\lambda)}{\int_0^1 \alpha(\lambda)\cdot f(\lambda) d\lambda} - 1 \right) \;\geq\; 0 \qquad \forall \lambda \in [0,\lambda^*)$$

$\forall x \in (\lambda^*, 1]$ note that $B'(x) > 0$ as

$$\left(\frac{\alpha(\lambda)}{\int_0^1 \alpha(\lambda)\cdot f(\lambda) d\lambda} - 1 \right) \;<\; 0 \qquad \forall \lambda \in (\lambda^*, 1]$$

Further, note that $B(1) = 0$. $B'(x) < 0$ and $B(1) = 0$ imply together that $B(x) \geq 0$ $\forall x \in [0, \lambda^*)$ This proves that the weight distribution in k+1 is first order stochastically dominated by the weight distribution in k. This implies that:

$$E_{\omega_k}[s(\lambda)] \;\geq\; E_{\omega_{k+1}}[s(\lambda)]$$

Thus, the subsidy function must be falling with time spent on STW! This also proves the guess-and-verify!

2.C.2.1 Optimal STW Benefits — Time dependent System

FOC for subsidy when only jobs at risk are observable:

$$\begin{aligned}\frac{\partial}{\partial \tau_{stw}(k)} &= -\int_0^1 \mu_R(\lambda,k)\cdot\left(\bar{h} - h_{stw,k}(R(\lambda,k))\right)d\lambda \\ &\quad - (1-\lambda)^{k-1}\cdot p(\lambda)\cdot\left(1+\frac{\mu_\theta}{n^s}\right)\cdot n^s\cdot\frac{\partial\Omega(\lambda,k)}{\partial\tau_{stw}(\lambda,k)} \;=\; 0\end{aligned}$$

Insert for the Lagrange multiplier:

$$\begin{aligned}&\int_0^1 (1-\lambda)^{k-1}\cdot g_\lambda(R(\lambda,k))\cdot p(\lambda)\cdot\frac{\partial R(\lambda,k)}{\partial\tau_{stw}(k)}\cdot n^s\cdot\left(1+\frac{\mu_\theta}{n^s}\right)\\ &\quad\cdot\left(\frac{\lambda}{f}\cdot b - \tau_{stw}(k)\cdot\left(\bar{h}-h_{stw}(R(\lambda,k))\right)\right)d\lambda\\ &-\int_0^1 (1-\lambda)^{k-1}\cdot p(\lambda)\cdot\left(1+\frac{\mu_\theta}{n^s}\right)\cdot n^s\cdot\frac{\partial\Omega(\lambda,k)}{\partial\tau_{stw}(\lambda,k)}\cdot d\lambda \;=\; 0\end{aligned}$$

Rearranging for the net transfers gives:

$$\begin{aligned}&\tau_{stw}(k)\cdot\left(\bar{h}-h_{stw}(R(\lambda,k))\right)\\ &= \int_0^1 \frac{(1-\lambda)^{k-1}\cdot g_\lambda(R(\lambda,k))\cdot p(\lambda)\cdot\frac{\partial R(\lambda,k)}{\partial\tau_{stw}(k)}\cdot n^s}{\int_0^1(1-\lambda)^{k-1}\cdot g_\lambda(R(\lambda,k))\cdot p(\lambda)\cdot\frac{\partial R(\lambda,k)}{\partial\tau_{stw}(k)}\cdot n^s\cdot d\lambda}\\ &\quad\cdot\left(\frac{\lambda}{f}\cdot b-\left(n_{stw}(\lambda,k)/\frac{\partial n(\lambda,k)}{\partial\tau_{stw}(k)}\right)\cdot\frac{\partial\tilde{\Omega}(\lambda,k)}{\partial\tau_{stw}(k)}\right)\cdot d\lambda\end{aligned}$$

Note that $s(\lambda) = \frac{\lambda}{f}\cdot b$. Denote $\omega_k(\lambda)$ as weight for the optimal subsidy:

$$\begin{aligned}\omega_k(\lambda) &= \frac{(1-\lambda)^{k-1}\cdot g_\lambda(R(\lambda,k))\cdot p(\lambda)\cdot\frac{\partial R(\lambda,k)}{\partial\tau_{stw}(k)}\cdot n^s}{\int_0^1(1-\lambda)^{k-1}\cdot g_\lambda(R(\lambda,k))\cdot p(\lambda)\cdot\frac{\partial R(\lambda,k)}{\partial\tau_{stw}(k)}\cdot n^s\cdot dP(\lambda)}\\ &= \frac{\frac{\partial n(\lambda,k)}{\partial\tau_{stw}(k)}}{\int_0^1\frac{\partial n(\lambda,k)}{\partial\tau_{stw}(k)}\cdot d\lambda}\end{aligned}$$

This gives:

$$\int_0^1 \omega_k(\lambda)\cdot\tau_{stw}(k)\cdot\left(\bar{h}-h_{stw}(R(\lambda,k))\right)d\lambda$$
$$= \int_0^1 \omega_k(\lambda)\left[\frac{\lambda}{f}\cdot b-\left(n_{stw}(\lambda,k)/\frac{\partial n(\lambda,k)}{\partial\tau_{stw}(k)}\right)\cdot\frac{\partial\tilde{\Omega}(\lambda,k)}{\partial\tau_{stw}(k)}\right]d\lambda$$

or equivalently:

$$E_{\omega_k}\left[\tau_{stw}(k)\cdot\left(\bar{h}-h_{stw}(R(\lambda,k))\right)\right]$$
$$= E_{\omega_k}\left[\frac{\lambda}{f}\cdot b-\left(n_{stw}(\lambda,k)/\frac{\partial n(\lambda,k)}{\partial\tau_{stw}(k)}\right)\cdot\frac{\partial\tilde{\Omega}(\lambda,k)}{\partial\tau_{stw}(k)}\right]$$

Again, we can express this dependent on the optimal STW benefits without information asymmetry:

$$E_{\omega_k}\left[\tau_{stw}(k)\cdot\left(\bar{h}-h_{stw}(R(\lambda,k))\right)\right] = E_{\omega_k}\left[\tau_{stw}(\lambda,k)\cdot\left(\bar{h}-h_{stw}(R(\lambda,k))\right)\right]$$

Accordingly, the optimal benefits for k = K can be calculated.

2.C.3 Ramsey Problem — Jobs at risk observable, weakly increasing benefits schedule

Assumption: $\beta=1$.
Guess: $s(\lambda,h_{stw,k}(R(\lambda,k)),k)(\lambda,k)\le s(\lambda,h_{stw,k+1}(R(\lambda,k+1)),k+1)$
This implies: $R(\lambda,k)\ge R(\lambda,k+1)$
Then the problem of the Ramsey planner becomes:

$$\max_{s(\lambda,h,l)} \quad n^s\cdot\int_0^1\left(\sum_{k=1}^{K-1}(1-\lambda)^{k-1}\cdot\int_{R(\lambda,1)}^{\infty} z(\epsilon)-\Omega(\lambda,k)dG_\lambda(\epsilon)\right)dP(\lambda)d\lambda$$
$$+n^s\cdot\int_0^1\left(\frac{(1-\lambda)^{K-1}}{\lambda}\cdot\left[\int_{R(\lambda,1)}^{\infty} z(\epsilon)dG_\lambda(\epsilon)-\Omega(\lambda,K)\right]\right)dP(\lambda)d\lambda$$
$$-(1-n)\cdot\theta\cdot k_v$$

subject to

1. The number of workers that draw a shock this period:

$$n^s = \theta\cdot q(\theta)\cdot(1-n)+\int_0^1\sum_{k=1}^{K-1}\lambda\cdot n(\lambda,k)+\lambda\cdot n(\lambda,K)dP(\lambda)$$

2. The number of firms with shock duration $1/\lambda$, in period k after shock arrival $\forall \lambda \in [0,1]$ and $k \in \mathbb{N}_+$, $k \leq K$:

$$n(\lambda,k) = (1-\lambda)^{k-1} \cdot (1 - G(R(\lambda,1)) \cdot p(\lambda) \cdot n^s$$

3. The number of firms with shock duration $1/\lambda$, in period K or more periods after shock arrival $\forall \lambda \in [0,1]$:

$$n(\lambda,K) = \frac{(1-\lambda)^{K-1}}{\lambda} \cdot (1 - G(R(\lambda,1)) \cdot p(\lambda) \cdot n^s$$

4. Total employment

$$n = \int_0^1 \sum_{k=1}^{K-1} n(\lambda,k) + n(\lambda,K) d\lambda$$

5. Job-creation condition:

$$\begin{aligned}
\frac{1}{1-\eta} \cdot \frac{k_v}{q(\theta)} = & \int_0^1 \sum_{k=1}^{K-1} (1-\lambda)^{k-1} \cdot \Bigg[\int_{R(\lambda,1)}^{\infty} z(\epsilon) dG_\lambda(\epsilon) - \Omega(\lambda,k) - \frac{b}{n} \\
& + \frac{\lambda - \eta \cdot \theta \cdot q(\theta)}{1-\eta} \cdot \frac{k_v}{q(\theta)} \Bigg] dP(\lambda) \\
& + \int_0^1 \frac{(1-\lambda)^{K-1}}{\lambda} \cdot \Bigg[\int_{R(\lambda,1)}^{\infty} z(\epsilon) dG_\lambda(\epsilon) - \Omega(\lambda,K) - \frac{b}{n} \\
& + \frac{\lambda - \eta \cdot \theta \cdot q(\theta)}{1-\eta} \cdot \frac{k_v}{q(\theta)} \Bigg] dP(\lambda)
\end{aligned}$$

6. Separation Conditions for firms without access to a subsidy $\forall \lambda \in [0,1]$:

$$z(\xi(\lambda)) - b + \frac{\lambda - \eta \cdot f}{1-\eta} \cdot \frac{k_v}{q(\theta)} = 0$$

7. Separation condition with access to a subsidy $\forall \lambda \in [0,1]$

$$\begin{aligned}
& \sum_{k=1}^{K-1} \lambda \cdot (1-\lambda)^{k-1} \cdot \left[z_{stw,k}(R(\lambda,1)) + \tau_{stw}(k) \cdot \left(\bar{h} - h_{stw,k}(R(\lambda,1)) \right) \right] \\
& + (1-\lambda)^{K-1} \cdot \left[z_{stw,K}(R(\lambda,1)) + \tau_{stw}(K) \cdot \left(\bar{h} - h_{stw,K}(R(\lambda,1)) \right) \right] \\
& - b + \frac{\lambda - \eta \cdot f}{1-\eta} \cdot \frac{k_v}{q(\theta)} = 0
\end{aligned}$$

2.C.3.1 Derive Lagrange multiplier for separation condition

Compared to our initial problem, the FOCs for the separation thresholds have changed. The other FOCs stay conceptually the same and are not replicated here. In contrast to a weakly decreasing subsidy system, separation thresholds $R(\lambda, k)$ for $k = 2, 3, ..., K$ in a weakly increasing system do not affect employment. Thus, their FOC is

$$\frac{\partial}{\partial R(\lambda, k)} = -\mu_R(\lambda, k) \cdot \left(A \cdot R(\lambda, 1) \cdot h_{stw,k}(R(\lambda, k))^{\alpha} + c_f\right) = 0$$

which implies that their Lagrange multiplier must be equal to zero: $\mu_R(\lambda, k) = 0$. Firms and workers decide to separate in the first period only. Thus, the FOC for $R(\lambda, 1)$ is:

$$\begin{aligned}
\frac{\partial}{\partial R(\lambda, 1)} = & -\sum_{k=1}^{K-1}(1-\lambda)^{k-1} \cdot g_{\lambda}(R(\lambda, k)) \cdot p(\lambda) \cdot n^s \cdot \left(z_{stw,k}(R(\lambda, k)) + \mu_n^s(\lambda, k)\right) \\
& - \mu_{\theta} \cdot \sum_{k=1}^{K-1}(1-\lambda)^{k-1} \cdot g_{\lambda}(R(\lambda, k)) \cdot p(\lambda) \\
& \cdot \left(z_{stw,k}(R(\lambda, k)) - \frac{b}{n} + \frac{\lambda - \eta \cdot \theta \cdot q(\theta)}{1-\eta} \cdot \frac{k_v}{q(\theta)}\right) \\
& - \frac{(1-\lambda)^{K-1}}{\lambda} \cdot g_{\lambda}(R(\lambda, K)) \cdot p(\lambda) \cdot n^s \cdot \left(z_{stw,K}(R(\lambda, K)) + \lambda_n^s(\lambda, K)\right) \\
& - \mu_{\theta} \cdot \frac{(1-\lambda)^{K-1}}{\lambda} \cdot g_{\lambda}(R(\lambda, K)) \cdot p(\lambda) \\
& \cdot \left(z_{stw,K}(R(\lambda, K)) - \frac{b}{n} + \frac{\lambda - \eta \cdot \theta \cdot q(\theta)}{1-\eta} \cdot \frac{k_v}{q(\theta)}\right) \\
& - \mu_R(\lambda, 1) \cdot \left(A \cdot h_{stw,1}(R(\lambda, 1))^{\alpha} + c_f\right) = 0
\end{aligned}$$

We can rewrite this to:

$$\begin{aligned}
& E_K\left[g_{\epsilon}(R(\lambda, 1)) \cdot \frac{p(\lambda)}{\lambda} \cdot n^s \cdot \left(z_{stw,k}(R(\lambda, 1)) + \mu_n^s(\lambda, k)\right)\right] \\
& + \mu_{\theta} \cdot E_K\left[g_{\epsilon}(R(\lambda, 1)) \cdot \frac{p(\lambda)}{\lambda} \cdot \left(z_{stw,k}(R(\lambda, 1)) - \frac{b}{n} + \frac{\lambda - \eta \cdot \theta \cdot q(\theta)}{1-\eta} \cdot \frac{k_v}{q(\theta)}\right)\right] \\
& + \mu_R(\lambda, 1) \cdot \left(A \cdot h_{stw,1}(R(\lambda, 1))^{\alpha} + c_f\right) = 0
\end{aligned}$$

Insert $\mu_n^s(\lambda, k)$:

$$E_K\Bigg[g_\epsilon(R(\lambda,1))\cdot\frac{p(\lambda)}{\lambda}\cdot n^s\cdot\Big(z_{stw,1}(R(\lambda,1))+\frac{\mu_\theta}{n^s}\cdot\frac{b}{n}+k_v\cdot\theta + (\lambda-\theta\cdot q(\theta))\cdot\mu_n^s\Big)\Bigg]$$
$$+\mu_\theta\cdot E_K\Bigg[g_\epsilon(R(\lambda,1))\cdot\frac{p(\lambda)}{\lambda}\cdot\Big(z_{stw,1}(R(\lambda,1))-\frac{b}{n}+\frac{\lambda-\eta\cdot\theta\cdot q(\theta)}{1-\eta}\cdot\frac{k_v}{q(\theta)}\Big)\Bigg]$$
$$+\mu_R(\lambda,1)\cdot\Big(A\cdot h_{stw,1}(R(\lambda,1))^\alpha+c_f\Big) = 0$$

Insert μ_n^s:

$$E_K\Bigg[g_\epsilon(R(\lambda,1))\cdot\frac{p(\lambda)}{\lambda}\cdot n^s\cdot\Bigg(z_{stw,1}(R(\lambda,1))+\frac{\mu_\theta}{n^s}\cdot\frac{b}{n} + \frac{\lambda\cdot\gamma-\theta\cdot q(\theta)+(\lambda-\theta\cdot q(\theta))\cdot\chi}{1-\gamma}\cdot\frac{k_v}{q(\theta)}\Bigg)\Bigg]$$
$$+\mu_\theta\cdot E_K\Bigg[g_\epsilon(R(\lambda,1))\cdot\frac{p(\lambda)}{\lambda}\cdot\Bigg(z_{stw,1}(R(\lambda,1))-\frac{b}{n}+\frac{\lambda-\eta\cdot\theta\cdot q(\theta)}{1-\eta}\cdot\frac{k_v}{q(\theta)}\Bigg)\Bigg]$$
$$+\mu_R(\lambda,1)\cdot\Big(A\cdot h_{stw,1}(R(\lambda,1))^\alpha+c_f\Big) = 0$$

Subtract the separation condition:

$$E_K\Bigg[g_\epsilon(R(\lambda,1))\cdot\frac{p(\lambda)}{\lambda}\cdot n^s\cdot\Big(1+\frac{\mu_\theta}{n^s}\Big)\cdot\Big(\frac{\lambda}{f}\cdot b-E_K\big[\tau_{stw}(k)\cdot\big(\bar{h}-h_{stw,k}(R(\lambda,1))\big)\big]\Big)\Bigg]$$
$$+\mu_R(\lambda,1)\cdot\big(A\cdot h_{stw,1}(R(\lambda,1))^\alpha+c_f\big) = 0$$

Rearrange for multiplier:

$$-\mu_R(\lambda,1) = \frac{\big(1+\frac{\mu_\theta}{n^s}\big)}{\big(A\cdot h_{stw,1}(R(\lambda,k))^\alpha+c_f\big)} \cdot g_\epsilon(R(\lambda,1))\cdot\frac{p(\lambda)}{\lambda}\cdot n^s\cdot\Big(\frac{\lambda}{f}\cdot b-E_K\big[\tau_{stw}(k)\cdot\big(\bar{h}-h_{stw,k}(R(\lambda,1))\big)\big]\Big)$$

2.C.3.2 Derive Lagrange multiplier for separation condition

Relative to the Ramsey problem with a weakly decreasing benefit schedule, the first order condition for STW benefits is altered under an increasing schedule. When the planner adjusts benefits for firms that have remained on STW for k periods, this policy change propagates back to period-1 decisions by affecting the expected present value of subsidies, thereby shifting the initial separation threshold. The FOC for the STW benefits becomes:

$$\frac{\partial}{\partial \tau_{stw}(k)} = -\int_0^1 \lambda \cdot (1-\lambda)^{k-1} \cdot \mu(\lambda,1) \cdot \left(\bar{h} - h_{stw,k}(R(\lambda,1))\right) \cdot d\lambda$$
$$-\int_0^1 (1-\lambda)^{k-1} \cdot p(\lambda) \cdot n^s \cdot \left(1 + \frac{\mu_\theta}{n^s}\right) \cdot \frac{\partial \Omega(\lambda,k)}{\partial \tau_{stw}(k)} \cdot d\lambda = 0$$

Further, let LS(λ) define the social value of the marginal match:

$$LS(\lambda) = \frac{\lambda}{f} \cdot b - E_K\left[\tau_{stw}(k) \cdot \left(\bar{h} - h_{stw,k}(R(\lambda,1))\right)\right]$$

Inserting $\mu(R(\lambda,1))$ gives:

$$\int_0^1 (1-\lambda)^{k-1} \cdot g_\epsilon(R(\lambda,1)) \cdot p(\lambda) \cdot \left(-\frac{\partial R(\lambda,1)}{\partial \tau_{stw}(k)}\right) \cdot n^s \cdot \left(1 + \frac{\mu_\theta}{n^s}\right) \cdot LS(\lambda) \cdot d\lambda =$$
$$= \int_0^1 (1-\lambda)^{k-1} \cdot p(\lambda) \cdot n^s \cdot \left(1 + \frac{\mu_\theta}{n^s}\right) \cdot \frac{\partial \Omega(\lambda,k)}{\partial \tau_{stw}(k)} \cdot d\lambda$$

Using the definition for employment, we can rewrite the condition as:

$$\int_0^1 \frac{\partial n(\lambda,k)}{\partial R(\lambda,1)} \cdot LS(\lambda) d\lambda = \int_0^1 \frac{\partial n(\lambda,k)}{\partial R(\lambda,1)} \cdot \left(n_{stw}(\lambda,k) / \frac{\partial n(\lambda,k)}{\partial \tau_{stw}(\lambda,k)}\right) \cdot \frac{\partial \tilde{\Omega}(\lambda,k)}{\partial \tau_{stw}(k)} \cdot d\lambda$$

with

$$\frac{\partial n(\lambda,k)}{\partial R(\lambda,1)} = (1-\lambda)^{k-1} \cdot g_\epsilon(R(\lambda,1)) \cdot p(\lambda) \cdot \left(-\frac{\partial R(\lambda,1)}{\partial \tau_{stw}(k)}\right) \cdot n^s$$

This lets us rewrite the expression as:

$$\int_0^1 \omega_k(\lambda) \cdot LS(\lambda) d\lambda = \int_0^1 \omega_k(\lambda) \cdot \left(n_{stw}(\lambda,k) / \frac{\partial n(\lambda,k)}{\partial \tau_{stw}(\lambda,k)}\right) \cdot \frac{\partial \tilde{\Omega}(\lambda,k)}{\partial \tau_{stw}(k)} \cdot d\lambda$$

with

$$\omega_k(\lambda) = \frac{\frac{\partial n(\lambda,k)}{\partial R(\lambda,1)}}{\int_0^1 \frac{\partial n(\lambda,k)}{\partial R(\lambda,1)} \cdot d\lambda}$$

This gives back Proposition 2

$$E_{\omega_k}\left[\frac{n_{stw}(\lambda)}{\frac{\partial n(\lambda)}{\partial \tau_{stw}}} \frac{\partial \tilde{\Omega}(\lambda,k)}{\partial \tau_{stw}(k)}\right] = E_{\omega_k}\Big[LS(\lambda)\Big]$$

Note that we can write the derivative for the welfare costs of STW as:

$$\begin{aligned}\frac{\partial\tilde{\Omega}(\lambda,k)}{\partial\tau_{stw}(k)} &= -\int_{R(\lambda,1)}^{R_{stw}(\lambda)} \frac{\partial h_{stw,k}(\epsilon)}{\partial\tau_{stw}(k)} \cdot \tau_{stw}(k) \cdot \frac{dG_{\lambda}(\epsilon)}{\rho_{stw}(\lambda,k)} \\ &= \tau_{stw}(k) \cdot E_{\epsilon}\left[-\frac{\partial h_{stw,k}(\epsilon)}{\partial\tau_{stw}(k)}\middle| R_{stw}(\lambda) \geq \epsilon \geq R(\lambda,1)\right]\end{aligned}$$

Thus:

$$\begin{aligned}&E_{\omega_k}\left[\frac{n_{stw}(\lambda,k)}{\partial n(\lambda,k)/\partial\tau_{stw}(k)} \cdot \tau_{stw}(k) \cdot E_{\epsilon}\left[-\frac{\partial h_{stw,k}(\epsilon)}{\partial\tau_{stw}(k)}\middle| R_{stw}(\lambda) \geq \epsilon \geq R(\lambda,1)\right]\right] \\ &= E_{\omega_k}\Big[E_K\left[LS_k(\lambda)\right]\Big]\end{aligned}$$

Note that the weighted average value of a marginal match pins down the STW benefits in period k. The idea is that only the expected benefit payments from STW influence the separation condition directly. Therefore, the government seeks to set STW benefits in order to minimize the distortions of the STW system:

$$\tau_{stw}(k) = \frac{E_{\omega_k}\Big[LS(\lambda)\Big]}{E_{\omega_k}\left[\frac{n_{stw}(\lambda)}{\frac{\partial n(\lambda)}{\partial\tau_{stw}}} \cdot E_{\epsilon}\left[-\frac{\partial h_{stw,k}(\epsilon)}{\partial\tau_{stw}(k)}\middle| R_{stw}(\lambda) \geq \epsilon \geq R(\lambda,1)\right]\right]}$$

2.C.4 Ramsey Problem — Only Working Hours Observable

Assume that $\beta \to 1$. Then the problem of the Ramsey planner becomes:

$$n^s \cdot \int_0^1 \int_{A(\lambda)} \Big[z(\epsilon)dG_{\lambda}(\epsilon) - \Omega(\lambda,k)\Big]dP(\lambda) - (1-n) \cdot \theta \cdot k_v$$

subject to

1. The number of workers that draw a shock this period:

$$n^s = \theta \cdot q(\theta) \cdot (1-n) + \int_0^1 \lambda \cdot n(\lambda,k)d\lambda$$

2. The number of firms with shock duration $1/\lambda$, in period K or more periods after shock arrival $\forall\lambda \in [0,1]$:

$$n(\lambda) = \frac{1}{\lambda} \cdot (1-\rho(R(\lambda)) \cdot p(\lambda) \cdot n^s$$

3. Separation rate $\forall\lambda \in [0,1]$:

$$\rho(\lambda) \;=\; G(R(\lambda)) + G\big(\max\{\xi(\lambda), R_{stw}(\lambda)\}\big) - G\big(\max\{R_{stw}(\lambda), R(\lambda)\}\big)$$

4. Total employment

$$n \;=\; \int_0^1 n(\lambda) d\lambda$$

5. Job-creation condition:

$$\frac{1}{1-\eta}\cdot\frac{k_v}{q(\theta)} \;=\; \int_0^1 \left(\int_{A(\lambda)} z(\epsilon) dG_\lambda(\epsilon) - \Omega(\lambda,k) - \frac{b}{n} + \frac{\lambda-\eta\cdot\theta\cdot q(\theta)}{1-\eta}\cdot\frac{k_v}{q(\theta)}\right) dP(\lambda)$$

6. Set $\forall\lambda \in [0,1]$:

$$A(\lambda) \;=[\max\{R(\lambda), \max\{R_{stw}, R(\lambda)\}\} \;\cup\; [\max\{R_{stw}, \xi(\lambda)\}, \infty)$$

7. Separation Conditions for firms without access to a subsidy $\forall\lambda \in [0,1]$:

$$z(\xi(\lambda)) + \frac{\lambda-\eta\cdot f}{1-\eta}\cdot\frac{k_v}{q(\theta)} \;=\; 0$$

8. Separation condition with access to a subsidy $\forall\lambda \in [0,1]$:

$$z_{stw}(R(\lambda)) + \tau_{stw}\cdot\big(\bar{h} - h_{stw}(R(\lambda))\big) + \frac{\lambda-\eta\cdot f}{1-\eta}\cdot\frac{k_v}{q(\theta)} \;=\; 0$$

2.C.4.1 Screening Effect of the STW Threshold

For the following, assume for simplicity that $P(\lambda)$ is a steady distribution. Regarding the eligibility condition, we need to consider 3 cases:

$$\begin{aligned}
&\text{case 1: } R_{stw} \;<\; R(\lambda)\\
&\text{case 2: } R_{stw} \;\in\; [\xi(\lambda), R(\lambda)]\\
&\text{case 3: } R_{stw} \;>\; \xi(\lambda)
\end{aligned}$$

In case 1, the eligibility condition is so strict that firms with shock duration $1/\lambda$ would fire a worker before they are allowed to enter the STW program. In case 2, some firms with shock duration $1/\lambda$ have such a high productivity that they are excluded from STW, but not high enough to survive on their own. Others with lower productivity can enter STW and are rescued by it. Finally, in case 3, every firm with shock duration $1/\lambda$ that can be saved by STW can enter STW. At the same time, firms with shock duration $1/\lambda$ that do not need STW support also enter the system.

Let us define

$$\lambda^R = R^{-1}(R_{stw}) \quad \text{and} \quad \lambda^\xi = \xi^{-1}(R_{stw})$$

Then we can order the cases according to the length of the shock $1/\lambda$:

$$\begin{aligned}
&\text{case 1: } \lambda < \lambda^R \\
&\text{case 2: } \lambda \in [R(\lambda), \xi(\lambda)] \\
&\text{case 3: } \lambda > \lambda^\xi
\end{aligned}$$

If the probability of becoming productive again is very low, then firms will separate for high productivity values. As a result, some of the firms will not be eligible for STW (case 1). Vice versa, if the probability of becoming productive again is high, then they would be willing to keep the worker employed even for large drops in productivity. As a result, a lot of those firms might enter STW even though they would not need the subsidy to survive (case 3).

We can use this insight to reformulate the Ramsey problem:

2.C.4.2 Ramsey Problem - simplified

Assumption: $\beta \to 1$. Then the problem of the Ramsey planner becomes:

$$\begin{aligned}
\max_{\tau_{stw}, D} \quad & n^s \cdot \int_0^{\lambda^R} \int_{\xi(\lambda)}^{\infty} z(\epsilon) dG_\lambda(\epsilon) dP(\lambda) \\
& + n^s \cdot \int_{\lambda^R}^{\lambda^\xi} \left(\int_{\xi(\lambda)}^{\infty} z(\epsilon) dG_\lambda(\epsilon) + \int_{R(\lambda)}^{R_{stw}} z_{stw}(\epsilon) dG_\lambda(\epsilon) \right) dP(\lambda) \\
& + n^s \cdot \int_{\lambda^\xi}^{1} \left(\int_{R_{stw}}^{\infty} z(\epsilon) dG_\lambda(\epsilon) + \int_{R(\lambda)}^{R_{stw}} z(\epsilon) dG_\lambda(\epsilon) - \Omega(\lambda) \right) dP(\lambda) \\
& - (1-n) \cdot \theta \cdot k_v
\end{aligned}$$

subject to

1. The number of workers that draw a shock this period:

$$n^s = \theta \cdot q(\theta) \cdot (1-n) + \int_0^1 \lambda \cdot n(\lambda, k) d\lambda$$

2. The number of firms with shock duration $1/\lambda$, in period K or more periods after shock arrival $\forall \lambda \in [0,1]$:

$$n(\lambda) = \begin{cases} \frac{1}{\lambda} \cdot (1 - G_\lambda(\xi(\lambda))) \cdot p(\lambda) \cdot n^s & \lambda < \lambda^R \\ \frac{1}{\lambda} \cdot (1 - G_\lambda(\xi(\lambda)) + G_\lambda(R_{stw}) - G_\lambda(R(\lambda))) \cdot p(\lambda) \cdot n^s & \lambda \in \left[\lambda^R, \lambda^\xi\right] \\ \frac{1}{\lambda} \cdot (1 - G_\lambda(R(\lambda))) \cdot p(\lambda) \cdot n^s & \lambda > \lambda^\xi \end{cases}$$

3. Total employment

$$n = \int_0^1 n(\lambda) d\lambda$$

4. Job-creation condition:

$$\begin{aligned} \frac{1}{1-\eta} \cdot \frac{k_v}{q(\theta)} = & \int_0^{\lambda^R} \left(\int_{\xi(\lambda)}^{\infty} z(\epsilon)\, dG_\lambda(\epsilon) - \frac{b}{n} + \frac{\lambda - \eta \cdot \theta \cdot q(\theta)}{1-\eta} \cdot \frac{k_v}{q(\theta)} \right) dP(\lambda)\, d\lambda \\ & + n^s \cdot \int_{\lambda^R}^{\lambda^\xi} \left(\int_{\xi(\lambda)}^{\infty} z(\epsilon)\, dG_\lambda(\epsilon) \right. \\ & \qquad \left. + \int_{R(\lambda)}^{R_{stw}} z_{stw}(\epsilon)\, dG_\lambda(\epsilon) - \frac{b}{n} + \frac{\lambda - \eta \cdot \theta \cdot q(\theta)}{1-\eta} \cdot \frac{k_v}{q(\theta)} \right) dP(\lambda)\, d\lambda \\ & + n^s \cdot \int_{\lambda^\xi}^{1} \left(\int_{R_{stw}}^{\infty} z(\epsilon)\, dG_\lambda(\epsilon) - \Omega(\lambda) - \frac{b}{n} \right. \\ & \qquad \left. + \frac{\lambda - \eta \cdot \theta \cdot q(\theta)}{1-\eta} \cdot \frac{k_v}{q(\theta)} \right) dP(\lambda)\, d\lambda \end{aligned}$$

5. Eligibility thresholds on shock persistence:

$$R(\lambda^R) = R_{stw}, \quad \xi(\lambda^\xi) = R_{stw}$$

6. Separation Conditions for firms without access to a subsidy $\forall \lambda \in [0,1]$:

$$z(\xi(\lambda)) + \frac{\lambda - \eta \cdot f}{1-\eta} \cdot \frac{k_v}{q(\theta)} = 0$$

7. Separation condition with access to a subsidy $\forall \lambda \in [0,1]$:

$$z_{stw}(R(\lambda)) + \tau_{stw} \cdot \left(\bar{h} - h_{stw}(R(\lambda))\right) + \frac{\lambda - \eta \cdot f}{1-\eta} \cdot \frac{k_v}{q(\theta)} = 0$$

2.C.4.3 Optimal Eligibility Condition

FOC eligibility condition:

$$\begin{aligned}
\frac{\partial}{\partial \epsilon_{stw}} &= \int_0^{\lambda^R} 0 \cdot dP(\lambda) \\
&+ n^s \cdot \int_{\lambda^R}^{\lambda^\xi} \frac{p(\lambda)}{\lambda} \cdot g_\lambda(R_{stw}) \cdot \left[z_{stw}(R_{stw}) + \mu_n^s(\lambda)\right] d\lambda \\
&- n^s \cdot \int_{\lambda^\xi}^{1} \frac{p(\lambda)}{\lambda} \cdot \frac{\partial \Omega(\lambda)}{\partial \tau_{stw}} \cdot d\lambda \\
&+ \mu_\theta \cdot \int_0^{\lambda^R} 0 \cdot dP(\lambda) \\
&+ \mu_\theta \cdot \int_{\lambda^R}^{\lambda^\xi} \frac{p(\lambda)}{\lambda} \cdot g_\lambda(R_{stw}) \cdot \left[z_{stw}(R_{stw}) - \frac{b}{n} + \frac{\lambda - \theta \cdot q(\theta) \cdot \eta}{1-\eta} \cdot \right] d\lambda \\
&- \mu_\theta \cdot \int_{\lambda^\xi}^{1} \frac{p(\lambda)}{\lambda} \cdot \frac{\partial \Omega(\lambda)}{\partial \epsilon_{stw}} \cdot d\lambda \;=\; 0
\end{aligned}$$

The expression simplifies to:

$$\begin{aligned}
& n^s \cdot \int_{\lambda^R}^{\lambda^\xi} \frac{p(\lambda)}{\lambda} \cdot g_\lambda(R_{stw}) \cdot \left[z_{stw}(R_{stw}) + \mu_n^s(\lambda)\right] d\lambda \\
&+ \mu_\theta \cdot \int_{\lambda^R}^{\lambda^\xi} \frac{p(\lambda)}{\lambda} \cdot g_\lambda(R_{stw}) \cdot \left[z_{stw}(R_{stw}) - \frac{b}{n} + \frac{\lambda - \theta \cdot q(\theta) \cdot \eta}{1-\eta} \cdot \right] d\lambda \\
&- n^s \cdot \left(1 + \frac{\mu_\theta}{n^s}\right) \cdot \int_{\lambda^\xi}^{1} \frac{p(\lambda)}{\lambda} \cdot \frac{\partial \Omega(\lambda)}{\partial \epsilon_{stw}} \cdot d\lambda \;=\; 0
\end{aligned}$$

Inserting $\mu_n^s(\lambda)$ gives:

$$\begin{aligned}
& n^s \cdot \int_{\lambda^R}^{\lambda^\xi} \frac{p(\lambda)}{\lambda} \cdot g_\lambda(R_{stw}) \cdot \left[z_{stw}(R_{stw}) + \frac{\mu_\theta}{n^s} \cdot \frac{b}{n} + k_v \cdot \theta + (\lambda - \theta \cdot q(\theta)) \cdot \mu_n^s\right] d\lambda \\
&+ \mu_\theta \cdot \int_{\lambda^R}^{\lambda^\xi} \frac{p(\lambda)}{\lambda} \cdot g_\lambda(R_{stw}) \cdot \left[z_{stw}(R_{stw}) - \frac{b}{n} + \frac{\lambda - \theta \cdot q(\theta) \cdot \eta}{1-\eta} \cdot \right] d\lambda \\
&- n^s \cdot \left(1 + \frac{\mu_\theta}{n^s}\right) \cdot \int_{\lambda^\xi}^{1} \frac{p(\lambda)}{\lambda} \cdot \frac{\partial \Omega(\lambda)}{\partial \epsilon_{stw}} \cdot d\lambda \;=\; 0
\end{aligned}$$

Inserting μ_n^s gives:

$$
\begin{aligned}
& n^s \cdot \int_{\lambda^R}^{\lambda^\xi} \frac{p(\lambda)}{\lambda} \cdot g_\lambda(R_{stw}) \\
& \quad \cdot \left[z_{stw}(R_{stw}) + \frac{\mu_\theta}{n^s} \cdot \frac{b}{n} + \frac{\lambda \cdot \gamma - \theta \cdot q(\theta)}{1-\gamma} \cdot \frac{k_v}{q(\theta)} + \frac{(\lambda - \theta \cdot q(\theta)) \cdot \chi}{1-\gamma} \cdot \frac{k_v}{q(\theta)} \right] d\lambda \\
& + \mu_\theta \cdot \int_{\lambda^R}^{\lambda^\xi} \frac{p(\lambda)}{\lambda} \cdot g_\lambda(R_{stw}) \cdot \left[z_{stw}(R_{stw}) - \frac{b}{n} + \frac{\lambda - \theta \cdot q(\theta) \cdot \eta}{1-\eta} \right] d\lambda \\
& - n^s \cdot \left(1 + \frac{\mu_\theta}{n^s}\right) \cdot \int_{\lambda^\xi}^{1} \frac{p(\lambda)}{\lambda} \cdot \frac{\partial \Omega(\lambda)}{\partial \epsilon_{stw}} d\lambda \; = \; 0
\end{aligned}
$$

Subtracting separation conditions gives:

$$
\begin{aligned}
& n^s \cdot \left(1 + \frac{\mu_\theta^s}{n^s}\right) \cdot \int_{\lambda^R}^{\lambda^\xi} \frac{p(\lambda)}{\lambda} \cdot g_\lambda(R_{stw}) \\
& \quad \cdot \left[z_{stw}(R_{stw}) - z_{stw}(R(\lambda)) + b - \tau_{stw} \cdot (\bar{h} - h_{stw}(R(\lambda))) \right] d\lambda \\
& + n^s \cdot \int_{\lambda^R}^{\lambda^\xi} \frac{p(\lambda)}{\lambda} \cdot g_\lambda(R_{stw}) \cdot (\lambda - \theta \cdot q(\theta)) \cdot \left(\chi - \frac{\eta - \gamma}{1-\eta}\right) \cdot \frac{k_v}{q(\theta)} \cdot d\lambda \\
& - n^s \cdot \left(1 + \frac{\mu_\theta}{n^s}\right) \cdot \int_{\lambda^\xi}^{1} \frac{p(\lambda)}{\lambda} \cdot \frac{\partial \Omega(\lambda)}{\partial \epsilon_{stw}} \cdot d\lambda \; = \; 0
\end{aligned}
$$

Inserting $\left(\chi - \frac{\eta-\gamma}{1-\eta}\right) \cdot \frac{k_v}{q(\theta)}$ gives:

$$
\begin{aligned}
& n^s \cdot \left(1 + \frac{\mu_\theta^s}{n^s}\right) \cdot \int_{\lambda^R}^{\lambda^\xi} \frac{p(\lambda)}{\lambda} \cdot g_\lambda(R_{stw}) \\
& \quad \cdot \left[z_{stw}(R_{stw}) - z_{stw}(R(\lambda)) + b - \tau_{stw} \cdot (\bar{h} - h_{stw}(R(\lambda))) \right] d\lambda \\
& + n^s \cdot \left(1 + \frac{\mu_\theta}{n^s}\right) \cdot \int_{\lambda^R}^{\lambda^\xi} \frac{p(\lambda)}{\lambda} \cdot g_\lambda(R_{stw}) \cdot (\lambda - \theta \cdot q(\theta)) \cdot \frac{b}{f} \cdot d\lambda \\
& - n^s \cdot \left(1 + \frac{\mu_\theta}{n^s}\right) \cdot \int_{\lambda^\xi}^{1} \frac{p(\lambda)}{\lambda} \cdot \frac{\partial \Omega(\lambda)}{\partial \epsilon_{stw}} \cdot d\lambda \; = \; 0
\end{aligned}
$$

This lets us simplify the expression to:

$$
\begin{aligned}
& \int_{\lambda^R}^{\lambda^\xi} \frac{p(\lambda)}{\lambda} \cdot g_\lambda(R_{stw}) \cdot \left[z_{stw}(R_{stw}) - z_{stw}(R(\lambda)) + \frac{\lambda}{f} \cdot b - \tau_{stw} \cdot (\bar{h} - h_{stw}(R(\lambda))) \right] d\lambda \\
& - \int_{\lambda^\xi}^{1} \frac{p(\lambda)}{\lambda} \cdot \frac{\partial \Omega(\lambda)}{\partial \epsilon_{stw}} \cdot d\lambda \; = \; 0
\end{aligned}
$$

Define $SW(\epsilon, \lambda)$ as the social value of a match with productivity ϵ and shock duration $1/\lambda$:

$$SW(\epsilon,\lambda) \;=\; z_{stw}(\epsilon) - z_{stw}(R(\lambda)) + \frac{\lambda}{f}\cdot b - \tau_{stw}\cdot(\bar{h} - h_{stw}(\epsilon))$$

Define weights as:

$$\omega(\lambda) \;=\; \frac{\frac{\partial n(\lambda)}{\partial \epsilon_{stw}}}{\int_0^1 \frac{\partial n(\lambda)}{\partial \epsilon_{stw}} d\lambda}, \quad \text{with} \quad \frac{\partial n(\lambda)}{\partial \epsilon_{stw}} \;=\; \frac{p(\lambda)}{\lambda}\cdot g_{\lambda}(\epsilon_{stw})\cdot n^s$$

Further define

$$p_{stw} \;=\; P(\lambda^{\xi}) - P(\lambda^{R}), \quad p_f \;=\; 1 - P(\lambda^{\xi})$$

as the probability of becoming a firm that entered STW in need of support of STW, respectively, a firm that entered STW without the need of STW support. This gives:

$$p_{stw}\cdot E_{\omega}\left[SW(\epsilon,\lambda)\middle|\lambda \in \left[\lambda^R,\lambda^{\xi}\right]\right] \;=\; p_f \cdot E_{\omega}\left[\frac{n(\lambda)}{\frac{\partial n(\lambda)}{\partial \epsilon_{stw}}}\cdot\frac{\partial\tilde{\Omega}(\lambda)}{\partial \tau_{stw}}\middle|\lambda \in [\lambda^{\xi},1]\right]$$

2.C.4.4 Optimal STW benefits:

Following the same steps as in the other Ramsey problems, we can derive the Lagrange multiplier for the separation condition under STW, dependent on the probability of recovery:

$$-\mu_R(\lambda) \;=\; \begin{cases} 0, & \lambda \;<\; \lambda^R \\ \frac{p(\lambda)}{\lambda}\cdot g_{\lambda}(R(\lambda))\cdot n^s\cdot\frac{\frac{\lambda}{f}\cdot b - \tau_{stw}\cdot\left(\bar{h} - h_{stw}(R(\lambda))\right)}{A\cdot h_{stw}(R(\lambda))^{\alpha} + c_f}, & \lambda \;\geq\; \lambda^R \end{cases}$$

Note that the Lagrange multiplier is zero if the recovery probability is low enough. The idea is that firms with $\lambda < \lambda^R$ are screened out of the STW system. The first order condition for the STW benefits is:

$$\frac{\partial}{\partial \tau_{stw}} \;=\; -\int_0^1 \mu_R(\lambda)\cdot\left(\bar{h} - h_{stw}(R(\lambda))\right)d\lambda - n^s\cdot\int_{\lambda^{\xi}}^1 \frac{p(\lambda)}{\lambda}\cdot\frac{\partial\Omega(\lambda)}{\partial \tau_{stw}}\cdot d\lambda \;=\; 0$$

Inserting the Lagrange multiplier for the separation condition on STW gives:

$$\int_{\lambda^R}^1 \frac{p(\lambda)}{\lambda}\cdot g_{\lambda}(R(\lambda))\cdot\left(-\frac{\partial R(\lambda)}{\partial \tau_{stw}}\right)\cdot n^s\cdot\left(\frac{\lambda}{f}\cdot b - \tau_{stw}\cdot\left(\bar{h} - h_{stw}(R(\lambda))\right)\right)d\lambda$$
$$-\int_{\lambda^R}^1 \frac{p(\lambda)}{\lambda}\cdot\rho_{stw}(\lambda)\cdot\frac{\partial\tilde{\Omega}(\lambda)}{\partial \tau_{stw}}\cdot d\lambda \;=\; 0$$

Define:

$$\omega(\lambda) = \frac{\frac{\partial n(\lambda)}{\partial \tau_{stw}}}{\int_0^1 \frac{\partial n(\lambda)}{\partial \tau_{stw}} d\lambda}, \quad \text{with} \quad \frac{\partial n(\lambda)}{\partial \tau_{stw}} = \frac{p(\lambda)}{\lambda} \cdot g_\lambda(R(\lambda)) \cdot \left(-\frac{\partial R(\lambda)}{\partial \tau_{stw}}\right) \cdot n^s$$

Then we can write:

$$(1 - G_\lambda(\lambda^R)) \cdot E_\omega \left[\frac{\lambda}{f} \cdot b - \tau_{stw} \cdot \left(\bar{h} - h_{stw}(R(\lambda))\right) - \frac{n(\lambda)}{\frac{\partial n(\lambda)}{\partial \tau_{stw}}} \cdot \frac{\partial \Omega(\lambda)}{\partial \tau_{stw}} \middle| \lambda \geq \lambda^R \right]$$
$$= 0$$

Inserting Lemma 3 gives:

$$E_\omega\left[\tau_{stw} \cdot \left(\bar{h} - h_{stw}(R(\lambda))\right) \middle| \lambda \geq \lambda^R\right] = E_\omega\left[\tau_{stw}(\lambda) \cdot \left(\bar{h} - h_{stw}(R(\lambda))\right) \middle| \lambda \geq \lambda^R\right]$$

Chapter 3

Carrots or Sticks? Short-Time Work or Lay-off Taxes?

Joint with Johannes Weber

3.1 Introduction

It has long been understood that unemployment insurance systems can inefficiently increase separations in the labor market. To counter this inefficiency, governments commonly have two main policy instruments at their disposal. On the one hand, there are lay-off taxes that, in effect, punish firms for firing workers. On the other, there is the short-time work system that rewards the retainment of endangered workers through subsidizing and enabling hours reductions[1].

This raises the natural question, which of the two governments should employ. Existing literature highlights that lay-off taxes have many desirable properties. (Cahuc and Zylberberg, 2008) and Blanchard and Tirole, 2008 show in implicit contract frameworks that lay-off taxes can implement the planner solution. Short-time work on the other hand, has the problem of introducing new inefficiencies into the economy through the distortion of working hours (Stiepelmann, 2024). However, these results crucially hinge on the absence of financial constraints for firms. This assumption is widespread in the search and matching literature, but clearly is at odds with reality.

In this study, we relax this assumption in a rich yet analytically tractable DMP framework and allow for a share of firms to be financially constrained. Using the

1. Short-time work systems are pervasive, e.g., in European countries where they were utilized during the Great Recession and the Covid period. Lay-off taxes are implemented, e.g., in the U.S. through an experience-rated unemployment insurance system.

Ramsey policy approach, we then determine the welfare consequences of optimally set short-time work benefits and optimally set lay-off taxes.

Our main result is that lay-off taxes are indeed superior when the share of financially constrained firms is small. Short-time work emerges as the superior policy instrument if the share of constrained firms becomes sufficiently large. This is a consequence of two main channels. Firstly, lay-off taxes have trouble deterring separations in financially constrained firms, while short-time work can still operate as effectively as before. Secondly, with financial constraints, firms lose their ability to insure risk-averse workers against negative income shocks. Short-time work can then partially mitigate this and provide insurance against income shocks in the firms' stead.

Quantitatively, after calibrating the model to the U.S. economy, we find that short-time work is the superior policy instrument if 40% of firms in the economy or more are financially constrained.

The backbone of our model is a canonical Diamond-Mortensen-Pissaridis type search and matching model (Mortensen and Pissarides, 1993, Mortensen and Pissarides, 1994) with idiosyncratic productivity shocks, endogenous separations, and generalized Nash-Bargaining between workers and firms. We augment the standard model with three realistic key ingredients: Risk aversion on the worker side, financial constraints on the firm side, and flexibly adjustable working hours. Workers and firms are randomly matched, form expectations over their match productivity, and bargain over wages, working hours, and separation productivity thresholds. Productivity shocks are i.i.d. and realize after firms and workers complete bargaining.

Government implements an unemployment insurance system under which unemployed workers are paid lump sum benefits and either a lay-off tax system or a short-time work scheme. Further, it collects taxes on firms to finance the systems.

In the presence of lay-off taxes, firms have to pay lump-sum lay-off taxes once worker and firm agree to separate after productivity falls below the separation threshold. Importantly, firms pay no lay-off taxes if the worker unilaterally leaves the match.

The Short-time work scheme consists of two main components: eligibility conditions and benefit payments. Workers become eligible for short-time work once they agree to reduce their working hours below a threshold specified by the government. Short-time work benefits compensate for lost income by providing fixed payments for each hour not worked relative to normal working hours. Short-time work effectively acts as a subsidy paid directly to workers. Working hours under short-time work are also determined through bargaining between firms and workers. In this setup, firms and workers have an incentive to reduce working hours inefficiently to attract more government support.

We assume that a share of firms is financially constrained. Specifically, we assume that these firms can never borrow more than the expected discounted value of the firm. These constraints have direct welfare implications. With risk-averse workers and risk-neutral firms, firms would like to offer workers insurance against low productivity shocks and commit to paying the worker a fixed (consumption equivalent) wage, no matter how low or high productivity turns out to be. The worker is then willing to accept a slightly lower wage in return for the insurance. However, if the firm is financially constrained and the borrowing constraint binds, shocks pass through to the worker's wage fully. Workers are not committed to staying in the match and quit once they hit their participation constraint (i.e. once the value from quitting, becoming unemployed, and looking for a new job is greater than staying). Therefore, the welfare effect of financial constraints is twofold: Firstly, firms cannot provide as much insurance to workers as they would like, and secondly, workers quit sooner.

We derive closed-form expressions for optimal lay-off taxes and optimal short-time work schemes, depending on how many firms are financially constrained. The theoretical results show that the sole purpose of lay-off taxes is to offset the fiscal externality of the unemployment insurance system on financially unconstrained firms. However, lay-off taxes cannot correct inefficient separations or address the lack of insurance for workers in financially constrained firms. The intuition is straightforward: once financial constraints become binding, firms can no longer absorb shocks, which are instead passed on to workers in the form of reduced income. If income falls sufficiently, workers eventually choose to quit unilaterally. In such cases, firms are neither able nor obligated to pay lay-off taxes, rendering them ineffective.

By contrast, short-time work is particularly effective in supporting financially constrained firms. By supplementing the income of workers who experience reduced hours, it incentivizes workers to remain attached to the firm. In addition, it provides income insurance.

When the share of financially constrained firms is low, lay-off taxes are clearly preferable to short-time work, as the distortionary cost of reduced working hours under short-time work outweighs the cost of insufficient support for financially constrained firms under a lay-off tax regime. However, as the prevalence of financially constrained firms increases, the trade-off shifts. Both theoretically and quantitatively, we demonstrate that short-time work becomes the superior policy instrument once a sufficiently large share (more than 40%) of firms face financial constraints.

Related literature. We contribute to four branches of the literature. Firstly, we add to the literature about optimal unemployment insurance. The trade-off between the benefits of unemployment insurance and adverse effects on job-search, and separations has been a recurring theme in the economic literature (e.g. Shavell and Weiss, 1979, Baily, 1978). How UI should be used optimally has therefore been examined in seminal papers such as Hopenhayn and Nicolini, 1997 and Chetty, 2006. More recently Landais, Michaillat, and Saez, 2018 how UI should vary with labor market tightness and Kroft and Notowidigdo, 2016 show evidence that moral hazard costs of UI are procyclical. We contribute to this literature by evaluating policy tools that can mitigate the adverse externalities of unemployment insurance.

Secondly, there is a branch of the literature that concerns itself with lay-off taxes, often emphasizing their effectiveness in overcoming UI fiscal externalities. Blanchard and Tirole, 2008 and Cahuc and Zylberberg, 2008 propose frameworks in which lay-off taxes can decentralize the planner solution. Closely related to our work, Jung and Kuester, 2015 and Michau, 2015 take a Ramsey planner approach and examine optimal unemployment insurance with lay-off taxes and vacancy subsidies in a DMP. Duggan, Guo, and Johnston, 2023 show that lay-off taxes can act as a stabilizer over the business cycle. Postel-Vinay and Turon, 2011, on the other hand, show that employers can use severance packages to coax workers into quitting and avoid lay-off taxes. Ratner, 2013 makes the point that the experience-rated UI system in the U.S. reduces layoffs but also hampers hires. Similarly, Johnston, 2021 finds that increases in lay-off taxes lead to less hiring. We contribute to this literature by introducing firm borrowing constraints and showing that they reduce the effectiveness of lay-off taxes.

Thirdly, there is a growing body of research on short-time work, though the literature remains divided on its overall usefulness. Some studies emphasize potential inefficiencies. For example, Burdett and Wright, 1989 argue that short-time work encourages inefficient reductions in working hours. Cooper, Meyer, and Schott, 2017 highlight the risk of subsidizing employment at unproductive firms, while Giupponi and Landais, 2022 raise concerns about impeding beneficial worker reallocations, though they also note that short-time work may support firms in maintaining efficient levels of labor hoarding. Cahuc, Kramarz, and Nevoux, 2021 raise concern about the potential windfall effects of short-time work.

Other contributions focus on the potential advantages of short-time work. Balleer, Gehrke, Lechthaler, and Merkl (2016) show that it can introduce valuable flexibility at the intensive margin of employment. Giupponi, Landais, and Lapeyre (2021) argue that short-time work complements unemployment insurance by insuring against different types of labor market shocks. Similarly, Braun and Brügemann (2017) analyze optimal unemployment insurance and short-time work jointly within an implicit contract model.

Despite these contributions, the relationship between short-time work and unemployment insurance remains insufficiently understood, as emphasized in a comprehensive review by Cahuc (2024). Addressing this gap, Stiepelmann (2024) introduces the analysis of optimal short-time work policy and optimal unemployment insurance into a search-and-matching framework. Building on this foundation, our paper contributes to a better understanding of the interplay between short-time work and unemployment insurance under financial constraints.

