Essays on Spatial Inequality, Mobility, and Labor Market Policy

Inauguraldissertation

zur Erlangung des Grades eines Doktors
der Wirtschaftswissenschaften

durch

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

vorgelegt von

Johannes Weber

aus Bonn, Deutschland

2026

Dekan:	Prof. Dr. Martin Böse
Erstreferent:	Prof. Dr. Moritz Kuhn
Zweitreferent:	Prof. Dr. Hans-Martin von Gaudecker
Tag der mündlichen Prüfung:	29. Juni 2026

Acknowledgements

I am deeply grateful to my advisors, Moritz Kuhn, Hans-Martin von Gaudecker, and Christian Bayer, for their guidance and support throughout my doctoral studies. Their feedback has been essential in shaping this dissertation and my way of thinking about research. I particularly appreciate the time and care they devoted to discussing my work and their support at important stages along the way. Their mentorship has had a lasting impact on my academic development.

I am grateful to the Bonn Graduate School of Economics, the Institute on Behavior and Inequality, and the Collaborative Research Center TR 224 for their generous support and for providing an outstanding research environment throughout my doctoral studies. I would also like to thank the many people who supported me along the way. In particular, I am deeply grateful to Kim Schneider, Beate Roters, Simone Jost, Ilona Krupp, and Silke Kinzig for their constant help and support.

One of the most rewarding parts of this journey was sharing it with a wonderful group of friends and colleagues, in particular Jae Youn, Gero, Lennard, and Jan, as well as many others whom I cannot name individually here. Navigating this intense period would not have been possible without their support. The long hours we spent working on problem sets, preparing for exams, and discussing our early research ideas were challenging at the time, but have since become some of my most valued and fond memories.

None of this would have been possible without the unconditional support of my family. I am deeply grateful for their encouragement and for always keeping me grounded throughout this journey.

Johannes Weber

Bonn, April 2026

Contents

List of Figures

List of Tables

Introduction

This dissertation consists of three self-contained essays that study spatial inequality, worker mobility, and labor market policy. The first two chapters focus on regional disparities in labor market conditions and their implications for workers' careers and mobility decisions. The third chapter examines the effectiveness of labor market policies aimed at mitigating inefficiently high levels of separations in light of differences in firms' ability to take on credits. Taken together, the three essays shed new light on key dimensions of labor market heterogeneity and their implications for policy design.

Chapter 1 studies how the careers of workers unfold depending on the conditions in the local labor market in which workers enter. A large literature has documented the persistent effects of entering the labor market during aggregate downturns, i.e., at a bad time. However, much less is known about the consequences of entering the labor market in a bad region with persistently high unemployment. Using rich administrative data from Germany, I show that workers who start their careers in high-unemployment regions experience systematically worse labor market outcomes. These workers spend more time unemployed, earn lower wages, and accumulate substantially less income over their lifetimes. Despite these persistent differences in labor market outcomes, worker mobility is low. Even when workers move, they usually move to nearby locations and do not systematically seek out better labor markets.

In Chapter 2, I investigate why worker mobility is so low, despite large regional differences in wages. I begin by documenting two key empirical facts. (1) The difference between high and low wage regions is not simply in the level but in wage growth over the lifecycle. Young workers in their 20s do not earn much more on average in high-wage regions than in low-wage regions. Taking regional price levels into account, it emerges that young workers actually earn less on average in high-wage regions. However, they experience faster wage growth, such that careers in high-wage regions pay off long-term. (2) Occupations are regionally concentrated, and workers in high and low-wage regions do different kinds of jobs. Since it is known in the literature that large parts of human capital are occupation-specific, this finding implies that human capital cannot easily be transferred across regions.

Motivated by these findings, I then develop and calibrate a general equilibrium life-cycle model with region-specific human capital accumulation and local housing markets. In the model, different regions emphasize different skills in production, and workers accumulate the locally relevant skills more rapidly through experience. The calibrated model shows that low mobility arises from the dynamic interaction between labor and housing markets. Young workers, who have little wealth, are deterred from moving to high-wage regions by high housing costs. Older workers, in contrast, can afford higher rents but are less willing to move because their accumulated human capital is region-specific and loses value upon relocation.

Given that high housing costs are the main barrier preventing young workers from moving to high-opportunity regions, I evaluate a policy that increases housing supply in these regions. I show that expanding housing supply lowers rents, increases mobility toward high-growth regions, and generates welfare gains in both regions.

In Chapter 3, joint with Gero Stiepelmann, I study the design of labor market policies aimed at mitigating inefficiently high separation rates. Excess separations are a well-known feature of economies with unemployment insurance systems. Governments typically rely on two main instruments to address this inefficiency: lay-off taxes, which penalize firms for separating from workers, and short-time work schemes, which subsidize reductions in working hours and partially compensate workers for lost earnings. While both instruments are widely used in different countries, their relative effectiveness remains not fully understood.

The existing literature emphasizes the efficiency of lay-off taxes, while short-time work is known to introduce distortions through inefficient reductions in working hours. However, these results rely on the implicit assumption that firms are not financially constrained. In this chapter, we relax this assumption and develop a search-and-matching model with endogenous working hours, risk-averse workers, and financial constraints affecting a share of firms. Using a Ramsey planner approach, we characterize optimal lay-off taxes and optimal short-time work, taking the share of financially constrained firms into account.

We find that lay-off taxes are more effective at improving welfare when only a small share of firms is financially constrained. However, as financial constraints become more prevalent, short-time work emerges as the preferred policy instrument, as it both reduces separations and provides income insurance where firms can no longer do so. Quantitatively, we find that short-time work becomes the superior policy once more than 27% of firms are financially constrained.

The three essays highlight the importance of labor market heterogeneity, both across regions and across firms. Differences in local labor market conditions and in firms' financial constraints shape workers' careers, mobility decisions, and exposure to risk. These mechanisms are central to understanding persistent inequality in labor market outcomes and how policies can be designed to address these disparities.

Chapter 1

Local Unemployment, Worker Mobility and Labor Market Outcomes: Evidence from Germany*

1.1 Introduction

It is well documented that the level of unemployment matters for the development of workers' labor market outcomes. Entering the labor market in a recession at a time of high unemployment can be detrimental to workers' wages even in the long run (e.g., Lisa B. Kahn, 2010b; Oreopoulos, Von Wachter, and Heisz, 2012a; Altonji, Kahn, and Speer, 2016; Berge, 2018). However, the level of unemployment that young workers face at the beginning of their careers does not only change over time. In most developed economies, there are large and highly persistent differences in unemployment between regional labor markets as well.[1]

Policymakers have long recognized this as a source of inequality. Various policy measures aimed at supporting bad local labor markets are put in place by local and national governments in most countries.[2]

* I thank Moritz Kuhn and Hans-Martin von Gaudecker for invaluable advice and guidance in this project. I also thank participants in seminars at Bonn and Yale as well as participants at various conferences for invaluable feedback and advice. I gratefully acknowledge financial support from the Bonn Graduate School of Economics. Support by the Deutsche Forschungsgemeinschaft (DFG) through CRC TR 224 (Project C01) is gratefully

1. For documentation of these differences, see e.g. Topel (1986) or Kline and Moretti (2013) for the U.S., Bilal (2023) for France, Kuhn, Manovskii, and Qiu (2026) for the UK, Germany, and the U.S., and OECD (2005) for an overview of OECD economies.

2. In Germany, too, a set of measures (Gesamtdeutsches Fördersystem) aimed at supporting structurally weak areas relies on local unemployment as a means of determining which area is structurally weak and eligible for support, including subsidies for young firms and plants in structurally weak areas.

Despite its evident relevance, the effect of entering a career in a bad local labor market has, unlike the effects of starting a career at a bad time, received only scant attention in the literature. In this study, I fill this gap and ask whether it matters for the careers of workers if they enter the labor market in a place of low or high unemployment.

There are two components to this question. On the one hand, it is a question about differences in labor market outcomes, i.e., time in unemployment, wages, and ultimately lifetime income. On the other hand, it is also a question about the spatial mobility of workers between local labor markets, since in principle, workers are free to move and look for a job in an area with more favorable conditions. To answer both of these components, I make use of high-quality administrative data from Germany that allow me to track workers over their careers. I study several cohorts of West German male workers born between 1960 and 1980 and their labor market outcomes over the course of their careers, depending on where they entered the labor market.

Germany provides the ideal setting to study this research question. It is a reasonably large country with many local labor markets and considerable differences in local unemployment. Behind the overall unemployment rate of 6.7% in 2014, for example, there were local unemployment rates as high as 14.7% in Gelsenkirchen, a city in the industrial Ruhr area, and as low as 2.6% in the Bavarian town of Dillingen a. d. Donau.

The main contribution of this study is to document two novel facts: Firstly, worker mobility between good and bad labor markets is fairly limited, and workers spend most of their careers in areas with similar relative degrees of unemployment. Secondly, workers from better labor markets earn higher wages, experience greater wage growth, and spend less time unemployed. I find that these differences translate into substantial differences in earnings accumulated over workers' careers.

As a first step, I turn to worker mobility and ask whether workers systematically leave bad local labor markets. This turns out not to be the case. Instead, I find that about 40% of workers never leave the commuting zone they entered the labor market in, and that workers who move to jobs in other commuting zones usually do not move very far away. Close to 90% of workers never move to a commuting zone that is 200 km or further away from their initial commuting zone. As a result, workers who do move usually move to areas with similar relative degrees of unemployment and spend most of their careers in local labor markets of similar levels of unemployment.

With this in mind, I compare how much time workers spend in unemployment depending on how high unemployment is in their initial local labor market. I find that workers who enter the labor market in worse local labor markets lose their jobs more often and have longer unemployment spells on average. Up to the age of 55, workers who enter the labor market in a commuting zone of the highest unemployment quintile spend about a year more time in unemployment than their counterparts from the lowest unemployment quintile.

I compare the wages of workers and find that the differences in wage levels and wage growth persist over the lifecycle. The wage gap for full-time employed workers amounts to about 7-10% at 25 and opens up to roughly 20% towards the end of the lifecycle, depending on the cohort. These differences cannot fully be explained by local price levels. In particular, I find that even controlling for worker fixed effects leaves an unexplained wage gap that opens up to about 11% over the lifecycle.

Taking both differences in wages and differences in unemployment together, I calculate average accumulated earnings up to the age of 55. I find that the difference between the highest and the lowest quintile amounts to about €250,000 (at 2010-CPI level) on average. Controlling for local price levels and taking away the part of accumulated earnings explained by observables (education, a worker's occupation, and industry and local urbanization) leaves an average residual gap in these accumulated earnings of about €50,000 (at 2010-CPI level).

I conclude that it does indeed matter considerably whether workers enter the labor market in a good or in a bad local labor market. If anything, these differences are even more important for younger cohorts.

The remainder of the paper is structured as follows: The following paragraph summarizes related literature. Section 1.2 describes the data used for the analysis. Section 1.3 gives an overview of local unemployment rates in Germany and local labor markets. Section 1.4 documents mobility patterns between the different local labor markets. Section 1.5 describes differences in labor market outcomes and accumulated earnings. Section 1.6 concludes.

Related literature While labor market outcomes over the lifecycle in the context of spatial unemployment differences have, to my knowledge, never been studied before, there is a body of literature that this work relates to. There are three main branches of related literature.

Firstly, there are papers concerned with spatial unemployment. Focusing on excessive hiring costs and different local productivity levels Kline and Moretti (2013) discusses circumstances under which place-based hiring subsidies may be beneficial. Bilal (2023) stresses the role of firms and attributes spatial unemployment gaps in France to the agglomeration of productive firms in certain places, leading to unproductive worker-firm matches and high job-losing probabilities in bad labor markets. On the empirical side, papers often exploit regional employment shocks (e.g., Hornbeck and Moretti, 2022). Perhaps most relevant to my work, Blanchard et al. (1992) document responses to averse regional employment shocks in the context of U.S. states and find only transitory effects of unemployment shocks on wages. A main difference between these papers and my work is that I can track individual workers over their careers. Blanchard et al. (1992), for instance, rely on state aggregate wage data.

More broadly, there is a literature on spatial economic activity that does not necessarily focus on unemployment. Moretti (2011) gives an overview of what is known about gaps in local labor market outcomes and possible explanations for them. Other papers develop models and conditions to explain economic activity differences across space (e.g., Allen and Arkolakis, 2014; Gaubert, 2018). Redding and Rossi-Hansberg (2017) provide an overview of modeling tools and model building blocks for spatial models. Another well-studied area of spatial economics is the one concerned with wage and productivity gaps between rural and urban areas and urban wage premia (e.g., Harris and Todaro, 1970; Glaeser and Mare, 2001; Gould, 2007; Baum-Snow and Pavan, 2012; Young, 2013; Dauth et al., 2022). Local mismatch is another subject more directly related to local unemployment that has received attention. Şahin et al. (2014) and Marinescu and Rathelot (2018) focus on the mismatch of workers and local labor markets and its contribution to overall unemployment, for example.

Secondly, there is a literature on worker mobility. Going back to Sjaastad (1962), a branch of spatial economic research focuses on workers' location decisions and worker migration and migration frictions. For example, Roback (1982) presents evidence that local amenities are reflected in local labor market outcomes such as wages. More recently, Amior and Manning (2018) attribute the persistence of local unemployment despite a strong moving response by workers to persistent labor demand shocks. Focusing on moving frictions Bryan and Morten (2019) and Morten and Oliveira (2016) estimate large effects of facilitating worker relocation on productivity in the context of developing countries. A number of papers exploit various employment shocks and study the mobility response (e.g., Blanchard et al., 1992; Levy, Mouw, and Perez, 2017). Nakamura, Sigurdsson, and Steinsson (2016) use a volcanic eruption in Iceland to study the inter-generational effect of worker relocation. Huttunen, Møen, and Salvanes (2018) find that displaced workers who decide to move close to their families or rural areas after displacement experience income losses. Again, a difference between this literature and my work is that I follow individual workers over their whole employed working lives.

Thirdly, a branch of literature concerns itself with wage structure and the German wage structure in particular. Dustmann, Ludsteck, and Schönberg (2009) document that wage inequality has been on the rise in Germany in the 1980s and 1990s, and Bönke, Corneo, and Lüthen (2015) document a rise of intra-cohort wage inequality. Card, Heining, and Kline (2013) find that the dispersion of task and firm-specific premiums has gone up in West Germany. Relating to regional labor market differences, a number of papers study differences between East and West Germany (e.g., Uhlig, 2006; Spitz-Oener, 2007; Uhlig, 2008). My work adds to this literature by documenting the additional spatial component of wage inequality in West Germany.

1.2 Data

To analyze labor market outcomes over the lifecycles of workers depending on the location a worker works, data that is both regional and allows me to track workers and their outcomes and whereabouts over time is needed. The main data source that I choose for this purpose is the Sample of Integrated Labor Market Histories (SIAB) - a 2% sample of workers covered by German social security records. Additionally, the SIAB contains information on short-term unemployment benefits workers receive. Since in Germany only employed (and unemployed) workers are covered by social security, the sample does not contain self-employed workers or workers who are civil servants in government positions. However, a large share of the German labor force - about 80% - is covered by social security. The data version that I use covers the years 1975-2017. Not only does this data meet the mentioned requirements, but it is highly credible administrative data as well. This is important because, in many non-administrative surveys, there could be the worry that people who move to another location have a high probability of dropping out of the sample. With administrative data, this is not a problem.

Because the SIAB is drawn from social security records, wages are only reported up to the maximum social security contribution level. To circumvent this problem, I follow the procedure proposed by Gartner (2005) and impute top-coded wages. I describe the details of this procedure in Section 1.A.1 of the appendix.