Finally, there is a literature on firms' financial constraints. Drechsel, 2023 argues that earnings-based borrowing constraints react more strongly to shocks. A part of the literature shows that financial constraints matter for monetary shock transmissions (e.g. Ottonello and Winberry, 2020) or innovation (e.g. Cascaldi-Garcia, Vukoti, and Zubairy, 2023). In the labor market, fewer financial constraints are empirically shown to lead to higher employment (e.g. Duygan-Bump, Levkov, and Montoriol-Garriga, 2015, Fonseca and Van Doornik, 2022). We contribute to this literature by incorporating financial constraints into a search and matching framework with endogenous employment response.

3.2 Model

3.2.1 Basic Assumptions

Set-Up. Time is discrete. There is a unit mass of workers. Workers can be either employed or unemployed. While unemployed, workers search for vacant jobs and receive unemployment benefits b. Unemployed workers and firms with vacancies are matched randomly. The number of contacts is governed by the Cobb-Douglas matching function $m(v,n) = \bar{m} \cdot v^{(1-\gamma)} \cdot (1-n)^{\gamma}$, where v is the mass of vacancies and n is the mass of employed workers. We denote the job-finding rate by f, the vacancy-filling rate by q, and the separation rate by ρ. Let θ denote labor market tightness, i.e. $\theta = \frac{v}{1-n}$.

A worker-firm match produces output y according to the production function

$$y(\epsilon, h) \;=\; A \cdot \epsilon \cdot h^{\alpha} - (\mu_{\epsilon} - \epsilon) \cdot c_f$$

where ϵ is realization of a productivity shocj $\tilde{\epsilon}$ satisfying shock $\log \tilde{\epsilon} \sim \mathcal{N}(\mu, \sigma^2)$ with cdf $G(\epsilon)$ and pdf $g(\epsilon)$. New productivity shocks arrive with probability λ. h is the number of working hours worked by the worker, and A is the total factor productivity. The term $(\mu_{\epsilon} - \epsilon) \cdot c_f$ is a cost shock. While workers are employed they receive a salary $w(\epsilon)$ and suffer dissutility $\phi(h) = \frac{h^{(1+\psi)}}{1+\psi}$ from working h hours.

Firm profits go to firm owners of whom there is a mass of ν_f.

Preferences. Workers are risk-averse with a concave flow utility function over consumption, net of work disutility $u(c-\phi(h))$. Because utility is defined over consumption-equivalent units, we introduce notation for production in consumption-equivalent units for use in later expressions and let $z(\epsilon,h) = y(\epsilon,h)-\phi(h)$. Firm owners have the same flow utility function u but draw utility only from their consumption c_f (i.e. $u(c_f)$).

Government Policy. The government can choose between two policy regimes: the Short-Time Work (STW) regime and the Lay-Off Tax (LT) regime. Since the chosen policy regime has implications for firm and worker value, bargaining, and equilibrium, the remaining model assumptions are described in two separate sections for each respective policy regime in the following.

3.2.2 The Model with Lay-off Taxes

Government. Under the LT regime, firms have to pay lay-off taxes F once the worker and firm jointly agree to separate. Importantly, firms pay no lay-off taxes if the worker unilaterally leaves the match. In the following, let ρ^F denote the probability that firms and workers separate, and the firm has to pay lay-off taxes. The lay-off tax F and the unemployment insurance benefits b are set by the government. To finance the policy regime, firms pay a lump-sum tax τ whenever their productivity changes. The government budget constraint is

$$n^s \cdot \tau \;=\; (1-n)\cdot b - n^s \cdot \rho^F \cdot F$$

where n^s is the mass of matches that receive a new productivity shock.

Firms. There are two types of firms - financially constrained and financially unconstrained firms. Firms are financially constrained with probability p. When constrained, once the productivity shock has realized to ϵ, a firm can borrow no more than its expected value, conditional on its realized productivity. Specifically,

$$y(\epsilon,h(\epsilon)) - w^c(\epsilon) \;\geq\; -\lambda \cdot \bar{J} - (1-\lambda)\cdot J^c(\epsilon)$$

must hold.$\bar{J}$ is the expected firm value once a new shock arrives, and $J^c(\epsilon)$ is the value of a constrained firm with productivity draw ϵ. The constrained (monthly) wage function $w^c(\epsilon)$ will be made explicit in the bargaining section below. The bargained-over hours functions will depend on productivity ϵ but not on the financial constraint. Anticipating this, we save on notation and denote $h^u(\epsilon)$ and $h^c(\epsilon)$ as $h(\epsilon)$. Note that with persistent shocks ($\lambda < 1$), a low realization of productivity will lead to a tighter borrowing constraint.

The value of an unconstrained firm *after* productivity realizes to ϵ is

$$J^u(\epsilon) \;=\; y(\epsilon, h(\epsilon)) - w^u(\epsilon) + \lambda \bar{J} + (1-\lambda) J^u(\epsilon)$$

and the value of a constrained firm *after* productivity realizes to ϵ that is

$$J^c(\epsilon) \;=\; y(\epsilon, h(\epsilon)) - w^c(\epsilon) + \lambda \bar{J} + (1-\lambda) J^c(\epsilon)$$

where $\bar{J}$ is the expected firm value *before* the shock has realized and *before* it becomes known if the firm is constrained. It is given by

$$\bar{J} \;=\; -\tau + (1-p)\left(\int_{\epsilon_s^u}^{\infty} J^u(\epsilon)\, dG(\epsilon)\right) + p\left(\int_{\epsilon_s^c}^{\infty} J^u(\epsilon)\, dG(\epsilon)\right) \quad - \rho^F \cdot F - \rho^L \cdot L$$

Firms separate from a worker once the productivity drops below the separation threshold ϵ_s^u or ϵ_s^c for the unconstrained and constrained case, respectively. With probability ρ^F, firms and workers agree on separation, and the firm has to pay a lay-off tax F. When firms become financially constrained, they also lack the means to pay lay-off taxes. In this case, we assume that they become bankrupt. In the event of bankruptcy, firm owners have to pay liquidation costs $0 < L < F$. Let ρ^L denote the probability of this event.

Firms can freely enter the labor market and post vacancies at cost k_v. The mass of vacancies v must therefore solve

$$\frac{k_v}{q} \;=\; \bar{J}$$

Workers. Employed workers can work at financially constrained and unconstrained firms. The value of a worker at an unconstrained firm *after* productivity realizes to ϵ is

$$V^u(\epsilon) \;=\; u(w^u(\epsilon) - \phi(h(\epsilon))) + \lambda \bar{V} + (1-\lambda) V^u(\epsilon).$$

The value of a worker at a constrained firm *after* productivity realizes to ϵ is

$$V^c(\epsilon) \;=\; u(w^c(\epsilon) - \phi(h(\epsilon))) + \lambda \bar{V} + (1-\lambda) V^c(\epsilon)$$

where $\bar{V}$ is the expected worker value *before* a new productivity shock has realized It is given by

$$\bar{V} \;=\; (1-p)\left(\int_{\epsilon_s^u}^{\infty} V^u(\epsilon)\, dG(\epsilon)\right) + p\left(\int_{\epsilon_s^c}^{\infty} V^c(\epsilon)\, dG(\epsilon)\right) + \rho \cdot U.$$

With probability ρ a worker becomes unemployed. The Value of an unemployed worker U is given by

$$U = u(b) + f \cdot \bar{V} + (1-f) \cdot U$$

Bargaining. With financial constraints, bargaining takes place before both - the realization of productivity ϵ and whether the firm is constrained or not - become known. For each possible state $\epsilon \in \mathbb{R}_+$ worker and firm agree on the total monthly wage functions $w^u(\epsilon)$ and $w^c(\epsilon)$, the hours functions $h^u(\epsilon)$ and $h^c(\epsilon)$, on the separation threshold when unconstrained ϵ_s^u, as well as the separation threshold when constrained ϵ_s^c.

Further, there is no commitment to the contract on the workers' side. This means that if the outside option of the worker becomes weakly better than staying with the match, the worker quits unitarily:

$$V^u(\epsilon) \leq U, \qquad V^c(\epsilon) \leq U$$

This will become especially prevalent in the case where firms become financially constrained. Note that when workers quit unitarily the firm does not have to pay the lay-off tax. The commitment problem of workers implies that we can denote the probability that firms and workers decide to separate and the firm does not have to pay the lay-off tax as:

$$\rho^F = (1-p) \cdot G(\epsilon_s^u) \cdot \mathbb{1}(V^u(\epsilon_s^u) > U)$$

Note that financially constrained firms do not have the means to pay lay-off taxes. In this case, they need to pay the liquidation cost L. With the commitment problem of the worker, the probability of this event is:

$$\rho^L = p \cdot G(\epsilon_s^c) \cdot \mathbb{1}(V^c(\epsilon_s^c) > U)$$

Formally, the bargaining outcome is the solution to the maximization problem

$$\max_{w^u(\epsilon), w^c(\epsilon), h^u(\epsilon), h^c(\epsilon), \epsilon_s^u, \epsilon_s^c} \bar{J}^{(1-\eta)} \left(\bar{V} - U\right)^{\eta}$$

subject to

$$\begin{aligned} \text{Commitment Problem:} \quad & V^u(\epsilon) > U, \; V^c(\epsilon) > U, \\ \text{Financial Constraints:} \quad & y(\epsilon, h^c(\epsilon)) - w^c(\epsilon) \geq -\lambda \cdot \bar{J} - (1-\lambda) \cdot J^c(\epsilon) \end{aligned}$$

where η denotes the bargaining power of workers. Note that firms and workers have to take the commitment problem of the worker and the financial constraints of the firm into account when writing their contracts. One concern raised by Postel-Vinay and Turon, 2011 is that firms may encourage workers to quit using severance packages, in order to circumvent lay-off taxes. To isolate the core trade-off, we abstract from such strategic behavior, effectively comparing short-time work to an idealized benchmark of a lay-off tax. Appendix 3.B.1 derives the bargaining outcomes in detail.

Since workers are risk-averse and firms are risk-neutral, the bargaining outcome is that firms insure workers against low-productivity states. Workers will accept a lower average wage in exchange for insurance. As long as the firm is not constrained (either because it is an unconstrained firm or because constraints do not bind) the firm will *fully* insure the worker against productivity risk and the total monthly wage $w(\epsilon)$ will be set to a constant consumption equivalent c^w, equating worker utility across all realizations of ϵ that do not lead to binding constraints. c^w is pinned down by the condition

$$u'(c(\epsilon)) \;=\; u'(c^w)$$

Once a firm becomes constrained and the constraint binds, it instead pays the maximum monthly wage $c^w(\epsilon)$ it can still afford:

$$c^w(\epsilon) \;=\; z(\epsilon) + \lambda \cdot \bar{J}$$

The hours function that worker and firm agree on is, as already mentioned, the same for constrained and unconstrained firms. It is pinned down by

$$A \cdot \alpha \cdot \epsilon \cdot h(\epsilon)^{\alpha-1} \;=\; h(\epsilon)^{\psi}$$

both, in the constrained and the unconstrained case. The equilibrium, therefore, only contains one general hours Function $h(\epsilon)$. Since $h(\epsilon)$ equates the marginal disutility of work and the marginal product of labor, working hours are always set efficiently.

Note that because under bargaining, each productivity level ϵ will imply a working hours level $h(\epsilon)$, we write $z(\epsilon, h(\epsilon))$ simply as $z(\epsilon)$. The job-destruction equations pin down the bargained-over separation thresholds of constrained and unconstrained matches

$$\begin{aligned} z(\epsilon_s^u) + F + \frac{u(c^v) - u(b)}{u'(c^w)} - c^w + \frac{\lambda - \eta \cdot f}{1-\eta} \cdot \frac{k_v}{q} \;&=\; 0 \\ \frac{u(c^w(\epsilon_s^c)) - u(b)}{u'(c^w)} + (\lambda - f) \cdot \frac{\eta}{1-\eta} \cdot \frac{k_v}{q} \;&=\; 0 \end{aligned}$$

Note that F enters the separation condition for unconstrained firms only. In the unconstrained case, firms insure workers' income against idiosyncratic productivity shocks such that a low idiosyncratic productivity level is not passed on to the worker. As a result, the worker has no incentive to quit unilaterally $V^u(\epsilon_s^u) > U$, and firms have to pay the lay-off tax. A low productivity level in a financially constrained firm is passed onto the income of the worker. At some point, the income received from the firm is so low that the worker quits unilaterally $V^c(\epsilon_s^c) \leq U$. Thus, the firm does not have to pay a lay-off tax in this case.

Labor Markets. The separation rate ρ is given by $\rho = (1-p) \cdot \rho^u + p \cdot \rho^c$ where $\rho^u = G(\epsilon_s^u)$ and $\rho^c = G(\epsilon_c^c)$. The steady-state law of motion for employment is

$$n = (1-\lambda) \cdot n + (1-\rho) \cdot n^s,$$

where n^s denotes the number of matches that received a new shock

$$n^s = f \cdot (1-n) + \lambda \cdot n$$

The mass of unemployment workers can be expressed as $u = 1 - n$. The mass of workers who work for unconstrained firms n^u, and the mass of workers who work at constrained firms n^c are given by:

$$\begin{aligned} n^u &= \frac{1-p}{\lambda} \cdot (1-\rho^u) \cdot n^s \\ n^c &= \frac{p}{\lambda} \cdot (1-\rho^c) \cdot n^s \end{aligned}$$

Equilibrium. A steady state Equilibrium consists of the working hours function $h(\epsilon)$, the consumption equivalent c^w paid as the monthly wage to workers without binding constraints, the consumption-equivalent monthly wage $c^w(\epsilon)$ paid when financial constraints bind, the separation thresholds ϵ_s^u and ϵ_s^c, the productivity level at which the borrowing constraint becomes binding ϵ^p and labor market flows, i.e the job-finding rate f, the vacancy-filling rate q and the separation rate ρ^u and ρ^c as well as n. The exact equations pinning down equilibrium and their derivation are delegated to Section 3.B in the appendix.

3.2.3 The Model with Short-Time Work

Government. Under the STW regime, firms become eligible for short-time work benefits if they set their working hours on short-time work $h_{\text{stw}}(\epsilon)$ below a threshold D. In this case, the worker receives $\tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon))$ worth of benefits. $\bar{h}$ is a parameter that reflects the normal working hour level. In essence, for every hour the worker works less than she would under normal circumstances, she receives compensation τ_{stw}. Choosing an hours threshold D is a close implementation of what short-time work rules mandate in e.g. Germany.[2] Enforcing working hours reduction can help screen out firms that are actually productive from receiving subsidies because only unproductive firms will find hours reductions optimal (Teichgräber, Žužek, and Hensel, 2022).

In principle, government chooses the policy parameters D, τ_{stw} and b. However, since the hours functions $h^u(\epsilon)$, $h^c(\epsilon)$, $h^u_{\text{stw}}(\epsilon)$ and $h^c_{\text{stw}}(\epsilon)$ will always be strictly decreasing in ϵ, choosing D is equivalent to setting the eligibility productivity threshold ϵ_{stw} that will induce the same hours threshold D. The resulting government budget constraint is

$$\begin{aligned} n^s \cdot \tau \;=\; & \frac{1-p}{\lambda} \cdot n^s \cdot \int_{\epsilon^u_s}^{\max\{\epsilon_{\text{stw}}, \epsilon^u_s\}} \tau_{\text{stw}} \cdot (\bar{h} - h_{\text{stw}}(\epsilon))\, dG(\epsilon) \\ & + \frac{p}{\lambda} \cdot n^s \cdot \int_{\epsilon^c_s}^{\max\{\epsilon_{\text{stw}}, \epsilon^c_s\}} \tau_{\text{stw}} \cdot (\bar{h} - h_{\text{stw}}(\epsilon))\, dG(\epsilon) \\ & + (1-n) \cdot b \end{aligned}$$

where n^s is the mass of matches that receive a new productivity shock. ϵ^u_s and ϵ^c_s are the separation thresholds of matches with access to short-time work of unconstrained and constrained matches, respectively. The max operators in the integral limits are necessary because the government can choose an eligibility condition that is so tight that all matches separate without gaining access to short-time work first.

Firms. Again, firms can be either financially constrained or unconstrained. With short-time work, they can also currently be on short-time work or not. Firms are financially constrained with probability p. When constrained, once the productivity shock has realized to ϵ, a firm can borrow no more than its expected value, conditional on its realized productivity. Specifically, without access to short-time work

$$y(\epsilon, h(\epsilon)) - w^c(\epsilon) \;\geq\; -\lambda \cdot \bar{J} - (1-\lambda) \cdot J^c(\epsilon)$$

and with short-time work

$$y(\epsilon, h_{\text{stw}}(\epsilon)) - w^c_{\text{stw}}(\epsilon) \;\geq\; -\lambda \cdot \bar{J} - (1-\lambda) \cdot J^c_{\text{stw}}(\epsilon)$$

must hold. $\bar{J}$ is the expected firm value once a new shock arrives and $J^c(\epsilon)$ is the value of a constrained firm without short-time work and $J^c_{\text{stw}}(\epsilon)$ is the value of a constrained firm on short-time work. The constrained monthly wage function

2. In practice, this is implemented as a minimum reduction of hours threshold - which in our model is given by $\frac{D-h}{h}$ and is implicitly chosen by the government through D.

$w^c(\epsilon)$ will be made explicit in the bargaining section below. Like in the model with lay-off taxes, the bargained-over hours functions will depend on productivity ϵ but not on the financial constraint. Anticipating this, we save on notation and denote $h^u(\epsilon)$ and $h^u(\epsilon)$ as $h(\epsilon)$ and $h^u_{\text{stw}}(\epsilon)$ and $h^c_{\text{stw}}(\epsilon)$ as $h_{\text{stw}}(\epsilon)$. Further we write $z(\epsilon)$ instead of $z(\epsilon, h(\epsilon))$ and $z_{\text{stw}}(\epsilon)$ instead of $z(\epsilon, h_{\text{stw}(\epsilon)})$. The value of firm *without* short-time work *after* productivity realizes to ϵ that is unconstrained is

$$J^u(\epsilon) = y(\epsilon, h(\epsilon)) - w^u(\epsilon) + \lambda \bar{J} + (1-\lambda) J^u(\epsilon)$$

and the value of an unconstrained firm *with* short-time work *after* productivity realizes is

$$J^u_{\text{stw}}(\epsilon) = y(\epsilon, h_{stw}(\epsilon)) - w^u_{\text{stw}}(\epsilon) + \lambda \bar{J} + (1-\lambda) J^u_{\text{stw}}(\epsilon).$$

The value of firm *without* short-time work *after* productivity realizes to ϵ that is constrained is

$$J^c(\epsilon) = y(\epsilon, h(\epsilon)) - w^c(\epsilon) + \lambda \bar{J} + (1-\lambda) J^c(\epsilon)$$

and the value of a constrained firm *with* short-time work *after* productivity realizes is

$$J^c_{\text{stw}}(\epsilon) = y(\epsilon, h_{stw}(\epsilon)) - w^c_{\text{stw}}(\epsilon) + \lambda \bar{J} + (1-\lambda) J^c_{\text{stw}}(\epsilon)$$

where $\bar{J}$ is the expected firm value *before* the shock has realized. It is given by

$$\begin{aligned}\bar{J} = -\tau + (1-p)\left(\int_{\max\{\xi^u_s, \epsilon_{\text{stw}}\}}^{\infty} J^u(\epsilon)\, dG(\epsilon) + \int_{\epsilon^u_s}^{\epsilon_{\max\{\epsilon^u_s, \text{stw}\}}} J^u_{\text{stw}}(\epsilon)\, dG(\epsilon)\right) \\ + p\left(\int_{\max\{\xi^c_s, \epsilon_{\text{stw}}\}}^{\infty} J^c(\epsilon)\, dG(\epsilon) + \int_{\epsilon^c_s}^{\max\{\epsilon^c_s, \epsilon_{\text{stw}}\}} J^c_{\text{stw}}(\epsilon)\, dG(\epsilon)\right)\end{aligned}$$

The max operators in the integral bounds $\max\{\xi^i_s, \epsilon_{\text{stw}}\}$ reflect that a sufficiently strict eligibility threshold ϵ_{stw} can exclude certain matches from accessing short-time work, even if their productivity would allow them to survive with such support but not without it. In this case, otherwise, viable matches are denied access to the short-time work system, leading to unnecessary separations. Firms can freely enter the labor market and post vacancies at cost k_v. The mass of vacancies v must therefore solve

$$\frac{k_v}{q} = \bar{J}$$

Workers. The value of a worker at an unconstrained firm *without* short-time work *after* productivity realizes to ϵ is

$$V^u(\epsilon) \; = \; u(w^u(\epsilon) - \phi(h(\epsilon))) + \lambda\bar{V} + (1-\lambda)V^u(\epsilon)$$

and the value of a worker at an unconstrained firm *with* short-time work *after* productivity realizes is

$$V^u_{\text{stw}}(\epsilon) \; = \; u(w^u_{\text{stw}}(\epsilon) + \tau_{\text{stw}} \cdot (\bar{h} - h_{\text{stw}}(\epsilon)) - \phi(h_{\text{stw}}(\epsilon))) + \lambda\bar{V} + (1-\lambda)V^u_{\text{stw}}(\epsilon)$$

The value of worker *without* short-time work *after* productivity realizes to ϵ who works at a constrained firm is

$$V^c(\epsilon) \; = \; u(w^c(\epsilon) - \phi(h(\epsilon))) + \lambda\bar{V} + (1-\lambda)V^c(\epsilon)$$

and the value of a worker *with* short-time work *after* productivity realizes who works at a constrained firm is

$$V^c_{\text{stw}}(\epsilon) \; = \; u(w^c_{\text{stw}}(\epsilon) + \tau_{\text{stw}} \cdot (\bar{h} - h_{\text{stw}}(\epsilon)) - \phi(h^c_{\text{stw}}(\epsilon))) + \lambda\bar{V} + (1-\lambda)V^c_{\text{stw}}(\epsilon)$$

where $\bar{V}$ is the expected worker value *before* a new productivity shock has realized. It is given by

$$\begin{aligned}\bar{V} \; = \; & (1-p)\left(\int_{\max_{\{\epsilon_{\text{stw}},\xi^u_s\}}}^{\infty} V^u(\epsilon)\, dG(\epsilon) + \int_{\epsilon^u_s}^{\max\{\epsilon_{\text{stw}},\epsilon^u_s\}} V^u_{\text{stw}}(\epsilon)\, dG(\epsilon)\right) \\ & + p\left(\int_{\max\{\epsilon_{\text{stw}},\xi^c_s\}}^{\infty} V^c(\epsilon)\, dG(\epsilon) + \int_{\epsilon^c_s}^{\max\{\epsilon_{\text{stw}},\epsilon^c_s\}} V^c_{\text{stw}}(\epsilon)\, dG(\epsilon)\right) \\ & + \rho \cdot U\end{aligned}$$

where ρ is the separation rate. The Value of an unemployed worker U is given by

$$U \; = \; u(b) + f \cdot \bar{V} + (1-f) \cdot U$$

Bargaining. Bargaining takes place before both - the realization of productivity ϵ and whether the firm is constrained or not - become known. With short-time work, there are a few more items that workers and firms bargain about than in the lay-off tax regime.

Further, there is no commitment to the contract on the workers' side. This implies that workers will unilaterally choose to leave the match once the outside option of becoming unemployed offers a weakly higher expected utility than remaining employed at the firm.

$$V^u(\epsilon) \leq U, \quad V^u_{stw}(\epsilon) \leq U$$
$$V^c(\epsilon) \leq U, \quad V^c_{stw}(\epsilon) \leq U$$

In the case with and without short-time work, respectively. For each possible state $\epsilon \in \mathbb{R}_+$ worker and firm agree on the total monthly wage functions $w^u(\epsilon)$ and $w^c(\epsilon)$, as well as the total monthly wages paid in both cases while on short-time work $w^u_{\text{stw}}(\epsilon)$ and $w^c_{\text{stw}}(\epsilon)$. The hours functions, likewise, are bargained over separately - for times in which the firm has no access to short-time work ($h^u(\epsilon)$ and $h^c(\epsilon)$) and for times on short-time work ($h^u_{\text{stw}}(\epsilon)$ and $h^c_{\text{stw}}(\epsilon)$). Lastly, the separation thresholds are also agreed upon for both cases: access to short-time work and no access to short-time work. We denote the separation thresholds without short-time work by ξ^u_s and ξ^c_s. ϵ^u_s, and ϵ^c_s denote the separation thresholds while on short-time work.

Formally, the bargaining outcome is the solution to the maximization problem

$$\max_{\substack{w^u(\epsilon), w^c(\epsilon), w^u_{\text{stw}}(\epsilon), w^c_{\text{stw}}(\epsilon), h^u(\epsilon), \\ h^c(\epsilon), h^u_{\text{stw}}(\epsilon), h^c_{\text{stw}}(\epsilon), \epsilon^u_s, \epsilon^c_s, \xi^u_s, \xi^c_s}} \bar{J}^{(1-\eta)} \left(\bar{V} - U\right)^{\eta}$$

subject to

$$\begin{aligned} \text{Commitment Problem:} \quad & V^u(\epsilon) > U, \; V^c(\epsilon) > U, \; V^u_{stw}(\epsilon) > U, \; V^c_{stw}(\epsilon) > U, \\ \text{Financial Constraints:} \quad & y(\epsilon, h^c(\epsilon)) - w^c(\epsilon) \geq -\lambda \cdot \bar{J} - (1-\lambda) \cdot J^c(\epsilon) \\ & y(\epsilon, h^c_{\text{stw}}(\epsilon)) - w^c_{\text{stw}}(\epsilon) \geq -\lambda \cdot J^c_{\text{stw}}(\epsilon) - (1-\lambda) \cdot \bar{J} \end{aligned}$$

where η denotes the bargaining power of workers. Again, firms and workers need to take the commitment problem of the worker and the financial constraints of the firm into account when writing contracts. Appendix 3.C.1 derives the bargaining outcomes.

Since workers are risk-averse and firms are risk-neutral, the bargaining outcome is still that firms insure workers against low-productivity states, while workers will accept a lower average wage in exchange for insurance. Much like in the lay-off tax model, as long as the firm is not constrained (either because it is an unconstrained firm or because constraints do not bind) the firm will *fully* insure the worker against productivity risk and the total monthly wage $w(\epsilon)$ will be set to a constant consumption equivalent c^w, equating worker utility across all realizations of ϵ that do not lead to binding constraints. This also includes the state where workers are on STW:

$$u'(c^w(\epsilon)) \;=\; u'(c^w_{stw}(\epsilon)) \;=\; u'(c^w)$$

Once a firm becomes constrained and the constraint binds, it instead pays the maximum monthly wage $c^w(\epsilon)$ it can still afford. This will depend on whether the firm has access to short-time work or not:

$$\begin{aligned} c^w(\epsilon) &= z(\epsilon) + \lambda \cdot \frac{k_v}{q} \\ c^w_{\text{stw}}(\epsilon) &= z_{\text{stw}}(\epsilon) + \tau_{\text{stw}} \cdot (\bar{h} - h_{\text{stw}}(\epsilon)) + \lambda \cdot \frac{k_v}{q}. \end{aligned}$$

When the firm goes on short-time work, the worker is compensated for the reduced working hours by the government. This means that when the firm has a binding borrowing constraint and can only afford to pay the worker $c_{\text{stw}}(\epsilon)$ in monthly wages, going on short-time work can still increase the worker's income. The hours function that worker and firm agree on also differs depending on whether the firm is on short-time work or not. The two cases are pinned down by the conditions

$$\begin{aligned} A \cdot \alpha \cdot \epsilon \cdot h(\epsilon)^{\alpha-1} &= h(\epsilon)^{\psi} \\ A \cdot \alpha \cdot \epsilon \cdot h_{\text{stw}}(\epsilon)^{\alpha-1} &= h_{\text{stw}}(\epsilon)^{\psi} + \tau_{\text{stw}}. \end{aligned}$$

Again, the hours functions turn out to be independent of financial constraints and equilibrium; therefore, they only contain the two general hours functions $h(\epsilon)$ and $h_{\text{stw}(\epsilon)}$. Note that without short-time work, the bargaining outcome for working hours equates to the marginal product of labor and the marginal disutility from labor and thus sets working hours efficiently. With short-time work, working hours take the subsidy τ_{stw} into account, and hours are set lower than the efficient number of working hours. This means that through inefficiently low working hours, the short-time work scheme introduces a distortion into the economy. The bargained-over separation thresholds are pinned down by the job-destruction equations

$$\begin{aligned} z(\xi^u_s) + \frac{u(c^w) - u(b)}{u'(c^w)} - c^w + \frac{\lambda - \eta \cdot f}{1-\eta} \cdot \frac{k_v}{q} &= 0 \\ z(\epsilon^u_s) + \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon^u_s)) + \frac{u(c^w) - u(b)}{u'(c^w)} - c^w + \frac{\lambda - \eta \cdot f}{1-\eta} \cdot \frac{k_v}{q} &= 0 \\ \frac{u(c(\xi^u_s)) - u(b)}{u'(c^w)} + (\lambda - f) \cdot \frac{\eta}{1-\eta} \cdot \frac{k_v}{q} &= 0 \\ \frac{u(c_{stw}(\epsilon^c_s)) - u(b)}{u'(c^w)} + (\lambda - f) \cdot \frac{\eta}{1-\eta} \cdot \frac{k_v}{q} &= 0. \end{aligned}$$

In unconstrained firms, the separation decision is determined as a bargaining outcome: firms and workers separate once their joint surplus becomes negative. Importantly, short-time work benefits are not included in the worker's utility function. The rationale is straightforward—under STW, the firm continues to insure the worker's income against idiosyncratic productivity shocks. The worker receives a constant consumption equivalent c^w, regardless of the realization of ϵ. As a result, the worker has no incentive to quit unilaterally. STW operates by reducing the wage burden on the firm required to sustain this consumption level. In effect, STW lowers the cost of providing income insurance, thereby reducing the firm's incentive to initiate separations.
In unconstrained firms, the worker's commitment problem becomes binding: low productivity realizations are passed through to the worker's income. If income falls sufficiently, the worker chooses to quit the match. Short-time work increases the worker's available income in such cases, thereby increasing the worker's incentive to remain with the firm.

Labor Markets. The separation rate ρ is given by $\rho = (1-p)\cdot\rho^u + p\cdot\rho^c$ where $\rho^u = G(\epsilon_s^u) + G(\max\{\xi_s^u, \epsilon_{\text{stw}}\}) - G(\max\{\epsilon_s^u, \epsilon_{\text{stw}}\})$ and $\rho^c = G(\epsilon_c^c) + G(\max\{\xi_s^c, \epsilon_{\text{stw}}\}) - G(\max\{\epsilon_s^c, \epsilon_{\text{stw}}\})$. The remaining worker flows are given by $f = \frac{m}{1-n}$ and $q = \frac{m}{v}$. The steady-state law of motion for employment is

$$n = (1-\lambda)\cdot n + (1-\rho)\cdot n^s,$$

where n^s denotes the number of matches that received a new shock

$$n^s = f\cdot(1-n) + \lambda\cdot n$$

Unemployment can be denoted as $u = 1-n$. The number of workers who work for unconstrained firms n^u, and the number of workers who work at constrained firms n^c are given by:

$$\begin{aligned} n^u &= \frac{1-p}{\lambda}\cdot(1-\rho^u)\cdot n^s \\ n^c &= \frac{p}{\lambda}\cdot(1-\rho^c)\cdot n^s \end{aligned}$$

Equilibrium. A steady state Equilibrium consists of the working hours functions $h(\epsilon)$, $h_{\text{stw}}(\epsilon)$, the consumption equivalent c^w paid as the monthly wage to workers without binding constraints, the consumption-equivalent wage $c^w(\epsilon)$ and c^w_{stw} paid when financial constraints bind, the separation thresholds ξ_s^u, ξ_s^c, ϵ_s^u and ϵ_s^c and labor market flows, i.e the job-finding rate f, the vacancy-filling rate q and the separation rates ρ^u and ρ^c as well as n. The exact equations pinning down equilibrium and their derivation are delegated to Section 3.C in the appendix.

3.3 Optimal Policy

3.3.1 The Ramsey Problem

To show the differences in how short-time work and lay-off taxes act against the fiscal externality of UI on separations, we take a Ramsey planner approach. We set up separate Ramsey problems for the lay-off tax regime and the short-time work regime and proceed to compare optimal policies and their implications. In both cases, the Ramsey planner maximizes welfare, subject to the equilibrium constraints stated in Sections 3.B and 3.C in the appendix.

To state the welfare function formerly, recall that there is a mass of firm owners ν^f. Firm owners have the same preferences over consumption as workers and receive equal shares of firm profits in the economy. Further, let Ω denote the *average* loss of production in consumption equivalent units due to hours distortions:

$$\Omega = \frac{1}{1-\rho}\left((1-p)\cdot \underbrace{\int_{\epsilon_s^u}^{\max\{\epsilon_{\text{stw}},\epsilon_s^u\}} \Omega(\epsilon)\, dG(\epsilon)}_{:=(1-\rho^u)\Omega^u} + p\cdot \underbrace{\int_{\epsilon_s^c}^{\max\{\epsilon_{\text{stw}},\epsilon_s^c\}} \Omega(\epsilon)\, dG(\epsilon)}_{:=(1-\rho^c)\Omega^c} \right)$$

where

$$\Omega(\epsilon) = z(\epsilon) - z_{\text{stw}}(\epsilon).$$

Stiepelmann, 2024 shows that $\Omega \geq 0$, $\frac{\partial \Omega}{\partial \tau_{\text{stw}}}$ and $\frac{\partial \Omega}{\partial \epsilon_{\text{stw}}}$.

The welfare function is given by

$$\begin{aligned} W(\pi) = {} & n^u \cdot u(c^w) + n^c \cdot u^c + (1-n)\cdot u(b) \\ & + \nu^f \cdot u((n\cdot(z-\Omega) - n^u\cdot c^w - n^c \cdot e^c - \tau(b) - \theta\cdot(1-n)\cdot k_v)/\nu^f) \end{aligned} \tag{3.3.1}$$

where n^u is the mass of workers employed at firms that do not hit their constraint and n^c is the mass of workers whose employers have a binding borrowing constraint. u^c is the *average* utility of a worker at a firm that has hit its borrowing constraint. z is the *average* production, net of work-disutility, and Ω is the *average* loss of production in consumption equivalent units due to hours distortions: Under the lay-off tax regime, $\Omega = 0$ will always hold. The *average* total monthly wage of a worker at a firm that has hit its borrowing constraint is e^c. The UI-tax $\tau(b) = (1-n)\cdot b$ is used to finance the UI system.

π is the vector of policy parameters the planner can choose, i.e. $\pi = (b, F)$ in the lay-off tax regime and $\pi = (b, \epsilon_{\text{stw}}, \tau_{\text{stw}})$ under the short-time work regime. The problem of the Ramsey planner is given by

$$\max_{\pi} W(\pi) \quad \text{s.t. equilibrium conditions are fulfilled}$$

The full problem is stated in Section 3.D in the appendix. Since the analysis does not focus on distributional effects between firm owners and workers, we set the mass of firm owners ν^f such that the consumption equivalent of workers always equals the consumption of firm owners.

3.3.2 The Optimal Lay-Off Tax Regime

Before we turn our attention to how F is set optimally, we first look at how the planner sets b in the presence of a lay-off tax. This allows us to illustrate how a lay-off tax interacts with the fiscal externalities caused by unemployment insurance.

Setting the unemployment insurance level b must balance a trade-off. On the one hand, the Ramsey planner would like to insure the risk-averse worker against income losses after a separation. If there were no welfare costs to increasing b, the Ramsey planner would like to set b to a level that fully equates the utility of employed and unemployed workers. However, as is well known, unemployment insurance leads to fiscal externalities.

Increasing b will lead to more separations as the workers' outside option (i.e., U) becomes more attractive. This will be true for workers at unconstrained and constrained firms. This decreases the expected firm value, and vacancy posting falls as well. On top of that, with risk-averse workers, constrained firms will lose some of their ability to insure workers against bad productivity shocks because with higher b, the continuation value of a firm will be smaller, further tightening the borrowing constraint. The resulting trade-off between more worker insurance and greater fiscal externalities is balanced by the Ramsey planner's first order condition.

Proposition 1
The optimal unemployment insurance benefit b given lay-off tax F must fulfill the following first-order condition:

$$\underbrace{(1-n)\cdot\left(\frac{u'(b)-u'(c^w)}{u'(c^w)}\right)}_{\text{MUIB}} = \underbrace{\left(-\frac{\mathrm{d}f}{\mathrm{d}b}\right)^{ge}\cdot u\cdot L_v(b)}_{\text{MLV}} + \underbrace{\left(\frac{\mathrm{d}\epsilon_s^u}{\mathrm{d}b}\right)^{ge}\cdot\frac{\partial n^u(p)}{\partial\epsilon_s^u}\cdot L_s^u(b,F)}_{\text{MLS}^u}$$
$$+\underbrace{\left(\frac{\mathrm{d}\epsilon_s^c}{\mathrm{d}b}\right)^{ge}\cdot\frac{\partial n^c(p)}{\partial\epsilon_s^c}\cdot L_s^c(b)}_{\text{MLS}^c}+\underbrace{n^c(p)\cdot\left(-\left(\frac{\mathrm{d}\theta}{\mathrm{d}b}\right)^{ge}\right)\cdot\hat{IE}_\theta}_{\text{MIE}^c}$$

where

$$\begin{aligned}
L_v &= \left(\frac{\eta-\gamma}{(\eta-\gamma)(1-\eta)}\cdot\bar{J}+\frac{1}{f}\cdot b\right)\\
L_s^u &= \lambda\cdot\left(\frac{1}{f}\cdot b-F\right)\\
L_s^c &= \frac{\lambda}{f}\cdot b\\
\hat{IE}_\theta &= \frac{1}{1-\rho^c}\left(\int_{\epsilon_s^c}^{\epsilon^p}\lambda\cdot\frac{\gamma}{f}\cdot\frac{u'(c(\epsilon))-u'(c^w)}{u'(c^w)}\,dG(\epsilon)\right)\cdot k_v\\
\frac{\partial n^u(p)}{\partial\epsilon_s^u} &= \frac{1-p}{\lambda}\cdot n^s\cdot g(\epsilon_s^u)\\
\frac{\partial n^c(p)}{\partial\epsilon_s^c} &= \frac{p}{\lambda}\cdot n^s\cdot g(\epsilon_s^c)
\end{aligned}$$

PROOF: See Section 3.E.6 in the appendix.

The term labeled *MUIB* represents the marginal social benefit of a higher unemployment benefits level b, stemming from improved income insurance for unemployed workers.

The term labeled *MLV* captures the marginal social loss associated with a decline in hiring induced by an increase in unemployment insurance benefits. The term comprises two components. First, L_v denotes the social value of hiring one additional worker. Second, $\left(-\frac{\mathrm{d}f}{\mathrm{d}b}\right)^{ge}\cdot u$ represents the reduction in the number of new hires. Intuitively, higher unemployment benefits raise the worker's outside option, which puts upward pressure on wages and reduces the value of a worker for the firm. As a result, firms find it less profitable to post vacancies, leading to a decline in job creation.

The *MLS* terms denote the marginal social loss resulting from increased separations caused by UI benefits. Specifically, MLS^u captures the social cost arising from increased separations in unconstrained firms, while MLS^c captures the corresponding cost in constrained firms. Intuitively, higher unemployment insurance benefits raise workers' outside options, exerting upward pressure on wages. As a result, unconstrained firms want to initiate separations early. In constrained firms, workers are less willing to accept income reductions and thus quit into unemployment sooner.

The term MLS^u consists of two components. First, L_s^u denotes the social value of the marginal match at an unconstrained firm. Second, $\left(\frac{\mathrm{d}\epsilon_s^u}{\mathrm{d}b}\right)^{ge} \cdot \frac{\partial n^u(p)}{\partial \epsilon_s^u}$ represents the number of additional separations in unconstrained firms.

Similarly, MLS^c consists of L_s^c, the social value of the marginal match at a constrained firm, and $\left(\frac{\mathrm{d}\epsilon_s^c}{\mathrm{d}b}\right)^{ge} \cdot \frac{\partial n^c(p)}{\partial \epsilon_s^c}$ which denotes the number of additional constrained matches lost due to higher UI benefits.

Finally, M^c labels the social loss of worker income insurance in financially constrained firms consisting of two parts: $\left(-\left(\frac{\mathrm{d}\theta}{\mathrm{d}b}\right)^{ge}\right) \cdot \hat{IE}_\theta$ denotes the average social loss arising from financially constrained firms' reduced capacity to provide income insurance[3] while n^c denotes the number of constrained firms. Intuitively, higher UI benefits reduce the continuation value of a worker for a firm. Firms can thus borrow less in bad periods, reducing their ability to insure the worker.

Note that, in general, all these loss terms consist of the marginal social effect of b on the respective threshold (or labor market tightness in the case of MIE^c) multiplied by the mass of affected workers.

The lay-off tax F enters the policy trade-off exclusively through the MLS^u term. Importantly, it does not affect any terms associated with financially *constrained* firms. The intuition is straightforward: once firms become financially constrained, they lose their ability to insure workers against adverse productivity shocks. As described in Section 3.2, low productivity is then fully passed through to the worker's income. If the income falls sufficiently, the value of remaining employed falls below the value of unemployment ($V^c(\epsilon) < U$) and the worker will choose to quit into unemployment. Since lay-off taxes are paid only in cases of mutual separation, they do not apply to separations initiated unilaterally by workers who leave due to insufficient total wage offers from financially constrained firms.

As a result, lay-off taxes can only mitigate the negative fiscal externalities of UI by stabilizing the number of separations in *unconstrained* firms. This means that when $p = 0$, lay-off taxes can induce the efficient number of separations in the

3. Note that ϵ^p in IE_θ denotes the productivity level at which financial constraints become binding for a constrained firm.

economy. As p increases, the lay-off taxes gradually lose their effectiveness until they lose their bite completely once $p = 1$.

On the upside, the planner can choose to completely eliminate this part of fiscal UI externalities. This can be done with the optimal Lay-off tax stated in Proposition 2.

Proposition 2 *The optimal level of F is determined by its FOC:*

$$F = \frac{1}{f} \cdot b + BE_{\mathrm{lt}}(b) \qquad (3.3.2)$$

PROOF: See Section 3.E.5 in the appendix.

The first component captures the fiscal externality of the unemployment insurance system. It reflects the expected UI benefits a worker receives upon entering unemployment, and thus represents the fiscal cost that each separation imposes on the UI system. BE_{lt} is a bargaining effect - an expression acknowledging that setting F will have general equilibrium effects on other equilibrium variables such as worker flows via the bargaining process. Note that the optimal lay-off tax implements the efficient number of separations in unconstrained firms as shown in Corollary 1.

Corollary 1
(i) Optimal Lay-off taxes fully eliminate socially inefficient separations in unconstrained *firms.*
(ii) If p = 0, then all inefficient separations are eliminated in the economy
PROOF: See Section 3.E in the appendix.

The proof shows that the optimal lay-off tax sets the Lagrange multiplier for the separation condition of unconstrained firms equal to zero.

Note that our model extends the insights of Cahuc and Zylberberg, 2008 and Blanchard and Tirole, 2008 by embedding optimal lay-off taxes within a modern search-and-matching framework of the labor market. Both of these earlier studies show that lay-off taxes can decentralize the social planner's allocation in settings that focus solely on the separation margin. By contrast, search-and-matching models feature both a hiring margin and a separation margin. In the absence of financial constraints (i.e., when $p = 0$), our model confirms that the logic underlying lay-off taxes - namely, their ability to correct the fiscal externality associated with unemployment insurance - carries over to this richer framework by eliminating inefficient separations.

3.3.3 The Optimal Short-Time Work Regime

Under the short-time work regime the Ramsey planner chooses τ_{stw} and ϵ_{stw} (through D). Again, we first turn to how the planner sets the level of UI benefits b in the presence of the short-time work scheme. Setting the unemployment insurance level b must balance the same trade-off as before:

Proposition 3
The optimal unemployment insurance benefit b given a short-time work scheme $(\tau_{stw}, \epsilon_{stw})$ must fulfill the following first-order condition:

$$\begin{aligned}
\underbrace{(1-n)\cdot\left(\frac{u'(b)-u'(c^w)}{u'(c^w)}\right)}_{\text{MUIB}} &= \underbrace{\left(-\frac{\mathrm{d}f}{\mathrm{d}b}\right)^{ge}\cdot u\cdot L_v(b)}_{\text{MLV}} + \underbrace{\left(\frac{\mathrm{d}\epsilon_s^u}{\mathrm{d}b}\right)^{ge}\cdot\frac{\partial n^u(p)}{\partial \epsilon_s^u}\cdot L_s^u(b,\tau_{\text{stw}})}_{\text{MLS}^u} \\
&+ \underbrace{\mathbb{1}(\epsilon_{\text{stw}} \geq \epsilon_s^c)\cdot\left(\frac{\mathrm{d}\epsilon_s^c}{\mathrm{d}b}\right)^{ge}\cdot\frac{\partial n^c(p)}{\partial \epsilon_s^c}\cdot L_{s,\epsilon}^c(b,\tau_{\text{stw}})}_{\text{MLS}^c} \\
&+ \underbrace{\mathbb{1}(\epsilon_{\text{stw}} \leq \xi_s^c)\cdot\left(\frac{\mathrm{d}\xi_s^c}{\mathrm{d}b}\right)^{ge}\cdot\frac{\partial n^c(p)}{\partial \xi_s^c}\cdot L_{s,\xi}^c(b,\tau_{\text{stw}})}_{\text{MLS}^c} \\
&+ n^c(p)\cdot\underbrace{\left(-\left(\frac{\mathrm{d}\theta}{\mathrm{d}b}\right)^{ge}\right)\cdot \hat{IE}_\theta(\tau_{\text{stw}},\epsilon_{\text{stw}})}_{\text{MIE}^c}
\end{aligned}$$

where

$$
\begin{aligned}
L_v &= \left(\frac{\eta-\gamma}{(\eta-\gamma)(1-n)}\cdot\bar{J}+\frac{\lambda}{f}\cdot b\right)\\
L_s^u &= \left(\frac{\lambda}{f}\cdot b-\tau_{stw}\cdot\left(\bar{h}-h_{stw}(\xi_s^u)\right)\right)\\
L_{s,\epsilon}^c &= \left(\frac{\lambda}{f}\cdot b-\tau_{stw}\cdot\left(\bar{h}-h_{stw}(\xi^c)\right)\right)\\
L_{s,\xi}^c &= \frac{\lambda}{f}\cdot b\\
\hat{IE}_\theta &= \frac{1}{1-\rho^c}\cdot\left(\int_{\max\{\epsilon_{\text{stw}},\xi_s^c\}}^{\epsilon^p}\lambda\cdot\frac{\gamma}{f}\cdot\frac{u'(c(\epsilon))-u'(c^w)}{u'(c^w)}\,dG(\epsilon)\right.\\
&\quad\left.+\int_{\epsilon_s^c}^{\max\{\epsilon_{\text{stw}},\epsilon_s^c\}}\lambda\cdot\frac{\gamma}{f}\cdot\frac{u'(c_{stw}(\epsilon))-u'(c^w)}{u'(c^w)}\,dG(\epsilon)\right)\cdot k_v\\
\frac{\partial n^u(p)}{\partial\epsilon_s^u} &= \frac{1-p}{\lambda}\cdot n^s\cdot g(\epsilon_s^u)\\
\frac{\partial n^c(p)}{\partial\epsilon_s^c} &= \frac{p}{\lambda}\cdot n^s\cdot g(\epsilon_s^c)
\end{aligned}
$$

PROOF: See Section 3.F.6 in the appendix.

Again, L_v is the social value of hiring one additional worker. L_s^u is the social value of the marginal match at an unconstrained firm. $L_{s,i}^c$ denotes the social value of the marginal match in a financially constrained firm. Importantly, constrained firms may feature two distinct separation margins: one at the threshold where matches separate with short-time work support, denoted by $i=\epsilon$, and another where they separate without STW support, denoted by $i=\xi$. The eligibility condition for STW determines which of these separation thresholds is operative. Finally, $\hat{IE}_\theta$ captures the social loss through constrained firms' reduced ability to provide workers with insurance. Each of the terms is weighted by the mass of workers it affects.

Like with the lay-off tax regime, the marginal social benefit of more generous unemployment benefits (*MUIB*) has to equal the marginal welfare losses, i.e. the marginal loss through too few vacancies being posted (*MLV*), too many separations in both unconstrained and constrained firms (MLS^u and MLS^c) and loss of insurance by constrained firms (MIE^c). However, there is a key difference between the FOCs under the lay-off tax regime (Proposition 1) and the FOCs under the short-time work regime. The lay-off tax parameter F entered only the MLS^u term, and lay-off tax could only counteract the adverse effect of unemployment benefits on separations in unconstrained firms. By contrast, the short-time work parameters ($\tau_{\text{stw}},\epsilon_{\text{stw}}$) appear in all the terms except the MLV term. Unlike lay-off taxes, short-time work has an effect not only on separations in unconstrained firms, but can also act on constrained firms and counteract inefficient separations.

The core problem for financially constrained firms is their inability to absorb negative productivity shocks, which forces them to pass the resulting income loss directly onto workers. If the decline in income is large enough, workers may choose to quit. Short-time work functions as a subsidy that partially replaces this lost income, thereby incentivizing workers to remain with the firm. Through this channel, STW not only reduces inefficient separations, captured by the marginal social loss (MLS) terms but also enhances income insurance by acting on the MIE^s term.

A second difference is that the eligibility threshold ϵ_{stw} determines which firms have access to STW, leading to the MLS^c-term being split into two parts. If the eligibility condition is strict, i.e., $\epsilon_{\text{stw}} < \xi_s^c$, then some firms that could otherwise survive with STW support are excluded from the system. Naturally, UI benefits also influence the separation threshold without STW support. Moreover, if ϵ_{srw} is set even strictly to $\epsilon_{\text{stw}} < \epsilon_s^c$, then constrained firms are entirely excluded from the STW system. As a result, unemployment insurance benefits can no longer influence separations under the STW regime, eliminating the corresponding term from the equation.

It is clear from Proposition 3 that short-time work has the advantage of acting on the loss terms of constrained firms. On the downside, as argued in the bargaining paragraph of Section 3.2.3, firms on short-time work will distort their working hours, in turn leading to welfare losses. The Ramsey planner has to trade off the benefits of short-time work as a tool that can counteract the fiscal externalities of UI and the adverse effects of hours distortions. This is reflected in how optimal short-time work is set.

We begin by examining how the eligibility of STW should be determined. The following Proposition describes the optimal eligibility condition:

Proposition 4
Assume that the Ramsey planner never wants to use short-time work to impede vacancy posting by destroying the efficiency of the match.[4] *Depending on where the optimal eligibility threshold* $\epsilon^{\star}_{stw}$ *is set, relative to the different separation thresholds,*

4. In principle, this could happen in extreme cases of deviations from the Hosios condition. In this case, too many vacancies could introduce inefficiencies to such an extent that the Ramsey planner would use short-time work to destroy vacancies. Since this is neither realistic nor the focus of our analysis, we exclude this case. The formal corresponding assumption is stated as Inequality in Section 3.F.5 in the appendix.

its optimality condition differs. There are three cases to distinguish. For each case, the following condition will pin down a candidate for $\epsilon^{\star}_{stw}$. Which of the candidates is the optimum can then be determined by evaluating the welfare function at the three candidate values.

Case 1: $\epsilon^c_s \geq \epsilon_{stw} \geq \xi^u_s$
In this case $\epsilon_{stw} = \xi^u_s$ is the candidate.

Case 2: let $I = \left[\max\{\epsilon^c_s, \xi^u_s\}, \xi^c_s\right]$ and $\epsilon_{stw} \in I$
In this case if

$$\underbrace{\frac{\partial n^c(p)}{\partial \epsilon_{\text{stw}}} \cdot L^c_{s,\epsilon}(\epsilon)}_{\substack{\text{preserve matches in}\\ \text{constrained firms}}} \quad > \quad \underbrace{n^u \cdot \frac{\partial \Omega^u}{\partial \epsilon_{\text{stw}}}}_{\substack{\text{distort working hours}\\ \text{in unconstrained firms}}} + BE_{\text{stw},2} \quad \forall \epsilon \in I, \tag{3.3.3}$$

then the candidate is $\epsilon_{stw} = \xi^c_s$. $BE_{stw,2}$ is a term that captures general equilibrium effects through the generalized Nash bargaining process. If

$$\frac{\partial n^c(p)}{\partial \epsilon_{\text{stw}}} \cdot L^c_{s,\epsilon}(\epsilon) \quad < \quad n^u \cdot \frac{\partial \Omega^u}{\partial \epsilon_{\text{stw}}} + BE_{\text{stw},2} \quad \forall \epsilon \in I, \tag{3.3.4}$$

then the candidate is $\epsilon_{stw} = \xi^u_s$. In the remaining case, i.e., the above inequality holds as an equality, this equality pins down ϵ_{stw}. In this case $\epsilon_{stw} \in \left[\xi^u_s, \xi^c_s\right]$.

Case 3: $\epsilon_{stw} \geq \xi^c_s$
In this case, if

$$\underbrace{n \cdot \frac{\partial \Omega}{\partial \epsilon_{\text{stw}}}}_{\text{distortion}} + BE_{\text{stw},2}$$

$$> \underbrace{\frac{\partial n^c(p)}{\partial \epsilon_{\text{stw}}} \cdot \frac{\frac{u(c_{\text{stw}}(\epsilon_{\text{stw}})) - u(c(\epsilon_{\text{stw}}))}{c_{\text{stw}}(\epsilon_{\text{stw}}) - c(\epsilon_{\text{stw}})} - u'(c^w)}{u'(c^w)} \cdot [c_{\text{stw}}(\epsilon_{\text{stw}}) - c(\epsilon_{\text{stw}})]}_{\text{additional insurance}}$$

then the candidate is $\epsilon_{stw} = \xi^c_s$. Otherwise, the above inequality holds as an equality and pins down $\epsilon_{stw} \geq \xi^c_s$.