The regional variable contained in the SIAB is the county (Landkreis) in which a worker's current job is registered. This means that for workers who are currently unemployed, there is no information about their whereabouts. To deal with this issue, I assume that workers who lose their jobs do not immediately move to another location but look for a new job first and move only when starting a new job in a different county. I then fill the gaps in the location variable using the last valid county reported for each worker.

German counties are relatively fine-grained regional units. It is likely that in some places many workers will live in one county and commute to work in another one. It is therefore questionable if counties are a good measure of local labor markets. To circumvent this potential problem, I turn to commuting zones published by the Federal Office for Building and Regional Planning (Bundesamt für Bauwesen und Raumordnung - *BBSR*). These commuting zones group counties into zones, taking into account commuting streams between the counties, such that the resulting commuting zones are as self-contained as possible.[3] There are 203 such commuting zones in West Germany.

I impose a couple of restrictions on the sample. First of all, workers from East Germany are recorded in the SIAB only from 1993, and after decades of socialism, labor markets in East Germany are likely to be inherently different along many dimensions. For the purpose of this study, I need as much comparability over time as possible and do not want to deal with these differences. I therefore restrict the sample to individuals who have no spells in any East German county. Secondly, because I track cohorts born in the '60s and '70s when female labor supply was much smaller than today, I restrict the sample to men. This also ensures comparability with the related literature that usually imposes the same restriction.

Apart from worker data, data on unemployment at a regional level is needed. As my primary data source, I use county-level yearly unemployment rates I obtained from the BA. This data reports both unemployment rates and the number of registered unemployed workers. I use this information to calculate unemployment rates at the commuting zone level. While this is the most credible unemployment data source for Germany, county-level data from the BA is unfortunately only available for the years after 1985. Before 1985, the SIAB can be used to directly estimate commuting zone-level unemployment. I show that the correlation between SIAB estimates and BA unemployment rates is 92% in 1985 in Figure 1.A.1 in the appendix.

It is well known that price levels can vary greatly between different regions within a country. To measure this, I use a regional CPI at the county level from the *BBSR* for the year 2007 to measure local price levels. Unfortunately, no comprehensive regional CPI with price measurements over time exists for Germany.[4] Of course, price levels can change over time, and calculating real wages using a CPI based on prices measured in 2007 for other years is potentially problematic. I take regional building land prices from the federal office for statistics, as well as a local rent index from the *BBSR*, to validate this CPI for other years Figures 1.A.3 and 1.A.2 in the appendix show that the correlations between the regional CPI and the prices in the available years that are the closest and the furthest apart from 2007 are very similar.

3. The counties reported in the SIAB are counties as of 2017. Because of reforms, counties' boundaries do change over time. In West Germany, however, this is relatively rare. There are several versions of these commuting zones from different years. The last version before 2017 is from 2011. Since there are only few differences between the 2011 and earlier versions and county codes in the SIAB are valid in 2017 I use the 2011 commuting zones. I use a county code conversion key, also obtained from the Federal Office for Building and Regional Planning, to translate the 2011 commuting zone key into 2017 county terms.

4. The BBSR did collect regional price data in 1994, but compared only 50 cities across Germany, not taking into account rents. Kosfeld, Eckey, and Schüßler (2009) propose an econometric model to estimate regional CPIs based on this data, but working with the comprehensive CPI from 2009 seems preferable.

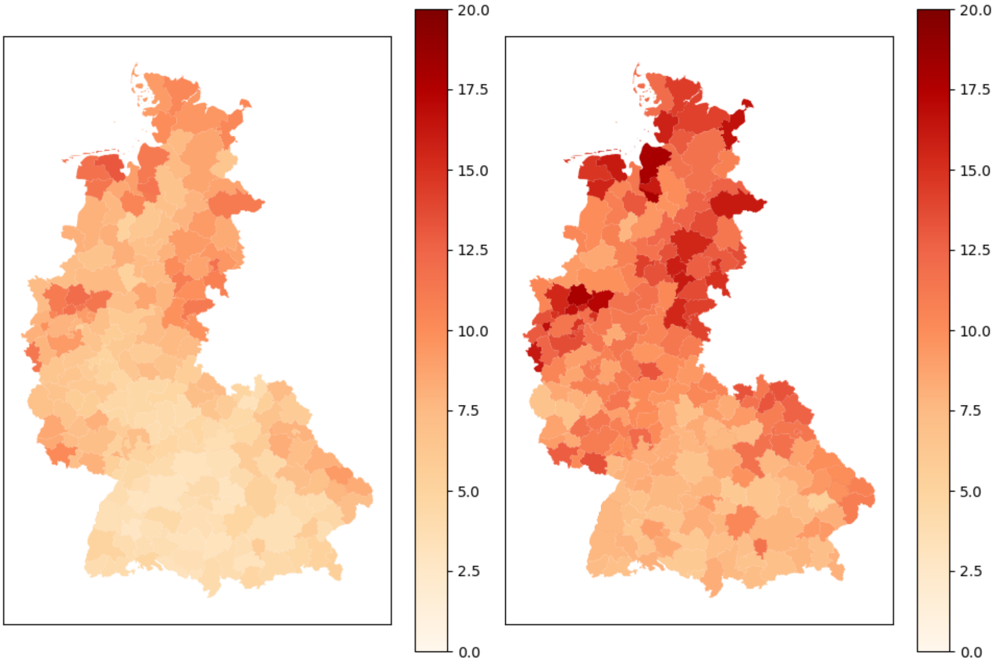
1.3 Local Unemployment in West-Germany and Definitions

1.3.1 Local Unemployment

Figure 1.1 shows unemployment rates across West German commuting zones for the years 1990 and 2005 - a year with comparatively low aggregate unemployment and a year with infamously high aggregate unemployment. In both years, the heterogeneity in unemployment is well visible with local unemployment rates ranging from 2.8% to 13.06% in 1990 and from 5.5% to as much as 21.5% in 2005.

Figure 1.1 also hints at another well-established fact. Local unemployment is typically very persistent over time. Figure 1.2 makes this explicit and plots commuting zone level unemployment rates of 1997 and 2005, as well as 1990 and 2010, against each other. 1997 and 2005 are two years with high aggregate unemployment, while 1990 and 2010 saw much less aggregate unemployment. In both cases, the green dots that represent the commuting zones align closely along the 45-degree line, indicating a very high level of persistence. Given that there are 20 years between 1990 and 2010, it is striking how high the persistence of local unemployment is, with a correlation of close to 80% between local unemployment rates in the two years.

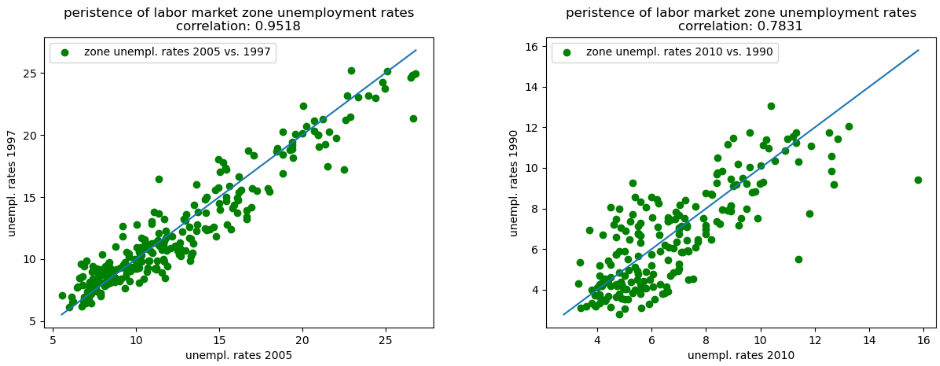
Figure 1.1. Unemployment across commuting zones



(a) 1990 *(b) 2005*

Notes: The scale is in %-terms. The maps shown are for the years 1990 (left) and 2005 (right). The regions shown are the commuting zones described in Section 1.2. Unemployment data comes from the BA.

Figure 1.2. Persistence of Unemployment



(a) 1997-2005 *(b) 1990-2010*

Notes: 1997 unemployment rates against 2005 unemployment rates (left) and 1990 unemployment rates against 2010 unemployment rates (right). Each green dot represents a commuting zone. The scale is in %-terms.

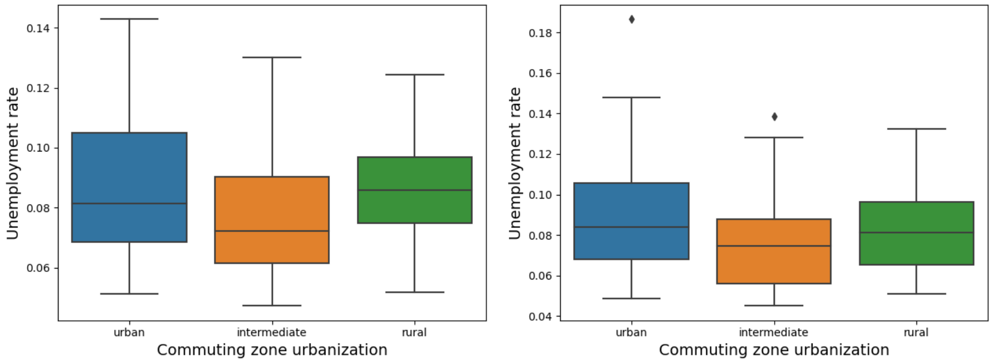
Given this persistence, a natural expectation could be that high and low-unemployment commuting zones are different in some fundamental way. However, in terms of key observables - urbanization and education - this is not the case. Figure 1.3 shows the distributions of average unemployment rates over the 1990s and 2000s decades in the three urbanization categories that the *BBSR* assigns to the commuting zones described in the previous section. Clearly, while it is true that the commuting zones with the most extreme unemployment are urban, it is not true that high unemployment exists only in urban commuting zones. In fact, in both decades, the median unemployment rate in rural areas is higher than in urban areas. Generally, the distributions cover a similar support.

Next, I turn to educational attainment. Figure 1.4 plots average unemployment rates against average shares of workers with college degrees in commuting zones in regression plots. To calculate the college share, I compute the share of workers in the SIAB who report a tertiary degree.[5] Figure 1.4 reveals that by and large, there is no systematic correlation between local unemployment and the college share.

As a consequence, the findings presented in the remainder of this study should not be regarded simply as an artifact of urban and rural differences or agglomeration of educated and uneducated workers. Instead, they are the result of structural differences in the quality of local labor markets.

5. I include both university and university of applied science degrees.

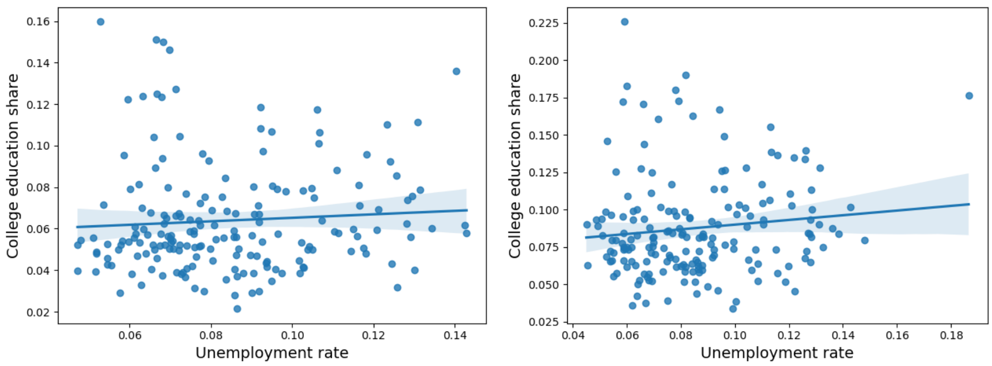
Figure 1.3. Unemployment and Urbanization



(a) 1990s *(b) 2000s*

Notes: Boxplots of average unemployment rates over the 1990s (left) and the 2000s (right) in the three urbanization categories for *BBSR* commuting zones. The boxes represent the quartiles of the distribution, the whiskers extend to the unemployment rate the furthest away but within 1.5 inter-quartile ranges from the box. Points beyond that are shown as diamonds. The line through the box represents the median.

Figure 1.4. Unemployment and Education



(a) 1990s *(b) 2000s*

Notes: Regression plots of average unemployment rates over the 1990s (left) and the 2000s (right) against the average share of workers with a college degree.

1.3.2 Definitions

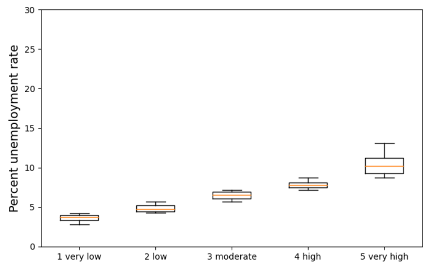
In order to measure differences between local labor markets with high and low unemployment, I need a notion of what "high" and "low" unemployment mean. In particular, what really matters for my analysis is not the absolute level of unemployment itself, but how high unemployment in a commuting zone is, relative to the other commuting zones. I decide to work with quintiles and assign the ranks *2 very low* unemployment, *3 low* unemployment, *4 moderate* unemployment, *5 high* unemployment, and *6 very high* unemployment to each commuting zone each year, taking quintiles over commuting zone-level unemployment rates. I refer to these ranks as the *plain ranks* in the remainder of the paper.

When comparing workers, however, I need one more step because, in principle, workers can move between the commuting zones. When I want to compare workers over their lifecycle, I use the most natural option: I simply use the *plain rank* of the commuting zone a worker enters the labor market in for the very first time, in the year the worker starts working there. I refer to this as the *entry rank*.

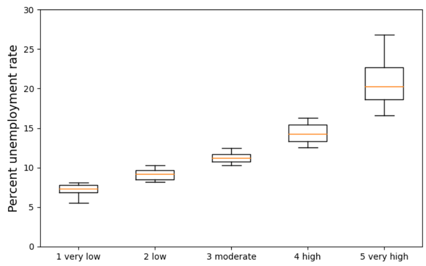
As Figure 1.5 demonstrates, depending on the overall level of unemployment in a given year, the *plain ranks* can be assigned to quite different absolute levels of unemployment. While I am interested in relative differences in local unemployment, it is important to keep these considerable differences in overall unemployment between different points in time in mind. Workers who enter the labor market at a time of high overall unemployment are likely to experience different labor market outcomes than a worker who enters under more favorable conditions.[6] To avoid these differences driving the results I measure, I take a cohort perspective and group the birth years of workers in the SIAB into cohorts. Since I want the cohorts to face similar conditions in the labor market over their lifecycles, I choose cohorts of 2-3 birth years such that the overall level of unemployment in West Germany is similar for all members of a cohort at age 25. Because unemployment was very low in the 1970s and early 80s, I focus on later cohorts who are faced with more substantial unemployment rates during their careers and consider the cohorts 1960-1962, 1963-1965, 1966-1967, 1968-1970, 1971-1973, 1974-1976, and 1977-1979. Many findings are very similar for all cohorts. Instead of showing results for all cohorts in the main text, unless labeled otherwise, I hence show results for the 1963-1965 and the 1971-1973 cohorts in the main text and results for the remaining cohorts in the appendix. I do not consider later cohorts because the SIAB data covers only the years until 2017, and workers born after the 1970s can only be observed until their 30s, which makes comparing outcomes over the lifecycle difficult.

6. See e.g. Lisa B Kahn (2010a) or Oreopoulos, Von Wachter, and Heisz (2012b).

Figure 1.5. Unemployment rates in the different ranks: 1990 vs. 2005



(a) 1990



(b) 2005

Notes: Boxplots of unemployment rates by rank in the commuting zones by the rank the commuting zones are assigned. The box plots display the upper and lower quartiles, median, and whiskers extending to the furthest data points within 1.5 times the interquartile range. The left side of the figure corresponds to the plots for the year 1990, while the right side represents the plots for the year 2005.