PROOF: See Section 3.F.5 in the appendix.

To explain Proposition 4, one intuitive additional result is crucial, namely that constrained firms will always separate at even higher productivity levels than their unconstrained counterparts. The following lemma confirms this:

Lemma 1

$$\epsilon_s^c < \epsilon_s^u \quad \text{and} \quad \xi_s^c < \xi_s^u$$

PROOF: See Section 3.G in the appendix.

This presents the planner with a trade-off. Remember that any mass of firms on short-time work introduces socially sub-optimal distortion of working hours into the economy. This makes it socially costly to allow firms that would not have separated even without short-time work benefits to access short-time work.

When choosing ϵ_{stw}, the planner can choose to set the threshold leniently enough to allow constrained firms that would otherwise separate onto short-time work - i.e to $\epsilon_{\text{stw}} \geq \xi_s^c$. In this case, however, there will be a mass of unconstrained firms that can access short-time work, even though it would not separate even without it, introducing costly working hours distortions.

Alternatively, the planner can choose to ignore constrained firms and set the eligibility threshold to $\epsilon_{\text{stw}} \in [\epsilon_s^u, \xi_s^c)$ where unconstrained firms separate without access to short-time work. In this case, some constrained firms are cut off from the STW system even though they are in even more need of support than the unconstrained firms, not because of their productivity draw, but because of financial constraints that force them into separation before they become eligible for short-time work.

Proposition 4 reflects this problem. First, let us consider case 3. In this case, the eligibility condition is so loose that all firms in need of support can access short-time work. To achieve this $\epsilon_{\text{stw}} = \xi_s^c$ is sufficient. The question remains whether an even looser eligibility condition can be optimal. Because short-time work can provide income insurance to workers in financially constrained firms even before the match reaches the separation margin, it may be socially efficient to adopt a more generous eligibility condition. However, this effect could be outweighed by the additional working hours distortion introduced to both constrained and unconstrained firms, in which case the threshold remains at ξ_s^c. Our numerical results presented in Section 3.4.2, indeed, show that quantitatively, the distortion effect dominates and that $\epsilon_{\text{stw}} = \xi_s^c$ will be optimal.

Case 2 considers the case where the eligibility condition is set so strict that some of the firms with financial constraints are excluded from the STW system, even though they would be in need of support. At the same time all unconstrained firms that are in need of support can enter the STW system. The Ramsey planner must trade off the additional hours distortion in excess unconstrained firms against the social gains from protecting constrained firms. This trade-off is reflected by Inequality 3.3.3 and 3.3.4 and can go either way, depending on the value of p. In case the distortion effect outweighs the benefits of rescuing additional firms with financial constraints[5], the planner will choose to ignore constrained firms and choose the eligibility thresholds ξ_s^u. Only if 3.3.3 holds as an equality and the two effects are in the balance could ϵ_{stw} be between the two separation thresholds.

In case 1, the planner excludes all constrained firms. As argued by Stiepelmann, 2024 $\epsilon_{\text{stw}} < \xi_s^u$ cannot be optimal as this would exclude the most productive firms in need of support from short-time work. Increasing the threshold cannot be optimal either, as this would not impede additional separations but introduce additional distortions in working hours.

Proposition 5 The optimal short-time work subsidy τ_{stw} is pinned down by the first order condition

$$\begin{aligned}\bar{\tau}_{\text{stw}} = \; & \underbrace{\frac{\lambda}{f} b}_{\text{Fiscal Ext.}} - \underbrace{\frac{n}{\varphi(p)} \cdot \frac{\partial \Omega}{\partial \tau_{\text{stw}}}}_{\text{Distortion}} \\ & + \underbrace{\frac{n^c(p)}{\varphi(p)} \cdot \frac{1}{1-\rho^c} \cdot \int_{\epsilon_s^c}^{\max\{\epsilon_s^c, \epsilon_{\text{stw}}\}} \left(\frac{u'(c_{\text{stw}}(\epsilon) - u'(c^w))}{u'(c^w)} \right) (\bar{h} - h_{\text{stw}}(\epsilon))\, dG(\epsilon)}_{\text{Insurance}} \\ & + \underbrace{BE_{\text{stw},3}}_{\text{Bargaining Effect}}\end{aligned}$$

with

$$\bar{\tau}_{\text{stw}} = \frac{\varphi^u(p)}{\varphi(p)} \tau_{\text{stw}} (\bar{h} - h_{\text{stw}}(\epsilon_s^u)) + \frac{\varphi^c(p)}{\varphi(p)} \tau_{\text{stw}} (\bar{h} - h_{\text{stw}}(\epsilon_s^c))$$

where $BE_{\text{stw, 3}}$ captures general equilibrium effects through wage setting in the generalized Nash bargaining process. $\varphi(p)$, $\varphi^c(p)$ and $\varphi^u(p)$ are weight-terms with $\varphi(p) = \varphi^u(p) + \varphi^c(p)$ that are explicitly stated in Section 3.F.5 in the appendix.

5. Here $L_s^c(\epsilon)$ denotes the social value of a financially constrained match with productivity level ϵ.

PROOF: See Section 3.F.5 in the appendix.

Proposition 5 reveals that the Ramsey planner sets the average net subsidy $\bar{\tau}_{\text{stw}}$ to reflect three forces (abstracting from the general equilibrium bargaining effect). The first part of the sum is exactly equal to the fiscal externality of unemployment insurance on separations, i.e. the cost that the marginal worker who separates into unemployment imposes on the UI system. The second term reflects the fact that higher subsidies enable constrained firms to offer more insurance to workers to the extent that they are eligible for short-time work. This increases the optimal level of $\bar{\tau}_{\text{stw}}$. However, larger short-time work benefits will also lead to larger working hours distortions, in turn lowering the optimal level of $\bar{\tau}_{\text{stw}}$. τ_{stw} is then determined by distributing the load of $\bar{\tau}$ on constrained and unconstrained firms, taking into account marginal welfare gains.

It is noteworthy that the Ramsey planner sets the *average net* subsidy $\bar{\tau}_{\text{stw}}$ optimally and then adjusts the actual benefits τ_{stw} to achieve this optimal level. Since constrained firms have higher separation thresholds ($\epsilon_s^c > \epsilon_s^u$), it will hold that $\bar{h} - h_{\text{stw}}(\epsilon_s^c) < \bar{h} - h_{\text{stw}}(\epsilon_s^u)$. Because $\varphi^u(p)$ is decreasing in p and $\varphi^c(p)$ is increasing in p, this means that holding ϵ_{stw} fixed, τ_{stw} will be increasing in p.

Having disentangled how the Ramsey planner uses lay-off taxes and short-time work to counteract the fiscal externalities created by unemployment insurance, it is now time to determine which is the superior policy tool. From a theoretical point of view, this cannot be definitively determined. While lay-off taxes do not introduce any inefficiencies into the economy and can in fact induce the efficient number of separations in unconstrained firms, they cannot counteract fiscal externalities on constrained firms. Short-time work, on the other hand, can act on constrained firms as well. This, though, comes at the price of distorted working hours. To determine which is more effective in practice, we focus on the quantitative analysis of the model in the remainder of the paper.

Corollary 2 *If there are no constrained firms in the economy (i.e., $p = 0$), the optimal lay-off tax increases welfare more than the optimal short-time work scheme.*

PROOF: This follows from the fact that lay-off taxes introduce no distortions but can fully eliminate inefficient separations in unconstrained firms. All inefficiencies that short-time work can act on but lay-off taxes cannot, drop out when $p = 0$.

It is also clear that short-time work is the superior policy instrument if all firms are constrained.

Corollary 3 *If there are no unconstrained firms in the economy (i.e., $p = 1$), the optimal short-time work scheme increases welfare more than the optimal lay-off tax.*

PROOF: This follows from the fact that the only inefficiency that short-time work can act on, i.e., inefficient separations in unconstrained firms, cannot occur anymore. Short-time work can still act against inefficient separations in constrained firms and provide worker insurance.

In all cases with $p \in (0, 1)$, it is ambiguous which of the two policy regimes is superior without quantitative analysis, to which we turn next. It is clear, however, that as p increases lay-off taxes gradually lose their advantage over short-time work.

3.4 Quantitative Results

3.4.1 Calibration

As the utility function, we choose $u(c) = \log(c)$. We calibrate the model without either a lay-off tax or short-time work to the U.S. economy at monthly frequency. In order to evaluate the implications of short-time work and lay-off taxes in the current U.S. economy, we fix the unemployment insurance system and target the wage replacement estimate reported by Engen and Gruber, 2001 for the U.S. Up to small numerical imprecisions, this can be done by reproducing the targeted moments exactly by solving a system of equations stated in Section 3.I in the appendix. Table 3.4.1 gives an overview of our selected targets.

Target	Description	Value
q	Vacancy-filling rate	0.3381
f	Job-finding rate	0.40
ρ	Separation rate	0.1
b^{rep}	UI replacement rate	0.45

Table 3.4.1. Model Calibration Targets

Table 3.4.2. Model Parameters: Calibrated and Set Values

Parameter	Description	Value
Calibrated Parameters		
$\bar{m}$	Matching efficiency parameter	0.3832
c_f	Cost shock parameter	3.3276
k_v	Vacancy posting cost	0.3315
b	Unemployment benefit level	0.9972
$\bar{h}$	Regular working hours level	1.1616
A	Total Factor Productivity (TFP)	1.8
Set Parameters		
λ	Shock arrival probability	0.3
α	Labor elasticity in production function	0.65
μ	Location of productivity shocks	0.094
σ	Scale of productivity shocks	0.12
ψ	Disutility of labor parameter	1.5
γ	Elasticity of matching w.r.t. unemployment	0.65
η	Worker bargaining power	0.65

Notes: The set parameters are based on the following sources: α: Christoffel and Linzert, 2010. ψ: targeting Frisch elasticity of 0.65 as in Domeij and Floden, 2006. σ: Krause and Lubik, 2007. μ: normalizing wage to 1. γ: standard value. η: Hosios condition (Hosios, 1990), reasonable set of parameters in Petrongolo and Pissarides, 2001.

The first part of Table 3.4.2 shows the calibrated parameters set to achieve the targets. As shown by Costain and Reiter, 2008, large-surplus calibrations are needed not to overestimate the elasticity of worker flows to policy changes. We therefore set TFP A to 1.8.

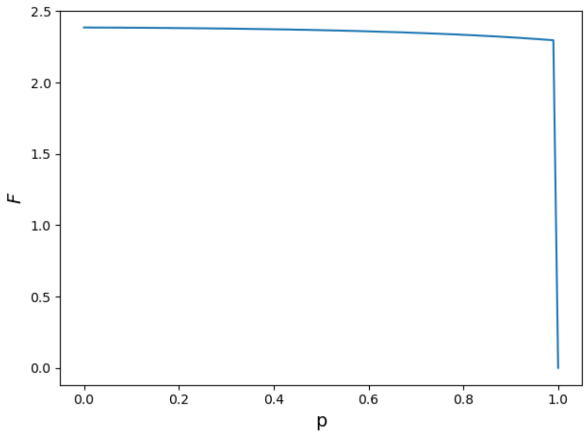
The remaining parameters are set to standard values shown in the second part of Table 3.4.2. We set $\lambda = 0.3$, so that a firm will remain at any drawn productivity level for $1/0.3 = 3.33$ months on average.

3.4.2 Results

To calculate the Ramsey optimal policy parameters, we maximize the welfare function (Equation 3.3.1) numerically. The details are described in Section 3.H in the appendix.

Figure 3.4.1 shows the results for the optimal lay-off tax F across $p \in [0, 1]$. The result is what its FOC (Equation 3.3.2) lets us expect. The planner exactly offsets the fiscal externality on unconstrained firms $\frac{\lambda}{f} \cdot b$. Since this does not depend on p in any way, the level of the optimal lay-off tax F hardly changes across p. The slight downward slope is explained by general-equilibrium feedback effects

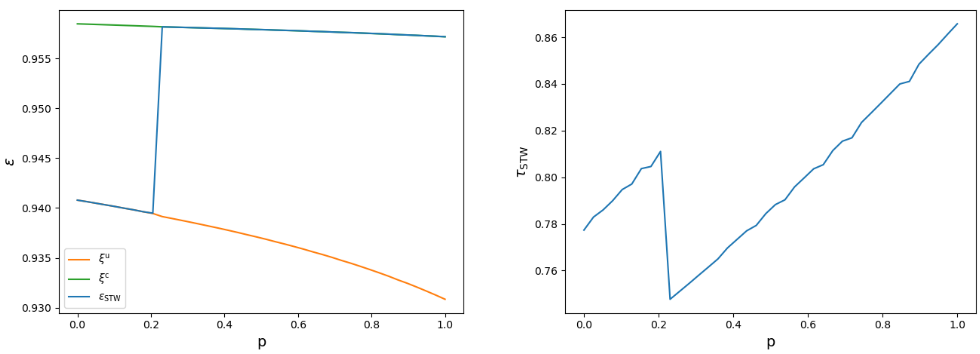
Figure 3.4.1. Optimal Lay-off Tax



Notes: Optimal lay-off tax F set by the Ramsey planner across $p \in [0, 1]$.

in $BE_{\text{lt}}(b)$. At $p = 1$, lay-off taxes lose their bite completely as they can no longer target any inefficiencies, and in fact, no firm will have to pay them. To illustrate this, we show the optimal lay-off tax to be $F = 0$. In fact, any lay-off tax $F \in \mathbb{R}_+$ is optimal.

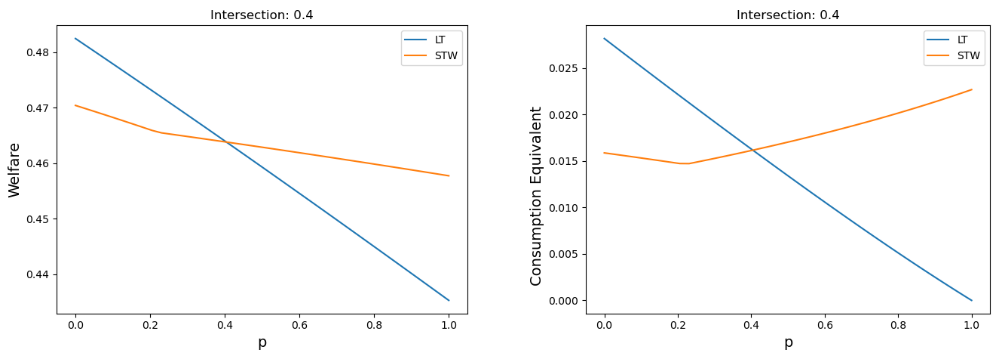
Figure 3.4.2. Optimal Short-time Work



(a) Optimal ε_{stw} **(b)** Optimal τ_{stw}

Notes: Optimal ε_{stw} and τ_{stw} set by the Ramsey planner across $p \in [0, 1]$. ξ_s^u is the separation threshold of an unconstrained firm without access to short-time work. ξ_s^c is the separation threshold of a constrained firm without access to short-time work.

Figure 3.4.2 shows the Ramsey optimal short-time work parameters across p. The way that the Ramsey planner sets the eligibility threshold reflects the trade-off described in Section 3.3.3 under Proposition 4: While there are only a few constrained firms in the economy, it is very costly to protect all firms, including constrained ones, from premature separations. Setting the eligibility constraint high enough to allow all constraint firms onto short-time work before they hit their separation threshold ξ_s^c will allow many unconstrained firms onto short-time work and collect windfall profits, even though they would not separate without it. As discussed, this introduces inefficient working hours distortions, and as long as only a few firms are constrained, these distortions outweigh the benefit of protecting these few constrained firms. The optimal eligibility threshold is therefore ξ_s^u. However, as the mass of constrained firms grows along the x-axis of Figure 3.4.2, the benefits from protecting constrained firms eventually outweigh the added distortion from protecting excess unconstrained firms. At this point ($p = 0.205$), the eligibility condition jumps to the threshold ξ_s^c at which constrained firms without short-time work separate.



(a) Welfare

(b) Consumption Equivalent

Figure 3.4.3. Welfare Comparison

Notes: Plot (a) on the left shows welfare across $p \in [0, 1]$ of the economy with the optimal lay-off tax and the optimal short-time work regimes, respectively. Plot (b) on the right shows improvements of optimal short-time work and optimal lay-off taxes over the economy without either of these two policies in terms of consumption equivalent variation.

The optimal generosity parameter τ_{stw} is increasing in p except at the kink at $p = 0.205$. This is because, as discussed in the context of Proposition 5 in Section 3.3.3, the planner targets the *average* net subsidy. As constrained firms have less margin to adjust their hours, this means that τ_{stw} has to rise in p as long as ϵ_{stw} does not change much. However, when ϵ_{stw} jumps to ξ_s^c, many more firms are suddenly on short-time work, introducing greater hours distortion to the economy. To counteract that distortion, the planner reduces the τ_{stw} as the more lenient eligibility threshold is introduced.

Having established how the Ramsey planner sets both lay-off taxes and short-time work, it is now time to discuss which is the superior policy tool. Figure 3.4.2 shows the respective welfare induced by the optimal lay-off tax and the optimal short-time work schemes across p on the left. Unsurprisingly, welfare is decreasing in p in both cases as higher shares of firms become constrained. As expected, when there are no or only a few constrained firms, the lay-off tax is the superior policy tool. Crucially, however, the two curves have different slopes and intersect at $p = 0.40$. The interpretation is simple: If more than 40% of firms are financially constrained, short-time work becomes the superior policy instrument; for smaller shares of constrained firms, lay-off taxes are the better tool.

For better interpretation, Figure 3.4.3 shows consumption equivalent variation, i.e., by what share consumption would need to be increased in the economy without either short-time work or lay-off taxes to induce the same welfare level as the economy with the respective optimal policy, on the right. To calculate consumption equivalent variation Δ^{policy} we solve

$$u(c_0 \cdot (1 + \Delta^{\text{policy}})) \; = \; u(c_{\text{policy}})$$

where $c_0 = u^{-1}(W(0))$ and $c_{\text{policy}} = u^{-1}(W(\pi^\star))$. $W(0)$ is welfare without either short-time work or lay-off taxes and $W(\pi^\star)$ is welfare under the respective optimal policy. While welfare improvements in terms of consumption equivalent variation intersect at the exact same point as raw welfare functions, they carry two additional messages. Firstly, the resulting improvements of the respective superior policy scheme are substantial at every p, ranging around 2%. Secondly, as p increases, short-time work becomes not only preferable to lay-off taxes, but improvements in terms of consumption equivalent variation increase. This stems from the effect that short-time work helps firms provide insurance, which becomes increasingly important for higher p.

3.5 Conclusion

We set out by asking the simple question, which of the two policy tools that policy makers in many countries already employ - lay-off taxes or short-time work - is better at counteracting the fiscal externalities of unemployment insurance. While existing literature emphasizes desirable properties of lay-off taxes, we show that their effectiveness is highly sensitive to the assumption that firms are unconstrained in their ability to smooth shocks.

By introducing firm-level financial constraints into a rich yet tractable Diamond-Mortensen-Pissaridis search-and-matching framework (DMP) with endogenous separations, risk-averse workers, and flexible working hours, we show that lay-off taxes struggle with counteracting these externalities when firms are constrained. Short-time work schemes, on the other hand, can act on constrained firms but have the disadvantage of introducing new inefficiencies into the economy through distorting working hours.

Our theoretical analysis, grounded in a Ramsey policy approach, delivers closed-form expressions for optimal lay-off taxes and STW subsidies as a function of the share of financially constrained firms. We show that lay-off taxes are the superior policy instrument with no or only a few financially constrained firms, but that short-time work benefits are better if sufficient firms are financially constrained. Quantitatively, we find that when 40% or more of firms are financially constrained, short-time work dominates lay-off taxes in welfare terms.

Policy recommendations must therefore depend on the prevalence of financial constraints in an economy. Anecdotally, it seems highly plausible that in the U.S., where a lay-off tax is implemented through experience-rated unemployment insurance, financial constraints play a smaller role than in countries like Germany or France with less developed financial sectors and smaller firms where short-time work is widely used. Empirically determining the exact extent of financial constraints in different countries and the implications for optimal policy is therefore key.

Another interesting way forward is to explore how our results change over the business cycle. If the extent to which firms are borrowing constrained or the share of constrained firms changes in downturns, there could be a case for relying more on short-time work during downturns and more on lay-off taxes during good times.

We view this as a fruitful ground for future research.

References

Baily, Martin Neil. 1978. "Some aspects of optimal unemployment insurance." *Journal of Public Economics* 10 (3): 379–402. [226]

Balleer, Almut, Britta Gehrke, Wolfgang Lechthaler, and Christian Merkl. 2016. "Does short-time work save jobs? A business cycle analysis." *European Economic Review* 84 (C): 99–122. [226]

Blanchard, Olivier J., and Jean Tirole. 2008. "The Joint Design of Unemployment Insurance and Employment Protection: A First Pass." *Journal of the European Economic Association* 6 (1): 45–77. [223, 226, 243]

Braun, Helge, and Björn Brügemann. 2017. "Welfare Effects of Short-Time Compensation." Tinbergen Institute Discussion Papers 17-010/VI. Tinbergen Institute. [226]

Burdett, Kenneth, and Randall Wright. 1989. "Unemployment Insurance and Short-Time Compensation: The Effects on Layoffs, Hours per Worker, and Wages." *Journal of Political Economy* 97 (6): 1479–96. [226]

Cahuc, Pierre. 2024. "The Micro and Macro Economics of Short-Time Work." IZA Discussion Papers 17111. Institute of Labor Economics (IZA). [227]

Cahuc, Pierre, Francis Kramarz, and Sandra Nevoux. 2021. "The Heterogeneous Impact of Short-Time Work: From Saved Jobs to Windfall Effects." IZA Discussion Papers 14381. Institute of Labor Economics (IZA). [226]

Cahuc, Pierre, and André Zylberberg. 2008. "Optimum income taxation and layoff taxes." *Journal of Public Economics* 92 (10-11): 2003–19. [223, 226, 243]

Cascaldi-Garcia, Danilo, Marija Vukoti, and Sarah Zubairy. 2023. "Innovation During Challenging Times." The Warwick Economics Research Paper Series (TWERPS) 1475. University of Warwick, Department of Economics. [227]

Chetty, Raj. 2006. "A general formula for the optimal level of social insurance." *Journal of Public Economics* 90 (10-11): 1879–901. [226]

Christoffel, Kai, and Tobias Linzert. 2010. "The Role of Real Wage Rigidity and Labor Market Frictions for Inflation Persistence." *Journal of Money, Credit and Banking* 42 (7): 1435–46. [252]

Cooper, Russell, Moritz Meyer, and Immo Schott. 2017. "The Employment and Output Effects of Short-Time Work in Germany." NBER Working Papers 23688. National Bureau of Economic Research, Inc. [226]

Costain, James S., and Michael Reiter. 2008. "Business cycles, unemployment insurance, and the calibration of matching models." *Journal of Economic Dynamics and Control* 32 (4): 1120–55. [252]

Domeij, David, and Martin Floden. 2006. "The Labor-Supply Elasticity and Borrowing Constraints: Why Estimates are Biased." *Review of Economic Dynamics* 9 (2): 242–62. [252]

Drechsel, Thomas. 2023. "Earnings-Based Borrowing Constraints and Macroeconomic Fluctuations." *American Economic Journal: Macroeconomics* 15 (2): 1–34. [227]

Duggan, Mark, Audrey Guo, and Andrew C. Johnston. 2023. "Experience Rating as an Automatic Stabilizer." *Tax Policy and the Economy* 37: 109–33. eprint: https://doi.org/10.1086/724354. [226]

Duygan-Bump, Burcu, Alexey Levkov, and Judit Montoriol-Garriga. 2015. "Financing constraints and unemployment: Evidence from the Great Recession." *Journal of Monetary Economics* 75: 89–105. [227]

Engen, Eric M., and Jonathan Gruber. 2001. "Unemployment insurance and precautionary saving." *Journal of Monetary Economics* 47 (3): 545–79. [251]

Fonseca, Julia, and Bernardus Van Doornik. 2022. "Financial development and labor market outcomes: Evidence from Brazil." *Journal of Financial Economics* 143 (1): 550–68. [227]

Giupponi, Giulia, and Camille Landais. 2022. "Subsidizing Labour Hoarding in Recessions: The Employment and Welfare Effects of Short-time Work." *Review of Economic Studies* 90 (4): 1963–2005. eprint: https://academic.oup.com/restud/article-pdf/90/4/1963/51440201/rdac069.pdf. [226]

Giupponi, Giulia, Camille Landais, and Alice Lapeyre. 2021. "Should We Insure Workers or Jobs During Recessions?" Report. London, Centre for Economic Policy Research. [226]

Hopenhayn, Hugo, and Juan Pablo Nicolini. 1997. "Optimal Unemployment Insurance." *Journal of Political Economy* 105 (2): 412–38. [226]

Hosios, Arthur J. 1990. "On the Efficiency of Matching and Related Models of Search and Unemployment." *Review of Economic Studies* 57 (2): 279–98. [252]

Johnston, Andrew C. 2021. "Unemployment Insurance Taxes and Labor Demand: Quasi-Experimental Evidence from Administrative Data." *American Economic Journal: Economic Policy* 13 (1): 266–93. [226]

Jung, Philip, and Keith Kuester. 2015. "Optimal Labor-Market Policy in Recessions." *American Economic Journal: Macroeconomics* 7 (2): 124–56. [226]

Krause, Michael U., and Thomas A. Lubik. 2007. "The (ir)relevance of real wage rigidity in the New Keynesian model with search frictions." *Journal of Monetary Economics* 54 (3): 706–27. [252]

Kroft, Kory, and Matthew J. Notowidigdo. 2016. "Should Unemployment Insurance Vary with the Unemployment Rate? Theory and Evidence." *Review of Economic Studies* 83 (3): 1092–124. eprint: https://academic.oup.com/restud/article-pdf/83/3/1092/11126211/rdw009.pdf. [226]

Landais, Camille, Pascal Michaillat, and Emmanuel Saez. 2018. "A Macroeconomic Approach to Optimal Unemployment Insurance: Theory." *American Economic Journal: Economic Policy* 10 (2): 152–81. [226]

Michau, Jean-Baptiste. 2015. "Optimal labor market policy with search frictions and risk-averse workers." *Labour Economics* 35 (C): 93–107. [226]

Mortensen, Dale T., and Christopher A. Pissarides. 1993. "Job Creation and Job Destruction in the Theory of Unemployment." CEP Discussion Papers dp0110. Centre for Economic Performance, LSE. [224]

Mortensen, Dale T., and Christopher A. Pissarides. 1994. "Job Creation and Job Destruction in the Theory of Unemployment." *Review of Economic Studies* 61 (3): 397–415. [224]

Ottonello, Pablo, and Thomas Winberry. 2020. "FINANCIAL HETEROGENEITY AND THE INVESTMENT CHANNEL OF MONETARY POLICY." *Econometrica* 88 (6): pp. 2473–2502. [227]

Petrongolo, Barbara, and Christopher A. Pissarides. 2001. "Looking into the Black Box: A Survey of the Matching Function." *Journal of Economic Literature* 39 (2): 390–431. [252]

Postel-Vinay, Fabien, and Hélène Turon. 2011. "Severance Packages." IZA Discussion Papers 6023. Institute of Labor Economics (IZA). [226, 231]

Ratner, David. 2013. "Unemployment insurance experience rating and labor market dynamics." [226]

Shavell, Steven, and Laurence Weiss. 1979. "The Optimal Payment of Unemployment Insurance Benefits over Time." *Journal of Political Economy* 87 (6): 1347–62. [226]

Stiepelmann, Gero. 2024. “Optimal Short-Time Work Policy in Recessions.” Working Papers. University of Bonn. [223, 227, 239, 249]

Teichgräber, Julian, Simon Žužek, and Jannik Hensel. 2022. “Optimal short-time work: screening for jobs at risk.” ECON - Working Papers 402. Department of Economics - University of Zurich. [233]

Appendix 3.A Table of Model Parameters, Variables and Functions

Symbol	Description
n	Mass of employed workers
$u = 1 - n$	Mass of unemployed workers
v	Mass of posted vacancies
$\theta = \frac{v}{1-n}$	Labor market tightness
f	Job-finding rate
q	Vacancy-filling rate
ρ	Separation rate
ρ^u, ρ^c	Separation rates (unconstrained/ constrained)
ρ^F	Separation with lay-off tax
ρ^L	Separation via bankruptcy
p	Probability firm is constrained
λ	Probability of productivity shock
$\tilde{\varepsilon}$	Productivity shock (lognormal)
$G(\varepsilon), g(\varepsilon)$	CDF and PDF of productivity shocks
$\bar{m}$, γ	Matching function parameters
A	Total factor productivity
α	Output elasticity (hours)
ψ	Inverse Frisch elasticity
v_f	Mass of firm owners
$\bar{h}$	Reference hours under STW
D	Eligibility threshold for STW
ε_{stw}	Productivity threshold for STW
b	Unemployment benefits
F	Lay-off tax
L	Liquidation cost
τ	Lump-sum tax per shock
τ_{stw}	STW subsidy per hour gap
k_v	Cost of posting a vacancy
η	Bargaining power of workers

Table 3.A.1. Model Parameters, Variables and Functions Part 1

Symbol	Description
u/ c	unconstrained/ constrained
$m(v, n)$	Matching function
$\phi(h)$	Disutility of labor: $\frac{h^{1+\psi}}{1+\psi}$
$u(c)$	Worker utility from consumption net of disutility
$y(\varepsilon, h)$	Output
$z(\varepsilon, h)$	Output net of disutility (cons.-eq. units)
$z(\varepsilon)$	Shorthand for $z(\varepsilon, h(\varepsilon))$
$z_{\text{stw}}(\varepsilon)$	Output net of disutility under STW hours
u^c	Average utility of a worker at a constrained firm
e^c	Average total wage net of disutility paid at a constrained firm
$\Omega(\varepsilon)$	Welfare Costs STW in match at productivity ε
Ω^u	Average welfare loss STW in unconstrained firms
Ω^c	Average welfare loss STW in constrained firms
Ω	Average welfare loss net of disutility of all firms
$h(\varepsilon)$	Hours worked (non-STW)
$h_{\text{stw}}(\varepsilon)$	Hours worked under STW
c^w	Constant consumption-equivalent wage
$c^w(\varepsilon)$	Constrained consumption wage
$c^w_{\text{stw}}(\varepsilon)$	STW-constrained wage
$w^u(\varepsilon), w^c(\varepsilon)$	(Monthly) Wage functions (u/ c, no STW)
$w^u_{\text{stw}}(\varepsilon), w^c_{\text{stw}}(\varepsilon)$	(Monthly) Wage (u/ c) under STW
$V^u(\varepsilon), V^c(\varepsilon)$	Worker value (u/ c, no STW)
$V^u_{\text{stw}}(\varepsilon), V^c_{\text{stw}}(\varepsilon)$	Worker value (STW)
$\bar{V}$	Expected worker value
U	Value of unemployment
$J^u(\varepsilon), J^c(\varepsilon)$	Firm value (u/ c, no STW)
$J^u_{\text{stw}}(\varepsilon), J^c_{\text{stw}}(\varepsilon)$	Firm value (u/ c) under STW
$\bar{J}$	Expected firm value
$\varepsilon^u_s, \varepsilon^c_s$	Separation thresholds (u/ c, under STW)
ξ^u_s, ξ^c_s	Separation thresholds (u/ c, no STW)
ε^p	Productivity where constraint binds
n^s	Mass of matches receiving new shock
n^u, n^c	Mass of workers at u/c firms

Table 3.A.2. Model Parameters, Variables and Functions Part 2

Appendix 3.B Equilibrium Lay-off Tax

3.B.1 Bargaining

Nash-Bargaining Problem:

$$\max_{w^u(\epsilon),w^c(\epsilon),h^u(\epsilon),h^c(\epsilon),\epsilon_s^u,\epsilon_s^c} \bar{J}^{(1-\eta)}\left(\bar{V}-U\right)^{\eta}$$

subject to

1.Financial constraint of financially constrained firms:

$$y(\epsilon, h(\epsilon)) - w^c(\epsilon) + \lambda \cdot J^c(\epsilon) + (1-\lambda)\cdot \bar{J} \;\geq\; 0$$

2.Commitment problem (Worker):

$$V^c(\epsilon) \;>\; U, \quad V^u(\epsilon) \;>\; 0$$

First of all, note that we can rewrite the financial constraint as

$$z(\epsilon) - c(\epsilon) + \lambda J^c(\epsilon) + (1-\lambda)\bar{J} \;>\; 0$$

Second, we can integrate the commitment problem of the worker into the value functions for the firm and worker surplus. For $i \in \{c,u\}$, we can rearrange the condition as:

$$u(c^i(\epsilon)) - u(b) + (\lambda - f)\cdot(\bar{V}-U) \;\geq\; 0$$

Let $\epsilon_s^{i,\text{cp}}$ denote the threshold at which the worker wants to separate on his own terms:

$$u(c(\epsilon_s^{i,\text{cp}})) - u(b) + (\lambda - f)\cdot(\bar{V}-U) \;=\; 0$$

We can rewrite the value functions of the firm and the surplus of the worker as:

$$
\begin{aligned}
\bar{J}^B \;=\; & -\bar{\tau}^c + (1-p)\cdot\int_{\max\{\epsilon_s^{u,B},\epsilon_s^{u,\text{cp}}\}}^{\infty}\frac{1}{\lambda}\cdot\left(z(\epsilon)-c^w(\epsilon)+\lambda\bar{J}\right)dG(\epsilon)\\
& + (1-p)\cdot G(\epsilon_s^{u,B})\cdot\mathbb{1}(\epsilon_s^{u,B} \;>\; \epsilon_s^{u,\text{cp}})\cdot F\\
& + p\cdot\int_{\max\{\epsilon_s^{u,B},\epsilon_s^{u,\text{cp}}\}}^{\infty}\frac{1}{\lambda}\cdot\left(z(\epsilon)-c^c(\epsilon)+\lambda\bar{J}\right)dG(\epsilon)\\
& + p\cdot G(\epsilon_s^{u,B})\cdot\mathbb{1}(\epsilon_s^{u,B} \;>\; \epsilon_s^{u,\text{cp}})\cdot L\\
(\bar{V}-U)^B \;=\; & (1-p)\cdot\int_{\max\{\epsilon_s^{u,B},\epsilon_s^{u,\text{cp}}\}}^{\infty}\frac{1}{\lambda}\cdot\left(u(c^u(\epsilon))-u(b)+\lambda(\bar{V}-U)\right)dG(\epsilon)\\
& + p\cdot\int_{\max\{\epsilon_s^{u,B},\epsilon_s^{u,\text{cp}}\}}^{\infty}\frac{1}{\lambda}\cdot\left(u(c^c(\epsilon))-u(b)+\lambda(\bar{V}-U)\right)dG(\epsilon)
\end{aligned}
$$

Note that $\epsilon_s^{i,B}$ denotes the separation threshold in the bargaining outcome. In the following, we rule out the case where the consumption of the worker can be set so low that the worker quits on their own terms to avoid the lay-off tax. This means that ϵ_s^{cp} is seen as exogenous to the bargaining participants. Several lawsuits in the US show that this might, in reality, actually be a problem of lay-off taxes that we abstract from.

Therefore, lay-off taxes in this model can be interpreted as marginally effective! Further, note that financially constrained firms cannot pay lay-off taxes and go bankrupt when they are constrained and have to pay the lay-off tax. In this case, we assume some liquidation costs $0 \leq L \leq F$, borne by the firm owners. Finally, note that it is equivalent to maximize over $w^i(\epsilon)$ and $c^i(\epsilon)$ as

$$c^i(z) \;=\; w^i(\epsilon) - \phi^i(h(\epsilon)).$$

Set up Kuhn-Tucker Problem:

$$\max_{c^c(\epsilon),c^u(\epsilon),h^c(\epsilon),h^u(\epsilon),\epsilon_s^{c,B},\epsilon_s^{u,B}} \mathscr{B}$$

where

$$\mathscr{B} \;=\; \left(\bar{J}^B\right)^{\eta}\cdot\left(\bar{V}-U\right)^{1-\eta} - \int_0^{\infty}\mathscr{M}(\epsilon)\left[z(\epsilon)-c(\epsilon)+\lambda\cdot\bar{J}+(\lambda-f)\cdot\bar{J}(\epsilon)\right]d\epsilon$$

With Kuhn-Tucker conditions for a maximum:

$$
\begin{aligned}
&(I) \quad \mathscr{M}(\epsilon) \;\leq\; 0\\
&(II) \quad \mathscr{M}(\epsilon)\cdot\left[z(\epsilon)-c(\epsilon)+\lambda\cdot\bar{J}+(\lambda-f)\cdot\bar{J}(\epsilon)\right] \;=\; 0
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial \mathscr{B}}{\partial c^u(\epsilon)} &= -(1-\eta)\cdot(1-p)\cdot g(\epsilon)\cdot\left(\frac{\bar{V}-U}{\bar{J}}\right)^{\eta} \\
&\quad + \eta\cdot(1-p)\cdot g(\epsilon)\cdot u'\big(c^u(\epsilon)\big)\cdot\left(\frac{\bar{J}}{\bar{V}-U}\right)^{1-\eta} \\
&= 0
\end{aligned}
$$

$$
\Longleftrightarrow \quad \frac{1-\eta}{\eta}\cdot\frac{\bar{V}-U}{\bar{J}} = u'\big(c^u(\epsilon)\big) \quad \Rightarrow \quad u'\big(c^u(\epsilon)\big) = u'(c^w)
$$

Note that the firm wants to perfectly insure workers against idiosyncratic productivity shocks as long as they are unconstrained.
Define the joint surplus as

$$
\bar{S} = \bar{J} + \frac{\bar{V}-U}{u'(c^w)}
$$

The resulting surplus splitting rule gives:

$$
J = (1-\eta)\cdot\bar{S}, \quad \frac{\bar{V}-U}{u'(c^u)} = \eta\cdot\bar{S}
$$

$$
\begin{aligned}
\frac{\partial \mathscr{B}}{\partial c^c(\epsilon)} &= -(1-p)\cdot g(\epsilon)\cdot(1-\eta)\cdot\left(\frac{\bar{V}-U}{\bar{S}}\right)^{n} \\
&\quad + (1-p)\cdot g(\epsilon)\cdot u'\big(c^u(\epsilon)\big)\cdot U\cdot\left(\frac{J}{\bar{V}-U}\right)^{1-\eta} \\
&\quad + \mathscr{M}(\epsilon) \overset{!}{=} 0
\end{aligned}
$$

Insert the surplus splitting rule:

$$
\begin{aligned}
\Rightarrow \quad & -(1-p)\cdot g(\epsilon)\cdot\left(\frac{\eta}{1-\eta}\cdot u'(c^w)\right)^{\eta} \\
& \quad + (1-p)\cdot g(\epsilon)\cdot u'\big(c^c(\epsilon)\big)\cdot\eta \\
& \quad + \mathscr{M}(\epsilon)\cdot\left(u'(c^w)\cdot\frac{1-\eta}{\eta}\right)^{1-\eta} = 0
\end{aligned}
$$

$$
\Leftrightarrow \; (1-p)\cdot g(\epsilon)\cdot\eta\cdot\big(u'(c^c(\epsilon)) - u'(c^w)\big) = -\mathscr{M}(\epsilon)\cdot\left(u'(c^u)\cdot\frac{\eta}{1-\eta}\right)^{1-\eta}
$$

Note that if $c^c(\epsilon) < c^w$, then the RHS of the equation is positive due to risk aversion. Thus, firms and workers gain joint surplus until $c^c(\epsilon) = c^w$.

If

$$
z(\epsilon) - c^w + \lambda\cdot\bar{J} + (1-\lambda)\cdot J^c(\epsilon) \geq 0
$$

then the constraint is non-binding and $\mathscr{M}(\epsilon) = 0$. ✓

If

$$z(\epsilon) - c^w + \lambda \cdot \bar{J} + (1-\lambda) \cdot J^c(\epsilon) < 0$$

then the constraint is binding and $\mathscr{M}(\epsilon) < 0$. ✓
Thus, it is optimal for the firm to pay as much as it can:

$$J^c(\epsilon) = z(\epsilon) - c^c(\epsilon) + \lambda \cdot \bar{J} + (1-\lambda) \cdot J^c(\epsilon) = 0$$

$$\iff \quad c^c(\epsilon) = z(\epsilon) + \lambda \cdot \bar{J}$$

Using the surplus splitting rule, the Nash-Bargaining problem can be simplified to

$$\max_{w^u(\epsilon), w^c(\epsilon), h^u(\epsilon), h^c(\epsilon), \epsilon_s^u, \epsilon_s^c} (1-\eta)^{(1-\eta)} \cdot (u'(c^w)\eta)^{\eta} \cdot \bar{S}$$

That is, we just have to maximize the joint surplus with the remaining contracts:

$$\begin{aligned}
\bar{S} &= -\bar{\tau}^c \\
&+ (1-p) \cdot \int_{\max\{\epsilon_s^{u,B}, \epsilon_s^{u,\mathrm{cp}}\}}^{\infty} \frac{1}{\lambda} \cdot \left(z(\epsilon) + \frac{u(c^w) - u(b)}{u'(c^w)} - c^w + (1 - \eta \cdot f) \cdot \bar{S} \right) dG(\epsilon) \\
&- (1-p) \cdot G(\epsilon_s^{u,B}) \cdot \mathbb{1}(\epsilon_s^{u,B} > \epsilon_s^{u,\mathrm{cp}}) \cdot F \\
&+ p \cdot \int_{\epsilon^p}^{\infty} \frac{1}{\lambda} \cdot \left(z(\epsilon) + \frac{u(c^w) - u(b)}{u'(c^w)} - c^w + (1 - \eta \cdot f) \cdot \bar{S} \right) dG(\epsilon) \\
&+ p \cdot \int_{\max\{\epsilon_s^{u,B}, \epsilon_s^{u,\mathrm{cp}}\}}^{\epsilon_p} \frac{1}{\lambda} \cdot \left(z(\epsilon) + \frac{u(c^c(\epsilon)) - u(b)}{u'(c^w)} - c^w + (1 - \eta \cdot f) \cdot \bar{S} \right) dG(\epsilon) \\
&- p \cdot G(\epsilon_s^{u,B}) \cdot \mathbb{1}(\epsilon_s^{u,B} > \epsilon_s^{u,\mathrm{cp}}) \cdot L
\end{aligned}$$

The FOC for working hours in unconstrained firms is:

$$\frac{\partial \mathscr{B}}{\partial h^u(\epsilon)} = (1-p) \cdot g(\epsilon) \cdot \left(\frac{\partial y(\epsilon, h^u(\epsilon))}{\partial h^u(\epsilon)} - \phi'(h^u(\epsilon)) \right) = 0$$

$$\iff \quad \frac{\partial y(\epsilon, h^u(\epsilon))}{\partial h^u(\epsilon)} = \phi'(h^u(\epsilon))$$

The FOC for working hours in constrained firms is:

$$\frac{\partial \mathscr{B}}{\partial h^c(\epsilon)} = p \cdot g(\epsilon) \cdot \left(\frac{\partial y(\epsilon, h^c(\epsilon))}{\partial h^c(\epsilon)} - \phi'(h^c(\epsilon)) \right) \cdot \frac{u'(c^c(\epsilon))}{u'(c^w)} = 0$$

$$\iff \quad \frac{\partial y(\epsilon, h^c(\epsilon))}{\partial h^c(\epsilon)} = \phi'(h^c(\epsilon))$$

Suppose that $\epsilon_s^{u^b} > \epsilon_s^{u^{cp}}$

$$\frac{\partial \mathcal{B}}{\partial \epsilon_s^{u,b}} = -(1-p) \cdot g(\epsilon_s^{u,b}) \cdot \left[\frac{1}{\lambda} \cdot \Big(z(\epsilon_s^{u,b}) + \frac{u(c^w) - u(b)}{u'(c^w)} - c^w + (\lambda - \eta f) \cdot \bar{S} \Big) + F \right] = 0$$

$$\iff \quad z(\epsilon_s^{u,b}) + \lambda \cdot F + \frac{u(c^w) - u(b)}{u'(c^w)} - c^w + (\lambda - \eta f) \cdot \bar{S} = 0$$

Insert free-entry Condition:

$$z(\epsilon_s^{u,b}) + \lambda \cdot F + \frac{u(c^w) - u(b)}{u'(c^w)} - c^w + (\lambda - \eta f) \cdot \left(\frac{k_v}{q} \right) = 0$$

Note that $\epsilon_s^{u^b} > \epsilon_s^{u^{cp}}$ must hold, as workers would never quit under the insurance constraint of the firm:

$$V(\epsilon) - U = u(c^w) - u(b) + (\lambda - f)(\bar{V} - U) > 0$$

It always gets positive surplus!
Suppose that $\epsilon_s^{c,B} \geq \epsilon_s^{u,cp}$

$$\frac{\partial \mathcal{B}}{\partial \epsilon_s^{c,B}} = -(1-p) \cdot g(\epsilon_s^{c,B}) \cdot \left[\frac{1}{\lambda} \cdot \Big(z(\epsilon_s^{c,B}) + \frac{u\big(c^c(\epsilon_s^{c,B})\big) - u(b)}{u'(c^w)} - c^c(\epsilon_s^{c,B}) + (\lambda - \eta f) \cdot \bar{S} \Big) + L \right] = 0$$

$$\iff \quad z(\epsilon_s^{c,B}) + \lambda \cdot L + \frac{u(c^c(\epsilon_s^{c,B})) - u(b)}{u'(c^w)} - c^c(\epsilon_s^{c,B}) + (\lambda - \eta f) \cdot \bar{S} = 0$$

Insert free-entry condition:

$$z(\epsilon_s^{c,B}) + \lambda \cdot L + \frac{u(c^c(\epsilon_s^{c,B})) - u(b)}{u'(c^w)} - c^c(\epsilon_s^{cp}) + \frac{\lambda - \eta f}{1 - \eta} \cdot \frac{k_v}{q} = 0$$

The participation constraint of workers can be written with free-entry condition as:

$$z(\epsilon_s^{c,cp}) + \frac{u(c^c(\epsilon_s^{c,cp})) - u(b)}{u'(c^w)} - c^c(\epsilon_s^{c,cp}) + \frac{(\lambda - \eta f)}{1 - \eta} \cdot \frac{k_v}{q} = 0$$

Note that to avoid liquidation costs, firms could like workers to stay within the firm for negative surplus values. Thus, the worker decides to leave the firm before the contractual separation threshold $\epsilon_s^{c,B} < \epsilon_s^{u,cp}$, avoiding lay-off taxes.

3.B.2 Job Creation and Wage Equation

The notation of this section follows Appendix 3.D for lay-off taxes.

(I)Value of an unconstrained firm after the idiosyncratic productivity shock has realized:

$$\begin{aligned} J^u(\epsilon) &= z(\epsilon) - c^w + \lambda \bar{J} + (1-\lambda) \cdot J^u(\epsilon) \\ \Leftrightarrow \quad J^u(\epsilon) &= \frac{1}{\lambda}\left(z(\epsilon) - c^w + \lambda \bar{J}\right) \end{aligned}$$

(II)Value of a constrained firm without binding constraints after the idiosyncratic productivity shock has realized:

$$\begin{aligned} J^c(\epsilon) &= z(\epsilon) - c^w + \lambda \bar{J} + (1-\lambda) \cdot J^c(\epsilon) \\ \Leftrightarrow \quad J^c(\epsilon) &= \frac{1}{\lambda}\left(z(\epsilon) - c^w + \lambda \bar{J}\right) \end{aligned}$$

(III)Value of a constrained firm with binding constraints after the idiosyncratic productivity shock has realized:

$$\begin{aligned} J^c(\epsilon) &= z(\epsilon) - c^w(\epsilon) + \lambda \bar{J} + (1-\lambda) \cdot J^c(\epsilon) \\ \Leftrightarrow \quad J^c(\epsilon) &= \frac{1}{\lambda}\left(z(\epsilon) - c^w(\epsilon) + \lambda \bar{J}\right) \end{aligned}$$

(IV)Value of a firm before the idiosyncratic productivity shock has realized:

$$\bar{J} = -\tau + (1-p) \cdot \int_{\epsilon_s^u}^{\infty} J^u(\epsilon)\, dG(\epsilon) + p \cdot \int_{\epsilon_s^c}^{\infty} J^c(\epsilon)\, dG(\epsilon) - (1-p) \cdot \rho^u \cdot F$$

Inserting (I)-(III) into (IV) gives:

$$\begin{aligned} \Leftrightarrow \quad \bar{J} = -\tau + \frac{1}{\lambda}\Big[&(1-\rho) \cdot z - (1-p) \cdot (1-\rho^u) \cdot c^w - p \cdot (1-\rho^c) \cdot e^c \\ &+ (1-p) \cdot \lambda \cdot \bar{J}\Big] - (1-p) \cdot \rho^u \cdot F \end{aligned}$$

(V)Value of an unconstrained worker after the idiosyncratic productivity shock has realized:

$$\begin{aligned} V^u(\epsilon) &= u(c^w) + \lambda \bar{V} + (1-\lambda) \cdot V^u(\epsilon) \\ \Leftrightarrow \quad V^u(\epsilon) &= \frac{1}{\lambda}\left(u(c^w) + \lambda \bar{V}\right) \end{aligned}$$

(VI)Value of a constrained worker without binding constraints after the idiosyncratic productivity shock has realized:

$$\begin{aligned} V^c(\epsilon) &= u(c^w) + \lambda \bar{V} + (1-\lambda) \cdot V^c(\epsilon) \\ \Leftrightarrow \quad V^c(\epsilon) &= \frac{1}{\lambda}\big(u(c^w) + \lambda \bar{V}\big) \end{aligned}$$

(VII)Value of a constrained worker with binding constraints after the idiosyncratic productivity shock has realized:

$$\begin{aligned} V^c(\epsilon) &= u(c^w(\epsilon)) + \lambda \bar{V} + (1-\lambda) \cdot V^c(\epsilon) \\ \Leftrightarrow \quad V^c(\epsilon) &= \frac{1}{\lambda}\big(u(c^w(\epsilon)) + \lambda \bar{V}\big) \end{aligned}$$

(VIII)Value of a worker before the idiosyncratic productivity shock has realized:

$$\bar{V} = (1-p) \cdot \int_{\epsilon_s^u}^{\infty} V^u(\epsilon)\, dG(\epsilon) + p \cdot \int_{\xi^c}^{\infty} V^c(\epsilon)\, dG(\epsilon) + \rho \cdot U$$

Inserting (V)-(VII) into (VIII) gives:

$$\Leftrightarrow \quad \bar{V} = \frac{1}{\lambda}\Big[(1-p)(1-\rho^u) \cdot u(c^w) + p \cdot (1-\rho^c) \cdot u^c + (1-\rho) \cdot \lambda \cdot \bar{V} + \rho \cdot U\Big]$$

(IX)Unemployment:

$$U = u(b) + f \cdot \bar{V} + (1-f) \cdot U$$

Next, we turn to calculating the joint surplus.

A)First, let us calculate the surplus of the worker after the idiosyncratic productivity shock has been realized:

$$\begin{aligned} V^i(\epsilon) - U &= u(c^i(\epsilon)) - u(b) + (\lambda - f) \cdot V^i(\epsilon) + (\lambda - f) \cdot \bar{V} - (\lambda - f) \cdot U \\ \Leftrightarrow \quad V^i(\epsilon) - U &= u(c^i(\epsilon)) - u(b) + (\lambda - f) \cdot (V^i(\epsilon) - U) + (\lambda - f) \cdot (\bar{V} - U) \\ \Leftrightarrow \quad V^i(\epsilon) - U &= \frac{1}{\lambda}\big(u(c^i(\epsilon)) - u(b) + (\lambda - f) \cdot (\bar{V} - U)\big) \end{aligned}$$

B)Next, we calculate the expected surplus of the worker, before the shocks have been realized:

$$\begin{aligned}
\bar{V} - U &= (1-\rho) \cdot \int_{\epsilon_s^u}^{\infty} \big(V^u(\epsilon) - U\big)\, dG(\epsilon) + p \cdot \int_{\epsilon_s^c}^{\infty} \big(V^c(\epsilon) - U\big)\, dG(\epsilon) \\
&= \frac{1}{\lambda}\Big[(1-\rho) \cdot (1-\rho^u) \cdot \big(u(c^u) - u(b)\big) \\
&\quad + p \cdot (1-\rho^c) \cdot \big(u(c^c) - u(b)\big) \\
&\quad + (1-\rho) \cdot (\lambda - f) \cdot (\bar{V} - U) \Big]
\end{aligned}$$

C)Next, we can calculate the joint surplus from the expected value of a worker for a firm and the expected surplus of the worker:

$$\begin{aligned}
S &= \bar{J} + \frac{\bar{V} - U}{u'(c^u)} \\
&= -\tau + \frac{1}{\lambda}\Big[(1-\rho) \cdot z \\
&\quad + (1-\rho) \cdot (1-\rho^u) \cdot \left(\frac{u(c^u) - u(b)}{u'(c^u)} - c^u \right) \\
&\quad + p \cdot (1-\rho^c) \cdot \left(\frac{u(c^c) - u(b)}{u'(c^u)} - e^c \right) \\
&\quad + (1-\rho) \cdot \lambda \cdot \bar{J} + (1-\rho)(\lambda - f) \cdot \frac{\bar{V} - U}{u'(c^u)} \Big] - (1-p) \cdot \rho^u \cdot F
\end{aligned}$$

Using the surplus splitting rule, we can rewrite the equation above as:

$$\begin{aligned}
S &= -\tau + \frac{1}{\lambda}\Big[(1-\rho) \cdot z \\
&\quad + (1-\rho) \cdot (1-\rho^u) \cdot \left(\frac{u(c^u) - u(b)}{u'(c^u)} - c^u \right) \\
&\quad + p \cdot (1-\rho^c) \cdot \left(\frac{u(c^c) - u(b)}{u'(c^u)} - e^c \right) \\
&\quad + (1-\rho) \cdot (\lambda - \eta f) \cdot S \Big] + (1-p) \cdot \rho^u \cdot F
\end{aligned}$$

Next, we want to replace the tax in the surplus equation. We can do this by using the budget constraint of the government:

$$\begin{aligned} n^s \cdot \tau &= n^s \cdot (1-p) \cdot \rho^u \cdot F + (1-n) \cdot b \\ \Leftrightarrow \qquad \tau &= (1-p) \cdot \rho^u \cdot F + \frac{1-n}{n^s} \cdot b \\ \text{Insert } n^s &= \frac{\lambda}{1-\rho} \cdot n \\ \Rightarrow \qquad \tau &= n^s \cdot (1-p) \cdot \rho^u \cdot F + \frac{1}{\lambda} \cdot (1-\rho) \cdot \frac{1-n}{n} \cdot b \end{aligned}$$

Inserting it into the equation for the joint surplus gives:

$$S = \frac{1}{\lambda}\Bigg[(1-\rho) \cdot \left(z - \frac{1-n}{n} \cdot b\right) + (1-p) \cdot (1-\rho^u) \cdot \left(\frac{u(c^w) - u(b)}{u'(c^w)} - c^w\right)$$

$$+ p \cdot (1-\rho^c) \cdot \left(\frac{u(c^c) - u(b)}{u'(c^w)} - e^c\right) + (1-\rho) \cdot (\lambda - \eta f) \cdot S\Bigg]$$

Next, insert the free-entry condition: $\frac{k_v}{q} = \bar{J}$

$$\begin{aligned} \frac{1}{1-\eta} \cdot \frac{k_v}{q} = \frac{1}{\lambda} \cdot \Bigg[& \\ & (1-\rho) \cdot \left(z - \frac{1-n}{n} \cdot b\right) \\ & + (1-p) \cdot (1-\rho^u) \cdot \left(\frac{u(c^w) - u(b)}{u'(c^w)} - c^w\right) \\ & + p \cdot (1-\rho^c) \cdot \left(\frac{u(c^c) - u(b)}{u'(c^w)} - e^c\right) \\ & + (1-\rho) \cdot \left(\frac{\lambda - \eta f}{1-\eta} \cdot \frac{k_v}{q}\right)\Bigg] \end{aligned}$$

The subsequent derivation of the wage equation is completely analogous to the derivation under STW regime described in Section 3.C.2. The wage equation thus reads:

$$(1-p)(1-\rho^u) \cdot (1-\eta) \cdot \left(\frac{u(c^w) - u(b)}{u'(c^w)} + \eta \cdot c^w\right) =$$

$$\eta \cdot \left[(1-\rho) \cdot z - (1-\rho) \cdot \frac{1-n}{n} \cdot b + (1-\rho) \cdot \theta \cdot k_v\right]$$

$$-p \cdot (1-\rho^c) \cdot (1-\eta) \cdot \left(\frac{u(c^c) - u(b)}{u'(c^w)} + \eta \cdot e^c\right)$$

Appendix 3.C Equilibrium STW

3.C.1 Bargaining

Nash-Bargaining Problem:

$$\max_{\substack{w^u(\epsilon),w^c(\epsilon),w^u_{\text{stw}}(\epsilon),w^c_{\text{stw}}(\epsilon),\\ h^u(\epsilon),h^c(\epsilon),h^u_{\text{stw}}(\epsilon),h^c_{\text{stw}}(\epsilon),\\ \epsilon^u_s,\epsilon^c_s,\xi^u_s,\xi^c_s}} \bar{J}^{1-\eta} \cdot (\bar{V} - U)^{\eta}$$

Subject to:

1.Financial constraints of financially constrained firms:

$$\begin{aligned} y(\epsilon, h(\epsilon)) - w^c(\epsilon) &\geq \lambda \cdot J^c(\epsilon) + (1-\lambda) \cdot \bar{J} \\ y(\epsilon, h_{\text{stw}}(\epsilon)) - w^c_{\text{stw}}(\epsilon) &\geq \lambda \cdot J^c_{\text{stw}}(\epsilon) + (1-\lambda) \cdot \bar{J} \end{aligned}$$

2.Workers have a commitment problem:

$$\begin{aligned} U &> V^u(\epsilon), \quad U > V^u_{\text{stw}}(\epsilon) \\ U &> V^c(\epsilon), \quad U > V^c_{\text{stw}}(\epsilon) \end{aligned}$$

Second, we can integrate the commitment problem of the worker into the value functions for the firm and worker surplus. For $i \in \{\text{stw}, \text{no stw}\}$ and $j \in \{u, c\}$, we can reformulate the condition as:

$$u(c^j_i(\epsilon)) - u(b) + (\lambda - f) \cdot (\bar{V} - U) \geq 0$$

Let $\epsilon^{j,\text{cp}}_s$ denote the threshold at which the worker wants to separate from a firm with STW support:

$$u(c^j_i(\epsilon^{j,\text{stw},\text{cp}}_s)) - u(b) + (\lambda - f)(\bar{V} - U) = 0$$

Likewise, we can determine the threshold at which workers leave firms without STW support, $\xi^{j,\text{cp}}_s$, as:

$$u(c^j_i(\xi^{j,\text{cp}}_s)) - u(b) + (\lambda - f)(\bar{V} - U) = 0$$

We can rewrite the value functions of the firm and the surplus of the worker as:

$$
\begin{aligned}
\bar{J}^B \;=\; & -\tau + (1-p)\cdot\Bigg[\int_{\max\{\xi_s^{u,B},\,\xi_s^{u,\mathrm{cp}},\,\epsilon_{\mathrm{stw}}\}}^{\infty} \frac{1}{\lambda}\Big(z(\epsilon) - c^u(\epsilon) + \lambda\bar{J}\Big)\,dG(\epsilon) \\
& \qquad + \int_{\max\{\epsilon_s^{u,B},\,\epsilon_s^{u,\mathrm{cp}}\}}^{\max\{\epsilon_s^{u,B},\epsilon_s^{u,cp},\epsilon_{\mathrm{stw}}\}} \frac{1}{\lambda}\Big(z_{\mathrm{stw}}(\epsilon) - c^u_{\mathrm{stw}}(\epsilon) + \lambda\bar{J}\Big)\,dG(\epsilon)\Bigg] \\
& + p\cdot\Bigg[\int_{\max\{\xi_s^{c,B},\,\xi_s^{c,\mathrm{cp}},\,\epsilon_{\mathrm{stw}}\}}^{\infty} \frac{1}{\lambda}\Big(z(\epsilon) - c^c(\epsilon) + \lambda\bar{J}\Big)\,dG(\epsilon) \\
& \qquad + \int_{\max\{\epsilon_s^{c,B},\,\epsilon_s^{c,\mathrm{cp}}\}}^{\max\{\epsilon_s^{c,B},\epsilon_s^{c,cp},\epsilon_{\mathrm{stw}}\}} \frac{1}{\lambda}\Big(z_{\mathrm{stw}}(\epsilon) - c^c_{\mathrm{stw}}(\epsilon) + \lambda\bar{J}\Big)\,dG(\epsilon)\Bigg]
\end{aligned}
$$

$$
\begin{aligned}
& (\bar{V}-U)^B \\
& = (1-p)\cdot\Bigg[\int_{\max\{\xi_s^{u,B},\,\xi_s^{u,\mathrm{cp}},\,\epsilon_{\mathrm{stw}}\}}^{\infty} \frac{1}{\lambda}\Big(u(c^u(\epsilon)) - u(b) + \lambda(\bar{V}-U)\Big)\,dG(\epsilon) \\
& \quad + \int_{\max\{\epsilon_s^{u,B},\,\epsilon_s^{u,\mathrm{cp}}\}}^{\max\{\epsilon_s^{u,B},\epsilon_s^{u,cp},\epsilon_{\mathrm{stw}}\}} \frac{1}{\lambda}\Big(u(c^u_{\mathrm{stw}}(\epsilon)) + \tau_{\mathrm{stw}}(\bar{h} - h_{\mathrm{stw}}(\epsilon)) - u(b) + \lambda(\bar{V}-U)\Big)\,dG(\epsilon)\Bigg] \\
& \quad + p\cdot\Bigg[\int_{\max\{\xi_s^{c,B},\,\xi_s^{c,\mathrm{cp}},\,\epsilon_{\mathrm{stw}}\}}^{\infty} \frac{1}{\lambda}\Big(u(c^w(\epsilon)) - u(b) + \lambda(\bar{V}-U)\Big)\,dG(\epsilon) \\
& \quad + \int_{\max\{\epsilon_s^{c,B},\,\epsilon_s^{c,\mathrm{cp}}\}}^{\max\{\epsilon_s^{c,B},\epsilon_s^{c,cp},\epsilon_{\mathrm{stw}}\}} \frac{1}{\lambda}\Big(u(c^c_{\mathrm{stw}}(\epsilon)) + \tau_{\mathrm{stw}}(\bar{h} - h_{\mathrm{stw}}(\epsilon)) - u(b) + \lambda(\bar{V}-U)\Big)\,dG(\epsilon)\Bigg]
\end{aligned}
$$

$$
\begin{aligned}
& (\bar{V}-U)^B \\
& = (1-p)\cdot\Bigg[\int_{\max\{\xi_s^{u,B},\,\xi_s^{u,\mathrm{cp}},\,\epsilon_{\mathrm{stw}}\}}^{\infty} \frac{1}{\lambda}\Big(u(c^u(\epsilon)) - u(b) + \lambda(\bar{V}-U)\Big)\,dG(\epsilon) \\
& \quad + \int_{\max\{\epsilon_s^{u,B},\,\epsilon_s^{u,\mathrm{cp}}\}}^{\max\{\epsilon_s^{u,B},\,\epsilon_s^{u,cp},\,\epsilon_{\mathrm{stw}}\}} \frac{1}{\lambda}\Big(u(c^u_{\mathrm{stw}}(\epsilon)) + \tau_{\mathrm{stw}}(\bar{h} - h_{\mathrm{stw}}(\epsilon)) - u(b) + \lambda(\bar{V}-U)\Big)\,dG(\epsilon)\Bigg] \\
& \quad + p\cdot\Bigg[\int_{\max\{\xi_s^{c,B},\,\xi_s^{c,\mathrm{cp}},\,\epsilon_{\mathrm{stw}}\}}^{\infty} \frac{1}{\lambda}\Big(u(c^w(\epsilon)) - u(b) + \lambda(\bar{V}-U)\Big)\,dG(\epsilon) \\
& \quad + \int_{\max\{\epsilon_s^{c,B},\,\epsilon_s^{c,\mathrm{cp}}\}}^{\max\{\epsilon_s^{c,B},\,\epsilon_s^{c,cp},\,\epsilon_{\mathrm{stw}}\}} \frac{1}{\lambda}\Big(u(c^c_{\mathrm{stw}}(\epsilon)) + \tau_{\mathrm{stw}}(\bar{h} - h_{\mathrm{stw}}(\epsilon)) - u(b) + \lambda(\bar{V}-U)\Big)\,dG(\epsilon)\Bigg]
\end{aligned}
$$

Note that $\xi_s^{u,B}, \xi_s^{c,\mathrm{B}}, \epsilon_s^{u,B}$ and $\epsilon_s^{c,\mathrm{B}}$ denote the separation thresholds at bargaining outcome.

To be consistent with the bargaining problem under lay-off taxes, we assume that $\epsilon_s^{i,\mathrm{cp}}$ and $\xi_s^{i,\mathrm{cp}}$ are exogenous to the bargaining problem.

Note that we can rewrite the problem so that we optimize over consumption equivalents instead of salaries. $c^i_{\mathrm{stw}}(\epsilon)$ denotes the consumption equivalent paid from the firm to the worker on STW. $c^i(\epsilon)$ is the consumption equivalent consumed by the worker.

$$\max_{\substack{c^u(\epsilon),c^c(\epsilon),c^u_{\text{stw}}(\epsilon),c^c_{\text{stw}}(\epsilon),\\ h^u(\epsilon),h^c(\epsilon),h^u_{\text{stw}}(\epsilon),h^c_{\text{stw}}(\epsilon),\\ \xi^{u,B}_s,\xi^{c,B}_s,\epsilon^{u,B}_s,\epsilon^{c,B}_s}} \left(\bar{J}^B\right)^{1-\eta} \cdot \left(\bar{V}-U\right)^{\eta}$$

1.Financial constraints of financially constrained firms

$$z(\epsilon) - c^c(\epsilon) \;\geq\; -\lambda \cdot J^c(\epsilon) - (1-\lambda)\cdot \bar{J}$$

$$z_{\text{stw}}(\epsilon) - c^c_{\text{stw}}(\epsilon) \;\geq\; -\lambda \cdot J^c_{\text{stw}}(\epsilon) - (1-\lambda)\cdot \bar{J}$$

Set up Kuhn-Tucker Conditions:

$$\left(\bar{J}^B\right)^{\eta} \cdot \left(\bar{V}-U\right)^{1-\eta} - \int_0^{\infty} \mathscr{M}(\epsilon)\left[z(\epsilon) - c^c(\epsilon) + \lambda\cdot\bar{J} + (1-\lambda)\cdot J^c(\epsilon)\right]dG(\epsilon)$$
$$-\int_0^{\infty} \mathscr{M}_{\text{stw}}(\epsilon)\left[z_{stw}(\epsilon) - c^c_{\text{stw}}(\epsilon) + \lambda\cdot\bar{J} + (1-\lambda)\cdot J^c_{\text{stw}}(\epsilon)\right]dG(\epsilon)$$

With Kuhn-Tucker conditions for a maximum:

(A) (i) $\mathscr{M}(\epsilon) \;\leq\; 0$

(ii) $\mathscr{M}(\epsilon)\left[z(\epsilon) - c^c(\epsilon) + \lambda\cdot\bar{J} + (1-\lambda)\cdot J^c(\epsilon)\right] \;=\; 0$

(B) (i) $\mathscr{M}_{\text{stw}}(\epsilon) \;\leq\; 0$

(ii) $\mathscr{M}_{\text{stw}}(\epsilon)\left[z_{stw}(\epsilon) - c^c_{\text{stw}}(\epsilon) + \lambda\cdot\bar{J} + (1-\lambda)\cdot J^c_{\text{stw}}(\epsilon)\right] \;=\; 0$

FOC for $c^u(\epsilon)$:

$$\frac{\partial \mathscr{B}}{\partial c^u(\epsilon)} \;=\; -(1-\eta)(1-p)\cdot g(\epsilon)\left(\frac{\bar{V}-U}{\bar{J}}\right)^{\eta}$$
$$+\;\eta\cdot(1-p)\cdot g(\epsilon)\cdot u'(c^u(\epsilon))\left(\frac{\bar{J}}{\bar{V}-U}\right)^{1-\eta} \;=\; 0$$

Unconstrained firms want to insure workers against idiosyncratic productivity shocks, outside STW:

$$\Longleftrightarrow \qquad \frac{1-\eta}{\eta}\cdot\frac{\bar{V}-U}{\bar{J}} \;=\; u'\left(c^u(\epsilon)\right) \quad \Rightarrow \quad u'(c^u(\epsilon)) \;=\; u'(c^u(\epsilon')) \;=\; u'(c^u)$$

FOC for $c^u_{\text{stw}}(\epsilon)$:

$$\frac{\partial \mathcal{B}}{\partial c^u_{\text{stw}}(\epsilon)} = -(1-\eta)(1-p)\cdot g(\epsilon)\left(\frac{\bar{V}-U}{\bar{J}}\right)^{\eta} + \eta\cdot(1-p)\cdot g(\epsilon)\cdot u'\left(c^u_{\text{stw}}(\epsilon)\right)\cdot\eta\cdot\left(\frac{\bar{J}}{\bar{V}-U}\right)^{1-\eta} = 0$$

Unconstrained firms want to insure workers against idiosyncratic productivity shocks, also on STW:

$$\iff \quad \frac{1-\eta}{\eta}\cdot\frac{\bar{V}-U}{\bar{J}} = u'\left(c^u(\epsilon)\right)$$

$$\Rightarrow \quad u'\left(c^{u,\text{f}}_{\text{stw}}(\epsilon)\right) = u'\left(c^u_{\text{stw}}(\epsilon')\right) = u'(c^w)$$

FOC for $c^c(\epsilon)$:

$$\frac{\partial \mathcal{B}}{\partial c^c(\epsilon)} = -(1-p)\cdot g(\epsilon)\cdot(1-\eta)\cdot\left(\frac{\bar{V}-U}{\bar{J}}\right)^{\eta} + (1-p)\cdot g(\epsilon)\cdot u'(c^c(\epsilon))\cdot\bar{N}\cdot\left(\frac{\bar{J}}{\bar{V}-U}\right)^{1-\eta} + \mathcal{M}(\epsilon) \overset{!}{=} 0$$

Insert the surplus splitting rule:

$$-(1-p)\cdot g(\epsilon)\cdot\left(\frac{\eta}{1-\eta}\cdot u'(c^w)\right)^{\eta} + (1-p)\cdot g(\epsilon)\cdot u'(c^c(\epsilon))\cdot\eta\cdot\left(\frac{1-\eta}{\eta}\cdot\frac{1}{u'(c^w)}\right) + \mathcal{M}(\epsilon) \overset{!}{=} 0$$

$$\iff \quad (1-p)\cdot g(\epsilon)\cdot\eta\cdot\left(u'(c^c(\epsilon))-u'(c^w)\right) = -\mathcal{M}(\epsilon)\cdot\left(u'(c^w)\cdot\frac{\eta}{1-\eta}\right)^{1-\eta}$$

Note: If $c^c(\epsilon) < u'(c^w)$, then the RHS of the equation is positive.
Thus, firms and workers gain joint surplus until $c^c(\epsilon) = c^w$.

If

$$z(\epsilon) - c^w + \lambda\cdot\bar{J} + (1-\lambda)\cdot J^c(\epsilon) \geq 0,$$

then the constraint is non-binding and $\mathcal{M}(\epsilon) = 0$. ✓

If

$$z(\epsilon) - c^w + \lambda \cdot \bar{J} + (1-\lambda) \cdot J^c(\epsilon) \ < \ 0,$$

then the constraint is binding and $\mathscr{M}(\epsilon) \ < \ 0$. ✓

Thus, it is optimal for the firm to pay as much as it can:

$$J^c(\epsilon) \ = \ z(\epsilon) - c^c(\epsilon) + \lambda \cdot \bar{J} + (1-\lambda) \cdot J^c(\epsilon) \ = \ 0$$

$$\iff \qquad c^c(\epsilon) \ = \ z(\epsilon) + \lambda \cdot \bar{J}$$

Using the same arguments, we get that under STW

$$c_{\text{stw}}(\epsilon) \ = \ c^w \quad \text{if the financial constraint is non-binding,}$$

and

$$c_{\text{stw}}(\epsilon) \ = \ \epsilon_{\text{stw}}(\epsilon) + \tau_{\text{stw}} \cdot (\bar{h} - h_{\text{stw}}(\epsilon)) + \lambda \cdot \bar{J} \quad \text{when it is binding.}$$

Using the surplus splitting rule, we can simplify the Nash-Bargaining problem to:

$$\max_{h^u(\epsilon), h^c(\epsilon), h^u_{\text{stw}}(\epsilon), h^c_{\text{stw}}(\epsilon), \, \xi^u_s, \xi^c_s, \xi^{uc}_s, \xi^{c,c}_s} (1-\eta)^{1-\eta} \cdot (u'(c^w) \cdot \eta)^{\eta} \cdot \bar{S}$$

That is, we just have to maximize the joint surplus:

$$\begin{aligned}
\bar{S} \ = \ & -\tau \\
& + (1-p) \cdot \int_{\max\{\xi^{u,B}_s, \xi^{u,cp}_s, \epsilon_{stw}\}}^{\infty} \frac{1}{\lambda}\Big(z(\epsilon) + \frac{u(c^w) - u(b)}{u'(c^w)} - c^w + (1-\eta \cdot f) \cdot \bar{S}\Big) dG(\epsilon) \\
& + (1-p) \cdot \int_{\max\{\epsilon^{u,B}_s, \epsilon^{u,cp}_s\}}^{\max\{\epsilon^{u,B}_s, \epsilon^{u,cp}_s, \epsilon_{stw}\}} \frac{1}{\lambda}\Big(z(\epsilon) + \frac{u(c^w) - u(b)}{u'(c^w)} - c^w \\
& \qquad + \tau_{stw}(\bar{h} - h_{stw}(\epsilon)) + (1-\eta \cdot f) \cdot \bar{S}\Big) dG(\epsilon) \\
& + p \cdot \int_{\epsilon^p}^{\infty} \frac{1}{\lambda}\Big(z(\epsilon) + \frac{u(c^w) - u(b)}{u'(c^w)} - c^w + (1-\eta \cdot f) \cdot \bar{S}\Big) dG(\epsilon) \\
& + p \cdot \int_{\max\{\xi^c_s, \xi^{c,cp}_s, \epsilon_{stw}\}}^{\epsilon^p} \frac{1}{\lambda}\Big(z(\epsilon) + \frac{u(c^c(\epsilon)) - u(b)}{u'(c^w)} - c^c(\epsilon) + (1-\eta \cdot f) \cdot \bar{S}\Big) dG(\epsilon) \\
& + p \cdot \int_{\max\{\epsilon^c_s, \epsilon^{c,cp}_s\}}^{\max\{\epsilon^c_s, \epsilon^{c,cp}_s, \epsilon_{stw}\}} \frac{1}{\lambda}\Big(z_{stw}(\epsilon) + \frac{u(c^c_{stw}(\epsilon)) - u(b)}{u'(c^w)} \\
& \qquad - c^c(\epsilon) + (1-\eta \cdot f) \cdot \bar{S}\Big) dG(\epsilon)
\end{aligned}$$

Note that without financial constraints, the firm smooths the consumption equivalent consumed by the worker:

$$c^w = c^u_{\text{stw}}(\epsilon) = c^u_{\text{stw}}(\epsilon) + \tau_{\text{stw}} \cdot (\bar{h} - h_{\text{stw}}(\epsilon))$$

$$\Leftrightarrow \quad c^u_{\text{stw}} = c^w - \tau_{\text{stw}} \cdot (\bar{h} - h_{\text{stw}}(\epsilon))$$

STW makes it less expensive to smooth consumption equivalents consumed by the worker.

The FOC for working hours in unconstrained firms outside STW is:

$$\begin{aligned} \frac{\partial \mathcal{B}}{\partial h^u(\epsilon)} &= (1-p) \cdot g(\epsilon) \cdot \left(\frac{\partial y(\epsilon, h^u(\epsilon))}{\partial h^u(\epsilon)} - \phi'(h^u(\epsilon)) \right) = 0 \\ \Leftrightarrow \quad \frac{\partial y(\epsilon, h^u(\epsilon))}{\partial h^u(\epsilon)} &= \phi'(h^u(\epsilon)) \end{aligned}$$

The FOC for working hours in constrained firms *outside* STW is:

$$\begin{aligned} \frac{\partial \mathcal{B}}{\partial h^c(\epsilon)} &= p \cdot g(\epsilon) \cdot \left(\frac{\partial y(\epsilon, h^c(\epsilon))}{\partial h^c(\epsilon)} - \phi'(h^c(\epsilon)) \right) \cdot \frac{u'(c^c(\epsilon))}{u'(c^w)} = 0 \\ \Leftrightarrow \quad \frac{\partial y(\epsilon, h^c(\epsilon))}{\partial h^c(\epsilon)} &= \phi'(h^c(\epsilon)) \end{aligned}$$

The FOC for working hours in unconstrained firms *on* STW is:

$$\begin{aligned} \frac{\partial \mathcal{B}}{\partial h^u_{\text{stw}}(\epsilon)} &= (1-p) \cdot g(\epsilon) \cdot \left(\frac{\partial y(\epsilon, h^u_{\text{stw}}(\epsilon))}{\partial h^u_{\text{stw}}(\epsilon)} - \phi'(h^u_{\text{stw}}(\epsilon)) - \tau_{\text{stw}} \right) = 0 \\ \Leftrightarrow \quad \frac{\partial y(\epsilon, h^u_{\text{stw}}(\epsilon))}{\partial h^u_{\text{stw}}(\epsilon)} &= \phi'(h^u_{\text{stw}}(\epsilon)) + \tau_{\text{stw}} \end{aligned}$$

The FOC for working hours in constrained firms *on* STW is:

$$\begin{aligned} \frac{\partial \mathcal{B}}{\partial h^c(\epsilon)} &= p \cdot g(\epsilon) \cdot \left(\frac{\partial y(\epsilon, h^c(\epsilon))}{\partial h^c(\epsilon)} - \phi'(h^c(\epsilon)) - \tau_{\text{stw}} \right) \cdot \frac{u'(c^c(\epsilon))}{u'(c^w)} = 0 \\ \Leftrightarrow \quad \frac{\partial y(\epsilon, h^c(\epsilon))}{\partial h^c(\epsilon)} &= \phi'(h^c(\epsilon)) + \tau_{\text{stw}} \end{aligned}$$

Suppose that

$$\max\{\xi_s^{u,B}, \xi_s^{u,\text{cp}}, \epsilon_{\text{stw}}\} = \xi_s^{u,B}$$

Then the FOC for the separation threshold of the unconstrained firm without STW support is:

$$\frac{\partial \mathcal{B}}{\partial \xi_s^{u,B}} = -(1-p) \cdot g(\xi_s^{u,B}) \cdot \frac{1}{\lambda} \left(z(\xi_s^{u,B}) + \frac{u(c^w) - u(b)}{u'(c^w)} - c^w + (\lambda - \eta f) \cdot \bar{S} \right) = 0$$

$$\iff \quad z(\xi_s^{u,B}) + \frac{u(c^w) - u(b)}{u'(c^w)} - c^w + (\lambda - \eta f) \cdot \bar{S} \;=\; 0$$

Insert free-entry Condition:

$$z(\xi_s^{u,B}) + z(\xi_s^{u,B}) + \frac{u(c^w) - u(b)}{u'(c^w)} - c^w + \frac{(\lambda - \eta f)}{1-\eta} \cdot \frac{k_v}{q} \;=\; 0$$

Note that $\xi_s^{u,B} > \xi_s^{u,\text{cp}}$ must hold, as workers would never quit under the insurance constraint of the firm:

$$V(\epsilon) - U \;=\; u(c^w) - u(b) + (\lambda - f)(\bar{V} - U) \;>\; 0$$

It always gets positive surplus!
Suppose that

$$\max\{\xi_s^{c,B}, \xi_s^{c,\text{cp}}, \epsilon_{\text{stw}}\} \;=\; \xi_s^{c,B}$$

Then the FOC for the separation threshold of the unconstrained firm without STW support is:

$$\frac{\partial \mathcal{B}}{\partial \xi_s^{c,B}} \;=\; -(1-p) \cdot g(\xi_s^{c,B}) \cdot \left[\frac{1}{\lambda} \cdot \left(z(\xi_s^{c,B}) + \frac{u\left(c^c(\xi_s^{c,B})\right) - u(b)}{u'(c^w)} - c^c(\xi_s^{c,B}) + (\lambda - \eta f) \cdot \bar{S} \right) + L \right] \;=\; 0$$

$$\iff \quad z(\xi_s^{c,B}) + \lambda \cdot F + \frac{u(c^c(\xi_s^{c,B})) - u(b)}{u'(c^w)} - c^c(\xi_s^{c,B}) + (\lambda - \eta f) \cdot \bar{S} \;=\; 0$$

Insert free-entry condition:

$$z(\xi_s^{c,B}) + \frac{u(c^c(\xi_s^{c,B})) - u(b)}{u'(c^w)} - c^c(\xi_s^{c,B}) + \frac{(\lambda - \eta f)}{1-\eta} \cdot \frac{k_v}{q} \;=\; 0$$

The participation constraint of workers can be written with free-entry condition as:

$$z(\xi_s^{c,\text{cp}}) + \frac{u(c^c(\xi_s^{c,\text{cp}})) - u(b)}{u'(c^w)} - c^c(\xi_s^{c,\text{cp}}) + \frac{(\lambda - \eta f)}{1-\eta} \cdot \frac{k_v}{q} \;=\; 0$$

Note that these are the same conditions. This implies that given the inability to insure workers against idiosyncratic productivity shocks, separations for a constrained firm without access to STW are efficient.
Suppose that

$$\max\{\epsilon_s^{u,B}, \epsilon_s^{u,\text{cp}}, \epsilon_{\text{stw}}\} \;=\; \epsilon_s^{u,B}$$

Then the FOC for the separation threshold of the unconstrained firm with STW support is:

$$\frac{\partial \mathcal{B}}{\partial \epsilon_s^{u,B}} = -(1-p)\cdot g(\epsilon_s^{u,B})\cdot\Bigg[\frac{1}{\lambda}\cdot\Big(z(\epsilon_s^{u,B}) + \frac{u(c^w)-u(b)}{u'(c^w)} - c^w + \tau_{\text{stw}}\cdot(\bar{h} - h_{\text{stw}}(\epsilon_s^{u,B})) + (\lambda - \eta f)\cdot\bar{S}\Big)\Bigg] = 0$$

$$\iff \quad z(\epsilon_s^{u,B}) + \frac{u(c^w)-u(b)}{u'(c^w)} - c^w + \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon_s^{u,B})) + (\lambda - \eta f)\cdot\bar{S} = 0$$

Insert free-entry Condition:

$$z(\epsilon_s^{u,B}) + \frac{u(c^w)-u(b)}{u'(c^w)} - c^w + \frac{\lambda-\eta f}{1-\eta}\cdot\frac{k_v}{q} = 0$$

Note that $\epsilon_s^{u,B} > \epsilon_s^{u,\text{cp}}$ must hold, as workers would never quit under the insurance constraint of the firm:

$$V(\epsilon) - U = u(c^w) - u(b) + (\lambda - f)(\bar{V} - U) > 0$$

It always gets positive surplus!
Suppose that

$$\max\{\epsilon_s^{c,B}, \epsilon_s^{c,\text{cp}}, \epsilon_{\text{stw}}\} = \epsilon_s^{c,B}$$

Then the FOC for the separation threshold of the constrained firm with STW support is:

$$\frac{\partial \mathcal{B}}{\partial \epsilon_s^{c,B}} = -(1-p)\cdot g(\epsilon_s^{c,B})\cdot\Bigg[\frac{1}{\lambda}\cdot\Big(z(\epsilon_s^{c,B}) + \frac{u\big(c_{\text{stw}}^c(\epsilon_s^{c,B})\big)-u(b)}{u'(c^w)} - c_{\text{stw}}^c(\epsilon_s^{c,B}) + (\lambda-\eta f)\cdot\bar{S}\Big) + L\Bigg] = 0$$

$$\iff \quad z(\epsilon_s^{c,B}) + \frac{u(c_{\text{stw}}^c(\epsilon_s^{c,B}))-u(b)}{u'(c^w)} - c_{\text{stw}}^c(\epsilon_s^{c,B}) + (\lambda-\eta f)\cdot\bar{S} = 0$$

Insert free-entry condition:

$$z(\epsilon_s^{c,B}) + \frac{u(c_{\text{stw}}^c(\epsilon_s^{c,B}))-u(b)}{u'(c^w)} - c_{\text{stw}}^c(\epsilon_s^{c,B}) + \frac{\lambda-\eta f}{1-\eta}\cdot\frac{k_v}{q} = 0$$

The participation constraint of workers can be written with free-entry condition as:

$$z(\epsilon_s^{c,\text{cp}}) + \frac{u(c_{\text{stw}}^c(\epsilon_s^{c,\text{cp}}))-u(b)}{u'(c^w)} - c_{\text{stw}}^c(\epsilon_s^{c,\text{cp}}) + \frac{\lambda-\eta f}{1-\eta}\cdot\frac{k_v}{q} = 0$$

Note that these are the same conditions. This implies that, given the inability to insure workers against idiosyncratic productivity shocks, separations for a constrained firm with access to STW are efficient.

3.C.2 Job Creation and Wage Equation

This section derives the job creation and the equation for STW. It follows its notation part 3.D of the appendix.

(I)Value of an unconstrained firm outside STW after the idiosyncratic productivity shock has realized:

$$\begin{aligned} J^u(\epsilon) &= z(\epsilon) - c^w + \lambda \bar{J} + (1-\lambda) J^u(\epsilon) \\ \Leftrightarrow \quad J^u(\epsilon) &= \frac{1}{\lambda}\left(z(\epsilon) - c^w + \lambda \bar{J}\right) \end{aligned}$$

(II)Value of an unconstrained firm on STW after the idiosyncratic productivity shock has realized:

$$\begin{aligned} J^u_{\text{stw}}(\epsilon) &= z_{\text{stw}}(\epsilon) + \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon)) - c^w + \lambda \bar{J} + (1-\lambda) J^u_{\text{stw}}(\epsilon) \\ \Leftrightarrow \quad J^u_{\text{stw}}(\epsilon) &= \frac{1}{\lambda}\left(z_{\text{stw}}(\epsilon) - c^w + \lambda \bar{J}\right) \end{aligned}$$

(III)Value of a constrained firm outside STW with non-binding constraints after the idiosyncratic productivity shock has realized:

$$\begin{aligned} J^c(\epsilon) &= z(\epsilon) - c^w + \lambda \bar{J} + (1-\lambda) J^c(\epsilon) \\ \Leftrightarrow \quad J^c(\epsilon) &= \frac{1}{\lambda}\left(z(\epsilon) - c^w + \lambda \bar{J}\right) \end{aligned}$$

(IV)Value of a constrained firm outside STW with binding constraints after the idiosyncratic productivity shock has realized:

$$\begin{aligned} J^c(\epsilon) &= z(\epsilon) - c^w(\epsilon) + \lambda \bar{J} + (1-\lambda) J^c(\epsilon) \\ \Leftrightarrow \quad J^c(\epsilon) &= \frac{1}{\lambda}\left(z(\epsilon) - c^w(\epsilon) + \lambda \bar{J}\right) \end{aligned}$$

(V)Value of a constrained firm on STW after the idiosyncratic productivity shock has realized:

$$\begin{aligned} J^c_{\text{stw}}(\epsilon) &= z_{\text{stw}}(\epsilon) - c^w_{stw}(\epsilon) + \lambda J + (1-\lambda) J^c_{\text{stw}}(\epsilon) \\ \Leftrightarrow \quad J^c_{\text{stw}}(\epsilon) &= \frac{1}{\lambda}\left(z_{\text{stw}}(\epsilon) - c^w_{stw}(\epsilon) + \lambda J\right) \end{aligned}$$

(VI)Value of a firm before the idiosyncratic productivity shock has realized:

$$\begin{aligned}\bar{J} \;&=\; -\tau + (1-p)\cdot\left(\int_{\epsilon_{\text{stw}}}^{\infty} J^{u}(\epsilon)\,dG(\epsilon) + \int_{\epsilon_s^u}^{\epsilon_{\text{stw}}} J^{u}_{\text{stw}}(\epsilon)\,dG(\epsilon)\right)\\ &\quad + p\cdot\left(\int_{\max\{\epsilon_{\text{stw}},\xi_s^c\}}^{\infty} J^{c}(\epsilon)\,dG(\epsilon) + \int_{\epsilon_s^c}^{\max\{\epsilon_{\text{stw}},\epsilon_s^c\}} J^{c}_{\text{stw}}(\epsilon)\,dG(\epsilon)\right)\end{aligned}$$

Inserting (I) - (IV) into (V) gives:

$$\begin{aligned}\bar{J} \;&=\; -\tau\\ &\quad + \frac{1}{\lambda}\Bigg[(1-\rho)\cdot(z-\Omega)\\ &\quad - (1-p)\cdot(1-\rho^{u})\cdot\left(c^{w} - \int_{\epsilon_s^u}^{\epsilon_{\text{stw}}} \tau_{\text{stw}}(\bar{h}-h_{\text{stw}}(\epsilon))\,dG(\epsilon)\right)\\ &\quad - p\cdot(1-\rho^{c})\cdot\left(e^{c} - \int_{\epsilon_s^c}^{\max\{\epsilon_{\text{stw}},\epsilon_s^c\}} \tau_{\text{stw}}(\bar{h}-h_{\text{stw}}(\epsilon))\,dG(\epsilon)\right)\\ &\quad + (1-\rho)\cdot\lambda\cdot\bar{J}\Bigg]\end{aligned}$$

(VII)Value of an unconstrained worker outside STW after the idiosyncratic productivity shock has realized:

$$\begin{aligned} V^{u}(\epsilon) \;&=\; u(c^{w}) + \lambda\bar{V} + (1-\lambda)\,V^{u}(\epsilon)\\ \Leftrightarrow\quad V^{u}(\epsilon) \;&=\; \frac{1}{\lambda}\big(u(c^{w}) + \lambda\bar{V}\big)\end{aligned}$$

(VIII)Value of an unconstrained worker on STW after the idiosyncratic productivity shock has realized:

$$\begin{aligned} V^{u}_{\text{stw}}(\epsilon) \;&=\; u(c^{w}) + \lambda\bar{V} + (1-\lambda)\,V^{u}_{\text{stw}}(\epsilon)\\ \Leftrightarrow\quad V^{u}_{\text{stw}}(\epsilon) \;&=\; \frac{1}{\lambda}\big(u(c^{w}) + \lambda\bar{V}\big)\end{aligned}$$

(IX)Value of a constrained worker outside STW with non-binding constraints after the idiosyncratic productivity shock has realized:

$$\begin{aligned} V^{c}(\epsilon) \;&=\; u(c^{w}) + \lambda\bar{V} + (1-\lambda)\,V^{c}(\epsilon)\\ \Leftrightarrow\quad V^{c}(\epsilon) \;&=\; \frac{1}{\lambda}\big(u(c^{w} + \lambda\bar{V}\big)\end{aligned}$$

(X)Value of a constrained worker outside STW with binding constraints after the idiosyncratic productivity shock has realized:

$$
\begin{aligned}
V^c(\epsilon) &= u(c^w(\epsilon)) + \lambda \bar{V} + (1-\lambda) V^c(\epsilon) \\
\Leftrightarrow \quad V^c(\epsilon) &= \frac{1}{\lambda}\left(u(c^w(\epsilon)) + \lambda \bar{V}\right)
\end{aligned}
$$

(XI)Value of a constrained worker on STW after the idiosyncratic productivity shock has realized:

$$
\begin{aligned}
V^c_{\text{stw}}(\epsilon) &= u(c_{\text{stw}}(\epsilon)) + \lambda \bar{V} + (1-\lambda) V^c_{\text{stw}}(\epsilon) \\
\Leftrightarrow \quad V^c_{\text{stw}}(\epsilon) &= \frac{1}{\lambda}\left(u(c_{\text{stw}}(\epsilon)) + \lambda \bar{V}\right)
\end{aligned}
$$

(XII)Value of a worker before the idiosyncratic productivity shock has realized:

$$
\begin{aligned}
\bar{V} = (1-p) \cdot \Bigg(& \int_{\epsilon_{stw}}^{\infty} V^u(\epsilon)\, dG(\epsilon) \\
& + \int_{\epsilon_s^u}^{\epsilon_{stw}} V^u_{\text{stw}}(\epsilon)\, dG(\epsilon) \Bigg) \\
& + p \cdot \Bigg(\int_{\epsilon_p}^{\infty} V^c(\epsilon)\, dG(\epsilon) \\
& + \int_{\max\{\xi_s^c, \epsilon_{stw}\}}^{\epsilon_p} V^c(\epsilon)\, dG(\epsilon) \\
& + \int_{\epsilon_s^c}^{\max\{\epsilon_s^c, \epsilon_{stw}\}} V^c_{\text{stw}}(\epsilon)\, dG(\epsilon) \Bigg) \\
& + \rho \cdot U
\end{aligned}
$$

Inserting (VII) - (XI) into (XII) gives:

$$
\begin{aligned}
\bar{V} = \frac{1}{\lambda}\Big[& (1-p)(1-\rho^u) \cdot u(c^w) + p \cdot (1-\rho^c) \cdot u^c \\
& + (1-\rho) \cdot \lambda \cdot \bar{V} + \rho \cdot U \Big]
\end{aligned}
$$

(XIII)Unemployment:

$$
U = u(b) + f \cdot \bar{V} + (1-f) \cdot U
$$

Next, we want to calculate the joint surplus of firms and workers.

(A)The surplus of the worker after the idiosyncratic productivity shock has realized is:

$$
\begin{aligned}
& V_i^j(\epsilon) - U \;=\; u(c_i^j(\epsilon)) - u(b) + (\lambda - f)\cdot V_i^j(\epsilon) + (\lambda - f)\cdot \bar{V} - (\lambda - f)\cdot U \\
\Leftrightarrow\quad & V_i^j(\epsilon) - U \;=\; u(c_i^j(\epsilon)) - u(b) + (\lambda - f)\cdot (V_i^j(\epsilon) - U) + (\lambda - f)\cdot (\bar{V} - U) \\
\Leftrightarrow\quad & V_i^j(\epsilon) - U \;=\; u(c_i^j(\epsilon)) - u(b) + (\lambda - f)\cdot (V_i^j(\epsilon) - U) + (\lambda - f)\cdot (\bar{V} - U) \\
\Leftrightarrow\quad & V_i^j(\epsilon) - U \;=\; \frac{1}{\lambda}\Big(u(c_i(\epsilon)) - u(b) + (\lambda - f)\cdot (\bar{V} - U)\Big)
\end{aligned}
$$

for $i \in \{\text{stw}, \text{no stw}\}$ and $j \in \{u, c\}$.

(B)The surplus of the worker before shock realizations are known:

$$
\begin{aligned}
\bar{V} - U \;=\; & (1-p)\cdot\left(\left(\int_{\epsilon_{\text{stw}}}^{\infty} V^u(\epsilon) - U\, dG(\epsilon) + \int_{\epsilon_s^u}^{\epsilon_{\text{stw}}} V_{\text{stw}}^u(\epsilon) - U\, dG(\epsilon)\right)\right. \\
& + p\cdot\left(\left(\int_{\epsilon_p}^{\infty} V(\epsilon) - U\, dG(\epsilon) + \int_{\max\{\xi_s^c, \epsilon_{\text{stw}}\}}^{\epsilon_p} V^c(\epsilon) - U\, dG(\epsilon)\right.\right. \\
& + \left.\int_{\epsilon_s^c}^{\max\{\epsilon_s^c, \epsilon_{\text{stw}}\}} V_{\text{stw}}^c(\epsilon) - U\, dG(\epsilon)\right)
\end{aligned}
$$

Inserting A into B gives:

$$
\begin{aligned}
\bar{V} - U \;=\; & \frac{1}{\lambda}\Big[(1-p)(1-\rho^u)\cdot(u(c^w) - u(b)) \\
& + p\cdot(1-\rho^c)\cdot(u^c - u(b)) \\
& + (\lambda - f)\cdot(\bar{V} - U)\Big]
\end{aligned}
$$

(C)Finally, we can calculate the joint surplus before shock realizations are known:

$$
\begin{aligned}
S \;=\; & J + \frac{\bar{V} - U}{u'(c^w)} \\
=\; & -\tau + \frac{1}{\lambda}\Bigg[(1-\rho)(1-\rho^u)\cdot\left(\frac{u(c^w) - u(b)}{u'(c^w)} - c^w\right. \\
& + \left.\frac{1}{1-\rho^u}\int_{\epsilon_s^u}^{\epsilon_{stw}} (\bar{h} - h_{\text{stw}}(\epsilon))\, dG(\epsilon)\right) \\
& + p(1-\rho^c)\cdot\left(\frac{u^c - u(b)}{u'(c^w)} - e^c\right. \\
& + \left.\frac{1}{1-\rho^c}\int_{\epsilon_s^c}^{\max\{\epsilon_{\text{stw}}, \epsilon_s^c\}} \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon))\, dG(\epsilon)\right) \\
& + (1-\rho)\lambda J + (1-\rho)(\lambda - f)\cdot\frac{\bar{V} - U}{u'(c^w)}\Bigg]
\end{aligned}
$$

Next, we start deriving the job-creation condition. Inserting the surplus splitting rule

$$J = \eta \cdot S, \quad \frac{\bar{V} - U}{u'(c^w)} = (1-\eta) \cdot S$$

into the equation for the joint surplus gives:

$$\begin{aligned}
S &= J + \frac{\bar{V} - U}{u'(c^w)} = \eta S + (1-\eta) S \\
&= -\tau + \frac{1}{\lambda}\bigg[(1-\rho)\left(z - \Omega - \frac{1-n}{n} b\right) \\
&\quad + (1-p)(1-\rho^u)\left(\frac{u(c^w) - u(b)}{u'(c^w)} - c^w\right) \\
&\quad + (1-p)\int_{\epsilon_s^u}^{\epsilon_{stw}} (\bar{h} - h_{\text{stw}}(\epsilon))\, dG(\epsilon) \\
&\quad + p(1-\rho^c)\left(\frac{u^c - u(b)}{u'(c^w)} - e^c\right) \\
&\quad + p\int_{\epsilon_s^c}^{\max\{\epsilon_{\text{stw}}, \epsilon_s^c\}} \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon))\, dG(\epsilon) \\
&\quad + (1-\rho)\cdot\lambda\cdot(1-\eta)\cdot S + (1-\rho)(\lambda - f)\eta \cdot S\bigg] \\
\Leftrightarrow \quad S &= -\tau + \frac{1}{\lambda}\bigg[(1-\rho)\left(z - \Omega - \frac{1-n}{n} b\right) \\
&\quad + (1-p)(1-\rho^u)\left(\frac{u(c^w) - u(b)}{u'(c^w)} - c^w\right) \\
&\quad + (1-p)\int_{\epsilon_s^u}^{\epsilon_{stw}} (\bar{h} - h_{\text{stw}}(\epsilon))\, dG(\epsilon) \\
&\quad + p(1-\rho^c)\left(\frac{u^c - u(b)}{u'(c^w)} - e^c\right) \\
&\quad + p\int_{\epsilon_s^c}^{\max\{\epsilon_{\text{stw}}, \epsilon_s^c\}} \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon))\, dG(\epsilon) \\
&\quad + (1-\rho)(\lambda - \eta \cdot f)\cdot S\bigg]
\end{aligned}$$

Next, we want to replace the tax. To do this, we need to know the budget constraint of the government:

$$n^s \cdot \tau = \frac{n}{1-\rho} \cdot \left((1-p) \int_{\epsilon_s^u}^{\epsilon_{stw}} \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon))\, dG(\epsilon) \right.$$
$$\left. + p \int_{\epsilon_s^c}^{\max\{\epsilon_{\text{stw}}, \epsilon_s^c\}} \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon))\, dG(\epsilon) \right)$$
$$+ (1-n) \cdot b$$

Inserting the employment equation

$$n = (1-\rho) \cdot \frac{n^s}{\lambda}$$

gives:

$$\tau = \frac{1}{\lambda} \cdot \left((1-p) \int_{\epsilon_s^u}^{\infty} \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon))\, dG(\epsilon) \right.$$
$$\left. + p \int_{\epsilon_s^c}^{\max\{\epsilon_{\text{stw}}, \epsilon_s^c\}} \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon))\, dG(\epsilon) \right)$$
$$+ \frac{(1-\rho)}{n^s} \cdot b$$

Inserting the expression for the number of firms that receive a shock

$$n^s = \frac{n}{1-\rho} \cdot \lambda$$

gives:

$$\tau = \frac{1}{\lambda} \cdot \left((1-p) \int_{\epsilon_s^u}^{\infty} \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon))\, dG(\epsilon) \right.$$
$$\left. + p \int_{\epsilon_s^c}^{\max\{\epsilon_{\text{stw}}, \epsilon_s^c\}} \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon))\, dG(\epsilon) \right)$$
$$+ \frac{1}{\lambda} \cdot (1-\rho) \cdot \left(\frac{1-n}{n} \cdot b \right)$$

Inserting the tax into the surplus equation gives:

$$
\begin{aligned}
S = \frac{1}{\lambda}\Big[&(1-\rho)\left(z-\Omega-\frac{1-n}{n}\cdot b\right)\\
&+(1-p)(1-\rho^u)\left(\frac{u(c^w)-u(b)}{u'(c^w)}-c^w\right)\\
&+(1-p)\int_{\epsilon_s^u}^{\epsilon_{stw}}(\bar{h}-h_{\text{stw}}(\epsilon))\,dG(\epsilon)\\
&+p(1-\rho^c)\left(\frac{u^c-u(b)}{u'(c^w)}-e^c\right)\\
&+p\int_{\epsilon_s^c}^{\max\{\epsilon_{\text{stw}},\epsilon_s^c\}}\tau_{\text{stw}}(\bar{h}-h_{\text{stw}}(\epsilon))\,dG(\epsilon)\\
&+(1-\rho)(\lambda-f)\cdot S\Big]
\end{aligned}
$$

Inserting the free-entry condition

$$\frac{k_v}{q} = J$$

gives:

$$
\begin{aligned}
\frac{1}{1-\eta}\cdot\frac{k_v}{q} = \frac{1}{\lambda}\Big[&(1-\rho)\left(z-\Omega-\frac{1-n}{n}\cdot b\right)\\
&+(1-p)(1-\rho^u)\left(\frac{u(c^w)-u(b)}{u'(c^w)}-c^w\right)\\
&+p(1-\rho^c)\left(\frac{u^c-u(b)}{u'(c^w)}-e^c\right)\\
&+(1-\rho)\cdot\frac{\lambda-\eta\cdot f}{1-\eta}\cdot\frac{k_v}{q}\Big]
\end{aligned}
$$

This is the job-creation condition used in the Ramsey problem for STW in Appendix 3.D.