1.4 Worker Mobility

A central question to assess how harmful starting a career in a bad local labor market is is how mobile workers are between local labor markets of different quality. If workers can freely migrate from places with high unemployment to places with better conditions, entering the labor market in a place with high unemployment should not have great effects on long-term labor market outcomes.

As a first step, ignoring local unemployment, I study the overall mobility of workers over the course of their entire careers. Table 1.1 shows the fraction of careers that take place within a given distance from the commuting zone a worker starts his career in for different cohorts. To ensure that results are not driven by workers who cannot be observed for sufficiently long, I restrict the sample to workers who are first observable before or at the age of 25 and who can still be observed at or after the age of 40. As column 1 indicates, about 40% of workers work in the same commuting zone for their entire careers. Column 2 reveals that almost 60% of workers never move further than to a commuting zone, adjacent to their initial commuting zone. Columns 3 and 4 show that about 80% of workers never move further than 100 km and almost 90% of workers never move further than 200 km away from their starting commuting zone.

The high figures presented in table 1.1 establish the fact that overall worker mobility is low and that most careers in West Germany take place within narrowly confined areas. Nevertheless, about 60% of workers do move at some point in their careers. This raises the question if workers who do move seek out good local labor markets and avoid bad ones. To provide an answer, I collect all moves between

Table 1.1. Fractions of workers who stay within a given distance from their first commuting zone

Birth Years	Stayers	Neighboring Zone	100 km	200 km
1960-1962	0.42	0.59	0.79	0.87
1963-1965	0.43	0.59	0.79	0.87
1966-1967	0.42	0.59	0.79	0.86
1968-1970	0.42	0.59	0.79	0.87
1971-1973	0.41	0.58	0.78	0.86
1974-1976	0.41	0.56	0.78	0.86
1977-1979	0.42	0.56	0.77	0.86

Notes: The sample is restricted to workers with observations at age 25 and at or after age 40. The first column shows the fraction of workers who stay in the same commuting zone for their whole observable career. The second column shows the fraction of workers who never move further than to a commuting zone adjacent to their initial commuting zone. The third and fourth columns show the fractions of workers who stay within a radius of 100 km and 200 km, respectively. The distance between two commuting zones is measured as the minimum distance between the borders of these commuting zones.

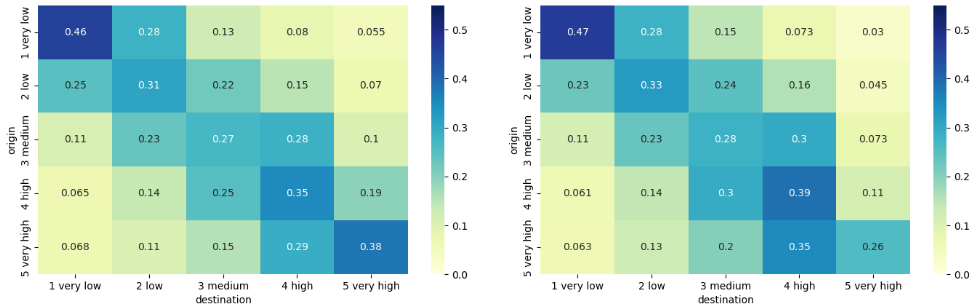
commuting zones that happen throughout the whole careers of workers[7] and check which of the quintile ranks the move is from and which of the ranks the move is to. Then, for each of the ranks of the commuting zones that the moves are from, I calculate the relative frequencies of moves that go to commuting zones with each of the 5 ranks a move can be to:

$$a_{ij} = \frac{\#\text{moves from rank } i \text{ to rank } j}{\#\text{moves from rank } i}, \quad i,j = 1,2,3,4,5$$

The resulting 5×5 matrix A is shown in Figure 1.6 for the 1960-1962 and 1971-1973 cohorts. For example, the top left corner of the left matrix states that 46% of all moves away from a commuting zone with the rank *1 very low*, among workers born between 1963 and 1965, were to another commuting zone with the same rank. By contrast, the top right corner of the same matrix states that only 5.5% of moves away from a *1 very low* commuting zone were to commuting zones with the rank *5 very high*. As the deeper shades of blue towards the diagonals of the two matrices suggest, workers, when they move, tend to move to commuting zones with similar levels of unemployment to the commuting zone they are leaving. Perhaps unsurprisingly, this is particularly pronounced for moves away from very good labor markets, more than 40% of which go to *1 very low*-commuting zones in all cohorts. But even for moves away from high unemployment zones, moves tend to be to places with similar degrees of unemployment. In all cohorts, over 60% of moves away from commuting zones with rank *5 very high* are to a zone with either the rank *5 very high*

7. Results are very similar when looking at moves at different ages separately.

Figure 1.6. Fractions of moves by unemployment rank of origin- and destination commuting zones



(a) 1960-1962 Cohort *(b) 1971-1973 Cohort*

Notes: Each cell in the matrices above represents the fraction of moves away from commuting zones with an unemployment rank indicated by the row to a commuting zone with an unemployment rank indicated by the column. E.g., the second cell in the first row of the left matrix indicates that 28% of moves away from zones with rank *1 very low* went to zones with rank *2 low*. Workers are ranked based on the unemployment quintile of the commuting zone where they initially enter the labor market. The matrix on the left shows the results for the 1960-1962 cohort, and the matrix on the right shows the results for the 1971-1973 cohort.

or *4 high*. Fewer than 7% of workers who leave commuting zones with the highest two ranks move to the very best labor markets with the lowest rank. While natural expectation could be that this pattern could be very different if only the moves of unemployed or employed workers were counted, I show in Figure 1.C.3 and Figure 1.C.4 in the appendix that this is not the case and that the pattern does not change much even when workers are unemployed just before moving.

Given that, as discussed, most workers never move over great distances, the question arises whether these findings are simply a product of agglomeration of good and bad local labor markets or whether these patterns persist conditional on the distance over which a move occurs. To formally test whether this is the case, I adopt a gravity view of migration flows and impose that the number of people f_{ijt} that move between zones i and j in year t are a function of the distance d_{ij} between the two zones, size of the two zones s_i and s_j and the numbers of unemployed workers in both zones u_i and u_j, as well as a year effect β_t. Specifically, I assume that

$$f_{ijt} = \beta_t \frac{s_{it}^{\beta_1} s_{jt}^{\beta_2} u_{it}^{\beta_3} u_{jt}^{\beta_4}}{d_{ij}^{-\beta_5}}. \tag{1.1}$$

Taking logs equation 1.1 can then be estimated as

$$\log f_{ijt} = \log \beta_t + \beta_1 \log s_{it} + \beta_2 \log s_{jt} + \beta_3 \log u_{it} + \beta_4 \log u_{jt} + \beta_5 \log d_{ij} + e_{jit} \tag{1.2}$$

where e_{ijt} is the error term, and the coefficients can be interpreted as moving flow elasticities. Since the SIAB is a 2% sample of the German labor force that is representative at the county level, I estimate the size of each commuting zone by multiplying the number of workers in each commuting zone by 50. f_{ijt} is the total number of moves from i to j for all commuting zones i and j in the SIAB. I use data from the years 1985-2017. Numbers of unemployed workers come from the BA. The distance between commuting zones is measured as the distance between the two centroids of the two commuting zones. Results are reported in column (1) of Table 1.2.

Table 1.2. Gravity regression results

	(1)	(2)
log size of origin zone	0.2210	0.2361
	(0.011)	(0.010)
log size of destination zone	0.2500	0.2364
	(0.010)	(0.010)
log distance	-0.7102	-0.6987
	(0.081)	(0.083)
log n. unempl. origin	0.0370	-
	(0.006)	-
log n. unempl. destination	0.0085	-
	(0.009)	-
log (n. unempl. or.) × (n. unempl. dest.)	-	0.0218
	-	(0.007)
year fixed effects	yes	yes

Notes: The outcome variable is the log of the number of workers who moved from the origin commuting zone to the destination commuting zone. Standard errors are clustered at the year level and reported in parentheses below each coefficient.

Unsurprisingly, the size of both the origin and the destination commuting zone is positively associated with the number of moves, and the elasticity with respect to distance is negative and sizeable, further emphasizing the point that workers rarely move over long distances. The smaller and positive coefficient for the log number of unemployed workers in the zone of origin is in line with more unemployed workers looking for and eventually finding jobs in different commuting zones. Crucially, though, the coefficient for the log number of unemployed workers in the destination zone is neither negative nor statistically significant. This means that even conditional on distance, it is not the case that workers seek out commuting zones with less unemployment.

In order to be able to reproduce the finding from Figure 1.6 that, moreover, workers who move tend to move to similar commuting zones in terms of local unemployment, I also estimate a gravity equation that includes an interaction term of the numbers of workers who are unemployed in both origin and destination zone. The gravity equation then becomes

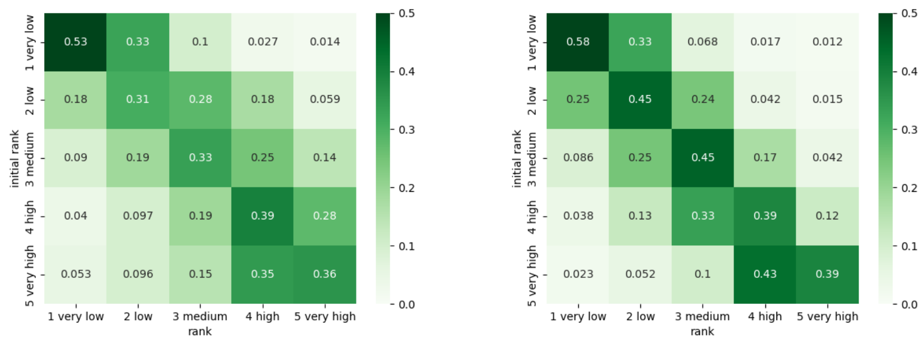
$$f_{ijt} = \beta_t \frac{s_{it}^{\beta_1} s_{jt}^{\beta_2} (u_{it}u_{jt})^{\beta_3}}{d_{ij}^{-\beta_4}} \tag{1.3}$$

Results from estimating this equation in logs are in column (2) of Table 1.2. The coefficient for the interaction term is indeed positive and statistically significant, if far smaller than the coefficient for distance. This means that even conditional on distance, workers who move tend to move to commuting zones with similar degrees of unemployment.

The persistence of local unemployment and the now documented lack of mobility of workers over long distances and between good and bad local labor markets together imply that workers spend most of their careers under similar conditions in terms of local unemployment. Figure 1.7 puts numbers to this fact. It shows the average fraction of days in a career spent in each of the local unemployment quintile ranks, conditional on the *plain rank* of the commuting zone a worker first entered the labor market in at the time of labor market entry. It is worth remembering that, while as demonstrated, regional mobility of workers is limited and local unemployment is quite persistent over time, in principle, over the course of a career, both locations and a location's unemployment rank can change. Nevertheless, the rank of a worker's initial commuting zone is a good predictor of the degree of unemployment that a worker is faced with throughout his career. Workers born between 1960 and 1962 who start their observable careers in a commuting zone with rank *5 very high* spend 36% of their time in commuting zones that have the same rank. Strikingly, as much as 71% of time is spent in commuting zones with either the rank *5 very high* or with the rank *4 high* on average. At the other extreme, more than half the time of workers who start their careers in the best labor markets with rank *1 very low* is spent in zones with that same rank, and more than 80% of these workers' time is spent in places with the lowest two ranks. Results for the other cohorts are quite similar.

The results presented in this section suggest that unemployment rates are persistent not only over time across locations, as demonstrated in Section 1.3. In fact, as a result of this regional persistence and limited worker mobility, they are persistent at the worker level as well. Workers who start their careers in depressed local labor markets are likely to spend most of their time in local labor markets with a high degree of local unemployment.

Figure 1.7. Fractions of time spent in the five unemployment ranks



(a) 1960-1962 Cohort *(b) 1971-1973 Cohort*

Notes: Fractions of time spent in commuting zones of each of the five ranks by the rank of the initial commuting zone where a worker first entered the labor market. The initial ranks are on the y-axis. For example, the second cell in the first row of the left graph shows that workers who start in a commuting zone with rank *1 very low* spend 32% of their time in commuting zones with rank *2 low* on average. Workers are ranked based on the unemployment quintile of the commuting zone where they initially enter the labor market. The matrix on the left shows results for the 1960–1962 cohort, and the matrix on the right shows results for the 1971–1973 cohort.

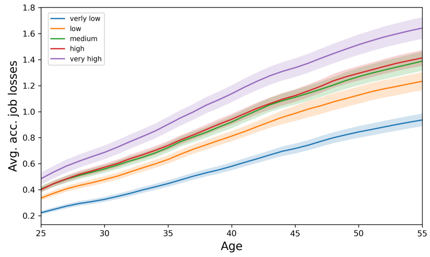
1.5 Workers' Labor Market Outcomes

1.5.1 Time in Unemployment

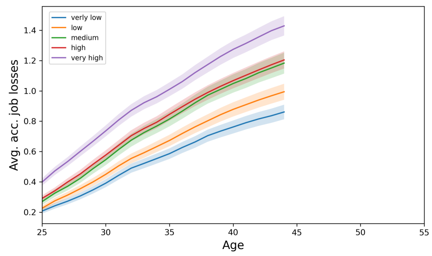
Having established that workers usually spend most of their time in similar local labor markets, I now turn to the question of what this means for their labor market outcomes. The most direct way in which long periods of time could be harmful to workers is leading to more unemployment. I turn my attention to this channel now. Because the SIAB reports observations exactly to the day I can not only detect a job loss in a worker's labor market history but also calculate how long this unemployment spell lasts. A caveat is that not all sources of unemployment spells exist for all years. Before 1997 only spells during which workers receive short-term unemployment benefits are recorded in the data.[8] From 1997 onwards workers show up as unemployed whenever they are registered as unemployed with the Federal Employment Agency (BA). As a result, in this Section numbers for years before 1997 exclude workers who are long-term unemployed.

8. In Germany, every employed worker contributes to unemployment insurance. Workers who were employed for 12 months prior to becoming unemployed receive unemployment benefits from the insurance for 6-12 months (up to 24 months for workers who are older than 50). Workers who do not meet this criterion or remain unemployed for longer periods of time receive social benefits (ALG II). Information about these unemployment benefits is contained in the SIAB. Unfortunately, there is no information about ALG II recipients before the year 2005 in the SIAB.

Figure 1.8. Average accumulated number of job losses over the lifecycle



(a) 1960-1962 Cohort



(b) 1971-1973 Cohort

Notes: The average number of job losses workers from each rank have had at each age. Workers are ranked based on the unemployment quintile of the commuting zone where they initially enter the labor market. Workers who have no observations after an age younger than 55 are assumed to lose no further jobs. Job losses are defined as transitions from an employed spell to an unemployed spell with a gap of fewer than 31 days between them. Birth years 1960-1962 on the left, birth years 1971-1973 on the right.

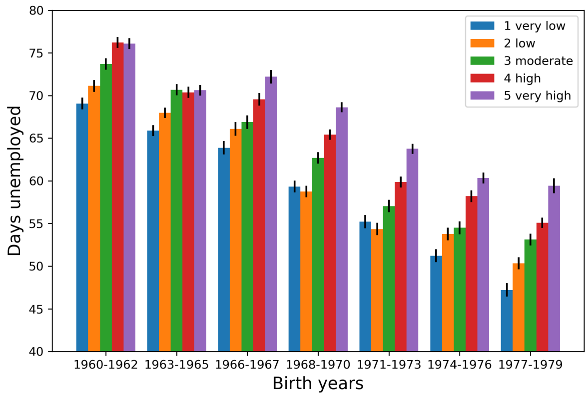
Figure 1.8 shows how often workers have lost their jobs on average at each age. I count any occurrence of an unemployment spell within a month after having been employed as a job loss to ensure that I include job losses where workers fail to register as unemployed immediately which could result in short gaps in the career history. I again use the *entry rank* described in section 1.3.2 to rank workers.