Next, we want to calculate the wage-equation. Remember the surplus splitting rule from Nash-Bargaining:

$$\eta\cdot\bar{J} = (1-\eta)\cdot\frac{\bar{V}-U}{u'(c^w)}$$

Inserting the tax into the value of a worker for a firm gives:

$$
\begin{aligned}
\bar{J} = \frac{1}{\lambda}\Big[&(1-\rho)(z-\Omega)-p(1-\rho^u)e^w-(1-p)(1-\rho^c)e\\
&-(1-\rho)\cdot\frac{1-n}{n}b+(1-\rho)\lambda\bar{J}\Big]
\end{aligned}
$$

Likewise, we can calculate the surplus of a worker:

$$\bar{V} - U = \frac{1}{\lambda}\Big[(1-p)(1-\rho^u)\big(u(c^w) - u(b)\big) + p(1-\rho^c)\big(u^c - u(b)\big) + (1-\rho)(\lambda - f)\cdot(\bar{V} - U)\Big]$$

Next, we can insert the value of a worker for a firm and the surplus of a worker into the surplus splitting rule:

$$\frac{\eta}{\lambda}\cdot\Big[\ (1-\rho)(z-\Omega) - p(1-\rho^u)\cdot c^w - (1-p)(1-\rho^c)\cdot e^c - (1-\rho)\cdot\frac{1-n}{n}\cdot b + (1-\rho)\cdot\lambda\cdot\bar{J} = \frac{1-\eta}{\lambda}\cdot\frac{1}{u'(c^w)}\cdot\Big[(1-p)(1-\rho^u)\cdot(u(c^w) - u(b)) + p(1-\rho^c)\cdot(u^c - u(b)) + (1-\rho)(\lambda - f)\cdot(\bar{V} - U)\Big]\Big]$$

Insert the surplus splitting rule again: $\eta\cdot\bar{J} = (1-\eta)\cdot\frac{\bar{V}-U}{u'(c^w)}$

$$\frac{\eta}{\lambda}\cdot\Big[\ (1-\rho)(z-\Omega) - p(1-\rho^u)\cdot c^w - (1-p)(1-\rho^c)\cdot e^c - (1-\rho)\cdot\frac{1-n}{n}\cdot b + (1-\rho)\cdot\lambda\cdot\bar{J}\Big] = \frac{1-\eta}{\lambda}\cdot\frac{1}{u'(c^w)}\cdot\Big[(1-p)(1-\rho^u)\cdot(u(c^w) - u(b)) + p(1-\rho^c)\cdot(u^c - u(b)) + (1-\rho)(\lambda - f)\cdot u'(c^w)\cdot\frac{\eta}{1-\eta}\cdot\bar{J}\Big]$$

$$\Leftrightarrow \frac{\eta}{\lambda}\Big[(1-\rho)(z-\Omega) - p(1-\rho^u)\cdot c^w - (1-p)(1-\rho^c)\cdot e^c - (1-\rho)\cdot\frac{1-n}{n}\cdot b + (1-\rho)\cdot\lambda\cdot f\cdot\bar{J}\Big] = \frac{1-\eta}{\lambda}\cdot\frac{1}{u'(c^w)}\Big[(1-p)(1-\rho^u)(u(c^w) - u(b)) + p(1-\rho^c)(u^c - u(b))\Big]$$

$$
\begin{aligned}
\Leftrightarrow \quad & \eta \cdot \left[(1-\rho)(z-\Omega) - (1-\rho) \cdot \frac{1-n}{n} \cdot b + (1-\rho) \cdot f \cdot \bar{J} \right] \\
&= (1-p)(1-\rho^u)(1-\eta) \cdot \left(\frac{u(c^w)-u(b)}{u'(c^w)} + \eta \cdot c^w \right) \\
&\quad + p(1-\rho^c)(1-\eta) \cdot \left(\frac{u^c - u(b)}{u'(c^w)} + \eta \cdot e^c \right)
\end{aligned}
$$

Inserting the free-entry condition again, we get the wage equation from Appendix 3.D for STW.

$$
\begin{aligned}
&(1-p)(1-\rho^u)(1-\eta) \cdot \frac{u(c^w)-u(b)}{u'(c^w)} + \eta \cdot c^w \\
&= \eta \cdot \left[(1-\rho)(z-\Omega) - (1-\rho) \cdot \frac{1-n}{n} \cdot b + (1-\rho) \cdot \theta \cdot k_v \right] \\
&\quad - p(1-\rho^c)(1-\eta) \cdot \left(\frac{u^c - u(b)}{u'(c^w)} + \eta \cdot e^c \right)
\end{aligned}
$$

Appendix 3.D The full Ramsey Problems

3.D.1 Lay-off Tax

The Ramsey planner's problem reads:

$$
\begin{aligned}
\max_{b, \tau_{\text{stw}}, \epsilon_{\text{stw}}} \; & n^u \cdot u(c^w) + n^c \cdot u^c + (1-n) \cdot u(b) \\
& + v^f \cdot u((n \cdot z - n^u \cdot c^w - n^c \cdot e^c - \tau(b) - \theta \cdot (1-n) \cdot k_v)/v^f)
\end{aligned}
$$

subject to the following constraints:

(I)Number of unconstrained workers:

$$
n^u = \frac{1-p}{\lambda} \cdot (1-\rho^u) \cdot n^s
$$

(II)Separation rate, unconstrained workers:

$$
\rho^u = G(\epsilon_s^u)
$$

(III)Number of constrained workers:

$$
n^c = \frac{p}{\lambda} \cdot (1-\rho^c) \cdot n^s
$$

(IV)Separation rate, constrained workers:

$$\rho^c = G(\epsilon_s^c)$$

(V)Aggregate separation rate:

$$\rho = (1-p) \cdot \rho^u + p \cdot \rho^c$$

(VI)Number of firms that are hit by a shock:

$$n^s = \theta \cdot q(\theta) \cdot (1-n) + \lambda \cdot n$$

(VII)Total employment:

$$n = n^u + n^c$$

(VIII)Average utility of constrained worker:

$$u^c = \frac{1}{1-\rho^c}\left((1-G(\epsilon^p)) \cdot u(c^w) + \int_{\epsilon_s^c}^{\epsilon^p} u(c(\epsilon))\, dG(\epsilon)\right)$$

(IX)Average cost of a constrained worker for a firm:

$$e^c = \frac{1}{1-\rho^c}\left[(1-G(\epsilon^p)) \cdot c^w + \int_{\epsilon_s^c}^{\epsilon^p} c^w(\epsilon)\, dG(\epsilon)\right]$$

(X)Average production (without distortions):

$$z = \frac{1}{1-\rho}\left((1-p)\int_{\epsilon_s^u}^{\infty} z(\epsilon)\, dG(\epsilon) + p\int_{\epsilon_s^c}^{\infty} z(\epsilon)\, dG(\epsilon)\right)$$

(XI)Job-creation condition:

$$\begin{aligned}\frac{1}{1-\eta} \cdot \frac{k_v}{q} = \frac{1}{\lambda}\Bigg[&(1-\rho) \cdot \left(z - \frac{1-n}{n} \cdot b\right) \\ &+ (1-p)(1-\rho^u) \cdot \left(\frac{u(c^w)-u(b)}{u'(c^w)} - c^w\right) \\ &+ (1-p)(1-\rho^c) \cdot \left(\frac{u^c-u(b)}{u'(c^w)} - e^c\right) \\ &+ (1-\rho) \cdot \frac{\lambda - f \cdot \eta}{1-\eta} \cdot \frac{k_v}{q}\Bigg]\end{aligned}$$

(XII)Wage:

$$\begin{aligned} WE =&(1-p)\cdot(1-\rho^u)\cdot\left(\eta\cdot c^w+(1-n)\cdot\frac{u(c^w)-u(b)}{u'(c^w)}\right)\\ &-\eta\cdot(1-\rho)\cdot\left(z-\frac{1-n}{n}\cdot b+\theta\cdot\frac{k_v}{q}\right)\\ &+p\cdot(1-\rho^c)\cdot\left[(1-n)\cdot\frac{u^c-u(b)}{u'(c^w)}+\eta\cdot e^c\right] = 0 \end{aligned}$$

(XIII)Separation condition unconstrained firm:

$$z(\epsilon_s^u)+F+\frac{u(c^w)-u(b)}{u'(c^w)}-c^w+\frac{\lambda-\eta\cdot f}{1-\eta}\cdot\frac{k_v}{q} = 0$$

(XIV)Threshold at which the constraint is binding:

$$z(\epsilon^p)+\lambda\cdot\frac{k_v}{q} = c^w$$

(XV)Separation condition constrained firm:

$$\frac{u(c^w(\epsilon_s^c))-u(b)}{u'(c^w)}+(\lambda-f)\cdot\frac{\eta}{1-\eta}\cdot\frac{k_v}{q} = 0$$

(XVI)Consumption of constrained worker:

$$c^w(\epsilon) = z(\epsilon)+\lambda\cdot\frac{k_v}{q}$$

3.D.2 Short-time Work

The Ramsey planner's problem reads:

$$\begin{aligned} \max_{b,\tau_{\text{stw}},\epsilon_{\text{stw}}}\ & n^u\cdot u(c^w)+n^c\cdot u^c+(1-n)\cdot u(b)\\ &+v^f\cdot u((n\cdot z-n^u\cdot c^w-n^c\cdot e^c-\tau(b)-\theta\cdot(1-n)\cdot k_v)/v^f) \end{aligned}$$

subject to the following constraints:

(I)Number of unconstrained workers:

$$n^u = \frac{1-p}{\lambda}\cdot(1-\rho^u)\cdot n^s$$

(II)Separation rate (unconstrained workers):

$$\rho^u = G(\epsilon_s^u)$$

(III)Number of constrained workers:

$$n^c = \frac{p}{\lambda} \cdot (1-\rho^c) \cdot n^s$$

(IV)Separation rate (constrained workers):

$$\rho^c = G\big(\max\{\xi_s^c, \epsilon_{\text{stw}}\}\big) - G\big(\max\{\epsilon_{\text{stw}}, \epsilon_s^c\}\big) + G(\xi_s^c)$$

(V)Aggregate separation rate:

$$\rho = (1-p) \cdot \rho^u + p \cdot \rho^c$$

(VI)Number of firms that are hit by a shock:

$$n^s = \theta \cdot q(\theta) \cdot (1-n) + \lambda \cdot n$$

(VII)Total employment:

$$n = n^u + n^c$$

(VIII)Average utility of a constrained worker:

$$u^c = \frac{1}{1-\rho^c}\Bigg((1-G(\epsilon^p)) \cdot u(c^w) + \int_{\max\{\epsilon_{\text{stw}},\xi_s^c\}}^{\epsilon_p} u(c^w(\epsilon))\,dG(\epsilon) + \int_{\epsilon_s^c}^{\max\{\epsilon_{\text{stw}},\epsilon_s^c\}} u(c_{\text{stw}}^w(\epsilon))dG(\epsilon)\Bigg)$$

(IX)Average cost of a constrained worker for a firm:

$$e^c = \frac{1}{1-\rho^c}\Bigg[(1-G(\epsilon^p)) \cdot c^w + \int_{\max\{\epsilon_{\text{stw}},\xi_s^c\}}^{\epsilon_p} c^w(\epsilon)\,dG(\epsilon) + \int_{\epsilon_s^c}^{\max\{\epsilon_{\text{stw}},\epsilon_s^c\}} \big(c_{\text{stw}}^w(\epsilon) - \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon))\big)\,dG(\epsilon)\Bigg]$$

(X)Average production (without distortions):

$$z = \frac{1}{1-\rho}\Bigg((1-p)\int_{\xi_s^u}^{\infty} z(\epsilon)\,dG(\epsilon) + p\int_{\max\{\xi_s^c,\epsilon_{\text{stw}}\}}^{\infty} z(\epsilon)\,dG(\epsilon) + \int_{\epsilon_s^c}^{\max\{\epsilon_s^c,\epsilon_{\text{stw}}\}} z(\epsilon)\,dG(\epsilon)\Bigg)$$

(XI)Average distortion of working hours:

$$\Omega = \frac{1}{1-\rho}\Bigg((1-p)\int_{\epsilon_s^u}^{\epsilon_{\text{stw}}} \Omega(\epsilon)\,dG(\epsilon) + p\int_{\epsilon_s^c}^{\max\{\epsilon_{\text{stw}},\epsilon_s^c\}} \Omega(\epsilon)\,dG(\epsilon)\Bigg)$$

$$\text{with} \quad \Omega(\epsilon) = \bar{z}(\epsilon) - z_{\text{stw}}(\epsilon)$$

(XII)Job-creation condition:

$$\frac{1}{1-\eta}\cdot\frac{k_v}{q} = \frac{1}{\lambda}\Bigg[(1-\rho)(z-\Omega-\frac{1-n}{n}b) + p(1-\rho^u)\left(\frac{u(c^w)-u(b)}{u'(c^w)} - c^w\right) + (1-p)(1-\rho^c)\left(\frac{u^c-u(b)}{u'(c^w)} - e^c\right) + (1-\rho)\cdot\frac{\lambda-\eta\cdot f}{\eta}\cdot\frac{k_v}{q}\Bigg]$$

(XIII)Wage:

$$WE = (1-p)\cdot(1-\rho^u)\left(\eta\cdot c^w + (1-\eta)\cdot\frac{u(c^w)-u(b)}{u'(c^w)}\right) - \eta(1-\rho)\left(z - \frac{1-n}{n}b + \theta\cdot k_v\right) + p\cdot(1-\rho^c)\left[(1-\eta)\cdot\frac{u^c-u(b)}{u'(c^w)} + \eta\cdot e^c\right] = 0$$

(XIV)Separation condition for unconstrained firm without STW:

$$z(\xi_s^u) + \frac{u(c^w)-u(b)}{u'(c^w)} - c^w + \frac{\lambda-\eta\cdot f}{1-\eta}\cdot\frac{k_v}{q} = 0$$

(XV)Separation condition for unconstrained firm with STW:

$$z_{stw}(\epsilon_s^u) + \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon_s^u)) + \frac{u(c^w)-u(b)}{u'(c^w)} - c^w + \frac{\lambda-\eta\cdot f}{1-\eta}\cdot\frac{k_v}{q} = 0$$

(XVI)Threshold at which the constraint is binding:

$$z(\epsilon_p) + \lambda \cdot \frac{k_v}{q} = c^w$$

(XVII)Separation condition, constrained firm without STW:

$$\frac{u(c^w(\xi_s^c)) - u(b)}{u'(c^w)} + (\lambda - f) \cdot \frac{\eta}{1-\eta} \cdot \frac{k_v}{q} = 0$$

(XVIII)Separation condition, constrained firm with STW:

$$\frac{u(c_{stw}^w(\epsilon_s^c)))}{u'(c^w)} + (\lambda - f) \cdot \frac{\eta}{1-\eta} \cdot \frac{k_v}{q} = 0$$

(XIX)Consumption, constrained worker outside STW

$$c^w(\epsilon) = z(\epsilon) + \lambda \cdot \frac{k_v}{q}$$

(XX)Consumption, constrained worker on STW

$$c_{stw}^w(\epsilon) = z_{stw}(\epsilon) + \tau_{stw}(\bar{h} - h_{stw}(\epsilon)) + \lambda \cdot \frac{k_v}{q}$$

Appendix 3.E Ramsey FOCs with Lay-off Tax

In the following, multipliers from the Lagrangian, implied by the maximization problem from the previous section, are denoted by λ_{idx}, the index depending on the constraint. Here, λ_n denotes the Lagrange multiplier for the total employment equation, λ_{n^u} for the number of unconstrained firms, λ_{n^c} for the number of constrained firms, λ_{n^s} for the number of firms that received a shock, λ_θ for the job-creation condition, λ_c for the wage equation, $\lambda_{\epsilon_s^u}$ for the separation condition of unconstrained firms and $\lambda_{\epsilon_s^c}$ for the separation condition of constrained firms. Every other equation listed in the Ramsey problem for the lay-off tax is assumed to be plugged in.

3.E.1 Employment (lay-off tax)

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial n^s} &= -\lambda_{n^s} + \frac{1-p}{\lambda} \cdot (1-\rho^u) \cdot \big(u(c^w) - u'(c^w) \cdot c^u\big) \\
&\quad + \frac{p}{\lambda} \cdot (1-\rho^c) \cdot \big(u(c^w) - u'(c^w) \cdot e^c\big) \\
&\quad + \frac{1-p}{\lambda} \cdot (1-\rho^u) \cdot \lambda_{n^u} + \frac{p}{\lambda} \cdot (1-\rho^c) \cdot \lambda_{n^c} = 0 \\
\frac{\partial \mathcal{L}}{\partial n^u} &= -\lambda_{n^u} + \lambda_n = 0 \\
\frac{\partial \mathcal{L}}{\partial n^c} &= -\lambda_{n^c} + \lambda_n = 0 \\
\frac{\partial \mathcal{L}}{\partial n} &= -\lambda_n + \lambda_{n^s} \cdot (\lambda - q(\theta) \cdot \theta) + u'(c^w) \cdot [z + b + \theta \cdot c] - u(b) \\
&\quad + \frac{1-\rho}{\lambda} \cdot \frac{b}{n^2} \cdot \lambda_\theta + \eta(1-\rho) \cdot \frac{b}{n^2} \cdot \lambda_c = 0 \\
\Leftrightarrow \quad \frac{\lambda_n}{u'(c^w)} &= (\lambda - q(\theta) \cdot \theta) \cdot \frac{\lambda_{n^s}}{u'(c^w)} + [z + b + \theta \cdot c] - \frac{u(b)}{u'(c^u)} \\
&\quad + \frac{1-\rho}{\lambda} \cdot \frac{b}{n^2} \cdot \frac{\lambda_\theta}{u'(c^w)} + \frac{\lambda \cdot b}{n^2} \cdot \frac{\lambda_c}{u'(c^w)} + \eta \cdot \frac{b}{n^2} \cdot \frac{\lambda_c}{u'(c^w)}
\end{aligned}$$

3.E.2 Optimal Job Creation Condition (lay-off tax)

Before we begin, let us define the insurance effect as:

$$\begin{aligned}
\tilde{IE}_\theta &= \frac{p}{\lambda}\left(\int_{\epsilon_s^c}^{\epsilon^p} \lambda \cdot \frac{\gamma}{f} \cdot \frac{u'(c(\epsilon)) - u'(c^w)}{u'(c^w)}\, dG(\epsilon)\right) \cdot k_v \\
IE_\theta &= n^s \cdot u'(c^w) \cdot \left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right) \cdot \tilde{IE}_\theta
\end{aligned}$$

The FOC for labor market tightness denotes:

$$
\begin{aligned}
\frac{\partial \mathscr{L}}{\partial \theta} &= -k_v(1-n)u'(c^w) + IE_\theta \cdot k_v + (1-\gamma)q(\theta)(1-n)\lambda_{n^s} \\
&\quad - \frac{\gamma}{1-\eta}k_v\lambda_\theta + \frac{\lambda\gamma - f\eta}{(1-\eta)f}\left(\frac{1-\rho}{\lambda}\lambda_\theta - \lambda_{\epsilon_s^u}\right) \\
&\quad - \left(\lambda \cdot \frac{\gamma}{f}u'(c(\epsilon_s^u)) + \frac{\lambda\gamma - f}{(1-\eta)f}\eta\right)\lambda_{\epsilon_s^c} \\
&\quad - \lambda_c \cdot \frac{\partial WE}{\partial \theta} \\
\Leftrightarrow \quad \frac{\lambda_{n^s}}{u'(c^w)} &= \frac{1+\chi}{1-\eta} \cdot \frac{k_v}{q}
\end{aligned}
$$

with

$$
\begin{aligned}
\chi = \frac{1}{u \cdot u'(c^w)} \cdot \Bigg[& \frac{1}{f} \cdot \frac{1}{1-\eta} \cdot \left(\gamma\lambda_\theta - (\lambda\gamma - f\eta)\left(\frac{1-\rho}{\lambda}\lambda_\theta - \lambda_{\epsilon_s^u}\right)\right) \\
& + \left(\lambda \cdot \frac{\gamma}{f} \cdot u'(c^w(\epsilon_s^u)) + \frac{\lambda\gamma - f}{(1-\eta)f} \cdot \eta\right)\lambda_{\epsilon_s^c} \\
& + \frac{1}{k_v} \cdot \left(\lambda_c \cdot \frac{\partial WE}{\partial \theta} - IE_\theta\right)\Bigg]
\end{aligned}
$$

Using the FOCs for employment and labor market tightness gives:

$$
\begin{aligned}
u'(c^w) \cdot \frac{1+\chi}{1-\gamma} \cdot \frac{k_v}{q} &= \frac{1-p}{\lambda} \cdot (1-\rho^u) \cdot \left(u(c^w) - u(b) - u'(c^w) \cdot c^w\right) \\
&\quad + \frac{p}{\lambda} \cdot (1-\rho^c) \cdot \left(u^c - u(b) - u'(c^w) \cdot e^c\right) \\
&\quad + u'(c^w) \cdot \frac{1-\rho}{\lambda} \cdot \left(z + b + \frac{\lambda - \gamma f + \chi(\lambda - f)}{1-\gamma} \cdot \frac{k_v}{q}\right) \\
&\quad + \frac{1-\rho}{\lambda} \cdot \frac{\lambda_\theta}{n^s} \cdot \frac{b}{n} + \frac{1-\rho}{n^s} \cdot \frac{b}{n} \cdot \lambda_c
\end{aligned}
$$

Rearranging gives the *Optimal Job Creation Condition:*

$$
\begin{aligned}
\frac{1+\chi}{1-\gamma} \cdot \frac{k_v}{q} &= \frac{1-\rho}{\lambda} \cdot (z+b) + \frac{1-\rho}{\lambda} \cdot \frac{\lambda_\theta}{n^s \cdot u'(c^w)} \cdot \frac{b}{n} + \frac{1-\rho}{\lambda} \cdot \frac{\lambda \cdot \lambda_c}{n^s \cdot u'(c^w)} \cdot \frac{b}{n} \\
&\quad + \frac{1-p}{\lambda} \cdot (1-\rho^u) \cdot \left(\frac{u(c^w) - u(b)}{u'(c^w)} - c^w\right) \\
&\quad + \frac{p}{\lambda} \cdot (1-\rho^c) \cdot \left(\frac{u^c - u(b)}{u'(c^w)} - e^c\right) \\
&\quad + \frac{1-\rho}{\lambda} \cdot \frac{\lambda - \gamma f + \chi \cdot (\lambda - f)}{1-\gamma} \cdot \frac{k_v}{q}
\end{aligned}
$$

Subtracting the decentralized job-creation condition from the optimal gives:

$$\left(\chi - \frac{\eta-\gamma}{1-\eta}\right)\cdot\frac{1}{1-\gamma}\cdot\frac{k_v}{q} = \left(1+\frac{\lambda_\theta+\eta+\lambda\cdot\lambda_c}{n^s\cdot u'(c^w)}\right)\cdot\frac{1-\rho}{\lambda}\cdot\frac{b}{n} + \frac{(1-\rho)\cdot(\lambda-f)}{\lambda}\cdot\left(\chi-\frac{\eta-\gamma}{1-\gamma}\right)\cdot\frac{k_v}{q}$$

Rearranging gives:

$$\left(\chi - \frac{\eta-\gamma}{1-\eta}\right)\cdot\frac{1}{1-\gamma}\cdot\frac{k_v}{q} = \left(1+\frac{\lambda_\theta+\eta\cdot\lambda\cdot\lambda_c}{n^s\cdot u'(c^w)}\right)\cdot\frac{\frac{1-\rho}{\lambda}}{\rho+(1-\rho)\cdot\frac{f}{\lambda}}\cdot\frac{b}{n} = \left(1+\frac{\lambda_\theta+\eta\cdot\lambda\cdot\lambda_c}{n^s\cdot u'(c^w)}\right)\cdot\frac{b}{f}$$

3.E.3 Optimal separation condition (unconstrained firms)

FOC for the separation threshold of unconstrained firms:

$$-\frac{\partial\mathscr{L}}{\partial\epsilon_s^u} = \frac{n}{1-\rho}\cdot(1-p)\cdot g(\epsilon_s^u)\cdot u'(c^w)\cdot\left(z(\epsilon_s^u)+\frac{u(c^w)}{u'(c^w)}-c^w-z\right) + \lambda_\theta\cdot\frac{1-p}{\lambda}\cdot g(\epsilon_s^u)\cdot\left[z(\epsilon_s^u)+\frac{u(c^w)-u(b)}{u'(c^w)}-c^w-\frac{1-n}{n}\cdot b+\frac{\lambda-\eta f}{1-\eta}\cdot\frac{k_v}{q}\right] + \lambda_c\cdot\frac{\partial WE}{\partial\epsilon_s^u}+\frac{1-p}{\lambda}\cdot g(\epsilon_s^u)\cdot n^s\cdot\lambda_{n^u}+\lambda_{\epsilon_s^u}\cdot\frac{\partial S_{\mathrm{lt}}^u(\epsilon_s^u)}{\partial\epsilon_s^u} = 0$$

Insert for λ_{n^u} and subtract decentralized separation condition of unconstrained firms:

$$g(\epsilon_s^u)\cdot\frac{1-p}{\lambda}\cdot n^s\cdot\left[z(\epsilon_s^u)+\frac{u(c^w)}{u'(c^w)}-c^w+\frac{\lambda_n}{u'(c^w)}\right] - \frac{\lambda_\theta}{u'(c^w)}\cdot\frac{1-p}{\lambda}\cdot g(\epsilon_s^u)\cdot\left[\lambda F+\frac{1-n}{n}\cdot b\right] + \frac{\lambda_c}{u'(c^w)}\cdot\frac{\partial WE}{\partial\epsilon_s^u}+\frac{\lambda_{\epsilon_s^u}}{u'(c^w)}\cdot\frac{\partial S_{\mathrm{lt}}^u(\epsilon_s^u)}{\partial\epsilon_s^u} = 0$$

Insert for λ_n:

$$g(\epsilon_s^u)\cdot\frac{1-p}{\lambda}\cdot n^s\left[z(\epsilon_s^u)+\frac{u(c^w)}{u'(c^w)}-c^w-z\right] + g(\epsilon_s^u)\cdot\frac{1-p}{\lambda}\cdot n^s\left[(\lambda-q(\theta)\cdot\theta)\cdot\frac{\lambda_{n^s}}{u'(c^w)}+(z+b+\theta\cdot c)-\frac{u(b)}{u'(c^w)} + \frac{1-\rho}{\lambda}\cdot\frac{b}{n^2}\cdot\frac{\lambda_\theta}{u'(c^w)}+\eta\cdot\frac{b}{n^2}\cdot\frac{\lambda_c}{u'(c^w)}\right] + \frac{\lambda_\theta}{u'(c^w)}\cdot\frac{1-p}{\lambda}\cdot g(\epsilon_s^u)\left[\lambda F+\frac{1-n}{n}\cdot b\right] + \frac{\lambda_c}{u'(c^w)}\cdot\frac{\partial WE}{\partial\epsilon_s^u}+\frac{\lambda_{\epsilon_s^u}}{u'(c^w)}\cdot\frac{\partial S_{\mathrm{stw}}^u(\epsilon_s^u)}{\partial\epsilon_s^u} = 0$$

$$\begin{aligned}\iff \quad & z(\epsilon_s^u) + \frac{u(c^w) - u(b)}{u'(c^w)} - c^w + b + \big(\lambda - q(\theta)\cdot\theta\big)\cdot\frac{\lambda_{n^s}}{u'(c^w)} + \theta\cdot c \\ & - \frac{\lambda_\theta}{n^s\cdot u'(c^w)}\left(\lambda F + \frac{1-n}{n}\cdot b - \frac{b}{n}\right) \\ & + \frac{\lambda_c}{n^s\cdot u'(c^w)}\left(\eta\cdot\frac{b}{n} + \frac{\lambda}{1-p}\cdot\frac{1}{g(\epsilon_s^u)}\cdot\frac{\partial WE}{\partial\epsilon_s^u}\right) \\ & + \frac{\lambda_{\epsilon_s^u}}{n^s\cdot u'(c^w)}\cdot\frac{\lambda}{1-p}\cdot\frac{1}{g(\epsilon_s^u)}\cdot\frac{\partial S_{\mathrm{lt}}^u(\epsilon_s^u)}{\partial\epsilon_s^u} \;=\; 0\end{aligned}$$

Inserting λ_n^s gives the *Optimal Separation Condition*:

$$\begin{aligned} & z(\epsilon_s^u) + \frac{u(c^w) - u(b)}{u'(c^w)} - c^w + b + \frac{\lambda - \gamma\cdot f + \chi\cdot(\lambda - f)}{1-\gamma}\cdot\frac{k_v}{q} \\ & - \frac{\lambda_\theta}{n^s\cdot u'(c^w)}\cdot(\lambda\cdot F - b) + \frac{\lambda_c}{n^s\cdot u'(c^w)}\cdot\left(\eta\cdot\frac{b}{n} + \frac{\lambda}{1-p}\cdot\frac{1}{g(\epsilon_s^u)}\cdot\frac{\partial WE}{\partial\epsilon_s^u}\right) \;=\; 0\end{aligned}$$

Subtracting the decentralized separation condition and rearranging gives:

$$\begin{aligned}\lambda\cdot F \;=\; & b + (\lambda - \theta q(\theta))\cdot\left(\chi - \frac{\eta-\gamma}{1-\gamma}\right)\cdot\frac{k_v}{q} \\ & + \frac{\lambda_{\epsilon_s^u}}{n^s\cdot u'(c^w)}\cdot\frac{\lambda}{1-p}\cdot\frac{1}{g(\epsilon_s^u)}\cdot\frac{\partial S_{\mathrm{lt}}^u(\epsilon_s^u)}{\partial\epsilon_s^u} \\ & - \frac{\lambda_\theta}{n^s\cdot u'(c^w)}\cdot(\lambda F - b) \\ & + \frac{\lambda_c}{n^s\cdot u'(c^w)}\left(\eta\cdot\frac{b}{n} + \frac{\lambda}{1-p}\cdot\frac{1}{g(\epsilon_s^u)}\cdot\frac{\partial WE}{\partial\epsilon_s^u}\right)\end{aligned}$$

$$\begin{aligned}\iff \quad 0 \;=\; & \left(1 + \frac{\lambda_\theta}{u'(c^w)\cdot n^s}\right)\cdot\left(b + (\lambda - f)\cdot\frac{\frac{1-\rho}{\lambda}}{\rho + (1-\rho)\cdot\frac{f}{\lambda}}\cdot\frac{b}{n} - \lambda\cdot F\right) \\ & + \frac{\lambda_c}{n^s\cdot u'(c^w)}\left(\eta\cdot\frac{b}{n} + \frac{\lambda}{1-p}\cdot\frac{1}{g(\epsilon_s^u)}\cdot\frac{\partial WE}{\partial\epsilon_s^u}\right) \\ & + \frac{\lambda}{1-p}\cdot\frac{1}{g(\epsilon_s^u)}\cdot\frac{\lambda_{\epsilon_s^u}}{n^s\cdot u'(c^w)}\cdot\frac{\partial S_{\mathrm{lt}}^u(\epsilon_s^u)}{\partial\epsilon_s^u}\end{aligned}$$

$$\begin{aligned}
\iff \quad 0 \;=\; & \left(1 + \frac{\lambda_\theta}{u'(c^w)\cdot n^s}\right)\cdot\left(b + (\lambda - f)\cdot\frac{b}{f} - \lambda F\right) \\
& + \frac{\lambda}{1-p}\cdot\frac{1}{g(\epsilon_s^u)}\cdot\frac{\lambda_{\epsilon_s^u}}{n^s\cdot u'(c^w)}\frac{\partial S^u_{\text{stw}}(\epsilon_s^u)}{\partial\epsilon_s^u} \\
& + \frac{\lambda_c}{n^s\cdot u'(c^w)}\cdot\left(\eta\cdot\frac{b}{n} + \frac{\lambda}{1-p}\cdot\frac{1}{g(\epsilon_s^u)}\cdot\frac{\partial WE}{\partial\epsilon_s^u}\right)
\end{aligned}$$

$$\begin{aligned}
\iff \quad 0 \;=\; & \left(1 + \frac{\lambda_\theta}{u'(c^w)\cdot n^s}\right)\cdot\left(\frac{\lambda}{f}\cdot b - \lambda F\right) \\
& + \frac{\lambda}{1-p}\cdot\frac{1}{g(\epsilon_s^u)}\cdot\frac{\lambda_{\epsilon_s^u}}{n^s\cdot u'(c^w)}\cdot\frac{\partial S^u_{\text{lt}}(\epsilon_s^u)}{\partial\epsilon_s^u} \\
& + \frac{\lambda_c}{n^s\cdot u'(c^w)}\cdot\left(\eta\cdot\frac{b}{n} + \frac{\lambda}{1-p}\cdot\frac{1}{g(\epsilon_s^u)}\cdot\frac{\partial WE}{\partial\epsilon_s^u}\right)
\end{aligned}$$

Rearranging gives the Lagrange multiplier for unconstrained firms:

$$\begin{aligned}
-\frac{\lambda_{\epsilon_s^u}}{n^s\cdot u'(c^w)} \;=\; & \frac{1-p}{\lambda}\cdot\frac{g(\epsilon_s^u)}{\frac{\partial S^u_{\text{stw}}(\epsilon_s^u)}{\partial\epsilon_s^u}}\cdot\left(\left(1 + \frac{\lambda_\theta}{n^s\cdot u'(c^w)}\right)\cdot\left(\frac{\lambda}{f}\cdot b - \lambda F\right)\right. \\
& \left. + \frac{\lambda_c}{n^s\cdot u'(c^w)}\cdot\left(\eta\cdot\frac{b}{n} + \frac{1}{g(\epsilon_s^u)}\cdot\frac{\lambda}{1-p}\cdot\frac{1}{g(\epsilon_s^u)}\cdot\frac{\partial WE}{\partial\epsilon_s^u}\right)\right)
\end{aligned}$$

3.E.4 Optimal separation condition (constrained firms)

The FOC for the separation threshold in the constrained firm is:

$$\begin{aligned}
-\frac{\partial\mathcal{L}}{\partial\epsilon_s^c} \;=\; & \frac{n}{1-\rho}\cdot p\cdot g(\epsilon_s^c)\cdot u'(c^w)\left(z(\epsilon_s^c) + \frac{u(c^w(\epsilon_s^c)) - u(c^w)}{u'(c^w)} - c^w(\epsilon_s^c) - z\right) \\
& + \lambda_\theta\cdot\frac{p}{\lambda}\cdot g(\epsilon_s^c)\left[z(\epsilon_s^c) + \frac{c^w(\epsilon_s^c) - u(b)}{u'(c^w)} - c^w(\epsilon_s^c) - \frac{1-n}{n}\cdot b + \frac{\lambda-\eta f}{1-\eta}\cdot\frac{k_v}{q}\right] \\
& + \frac{p}{\lambda}\cdot g(\epsilon_s^c)\cdot n^s\cdot\lambda_{n^c} \\
& + \lambda_c\cdot\frac{\partial WE}{\partial\epsilon_s^c} \\
& + \lambda_{\epsilon_s^c}\cdot\frac{\partial S^c_{\text{lt}}(\epsilon_s^c)}{\partial\epsilon_s^c} \;=\; 0
\end{aligned}$$

Note that we can reformulate the decentralized separation threshold of a worker in a constrained firm as:

$$
\begin{aligned}
& \frac{u(c^w(\epsilon_s^c)) - u(b)}{u'(c^w)} + (\lambda - f) \cdot \frac{\eta}{1-\eta} \cdot \frac{k_v}{q} = 0 \\
\iff \quad & \frac{u(c^w(\epsilon_s^c)) - u(b)}{u'(c^w)} + (\lambda - f) \cdot \frac{\eta}{1-\eta} \cdot \frac{k_v}{q} = c^w(\epsilon_s^c) - z(\epsilon_s^c) - \lambda \cdot \frac{k_v}{q} \\
\iff \quad & z(\epsilon_s^c) + \frac{u(c^w(\epsilon_s^c)) - u(b)}{u'(c^w)} - c^w(\epsilon_s^c) + \frac{\lambda - \eta f}{1-\eta} \cdot \frac{k_v}{q} = 0
\end{aligned}
$$

This simplifies the equation to:

$$
\begin{aligned}
-\frac{\partial \mathscr{L}}{\partial \epsilon_s^c} = & \frac{n}{1-\rho} \cdot p \cdot g(\epsilon_s^c) \cdot \left(z(\epsilon_s^c) + \frac{u(c^w(\epsilon_s^c))}{u'(c^w)} - c^w(\epsilon_s^c) - z \right) \\
& + \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot n^s \cdot \frac{\lambda_{n^c}}{u'(c^w)} \\
& - \frac{\lambda_\theta}{u'(c^w)} \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot \frac{1-n}{n} \cdot b \\
& + \frac{\lambda_c}{u'(c^w)} \cdot \frac{\partial WE}{\partial \epsilon_s^c} + \frac{\lambda_{\epsilon_s^c}}{u'(c^w)} \cdot \frac{\partial S^c(\epsilon_s^c)}{\partial \epsilon_s^c} \overset{!}{=} 0
\end{aligned}
$$

Inserting λ_n^s gives the *Optimal Separation Condition*

$$
\begin{aligned}
-\frac{\partial \mathscr{L}}{\partial \epsilon_s^c} = & z(\epsilon_s^c) + \frac{u(c^w(\epsilon_s^c))}{u'(c^w)} - c^w(\epsilon_s^c) + b + \frac{\lambda - \gamma \cdot f + \chi \cdot (\lambda - f)}{1-\gamma} \cdot \frac{k_v}{q} \\
& + \frac{\lambda_\theta}{n^s \cdot u'(c^w)} \cdot b \\
& + \frac{\lambda_c}{n^s \cdot u'(c^w)} \left(\eta \cdot \frac{b}{n} + \frac{1}{g(\epsilon_s^c)} \cdot \frac{\lambda}{p} \cdot \frac{\partial WE}{\partial \epsilon_s^c} \right) \\
& + \frac{\lambda}{p} \cdot \frac{1}{n^s \cdot g(\epsilon_s^c)} \cdot \frac{\lambda_{\epsilon_s^c}}{u'(c^w)} \cdot \frac{\partial S_{\mathrm{lt}}^c(\epsilon_s^c)}{\partial \epsilon_s^c} = 0
\end{aligned}
$$

Subtracting the decentralized separation condition gives:

$$
\begin{aligned}
0 = & b + (\lambda - \theta q(\theta)) \cdot \left(\chi - \frac{\eta - \gamma}{1-\eta} \right) \cdot \frac{1}{1-\gamma} \cdot \frac{k_v}{q} \\
& + \frac{\lambda \cdot \epsilon_s^c}{n^s \cdot u'(c^w)} \cdot \frac{1}{\rho} \cdot \frac{1}{\lambda} \cdot \frac{\partial S_{\mathrm{lt}}^c(\epsilon_s^c)}{\partial \epsilon_s^c} \cdot \frac{1}{g(\epsilon_s^c)} \\
& + \frac{\lambda \cdot \theta}{n^s \cdot u'(c^w)} \cdot b + \frac{\lambda_c}{n^s \cdot u'(c^w)} \left(\eta \cdot \frac{b}{n} + \frac{1}{g(\epsilon_s^c)} \cdot \frac{\lambda}{p} \cdot \frac{\partial WE}{\partial \epsilon_s^c} \right)
\end{aligned}
$$

Rearranging gives back the Lagrange multiplier:

$$-\frac{\lambda_{\epsilon_s^c}}{n^s \cdot u'(c^w)} = \frac{p}{\lambda} \cdot \frac{g(\epsilon_s^c)}{\frac{\partial S_{\text{stw}}^c(\epsilon_s^c)}{\partial \epsilon_s^c}} \cdot \left(\left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right) \cdot \frac{\lambda}{f} \cdot b + \frac{\lambda_c}{n^s \cdot u'(c^w)} \cdot \left(\eta \cdot \frac{b}{n} + \frac{1}{g(\epsilon_s^c)} \cdot \frac{\lambda}{p} \cdot \frac{\partial WE}{\partial \epsilon_s^c}\right)\right)$$

3.E.5 The Optimal Lay-Off Tax

We start from

$$\frac{\partial \mathscr{L}}{\partial F} = -\lambda_{\epsilon_s^u} = 0$$

Note that the optimal lay-off tax sets the Lagrange multiplier for the separation condition equal to zero. This implies that the lay-off tax implements the optimal separation threshold for unconstrained firms. Then

$$-\frac{\lambda_{\epsilon_s^u}}{n^s \cdot u'(c^w)} = \frac{p}{\lambda} \cdot \frac{g(\epsilon_s^u)}{\frac{\partial S_{\text{lt}}^u(\epsilon_s^u)}{\partial \epsilon_s^u}} \cdot \left(\left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right) \cdot \left(\frac{\lambda}{f} \cdot b - \lambda \cdot F\right) + \frac{\lambda_c}{n^s \cdot u'(c^w)} \left(\eta \cdot \frac{b}{n} + \frac{\lambda}{g(\epsilon_s^u)} \cdot \frac{\lambda}{1-p} \cdot \frac{\partial WE}{\partial \epsilon_s^u}\right)\right)$$

This leads to:

$$F = \frac{1}{f} \cdot b + BE_{lt}$$

$$BE_{lt} = \frac{1}{\lambda} \frac{1}{\left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right)} \cdot \frac{\lambda_c}{n^s \cdot u'(c^w)} \cdot \left(\eta \cdot \frac{b}{n} + \frac{1}{g(\epsilon_s^u)} \cdot \frac{\lambda}{1-p} \cdot \frac{\partial WE}{\partial \epsilon_s^u}\right)$$

3.E.6 Optimal UI given Lay-off Tax

From the FOC of UI, we can derive the optimality condition for unemployment insurance. With an optimal lay-off tax, the Lagrange multiplier for the separation condition of the unconstrained firm is equal to zero.

$$(1-\eta) \cdot \left(u'(b) - u'(c^w)\right) = \lambda_\theta \cdot (1-\rho) \cdot \left(\frac{1-n}{n} + \frac{u'(b)}{u'(c^w)}\right) + \left(\lambda_{\epsilon_s^c} + \lambda_{\xi_s^c}\right) \cdot \left(-\frac{u'(b)}{u'(c^w)}\right) + \lambda_c \cdot (1-\rho) \cdot \left(\eta \cdot \frac{1-n}{n} + (1-\eta) \cdot \frac{u'(b)}{u'(c^w)}\right)$$

To get better insight, we need to determine the Lagrange multipliers for λ_θ, λ_c. We already know the Lagrange multipliers for the separation conditions of constrained firms:

$$-\frac{\lambda_{\epsilon_s^u}}{n^s \cdot u'(c^w)} = \frac{1-p}{\lambda} \cdot g(\epsilon_s^u) \cdot \frac{1}{\frac{\partial S_{\text{stw}}^u(\epsilon_s^u)}{\partial \epsilon_s^u}} \cdot \left[\left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right) \cdot \left(\frac{\lambda}{f} \cdot b - \lambda \cdot F\right) + \frac{\lambda_c}{n^s \cdot u'(c^w)} \cdot \left(\eta \cdot \frac{b}{n} + \frac{1}{g(\epsilon_s^u)} \cdot \frac{\lambda}{1-p} \cdot \frac{\partial WE}{\partial \epsilon_s^u}\right)\right]$$

$$-\frac{\lambda_{\epsilon_s^c}}{n^s \cdot u'(c^w)} = \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot \frac{1}{\frac{\partial S_{\text{stw}}^c(\epsilon_s^c)}{\partial \epsilon_s^c}} \cdot \left[\left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right) \cdot \frac{\lambda}{f} \cdot b + \frac{\lambda_c}{n^s \cdot u'(c^w)} \cdot \left(\eta \cdot \frac{b}{n} + \frac{1}{g(\epsilon_s^c)} \cdot \frac{\lambda}{p} \cdot \frac{\partial WE}{\partial \epsilon_s^c}\right)\right]$$

First, let us find an expression for λ_c:

$$\begin{aligned}\frac{\partial \mathscr{L}}{\partial c^w} = &-\Bigg((1-p) \cdot (1-\rho^u) \cdot \frac{u(c^w) - u(b)}{u'(c^w)} \cdot \frac{u''(c^w)}{u'(c^w)} \\ &+ p \cdot (1-\rho^c) \cdot \frac{u^c - u(b)}{u'(c^w)} \cdot \frac{u''(c^w)}{u'(c^w)}\Bigg) \cdot \lambda_\theta \\ &- (1-\eta) \cdot \Bigg((1-p) \cdot (1-\rho^u) \cdot \left(1 - (1-\eta) \cdot \frac{u(c^w) - u(b)}{u'(c^w)} \cdot \frac{u''(c^w)}{u'(c^w)}\right) \\ &- p \cdot (1-\rho^c) \cdot (1-\eta) \cdot \frac{u^c - u(b)}{u'(c^w)} \cdot \frac{u''(c^w)}{u'(c^w)}\Bigg) \cdot \lambda_c \\ &+ \frac{u(c^w) - u(b)}{u'(c^w)} \cdot \frac{u''(c^w)}{u'(c^w)} \cdot \lambda_{\epsilon_s^u} \\ &+ \frac{u(c_{\text{stw}}(\epsilon_s^c)) - u(b)}{u'(c^w)} \cdot \frac{u''(c^w)}{u'(c^w)} \cdot \lambda_{\epsilon_s^c} = 0\end{aligned}$$

Next we can solve for the Lagrange multiplier of the wage equation:

$$\lambda_c = -\frac{1}{\left(\frac{\partial WE}{\partial c^w}\right)^{\text{total}}} \left[\frac{\partial \epsilon_s^u}{\partial c^w} \cdot \frac{1-p}{\lambda} \cdot g(\epsilon_s^u) \cdot n^s \cdot u'(c^w) \cdot \left(\frac{\lambda}{f} \cdot b - \lambda \cdot F\right) + \frac{\partial \epsilon_s^c}{\partial c^w} \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot n^s \cdot u'(c^w) \cdot \frac{\lambda}{f} \cdot b + \left(-\frac{\partial S^{\text{total}}}{\partial c^w}\right) \cdot \lambda_\theta\right]$$

Insert λ_c into the Lagrange multipliers for the separation conditions. To do that, let us rewrite the Lagrange multipliers of the separation conditions:

$$\begin{aligned}
\lambda_{\epsilon_s^u} \;=\; & -\frac{1}{\frac{\partial S^u_{\text{stw}}(\epsilon_s^u)}{\partial \epsilon_s^u}} \cdot \left(n^s \cdot u'(c^w) \cdot \frac{1-p}{\lambda} \cdot g(\epsilon_s^u) \cdot \left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right) \cdot \left(\frac{\lambda}{f} \cdot b - \lambda \cdot F\right) \right. \\
& \left. + \lambda_c \cdot \left(\frac{\partial n}{\partial \epsilon_s^u} \cdot \left(-\frac{\partial WE}{\partial n}\right) + \frac{\partial WE}{\partial \epsilon_s^u} \right) \right) \\
\lambda_{\epsilon_s^c} \;=\; & -\frac{1}{\frac{\partial S^c_{\text{stw}}(\epsilon_s^c)}{\partial \epsilon_s^c}} \cdot \left(n^s \cdot u'(c^w) \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot \left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right) \cdot \frac{\lambda}{f} \cdot b \right. \\
& \left. + \lambda_c \cdot \left(\frac{\partial n}{\partial \epsilon_s^c} \cdot \left(-\frac{\partial WE}{\partial n}\right) + \frac{\partial WE}{\partial \epsilon_s^c} \right) \right)
\end{aligned}$$

Inserting λ_c gives:

$$\begin{aligned}
\lambda_{\epsilon_s^u} \;=\; & -\frac{1}{\frac{\partial S^u_{\text{stw}}(\epsilon_s^u)}{\partial \epsilon_s^u}} \cdot \left(n^s \cdot u'(c^w) \cdot \frac{1-p}{\lambda} \cdot g(\epsilon_s^u) \cdot \left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right) \cdot \left(\frac{\lambda}{f} \cdot b - \lambda \cdot F\right) \right. \\
& + \left(\frac{\partial n}{\partial \epsilon_s^u} \cdot \left(-\frac{\partial c^w}{\partial n}\right)^{\text{total}} + \left(\frac{\partial c^w}{\partial \epsilon_s^u}\right)^{\text{total}} \right) \\
& \cdot \left[\frac{\partial \epsilon_s^u}{\partial c^w} \cdot \frac{1-p}{\lambda} \cdot g(\epsilon_s^u) \cdot n^s \cdot u'(c^w) \cdot \left(\frac{\lambda}{f} \cdot b - \lambda \cdot F\right) \right. \\
& \left.\left. + \frac{\partial \epsilon_s^c}{\partial c^w} \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot n^s \cdot u'(c^w) \cdot \frac{\lambda}{f} \cdot b + \left(-\frac{\partial S^{\text{total}}}{\partial c^w}\right) \cdot \lambda_\theta \right] \right)
\end{aligned}$$

$$\begin{aligned}
\lambda_{\epsilon_s^c} \;=\; & \frac{1}{\frac{\partial S^c_{\text{stw}}(\epsilon_s^c)}{\partial \epsilon_s^c}} \cdot \left(n^s \cdot u'(c^w) \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot \left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right) \cdot \frac{\lambda}{f} \cdot b \right. \\
& + \left(\left(\frac{\partial n}{\partial \epsilon_s^c}\right) \cdot \left(-\frac{\partial c^w}{\partial n}\right)^{\text{total}} + \left(\frac{\partial c^w}{\partial \epsilon_s^c}\right)^{\text{total}} \right) \\
& \cdot \left[\frac{\partial \epsilon_s^u}{\partial c^w} \cdot \frac{1-p}{\lambda} \cdot g(\epsilon_s^u) \cdot n^s \cdot u'(c^w) \cdot \left(\frac{\lambda}{f} \cdot b - \lambda \cdot F\right) \right. \\
& \left.\left. + \frac{\partial \epsilon_s^c}{\partial c^w} \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot n^s \cdot u'(c^w) \cdot \frac{\lambda}{f} \cdot b + \left(-\frac{\partial S^{\text{total}}}{\partial c^w}\right) \cdot \lambda_\theta \right] \right)
\end{aligned}$$

Insert λ_c into the Lagrange multipliers for the separation conditions. To do that, let us rewrite the Lagrange multipliers of the separation conditions:

$$\lambda_{\epsilon_s^u} = -\frac{1}{\frac{\partial S_{\text{stw}}^u(\epsilon_s^u)}{\partial \epsilon_s^u}} \cdot \Bigg(- n^s \cdot u'(c^w) \cdot \frac{1-p}{\lambda} \cdot g(\epsilon_s^u) \cdot \left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right) \cdot \left(\frac{\lambda}{f} \cdot b - \lambda \cdot F\right) + \lambda_c \cdot \left(\frac{\partial n}{\partial \epsilon_s^u} \cdot \left(-\frac{\partial WE}{\partial n}\right) + \frac{\partial WE}{\partial \epsilon_s^u}\right)\Bigg)$$

$$\lambda_{\epsilon_s^c} = -\frac{1}{\frac{\partial S_{\text{stw}}^c(\epsilon_s^c)}{\partial \epsilon_s^c}} \cdot \Bigg(n^s \cdot u'(c^w) \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot \left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right) \cdot \frac{\lambda}{f} \cdot b + \lambda_c \cdot \left(\frac{\partial n}{\partial \epsilon_s^c} \cdot \left(-\frac{\partial WE}{\partial n}\right) + \frac{\partial WE}{\partial \epsilon_s^c}\right)\Bigg)$$

Inserting λ_c gives:

$$\begin{aligned}\lambda_{\epsilon_s^u} = -\frac{1}{\frac{\partial S_{\text{stw}}^u(\epsilon_s^u)}{\partial \epsilon_s^u}} \cdot \Bigg(& n^s \cdot u'(c^w) \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^u) \cdot \left(\lambda + \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right) \cdot \left(\frac{\lambda}{f} \cdot b - \lambda \cdot F\right) \\ &+ \left(\frac{\partial n}{\partial \epsilon_s^u} \cdot \left(-\frac{\partial c^w}{\partial n}\right)^{\text{total}} + \left(\frac{\partial c^w}{\partial \epsilon_s^u}\right)^{\text{total}}\right) \\ &\cdot \Bigg[\frac{\partial \epsilon_s^u}{\partial c^w} \cdot \frac{1-p}{\lambda} \cdot g(\epsilon_s^u) \cdot n^s \cdot u'(c^w) \cdot \left(\frac{\lambda}{f} \cdot b - \lambda \cdot F\right) \\ &+ \frac{\partial \epsilon_s^c}{\partial c^w} \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot n^s \cdot u'(c^w) \cdot \frac{\lambda}{f} \cdot b + \left(-\frac{\partial S^{\text{total}}}{\partial c^w}\right) \cdot \lambda_\theta\Bigg]\Bigg)\end{aligned}$$

$$\begin{aligned}\lambda_{\epsilon_s^c} = -\frac{1}{\frac{\partial S_{\text{stw}}^c(\epsilon_s^c)}{\partial \epsilon_s^c}} \cdot \Bigg(& n^s \cdot u'(c^w) \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot \left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right) \cdot \frac{\lambda}{f} \cdot b \\ &+ \left(\left(\frac{\partial n}{\partial \epsilon_s^c}\right) \cdot \left(-\frac{\partial c^w}{\partial n}\right)^{\text{total}} + \left(\frac{\partial c^w}{\partial \epsilon_s^c}\right)^{\text{total}}\right) \\ &\cdot \Bigg[\frac{\partial \epsilon_s^u}{\partial c^w} \cdot \frac{1-p}{\lambda} \cdot g(\epsilon_s^u) \cdot n^s \cdot u'(c^w) \cdot \left(\frac{\lambda}{f} \cdot b - \lambda \cdot F\right) \\ &+ \frac{\partial \epsilon_s^c}{\partial c^w} \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot n^s \cdot u'(c^w) \cdot \frac{\lambda}{f} \cdot b + \left(-\frac{\partial S^{\text{total}}}{\partial c^w}\right) \cdot \lambda_\theta\Bigg]\Bigg)\end{aligned}$$

Now we have everything to calculate λ_θ from the two equations:

$$\text{(1)} \quad \left(\chi - \frac{\eta-\gamma}{1-\eta}\right)\cdot\frac{1}{1-\gamma}\cdot\frac{k_v}{q} \;=\; \left(1+\frac{\lambda_\theta+\eta\cdot\lambda\cdot\lambda_c}{n^s\cdot u'(c^w)}\right)\cdot\frac{b}{f}$$

$$\begin{aligned}\text{(2)} \quad \chi \;=\; & \frac{1}{1-\eta}\cdot\frac{1}{u\cdot f}\cdot\frac{1}{u'(c^w)}\cdot\left[\gamma-(\lambda\gamma-f\eta)\cdot\left(\frac{1}{\lambda}(1-\rho)\lambda_\theta-\lambda_{\epsilon_s^u}\right)\right]\\ &+\frac{1}{u}\cdot\frac{1}{u'(c^w)}\cdot\left(\lambda\cdot\frac{\gamma}{f}\cdot u'(c^w)+\left(\frac{1}{1-\eta}\cdot\frac{\lambda\gamma-f}{f}\right)\cdot\eta\right)\cdot\lambda_{\epsilon_s^c}\\ &+\frac{1}{u}\cdot\frac{1}{u'(c^w)}\cdot\lambda_c\cdot\frac{\partial WE}{\partial\theta}\cdot\frac{1}{k_v}\\ &-\frac{1}{u}\cdot\frac{1}{u'(c^w)}\cdot IE_\theta\end{aligned}$$

We can rearrange (1) to:

$$\chi\cdot k_v \;=\; (1-\gamma)\cdot q\cdot\left(\frac{\eta-\gamma}{(1-\gamma)(1-\eta)}\cdot\frac{k_v}{q}+\left(1+\frac{\lambda_\theta+\eta\cdot\lambda\cdot\lambda_c}{n^s\cdot u'(c^w)}\right)\cdot\frac{b}{f}\right)$$

Note that

$$(1-\gamma)\cdot q \;=\; \frac{\partial f}{\partial\theta}$$

This follows from the fact that:

$$f'(\theta) \;=\; q(\theta)+\theta\cdot q'(\theta) \;=\; q(\theta)\cdot\left(1+\frac{q'(\theta)\cdot\theta}{q(\theta)}\right) \;=\; q(\theta)\cdot(1-\gamma)$$

So we can write:

$$\chi\cdot k_v \;=\; \left(\frac{\partial f}{\partial\theta}\right)\cdot\left(\frac{\eta-\gamma}{(1-\gamma)(1-\eta)}\cdot\frac{k_v}{q}+\left(1+\frac{\lambda_\theta+\eta\cdot\lambda\cdot\lambda_c}{n^s\cdot u'(c^w)}\right)\cdot\frac{b}{f}\right)$$

Inserting χ gives:

$$\left[\gamma-(\lambda\gamma-f\eta)\cdot\frac{1}{\lambda}\cdot(1-\rho)\right]\cdot\frac{1}{1-\eta}\cdot\frac{k_v}{f}\cdot\lambda_\theta$$

$$= u'(c^w)\cdot\left(\frac{\partial f}{\partial\theta}\right)\cdot u\cdot\left(\frac{\eta-\gamma}{(1-\gamma)(1-\eta)}\cdot\frac{k_v}{q}+\left(1+\frac{\lambda_\theta+\eta\cdot\lambda\cdot\lambda_c}{n^s\cdot u'(c^w)}\right)\cdot\frac{b}{f}\right)$$

$$+(\lambda\gamma-f\eta)\cdot\frac{1}{1-\eta}\cdot\frac{k_v}{f}\cdot(-\lambda_{\epsilon_s^u})$$

$$+\left(\lambda\cdot\frac{\gamma}{f}\cdot u'(c^w)+\left(\frac{1}{1-\eta}\cdot\frac{\lambda\gamma-f}{f}\right)\cdot\eta\right)\cdot k_v\cdot(-\lambda_{\epsilon_s^c})$$

$$+\frac{\partial WE}{\partial\theta}\cdot(-\lambda_c) \quad +IE_\theta$$

Inserting $\lambda_{\epsilon_s^u}, \lambda_{\epsilon_s^c}, IE_\theta$ gives

$$\begin{aligned}
&\left[\gamma - (\lambda\gamma - f\eta)\cdot\frac{1}{\lambda}\cdot(1-\rho)\right]\cdot\frac{1}{1-\eta}\cdot\frac{k_v}{f}\cdot\lambda_\theta \\
&= u'(c^w)\cdot\frac{\partial f}{\partial\theta}\cdot u\cdot\left(\frac{\eta-\gamma}{(1-\gamma)(1-\eta)}\cdot\frac{k_v}{q}+\left(1+\frac{\lambda_\theta}{n^s u'(c^w)}\right)\cdot\frac{b}{f}\right) \\
&\quad + u'(c^w)\cdot n^s\cdot\left(-\frac{\partial\epsilon_s^u}{\partial\theta}\right)^{\text{total}}\cdot g(\epsilon_s^u)\cdot\frac{1-p}{\lambda}\cdot\left(1+\frac{\lambda_\theta}{n^s u'(c^w)}\right)\cdot\left(\frac{\lambda}{f}\cdot b-\lambda F\right) \\
&\quad + u'(c^w)\cdot n^s\cdot\left(-\frac{\partial\epsilon_s^c}{\partial\theta}\right)^{\text{total}}\cdot g(\epsilon_s^c)\cdot\frac{p}{\lambda}\cdot\left(1+\frac{\lambda_\theta}{n^s u'(c^w)}\right)\cdot\frac{\lambda}{f}\cdot b \\
&\quad + \lambda_\theta\cdot\left(-\frac{\partial c^w}{\partial\theta}\right)^{\text{total},2}\cdot\left(-\frac{\partial S}{\partial c^w}\right)^{\text{total}} \\
&\quad + u'(c^w)\cdot n^s\cdot\left(1+\frac{\lambda_\theta}{n^s u'(c^w)}\right)\cdot\tilde{IE}_\theta
\end{aligned}$$

with

$$\begin{aligned}
\left(-\frac{\partial\epsilon_s^u}{\partial\theta}\right)^{\text{total}} &= \left(-\frac{\partial\epsilon_s^u}{\partial\theta}\right)+\left(\frac{\partial c^w}{\partial\theta}\right)^{\text{total}}\cdot\left(-\frac{\partial\epsilon_s^u}{\partial c^w}\right) \\
&\quad + \left(-\frac{\partial\epsilon_s^u}{\partial\theta}\right)\cdot\left[\left(\frac{\partial n}{\partial\epsilon_s^u}\right)\cdot\left(-\frac{\partial c^w}{\partial n}\right)^{\text{total}}+\left(\frac{\partial c^w}{\partial\epsilon_s^u}\right)^{\text{total}}\right]\cdot\left(\frac{\partial\epsilon_s^u}{\partial c^w}\right) \\
&\quad + f'(\theta)\cdot\frac{\partial n}{\partial f}\cdot\frac{\partial c^w}{\partial n}\cdot\frac{\partial\epsilon_s^u}{\partial c^w}
\end{aligned}$$

$$\begin{aligned}
\left(-\frac{\partial\epsilon_s^c}{\partial\theta}\right)^{\text{total}} &= \left[\left(-\frac{\partial\epsilon_s^c}{\partial\theta}\right)+\left(\frac{\partial c^w}{\partial\theta}\right)^{\text{total}}\cdot\left(-\frac{\partial\epsilon_s^c}{\partial c^w}\right)\right. \\
&\quad + \left(-\frac{\partial\epsilon_s^c}{\partial\theta}\right)\cdot\left[\left(\frac{\partial n}{\partial\epsilon_s^c}\right)\cdot\left(-\frac{\partial c^w}{\partial n}\right)^{\text{total}}+\left(\frac{\partial c^w}{\partial\epsilon_s^c}\right)^{\text{total}}\right]\cdot\frac{\partial\epsilon_s^c}{\partial c^w} \\
&\quad \left. + f'(\theta)\cdot\frac{\partial n}{\partial f}\cdot\frac{\partial c^w}{\partial n}\cdot\frac{\partial\epsilon_s^c}{\partial c^w}\right]
\end{aligned}$$

$$\begin{aligned}
\left(-\frac{\partial c^w}{\partial\theta}\right)^{\text{total},2} &= \left(-\frac{\partial c^w}{\partial\theta}\right)^{\text{total}}+\left(-\frac{\partial\epsilon_s^u}{\partial\theta}\right)\cdot\left(\frac{\partial n}{\partial\epsilon_s^u}\cdot\left(-\frac{\partial c^w}{\partial n}\right)^{\text{total}}+\left(\frac{\partial c^w}{\partial\epsilon_s^u}\right)^{\text{total}}\right) \\
&\quad + \left(-\frac{\partial\epsilon_s^c}{\partial\theta}\right)\cdot\left(\frac{\partial n}{\partial\epsilon_s^c}\cdot\left(-\frac{\partial c^w}{\partial n}\right)^{\text{total}}+\left(\frac{\partial c^w}{\partial\epsilon_s^c}\right)^{\text{total}}\right) \\
&\quad + f'(\theta)\cdot\frac{\partial n}{\partial f}\cdot\frac{\partial c^w}{\partial n}
\end{aligned}$$

This allows us to calculate the Lagrange multiplier for the job-creation condition

$$M \cdot \lambda_\theta = u'(c^w) \cdot \frac{\partial f}{\partial \theta} \cdot u \cdot \left(\frac{\eta - \gamma}{(1-\gamma)(1-\eta)} \cdot \frac{k_v}{q} + \left(1 + \frac{\lambda_\theta}{n^s u'(c^w)}\right) \cdot \frac{b}{f} \right)$$
$$+ u'(c^w) \cdot n^s \cdot \left(-\frac{\partial \epsilon_s^u}{\partial \theta} \right)^{\text{total}} \cdot g(\epsilon_s^u) \cdot \frac{1-p}{\lambda} \left(\frac{\lambda}{f} b - \lambda F \right)$$
$$+ u'(c^w) \cdot n^s \cdot \left(-\frac{\partial \epsilon_s^c}{\partial \theta} \right)^{\text{total}} \cdot g(\epsilon_s^c) \cdot \frac{p}{\lambda} \cdot \frac{\lambda}{f} b$$
$$+ u'(c^w) \cdot n^s \cdot \tilde{IE}_\theta$$

with

$$M = \left[\gamma - (\lambda\gamma - f \cdot \eta) \cdot \frac{1}{\lambda} \cdot (1-\rho) \right] \cdot \frac{1}{1-\eta} \cdot \frac{k_v}{f}$$
$$+ \left(-\frac{\partial f}{\partial \theta} \cdot u \cdot \frac{b}{f} \right)$$
$$+ \left(\frac{\partial \epsilon_s^u}{\partial \theta} \right)^{\text{total}} \cdot g(\epsilon_s^u) \cdot \frac{1-p}{\lambda} \cdot \left(\frac{\lambda}{f} b - \lambda \cdot F \right)$$
$$+ \left(\frac{\partial \epsilon_s^c}{\partial \theta} \right)^{\text{total}} \cdot g(\epsilon_s^c) \cdot \frac{p}{\lambda} \cdot \frac{\lambda}{f} b$$
$$+ \left(\frac{\partial c^w}{\partial \theta} \right)^{\text{total}} \cdot \left(-\frac{\partial S}{\partial c^w} \right)^{\text{total}}$$
$$+ \tilde{IE}_\theta$$

Note: $\frac{1}{M}$ denotes the general equilibrium effect of an increase of the joint surplus on θ:

$$\left(\frac{\partial \theta}{\partial S} \right)^{\text{ge}} = \frac{1}{M}$$

Rearranging for λ_θ gives:

$$\lambda_\theta = u'(c^w) \cdot \left(\frac{\partial f}{\partial S} \right)^{\text{ge}} \cdot u \cdot \left(\frac{\eta - \gamma}{(1-\gamma)(1-\eta)} \cdot \frac{k_v}{q} + \frac{\lambda}{f} \cdot b \right)$$
$$+ u'(c^w) \cdot \left(-\frac{\partial \epsilon_s^u}{\partial S} \right)^{\text{ge}} \cdot \frac{\partial n^u}{\partial \epsilon_s^u} \cdot \left(\frac{\lambda}{f} b - \lambda F \right)$$
$$+ u'(c^w) \cdot \left(-\frac{\partial \epsilon_s^c}{\partial S} \right)^{\text{ge}} \cdot \frac{\partial n^c}{\partial \epsilon_s^c} \cdot \frac{\lambda}{f} b$$
$$+ u'(c^w) \cdot n^c \cdot \left(\frac{\partial \theta}{\partial S} \right)^{\text{ge}} \cdot \hat{IE}_\theta$$

Note that

$$\left(\frac{\partial f}{\partial S} \right)^{\text{ge}} = \frac{1}{M} \cdot \frac{\partial f}{\partial S}$$

$$\left(\frac{\partial \epsilon_s^u}{\partial S}\right)^{\text{ge}} = \frac{1}{M} \cdot \left(\frac{\partial \epsilon_s^u}{\partial S}\right)^{\text{total}}$$

$$\left(\frac{\partial \epsilon_s^c}{\partial S}\right)^{\text{ge}} = \frac{1}{M} \cdot \left(\frac{\partial \epsilon_s^c}{\partial S}\right)^{\text{total}}$$

Further define:

$$\hat{IE}_\theta = \frac{1}{1-\rho^c} \cdot \int_{\epsilon_s^c}^{\epsilon^p} \lambda \cdot \frac{\gamma}{f} \cdot \frac{u'(c_{\text{stw}}(\epsilon)) - u'(c^w)}{u'(c^w)} dG(\epsilon)$$

Finally, we can calculate the Lagrange multiplier for λ_c by inserting into λ_θ:

$$\begin{aligned}
\lambda_c = \frac{-1}{\left(\frac{\partial S}{\partial c^w}\right)^{\text{total}}} \cdot \Bigg\{ \\
&\left(\frac{\partial \epsilon_s^u}{\partial c^w} + \left(\frac{\partial S}{\partial c^w}\right)^{\text{total}} \cdot \left(\frac{\partial \epsilon_s^u}{\partial S}\right)^{\text{ge}}\right) \cdot \frac{1-p}{\lambda} \cdot g(\epsilon_s^u) \cdot n^s \cdot u'(c^w) \cdot \left(\frac{\lambda}{f} b - \lambda F\right) \\
&+ \left(\frac{\partial \epsilon_s^c}{\partial c^w} + \left(\frac{\partial S}{\partial c^w}\right)^{\text{total}} \cdot \left(\frac{\partial \epsilon_s^c}{\partial S}\right)^{\text{ge}}\right) \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot n^s \cdot u'(c^w) \cdot \frac{\lambda}{f} b \\
&+ \left(\frac{\partial S}{\partial c^w}\right)^{\text{total}} \cdot \left(-\frac{\partial f}{\partial S}\right)^{\text{ge}} \cdot u \cdot \left(\frac{\eta - \gamma}{(1-\gamma)(1-\eta)} \cdot \frac{k_v}{q} + \left(1 + \frac{\lambda_\theta}{n^s u'(c^w)}\right) \cdot \frac{b}{f}\right) \\
&- n^c \cdot \left(\frac{\partial S}{\partial c^w}\right)^{\text{ge}} \cdot \left(\frac{\partial \theta}{\partial S}\right) \cdot \hat{IE}_\theta \Bigg\}.
\end{aligned}$$

Simplifying:

$$\begin{aligned}
\lambda_c = -\frac{u'(c^w)}{\left(\frac{\partial WE}{\partial c^w}\right)^{\text{total}}} \Bigg[&\left(\frac{\partial \epsilon_s^u}{\partial c^w}\right)^{\text{ge}} \cdot \frac{\partial n^u}{\partial \epsilon_s^u} \cdot \left(\frac{\lambda}{f} b - \lambda F\right) + \left(\frac{\partial \epsilon_s^c}{\partial c^w}\right)^{\text{ge}} \cdot \frac{\partial n^c}{\partial \epsilon_s^c} \cdot \frac{\lambda}{f} b \\
&+ \left(-\frac{\partial f}{\partial c^w}\right)^{\text{ge}} \cdot u \cdot \left(\frac{\eta - \gamma}{(1-\gamma)(1-\eta)} \cdot \frac{k_v}{q} + \frac{b}{f}\right) + n^c \cdot \left(-\frac{\partial \theta}{\partial c^w}\right)^{\text{ge}} \cdot \hat{IE}_\theta \Bigg]
\end{aligned}$$

Following the same arguments, we can express the Lagrange multiplier for the separation conditions as:

$$\begin{aligned}
\lambda_{\epsilon_s^u} = -\frac{u'(c^w)}{\left(\frac{\partial S^u_{\text{stw}}(\epsilon_s^u)}{\partial \epsilon_s^u}\right)^{\text{total}}} \Bigg[&\left(\frac{\partial n^u}{\partial \epsilon_s^u}\right)^{\text{ge}} \cdot \left(\frac{\lambda}{f} b - \lambda F\right) + \left(\frac{\partial \epsilon_s^c}{\partial \epsilon_s^u}\right)^{\text{ge}} \cdot \frac{\partial n^c}{\partial \epsilon_s^c} \cdot \frac{\lambda}{f} b \\
&+ \frac{\partial c^w}{\partial \epsilon_s^u} \cdot \left(-\frac{\partial f}{\partial c^w}\right)^{\text{ge}} \cdot u \cdot \left(\frac{\eta - \gamma}{(1-\gamma)(1-\eta)} \cdot \frac{k_v}{q} + \frac{b}{f}\right) \\
&+ n^c \cdot \frac{\partial c^w}{\partial \epsilon_s^u} \cdot \left(-\frac{\partial \theta}{\partial c^w}\right)^{\text{ge}} \cdot \hat{IE}_\theta \Bigg]
\end{aligned}$$

$$\lambda_{\epsilon_s^c} = -\frac{u'(c^w)}{\left(\frac{\partial S^c_{\text{stw}}(\epsilon_s^c)}{\partial \epsilon_s^c}\right)^{\text{total}}}\Bigg[\frac{\partial \epsilon_s^u}{\partial \epsilon_s^c} \cdot \left(\frac{\partial n^u}{\partial \epsilon_s^u}\right)^{\text{ge}} \cdot \left(\frac{\lambda}{f}b - \lambda F\right)$$
$$+ \left(\frac{\partial n^c}{\partial \epsilon_s^c}\right)^{\text{ge}} \cdot \frac{\lambda}{f}b + \frac{\partial c^w}{\partial \epsilon_s^c} \cdot \left(-\frac{\partial f}{\partial c^w}\right)^{\text{ge}} \cdot u \cdot \left(\frac{\eta - \gamma}{(1-\gamma)(1-\eta)} \cdot \frac{k_v}{q} + \frac{b}{f}\right)$$
$$+ n^c \cdot \frac{\partial c^w}{\partial \epsilon_s^c} \cdot \left(-\frac{\partial \theta}{\partial c^w}\right)^{\text{ge}} \cdot \hat{IE}_\theta\Bigg]$$

For convenience, the optimal UI benefits are replicated here:

$$(1-n) \cdot \left(u'(b) - u'(c^w)\right) = \lambda_\theta \cdot (1-\rho) \cdot \left(\frac{1-n}{n} + \frac{u'(b)}{u'(c^w)}\right) \cdot \left(-\frac{u'(b)}{u'(c^w)}\right)$$
$$+ (\lambda_{\epsilon_s^u} + \lambda_{\epsilon_s^c}) \cdot \left(-\frac{u'(b)}{u'(c^w)}\right)$$
$$+ \lambda_c \cdot (1-\rho) \cdot \left(\eta \cdot \frac{1-n}{n} + (1-\eta) \cdot \frac{u'(b)}{u'(c^w)}\right)$$

Inserting the Lagrange multiplier gives:

$$(1-n)\cdot\left(u'(b) - u'(c^w)\right)$$
$$= u'(c^w) \cdot \left[\frac{\partial c^w}{\partial \epsilon_s^u} \cdot \left(-\frac{\partial f}{\partial c^w}\right)^{\text{ge}} \cdot u \cdot \left(\frac{\eta - \gamma}{(1-\gamma)(1-\eta)} \cdot \frac{k_v}{q} + \frac{b}{f}\right)\right]$$
$$+ u'(c^w) \cdot \left[\left(\frac{\partial \epsilon_s^u}{\partial b}\right)^{\text{ge}} \cdot \frac{\partial n^u}{\partial \epsilon_s^u} \cdot \left(\frac{\lambda}{f}b - \lambda F\right)\right]$$
$$+ u'(c^w) \cdot \left[\left(\frac{\partial \epsilon_s^c}{\partial b}\right)^{\text{ge}} \cdot \frac{\partial n^c}{\partial \epsilon_s^c} \cdot \frac{\lambda}{f}b\right]$$
$$+ u'(c^w) \cdot \left[n^c \cdot \left(-\frac{\partial \theta}{\partial b}\right)^{\text{ge}} \cdot \hat{IE}_\theta\right].$$

With:

$$\left(\frac{\partial \epsilon_s^u}{\partial b}\right)^{\text{ge}} \cdot \frac{\partial n^u}{\partial \epsilon_s^u} = \underbrace{\left(\frac{\partial \epsilon_s^u}{\partial b}\right)^{\text{ge}} \cdot \frac{\partial n^u}{\partial \epsilon_s^u}}_{\text{direct effects}}$$
$$+ \underbrace{\left[\frac{\partial S}{\partial b} \cdot \left(\frac{\partial \epsilon_s^u}{\partial S}\right)^{\text{ge}} + \frac{\partial \epsilon_s^c}{\partial b} \cdot \left(\frac{\partial \epsilon_s^u}{\partial \epsilon_s^c}\right)^{\text{ge}} + \frac{\partial c^w}{\partial b} \cdot \left(\frac{\partial \epsilon_s^u}{\partial c^w}\right)^{\text{ge}}\right] \cdot \frac{\partial n^u}{\partial \epsilon_s^u}}_{\text{indirect effect}}$$

$$\left(\frac{\partial \epsilon_s^c}{\partial b}\right)^{\text{ge}} \cdot \frac{\partial n^c}{\partial \epsilon_s^c} = \underbrace{\left(\frac{\partial \epsilon_s^c}{\partial b}\right)^{\text{ge}} \cdot \frac{\partial n^c}{\partial \epsilon_s^c}}_{\text{direct effects}} + \underbrace{\left[\frac{\partial S}{\partial b} \cdot \left(\frac{\partial \epsilon_s^c}{\partial S}\right)^{\text{ge}} + \frac{\partial \epsilon_s^u}{\partial b} \cdot \left(\frac{\partial \epsilon_s^c}{\partial \epsilon_s^u}\right)^{\text{ge}} + \frac{\partial c^w}{\partial b} \cdot \left(\frac{\partial \epsilon_s^c}{\partial c^w}\right)^{\text{ge}}\right] \cdot \frac{\partial n^c}{\partial \epsilon_s^c}}_{\text{indirect effect}}$$

$$\left(\frac{\partial f}{\partial b}\right)^{\text{ge}} = \underbrace{\left(\frac{\partial S}{\partial b} \cdot \left(-\frac{\partial f}{\partial S}\right)^{\text{ge}}\right)}_{\text{direct effect}} + \underbrace{\left[\frac{\partial \epsilon_s^u}{\partial b} \cdot \frac{\partial c^w}{\partial \epsilon_s^u} + \frac{\partial \epsilon_s^c}{\partial b} \cdot \frac{\partial c^w}{\partial \epsilon_s^c} + \frac{\partial c^w}{\partial b}\right] \cdot \left(\frac{\partial f}{\partial c^w}\right)^{\text{ge}}}_{\text{indirect effect}}$$

$$\left(\frac{\partial \theta}{\partial b}\right)^{\text{ge}} = \underbrace{\frac{\partial S}{\partial b} \cdot \left(\frac{\partial \theta}{\partial S}\right)^{\text{ge}}}_{\text{direct effect}} + \underbrace{\left[\frac{\partial \epsilon_s^u}{\partial b} \cdot \frac{\partial c^w}{\partial \epsilon_s^u} + \frac{\partial \epsilon_s^c}{\partial b} \cdot \frac{\partial c^w}{\partial \epsilon_s^c} + \frac{\partial c^w}{\partial b}\right] \cdot \left(\frac{\partial \theta}{\partial c^w}\right)^{\text{ge}}}_{\text{indirect effect}}$$

3.E.7 Optimal UI under Optimal Lay-off Tax

From the FOC of UI, we can derive the optimality condition for unemployment insurance. In contrast to the section where we take lay-off taxes as given, the government implements the optimal lay-off tax. This implies that the Lagrange multiplier of the separation condition in unconstrained firms is equal to zero. This is also fundamental for the optimality condition for the UI benefits:

$$(1-\eta) \cdot \left(u'(b) - u'(c^w)\right) = \lambda_\theta \cdot (1-\rho) \cdot \left(\frac{1-n}{n} + \frac{u'(b)}{u'(c^w)}\right) + \lambda_{\epsilon_s^u} \cdot \left(-\frac{u'(b)}{u'(c^w)}\right) + \lambda_c \cdot (1-\rho) \cdot \left(\eta \cdot \frac{1-n}{n} + (1-\eta) \cdot \frac{u'(b)}{u'(c^w)}\right)$$

The Lagrange multiplier for the separation condition of unconstrained firms is equal to

$$-\frac{\lambda_{\epsilon_s^c}}{n^s \cdot u'(c^w)} = \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot \frac{1}{\frac{\partial S^c_{\text{stw}}(\epsilon_s^c)}{\partial \epsilon_s^c}} \cdot \left[\left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right) \cdot \frac{\lambda}{f} \cdot b \right.$$
$$\left. + \frac{\lambda_c}{n^s \cdot u'(c^w)} \cdot \left(\eta \cdot \frac{b}{n} + \frac{1}{g(\epsilon_s^c)} \cdot \frac{\lambda}{p} \cdot \frac{\partial WE}{\partial \epsilon_s^c}\right)\right]$$

First, let us find an expression for λ_c:

$$\frac{\partial \mathcal{L}}{\partial c^w} = -\Bigg((1-p) \cdot (1-\rho^u) \cdot \frac{u(c^w) - u(b)}{u'(c^w)} \cdot \frac{u''(c^w)}{u'(c^w)}$$
$$+ p \cdot (1-\rho^c) \cdot \frac{u^c - u(b)}{u'(c^w)} \cdot \frac{u''(c^w)}{u'(c^w)}\Bigg) \cdot \lambda_\theta$$
$$- (1-\eta) \cdot \Bigg((1-p) \cdot (1-\rho^u) \cdot \left(1 - (1-\eta) \cdot \frac{u(c^w) - u(b)}{u'(c^w)} \cdot \frac{u''(c^w)}{u'(c^w)}\right)$$
$$- p \cdot (1-\rho^c) \cdot (1-\eta) \cdot \frac{u^c - u(b)}{u'(c^w)} \cdot \frac{u''(c^w)}{u'(c^w)}\Bigg) \cdot \lambda_c$$
$$+ \frac{u(c_{\text{stw}}(\epsilon_s^c)) - u(b)}{u'(c^w)} \cdot \frac{u''(c^w)}{u'(c^w)} \cdot \lambda_{\epsilon_s^c} = 0$$

Insert Lagrange multipliers for separation condition:

$$-\Bigg((1-p) \cdot (1-\rho^u) \cdot \frac{u(c^w) - u(b)}{u'(c^w)} \cdot \frac{u''(c^w)}{u'(c^w)}$$
$$+ p \cdot (1-\rho^c) \cdot \frac{u^c - u(b)}{u'(c^w)} \cdot \frac{u''(c^w)}{u'(c^w)}\Bigg) \cdot \lambda_\theta$$
$$- (1-\eta) \cdot (1-p) \cdot (1-\rho^u) \cdot \left(1 - (1-\eta) \cdot \frac{u(c^w) - u(b)}{u'(c^w)} \cdot \frac{u''(c^w)}{u'(c^w)}\right)$$
$$\cdot\, p \cdot (1-\rho^c) \cdot (1-\eta) \cdot \frac{u^c - u(b)}{u'(c^w)} \cdot \frac{u''(c^w)}{u'(c^w)} \cdot \lambda_c$$
$$+ \frac{\partial \epsilon_s^c}{\partial c^w} \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot \left(\left(n^s \cdot u'(c^w) + \lambda_\theta\right) \cdot \left(\frac{\lambda}{f} \cdot b\right)\right.$$
$$\left. + \lambda_c \cdot \left(\eta \cdot \frac{b}{n} + \frac{1}{g(\epsilon_s^c)} \cdot \frac{\lambda}{p} \cdot \frac{\partial WE}{\partial \epsilon_s^c}\right)\right) = 0$$

Rearranging gives:

$$0 = \lambda_\theta \cdot \frac{\partial S^{\text{total}}}{\partial c^w}$$
$$+ \lambda_c \cdot \frac{\partial WE^{\text{total}}}{\partial c^w}$$
$$+ \frac{\partial \epsilon_s^c}{\partial c^w} \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot n^s \cdot u'(c^w) \cdot \frac{\lambda}{f} b$$

with

$$\frac{\partial S^{\text{total}}}{\partial c^w} = \Bigg[p(1-\rho^u)\frac{u(c^w)-u(b)}{u'(c^w)}\frac{u''(c^w)}{u'(c^w)} - (1-p)(1-\rho^c)\frac{u^c-u(b)}{u'(c^w)}\frac{u''(c^w)}{u'(c^w)} + \frac{\partial \epsilon_s^c}{\partial \theta}\cdot\frac{p}{\lambda}\cdot g(\epsilon_s^c)\cdot\frac{\lambda}{f}b\Bigg]$$

and

$$\frac{\partial WE^{\text{total}}}{\partial c^w} = \Bigg[-(1-\eta)(1-p\rho^u)\left(1-(1-\eta)\frac{u(c^w)-u(b)}{u'(c^w)}\frac{u''(c^w)}{u'(c^w)}\right) + (1-(1-p)\rho^c)(1-\eta)\frac{u^c-u(b)}{u'(c^w)}\frac{u''(c^w)}{u'(c^w)} + \frac{\partial \epsilon_s^c}{\partial c^w}\cdot\frac{p}{\lambda}\cdot g(\epsilon_s^c)\cdot\lambda_c\left(\eta\frac{b}{n}+\frac{1}{g(\epsilon_s^c)}\cdot\frac{\lambda}{p}\cdot\frac{\partial WE}{\partial \epsilon_s^c}\right)\Bigg]$$

Next we can solve for the Lagrange multiplier of the wage equation:

$$\lambda_c = -\frac{1}{\left(\frac{\partial WE}{\partial c^w}\right)^{\text{total}}}\Bigg[\frac{\partial \epsilon_s^c}{\partial c^w}\cdot\frac{p}{\lambda}\cdot g(\epsilon_s^c)\cdot n^s\cdot u'(c^w)\cdot\frac{\lambda}{f}\cdot b + \left(-\frac{\partial S^{\text{total}}}{\partial c^w}\right)\cdot\lambda_\theta\Bigg]$$

Now we have everything to calculate λ_θ from the two equations:

$$(1)\quad \left(\chi-\frac{\eta-\gamma}{1-\eta}\right)\cdot\frac{1}{1-\gamma}\cdot\frac{k_v}{q} = \left(1+\frac{\lambda_\theta+\eta\cdot\lambda\cdot\lambda_c}{n^s\cdot u'(c^w)}\right)\cdot\frac{b}{f}$$

$$(2)\quad \chi = \frac{1}{1-\eta}\cdot\frac{1}{u\cdot f}\cdot\frac{1}{u'(c^w)}\cdot\left[\gamma-(\lambda\gamma-f\eta)\cdot\left(\frac{1}{\lambda}(1-\rho)\lambda_\theta\right)\right] + \frac{1}{u}\cdot\frac{1}{u'(c^w)}\cdot\left(\lambda\cdot\frac{\gamma}{f}\cdot u'(c^w)+\left(\frac{1}{1-\eta}\cdot\frac{\lambda\gamma-f}{f}\right)\cdot\eta\right)\cdot\lambda_{\epsilon_s^c} + \frac{1}{u}\cdot\frac{1}{u'(c^w)}\cdot\lambda_c\cdot\frac{\partial WE}{\partial\theta}\cdot\frac{1}{k_v} - \frac{1}{u}\cdot\frac{1}{u'(c^w)}\cdot IE_\theta$$

We can rearrange (1) to:

$$\chi\cdot k_v = (1-\gamma)\cdot q\cdot\left(\frac{\eta-\gamma}{(1-\gamma)(1-\eta)}\cdot\frac{k_v}{q}+\left(1+\frac{\lambda_\theta+\eta\cdot\lambda\cdot\lambda_c}{n^s\cdot u'(c^w)}\right)\cdot\frac{b}{f}\right)$$

Note that

$$(1-\gamma)\cdot q \;=\; \frac{\partial f}{\partial \theta}$$

This follows from the fact that:

$$f'(\theta) \;=\; q(\theta)+\theta\cdot q'(\theta) \;=\; q(\theta)\cdot\left(1+\frac{q'(\theta)\cdot\theta}{q(\theta)}\right) \;=\; q(\theta)\cdot(1-\gamma)$$

So we can write:

$$\chi\cdot k_v \;=\; \left(\frac{\partial f}{\partial \theta}\right)\cdot\left(\frac{\eta-\gamma}{(1-\gamma)(1-\eta)}\cdot\frac{k_v}{q}+\left(1+\frac{\lambda_\theta+\eta\cdot\lambda\cdot\lambda_c}{n^s\cdot u'(c^w)}\right)\cdot\frac{b}{f}\right)$$

Inserting χ gives:

$$\left[\gamma-(\lambda\gamma-f\eta)\cdot\frac{1}{\lambda}\cdot(1-\rho)\right]\cdot\frac{1}{1-\eta}\cdot\frac{k_v}{f}\cdot\lambda_\theta$$

$$=\; u'(c^w)\cdot\left(\frac{\partial f}{\partial \theta}\right)\cdot u\cdot\left(\frac{\eta-\gamma}{(1-\gamma)(1-\eta)}\cdot\frac{k_v}{q}+\left(1+\frac{\lambda_\theta+\eta\cdot\lambda\cdot\lambda_c}{n^s\cdot u'(c^w)}\right)\cdot\frac{b}{f}\right)$$

$$+\left(\lambda\cdot\frac{\gamma}{f}\cdot u'(c^w)+\left(\frac{1}{1-\eta}\cdot\frac{\lambda\gamma-f}{f}\right)\cdot\eta\right)\cdot k_v\cdot(-\lambda_{\epsilon_s^c})$$

$$+\frac{\partial WE}{\partial \theta}\cdot(-\lambda_c)\quad +IE_\theta$$

Inserting $\lambda_{\epsilon_s^c}, IE_\theta$ gives

$$\begin{aligned}
&\left[\gamma-(\lambda\gamma-f\eta)\cdot\frac{1}{\lambda}\cdot(1-\rho)\right]\cdot\frac{1}{1-\eta}\cdot\frac{k_v}{f}\cdot\lambda_\theta\\
&=\; u'(c^w)\cdot\frac{\partial f}{\partial \theta}\cdot u\cdot\left(\frac{\eta-\gamma}{(1-\gamma)(1-\eta)}\cdot\frac{k_v}{q}+\left(1+\frac{\lambda_\theta}{n^s u'(c^w)}\right)\cdot\frac{b}{f}\right)\\
&\quad+u'(c^w)\cdot n^s\cdot\left(-\frac{\partial\epsilon_s^c}{\partial\theta}\right)^{\text{total}}\cdot g(\epsilon_s^c)\cdot\frac{p}{\lambda}\cdot\left(1+\frac{\lambda_\theta}{n^s u'(c^w)}\right)\cdot\frac{\lambda}{f}\cdot b\\
&\quad+\lambda_\theta\cdot\left(-\frac{\partial c^w}{\partial\theta}\right)^{\text{total,2}}\cdot\left(-\frac{\partial S}{\partial c^w}\right)^{\text{total}}\\
&\quad+u'(c^w)\cdot n^s\cdot\left(1+\frac{\lambda_\theta}{n^s u'(c^w)}\right)\cdot\tilde{IE}_\theta
\end{aligned}$$

with

$$
\begin{aligned}
\left(-\frac{\partial \epsilon_s^c}{\partial \theta}\right)^{\text{total}} =& \Bigg[\left(-\frac{\partial \epsilon_s^c}{\partial \theta}\right) + \left(\frac{\partial c^w}{\partial \theta}\right)^{\text{total}} \cdot \left(-\frac{\partial \epsilon_s^c}{\partial c^w}\right) \\
&+ \left(-\frac{\partial \epsilon_s^c}{\partial \theta}\right) \cdot \left(\frac{\partial n}{\partial \epsilon_s^c} \cdot \left(-\frac{\partial c^w}{\partial n}\right)^{\text{total}} + \left(\frac{\partial c^w}{\partial \epsilon_s^c}\right)^{\text{total}}\right)\Bigg] \cdot \frac{\partial \epsilon_s^c}{\partial c^w} \\
&+ f'(\theta) \cdot \frac{\partial n}{\partial f} \cdot \frac{\partial c^w}{\partial n} \cdot \frac{\partial \epsilon_s^c}{\partial c^w}
\end{aligned}
$$

$$
\begin{aligned}
\left(-\frac{\partial c^w}{\partial \theta}\right)^{\text{total,2}} =& \left(-\frac{\partial c^w}{\partial \theta}\right)^{\text{total}} \\
&+ \left(-\frac{\partial \epsilon_s^u}{\partial \theta}\right) \cdot \left(\frac{\partial n}{\partial \epsilon_s^c} \cdot \left(-\frac{\partial c^w}{\partial n}\right)^{\text{total}} + \left(\frac{\partial c^w}{\partial \epsilon_s^c}\right)^{\text{total}}\right) \\
&+ f'(\theta) \cdot \frac{\partial n}{\partial f} \cdot \frac{\partial c^w}{\partial n}
\end{aligned}
$$

This allows us to calculate the Lagrange multiplier for the job-creation condition

$$
\begin{aligned}
M \cdot \lambda_\theta \;=\; & u'(c^w) \cdot \frac{\partial f}{\partial \theta} \cdot u \cdot \left(\frac{\eta - \gamma}{(1-\gamma)(1-\eta)} \cdot \frac{k_v}{q} + \left(1 + \frac{\lambda_\theta}{n^s u'(c^w)}\right) \cdot \frac{b}{f}\right) \\
&+ u'(c^w) \cdot n^s \cdot \left(-\frac{\partial \epsilon_s^c}{\partial \theta}\right)^{\text{total}} \cdot g(\epsilon_s^c) \cdot \frac{p}{\lambda} \cdot \frac{\lambda}{f} b \\
&+ u'(c^w) \cdot n^s \cdot \tilde{IE}_\theta
\end{aligned}
$$

with

$$
\begin{aligned}
M =& \left[\gamma - (\lambda\gamma - f \cdot \eta) \cdot \frac{1}{\lambda} \cdot (1-\rho)\right] \cdot \frac{1}{1-\eta} \cdot \frac{k_v}{f} \\
&+ \left(-\frac{\partial f}{\partial \theta} \cdot u \cdot \frac{b}{f}\right) \\
&+ \left(\frac{\partial \epsilon_s^c}{\partial \theta}\right)^{\text{total}} \cdot g(\epsilon_s^c) \cdot \frac{p}{\lambda} \cdot \frac{\lambda}{f} b \\
&+ \left(\frac{\partial c^w}{\partial \theta}\right)^{\text{total}} \cdot \left(-\frac{\partial S}{\partial c^w}\right)^{\text{total}} \\
&+ \tilde{IE}_\theta
\end{aligned}
$$

Note: $\frac{1}{M}$ denotes the general equilibrium effect of an increase of the joint surplus on θ:

$$
\left(\frac{\partial \theta}{\partial S}\right)^{\text{ge}} = \frac{1}{M}
$$

Rearranging for λ_θ gives:

$$\begin{aligned}\lambda_\theta \;=\; & u'(c^w)\cdot\left(\frac{\partial f}{\partial S}\right)^{\text{ge}}\cdot u\cdot\left(\frac{\eta-\gamma}{(1-\gamma)(1-\eta)}\cdot\frac{k_v}{q}+\frac{\lambda}{f}\cdot b\right)\\ & +u'(c^w)\cdot\left(-\frac{\partial\epsilon_s^c}{\partial S}\right)^{\text{ge}}\cdot\frac{\partial n^c}{\partial\epsilon_s^c}\cdot\frac{\lambda}{f}b\\ & +u'(c^w)\cdot n^c\cdot\left(\frac{\partial\theta}{\partial S}\right)^{\text{ge}}\cdot\hat{IE}_\theta\end{aligned}$$

Note that

$$\left(\frac{\partial f}{\partial S}\right)^{\text{ge}} \;=\; \frac{1}{M}\cdot\frac{\partial f}{\partial S}$$

$$\left(\frac{\partial\epsilon_s^c}{\partial S}\right)^{\text{ge}} \;=\; \frac{1}{M}\cdot\left(\frac{\partial\epsilon_s^c}{\partial S}\right)^{\text{total}}$$

Further define:

$$\hat{IE}_\theta \;=\; \frac{1}{1-\rho^c}\cdot\int_{\epsilon_s^c}^{\epsilon^p}\lambda\cdot\frac{\gamma}{f}\cdot\frac{u'(c_{\text{stw}}(\epsilon))-u'(c^w)}{u'(c^w)}\,dG(\epsilon)$$

Finally, we can calculate the Lagrange multiplier for λ_c by inserting into λ_θ:

$$\begin{aligned}\lambda_c \;=\; & \frac{-1}{\left(\frac{\partial S}{\partial c^w}\right)^{\text{total}}}\Bigg[\left(\frac{\partial\epsilon_s^c}{\partial c^w}+\left(\frac{\partial S}{\partial c^w}\right)^{\text{total}}\cdot\left(\frac{\partial\epsilon_s^c}{\partial S}\right)^{\text{ge}}\right)\cdot\frac{p}{\lambda}\cdot g(\epsilon_s^c)\cdot n^s\cdot u'(c^w)\cdot\frac{\lambda}{f}b\\ & +\left(\frac{\partial S}{\partial c^w}\right)^{\text{total}}\cdot\left(-\frac{\partial f}{\partial S}\right)^{\text{ge}}\cdot u\cdot\left(\frac{\eta-\gamma}{(1-\gamma)(1-\eta)}\cdot\frac{k_v}{q}+\left(1+\frac{\lambda_\theta}{n^s u'(c^w)}\right)\cdot\frac{b}{f}\right)\\ & +n^c\cdot\left(\frac{\partial S}{\partial c^w}\right)^{\text{ge}}\cdot\left(-\frac{\partial\theta}{\partial c^w}\right)\cdot\hat{IE}_\theta\Bigg]\end{aligned}$$

Simplifying:

$$\begin{aligned}\lambda_c \;=\; & -\frac{u'(c^w)}{\left(\frac{\partial WE}{\partial c^w}\right)^{\text{total}}}\Bigg[\left(\frac{\partial\epsilon_s^u}{\partial c^w}\right)^{\text{ge}}\cdot\frac{\partial n^u}{\partial\epsilon_s^u}\cdot\left(\frac{\lambda}{f}b-\lambda F\right)+\left(\frac{\partial\epsilon_s^c}{\partial c^w}\right)^{\text{ge}}\cdot\frac{\partial n^c}{\partial\epsilon_s^c}\cdot\frac{\lambda}{f}b\\ & +\left(-\frac{\partial f}{\partial c^w}\right)^{\text{ge}}\cdot u\cdot\left(\frac{\eta-\gamma}{(1-\gamma)(1-\eta)}\cdot\frac{k_v}{q}+\frac{b}{f}\right)+n^c\cdot\left(-\frac{\partial\theta}{\partial c^w}\right)^{\text{ge}}\cdot\hat{IE}_\theta\Bigg]\end{aligned}$$

Following the same arguments, we can express the Lagrange multiplier for the separation condition as:

$$\lambda_{\epsilon_s^c} = -\frac{u'(c^w)}{\left(\frac{\partial S^c_{\text{stw}}(\epsilon_s^c)}{\partial \epsilon_s^c}\right)^{\text{total}}} \left[\frac{\partial \epsilon_s^u}{\partial \epsilon_s^c} \cdot \left(\frac{\partial n^u}{\partial \epsilon_s^u}\right)^{\text{ge}} \cdot \left(\frac{\lambda}{f} b - \lambda F\right) + \left(\frac{\partial n^c}{\partial \epsilon_s^c}\right)^{\text{ge}} \cdot \frac{\lambda}{f} b \right.$$
$$\left. + \frac{\partial c^w}{\partial \epsilon_s^c} \cdot \left(-\frac{\partial f}{\partial c^w}\right)^{\text{ge}} \cdot u \cdot \left(\frac{\eta - \gamma}{(1-\gamma)(1-\eta)} \cdot \frac{k_v}{q} + \frac{b}{f}\right) + \frac{\partial c^w}{\partial \epsilon_s^c} \cdot \left(-\frac{\partial \theta}{\partial c^w}\right)^{\text{ge}} \cdot \hat{IE}_\theta \right]$$

For convenience, the optimal UI benefits are replicated here:

$$\begin{aligned}(1-n) \cdot \left(u'(b) - u'(c^w)\right) &= \lambda_\theta \cdot (1-\rho) \cdot \left(\frac{1-n}{n} + \frac{u'(b)}{u'(c^w)}\right) \cdot \left(-\frac{u'(b)}{u'(c^w)}\right) \\ &\quad + \lambda_{\epsilon_s^c} \cdot \left(-\frac{u'(b)}{u'(c^w)}\right) \\ &\quad + \lambda_c \cdot (1-\rho) \cdot \left(\eta \cdot \frac{1-n}{n} + (1-\eta) \cdot \frac{u'(b)}{u'(c^w)}\right)\end{aligned}$$

Inserting the Lagrange multiplier gives

$$\begin{aligned}(1-n) \cdot \left(u'(b) - u'(c^w)\right) &= u'(c^w) \cdot \left[\left(-\frac{\partial f}{\partial b}\right)^{\text{ge}} \cdot u \cdot \left(\frac{\eta - \gamma}{(1-\gamma)(1-\eta)} \cdot \frac{k_v}{q} + \frac{b}{f}\right) \right. \\ &\quad + \left(\frac{\partial \epsilon_s^c}{\partial b}\right)^{\text{ge}} \cdot \frac{\partial n^c}{\partial \epsilon_s^c} \cdot \frac{\lambda}{f} b \\ &\quad \left. + n^c \cdot \left(-\frac{\partial \theta}{\partial b}\right)^{\text{ge}} \cdot \hat{IE}_\theta \right]\end{aligned}$$

with

$$\left(\frac{\partial \epsilon_s^c}{\partial b}\right)^{\text{ge}} \cdot \frac{\partial n^c}{\partial \epsilon_s^c} = \underbrace{\left(\frac{\partial \epsilon_s^c}{\partial b}\right)^{\text{ge}} \cdot \frac{\partial n^c}{\partial \epsilon_s^c}}_{\text{direct effects}} + \underbrace{\left[\frac{\partial S}{\partial b} \cdot \left(\frac{\partial \epsilon_s^c}{\partial S}\right)^{\text{ge}} + \frac{\partial c^w}{\partial b} \cdot \left(\frac{\partial \epsilon_s^c}{\partial c^w}\right)^{\text{ge}}\right] \cdot \frac{\partial n^c}{\partial \epsilon_s^c}}_{\text{indirect effect}}$$

$$\left(\frac{\partial f}{\partial b}\right)^{\text{ge}} = \underbrace{\left(\frac{\partial S}{\partial b} \cdot \left(-\frac{\partial f}{\partial S}\right)^{\text{ge}}\right)}_{\text{direct effect}} + \underbrace{\left[\frac{\partial \epsilon_s^c}{\partial b} \cdot \frac{\partial c^w}{\partial \epsilon_s^c} + \frac{\partial c^w}{\partial b}\right] \cdot \left(\frac{\partial f}{\partial c^w}\right)^{\text{ge}}}_{\text{indirect effect}}$$

$$\left(\frac{\partial \theta}{\partial b}\right)^{\text{ge}} = \underbrace{\frac{\partial S}{\partial b} \cdot \left(\frac{\partial \theta}{\partial S}\right)^{\text{ge}}}_{\text{direct effect}} + \underbrace{\left[\frac{\partial \epsilon_s^c}{\partial b} \cdot \frac{\partial c^w}{\partial \epsilon_s^c} + \frac{\partial c^w}{\partial b}\right] \cdot \left(\frac{\partial \theta}{\partial c^w}\right)^{\text{ge}}}_{\text{indirect effect}}$$

Inserting the Lagrange multiplier gives

$$(1-n)\cdot\frac{\left(u'(b)-u'(c^w)\right)}{u'(c^w)} = \left(-\frac{\partial f}{\partial b}\right)^{\text{ge}}\cdot u\cdot L_v + \left(\frac{\partial \epsilon_s^c}{\partial b}\right)^{\text{ge}}\cdot\frac{\partial n^c}{\partial \epsilon_s^c}\cdot L_s^c + n^c\cdot\left(-\frac{\partial \theta}{\partial b}\right)^{\text{ge}}\cdot \hat{IE}_\theta$$

Note that when we set the lay-off tax optimally, then UI will not impose any distortions on the separation condition of unconstrained firms. In the context of Proposition 1, this means that $MLS^u = 0$, bolstering the fact that optimal lay-off taxes can implement the optimal number of separations.

Appendix 3.F Ramsey FOCs with STW

In the following, multipliers from the Lagrangian, implied by the maximization problem from the previous section, are denoted by λ_{idx}, the index depending on the constraint. Here, λ_n denotes the Lagrange multiplier for the total employment equation, λ_{n^u} for the number of unconstrained firms, λ_{n^c} for the number of constrained firms, λ_{n^s} for the number of firms that received a shock, λ_θ for the job-creation condition, λ_c for the wage equation, $\lambda_{\epsilon_s^u}$ for the separation condition of unconstrained firms with STW support, $\lambda_{\xi_s^c}$ for the separation condition of constrained firms without STW support, and $\lambda_{\epsilon_s^c}$ for the separation condition of constrained firms with STW support. Every other equation listed in the Ramsey problem for STW is assumed to be plugged in.

3.F.1 Employment

$$\begin{aligned}\frac{\partial \mathscr{L}}{\partial n^s} &= -\lambda_{n^s} + \frac{1-p}{\lambda}(1-\rho^u)\cdot\left(u(c^w) - u'(c^w)\cdot c^w\right)\\ &\quad + \frac{p}{\lambda}(1-\rho^c)\cdot\left(u^c - u'(c^w)\cdot e^c\right)\\ &\quad + \frac{1-p}{\lambda}\cdot(1-\rho^c)\cdot\lambda_{n^u} + \frac{p}{\lambda}(1-\rho^c)\cdot\lambda_{n^c} \;=\; 0\end{aligned}$$

$$\frac{\partial \mathscr{L}}{\partial n^u} = -\lambda_{n^u} + \lambda_n \;=\; 0$$

$$\frac{\partial \mathscr{L}}{\partial n^c} = -\lambda_{n^c} + \lambda_n \;=\; 0$$

$$\begin{aligned}\frac{\partial \mathscr{L}}{\partial n} &= -\lambda_n + \lambda_{n^s}\cdot(\lambda - q(\theta)\cdot\theta)\\ &\quad + u'(c^w)\cdot[(z-\Omega) + b + \theta\cdot k_v] - u(b)\\ &\quad + \frac{(1-\rho)}{\lambda}\cdot\frac{b}{n^2}\cdot\lambda_\theta + \eta(1-\rho)\cdot\frac{b}{n^2}\cdot\lambda_c\end{aligned}$$

$$\begin{aligned}\Leftrightarrow\quad \frac{\lambda_n}{u'(c^w)} &= (\lambda - q(\theta)\cdot\theta) + [z - \Omega + b + \theta\cdot k_v] - \frac{u(b)}{u'(c^w)}\\ &\quad + \frac{1-\rho}{\lambda}\cdot\frac{b}{n^2}\cdot\frac{\lambda_\theta}{u'(c^w)} + \eta\cdot(1-\rho)\cdot\frac{b}{n^2}\cdot\frac{\lambda_c}{u'(c^w)}\end{aligned}$$

3.F.2 Job Creation

Before we start, let us define the Insurance effect:

$$IE_\theta \;=\; n^s\cdot u'(c^w)\cdot\left(1 + \frac{\lambda_\theta}{n^s\cdot u'(c^w)}\right)\cdot \tilde{IE}_\theta$$

$$\tilde{IE}_\theta \;=\; \frac{p}{\lambda}\left(\int_{\max\{\xi_s^c,\epsilon_{stw}\}}^{\epsilon_p}\frac{\lambda}{f}\cdot\gamma\cdot\frac{u'(c(\epsilon)) - u'(c^w)}{u'(c^w)}\,dG(\epsilon)\right.$$

$$\left. + \int_{\epsilon_s^c}^{\max\{\epsilon_{\text{stw}},\epsilon_s^c\}}\frac{\lambda}{f}\cdot\gamma\cdot\frac{u'(c_{\text{stw}}(\epsilon)) - u'(c^w)}{u'(c^w)}\,dG(\epsilon)\right)\cdot k_v$$

The FOC for the labor market tightness can be written as:

$$\begin{aligned}
\frac{\partial \mathscr{L}}{\partial \theta} \;=\; & -k_v(1-n)\cdot u'(c^w) + IE_\theta + (1-\gamma)\cdot q(\theta)\cdot(1-n)\cdot\lambda_{n^s} \\
& - \frac{1}{1-\eta}\cdot\gamma\cdot k_v + \frac{\lambda_\theta}{f} + \frac{1}{1-\eta}\cdot\frac{\lambda\gamma - f\eta}{f}\cdot\left(\frac{1}{\lambda}\cdot(1-\rho)\cdot\lambda_\theta - \lambda_{\epsilon_s^u}\right) \\
& - \left(\lambda\cdot\frac{\gamma}{f}\cdot u'(c^w_{stw}(\epsilon_s^c)) + \left(\frac{1}{1-\eta}\cdot\frac{\lambda\gamma - f}{f}\right)\eta\right)\cdot\lambda_{\epsilon_s^c} \\
& - \left(\lambda\cdot\frac{\gamma}{f}\cdot u'(c^w(\xi_s^c)) + \frac{1}{\lambda}\cdot\left(\frac{1}{1-\eta}\cdot\frac{\lambda\gamma - f}{f}\right)\eta\right)\cdot\lambda_{\xi_s^c} \\
& - \lambda_c\cdot\frac{\partial WE}{\partial\theta}
\end{aligned}$$

$$\iff \qquad \frac{\lambda_{n^s}}{u'(c^w)} \;=\; \frac{1+\chi}{1-\gamma}\cdot\frac{k_v}{q}$$

$$\begin{aligned}
\text{with}\quad \chi \;=\; & \frac{1}{(1-n)\cdot u'(c^w)}\cdot\frac{1}{f} \\
\cdot\; & \frac{1}{1-\eta}\cdot\left(\gamma\cdot\lambda_\theta - (\lambda\cdot\gamma - f\cdot\eta)\cdot\left(\frac{1}{\lambda}\cdot(1-\rho)\cdot\lambda_\theta - \lambda_{\epsilon_s^u}\right)\right) \\
& + \frac{1}{(1-n)\cdot u'(c^w)}\left(\lambda\cdot\frac{\gamma}{f}u'(c^w_{stw}(\epsilon_s^c)) + \left(\frac{1}{1-\eta}\cdot\frac{\lambda\gamma - f}{f}\right)\eta\right)\cdot\lambda_{\epsilon_s^c} \\
& + \frac{1}{(1-n)\cdot u'(c^w)}\left(\lambda\cdot\frac{\gamma}{f}u'(c^w(\xi_s^c)) + \left(\frac{1}{1-\eta}\cdot\frac{\lambda\gamma - f}{f}\right)\eta\right)\cdot\lambda_{\xi_s^c} \\
& + \frac{\lambda_c}{(1-n)\cdot u'(c^w)}\cdot\frac{\partial WE}{\partial\theta}/k_v \\
& - \frac{1}{(1-n)\cdot u'(c^w)}\cdot IE_\theta/k_v
\end{aligned}$$

Collecting terms from FOCs for employment and job-creation conditions gives:

$$\begin{aligned}
u'(c^w)\cdot\frac{1+\chi}{1-\gamma}\cdot\frac{k_v}{q} \;=\; & \frac{1-p}{\lambda}\cdot(1-\rho^u)\cdot\left(u(c^w) - u(b) - u'(c^w)\cdot c^w\right) \\
& + \frac{p}{\lambda}\cdot(1-\rho^c)\cdot\left(u^c - u(b) - u'(c^w)\cdot e^c\right) \\
& + \left(\frac{1-\rho}{\lambda}\cdot\frac{b}{n}\cdot\frac{\lambda_\theta}{n^s}\right) + \left(\frac{1-\rho}{\lambda}\cdot\frac{\lambda_c}{n^s}\cdot\eta\cdot\lambda_c\cdot\frac{b}{n}\right) \\
& + u'(c^w)\cdot\frac{1-\rho}{\lambda}\cdot\left(z - \Omega + b + \frac{\lambda - \gamma f + \chi(1-f)}{1-\gamma}\cdot\frac{k_v}{q}\right)
\end{aligned}$$

Optimal Job Creation Condition:

$$
\begin{aligned}
\frac{1+\chi}{1-\gamma}\cdot\frac{k_v}{q} = &\ \frac{1-\rho}{\lambda}\cdot(z-\Omega+b)\\
&+\frac{1-\rho}{\lambda}\cdot\frac{b}{n}\cdot\frac{\lambda_\theta}{n^s\cdot u'(c^w)}\\
&+\frac{1-\rho}{\lambda}\cdot\frac{b}{n}\cdot\frac{\lambda_c\cdot\lambda}{n^s\cdot u'(c^w)}\\
&+\frac{1-p}{\lambda}\cdot(1-\rho^u)\cdot\left(\frac{u(c^w)-u(b)}{u'(c^w)}-c^w\right)\\
&+\frac{p}{\lambda}\cdot(1-\rho^c)\cdot\left(\frac{u^c-u(b)}{u'(c^w)}-e^c\right)\\
&+\frac{1-\rho}{\lambda}\cdot\frac{\lambda-\gamma f+\chi(\lambda-f)}{1-\gamma}\cdot\frac{k_v}{q}
\end{aligned}
$$

Subtracting the decentralized job-creation condition from the optimal gives:

$$
\begin{aligned}
\left(\chi-\frac{\eta-\gamma}{1-\eta}\right)\cdot\frac{1}{1-\gamma}\cdot\frac{k_v}{q} = &\left(1+\frac{\lambda_\theta+\eta\cdot\lambda\cdot\lambda_c}{n^s\cdot u'(c^w)}\right)\cdot\frac{1-\rho}{\lambda}\cdot\frac{b}{n}\\
&+\frac{(1-\rho)(\lambda-f)}{\lambda}\cdot\left(\chi-\frac{\eta-\gamma}{1-\eta}\right)\cdot\frac{1}{1-\gamma}\cdot\frac{k_v}{q}
\end{aligned}
$$