It is evident from Figure 1.8 that the higher the unemployment rank, the more job losses workers accumulate on average. This is true for all other cohorts as well, as demonstrated in Figure 1.C.7 in the appendix. At age 55, workers who started in the rank *1 very low* accumulated about 0.8 job losses on average while their counterparts who started in the rank *5 very high* accumulated about 0.6. At the end of their observable career, (i.e. at age 43) workers with the worst rank have accumulated 0.6 more job losses on average than their peers with the best rank. In all other cohorts, the difference is at a similar order of magnitude - confer Figure 1.C.7.

Apart from losing their jobs more often, another channel through which high local unemployment may be harmful is greater difficulty in finding a new job, once unemployed. Figure 1.9 shows the average length of an unemployment spell in each cohort depending on the rank. Despite the clear trend of unemployment spells becoming shorter on average for younger cohorts, it is clear that workers with higher unemployment ranks have longer unemployment spells. Workers with the highest unemployment rank spend more than 10 days longer in unemployment than their peers with the lowest rank.

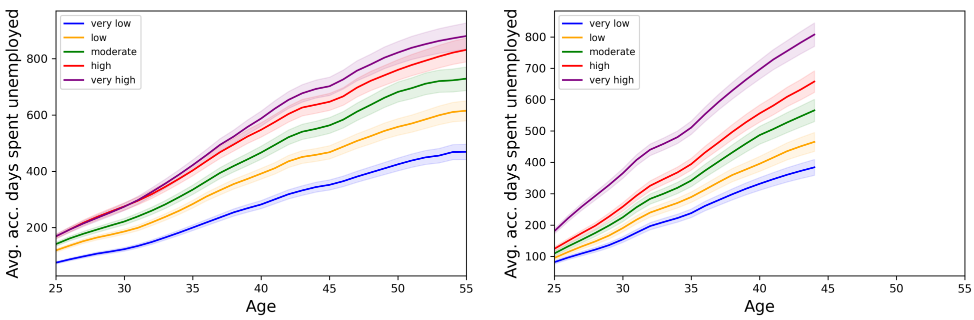
Figure 1.10 shows the result of combining the results presented in Figure 1.8 and Figure 1.9 by showing how much time in unemployment workers accumulate on average over the course of their careers by age. As expected, the plots for the different ranks appear in the expected order as workers with higher unemployment

Figure 1.9. Average length of unemployment spells



Notes: The averages presented in this figure represent the average number of days, of unemployment spells for workers aged 25-55. Workers are ranked based on the unemployment quintile of the commuting zone where they initially enter the labor market. The confidence intervals displayed are standard 95% confidence intervals.

Figure 1.10. Average accumulated time spent in unemployment



(a) 1960-1962 Cohort

(b) 1971-1973 Cohort

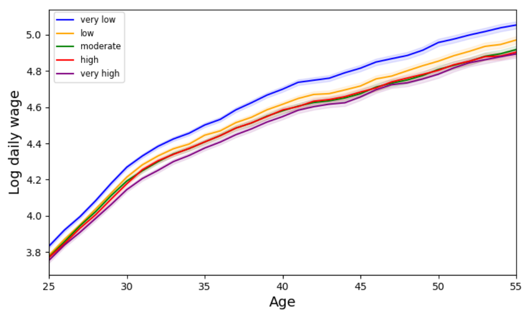
Notes: The average number of days that workers have accumulated up to each age. Workers who have no observations after an age younger than 55 are assumed to accumulate no further time in unemployment. Workers are ranked based on the unemployment quintile of the commuting zone where they initially enter the labor market. The confidence intervals displayed are standard 95% confidence intervals around the mean at each age. Results for the 1960-1962 cohort are on the left, and results for the 1971-1973 cohort are on the right.

ranks lose their jobs more often and have longer unemployment spells on average. At any age workers with higher unemployment ranks accumulate more time in unemployment. At the end of the observable lifecycle workers with the highest unemployment rank accumulate more than a year more time in unemployment on average than their peers with the lowest unemployment rank.

1.5.2 Wages

Having documented that workers who enter the labor markets in high-unemployment regions spend more time in unemployment, I now turn to wages and ask if workers in depressed local labor markets face lower wages than their peers as well.

Figure 1.11. Log-wage lifecycle profiles of workers born 1960-1962

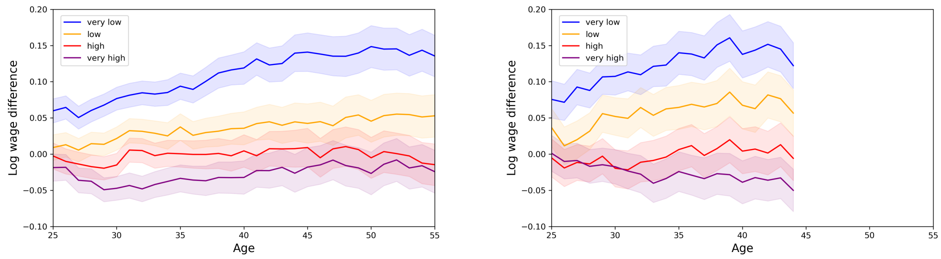


Notes: Workers are ranked based on the unemployment quintile of the commuting zone where they initially enter the labor market. Confidence bands represent standard 95% confidence intervals around the mean log-wage at each age. The sample is restricted to workers who work in full-time employment at each age.

Figure 1.11 shows log wage profiles for the 1960-1962 cohort. Since I only observe daily wages and cannot observe the number of hours a part-time worker works, I restrict the sample to full-time employed workers to ensure comparability. I use the *entry ranks* described in Section 1.3.2 to rank workers. While the shown log wage profiles all exhibit the familiar hump shape associated with log-wage lifecycle profiles, it is difficult to see the difference between the profiles of the different ranks. For this reason, in the remainder of the paper, I use the log wage profile of the *3 moderate* rank as a benchmark and subtract it from the other four profiles when comparing wages over the lifecycle. Figure 1.12 shows the result for the 1963-1965 and 1971-1973 cohorts.

For both cohorts, there is a clear difference in levels. At age 25, there is a difference of about 7 log points (or about 7%) for the 1960-1962 cohort and a difference of about 12 log points for the 1971-1973 cohort. A similar difference of about 10% exists for all the other cohorts as well.

Figure 1.12. Log-wage lifecycle profiles relative to rank *3 moderate*



(a) 1960-1962 Cohort

(b) 1971-1973 Cohort

Notes: Average log-wage profiles of workers by unemployment rank. Workers are ranked based on the unemployment quintile of the commuting zone where they initially enter the labor market. From each of the shown log wage lifecycle profiles, the profile of the *3 moderate* rank has been subtracted. Confidence intervals are constructed at the 95% level, at each age point. The left side of the figure represents birth years 1960-1962, while the right side represents birth years 1971-1973.

A second observation is that for most cohorts not only a difference in level exists, but that a difference in growth exists as well. For the 1971-1973 cohort, this is very pronounced as the wage gap opens up from about 7% at 25 to about 17% at 44. The wage gap for the 1960-1962 cohort opens up to about 15%. However, for all remaining cohorts, the growth level is also visible, and the wage gap opens up over the lifecycle (the remaining plots are in Figure 1.C.9 in the appendix).

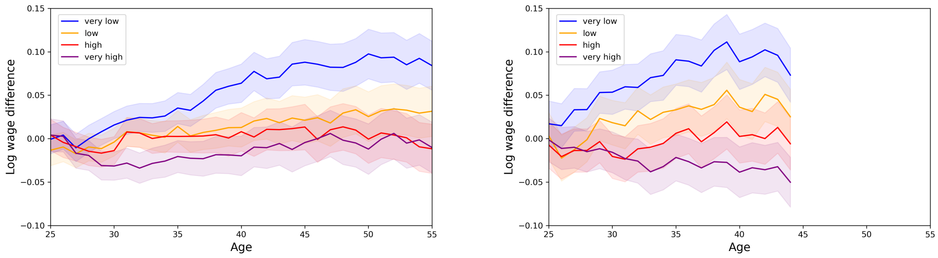
Given that I neither condition the sample on workers who stay in the same labor market nor on workers who are directly affected by unemployment, this relationship between local unemployment at labor market entry and the development of lifecycle profiles is striking. Not only do workers spend a lot of time in labor markets with a similar relative degree of unemployment, but this is clearly reflected in the development of wages over the lifecycle as well, stressing the important role of the local labor markets where workers start their careers.

Even though the wage gaps shown in Figure 1.12 are sizable, as discussed in Section 1.4, workers do not seem to systematically leave bad labor markets and seek out the commuting zones with lower unemployment and higher average wages. One way of explaining this could be that, while nominal wages are higher in low-unemployment labor markets, so are local prices. Indeed, some low-unemployment labor markets in Germany, such as Munich, are infamous for high prices and rents.

To quantify how much of the result is driven by these price differences, I reproduce the same figure using the regional CPI from the *BBSR* described in Section 1.2. Figure 1.13 shows the result.

Clearly, a good part of the differences is absorbed by the CPI correction. The 1960-1962 cohort wage profiles exhibit no difference at 25, and the gap opens up to only 10% at 55. The wage gaps in the 1971-1973 cohort, too, shrink to close to

Figure 1.13. Regional CPI-adjusted log-wage lifecycle profiles relative to rank *3 moderate*



(a) 1960-1962 Cohort *(b) 1971-1973 Cohort*

Notes: Average regional CPI-adjusted log-wage profiles of workers by unemployment rank. The CPI adjustment was calculated using the regional CPI from the BBSR. Workers are ranked based on the unemployment quintile of the commuting zone where they initially enter the labor market. From each of the shown log wage lifecycle profiles, the profile of the *3 moderate* rank has been subtracted. Confidence intervals are constructed at the 95% level, at each age point. The left side of the figure represents birth years 1963-1965, while the right side represents birth years 1971-1973.

indistinguishable instead of 10% at 25, opening up to about 15% instead of about 20% at 43. Nevertheless, a good part of the wage gaps persist and cannot be explained fully by the local price level, and for the younger cohorts in particular, the wage gap remains quite substantial.

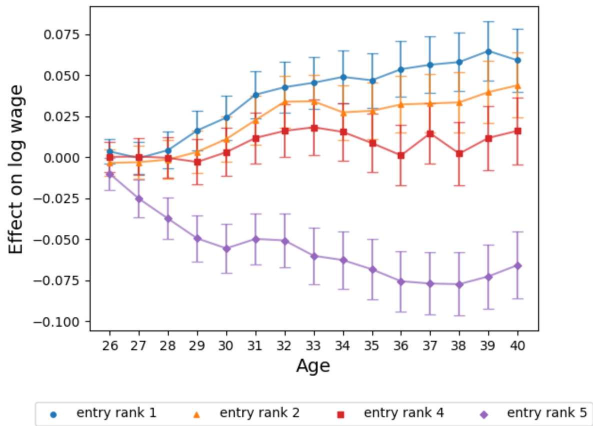
An explanation for spatial wage gaps that is common in the literature is worker and firm sorting, where wage differences are attributed to the spatial concentration of productive firms or productive workers. To demonstrate that plain sorting of worker types is not the driver of the observed wage gaps, I make further use of my ability to track workers along their careers in the SIAB and estimate effects on wages of having started in the different quintile ranks, conditional on worker fixed effects. I pool the data of all cohorts of age 25-40 together and estimate

$$\log \text{wage}_{ia} = \sum_{\substack{k=1,2,4,5 \\ j=26,\ldots,40}} \mathbb{1}\{a = j\}\mathbb{1}\{\textit{entry rank} = k\}\beta_{jk} + a\gamma_1 + a^2\gamma_2 + \alpha_i + e_{ia} \quad (1.4)$$

where α_i is a worker fixed effect for worker i, $\beta_{aq(i)}$ is the effect on the log-wage of being a years old and having first entered the labor market in a commuting zone of rank $q(i)$. Since I can only observe daily and not hourly wages, I exclude all workers who ever work part-time.[9] In order to estimate differential wage growth patterns, I control for age and age squared. I take age 25 and the middle rank (*3 moderate*) as benchmarks to avoid multicollinearity. Wages are adjusted using the regional CPI

9. I do not keep workers who work part-time only for some time because having worked part-time may have effects on full-time wages later in a career.

Figure 1.14. Fixed effect regression results



Notes: Estimates for $\{\beta_{a1}\}_{a=25,...,40}$ from Equation 1.4 in blue and $\{\beta_{a5}\}_{a=25,...,40}$ in purple. They represent wage effects of age, having started a career in the ranks *1 very low* and *5 very high* respectively, conditional on worker fixed effects. Wages are adjusted using the regional CPI from the *BBSR* before estimating the regression. Confidence intervals are at the 95% confidence level. Standard errors are clustered at the worker level.

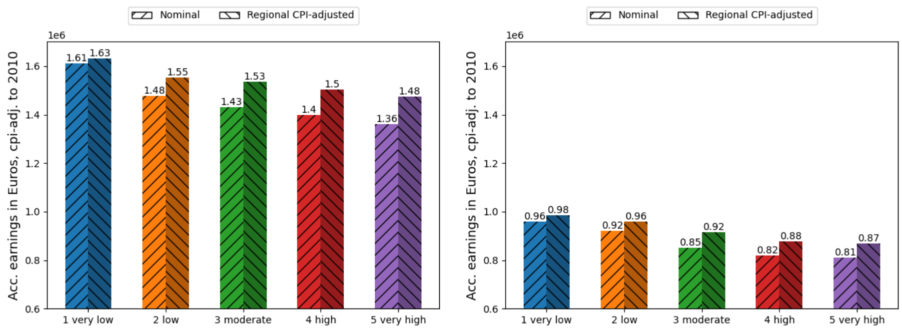
from the *BBSR*. The resulting fixed effects $\{\beta_{aq(i)}\}_{a=25,...,40}$ for the extreme ranks *1 very low* and *5 very high* are plotted in Figure 1.14.

The pattern of differential wage growth from Figures 1.12 and 1.13 is preserved. At the beginning of the career at 26, there is no difference in effects on wages for the two entry ranks. At age 40, however, this gap in wage effects has opened up to about 11%. The order of magnitude of this gap is similar to the one exhibited by the raw regional CPI adjusted shown in Figure 1.13. An important difference is that wage effects for the top four ranks are very similar, whereas having started a career in the worst rank leads to increasingly less wage growth, compared to the other ranks. Nevertheless, at least from the worker side, the selection of types does not explain the wage gap between the *5 very high* and *1 very low* unemployment regions.

1.5.3 Accumulated Earnings

Section 1.5.1 documented that workers with higher *entry ranks* accumulate more time in unemployment, and in Section 1.5.2, I documented that workers with higher *entry ranks* earn lower average wages. Both have implications for lifetime earnings. Lower wages directly lead to lower lifetime earnings. Time in unemployment leads to additional wage losses since the unemployment benefits workers receive are lower than the wage they would have earned in employment. On top of that, it is well documented that time in unemployment leads to reduced wages after

Figure 1.15. Average accumulated earnings by rank



Notes: Accumulated lifetime earnings by unemployment rank of the commuting zone workers started their first job in. The left side of the figure represents birth years 1960-1962, while the right side represents birth years 1971-1973. Earnings are accumulated up to the age of 55 in the left graph and up to the age of 44 on the right. The sample has been restricted to workers who are observable for at least 90% of the time between the age of 25 and the last observable age in each cohort. The price level has been adjusted to the level of 2010 using the CPI from the Federal Statistics Office. The darker bars have additionally been adjusted to the price level of Bonn using the regional CPI from the BBSR.

finding a new job - e.g., through less human capital accumulation or hampered further job ladder climbing. [10] More accumulation of unemployment, therefore, lowers lifetime earnings through both channels. In this section, leveraging the fact that I can observe daily wages in the SIAB, I calculate accumulated earnings over the life-cycle depending on the *entry ranks*.