Rearranging gives:

$$
\begin{aligned}
\left(\chi-\frac{\eta-\gamma}{1-\eta}\right)\cdot\frac{1}{1-\gamma}\cdot\frac{k_v}{q} &= \left(1+\frac{\lambda_\theta+\eta\cdot\lambda\cdot\lambda_c}{n^s\cdot u'(c^w)}\right)\cdot\frac{\frac{1-\rho}{\lambda}}{\rho+(1-\rho)\cdot\frac{f}{\lambda}}\cdot\frac{b}{n}\\
&= \left(1+\frac{\lambda_\theta+\eta\cdot\lambda\cdot\lambda_c}{n^s\cdot u'(c^w)}\right)\cdot\frac{b}{f}
\end{aligned}
$$

3.F.3 Lagrange Multipliers for Separation Condition

Optimal separation condition unconstrained firms

$$
\begin{aligned}
&-\frac{\partial\mathcal{L}}{\partial\epsilon_s^u}\\
&=\frac{n}{1-\rho}\cdot(1-p)\cdot g(\epsilon_s^u)\cdot u'(c^w)\cdot\left(z(\epsilon_s^u)-\Omega(\epsilon_s^u)+\frac{u(c^w)}{u'(c^w)}-c^w-(z+\Omega)\right)\\
&+\lambda_\theta\cdot\frac{1-p}{\lambda}\cdot g(\epsilon_s^u)\cdot\Bigg[z(\epsilon_s^u)-\Omega(\epsilon_s^u)+\frac{u(c^w)-u(b)}{u'(c^w)}-c^w\\
&-\frac{1-n}{n}\cdot b+\frac{\lambda-\eta f}{1-\eta}\cdot\frac{k_v}{q}\Bigg]\\
&+\lambda_c\cdot\frac{\partial WE}{\partial\epsilon_s^u}+\frac{1-p}{\lambda}\cdot g(\epsilon_s^u)\cdot n^s\cdot\lambda_{n^u}+\lambda_{\epsilon_s^u}\cdot\frac{\partial S^u_{\text{stw}}(\epsilon_s^u)}{\partial\epsilon_s^u}\overset{!}{=}0
\end{aligned}
$$

Insert for λ_{n^u}:

$$\begin{aligned}
\iff \quad & g(\epsilon_s^u) \cdot \frac{1-p}{\lambda} \cdot n^s \\
& \cdot \left[z(\epsilon_s^u) - \Omega(\epsilon_s^u) + \frac{u(c^w) - u(b)}{u'(c^w)} - c^w - (z + \Omega) + \frac{\lambda_{n^u}}{u'(c^w)} \right] \\
& - \frac{\lambda_\theta}{u'(c^w)} \cdot \frac{1-p}{\lambda} \cdot g(\epsilon_s^u) \cdot \left[\tau_{\text{stw}} \cdot (\bar{h} - h_{\text{stw}}(\epsilon_s^u)) + \frac{1-n}{n} \cdot b \right] \\
& + \frac{\lambda_c}{u'(c^w)} \cdot \frac{\partial WE}{\partial \epsilon_s^u} + \frac{\lambda_{\epsilon_s^u}}{u'(c^w)} \cdot \frac{\partial S_{\text{stw}}^u(\epsilon_s^u)}{\partial \epsilon_s^u} \overset{!}{=} 0
\end{aligned}$$

$$\begin{aligned}
\iff \quad & z(\epsilon_s^u) - \Omega(\epsilon_s^u) + \frac{u(c^w) - u(b)}{u'(c^w)} - c^w + b + \big(\lambda - q(\theta) \cdot \theta\big) \cdot \frac{\lambda_{n^s}}{u'(c^w)} + \theta \cdot c \\
& - \frac{\lambda_\theta}{n^s \cdot u'(c^w)} \cdot \left(\tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon_s^u)) + \frac{1-n}{n} \cdot b - \frac{b}{n} \right) \\
& + \lambda_c \cdot \frac{1}{n^s \cdot u'(c^w)} \cdot \left(\eta \cdot \frac{b}{n} + \frac{\lambda}{1-p} \cdot \frac{1}{g(\epsilon_s^u)} \cdot \frac{\partial WE}{\partial \epsilon_s} \right) \\
& + \lambda_{\epsilon_s^u} \cdot \frac{1}{n^s \cdot u'(c^w)} \cdot \frac{\lambda}{1-p} \cdot \frac{1}{g(\epsilon_s^u)} \cdot \frac{\partial S_{\text{stw}}^u(\epsilon_s^u)}{\partial \epsilon_s^u} = 0
\end{aligned}$$

Insert for λ_{n^s} :

$$\begin{aligned}
\iff \quad & z(\epsilon_s^u) - \Omega(\epsilon_s^u) + \frac{u(c^w) - u(b)}{u'(c^w)} - c^w + b \\
& + (\lambda - q(\theta) \cdot \theta) \cdot \frac{\lambda_{n^s}}{u'(c^w)} + \theta \cdot c \\
& - \frac{\lambda_\theta}{n^s \cdot u'(c^w)} \left(\tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon_s^u)) + \frac{1-n}{n} \cdot b - \frac{b}{n} \right) \\
& + \frac{\lambda_c}{n^s \cdot u'(c^w)} \left(\eta \cdot \frac{b}{n} + \frac{\lambda}{1-p} \cdot \frac{1}{g(\epsilon_s^u)} \cdot \frac{\partial WE}{\partial \epsilon_s^u} \right) \\
& + \frac{\lambda_{\epsilon_s^u}}{n^s \cdot u'(c^w)} \cdot \frac{1}{n^s} \cdot \frac{\lambda}{1-p} \cdot \frac{1}{g(\epsilon_s^u)} \cdot \frac{\partial S_{\text{stw}}^u(\epsilon_s^u)}{\partial \epsilon_s^u} = 0
\end{aligned}$$

Insert λ_{n^s}

$$
\begin{aligned}
\Longleftrightarrow\ & z(\epsilon_s^u) - \Omega(\epsilon_s^u) + \frac{u(c^w) - u(b)}{u'(c^w)} - c^w + b \\
& + \frac{\lambda - \gamma f + \chi(\lambda - f)}{1-\gamma} \cdot \frac{k_v}{q} \\
& - \frac{\lambda_\theta}{n^s \cdot u'(c^w)} \left(\tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon_s^u)) - b\right) \\
& + \lambda \cdot \frac{\lambda_c}{n^s \cdot u'(c^w)} \left(\eta \cdot \frac{b}{n} + \frac{\lambda}{1-p} \cdot \frac{1}{g(\epsilon_s^u)} \cdot \frac{\partial WE}{\partial \epsilon_s^u}\right) \\
& + \frac{\lambda_{\epsilon_s^u}}{n^s \cdot u'(c^w)} \cdot \frac{\lambda}{1-p} \cdot \frac{1}{g(\epsilon_s^u)} \cdot \frac{\partial S_{\text{stw}}^u(\epsilon_s^u)}{\partial \epsilon_s^u} \ = \ 0
\end{aligned}
$$

Subtract decentralized separation condition:

$$
\begin{aligned}
\tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon_s^u)) \ = \ & b + \left(\lambda - \theta q(\theta)\right) \cdot \left(\chi - \frac{\eta - \gamma}{1-\eta}\right) \cdot \frac{1}{1-\gamma} \cdot \frac{k_v}{q} \\
& + \frac{\lambda_{\epsilon_s^u}}{n^s \cdot u'(c^w)} \cdot \frac{\lambda}{1-p} \cdot \frac{1}{g(\epsilon_s^u)} \cdot \frac{\partial S_{\text{stw}}^u(\epsilon_s^u)}{\partial \epsilon_s^u} \\
& - \frac{\lambda_\theta \cdot \theta}{n^s \cdot u'(c^w)} \cdot \left(\tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon_s^u)) - b\right) \\
& + \frac{\lambda_c}{n^s \cdot u'(c^w)} \left(\eta \cdot \frac{b}{n} + \frac{\lambda}{1-p} \cdot \frac{1}{g(\epsilon_s^u)} \cdot \frac{\partial WE}{\partial \epsilon_s^u}\right)
\end{aligned}
$$

Hence

$$
\begin{aligned}
0 \ = \ & \left(1 + \frac{\lambda_\theta}{u'(c^w) \cdot n^s}\right) \\
& \cdot \left(b + (\lambda - f) \cdot \frac{(1 - \rho/\lambda)}{(\rho + (1-\rho) \cdot f/\lambda)} \cdot \frac{b}{n} - \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon_s^u))\right) \\
& + \frac{\lambda}{1-p} \cdot \frac{1}{g(\epsilon_s^u)} \cdot \frac{\lambda_{\epsilon_s^u}}{n^s \cdot u'(c^w)} \cdot \frac{\partial S_{\text{stw}}^u(\epsilon_s^u)}{\partial \epsilon_s^u} \\
& + \frac{\lambda_c}{n^s \cdot u'(c^w)} \left(\eta \cdot \frac{b}{n} + \frac{\lambda}{1-p} \cdot \frac{1}{g(\epsilon_s^u)} \cdot \frac{\partial WE}{\partial \epsilon_s^u}\right) \\
\Longleftrightarrow \qquad 0 \ = \ & \left(1 + \frac{\lambda_\theta}{u'(c^w) \cdot n^s}\right) \cdot \left(\frac{\lambda}{f} \cdot b - \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon_s^u))\right) \\
& + \frac{\lambda}{1-p} \cdot \frac{1}{g(\epsilon_s^u)} \cdot \frac{\lambda_{\epsilon_s^u}}{n^s \cdot u'(c^w)} \cdot \frac{\partial S_{\text{stw}}^u(\epsilon_s^u)}{\partial \epsilon_s^u} \\
& + \frac{\lambda_c}{n^s \cdot u'(c^w)} \left(\eta \cdot \frac{b}{n} + \frac{\lambda}{1-p} \cdot \frac{1}{g(\epsilon_s^u)} \cdot \frac{\partial WE}{\partial \epsilon_s^u}\right)
\end{aligned}
$$

And finally:

$$
\begin{aligned}
-\frac{\lambda_{\epsilon_s^u}}{n^s \cdot u'(c^w)} \;=\; & \frac{(1-p)\cdot g(\epsilon_s^u)}{\lambda} \cdot \frac{1}{\frac{\partial S^u_{\text{stw}}(\epsilon_s^u)}{\partial \epsilon_s^u}} \cdot \left(\left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)} \right) \right. \\
& \cdot \left(\frac{\lambda}{f} \cdot b - \tau_{\text{stw}} \cdot \left(\bar{h} - h_{\text{stw}}(\epsilon_s^u) \right) \right) \\
& \left. + \frac{\lambda_c}{n^s \cdot u'(c^w)} \cdot \left(\eta \cdot \frac{b}{n} + \frac{\lambda}{1-p} \cdot \frac{1}{g(\epsilon_s^u)} \cdot \frac{\partial WE}{\partial \epsilon_s^u} \right) \right)
\end{aligned}
$$

Optimal separation condition constrained firms

$$
\begin{aligned}
-\frac{\partial \mathscr{L}}{\partial \epsilon_s^c} \;=\; & \left[\frac{n}{1-\rho} \cdot p \cdot g(\epsilon_s^c) \cdot u'(c^w) \left(z(\epsilon_s^c) - \Omega(\epsilon_s^c) \right. \right. \\
& \left. + \frac{u(c_{\text{stw}}(\epsilon_s^c))}{u'(c^w)} - c_{\text{stw}}(\epsilon_s^c) - (z + \Omega) \right) \\
& + \lambda_\theta \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^c) \left[z(\epsilon_s^c) - \Omega(\epsilon_s^c) + \frac{u(c_{\text{stw}}(\epsilon_s^c)) - u(b)}{u'(c^w)} - c^w(\epsilon_s^c) \right. \\
& \left. - \frac{1-n}{n} \cdot b + \frac{\lambda - \eta f}{1-\eta} \cdot \frac{k_v}{q} \right] \\
& \left. - \lambda_c \cdot \frac{\partial WE}{\partial \epsilon_s^c} + \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot n^s \cdot \lambda_{n^c} \right] \cdot \mathbb{1}(\epsilon_{\text{stw}} \;\geq\; \epsilon_s^c) + \lambda^c_{\epsilon_s^c} \cdot \frac{\partial S^u_{\text{stw}}(\epsilon_s^c)}{\partial \epsilon_s^c} \;=\; 0
\end{aligned}
$$

Subtract decentralized separation condition (constrained firms):

$$
\begin{aligned}
-\frac{\partial \mathscr{L}}{\partial \epsilon_s^c} \;=\; & \left[\frac{n}{1-\rho} \cdot p \cdot g(\epsilon_s^c) \left(z_{\text{stw}}(\epsilon_s^c) - \Omega(\epsilon_s^c) + \frac{u(c_{\text{stw}}(\epsilon_s^c)) - u(c^w)}{u'(c^w)} - (z + \Omega) \right) \right. \\
& + \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot n^s \cdot \frac{\lambda_{n^c}}{u'(c^w)} \\
& - \frac{\lambda_\theta}{u'(c^w)} \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^c) \left[\tau_{\text{stw}} \cdot (\bar{h} - h_{\text{stw}}(\epsilon_s^c)) + \frac{1-n}{n} \cdot b \right] \\
& \left. + \frac{\lambda_c}{u'(c^w)} \cdot \frac{\partial WE}{\partial \epsilon_s^c} \right] \cdot \mathbb{1}(\epsilon_{\text{stw}} \;\geq\; \epsilon_s^c) + \frac{\lambda^c_{\epsilon_s^c}}{u'(c^w)} \cdot \frac{\partial S^c_{\text{stw}}(\epsilon_s^c)}{\partial \epsilon_s^c} \;=\; 0
\end{aligned}
$$

Insert λ_{n^c} and λ_n

$$\iff \quad -\frac{\partial \mathscr{L}}{\partial \epsilon_s^c} = \Bigg[\frac{p}{\lambda} \cdot n^s g(\epsilon_s^c)\Bigg(z_{\text{stw}}(\epsilon_s^c) - \Omega(\epsilon_s^c) + \frac{u(c_{\text{stw}}(\epsilon_s^c))}{u'(c^w)} - u(c^w) - (z + \Omega)\Bigg)$$
$$+ \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot n^s \cdot \Bigg((\lambda - q(\theta) \cdot \theta) \cdot \frac{\lambda_{n^s}}{u'(c^w)}$$
$$+ [z - \Omega + b + \theta \cdot c] - \frac{u(b)}{u'(c^w)}$$
$$+ \frac{1-\rho}{\lambda} \cdot \frac{b}{n^2} \cdot \lambda_\theta \cdot \frac{1}{u'(c^w)} + \eta \cdot \frac{b}{n^2} \cdot \frac{\lambda_c}{u'(c^w)}\Bigg)\Bigg]$$
$$- \frac{\lambda_\theta}{u'(c^w)} \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot \Bigg[\tau_{\text{stw}} \cdot (\bar{h} - h_{\text{stw}}(\epsilon_s^c)) + \frac{1-n}{n} \cdot b\Bigg]$$
$$+ \frac{\lambda_c}{u'(c^w)} \cdot \frac{\partial WE}{\partial \epsilon_s^c} \cdot \mathbb{1}(\epsilon_{\text{stw}} \geq \epsilon_s^c) + \frac{\lambda_{\epsilon_s^c}^c}{u'(c^w)} \cdot \frac{\partial S_{\text{stw}}^c(\epsilon_s^c)}{\partial \epsilon_s^c} = 0$$

Insert λ_{n^s}

$$-\frac{\partial \mathscr{L}}{\partial \epsilon_s^c} = \Bigg[z_{\text{stw}}(\epsilon_s^c) - \Omega(\epsilon_s^c) + \frac{u(c_{\text{stw}}(\epsilon_s^c))}{u'(c^w)} - c_{\text{stw}}(\epsilon_s^c) + b$$
$$+ \frac{\lambda - \gamma \cdot f + \chi \cdot (\lambda - f)}{1-\gamma} \cdot \frac{k_v}{q}$$
$$- \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\left[\tau_{\text{stw}} \cdot (\bar{h} - h_{\text{stw}}(\epsilon_s^c)) - b\right]$$
$$+ \frac{\lambda_c}{n^s \cdot u'(c^w)}\Bigg(\eta \cdot \frac{b}{n} + \frac{1}{g(\epsilon_s^c)} \cdot \frac{\lambda}{p} \cdot \frac{\partial WE}{\partial \epsilon_s^c}\Bigg)\Bigg] \cdot \mathbb{1}(\epsilon_{\text{stw}} \geq \xi_s^c)$$
$$+ \frac{\lambda}{p} \cdot \frac{1}{n^s \cdot g(\epsilon_s^c)} \cdot \frac{\lambda_{\epsilon_s^c}^c}{u'(c^w)} \cdot \frac{\partial S_{\text{stw}}^c(\epsilon_s^c)}{\partial \epsilon_s^c} = 0$$

And finally:

$$-\frac{\lambda_{\epsilon_s^c}}{n^s \cdot u'(c^w)} = \frac{p}{\lambda} \cdot \frac{g(\epsilon_s^c)}{\frac{\partial s_{\text{stw}}^c(\epsilon_s^c)}{\partial \epsilon_s^c}} \cdot \Bigg(\Bigg(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\Bigg) \cdot \Bigg(\frac{\lambda}{f} \cdot b - \tau_{\text{stw}} \cdot (\bar{h} - h_{\text{stw}}(\epsilon_s^c))\Bigg)$$
$$+ \frac{\lambda_c}{n^s \cdot u'(c^w)}\Bigg(\eta \cdot \frac{b}{n} + \frac{1}{g(\epsilon_s^c)} \cdot \frac{\lambda}{p} \cdot \frac{\partial WE}{\partial \epsilon_s^c}\Bigg)\Bigg) \cdot \mathbb{1}(\epsilon_{\text{stw}} \geq \epsilon_s^c)$$

Analogously, it can be derived that

$$-\frac{\lambda_{\xi_s^c}}{n^s \cdot u'(c^w)} = \frac{p}{\lambda} \cdot \frac{g(\xi_s^c)}{\frac{\partial s_{\text{stw}}^c(\xi_s^c)}{\partial \xi^c}} \cdot \Bigg(\Bigg(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\Bigg) \cdot \Bigg(\frac{\lambda}{f} \cdot b\Bigg)$$
$$+ \frac{\lambda \cdot \lambda_c}{n^s \cdot u'(c^w)}\Bigg(\eta \cdot \frac{b}{n} + \frac{1}{g(\xi_s^c)} \cdot \frac{\lambda}{p} \cdot \frac{\partial WE}{\partial \xi^c}\Bigg)\Bigg) \cdot \mathbb{1}(\epsilon_{\text{stw}} \leq \xi_s^c)$$

3.F.4 Optimal STW Benefits

$$
\begin{aligned}
\frac{\partial \mathscr{L}}{\partial \tau_{\text{stw}}} = {} & \frac{p}{\lambda} \cdot n^s \cdot \int_{\epsilon_s^c}^{\max\{\epsilon_{\text{stw}}} \left(u'(c_{\text{stw}}(\epsilon)) - u'(c^w)\right)\left(\bar{h} - h_{\text{stw}}(\epsilon)\right) dG(\epsilon) \\
& + \lambda_\theta \cdot \frac{p}{\lambda} \cdot \int_{\epsilon_s^c}^{\max\{\epsilon_{\text{stw}},\epsilon_s^c\}} \left(\frac{u'(c_{\text{stw}}(\epsilon)) - u'(c^w)}{u'(c^w)}\right)\left(\bar{h} - h_{\text{stw}}(\epsilon)\right) dG(\epsilon) \\
& - \frac{n^s \cdot (1-\rho)}{\lambda} \cdot u'(c^w) \cdot \frac{\partial \Omega}{\partial \tau_{\text{stw}}} - \lambda_\theta \cdot \frac{(1-\rho)}{\lambda} \cdot \frac{\partial \Omega}{\partial \tau_{\text{stw}}} - \lambda_c \cdot \frac{\partial WE}{\partial \tau_{\text{stw}}} \\
& - \lambda_{\epsilon_s^u} \cdot \frac{\partial S_{\text{stw}}^u(\epsilon_s^u)}{\partial \tau_{\text{stw}}} - \lambda_{\epsilon_s^c} \cdot \frac{\partial S_{\text{stw}}^c(\epsilon_s^c)}{\partial \tau_{\text{stw}}} = 0
\end{aligned}
$$

Insert Lagrange multiplier separation condition:

$$
\begin{aligned}
& \left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right) \cdot \left(\frac{p}{\lambda} \int_{\epsilon_s^c}^{\max\{\epsilon_s^c,\epsilon_{\text{stw}}\}} \frac{u'(c_{\text{stw}}(\epsilon)) - u'(c^w)}{u'(c^w)} \cdot (\bar{h} - h_{\text{stw}}(\epsilon))\, dG(\epsilon) \right. \\
& \qquad \left. -(1-\rho) \cdot \frac{\partial \Omega}{\partial \tau_{\text{stw}}}\right) - \frac{\lambda_c}{n^s \cdot u'(c^w)} \cdot \frac{\partial WE}{\partial \tau_{\text{stw}}} \\
& + \frac{(1-p)}{\lambda} \cdot g(\epsilon_s^u) \cdot \frac{\partial \epsilon_s^u}{\partial \tau_{\text{stw}}} \cdot \left(\left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right) \cdot \left(\frac{\lambda}{f} \cdot b - \tau_{\text{stw}} \cdot (\bar{h} - h_{\text{stw}}(\epsilon_s^u))\right)\right. \\
& \qquad \left. + \frac{\lambda_c}{n^s \cdot u'(c^w)} \cdot \left(\eta \cdot \frac{b}{n} + \frac{1}{g(\epsilon_s^u)} \cdot \frac{\lambda}{1-p} \cdot \frac{\partial WE}{\partial \epsilon_s^u}\right)\right) \\
& + \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot \frac{\partial \epsilon_s^c}{\partial \tau_{\text{stw}}} \cdot \left(\left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right) \cdot \left(\frac{\lambda}{f} \cdot b - \tau_{\text{stw}} \cdot (\bar{h} - h_{\text{stw}}(\epsilon_s^c))\right)\right. \\
& \qquad \left. + \frac{\lambda_c}{n^s \cdot u'(c^w)} \cdot \left(\eta \cdot \frac{b}{n} + \frac{1}{g(\epsilon_s^c)} \cdot \frac{\lambda}{p} \cdot \frac{\partial WE}{\partial \epsilon_s^c}\right)\right) = 0
\end{aligned}
$$

Finally:

$$
\begin{aligned}
\bar{\tau}_{\text{stw}} = {} & \underbrace{\frac{\lambda}{f} b}_{\text{Fiscal Ext.}} \\
& + \underbrace{\frac{\mathbb{1}\{\epsilon_{\text{stw}} \geq \epsilon_s^c\}}{\varphi(p)} \cdot n^s \cdot \frac{p}{\lambda} \int_{\epsilon_s^c}^{\epsilon_{\text{stw}}} \left(\frac{u'(c_{\text{stw}}(\epsilon)) - u'(c^w))}{u'(c^w)}\right)(\bar{h} - h_{\text{stw}}(\epsilon))\, dG(\epsilon)}_{\text{Insurance}} \\
& \underbrace{- \frac{n^s}{\varphi(p)} \cdot \frac{1}{\lambda} \cdot (1-\rho) \frac{\partial \Omega}{\partial \tau_{\text{stw}}}}_{\text{Distortion}} + \underbrace{BE_{\text{stw},3}}_{\text{Bargaining Effect}}
\end{aligned}
$$

$$
\bar{\tau}_{\text{stw}} = \frac{\varphi^u(p)}{\varphi(p)} \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon_s^u)) + \frac{\varphi^c(p)}{\varphi(p)} \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon_s^c))
$$

where

$$\varphi^u = \frac{(1-p)}{\lambda} \cdot g(\epsilon_s^u) \cdot \frac{\partial \epsilon_s^u}{\partial \tau_{\text{stw}}} \cdot n^s, \quad \varphi^c = \mathbb{1}(\epsilon_{\text{stw}} \geq \epsilon_s^c) \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot \frac{\partial \epsilon_s^c}{\partial \tau_{\text{stw}}} \cdot n^s,$$

$$\varphi = \varphi^c + \varphi^u$$

and

$$\begin{aligned} BE_{\text{stw},3} = & \frac{1}{\left(1 + \frac{\lambda_\theta}{n^s u'(c^w)}\right)} \\ & \times \left[-\frac{\lambda_c}{\varphi\, n^s u'(c^w)} \frac{\partial WE}{\partial \tau_{\text{stw}}} + \frac{\varphi^u}{\varphi} \frac{\lambda_c}{n^s u'(c^w)} \left(\eta \frac{b}{n} + \frac{1}{g(\epsilon_s^u)} \frac{\lambda}{1-p} \frac{\partial WE}{\partial \epsilon_s^u} \right) \right. \\ & \left. + \frac{\varphi^c}{\varphi} \frac{\lambda_c}{n^s u'(c^w)} \left(\eta \frac{b}{n} + \frac{1}{g(\epsilon_s^c)} \frac{\lambda}{p} \frac{\partial WE}{\partial \epsilon_s^c} \right) \right] \end{aligned}$$

3.F.5 Optimal Eligibility Condition

Assumption throughout: $\epsilon^p \geq \epsilon_{\text{stw}} \geq \epsilon_s^u$
Further assume that:

$$n^s \cdot u'(c^w) + \lambda_\theta - \eta \cdot \lambda_c > 0, \qquad n^s \cdot u'(c^w) + \lambda_\theta > 0$$

These conditions exclude the case that the Planner would want to use STW's hours distortions to destroy production and thus reduce vacancy posting. The Planner might want to do that when the Hosios condition is not fulfilled. For reasonable parameter values, however, the condition should hold anyway. Recall that by Lemma 1 $\epsilon_s^c > \epsilon_s^u$ and $\xi_s^c > \xi_s^u$.
Case I:
FOC for the eligibility threshold:

$$\begin{aligned} \frac{\partial \mathscr{L}}{\partial \epsilon_{stw}} &= -n^s \cdot u'(c^w) \cdot \frac{1-p}{\lambda} \cdot (1-\rho^u) \cdot \frac{\partial \Omega_u}{\partial \epsilon_{stw}} u'(c^w) \\ & \quad - (\lambda_\theta - \eta \cdot \lambda_c) \cdot \frac{1-p}{\lambda} \cdot (1-\rho^u) \cdot \frac{\partial \Omega_u}{\partial \epsilon_{stw}} \\ &= -\left(n^s \cdot u'(c^w) + \lambda_\theta - \eta \cdot \lambda_c\right) \cdot \frac{1-p}{\lambda} \cdot (1-\rho^u) \cdot \frac{\partial \Omega_u}{\partial \epsilon_{stw}} \end{aligned}$$

Under reasonable parameter values, we get:

$$\frac{\partial \mathscr{L}}{\partial \epsilon_{stw}} < 0 \quad \text{thus it is optimal to set} \quad \epsilon_{stw} = \epsilon_s^u.$$

Case II:
FOC for the eligibility threshold:

$$
\begin{aligned}
\frac{\partial \mathscr{L}}{\partial \epsilon_{stw}} = & -n^u \cdot \frac{\partial \Omega_u}{\partial \epsilon_{stw}} \cdot u'(c^w) \\
& + \frac{n}{1-\rho} \cdot p \cdot g(\epsilon_{stw}) \cdot u'(c^w) \\
& \quad \cdot \left(z_{\text{stw}}(\epsilon_{stw}) + \frac{u(c_{\text{stw}}(\epsilon_{stw}))}{u'(c^w)} - c_{\text{stw}}(\epsilon_{stw}) - (z + \Omega) \right) \\
& + \lambda_\theta \cdot \frac{(1-p)}{\lambda} \cdot (1-\rho^u) \cdot \frac{\partial \Omega_u}{\partial \epsilon_{stw}} \\
& + \frac{p}{\lambda} \cdot g(\epsilon_{stw}) \\
& \quad \cdot \left[z(\epsilon_{\text{stw}}) + \frac{u(c_{stw}(\epsilon_{\text{stw}})) - u(b)}{u'(c^w)} \right. \\
& \qquad \left. - c_{stw}(\epsilon_{stw}) - \frac{1-n}{n} \cdot b + \frac{\lambda - \eta \cdot f}{1-\eta} \cdot \frac{k_v}{q} \right] \\
& - \lambda_c \cdot \frac{\partial WE}{\partial \epsilon_{\text{stw}}} + \frac{p}{\lambda} \cdot g(\epsilon_{\text{stw}}) \cdot n^s \cdot \lambda_{n^c} \overset{!}{>} 0.
\end{aligned}
$$

Insert n^u

$$
\begin{aligned}
\frac{\partial \mathscr{L}}{\partial \epsilon_{\text{stw}}} = & -n^s \cdot (1-\rho^u) \cdot \left(\frac{1-p}{\lambda} \right) \cdot \frac{\partial \Omega_u}{\partial \epsilon_{\text{stw}}} \cdot u'(c^w) \\
& + n^s \cdot \frac{p}{\lambda} \cdot g(\epsilon_{\text{stw}}) \cdot u'(c^w) \cdot \left(z_{\text{stw}}(\epsilon_{\text{stw}}) + \frac{u(c_{\text{stw}}(\epsilon_{\text{stw}}))}{u'(c^w)} \right) \\
& - n^s \cdot \frac{p}{\lambda} \cdot g(\epsilon_{\text{stw}}) \cdot u'(c^w) \cdot \left(c_{\text{stw}}(\epsilon_{\text{stw}}) + z + \Omega \right) \\
& + \lambda_\theta \cdot \frac{g(\epsilon_{\text{stw}})}{\lambda} \cdot (1-p)(1-\rho^u) \cdot \frac{\partial \Omega_u}{\partial \epsilon_{\text{stw}}} \\
& + \frac{p}{\lambda} \cdot g(\epsilon_{\text{stw}}) \cdot \left(z_{\text{stw}}(\epsilon_{\text{stw}}) + \frac{u(c_{\text{stw}}(\epsilon_{\text{stw}})) - u(b)}{u'(c^w)} - c_{\text{stw}}(\epsilon_{\text{stw}}) - \frac{1-n}{n} \cdot b \right. \\
& \qquad \left. + \frac{\lambda - \eta \cdot f}{1-\eta} \cdot \left(\frac{k_v}{q} \right) \right) \\
& - \lambda_c \cdot \frac{\partial WE}{\partial \epsilon_{\text{stw}}} + \frac{p}{\lambda} \cdot g(\epsilon_{\text{stw}}) \cdot n^s \cdot \lambda_{n^c} \overset{!}{>} 0
\end{aligned}
$$

$$
\begin{aligned}
\Leftrightarrow\ & -(n^s \cdot u'(c^w) + \lambda_\theta) \cdot (1-\rho^u) \cdot \left(\frac{1-p}{\lambda}\right) \cdot \frac{\partial \Omega_u}{\partial \epsilon_{\text{stw}}} \\
& + n^s \cdot \frac{p}{\lambda} \cdot g(\epsilon_{\text{stw}}) \cdot u'(c^w) \\
& \cdot \left(z_{\text{stw}}(\epsilon_{\text{stw}}) + \frac{u(c_{\text{stw}}(\epsilon_{\text{stw}}))}{u'(c^w)} - c_{\text{stw}}(\epsilon_{\text{stw}}) - (z+\Omega) + \frac{\lambda_{n^c}}{u'(c^w)} \right) \\
& + \frac{p}{\lambda} \cdot g(\epsilon_{\text{stw}}) \cdot \left[z_{\text{stw}}(\epsilon_{\text{stw}}) + \frac{u(c_{\text{stw}}(\epsilon_{\text{stw}})) - u(b)}{u'(c^w)} - c_{\text{stw}}(\epsilon_{\text{stw}}) \right. \\
& \qquad \left. - \frac{1-n}{n} \cdot b + \frac{\lambda - \eta \cdot f}{1-\eta} \cdot \left(\frac{k_v}{q}\right) \right] \\
& - \lambda_c \cdot \frac{\partial WE}{\partial \epsilon_{\text{stw}}} \overset{!}{>} 0
\end{aligned}
$$

Insert λ_{n^c} and subtract decentralized separation condition for constrained firms:

$$
\begin{aligned}
\Leftrightarrow\ & -\left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right) \cdot \frac{(1-p)}{\lambda} \cdot (1-\rho^u) \cdot \frac{\partial \Omega_u}{\partial \epsilon_{\text{stw}}} \cdot \frac{1}{g(\epsilon_{\text{stw}})} \\
& + \frac{p}{\lambda} \cdot \left[b + z_{\text{stw}}(\epsilon_{stw}) + \frac{u(c_{\text{stw}}(\epsilon_{stw})) - u(b)}{u'(c^w)} - c_{\text{stw}}(\epsilon_{stw}) \right. \\
& \left. + \frac{\lambda - \eta \cdot f + \chi \cdot (\lambda - f)}{1-\eta} \cdot \left(\frac{k_v}{q}\right) \right] \\
& - \frac{\lambda_\theta}{n^s \cdot u'(c^w)} \cdot \frac{p}{\lambda} \cdot [\tau_{\text{stw}} \cdot (h - h_{\text{stw}}(\epsilon_{stw})) - b] \\
& - \frac{\lambda_c}{n^s \cdot u'(c^w)} \cdot \left(\frac{\partial WE}{\partial \epsilon_{\text{stw}}} \cdot \frac{1}{g(\epsilon_{\text{stw}})} \cdot \frac{\partial S_{\text{stw}}(\epsilon_{\text{stw}})}{\partial \epsilon_{\text{stw}}} + \frac{p}{\lambda} \cdot \eta \cdot \frac{b}{n} \right) \overset{!}{>} 0
\end{aligned}
$$

Subtract decentralized separation condition:

$$
\begin{aligned}
\Leftrightarrow\ & -\left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right) \cdot \frac{(1-p)}{\lambda} \cdot (1-\rho^u) \cdot \frac{\partial \Omega_u}{\partial \epsilon_{\text{stw}}} \cdot \frac{1}{g(\epsilon_{\text{stw}})} \\
& + \frac{p}{\lambda} \cdot \left[z_{\text{stw}}(\epsilon_{\text{stw}}) - z_{\text{stw}}(\epsilon_s) + \frac{u(c_{\text{stw}}(\epsilon_{\text{stw}})) - u(c_{\text{stw}}(\epsilon_s^c))}{u'(c^w)} \right. \\
& \quad \left. - \tau_{\text{stw}} \cdot (h - h_{\text{stw}}(\epsilon_s^c)) + b + (1-\rho) \cdot (\lambda - f) \cdot \left(\chi - \frac{\eta - \gamma}{1-\eta}\right) \cdot \frac{1}{1-\gamma} \cdot \frac{k_v}{q} \right] \\
& + \frac{\lambda_\theta}{n^s \cdot u'(c^w)} \cdot \frac{p}{\lambda} \cdot \left[z_{\text{stw}}(\epsilon_{\text{stw}}) - z_{\text{stw}}(\epsilon_s^c) + \frac{u(c_{\text{stw}}(\epsilon_{\text{stw}})) - u(c_{stw}(\epsilon_s^c))}{u'(c^w)} \right. \\
& \left. - \tau_{\text{stw}} \cdot (h - h_{\text{stw}}(\epsilon_s^c)) + b \right] \\
& - \frac{\lambda_c}{n^s \cdot u'(c^w)} \cdot \left(\frac{\partial WE}{\partial \epsilon_{\text{stw}}} \cdot \frac{1}{g(\epsilon_{\text{stw}})} \cdot \frac{\partial S_{\text{stw}}(\epsilon_{\text{stw}})}{\partial \epsilon_{\text{stw}}} + \frac{p}{\lambda} \cdot \eta \cdot \frac{b}{n} \right) \overset{!}{>} 0
\end{aligned}
$$

Inserting for $(\chi - \frac{\eta-\gamma}{1-\eta}) \cdot \frac{1}{1-\gamma} \cdot \frac{k_v}{q}$ gives us the final condition. We can distinguish between three cases:

(A)

$$
\begin{aligned}
& g(\epsilon_{\text{stw}}) \cdot \frac{p}{\lambda} \cdot n^s \cdot \Bigg[z_{\text{stw}}(\epsilon_{\text{stw}}) - z_{\text{stw}}(\epsilon_s^c) \\
& \quad + \frac{u(c_{\text{stw}}(\epsilon_{\text{stw}})) - u(c_{\text{stw}}(\epsilon_s^c))}{u'(c^w)} + \frac{\lambda}{f} \cdot b - \tau_{\text{stw}} \cdot \big(h - h_{\text{stw}}(\epsilon_s^c)\big) \Bigg] \\
& \quad > n^s \cdot \frac{1-p}{\lambda} \cdot (1-\rho) \cdot \frac{\partial \Omega_u}{\partial \epsilon_{\text{stw}}} + BE \\
& \forall\, \epsilon_{\text{stw}} \text{ in case 2} \quad \Rightarrow \quad \epsilon_{\text{stw}} \;=\; \xi_s^c
\end{aligned}
$$

(B)

$$
\begin{aligned}
& g(\epsilon_{\text{stw}}) \cdot \frac{p}{\lambda} \cdot n^s \cdot \Bigg[z_{\text{stw}}(\epsilon_{\text{stw}}) - z_{\text{stw}}(\epsilon_s^c) \\
& \quad + \frac{u(c_{\text{stw}}(\epsilon_{\text{stw}})) - u(c_{\text{stw}}(\epsilon_s^c))}{u'(c^w)} + \frac{\lambda}{f} \cdot b - \tau_{\text{stw}} \cdot \big(h - h_{\text{stw}}(\epsilon_s^c)\big) \Bigg] \\
& \quad < n^s \cdot \frac{1-p}{\lambda} \cdot (1-\rho) \cdot \frac{\partial \Omega_u}{\partial \epsilon_{\text{stw}}} + BE \\
& \forall\, \epsilon_{\text{stw}} \text{ in case 2} \quad \Rightarrow \quad \epsilon_{\text{stw}} \;=\; \max\{\xi_s^c, \xi_s^u\}
\end{aligned}
$$

(C)

$$
\begin{aligned}
& g(\epsilon_{\text{stw}}) \cdot \frac{p}{\lambda} \cdot n^s \cdot \Bigg[z_{\text{stw}}(\epsilon_{\text{stw}}) - z_{\text{stw}}(\epsilon_s^c) \\
& \quad + \frac{u(c_{\text{stw}}(\epsilon_{\text{stw}})) - u(c_{\text{stw}}(\epsilon_s^c))}{u'(c^w)} + \frac{\lambda}{f} \cdot b - \tau_{\text{stw}} \cdot \big(h - h_{\text{stw}}(\epsilon_s^c)\big) \Bigg] \\
& \quad = n^s \cdot \frac{1-p}{\lambda} \cdot (1-\rho) \cdot \frac{\partial \Omega_u}{\partial \epsilon_{\text{stw}}} + BE
\end{aligned}
$$

BE is defined as:

$$
\begin{aligned}
BE \;=\; & g(\epsilon_{\text{stw}}) \cdot \frac{\lambda_c}{u'(c^w)} \cdot \left(\frac{\partial WE}{\partial \epsilon_{\text{stw}}} \cdot \frac{1}{g(\epsilon_{\text{stw}})} \cdot \frac{\partial S_{\text{stw}}(\epsilon_{\text{stw}})}{\partial \epsilon_{\text{stw}}} + \eta \cdot \frac{b}{n} \right) \\
& \Bigg/ \left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)} \right)
\end{aligned}
$$

Case III:

$$\begin{aligned}\frac{\partial \mathscr{L}}{\partial \epsilon_{\text{stw}}} = {} & g(\epsilon_{\text{stw}}) \cdot \frac{p}{\lambda} \cdot n^s \cdot \Big(u(c_{\text{stw}}(\epsilon_{\text{stw}})) - u(c(\epsilon_{\text{stw}})) \\ & - u'(c^w) \cdot [c_{\text{stw}}(\epsilon_{\text{stw}}) - c(\epsilon_{\text{stw}})] \Big) \\ & - n \cdot \frac{\partial \Omega}{\partial \epsilon_{\text{stw}}} \\ & + \lambda_\theta \cdot \frac{p}{\lambda} \cdot \Big(u(c_{\text{stw}}(\epsilon_{\text{stw}})) - u(c(\epsilon_{\text{stw}})) \\ & - u'(c^w) \cdot [c_{\text{stw}}(\epsilon_{\text{stw}}) - c(\epsilon_{\text{stw}})] \Big) \\ & - \lambda_\theta \cdot \frac{1-\rho}{\lambda} \cdot \frac{\partial \Omega}{\partial \epsilon_{\text{stw}}} \\ & + \lambda_c \cdot \frac{\partial WE}{\partial \epsilon_{\text{stw}}} \overset{!}{\geq} 0\end{aligned}$$

$$\begin{aligned}\Leftrightarrow \quad & \left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right) \cdot \left(\frac{1-\rho}{\lambda} \cdot n^s \cdot \frac{\partial \Omega}{\partial \epsilon_{\text{stw}}} + BE\right) \\ & \overset{!}{\geq} \left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right) \cdot \frac{p}{\lambda} \cdot g(\epsilon_{\text{stw}}) \cdot n^s \\ & \cdot \left(\frac{\dfrac{u(c_{\text{stw}}(\epsilon_{\text{stw}})) - u(c(\epsilon_{\text{stw}}))}{c_{\text{stw}}(\epsilon_{\text{stw}}) - c(\epsilon_{\text{stw}})} - u'(c^w)}{u'(c^w)} \right) \\ & \cdot [c_{\text{stw}}(\epsilon_{\text{stw}}) - c(\epsilon_{\text{stw}})]\end{aligned}$$

$$\begin{aligned}\Leftrightarrow \quad & \left(\frac{1-\rho}{\lambda} \cdot n^s \cdot \frac{\partial \Omega}{\partial \epsilon_{\text{stw}}} + BE\right) \overset{!}{\geq} \cdot \frac{p}{\lambda} \cdot g(\epsilon_{\text{stw}}) \cdot n^s \\ & \cdot \left(\frac{\dfrac{u(c_{\text{stw}}(\epsilon_{\text{stw}})) - u(c(\epsilon_{\text{stw}}))}{c_{\text{stw}}(\epsilon_{\text{stw}}) - c(\epsilon_{\text{stw}})} - u'(c^w)}{u'(c^w)} \right) \\ & \cdot [c_{\text{stw}}(\epsilon_{\text{stw}}) - c(\epsilon_{\text{stw}})]\end{aligned}$$

If the equation holds with strict inequality, the cost of hours distortion exceeds the benefit of providing additional insurance to workers in constrained firms. Consequently, the Ramsey planner chooses not to allow firms and workers to enter short-time work (STW) when they could survive without reaching the threshold $\epsilon_{\text{stw}} = \xi_s^c$. Conversely, if the equation holds with exact equality, the STW threshold is determined by balancing the additional cost of hours distortions against the benefit of providing extra insurance.
This concludes the proof of Proposition 4.

3.F.6 Optimal Unemployment Insurance with STW

From the FOC of UI, we can derive the optimality condition of unemployment insurance:

$$
\begin{aligned}
(1-n)\cdot\left(u'(b)-u'(c^w)\right) \;=\;& \lambda_\theta(1-\rho)\cdot\left(\frac{1-n}{n}+\frac{u'(b)}{u'(c^w)}\right)\\
&+\left(\lambda_{\epsilon_s^u}+\lambda_{\epsilon_s^c}+\lambda_{\epsilon^g}\right)\cdot\left(-\frac{u'(b)}{u'(c^w)}\right)\\
&+\lambda_c(1-\rho)\cdot\left(\eta\cdot\frac{1-n}{n}+(1-\eta)\cdot\frac{u'(b)}{u'(c^w)}\right)
\end{aligned}
$$

To get better insight, we need to determine the Lagrange multipliers for λ_θ, λ_c. We already know the Lagrange multipliers for the separation conditions:

$$
\begin{aligned}
-\frac{\lambda_{\epsilon_s^u}}{n^s\cdot u'(c^w)} \;=\;& \frac{1-p}{\lambda}\cdot\frac{g(\epsilon_s^u)}{\frac{\partial S_{\text{stw}}^u(\epsilon_s^u)}{\partial\epsilon_s^u}}\\
&\cdot\left(\left(1+\frac{\lambda_\theta}{n^s\cdot u'(c^w)}\right)\cdot\left(\frac{\lambda}{f}\cdot b-\tau_{stw}\cdot\left(\bar{h}-h_{\text{stw}}(\epsilon_s^u)\right)\right)\right.\\
&\left.+\frac{\lambda_c}{n^s\cdot u'(c^w)}\cdot\left(\eta\cdot\frac{b}{n}+\frac{1}{g(\epsilon_s^u)}\cdot\frac{\lambda}{1-p}\cdot\frac{\partial WE}{\partial\epsilon_s^u}\right)\right)
\end{aligned}
$$

$$
\begin{aligned}
-\frac{\lambda_{\xi_s^c}}{n^s\cdot u'(c^w)} \;=\;& \frac{p}{\lambda}\cdot\frac{g(\xi_s^c)}{\frac{\partial S_{\text{stw}}^c(\xi_s^c)}{\partial\xi_s^c}}\cdot\left(\left(1+\frac{\lambda_\theta}{n^s\cdot u'(c^w)}\right)\cdot\frac{\lambda}{f}\cdot b\right.\\
&\left.+\frac{\lambda_c}{n^s\cdot u'(c^w)}\left(\eta\cdot\frac{b}{n}+\frac{1}{g(\xi_s^c)}\cdot\frac{\lambda}{p}\cdot\frac{\partial WE}{\partial\xi_s^c}\right)\right)\cdot\mathbb{1}(\epsilon_{\text{stw}}\;\leq\;\xi_s^c)
\end{aligned}
$$

$$
\begin{aligned}
-\frac{\lambda_{\epsilon_s^c}}{n^s\cdot u'(c^w)} \;=\;& \frac{p}{\lambda}\cdot\frac{g(\epsilon_s^c)}{\frac{\partial S_{\text{stw}}^c(\epsilon_s^c)}{\partial\epsilon_s^c}}\cdot\left(\left(1+\frac{\lambda_\theta}{n^s\cdot u'(c^w)}\right)\cdot\left(\frac{\lambda}{f}\cdot b-\tau_{stw}\cdot\left(\bar{h}-h_{\text{stw}}(\epsilon_s^c)\right)\right)\right.\\
&\left.+\frac{\lambda_c}{n^s\cdot u'(c^w)}\cdot\left(\eta\cdot\frac{b}{n}+\frac{1}{g(\epsilon_s^c)}\cdot\frac{\lambda}{p}\cdot\frac{\partial WE}{\partial\epsilon_s^c}\right)\right)\cdot\mathbb{1}(\epsilon_{\text{stw}}\;\geq\;\epsilon_s^c)
\end{aligned}
$$

First, let us find an expression for λ_c:

$$
\begin{aligned}
\frac{\partial \mathscr{L}}{\partial c^w} \;=\; & -\bigg((1-p)\cdot(1-\rho^u)\cdot\frac{u(c^w)-u(b)}{u'(c^w)}\cdot\frac{u''(c^w)}{u'(c^w)} \\
& \quad + p\cdot(1-\rho^c)\cdot\frac{u^c-u(b)}{u'(c^w)}\cdot\frac{u''(c^w)}{u'(c^w)}\bigg)\cdot\lambda_\theta \\
& -(1-\eta)\cdot(1-p)\cdot(1-\rho^u)\cdot\left(1-(1-\eta)\cdot\frac{u(c^w)-u(b)}{u'(c^w)}\cdot\frac{u''(c^w)}{u'(c^w)}\right) \\
& \quad -(1-p)\cdot\big(1-\cdot\rho^c\big)\cdot(1-\eta)\cdot\frac{u^c-u(b)}{u'(c^w)}\cdot\frac{u''(c^w)}{u'(c^w)}\cdot\lambda_c \\
& + \frac{u(c^w)-u(b)}{u'(c^w)}\cdot\frac{u''(c^w)}{u'(c^w)}\cdot\lambda_{\epsilon_s^u} \\
& + \frac{u(c(\xi_s^c))-u(b)}{u'(c^w)}\cdot\frac{u''(c^w)}{u'(c^w)}\cdot\lambda_{\xi_s^c} \\
& + \frac{u(c_{\text{stw}}(\epsilon_s^c))-u(b)}{u'(c^w)}\cdot\frac{u''(c^w)}{u'(c^w)}\cdot\lambda_{\epsilon_s^c} \;=\; 0
\end{aligned}
$$

Insert Lagrange multipliers for separation conditions

$$
\begin{aligned}
& -\left(p(1-\rho^u)\cdot\frac{u(c^w)-u(b)}{u'(c^w)}\cdot\frac{u''(c^w)}{u'(c^w)}+(1-p)(1-\rho^c)\cdot\frac{u^c-u(b)}{u'(c^w)}\cdot\frac{u''(c^w)}{u'(c^w)}\right)\lambda_\theta \\
& -(1-\eta)\cdot\left((1-p)\cdot(1-\rho^u)\cdot\left(1-(1-\eta)\cdot\frac{u(c^w)-u(b)}{u'(c^w)}\cdot\frac{u''(c^w)}{u'(c^w)}\right)\right. \\
& \left. -p\cdot(1-\rho^c)\cdot(1-\eta)\cdot\frac{u^c-u(b)}{u'(c^w)}\cdot\frac{u''(c^w)}{u'(c^w)}\right)\lambda_c \\
& + \frac{\partial\epsilon_s^u}{\partial c^w}\cdot\frac{1-p}{\lambda}\cdot g(\epsilon_s^u)\big(n^s\cdot u'(c^w)+\lambda_\theta\big)\cdot\left(\frac{\lambda}{f}\cdot b-\tau_{\text{stw}}\cdot(\bar{h}-h_{stw}(\epsilon_s^u))\right) \\
& + \lambda_c\cdot\left(\eta\cdot\frac{b}{n}+\frac{1}{g(\epsilon_s^u)}\cdot\frac{\lambda}{1-p}\cdot\frac{\partial WE}{\partial\epsilon_s^u}\right) \\
& + \frac{\partial\xi_s^c}{\partial c^w}\cdot\frac{p}{\lambda}\cdot g(\xi_s^c)\big(n^s\cdot u'(c^w)+\lambda_\theta\big)\cdot\left(\frac{\lambda}{f}\cdot b+\lambda_c\cdot\left(\eta\cdot\frac{b}{n}-\frac{1}{g(\xi_s^c)}\cdot\frac{\lambda}{p}\cdot\frac{\partial WE}{\partial\xi_s^c}\right)\right) \\
& \cdot\mathbb{1}(\epsilon_{stw}\;\le\;\xi_s^c) \\
& + \frac{\partial\xi_s^c}{\partial c^w}\cdot\frac{p}{\lambda}\cdot g(\xi_s^c)\big(n^s\cdot u'(c^w)+\lambda_\theta\big)\cdot\left(\frac{\lambda}{f}\cdot b-\tau_{\text{stw}}\cdot(\bar{h}-h_{stw}(\xi_s^c))\right) \\
& + \lambda_c\cdot\left(\eta\cdot\frac{b}{n}+\frac{1}{g(\xi_s^c)}\cdot\frac{\lambda}{p}\cdot\frac{\partial WE}{\partial\xi_s^c}\right)\cdot\mathbb{1}(\epsilon_{stw}\;\ge\;\epsilon_s^c)\;=\;0
\end{aligned}
$$

Rearranging yields

$$
\begin{aligned}
0 \ = & \lambda_\theta \cdot \frac{\partial S^{\text{total}}}{\partial c^w} + \lambda_c \cdot \frac{\partial WE^{\text{total}}}{\partial c^w} \\
& + \frac{\partial \epsilon_s^u}{\partial c^w} \cdot \frac{1-p}{\lambda} \cdot g(\epsilon_s^u) \cdot n^s \cdot u'(c^w) \cdot \left(\frac{\lambda}{f} \cdot b - \tau_{stw} \cdot (\bar{h} - h_{\text{stw}}(\epsilon_s^u)) \right) \\
& + \frac{\partial \xi_s^c}{\partial c^w} \cdot \frac{p}{\lambda} \cdot g(\xi_s^c) \cdot n^s \cdot u'(c^w) \cdot \frac{\lambda}{f} \cdot b \cdot \mathbb{1}(\epsilon_{\text{stw}} \ \leq \ \xi_s^c) \\
& + \frac{\partial \epsilon_s^c}{\partial c^w} \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot n^s \cdot u'(c^w) \cdot \left(\frac{\lambda}{f} \cdot b - \tau_{stw} \cdot (\bar{h} - h_{\text{stw}}(\epsilon_s^c)) \right) \cdot \mathbb{1}(\epsilon_{\text{stw}} \ \geq \ \epsilon_s^c)
\end{aligned}
$$

with

$$
\begin{aligned}
\frac{\partial S^{\text{total}}}{\partial c^w} \ = & \ (1-p) \cdot (1-\rho^u) \cdot \frac{u(c^w) - u(b)}{u'(c^w)} \cdot \frac{u''(c^w)}{u'(c^w)} \\
& - p \cdot (1-\rho^c) \cdot \frac{u^c - u(b)}{u'(c^w)} \cdot \frac{u''(c^w)}{u'(c^w)} \\
& + \frac{\partial \epsilon_s^u}{\partial c^w} \cdot \frac{1-p}{\lambda} \cdot g(\epsilon_s^u) \cdot \left(\frac{\lambda}{f} b - \tau_{\text{stw}} \cdot (\bar{h} - h_{\text{stw}}(\epsilon_s^u)) \right) \\
& + \frac{\partial \xi_s^c}{\partial c^w} \cdot \frac{p}{\lambda} \cdot g(\xi_s^c) \cdot \frac{\lambda}{f} \cdot b \cdot \mathbb{1}(\epsilon_{\text{stw}} \ \leq \ \xi_s^c) \\
& + \frac{\partial \epsilon_s^c}{\partial c^w} \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot \left(\frac{\lambda}{f} b - \tau_{\text{stw}} \cdot (\bar{h} - h_{\text{stw}}(\epsilon_s^c)) \right) \cdot \mathbb{1}(\epsilon_{\text{stw}} \ \geq \ \epsilon_s^c)
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{\partial WE^{\text{total}}}{\partial c^w} \ = & \ -(1-\eta) \cdot (1-p) \cdot (1-\rho^u) \cdot \left(1 - (1-\eta) \cdot \frac{u(c^w) - u(b)}{u'(c^w)} \cdot \frac{u''(c^w)}{u'(c^w)} \right) \\
& + p \cdot (1-\rho^c) \cdot (1-\eta) \cdot \frac{u^c - u(b)}{u'(c^w)} \cdot \frac{u''(c^w)}{u'(c^w)} \\
& + \frac{\partial \epsilon_s^u}{\partial c^w} \cdot \frac{1-p}{\lambda} \cdot g(\epsilon_s^u) \cdot \lambda_c \cdot \left(\eta \cdot \frac{b}{n} + \frac{1}{g(\epsilon_s^u)} \cdot \frac{\lambda}{1-p} \cdot \frac{\partial WE}{\partial \epsilon_s^u} \right) \\
& + \frac{\partial \xi_s^c}{\partial c^w} \cdot \frac{p}{\lambda} \cdot g(\xi_s^c) \cdot \lambda_c \cdot \left(\eta \cdot \frac{b}{n} - \frac{1}{g(\xi_s^c)} \cdot \frac{\lambda}{p} \cdot \frac{\partial WE}{\partial \epsilon_s^c} \right) \cdot \mathbb{1}(\epsilon_{\text{stw}} \ \leq \ \xi_s^c) \\
& + \frac{\partial \epsilon_s^c}{\partial c^w} \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot \lambda_c \cdot \left(\eta \cdot \frac{b}{n} - \frac{1}{g(\epsilon_s^c)} \cdot \frac{1}{p} \cdot \frac{\partial WE}{\partial \epsilon_s^c} \right) \cdot \mathbb{1}(\epsilon_{\text{stw}} \ \geq \ \epsilon_s^c)
\end{aligned}
$$