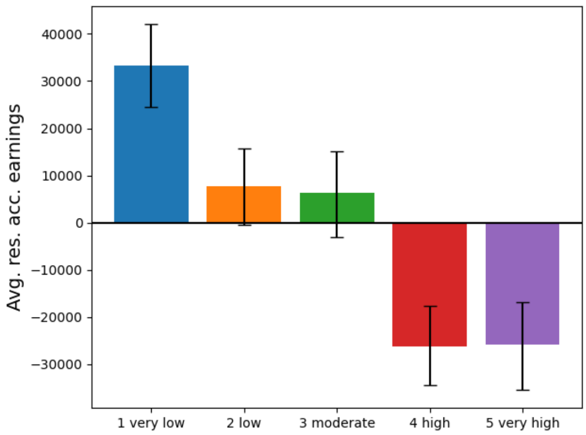
Figure 1.15 shows accumulated lifetime earnings in 2010 € by *entry rank*, both unadjusted and corrected for regional price differences using the regional CPI. To make wage levels comparable over cohorts, I CPI-adjust all wages to the level of 2010, using the CPI from the federal statistics office. Since I can only observe workers up to 2017, only the 1960-1962 cohort can be observed for close to a full lifecycle (up to age 55). Results for the other cohorts are therefore only partial lifetime earnings, and I will focus on the first cohort in this section. In order to make sure no results are driven by workers dropping out of the sample[11], I restrict the sample to workers who are unobservable for no longer than 10% of time between ages 25 and 55 (or the cohort's last observable age).

The gap between the highest and the lowest unemployment *entry rank* for the oldest cohort amounts to €250,000 or about 18% higher accumulated earnings.

10. E.g. Jacobson, LaLonde, and Sullivan (1993) and Lachowska, Mas, and Woodbury (2020).

11. This can be due to a number of reasons. Workers may genuinely leave the labor force, but they may also become civil servants, soldiers, become self-employed, or take up a job abroad. Additionally, the data before 1997 has no information on workers under long-term unemployment who receive ALG II unemployment benefits. Unfortunately, there is no way of telling any of these gaps apart, and gaps without any information on earnings, mechanically lower accumulated earnings over the lifecycle.

Figure 1.16. Average residualized accumulated earnings



Notes: Average residualized accumulated earnings until the age of 55 by unemployment rank of the commuting zone workers started their first job in. Residualization is achieved by projecting accumulated earnings on the specification represented by Equation 1.5. The figure shows results for the birth years 1960-1962.

When correcting for region prices, the difference shrinks to about €150,000 in Bonn prices or a difference of about 10%. To put these numbers into perspective, the average price for a new car in Germany in 2010 was about €26,000. These large numbers emphasize the long-term consequences of a career in a depressed labor market. The same gap for the younger 1971-1972 cohort (shown on the right of Figure 1.15) in accumulated earnings is about €150,000 in nominal terms or about €90,000 in Bonn prices at age 44. Since the wage gaps between the unemployment ranks open up with age, as documented in Section 1.5.2, this difference will likely become even larger until workers in this cohort retire.

To determine how much of these gaps in accumulated earnings are individual effects, it would be ideal to take out worker fixed effects. This is not feasible, however, since the unemployment rank a worker enters the labor market in and worker fixed effects are collinear. Instead, I project the accumulated, regional-CPI-adjusted earnings reported in Figure 1.15 on observables. Specifically, I estimate

$$w_i = X_i\beta + e_i, \tag{1.5}$$

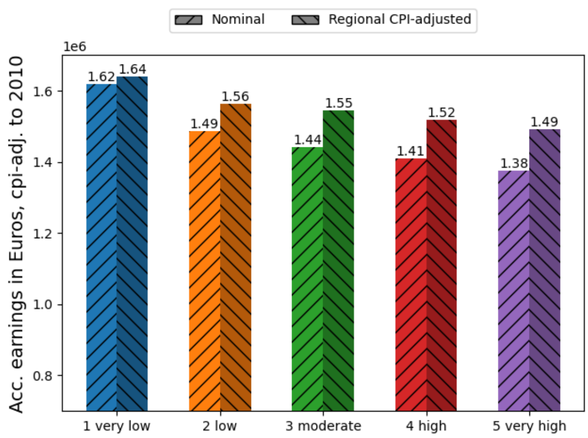
where w_i denotes accumulated wages and X_i includes controls for education, industry, occupation, and urbanization. I use the highest educational attainment ever reported for each individual i and the occupation each individual i spent the most time working in. The SIAB reports 120 levels of occupations based on the German KldB 3-digit classification and 8 levels of industries based on the WZ08 classification. urbanization$_i$ is a 3-level index that classifies the worker i's first-ever commuting zone into urban, rural, and in-between. It is taken from the *BBSR*. Figure 1.16

shows average residuals from this regression by unemployment rank for the 1960-1962 cohort.

The difference in the 1960-1962 cohort between the highest and the lowest rank is about €50,000. Although the difference between the residualized accumulated earnings is much smaller than between the raw accumulated earnings, a sizeable part of lifetime income is not absorbed by the rich observables I project on. This emphasizes the predictive power of initial local labor market conditions for the career outcomes of workers even further and raises the question of why workers do not move away from bad local labor markets.

The results shown in Figure 1.15 are based solely on labor income. However, unemployed workers receive unemployment insurance benefits. Given the persistence of local unemployment, unemployment insurance benefits redistribute between regions as every employed worker pays contributions, but only unemployed workers receive benefits. In the remainder of this section, I show that this is the case to a measurable degree, but not enough to change the situation of workers in high-unemployment regions meaningfully.

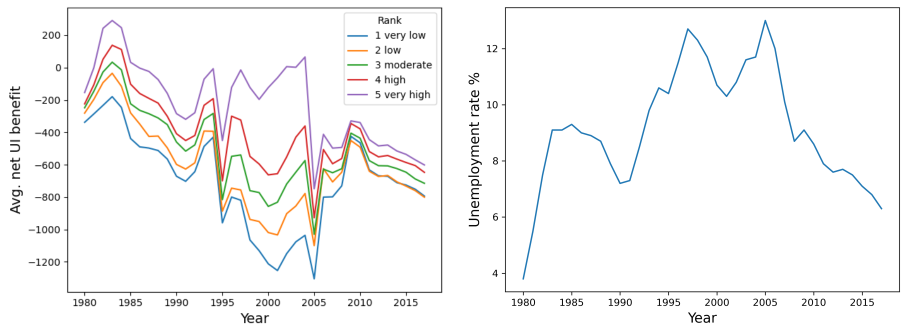
Figure 1.17. Average accumulated earnings and benefits by rank



Notes: Accumulated lifetime earnings by unemployment rank of the commuting zone workers started their first job in for the 1960-1962 cohort. Earnings are accumulated up to the age of 55 in the left graph and up to the age of 44 on the right. The sample has been restricted to workers who are observable for at least 90% of the time between the age of 25 and the last observable age in each cohort. The price level has been adjusted to the level of 2010 using the CPI from the Federal Statistics Office. The darker bars have additionally been adjusted to the price level of Bonn using the regional CPI from the BBSR. The numbers include (short-term) unemployment benefits.

Figure 1.17 shows accumulated earnings of the 1960-1962 cohort when unemployment benefits are included. The difference between the highest and lowest unemployment rank in accumulated earnings is hardly different when unemployment

Figure 1.18. Net unemployment benefits by unemployment ranks over time



Notes: The plot on the left shows contributions to social security subtracted from the sum of benefits as observable in the SIAB by unemployment rank. The plot on the right shows the overall unemployment rate in West Germany for comparison.

benefits are included. A likely driver is that wages in low-unemployment commuting zones are higher, as demonstrated in Section 1.5.2, and therefore so are short-term benefits, which replace a fixed part of a worker's last wage.

A caveat is that only short-term unemployment benefits are observable in the SIAB. In Section 1.B in the appendix, I describe a procedure to impute missing unemployment benefits and show that it does not meaningfully change the difference in accumulated income between the unemployment ranks either.

Nonetheless, the regional nature of unemployment makes unemployment insurance redistributive between regions on a small scale. Figure 1.18 plots net benefits (i.e., contributions to social security subtracted from the sum of benefits as observable in the SIAB) and the overall unemployment rate in West Germany. As is evident, particularly in high recessions, net benefits paid in high unemployment regions are higher[12]. This is of interest from a policy perspective in particular because local unemployment is so persistent and the regions that benefit the most from unemployment insurance are the same regions for extensive time periods.

1.6 Conclusion

Starting from the well-known fact that there is sizable and persistent heterogeneity in local unemployment, I documented that workers are not very mobile between good and bad local labor markets. Instead, I find that, driven by a lack of mobility over long distances, the degree of relative unemployment in a worker's initial

12. Social security contribution rates are not fixed over time, so the numbers in Figure 1.18 depend on these rates as well. I show these rates together with more details about the unemployment insurance system in Germany in Section 1.B in the appendix.

commuting zone is a very good predictor of the degree of local unemployment the worker will face throughout his working life.

I first document that workers who start in worse labor markets also accumulate more time in unemployment, losing their jobs more often, and finding new jobs more slowly on average.

I then proceed to demonstrate that workers earn different wages, depending on whether they enter the labor market in a good or a bad local labor market. I find wage gaps of about 7-10% at age 25 that open up to about 20% toward the end of the observable lifecycle between workers who started in the highest and workers who started in the lowest local unemployment quintile. I demonstrate that adjusting for the local price level and controlling for worker fixed effects does not do away with these differences in wage growth. At age 40, a difference in wage effects in the order of magnitude of 11% remains.

Controlling for local prices, until the age of 55 there is a difference in lifetime earnings of about €150,000 in terms of regionally adjusted 2010-prices. Residualizing lifetime earnings with rich observables, a gap of about €50,000 remains.

I conclude that it does matter considerably where a worker starts his working life. The natural way forward for future research would be to establish the mechanism that drives my findings. I view this as a promising avenue for future research.

1. Appendix

1.A Data

1.A.1 Wage Imputation

Employers in Germany report employers' wages only up to the contribution limit to social security accounts. The wage variable is therefore right-censored. To circumvent this problem I follow the procedure recommended by Gartner (2005) and impute wages in the following way: For every year I estimate the model $\log \text{wage}_{it} = x_{it}'\beta + e_{it}$, where x_{it} is a vector containing *imputed education level, labor market status* (full-time employed, part-time employed, marginally employed or apprentice), *the fraction of spells where an individuals wage is censored in other years* and the *mean reported wage from other years* for individual i at time t using TOBIT estimation.

In order not to underestimate the variance of wages above the censoring level I then use the predicted wage $x_{it}'\hat{\beta}_{\text{tobit}}$ and the estimate of the standard deviation $\hat{\sigma}_{\text{tobit}}$ to draw the imputed wage from the truncated normal distribution with mean $x_{it}'\hat{\beta}_{\text{tobit}}$ and standard deviation $\hat{\sigma}_{\text{tobit}}$ truncated to the left at the censoring level.

As the results presented below are mainly based on samples that are restricted to men I impute wages for men and women separately to allow for different wage variances for men and women.

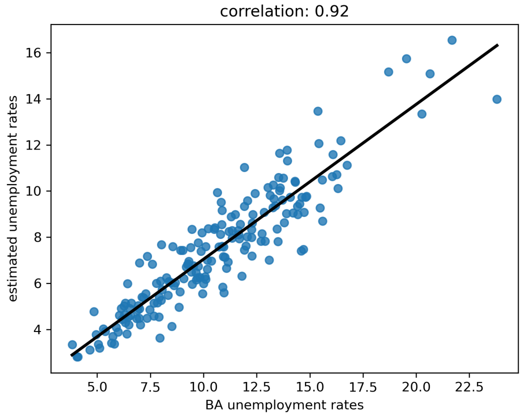
1.A.2 Education Imputation

The BeH data is taken from reports that employers make about their employees. While this official nature of the data source lends it a great deal of its credibility there is room for interpretation as to what exactly should be reported that may be handled differently by different employers. The education variable in particular suffers from inconsistencies such as reporting lower levels of education after higher levels have been reported for the same individual in previous years. One reason for this may be that employers tend to report only the minimum education level required for a particular job. To deal with these inconsistencies I employ the Education Imputation Procedure 1 proposed by Fitzenberger, Osikominu, and Völter (2005) and extrapolate reported education such that I am left with 6 consistent education variables: *1 No Degree, 2 High School, 3 Vocational Training, 4 High School and Vocational Training, 5 Technical College* and *6 University*.

1.A.3 Spell Error Imputation

In some cases, there are gaps in the reported unemployment history of an individual. Because employers report information about employees in continuous ongoing employment once yearly I follow Böhm, Gaudecker, and Schran (2023 forthcom-

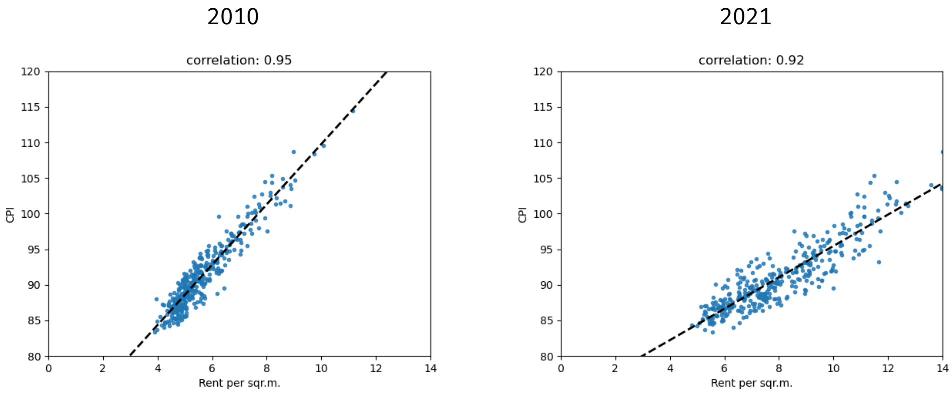
Figure 1.A.1. SIAB vs BA unemployment rates (1985)



Notes: BA unemployment rates or on the x-axis and unemployment rates estimated from the SIAB are on the y-axis. Each blue dot represents a commuting zone.

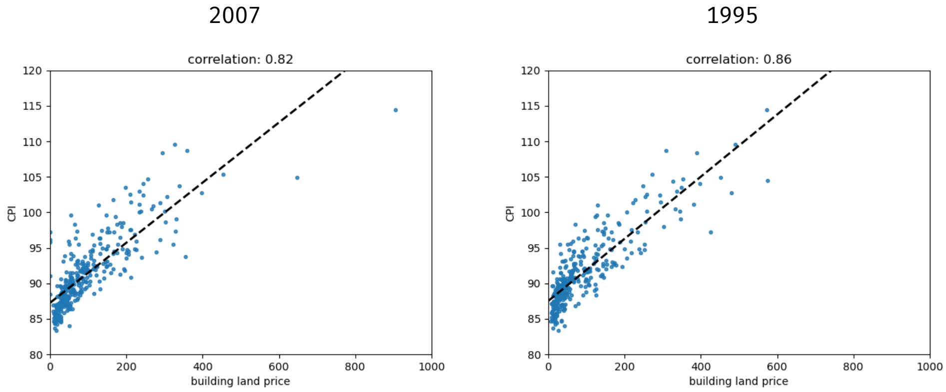
ing) and impute such gaps if there is a one-year gap between two spells with the same employer by filling it with the last valid spell. Other gaps are interpreted as spells during which an individual is not part of the labor force.