Collecting terms:

$$\lambda_c = \frac{-1}{\left(\frac{\partial WE}{\partial c^w}\right)^{\text{total}}}\Bigg[\frac{\partial \epsilon_s^u}{\partial c^w}\cdot\frac{1-p}{\lambda}\cdot g(\epsilon_s^u)\cdot n^s\cdot u'(c^w)\left(\frac{\lambda}{f}\cdot b-\tau_{\text{stw}}\cdot\left(\bar{h}-h_{\text{stw}}(\epsilon_s^u)\right)\right)$$
$$+\frac{\partial \xi_s^c}{\partial c^w}\cdot\frac{p}{\lambda}\cdot g(\xi_s^c)\cdot n^s\cdot u'(c^w)\left(\frac{\lambda}{f}\cdot b\right)\cdot\mathbb{1}(\epsilon_{\text{stw}}\leq\xi_s^c)$$
$$+\frac{\partial \epsilon_s^c}{\partial c^w}\cdot\frac{p}{\lambda}\cdot g(\epsilon_s^c)\cdot n^s\cdot u'(c^w)\left(\frac{\lambda}{f}\cdot b-\tau_{\text{stw}}\cdot\left(\bar{h}-h_{\text{stw}}(\epsilon_s^c)\right)\right)\cdot\mathbb{1}(\epsilon_{\text{stw}}\geq\epsilon_s^c)$$
$$+\left(-\frac{\partial S_{\text{total}}}{\partial c^w}\right)\cdot\lambda_\theta\Bigg]$$

Insert λ_c into the Lagrange multipliers for the separation conditions. To do that, let us rewrite the Lagrange multipliers of the separation conditions:

$$\lambda_{\epsilon_s^u} = -\frac{1}{\frac{\partial S_{\text{stw}}^u(\epsilon_s^u)}{\partial \epsilon_s^u}}\Bigg(n^s\cdot u'(c^w)\cdot\frac{p}{\lambda}\cdot g(\epsilon_s^u)\left(1+\frac{\lambda_\theta}{n^s\cdot u'(c^w)}\right)$$
$$\cdot\left(\frac{\lambda}{f}\cdot b-\tau_{\text{stw}}\cdot(\bar{h}-h_{\text{stw}}(\epsilon_s^u))\right)$$
$$+\lambda_c\left(\frac{\partial n}{\partial \epsilon_s^u}\cdot\left(-\frac{\partial WE}{\partial n}\right)+\frac{\partial WE}{\partial \epsilon_s^u}\right)\Bigg)$$

$$\lambda_{\xi_s^c} = \frac{\mathbb{1}(\epsilon_{\text{stw}}\leq\xi_s^c)}{\frac{\partial S_{\text{stw}}^c(\epsilon_s^c)}{\partial \epsilon_s^c}}\Bigg(n^s\cdot u'(c^w)\cdot\frac{p}{\lambda}\cdot g(\xi_s^c)\left(1+\frac{\lambda_\theta}{n^s\cdot u'(c^w)}\right)\cdot\frac{\lambda}{f}\cdot b$$
$$+\lambda_c\left(\frac{\partial n}{\partial \xi_s^c}\cdot\left(-\frac{\partial WE}{\partial n}\right)+\frac{\partial WE}{\partial \xi_s^c}\right)$$

$$\lambda_{\epsilon_s^c} = -\frac{\mathbb{1}(\epsilon_{\text{stw}}\geq\epsilon_s^c)}{\frac{\partial S_{\text{stw}}^c(\epsilon_s^c)}{\partial \epsilon_s^c}}\Bigg(n^s\cdot u'(c^w)\cdot\frac{p}{\lambda}\cdot g(\epsilon_s^c)\left(1+\frac{\lambda_\theta}{n^s\cdot u'(c^w)}\right)$$
$$\cdot\left(\frac{\lambda}{f}\cdot b-\tau_{\text{stw}}\cdot(\bar{h}-h_{\text{stw}}(\epsilon_s^c))\right)$$
$$+\lambda_c\left(\frac{\partial n}{\partial \epsilon_s^c}\cdot\left(-\frac{\partial WE}{\partial n}\right)+\frac{\partial WE}{\partial \epsilon_s^c}\right)\Bigg)$$

Inserting λ_c gives:

$$
\begin{aligned}
\lambda_{\epsilon_s^u} \; = & \\
& - \frac{1}{\frac{\partial S_{\text{stw}}^u(\epsilon_s^u)}{\partial \epsilon_s^u}} \Bigg[n^s u'(c^w) \cdot \frac{1-p}{\lambda} \cdot g(\epsilon_s^u) \left(1 + \frac{\lambda_\theta}{n^s u'(c^w)}\right) \left(\frac{\lambda}{f} b - \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon_s^u))\right) \\
& + \left(\frac{\partial n}{\partial \epsilon_s^u} \left(-\frac{\partial c^w}{\partial n}\right)^{\text{total}} + \left(\frac{\partial c^w}{\partial \epsilon_s^u}\right)^{\text{total}} \right) \\
& \cdot \left(\frac{\partial \epsilon_s^u}{\partial c^w} \cdot \frac{1-p}{\lambda} \cdot g(\epsilon_s^u) \cdot n^s u'(c^w) \left(\frac{\lambda}{f} b - \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon_s^u))\right) \right. \\
& + \frac{\partial \xi_s^c}{\partial c^w} \cdot \frac{p}{\lambda} \cdot g(\xi_s^c) \cdot n^s u'(c^w) \cdot \frac{\lambda}{f} b \cdot \mathbb{1}(\epsilon_{\text{stw}} \; \leq \; \xi_s^c) \\
& + \frac{\partial \epsilon_s^c}{\partial c^w} \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot n^s u'(c^w) \left(\frac{\lambda}{f} b - \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon_s^c))\right) \cdot \mathbb{1}(\epsilon_{\text{stw}} \; \geq \; \epsilon_s^c) \\
& \left. + \left(-\frac{\partial S^{\text{total}}}{\partial c^w}\right) \cdot \lambda_\theta \right) \Bigg]
\end{aligned}
$$

$$
\begin{aligned}
\lambda_{\xi_s^c} \; = & \\
& \frac{\mathbb{1}(\epsilon_{\text{stw}} \; \leq \; \xi_s^c)}{\frac{\partial S_{\text{stw}}^c(\xi_s^c)}{\partial \xi_s^c}} \Bigg[- n^s u'(c^w) \cdot \frac{p}{\lambda} \cdot g(\xi_s^c) \left(1 + \frac{\lambda_\theta}{n^s u'(c^w)}\right) \left(\frac{\lambda}{f} b - \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\xi_s^c))\right) \\
& + \left(\frac{\partial n}{\partial \xi_s^c} \left(-\frac{\partial c^w}{\partial n}\right)^{\text{total}} + \left(\frac{\partial c^w}{\partial \xi_s^c}\right)^{\text{total}} \right) \\
& \cdot \left(\frac{\partial \epsilon_s^u}{\partial c^w} \cdot \frac{1-p}{\lambda} \cdot g(\epsilon_s^u) \cdot n^s u'(c^w) \left(\frac{\lambda}{f} b - \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon_s^u))\right) \right. \\
& + \frac{\partial \xi_s^c}{\partial c^w} \cdot \frac{p}{\lambda} \cdot g(\xi_s^c) \cdot n^s u'(c^w) \cdot \frac{\lambda}{f} b \cdot \mathbb{1}(\epsilon_{\text{stw}} \; \leq \; \xi_s^c) \\
& + \frac{\partial \epsilon_s^c}{\partial c^w} \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot n^s u'(c^w) \left(\frac{\lambda}{f} b - \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon_s^c))\right) \cdot \mathbb{1}(\epsilon_{\text{stw}} \; \geq \; \epsilon_s^c) \\
& \left. + \left(-\frac{\partial S^{\text{total}}}{\partial c^w}\right) \cdot \lambda_\theta \right) \Bigg]
\end{aligned}
$$

$$
\begin{aligned}
&\lambda_{\epsilon_s^c} = \\
&-\frac{\mathbb{1}(\epsilon_{\text{stw}} \geq \epsilon_s^c)}{\frac{\partial S_{\text{stw}}^c(\epsilon_s^c)}{\partial \epsilon_s^c}}\Bigg[n^s u'(c^w)\cdot\frac{p}{\lambda}\cdot g(\epsilon_s^c)\left(1+\frac{\lambda_\theta}{n^s u'(c^w)}\right)\left(\frac{\lambda}{f}b-\tau_{\text{stw}}(\bar{h}-h_{\text{stw}}(\epsilon_s^c))\right)\\
&+\left(\frac{\partial n}{\partial \epsilon_s^c}\left(-\frac{\partial c^w}{\partial n}\right)^{\text{total}}+\left(\frac{\partial c^w}{\partial \epsilon_s^c}\right)^{\text{total}}\right)\\
&\cdot\Bigg(\frac{\partial \epsilon_s^u}{\partial c^w}\cdot\frac{1-p}{\lambda}\cdot g(\epsilon_s^u)\cdot n^s u'(c^w)\left(\frac{\lambda}{f}b-\tau_{\text{stw}}(\bar{h}-h_{\text{stw}}(\epsilon_s^u))\right)\\
&+\frac{\partial \xi_s^c}{\partial c^w}\cdot\frac{p}{\lambda}\cdot g(\xi_s^c)\cdot n^s u'(c^w)\cdot\frac{\lambda}{f}b\cdot\mathbb{1}(\epsilon_{\text{stw}} \leq \xi_s^c)\\
&+\frac{\partial \epsilon_s^c}{\partial c^w}\cdot\frac{p}{\lambda}\cdot g(\epsilon_s^c)\cdot n^s u'(c^w)\left(\frac{\lambda}{f}b-\tau_{\text{stw}}(\bar{h}-h_{\text{stw}}(\epsilon_s^c))\right)\cdot\mathbb{1}(\epsilon_{\text{stw}} \geq \epsilon_s^c)\\
&+\left(-\frac{\partial S^{\text{total}}}{\partial c^w}\right)\cdot\lambda_\theta\Bigg)\Bigg]
\end{aligned}
$$

Now we have everything to calculate λ_θ from two equations:

$$
\text{(1)}\quad \left(\chi-\frac{\eta-\gamma}{1-\eta}\right)\cdot\frac{1}{1-\gamma}\cdot\frac{k_v}{q} = \left(1+\frac{\lambda_\theta+\eta\cdot\lambda\cdot\lambda_c}{n^s\cdot u'(c^w)}\right)\cdot\frac{b}{f}
$$

$$
\begin{aligned}
\text{(2)}\quad \chi = {}& \frac{1}{1-\eta}\cdot\frac{1}{u\cdot f}\cdot\frac{1}{u'(c^w)}\cdot\left[\gamma-(\lambda\gamma-f\eta)\cdot\left(\frac{1}{\lambda}(1-\rho)\lambda_\theta-\lambda_{\epsilon_s^u}\right)\right]\\
&+\frac{1}{u}\cdot\frac{1}{u'(c^w)}\cdot\left(\lambda\cdot\frac{\gamma}{f}\cdot u'(c^w)+\left(\frac{1}{1-\eta}\cdot\frac{\lambda\gamma-f}{f}\right)\cdot\eta\right)\cdot\lambda_{\xi_s^c}\\
&+\frac{1}{u}\cdot\frac{1}{u'(c^w)}\cdot\left(\lambda\cdot\frac{\gamma}{f}\cdot u'(c^w)+\left(\frac{1}{1-\eta}\cdot\frac{\lambda\gamma-f}{f}\right)\cdot\eta\right)\cdot\lambda_{\epsilon_s^c}\\
&+\frac{1}{u}\cdot\frac{1}{u'(c^w)}\cdot\lambda_c\cdot\frac{\partial WE}{\partial\theta}\cdot\frac{1}{k_v}\\
&-\frac{1}{u}\cdot\frac{1}{u'(c^w)}\cdot IE_\theta
\end{aligned}
$$

We can rearrange (1) to:

$$
\chi\cdot k_v = (1-\gamma)\cdot q\left(\frac{\eta-\gamma}{(1-\gamma)(1-\eta)}\cdot\frac{k_v}{q}+\left(1+\frac{\lambda_\theta+\eta\cdot\lambda\cdot\lambda_c}{n^s\cdot u'(c^w)}\right)\cdot\frac{b}{f}\right)
$$

Note that

$$
(1-\gamma)\cdot q = -\frac{\partial f}{\partial\theta}
$$

with

$$\begin{aligned} f'(\theta) &= q(\theta) + \theta \cdot q'(\theta) \\ &= q(\theta) \cdot \left(1 + \frac{q'(\theta) \cdot \theta}{q(\theta)}\right) \\ &= q(\theta) \cdot (1 - \gamma) \end{aligned}$$

Inserting χ gives:

$$\begin{aligned} &[\gamma - (\lambda\gamma - f \cdot \eta)] \cdot \frac{1}{\lambda} \cdot (1 - \rho) \cdot \frac{1}{1-\eta} \cdot \frac{k_v}{f} \cdot \lambda_\theta \\ &= u'(c^w) \cdot \left(\frac{\partial f}{\partial \theta}\right) \cdot u \cdot \left(\frac{\eta - \gamma}{(1-\gamma)(1-\eta)} \cdot \frac{k_v}{q} + \left(1 + \frac{\lambda_\theta + \eta \cdot \lambda \cdot \lambda_c}{n^s \cdot u'(c^w)}\right) \cdot \frac{b}{f}\right) \\ &\quad + \frac{1}{f} \cdot (\lambda\gamma - f \cdot \eta) \cdot k_v \cdot (-\lambda_{\epsilon_s^u}) \\ &\quad + \left(\lambda \cdot \frac{\gamma}{f} \cdot u'(c^w(\epsilon_s^c)) + \left(\frac{1}{1-n} \cdot \frac{\lambda\gamma - f}{f}\right) \cdot \eta\right) \cdot k_v \cdot (-\lambda_{\xi_s^c}) \\ &\quad + \left(\lambda \cdot \frac{\gamma}{f} \cdot u'(c(\xi_s^c)) + \left(\frac{1}{1-n} \cdot \frac{\lambda\gamma - f}{f}\right) \cdot \eta\right) \cdot k_v \cdot (-\lambda_{\epsilon_s^c}) \\ &\quad + \frac{\partial WE}{\partial \theta} \cdot (-\lambda_c) \\ &\quad + IE_\theta \end{aligned}$$

Insert $\lambda_{\epsilon_s^u}, \lambda_{\epsilon_s^c}, \lambda_{\xi_s^c}, IE_\theta$:

$$
\begin{aligned}
&\left[\gamma - (\lambda\gamma - f \cdot \eta) \cdot \frac{1}{\lambda} \cdot (1-\rho)\right] \cdot \frac{1}{1-\eta} \cdot \frac{k_v}{f} \cdot \lambda_\theta \\
&= u'(c^w) \cdot \left(\frac{\partial f}{\partial \theta}\right) \cdot u \cdot \left(\frac{\eta - \gamma}{(1-\gamma)(1-\eta)} \cdot \frac{k_v}{q} + \left(1 + \frac{\lambda_\theta}{n^s u'(c^w)}\right) \cdot \frac{b}{f}\right) \\
&\quad + u'(c^w) \cdot n^s \cdot \left(-\frac{\partial \epsilon_s^u}{\partial \theta}\right)^{\text{total}} \cdot g(\epsilon_s^u) \\
&\quad \cdot \frac{1-p}{\lambda}\left(1 + \frac{\lambda_\theta}{n^s u'(c^w)}\right)\left(\frac{\lambda}{f} b - \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon_s^u))\right) \\
&\quad + u'(c^w) \cdot n^s \cdot \left(-\frac{\partial \xi_s^c}{\partial \theta}\right)^{\text{total}} \cdot g(\xi_s^c) \cdot \frac{p}{\lambda}\left(1 + \frac{\lambda_\theta}{n^s u'(c^w)}\right) \cdot \frac{\lambda}{f} b \cdot \mathbb{1}(\epsilon_{\text{stw}} \leq \xi_s^c) \\
&\quad + u'(c^w) \cdot n^s \cdot \left(-\frac{\partial \epsilon_s^c}{\partial \theta}\right)^{\text{total}} \cdot g(\epsilon_s^c) \\
&\quad \cdot \frac{p}{\lambda}\left(1 + \frac{\lambda_\theta}{n^s u'(c^w)}\right)\left(\frac{\lambda}{f} b - \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon_s^c))\right) \cdot \mathbb{1}(\epsilon_{\text{stw}} \geq \epsilon_s^c) \\
&\quad + \lambda_\theta \cdot \left(-\frac{\partial c^w}{\partial \theta}\right)^{\text{total}} \cdot \left(-\frac{\partial S}{\partial c^w}\right)^{\text{total}} \\
&\quad + n^s \cdot u'(c^w) \cdot \left(1 + \frac{\lambda_\theta}{n^s u'(c^w)}\right) \cdot \tilde{IE}_\theta
\end{aligned}
$$

Denote:

$$
\begin{aligned}
\left(-\frac{\partial \epsilon_s^u}{\partial \theta}\right)^{\text{total}} &= \left(-\frac{\partial \epsilon_s^u}{\partial \theta}\right) + \left(\frac{\partial c^w}{\partial \theta}\right)^{\text{total}} \cdot \left(-\frac{\partial \epsilon_s^u}{\partial c^w}\right) \\
&\quad + \left(-\frac{\partial \epsilon_s^u}{\partial \theta}\right) \cdot \left[\left(\frac{\partial n}{\partial \epsilon_s^u}\right) \cdot \left(-\frac{\partial c^w}{\partial n}\right)^{\text{total}} + \left(\frac{\partial c^w}{\partial \epsilon_s^u}\right)^{\text{total}}\right] \cdot \left(\frac{\partial \epsilon_s^u}{\partial c^w}\right) \\
&\quad + f'(\theta) \cdot \frac{\partial n}{\partial f} \cdot \frac{\partial c^w}{\partial n} \cdot \frac{\partial \epsilon_s^u}{\partial c^w}
\end{aligned}
$$

$$
\begin{aligned}
\left(-\frac{\partial \xi_s^c}{\partial \theta}\right)^{\text{total}} &= \left[\left(-\frac{\partial \xi_s^c}{\partial \theta}\right) + \left(\frac{\partial c^w}{\partial \theta}\right)^{\text{total}} \cdot \left(-\frac{\partial \xi_s^c}{\partial c^w}\right)\right. \\
&\quad + \left(-\frac{\partial \xi_s^c}{\partial \theta}\right) \cdot \left[\left(\frac{\partial n}{\partial \xi_s^c}\right) \cdot \left(-\frac{\partial c^w}{\partial n}\right)^{\text{total}} + \left(\frac{\partial c^w}{\partial \xi_s^c}\right)^{\text{total}}\right] \cdot \left(\frac{\partial \xi_s^c}{\partial c^w}\right) \\
&\quad \left. + f'(\theta) \cdot \frac{\partial n}{\partial f} \cdot \frac{\partial c^w}{\partial n} \cdot \frac{\partial \xi_s^c}{\partial c^w}\right] \cdot \mathbb{1}(\epsilon_{\text{stw}} \leq \xi_s^c)
\end{aligned}
$$

$$\begin{aligned}
\left(-\frac{\partial \epsilon_s^c}{\partial \theta}\right)^{\text{total}} = & \Bigg[\left(-\frac{\partial \epsilon_s^c}{\partial \theta}\right) + \left(\frac{\partial c^w}{\partial \theta}\right)^{\text{total}} \cdot \left(-\frac{\partial \epsilon_s^c}{\partial c^w}\right) \\
& + \left(-\frac{\partial \epsilon_s^c}{\partial \theta}\right) \cdot \left[\left(\frac{\partial n}{\partial \epsilon_s^c}\right) \cdot \left(-\frac{\partial c^w}{\partial n}\right)^{\text{total}} + \left(\frac{\partial c^w}{\partial \epsilon_s^c}\right)^{\text{total}}\right] \cdot \frac{\partial \epsilon_s^c}{\partial c^w} \\
& + f'(\theta) \cdot \frac{\partial n}{\partial f} \cdot \frac{\partial c^w}{\partial n} \cdot \frac{\partial \epsilon_s^c}{\partial c^w}\Bigg] \cdot \mathbb{1}(\epsilon_{\text{stw}} \geq \epsilon_s^c)
\end{aligned}$$

$$\begin{aligned}
\left(-\frac{\partial c^w}{\partial \theta}\right)^{\text{total,2}} = & \left(-\frac{\partial c^w}{\partial \theta}\right)^{\text{total}} \\
& + \left(-\frac{\partial \epsilon_s^u}{\partial \theta}\right) \cdot \left[\left(\frac{\partial n}{\partial \epsilon_s^u}\right) \cdot \left(-\frac{\partial c^w}{\partial n}\right)^{\text{total}} + \left(\frac{\partial c^w}{\partial \epsilon_s^u}\right)^{\text{total}}\right] \\
& + \left(-\frac{\partial \xi_s^c}{\partial \theta}\right) \cdot \left[\left(\frac{\partial n}{\partial \xi_s^c}\right) \cdot \left(-\frac{\partial c^w}{\partial n}\right)^{\text{total}} + \left(\frac{\partial c^w}{\partial \xi_s^c}\right)^{\text{total}}\right] \\
& + \left(-\frac{\partial \epsilon_s^c}{\partial \theta}\right) \cdot \left[\left(\frac{\partial n}{\partial \epsilon_s^c}\right) \cdot \left(-\frac{\partial c^w}{\partial n}\right)^{\text{total}} + \left(\frac{\partial c^w}{\partial \epsilon_s^c}\right)^{\text{total}}\right] \\
& + f'(\theta) \cdot \frac{\partial n}{\partial f} \cdot \frac{\partial c^w}{\partial n}
\end{aligned}$$

This allows us to calculate the Lagrange multiplier for the job-creation condition:

$$\begin{aligned}
& M \cdot \lambda_\theta = \\
& u'(c^w) \cdot \frac{\partial f}{\partial \theta} \cdot u \cdot \left(\frac{\eta - \gamma}{(1-\gamma)(1-\eta)} \cdot \frac{k_v}{q} + \left(1 + \frac{\lambda_\theta}{n^s u'(c^w)}\right) \cdot \frac{b}{f}\right) \\
& \quad + u'(c^w) \cdot n^s \cdot \left(\frac{\partial \epsilon_s^u}{\partial \theta}\right)^{\text{total}} \cdot g(\epsilon_s^u) \cdot \frac{1-p}{\lambda} \left(\frac{\lambda}{f} b - \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon_s^u))\right) \\
& \quad + u'(c^w) \cdot n^s \cdot \left(-\frac{\partial \xi_s^c}{\partial \theta}\right)^{\text{total}} \cdot g(\xi_s^c) \cdot \frac{p}{\lambda} \cdot \frac{\lambda}{f} b \cdot \mathbb{1}(\epsilon_{\text{stw}} \leq \xi_s^c) \\
& \quad + u'(c^w) \cdot n^s \cdot \left(-\frac{\partial \epsilon_s^c}{\partial \theta}\right)^{\text{total}} \cdot g(\epsilon_s^c) \\
& \quad \cdot \frac{p}{\lambda} \left(\frac{\lambda}{f} b - \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon_s^c))\right) \cdot \mathbb{1}(\epsilon_{\text{stw}} \geq \epsilon_s^c) \\
& \quad + u'(c^w) \cdot n^s \cdot \tilde{IE}_\theta
\end{aligned}$$

with

$$
\begin{aligned}
M \ = \ & \left[\gamma - (\lambda\gamma - f \cdot \eta) \cdot \frac{1}{\lambda} \cdot (1-\rho)\right] \cdot \frac{1}{1-\eta} \cdot \frac{k_v}{f} \\
& + \left(-\frac{\partial f}{\partial \theta} \cdot u \cdot \frac{b}{f}\right) \\
& + \left(\frac{\partial \epsilon_s^u}{\partial \theta}\right)^{\text{total}} \cdot g(\epsilon_s^u) \cdot \frac{1-p}{\lambda}\left(\frac{\lambda}{f}b - \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon_s^u))\right) \\
& + \left(\frac{\partial \xi_s^c}{\partial \theta}\right)^{\text{total}} \cdot g(\xi_s^c) \cdot \frac{p}{\lambda} \cdot \frac{\lambda}{f} b \cdot \mathbb{1}(\epsilon_{\text{stw}} \ \leq \ \xi_s^c) \\
& + \left(\frac{\partial \epsilon_s^c}{\partial \theta}\right)^{\text{total}} \cdot g(\epsilon_s^c) \cdot \frac{p}{\lambda}\left(\frac{\lambda}{f}b - \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon_s^c))\right) \cdot \mathbb{1}(\epsilon_{\text{stw}} \ \geq \ \epsilon_s^c) \\
& + \left(\frac{\partial c^w}{\partial \theta}\right)^{\text{total}} \cdot \left(-\frac{\partial S}{\partial c^w}\right)^{\text{total}} \\
& + \tilde{IE}_\theta
\end{aligned}
$$

Note. $\frac{1}{M}$ denotes the general equilibrium effect of an increase of the joint surplus on θ:

$$
\left(\frac{\partial \theta}{\partial S}\right)^{\text{ge}} \ = \ \frac{1}{M}
$$

Rearranging for λ_θ gives:

$$
\begin{aligned}
& M \cdot \lambda_\theta \ = \\
& u'(c^w) \cdot \left(\frac{\partial f}{\partial S}\right)^{\text{ge}} \cdot u \cdot \left(\frac{\eta - \gamma}{(1-\gamma)(1-\eta)} \cdot \frac{k_v}{q} + \frac{\lambda}{f} \cdot b\right) \\
& \quad + u'(c^w) \cdot n^s \cdot \left(\left(-\frac{\partial \epsilon_s^u}{\partial S}\right)^{\text{ge}} \cdot g(\epsilon_s^u) \cdot \frac{1-p}{\lambda}\left(\frac{\lambda}{f} \cdot b - \tau_{\text{stw}} \cdot (\bar{h} - h_{\text{stw}}(\epsilon_s^u))\right)\right) \\
& \quad + u'(c^w) \cdot n^s \cdot \left(\left(-\frac{\partial \xi_s^c}{\partial S}\right)^{\text{ge}} \cdot g(\xi_s^c) \cdot \frac{p}{\lambda} \cdot \frac{\lambda}{f} \cdot b \cdot \mathbb{1}(\epsilon_{\text{stw}} \ \leq \ \xi_s^c)\right) \\
& \quad + u'(c^w) \cdot n^s \\
& \quad \cdot \left(\left(-\frac{\partial \epsilon_s^c}{\partial S}\right)^{\text{ge}} \cdot g(\epsilon_s^c) \cdot \frac{p}{\lambda} \cdot \left(\frac{\lambda}{f} \cdot b - \tau_{\text{stw}} \cdot (\bar{h} - h_{\text{stw}}(\epsilon_s^c))\right) \cdot \mathbb{1}(\epsilon_{\text{stw}} \ \geq \ \epsilon_s^c)\right) \\
& \quad + u'(c^w) \cdot n^s \cdot \left(\frac{\partial \theta}{\partial S}\right)^{\text{ge}} \cdot \tilde{IE}_\theta
\end{aligned}
$$

Define:

$$
\begin{aligned}
\hat{IE}_\theta \ = \ & \frac{1}{1-\rho^c} \cdot \left[\int_{\max\{\epsilon_{\text{stw}}, \xi_s^c\}}^{\epsilon^p} \lambda \cdot \frac{\gamma}{f} \cdot \frac{u'(c^w(\epsilon)) - u'(c^w)}{u'(c^w)} \, dG(\epsilon) \right. \\
& \left. - \int_{\epsilon_s^c}^{\max\{\epsilon_{\text{stw}}, \epsilon_s^c\}} \lambda \cdot \frac{\gamma}{f} \cdot \frac{u'(c_{\text{stw}}(\epsilon)) - u'(c^w)}{u'(c^w)} \, dG(\epsilon)\right] \cdot k_v
\end{aligned}
$$

Note:

$$\left(\frac{\partial f}{\partial S}\right)^{\text{ge}} = \frac{1}{M}\cdot\frac{\partial f}{\partial S}$$
$$\left(\frac{\partial \epsilon_s^u}{\partial S}\right)^{\text{ge}} = \frac{1}{M}\cdot\left(\frac{\partial \epsilon_s^u}{\partial S}\right)^{\text{total}}$$
$$\left(\frac{\partial \xi_s^c}{\partial S}\right)^{\text{ge}} = \frac{1}{M}\cdot\left(\frac{\partial \xi_s^c}{\partial S}\right)^{\text{total}}$$
$$\left(\frac{\partial \epsilon_s^c}{\partial S}\right)^{\text{ge}} = \frac{1}{M}\cdot\left(\frac{\partial \epsilon_s^c}{\partial S}\right)^{\text{total}}$$

$$\begin{aligned}
\lambda_c = & -\frac{1}{\left(\frac{\partial WE}{\partial c^w}\right)^{\text{total}}}\Bigg[\left(\frac{\partial \epsilon_s^u}{\partial c^w}+\left(\frac{\partial S}{\partial c^w}\right)^{\text{total}}\left(\frac{\partial \epsilon_s^u}{\partial S}\right)^{\text{ge}}\right)\\
& \cdot\frac{1-p}{\lambda}\cdot g(\epsilon_s^u)\cdot n^s u'(c^w)\left(\frac{\lambda}{f}b-\tau_{\text{stw}}(\bar{h}-h_{\text{stw}}(\epsilon_s^u))\right)\\
& +\left(\frac{\partial \xi_s^c}{\partial c^w}+\left(\frac{\partial S}{\partial c^w}\right)^{\text{total}}\left(\frac{\partial \xi_s^c}{\partial S}\right)^{\text{ge}}\right)\\
& \cdot\frac{p}{\lambda}\cdot g(\xi_s^c)\cdot n^s u'(c^w)\cdot\frac{\lambda}{f}b\cdot\mathbb{1}(\epsilon_{\text{stw}}\ \leq\ \xi_s^c)\\
& +\left(\frac{\partial \epsilon_s^c}{\partial c^w}+\left(\frac{\partial S}{\partial c^w}\right)^{\text{total}}\left(\frac{\partial \epsilon_s^c}{\partial S}\right)^{\text{ge}}\right)\cdot\frac{p}{\lambda}\cdot g(\epsilon_s^c)\\
& \cdot n^s\cdot u'(c^w)\left(\frac{\lambda}{f}b-\tau_{\text{stw}}(\bar{h}-h_{\text{stw}}(\epsilon_s^c))\right)\cdot\mathbb{1}(\epsilon_{\text{stw}}\ \geq\ \epsilon_s^c)\\
& +\left(\frac{\partial S}{\partial c^w}\right)^{\text{total}}\left(\frac{\partial f}{\partial S}\right)^{\text{ge}}\cdot u\cdot\left(\frac{\eta-\gamma}{(1-\gamma)(1-\eta)}\cdot\frac{k_v}{q}+\left(1+\frac{\lambda_\theta}{n^s u'(c^w)}\right)\cdot\frac{b}{f}\right)\\
& -n^c\cdot\left(\frac{\partial S}{\partial c^w}\right)^{\text{total}}\left(\frac{\partial \theta}{\partial S}\right)^{\text{ge}}\cdot\hat{IE}_\theta\Bigg]
\end{aligned}$$

Following the same arguments, we can express the Lagrange multiplier for the separation conditions as:

$$
\begin{aligned}
\lambda_{\epsilon_s^u} = & -\frac{u'(c^w)}{\left(\frac{\partial S^u_{\text{stw}}(\epsilon_s^u)}{\partial \epsilon_s^u}\right)^{\text{total}}}\Bigg[\left(\frac{\partial n}{\partial \epsilon_s^u}\right)^{\text{ge}} \cdot \left(\frac{\lambda}{f}b - \tau_{\text{stw}}\left(\bar{h} - h_{\text{stw}}(\epsilon_s^u)\right)\right) \\
& + \left(\frac{\partial \xi_s^c}{\partial \epsilon_s^u}\right)^{\text{ge}} \cdot \frac{\partial n}{\partial \xi_s^c} \cdot \frac{\lambda}{f}b \cdot \mathbb{1}(\epsilon_{\text{stw}} \leq \xi_s^c) \\
& + \left(\frac{\partial \epsilon_s^c}{\partial \epsilon_s^u}\right)^{\text{ge}} \cdot \frac{\partial n}{\partial \epsilon_s^c}\left(\frac{\lambda}{f}b - \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon_s^c))\right) \cdot \mathbb{1}(\epsilon_{\text{stw}} \geq \epsilon_s^c) \\
& + \frac{\partial c^w}{\partial \epsilon_s^u}\left(\frac{\partial f}{\partial c^w}\right)^{\text{ge}} \cdot u \cdot \left(\frac{\eta - \gamma}{(1-\gamma)(1-\eta)} \cdot \frac{k_v}{q} + \frac{b}{f}\right) \\
& + n^c \cdot \frac{\partial c^w}{\partial \epsilon_s^u}\left(\frac{\partial \theta}{\partial c^w}\right)^{\text{ge}} \cdot \hat{IE}_\theta \Bigg]
\end{aligned}
$$

$$
\begin{aligned}
\lambda_{\xi_s^c} = & -\frac{u'(c^w)}{\left(\frac{\partial S^c_{\text{stw}}(\xi_s^c)}{\partial \xi_s^c}\right)^{\text{total}}}\Bigg[\left(\frac{\partial \xi_s^u}{\partial \xi_s^c}\right)^{\text{ge}} \cdot \frac{\partial n}{\partial \xi_s^u}\left(\frac{\lambda}{f}b - \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\xi_s^u))\right) \\
& + \left(\frac{\partial n}{\partial \xi_s^c}\right)^{\text{ge}} \cdot \frac{\lambda}{f}b \cdot \mathbb{1}(\epsilon_{\text{stw}} \leq \xi_s^c) \\
& + \left(\frac{\partial \epsilon_s^c}{\partial \xi_s^c}\right)^{\text{ge}} \cdot \frac{\partial n}{\partial \epsilon_s^c}\left(\frac{\lambda}{f}b - \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon_s^c))\right) \cdot \mathbb{1}(\epsilon_{\text{stw}} \geq \epsilon_s^c) \\
& + \frac{\partial c^w}{\partial \xi_s^c}\left(\frac{\partial f}{\partial c^w}\right)^{\text{ge}} \cdot u \cdot \left(\frac{\eta - \gamma}{(1-\gamma)(1-\eta)} \cdot \frac{k_v}{q} + \frac{b}{f}\right) \\
& + n^c \cdot \frac{\partial c^w}{\partial \xi_s^c}\left(\frac{\partial \theta}{\partial c^w}\right)^{\text{ge}} \cdot \hat{IE}_\theta \Bigg]
\end{aligned}
$$

$$
\begin{aligned}
\lambda_{\epsilon_s^c} = & -\frac{u'(c^w)}{\left(\frac{\partial S^c_{\text{stw}}(\epsilon_s^c)}{\partial \epsilon_s^c}\right)^{\text{total}}}\Bigg[\left(\frac{\partial \epsilon_s^u}{\partial \epsilon_s^c}\right)^{\text{ge}} \cdot \frac{\partial n}{\partial \epsilon_s^u}\left(\frac{\lambda}{f}b - \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\xi_s^u))\right) \\
& + \left(\frac{\partial \xi_s^c}{\partial \epsilon_s^c}\right)^{\text{ge}} \cdot \frac{\partial n}{\partial \epsilon_s^c} \cdot \frac{\lambda}{f}b \cdot \mathbb{1}(\epsilon_{\text{stw}} \leq \epsilon_s^c) \\
& + \frac{\partial n}{\partial \epsilon_s^c}\left(\frac{\lambda}{f}b - \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon_s^c))\right) \cdot \mathbb{1}(\epsilon_{\text{stw}} \geq \epsilon_s^c) \\
& + \frac{\partial c^w}{\partial \epsilon_s^c}\left(\frac{\partial f}{\partial c^w}\right)^{\text{ge}} \cdot u \cdot \left(\frac{\eta - \gamma}{(1-\gamma)(1-\eta)} \cdot \frac{k_v}{q} + \frac{b}{f}\right) \\
& + n^c \cdot \frac{\partial c^w}{\partial \epsilon_s^c}\left(\frac{\partial \theta}{\partial c^w}\right)^{\text{ge}} \cdot \hat{IE}_\theta \Bigg]
\end{aligned}
$$

For convenience, the optimal UI benefits are reproduced here:

$$
\begin{aligned}
(1-\eta)\cdot(u'(b)-u'(c^w)) \;=\; & \lambda_\theta(1-\rho)\cdot\left(\frac{1-n}{n}+\frac{u'(b)}{u'(c^w)}\right)\\
& +(\lambda_{\epsilon_s^u}+\lambda_{\xi_s^c}+\lambda_{\epsilon_s^c})\cdot\left(-\frac{u'(b)}{u'(c^w)}\right)\\
& +\lambda_c\cdot(1-\rho)\cdot\left(\eta\cdot\frac{1-n}{n}+(1-\eta)\cdot\frac{u'(b)}{u'(c^w)}\right)
\end{aligned}
$$

Inserting the Lagrange multiplier gives:

$$
\begin{aligned}
&(1-n)(u'(b)-u'(c^w))\\
&= u'(c^w)\Bigg[\left(\frac{\partial c^w}{\partial b}\right)^{\text{ge}}\left(-\frac{\partial f}{\partial c^w}\right)\cdot u\cdot\left(\frac{\eta-\gamma}{(1-\gamma)(1-\eta)}\cdot\frac{k_v}{q}+\frac{b}{f}\right)\\
&\quad+\left(\frac{\partial \epsilon_s^u}{\partial b}\right)^{\text{ge}}\cdot\frac{\partial n}{\partial \epsilon_s^u}\cdot\left(\frac{\lambda}{f}\cdot b-\tau_{\text{stw}}\cdot(\bar{h}-h_{\text{stw}}(\epsilon_s^u))\right)\\
&\quad+\left(\frac{\partial \xi_s^c}{\partial b}\right)^{\text{ge}}\cdot\frac{\partial n}{\partial \xi_s^c}\cdot\frac{\lambda}{f}\cdot b\cdot\mathbb{1}(\epsilon_{\text{stw}}\;\leq\;\xi_s^c)\\
&\quad+\left(\frac{\partial \epsilon_s^c}{\partial b}\right)^{\text{ge}}\cdot\frac{\partial n}{\partial \epsilon_s^c}\cdot\left(\frac{\lambda}{f}\cdot b-\tau_{\text{stw}}\cdot(\bar{h}-h_{\text{stw}}(\epsilon_s^c))\right)\cdot\mathbb{1}(\epsilon_{\text{stw}}\;\geq\;\epsilon_s^c)\\
&\quad+n^c\left(-\frac{\partial\theta}{\partial b}\right)^{\text{ge}}\cdot\hat{IE}_\theta\Bigg]
\end{aligned}
$$

with

$$
\begin{aligned}
\left(\frac{\partial \epsilon_s^u}{\partial b}\right)^{\text{ge}}\cdot\frac{\partial n}{\partial \epsilon_s^u} \;=\; & \underbrace{\left(\frac{\partial \epsilon_s^u}{\partial b}\right)^{\text{ge}}\left(\frac{\partial n}{\partial \epsilon_s^u}\right)}_{\text{direct effects}}\\
& \underbrace{+\left[\frac{\partial S}{\partial b}\left(\frac{\partial \epsilon_s^u}{\partial S}\right)^{\text{ge}}+\frac{\partial \xi_s^c}{\partial b}\left(\frac{\partial c^w}{\partial \xi_s^c}\right)^{\text{ge}}+\frac{\partial \epsilon_s^c}{\partial b}\left(\frac{\partial c^w}{\partial \epsilon_s^c}\right)^{\text{ge}}+\frac{\partial c^w}{\partial b}\left(\frac{\partial \epsilon_s^u}{\partial c^w}\right)^{\text{ge}}\right]\cdot\frac{\partial n}{\partial \epsilon_s^u}}_{\text{indirect effect}}
\end{aligned}
$$

$$\left(\frac{\partial \xi_s^c}{\partial b}\right)^{\text{ge}} \cdot \frac{\partial n}{\partial \xi_s^c} = \underbrace{\left(\frac{\partial \xi_s^c}{\partial b}\right)^{\text{ge}} \left(\frac{\partial n}{\partial \xi_s^c}\right)}_{\text{direct effects}}$$
$$+ \underbrace{\left[\frac{\partial S}{\partial b}\left(\frac{\partial \xi_s^c}{\partial S}\right)^{\text{ge}} + \frac{\partial \epsilon_s^u}{\partial b}\left(\frac{\partial c^w}{\partial \epsilon_s^u}\right)^{\text{ge}} + \frac{\partial \epsilon_s^c}{\partial b}\left(\frac{\partial c^w}{\partial \epsilon_s^c}\right)^{\text{ge}} + \frac{\partial c^w}{\partial b}\left(\frac{\partial \xi_s^c}{\partial c^w}\right)\right] \cdot \frac{\partial n}{\partial \xi_s^c}}_{\text{indirect effect}}$$

$$\left(\frac{\partial \epsilon_s^c}{\partial b}\right)^{\text{ge}} \cdot \frac{\partial n}{\partial \epsilon_s^c} = \underbrace{\left(\frac{\partial \epsilon_s^c}{\partial b}\right)^{\text{ge}} \left(\frac{\partial n}{\partial \epsilon_s^c}\right)}_{\text{direct effects}}$$
$$+ \underbrace{\left[\frac{\partial S}{\partial b}\left(\frac{\partial \epsilon_s^c}{\partial S}\right)^{\text{ge}} + \frac{\partial \epsilon_s^u}{\partial b}\left(\frac{\partial c^w}{\partial \epsilon_s^u}\right)^{\text{ge}} + \frac{\partial \xi_s^c}{\partial b}\left(\frac{\partial c^w}{\partial \xi_s^c}\right)^{\text{ge}} + \frac{\partial c^w}{\partial b}\left(\frac{\partial \epsilon_s^c}{\partial c^w}\right)\right] \cdot \frac{\partial n}{\partial \epsilon_s^c}}_{\text{indirect effect}}$$

$$\left(-\frac{\partial f}{\partial b}\right)^{\text{ge}} = \underbrace{\frac{\partial S}{\partial b}\left(-\frac{\partial f}{\partial S}\right)^{\text{ge}}}_{\text{direct effect}}$$
$$+ \underbrace{\left[\frac{\partial \epsilon_s^u}{\partial b} \cdot \frac{\partial c^w}{\partial \epsilon_s^u} + \frac{\partial \xi_s^c}{\partial b} \cdot \frac{\partial c^w}{\partial \xi_s^c} + \frac{\partial \epsilon_s^c}{\partial b} \cdot \frac{\partial c^w}{\partial \epsilon_s^c} + \frac{\partial c^w}{\partial b}\right] \cdot \left(-\frac{\partial f}{\partial c^w}\right)^{\text{ge}}}_{\text{indirect effect}}$$

$$\left(-\frac{\partial \theta}{\partial b}\right)^{\text{ge}} = \underbrace{\frac{\partial S}{\partial b}\left(-\frac{\partial \theta}{\partial S}\right)^{\text{ge}}}_{\text{direct effect}}$$
$$+ \underbrace{\left[\frac{\partial \epsilon_s^u}{\partial b} \cdot \frac{\partial c^w}{\partial \epsilon_s^u} + \frac{\partial \xi_s^c}{\partial b} \cdot \frac{\partial c^w}{\partial \xi_s^c} + \frac{\partial \epsilon_s^c}{\partial b} \cdot \frac{\partial c^w}{\partial \epsilon_s^c} + \frac{\partial c^w}{\partial b}\right] \cdot \left(-\frac{\partial \theta}{\partial c^w}\right)^{\text{ge}}}_{\text{indirect effect}}$$

Appendix 3.G Proof of Lemma 1

Separation condition: unconstrained matches

$$z_{\text{stw}}(\epsilon_s^u) + \frac{u(c^w) - u(b)}{u'(c^w)} - c^w + \tau_{\text{stw}}(\bar{h} - h(\epsilon_s^u)) + \frac{\lambda - \eta \cdot f}{1 - \eta} \cdot \frac{k_v}{q} = 0$$

Separation condition: constrained matches

$$\frac{u(c_{\text{stw}}(\epsilon_s^c)) - u(b)}{u'(c^w)} + \frac{\eta}{1 - \eta} \cdot (\lambda - f) \cdot \frac{k_v}{q}$$

Remember:

$$c_{\text{stw}}(\epsilon) = z_{\text{stw}}(\epsilon_s^c) + \tau_{\text{stw}}(\bar{h} - h(\epsilon_s^c)) + \lambda \cdot \frac{k_v}{q}$$

This allows us to rewrite the separation condition of constrained firms on *STW* as:

$$z_{\text{stw}}(\epsilon_s^c) + \frac{u(c_{\text{stw}}(\epsilon_s^c)) - u(b)}{u'(c^w)} - c_{\text{stw}}(\epsilon_s^c) + \tau_{\text{stw}}(\bar{h} - h(\epsilon_s^c)) + \frac{\lambda - \eta \cdot f}{1-\eta} \cdot \frac{k_v}{q} = 0$$

T.b.s.: $\epsilon_s^c > \epsilon_s^u$ **(the proof for $\xi_s^c > \xi_s^u$ is completely analogous)**
Note: $c^w > c_{\text{stw}}(\epsilon)$

$$\begin{aligned}\epsilon_s^c > \epsilon_s^u \iff\ & z_{\text{stw}}(\epsilon) + \frac{u(c^w) - u(b)}{u'(c^w)} - c^w + \tau_{\text{stw}}(\bar{h} - h(\epsilon)) + \frac{\lambda - \eta f}{1-\eta} \cdot \frac{k_v}{q} \\ & \geq z_{\text{stw}}(\epsilon) + \frac{u(c_{\text{stw}}(\epsilon)) - u(b)}{u'(c^w)} - c_{\text{stw}}(\epsilon) \\ & + \tau_{\text{stw}}(\bar{h} - h(\epsilon)) + \frac{\lambda - \eta f}{1-\eta} \cdot \frac{k_v}{q}\end{aligned}$$

$$\iff \frac{u(c^w)}{u'(c^w)} - c^w > \frac{u(c_{\text{stw}}(\epsilon))}{u'(c^w)} - c_{\text{stw}}(\epsilon)$$

$$\iff \frac{u(c^w) - u(c_{\text{stw}}(\epsilon))}{u'(c^w)} > c^w - c_{\text{stw}}(\epsilon)$$

$$\iff u(c^w) - u(c_{\text{stw}}(\epsilon)) > (c^w - c_{\text{stw}}(\epsilon)) \cdot u'(c^w)$$

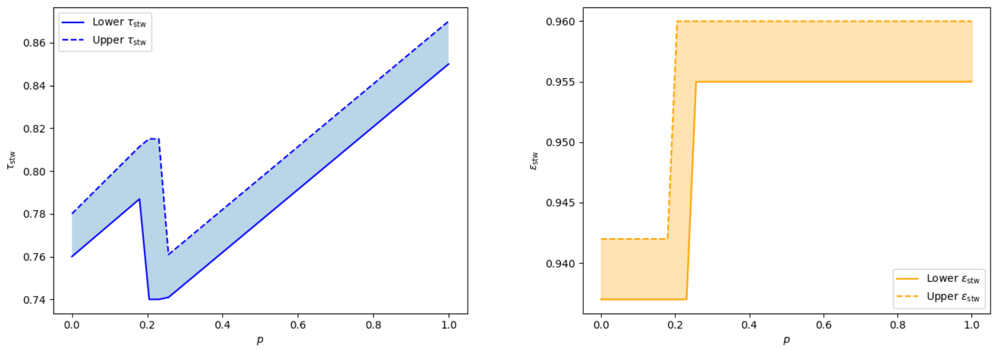
$$\iff \int_{c_{\text{stw}}(\epsilon)}^{c^w} u'(c)\, dc > \int_{c_{\text{stw}}(\epsilon)}^{c^w} u'(c^w)\, dc$$

$$\iff \int_{c_{\text{stw}}(\epsilon)}^{c^w} \left[u'(c) - u'(c^w)\right] dc > 0 \qquad \checkmark \quad \text{(due to risk aversion)}$$

Appendix 3.H Optimization

We rely on gradually refined grid search of welfare, as given by Equation 3.3.1. For each grid-point, we compute welfare and search for the maximum with each policy regime. All grids are equidistant. We rely on a grid for p of 40 points between 0 and 1.

For F, τ_{stw}, and ϵ_{stw}, we gradually narrow down the bounds of the intervals in which we search for a new optimum. For F, we start with a grid between 0 and 1.5. The results reported in the paper are based on a grid with limits 0.65 and 0.8. The F grid has 1000 points.



(a) Grid for τ_{stw} **(b)** Grid for ε_{stw}

Figure 3.H.1. Grid bounds for τ_{stw} and ε_{stw} across $p \in [0, 1]$.

The grids for τ_{stw} and ϵ_{stw} have 400 points each. We start with grids between 0.1 and 1.5 at every point on the p-grid for both parameters. The grids, resulting from sequential narrowing down bounds, on which the results in the paper are based, are shown in Figure 3.H.1.

Appendix 3.I Calibration

To calibrate the model, we solve the following system of equations for parameters, taking targets as exogenous. Note that we solve for some endogenous model variables, too. This works because the system of equations pins down the model solution for these variables, along with the parameters to be calibrated.

Exogenous:

$$f, \ q, \rho, b_{\text{rep}}, n$$

Endogenous:

$$c_f, \bar{m}, b, \epsilon_s^c, \epsilon_s^u, c^w, \omega$$

(I)

$$\rho \;=\; (1-p)\cdot G(\epsilon_s^u) + p\cdot G(\epsilon_s^c)$$

(II)

$$0 \;=\; z(\epsilon_s^u) + \frac{u(c^w)-u(b)}{u'(c^w)} - c^w + \frac{\lambda-\eta f}{1-\eta}\cdot\frac{k_v}{q}$$

(III)

$$0 \;=\; z(\epsilon_p) + \lambda\cdot\frac{k_v}{q} - c^w$$

(IV)

$$\frac{u(c(\epsilon_s^u))-u(b)}{u'(c^w)} + (\lambda - f)\cdot\frac{\eta}{1-\eta}\cdot\frac{k_v}{q} \;=\; 0$$

(V)

$$\begin{aligned}\frac{1}{1-\eta}\cdot\frac{k_v}{q} \;=\; \frac{1}{\lambda}\Bigg[&(1-\rho)\cdot\left(z-\frac{1-\eta}{\eta}b\right)\\ &+(1-p)(1-\rho^u)\cdot\left(\frac{u(c^w)-u(b)}{u'(c^w)}-c^w\right)\\ &+p(1-\rho^c)\cdot\left(\frac{u(c^c)-u(b)}{u'(c^w)}-e^c\right)\\ &+(1-\rho)\cdot\frac{\lambda-f\eta}{1-\eta}\cdot\frac{k_v}{q}\Bigg]\end{aligned}$$

(VI)

$$\begin{aligned}&(1-p)\cdot(1-\rho^u)\cdot\left(\eta\cdot c^w+(1-\eta)\cdot\frac{u(c^w)-u(b)}{u'(c^w)}\right)\\ &=\;\eta\cdot(1-\rho)\cdot\left(z+\frac{1-n}{n}\cdot b+\theta\cdot k_v\right)\\ &\quad-p\cdot(1-\rho^c)\cdot\left[(1-\eta)\cdot\frac{u^c-u(b)}{u'(c^w)}+\eta\cdot e^c\right]\end{aligned}$$

(VII)

$$\omega = \frac{1}{1-\rho} \cdot \Bigg((1-p) \cdot \int_{\epsilon_s^u}^{\infty} [c^w + \phi(h(\epsilon))] \, dG(\epsilon) + p \cdot \int_{\epsilon_p}^{\infty} [c^w + \phi(h(\epsilon))] \, dG(\epsilon) + p \cdot \int_{\epsilon_s^c}^{\epsilon_p} [c(\epsilon) + \phi(h(\epsilon))] \, dG(\epsilon) \Bigg)$$

(VIII)

$$b = b_{\text{rep}} \cdot \omega$$