Figure 1.A.2. Regional CPI vs INKAR rent-index



Notes: The INKAR rent-index is provided by the BBSR on the INKAR wbesite. It is based on rents of flats of 40-100 m^2 in average or good neighborhoods advertised on a variety of platforms. It is meant to reflect offers that people who look for a flat online in a given county are typically faced with. It is reported in €1 bins.

Figure 1.A.3. Regional CPI vs building land prices



Notes: Building land prices are average prices per square-meter of sold building land within a county. The data comes from Regionaldatenbank Deutschland - the regional data base of the federal and state level statistical offices of Germany.

1.B Unemployment Benefit Imputation

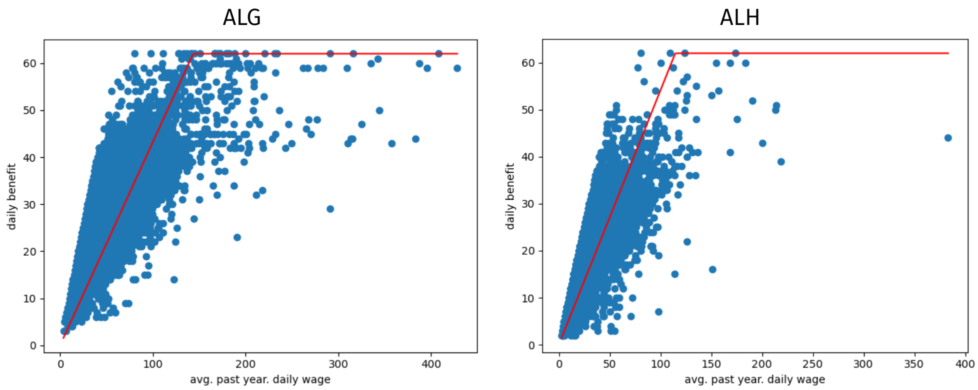
The German unemployment benefit system works as follows: Workers who were employed for at least 24 month with the last 5 years before entering unemployment receive 60% of their after tax wages up to the maximum social security contribution level for up to two years. This means that the benefit depends on the tax class and other particularities of the German tax system that are impossible to reconstruct for each individual worker because they in turn depend on unobserved characteristics such as e.g. marital status and the number of children. I therefore impute these benefits in the following way: Each year, I use the sample of workers for whom I can observe unemployment benefits and run a TOBIT regression, regressing the observed benefit on the last observed wage. I then use the fitted TOBIT model to predict the benefit for unemployed workers with unobservable unemployment benefit. Benefits paid under this scheme are called *Arbeitslosengeld (ALG)*.

Until 2005, as a fallback, workers who were employed for at least 12 months within the past 30 months before entering unemployment received lower benefits called *Arbeitslosenhilfe (ALH)*. ALH was based on a similar system as ALG and I impute it in the same way for workers without observable benefits who are eligible for ALH bit not for ALG. Figure 1.B.1 shows an illustration of the TOBIT predictions.

Until 2005 Long-term unemployment benefits were part of social benefits regulated and paid by local governments. I am unaware of a feasible imputation strategy for these payments. Since 2005 there are nation-wide fixed long-term unemployment benefits. I simply use the official benefit figures for long-term unemployment workers reported through the ASU.

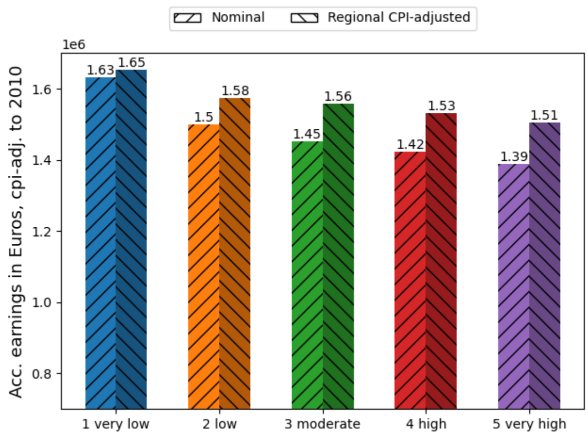
Adding these imputations to the workers observable income results in slightly higher accumulated income than shown in Figure 1.17. However, the difference

Figure 1.B.1. Tobit Prediction of Unemployment Benefit 2000



Notes: The blue dots plot actually observed benefits on the y-axis against the respective worker's last wage on the x-axis. The red line shows the fitted TOBIT model that I use to impute benefits for unemployed workers without observable benefits.

Figure 1.B.2. Average accumulated earnings and benefits by rank



Notes: Accumulated life-time earnings by unemployment rank of the commuting zone workers started their first job in for the 1960-1962 cohort. Earnings are accumulated up to the age of 55 for in the left graph and up to the age of 44 on the right. The sample has been restricted to workers who are observable for at least 90% of the time between the age of 25 and the last observable age in each cohort. The price level has been adjusted to the level of 2010 using the CPI from the Federal Statistics Office. The darker bars have additionally been adjusted to the price level of Bonn using the regional CPI from the BBSR. The numbers include imputed unemployment benefits.

between the high and low unemployment ranks doesn't change much. Results for the 1960-1962 cohort are shown in Figure 1.B.2.

Unemployment insurance is financed using social security contributions. Employed workers pay a share of their wage to social security. These contribution rates for the years covered by the SIAB are given in Table 1.B.1.

Table 1.B.1. Social Security Contribution Rates

Year	Rate	Year	Rate	Year	Rate	Year	Rate
1975	0.020	1986	0.041	1997	0.065	2008	0.038
1976	0.020	1987	0.041	1998	0.065	2009	0.033
1977	0.020	1988	0.041	1999	0.065	2010	0.028
1978	0.020	1989	0.041	2000	0.065	2011	0.028
1979	0.020	1990	0.043	2001	0.065	2012	0.030
1980	0.030	1991	0.043	2002	0.065	2013	0.030
1981	0.030	1992	0.043	2003	0.065	2014	0.030
1982	0.030	1993	0.043	2004	0.065	2015	0.030
1983	0.030	1994	0.043	2005	0.065	2016	0.030
1984	0.030	1995	0.065	2006	0.065	2017	0.030
1985	0.041	1996	0.065	2007	0.042		

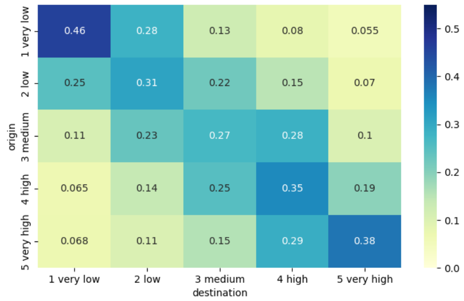
1.C Additional Figures and Tables

In the main text, in many cases, I only show results for two cohorts. Since the results are very similar across cohorts, I present the remaining cohort-specific results in this appendix.

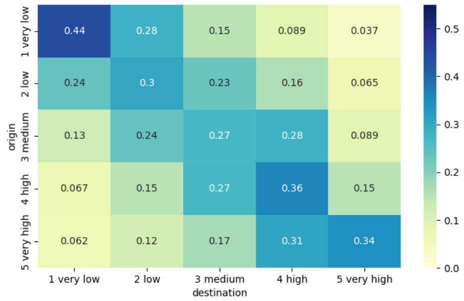
Figure 1.C.1 shows the fractions of moves by unemployment rank of origin and destination commuting zones across the additional cohorts. Figure 1.C.2 reports the same patterns for moves over at least 200 kilometers. Figures 1.C.3 and 1.C.4 show the corresponding transition matrices separately for workers moving after unemployment and after employment. Figure 1.C.5 shows how much time workers from each initial unemployment rank spend in commuting zones of the different ranks over their working lives. Figure 1.C.6 presents the same evidence for workers who entered the labor market as apprentices before age 20. Figure 1.C.7 shows accumulated job losses. Figure 1.C.8 shows accumulated time spent in unemployment. Figure 1.C.9 shows wage profiles relative to the *3 moderate* rank for the additional cohorts. Figure 1.C.10 shows the same wage profiles with the additional regional CPI adjustment. Figure 1.C.11 shows accumulated earnings by rank for the additional cohorts. Since the data is only available up to 2017, the 1977-1979 cohort is omitted from this figure.

Table 1.C.1 shows that workers who can be tracked in the data already as apprentices start their first job close to where they were apprenticed. This suggests that the place of labor market entry that I use to rank workers in the main analysis is likely highly correlated with where workers grew up.

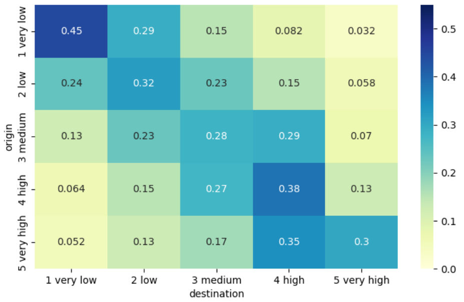
Figure 1.C.1. Fractions of moves by unemployment rank of origin- and destination commuting zones



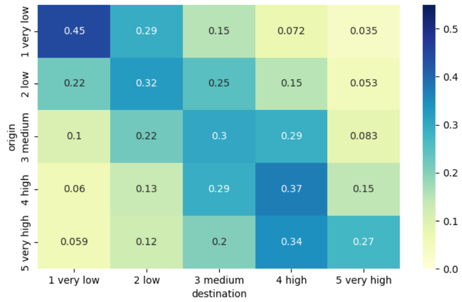
(a) 1960-1962 Cohort



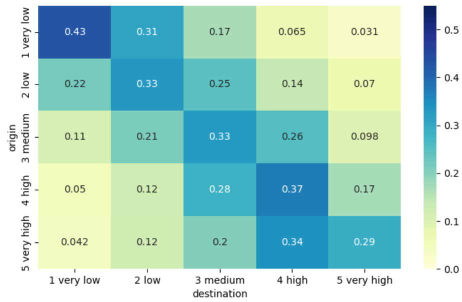
(b) 1966-1967 Cohort



(c) 1968-1970 Cohort



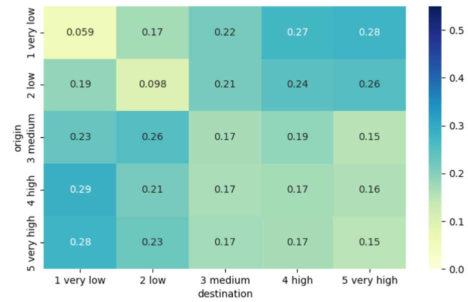
(d) 1974-1976 Cohort



(e) 1977-1979 Cohort

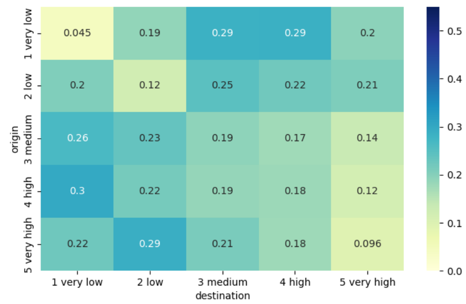
Notes: Each cell in the matrices above represents the fraction of moves away from commuting zones with an unemployment rank indicated by the row to a commuting zone with an unemployment rank indicated by the column. E.g., the second cell in the first row of the left matrix indicates that 29% of moves away from zones with rank *1 very low* went to zones with rank *2 low*. Workers are ranked based on the unemployment quintile of the commuting zone where they initially enter the labor market. Cohorts are arranged as follows: 1960-1962 - top left, 1966-1967 - top right, 1968-1970 - middle left, 1974-1976 - middle right, 1977-1979 bottom.

Figure 1.C.2. Fractions of moves over at least 200 km by unemployment rank of origin- and destination commuting zones

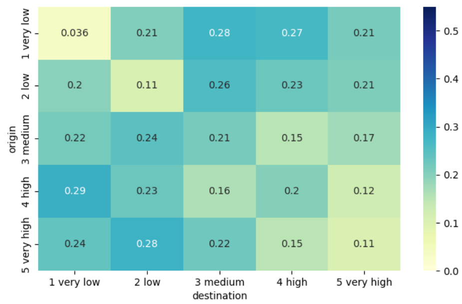


(a) 1960-1962 Cohort

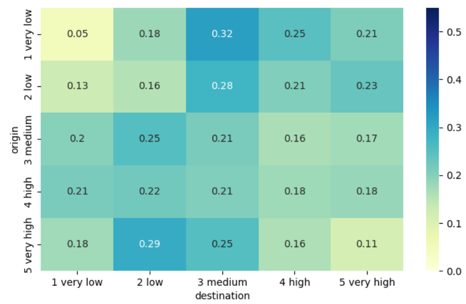
(b) 1966-1967 Cohort



(c) 1968-1970 Cohort



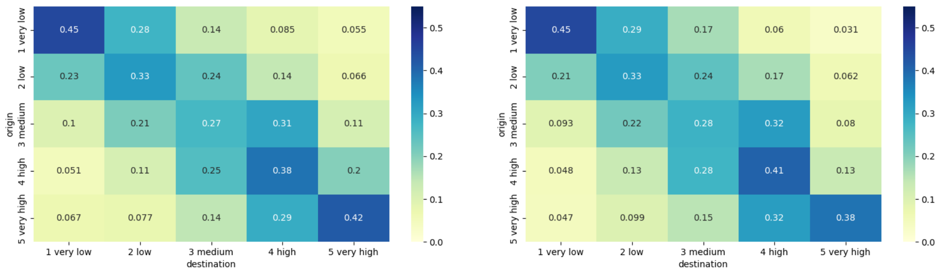
(d) 1974-1976 Cohort



(e) 1977-1979 Cohort

Notes: Each cell in the above matrices represents the fraction of moves away from commuting zones with an unemployment rank indicated by the row to a commuting zone with an unemployment rank indicated by the column. Only moves over a distance of at least 200 km are considered. Distance between two commuting zones is measured as the minimum distance between the borders of these commuting zones. E.g., the second cell in the first row of the left matrix indicates that 18% of moves away from zones with rank *1 very low* went to zones with rank *2 low*. Workers are ranked based on the unemployment quintile of the commuting zone where they initially enter the labor market. Cohorts are arranged as follows: 1960-1962 - top left, 1966-1967 - top right, 1968-1970 - middle left, 1974-1976 - middle right, 1977-1979 bottom.

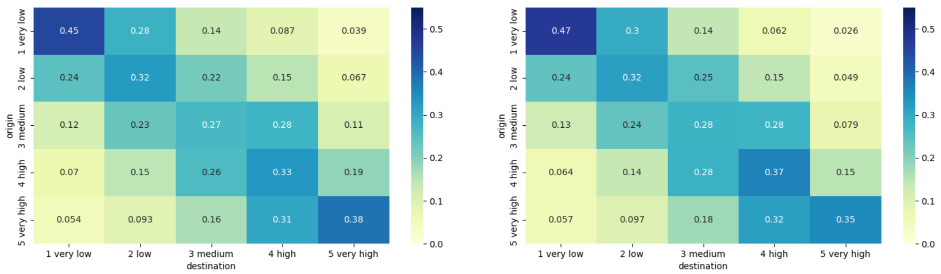
Figure 1.C.3. Fractions of moves by unemployment rank of origin- and destination commuting zones - unemployment workers



(a) 1963-1965 Cohort *(b) 1971-1973 Cohort*

Notes: Each cell in the matrices above represents the fraction of moves away from commuting zones with an unemployment rank indicated by the row to a commuting zone with an unemployment rank indicated by the column. Only moves that happen immediately after an unemployment spell are considered. E.g., the second cell in the first row of the left matrix indicates that 29% of moves away from zones with rank *1 very low* went to zones with rank *2 low*. Workers are ranked based on the unemployment quintile of the commuting zone where they initially enter the labor market. The matrix on the left shows the results for the 1963-1965 cohort, and the matrix on the right shows the results for the 1971-1973 cohort.

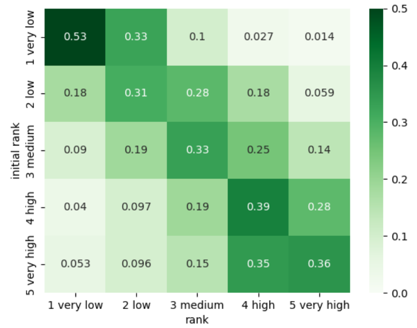
Figure 1.C.4. Fractions of moves by unemployment rank of origin- and destination commuting zones - unemployment workers



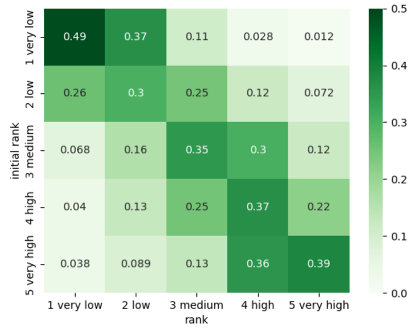
(a) 1963-1965 Cohort *(b) 1971-1973 Cohort*

Notes: Fractions of moves of workers who were employed before moving that occur over the whole life by the rank of origin and rank of destination. The matrix on the left shows the results for the 1963-1965 cohort, the matrix on the right shows the results for the 1971-1973 cohort. Each cell in the matrices above represents the fraction of moves away from commuting zones with an unemployment rank indicated by the row to a commuting zone with an unemployment rank indicated by the column. Only moves that happen immediately after a spell under employment are considered. E.g., the second cell in the first row of the left matrix indicates that 29% of moves away from zones with rank *1 very low* went to zones with rank *2 low*. Workers are ranked based on the unemployment quintile of the commuting zone where they initially enter the labor market. The matrix on the left shows the results for the 1963-1965 cohort, and the matrix on the right shows the results for the 1971-1973 cohort.

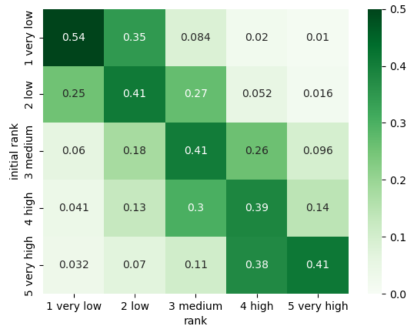
Figure 1.C.5. Fractions of time spent in the five unemployment ranks



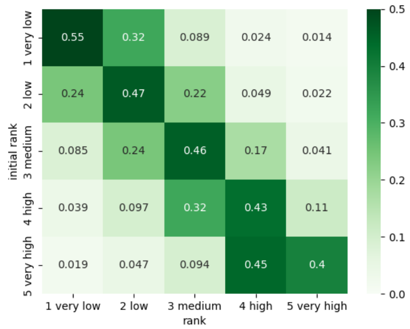
(a) 1960-1962 Cohort



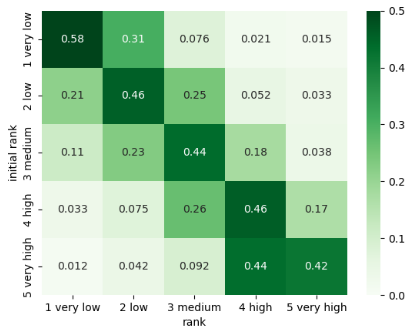
(b) 1966-1967 Cohort



(c) 1968-1970 Cohort



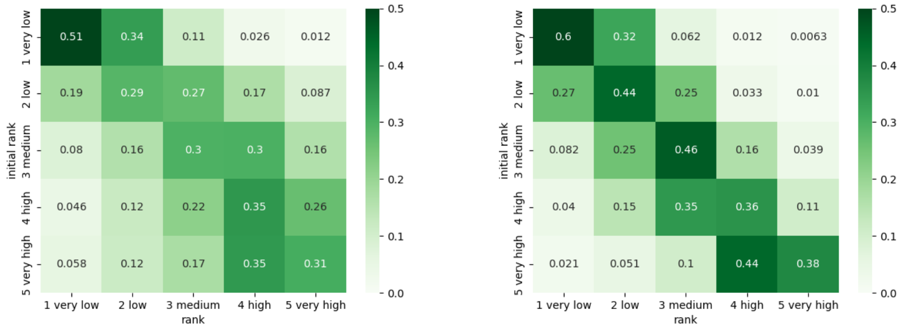
(d) 1974-1976 Cohort



(e) 1977-1979 Cohort

Notes: Fractions of time spent in commuting zones of each of the five ranks by the rank of the initial commuting zone of a worker where he first entered the labor market. The initial ranks are on the y-axis. For example, the second cell in the first row of the left graph states that workers who start in a commuting zone with rank *1 very low* spend 32% of their time in commuting zones with rank *2 low* on average. Workers are ranked based on the unemployment quintile of the commuting zone where they initially enter the labor market. Cohorts are arranged as follows: 1960-1962 - top left, 1966-1967 - top right, 1968-1970 - middle left, 1974-1976 - middle right, 1977-1979 bottom.

Figure 1.C.6. Fractions of time spent in the five unemployment ranks, apprentices only



(a) 1963-1965 Cohort *(b) 1971-1973 Cohort*

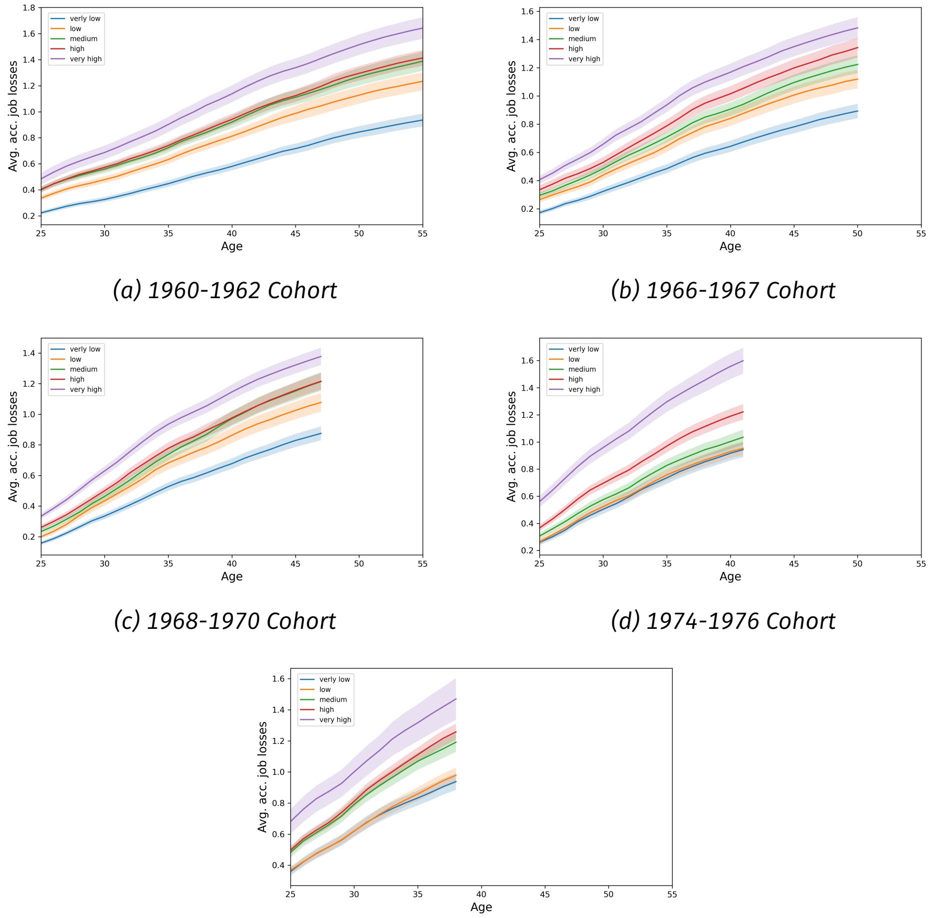
Notes: Fractions of time spent in commuting zones of each of the five ranks by rank of the initial commuting zone of a worker entered the labor market in. The sample is restricted to workers who enter the labor market as apprentices before the age of 20. The initial ranks are on the y-axis. For example, the second cell in the first row of the left graph states that workers who start in a commuting zone with rank *1 very low* spend 32% of their time in commuting zones with rank *2 low* on average. Workers are ranked based on the unemployment quintile of the commuting zone where they initially enter the labor market. The matrix on the left shows the results for the 1963-1965 cohort, and the matrix on the right shows the results for the 1971-1973 cohort.

Table 1.C.1. Fractions of workers who start a first job within a given distance from the commuting zone they were apprenticed in

cohort	stayers	neighboring zone	100 km	200km
1960-1962	0.886477	0.944741	0.980467	0.989316
1963-1965	0.871097	0.938801	0.978383	0.987640
1966-1967	0.863239	0.932684	0.975736	0.986004
1968-1970	0.866690	0.940537	0.980296	0.988598
1971-1973	0.857920	0.936378	0.977938	0.987726
1974-1976	0.851978	0.932327	0.974109	0.986430
1977-1979	0.845169	0.930024	0.976455	0.987192

Notes: Sample consists of workers who entered the labor market as apprentices before the age of 20. The first column shows the fraction of such workers who start their first non-apprentice-job in the same commuting zone that they were apprenticed in. The second column shows the fraction of workers that start their first non-apprentice job in the same or a neighboring commuting zone. The third and fourth columns show the fractions of workers that stay within a radius of 100 km and 200 km, respectively, from where they were apprenticed when they start their first non-apprentice-job. Distance between two commuting zones is measured as the minimum distance between the borders of these commuting zones.

Figure 1.C.7. Average accumulated job losses



(a) 1960-1962 Cohort

(b) 1966-1967 Cohort

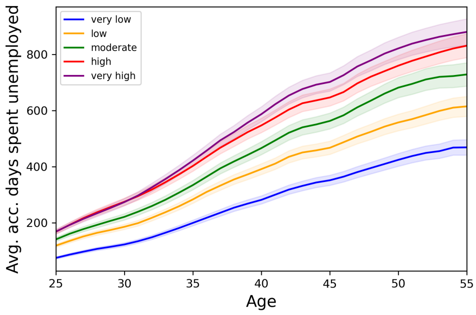
(c) 1968-1970 Cohort

(d) 1974-1976 Cohort

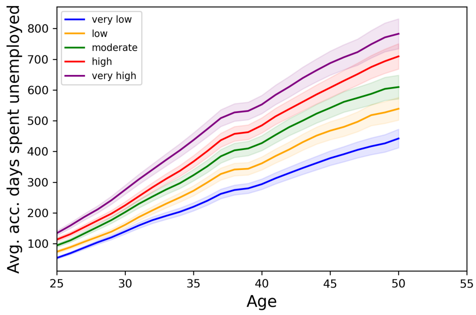
(e) 1977-1979 Cohort

Notes: The average number of job losses workers from each rank have had at each age. Workers are ranked based on the unemployment quintile of the commuting zone where they initially enter the labor market. Workers who have no observations after an age younger than 55 are assumed to lose no further jobs. Job losses are defined as transitions from an employed spell to an unemployed spell with a gap of fewer than 31 days between them. Cohorts are arranged as follows: 1960-1962 - top left, 1966-1967 - top right, 1968-1970 - middle left, 1974-1976 - middle right, 1977-1979 bottom.

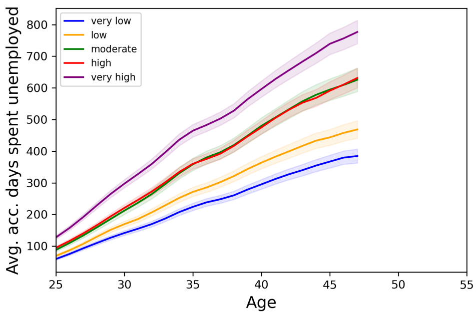
Figure 1.C.8. Average accumulated time spent in unemployment



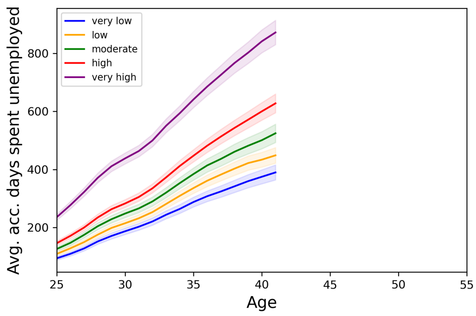
(a) 1960-1962 Cohort



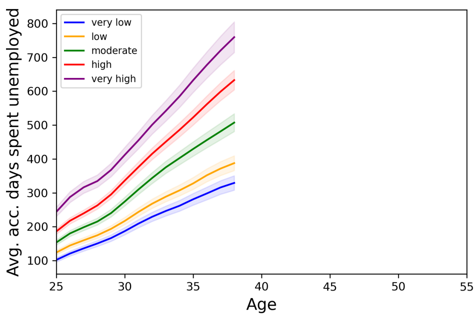
(b) 1966-1967 Cohort



(c) 1968-1970 Cohort



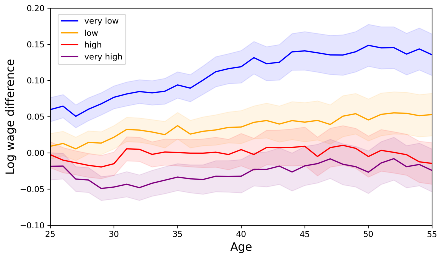
(d) 1974-1976 Cohort



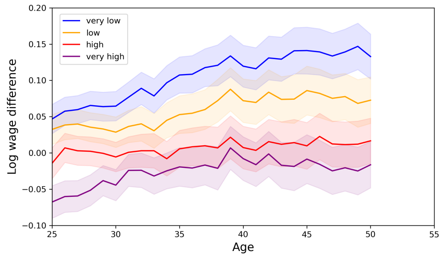
(e) 1977-1979 Cohort

Notes: The average number of days that workers have accumulated up to each age. Workers who have no observations after an age younger than 55 are assumed to accumulate no further time in unemployment. Workers are ranked based on the unemployment quintile of the commuting zone where they initially enter the labor market. The confidence intervals displayed are standard 95% confidence intervals around the mean at each age. Cohorts are arranged as follows: 1960-1962 - top left, 1966-1967 - top right, 1968-1970 - middle left, 1974-1976 - middle right, 1977-1979 bottom.

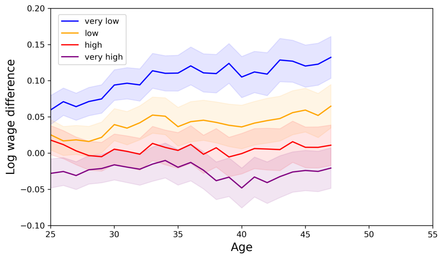
Figure 1.C.9. Average Log Wage lifecycle Profiles relative to rank *3 moderate*



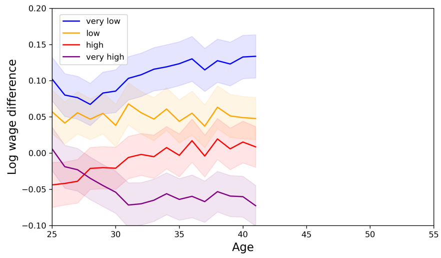
(a) 1960-1962 Cohort



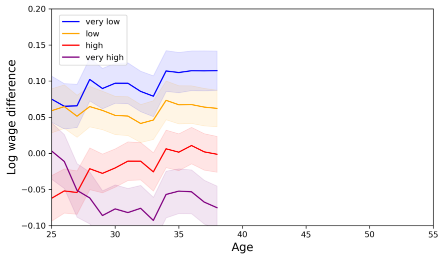
(b) 1966-1967 Cohort



(c) 1968-1970 Cohort



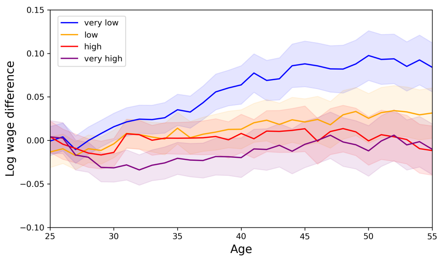
(d) 1974-1976 Cohort



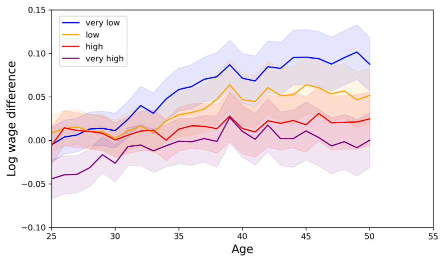
(e) 1977-1979 Cohort

Notes: Average log-wage profiles of workers by unemployment rank. Workers are ranked based on the unemployment quintile of the commuting zone where they initially enter the labor market. From each of the shown log wage lifecycle profiles, the profile of the *3 moderate* rank has been subtracted. Confidence intervals are computed using two-sided t-tests to assess whether the resulting difference is statistically significant from 0 at the 95% level at each age. Cohorts are arranged as follows: 1960-1962 - top left, 1966-1967 - top right, 1968-1970 - middle left, 1974-1976 - middle right, 1977-1979 bottom.

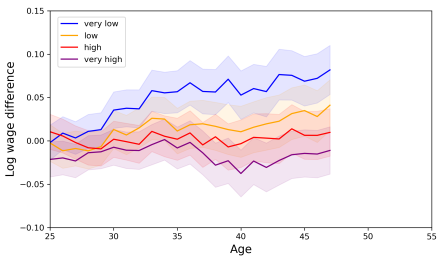
Figure 1.C.10. Regional CPI-adjusted Log Wage lifecycle Profiles relative to rank *3 moderate*



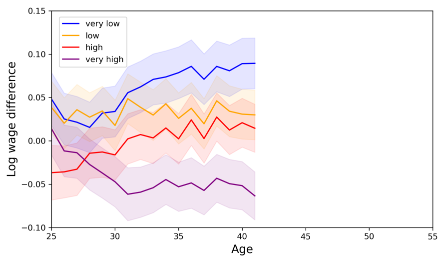
(a) 1960-1962 Cohort



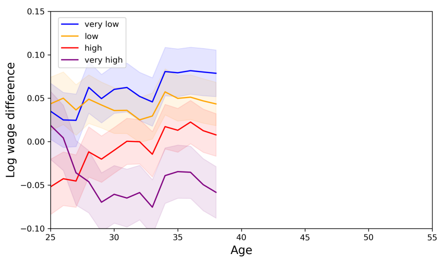
(b) 1966-1967 Cohort



(c) 1968-1970 Cohort



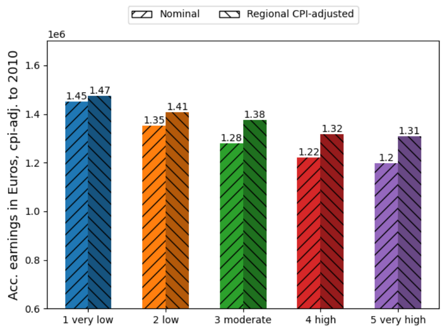
(d) 1974-1976 Cohort



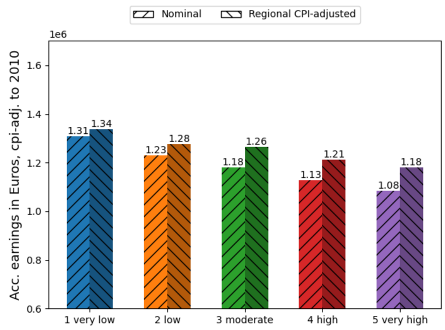
(e) 1977-1979 Cohort

Notes: Average regional CPI-adjusted log-wage profiles of workers by unemployment rank. The CPI adjustment was calculated using the regional CPI from the BBSR. Workers are ranked based on the unemployment quintile of the commuting zone where they initially enter the labor market. From each of the shown log wage lifecycle profiles, the profile of the *3 moderate* rank has been subtracted. Confidence intervals are computed using two-sided t-tests to assess whether the resulting difference is statistically significant from o at the 95% level at each age. Cohorts are arranged as follows: 1960-1962 - top left, 1966-1967 - top right, 1968-1970 - middle left, 1974-1976 - middle right, 1977-1979 bottom.

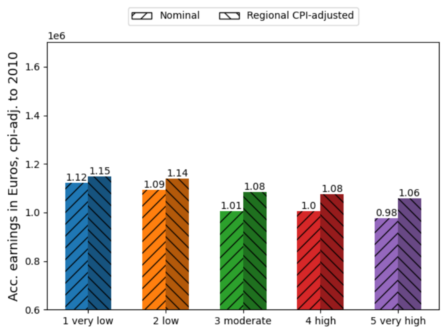
Figure 1.C.11. Average accumulated earnings



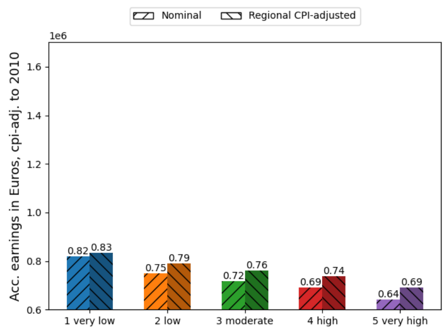
(a) 1963-1965 Cohort



(b) 1966-1967 Cohort



(c) 1968-1970 Cohort



(d) 1974-1976 Cohort

Notes: Average accumulated earnings up to the age of 40. The sample includes workers with gaps in their observable histories that comprise no more than 10% of the time between the ages of 25 and 55. Cohorts are arranged as follows: 1963-1965 - top left, 1966-1967 - top right, 1968-1970 - middle left, 1974-1976 - middle right.

References

Allen, Treb, and Costas Arkolakis. 2014. " Trade and the Topography of the Spatial Economy." *Quarterly Journal of Economics* 129 (3): 1085–140. [6]

Altonji, Joseph G, Lisa B Kahn, and Jamin D Speer. 2016. "Cashier or consultant? Entry labor market conditions, field of study, and career success." *Journal of Labor Economics* 34 (S1): S361–S401. [3]

Amior, Michael, and Alan Manning. 2018. "The Persistence of Local Joblessness." *American Economic Review* 108 (7): 1942–70. [6]

Baum-Snow, Nathaniel, and Ronni Pavan. 2012. "Understanding the city size wage gap." *Review of economic studies* 79 (1): 88–127. [6]

Berge, Wiljan van den. 2018. "Bad start, bad match? The early career effects of graduating in a recession for vocational and academic graduates." *Labour Economics* 53: 75–96. [3]

Bilal, Adrien. 2023. "The Geography of Unemployment." *Quarterly Journal of Economics* 138 (3): 1507–76. [3, 5]

Blanchard, Olivier Jean, Lawrence F Katz, Robert E Hall, and Barry Eichengreen. 1992. "Regional evolutions." *Brookings papers on economic activity* 1992 (1): 1–75. [5, 6]

Böhm, Michael J., Hans-Martin von Gaudecker, and Felix Schran. 2023 forthcoming. "Occupation Growth, Skill Prices, and Wage Inequality." *Journal of Labor Economics.* [30]

Bönke, Timm, Giacomo Corneo, and Holger Lüthen. 2015. "Lifetime Earnings Inequality in Germany." *Journal of Labor Economics* 33 (1): 171–208. [6]

Bryan, Gharad, and Melanie Morten. 2019. "The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia." *Journal of Political Economy* 127 (5): 2229–68. [6]

Card, David, Jörg Heining, and Patrick Kline. 2013. "Workplace Heterogeneity and the Rise of West German Wage Inequality." *Quarterly Journal of Economics* 128 (3): 967–1015. [6]

Dauth, Wolfgang, Sebastian Findeisen, Enrico Moretti, and Jens Suedekum. 2022. "Matching in cities." *Journal of the European Economic Association* 20 (4): 1478–521. [6]

Dustmann, Christian, Johannes Ludsteck, and Uta Schönberg. 2009. "Revisiting the German Wage Structure." *Quarterly Journal of Economics* 124 (2): 843–81. [6]

Fitzenberger, Bernd, Aderonke Osikominu, and Robert Völter. 2005. "Imputation rules to improve the education variable in the IAB employment subsample." *FDZ-Methodenreport* 03/2005. [30]

Gartner, Hermann. 2005. "The imputation of wages above the contribution limit with the German IAB employment sample." *FDZ-Methodenreport* 02/2005. [7, 30]

Gaubert, Cecile. 2018. "Firm sorting and agglomeration." *American Economic Review* 108 (11): 3117–53. [6]

Glaeser, Edward L, and David C Mare. 2001. "Cities and skills." *Journal of labor economics* 19 (2): 316–42. [6]

Gould, Eric D. 2007. "Cities, workers, and wages: A structural analysis of the urban wage premium." *Review of Economic Studies* 74 (2): 477–506. [6]

Harris, John R, and Michael P Todaro. 1970. "Migration, unemployment and development: a two-sector analysis." *American economic review,* 126–42. [6]

Hornbeck, Richard, and Enrico Moretti. 2022. "Estimating Who Benefits from Productivity Growth: Local and Distant Effects of City Productivity Growth on Wages, Rents, and Inequality." *Review of Economics and Statistics,* 1–49. [5]

Huttunen, Kristiina, Jarle Møen, and Kjell G Salvanes. 2018. "Job Loss and Regional Mobility." *Journal of Labor Economics* 36 (2): 479–509. [6]

Jacobson, Louis S, Robert J LaLonde, and Daniel G Sullivan. 1993. "Earnings losses of displaced workers." *American economic review,* 685–709. [25]

Kahn, Lisa B. 2010a. "The long-term labor market consequences of graduating from college in a bad economy." *Labour economics* 17 (2): 303–16. [12]

Kahn, Lisa B. 2010b. "The long-term labor market consequences of graduating from college in a bad economy." *Labour Economics* 17 (2): 303–16. [3]

Kline, Patrick, and Enrico Moretti. 2013. "Place based policies with unemployment." *American Economic Review* 103 (3): 238–43. [3, 5]

Kosfeld, Reinhold, Hans-Friedrich Eckey, and Martina Schüßler. 2009. "Ökonometrische Messung regionaler Preisniveaus auf der Basis örtlich beschränkter Erhebungen." RatSWD Research Notes 33. German Data Forum (RatSWD). [8]

Kuhn, Moritz, Iourii Manovskii, and Xincheng Qiu. 2026. "The Geography of job creation and job destruction." *Journal of Monetary Economics* 158: 103898. [3]

Lachowska, Marta, Alexandre Mas, and Stephen A Woodbury. 2020. "Sources of displaced workers' long-term earnings losses." *American Economic Review* 110 (10): 3231–66. [25]

Levy, Brian L., Ted Mouw, and Anthony Daniel Perez. 2017. "Why Did People Move During the Great Recession? The Role of Economics in Migration Decisions." *RSF: The Russell Sage Foundation Journal of the Social Sciences* 3 (3): 100–125. [6]

Marinescu, Ioana, and Roland Rathelot. 2018. "Mismatch unemployment and the geography of job search." *American Economic Journal: Macroeconomics* 10 (3): 42–70. [6]

Moretti, Enrico. 2011. "Local Labor Markets." Chap. 14, 1st ed., edited by O. Ashenfelter and D. Card, vol. 4B, 1237–313. Elsevier. [6]

Morten, Melanie, and Jaqueline Oliveira. 2016. "The effects of roads on trade and migration: Evidence from a planned capital city." *NBER Working Paper* 22158: 43. [6]

Nakamura, Emi, Jósef Sigurdsson, and Jón Steinsson. 2016. "The gift of moving: Intergenerational consequences of a mobility shock." Working paper. National Bureau of Economic Research. [6]

OECD. 2005. "How Persistent Are Regional Disparities in Employment? The Role of Geographic Mobility." *OECD Employment Outlook 2005 Edition.* [3]

Oreopoulos, Philip, Till Von Wachter, and Andrew Heisz. 2012a. "The short-and long-term career effects of graduating in a recession." *American Economic Journal: Applied Economics* 4 (1): 1–29. [3]

Oreopoulos, Philip, Till Von Wachter, and Andrew Heisz. 2012b. "The short-and long-term career effects of graduating in a recession." *American Economic Journal: Applied Economics* 4 (1): 1–29. [12]

Redding, Stephen J., and Esteban Rossi-Hansberg. 2017. "Quantitative Spatial Economics." *Annual Review of Economics* 9 (1): 21–58. [6]

Roback, Jennifer. 1982. "Wages, Rents, and the Quality of Life." *Journal of Political Economy* 90 (6): 1257–78. [6]

Şahin, Ayşegül, Joseph Song, Giorgio Topa, and Giovanni L Violante. 2014. "Mismatch unemployment." *American Economic Review* 104 (11): 3529–64. [6]

Sjaastad, Larry A. 1962. "The Costs and Returns of Human Migration." *Journal of Political Economy* 70 (5): 80–93. [6]

Spitz-Oener, Alexandra. 2007. "Human capital, job tasks and technology in East Germany after reunification." *National Institute Economic Review* 201: 97–106. [6]

Topel, Robert H. 1986. "Local labor markets." *Journal of Political economy* 94 (3, Part 2): S111–S143. [3]

Uhlig, Harald. 2006. "Regional labor markets, network externalities and migration: The case of German reunification." *American Economic Review* 96 (2): 383–87. [6]

Uhlig, Harald. 2008. "The slow decline of East Germany." *Journal of Comparative Economics* 36 (4): 517–41. [6]

Young, Alwyn. 2013. "Inequality, the Urban-Rural Gap, and Migration." *Quarterly Journal of Economics* 128 (4): 1727–85. [6]

Chapter 2

Regional Occupations, Local Rents and Worker Mobility*

2.1 Introduction

There are large differences between regional wage levels in developed economies. Yet, worker mobility remains low. This raises the question of why workers do not move to high-wage regions in large numbers.

In the literature, the most prevalent explanations are high moving costs, sorting of workers into regions by their productivity, and high regional living costs. Two important considerations have evaded the spotlight of the discussion, however. Firstly, local labor markets differ not only in productivity and wage levels, but also in the type of work workers perform. Secondly, moving decisions that take into account career opportunities depend crucially on the stage of a worker's career and therefore on age.

In this study, I combine these two points and argue that worker mobility is low because, over their careers, workers accumulate human capital that is tailored to their local labor market and is therefore not easily transferable to other regions. First, I empirically document large differences in real wage growth between regions and show that large parts of these differences can be accounted for by regional occupation compositions. Building on these findings, I develop a general equilibrium model in which the skills workers accumulate through experience reflect their labor market. In the calibrated model, young workers are discouraged from moving to the high-wage region primarily because of high rents, while older workers have little incentive to move since it would render the skills they have accumulated less

* I thank Moritz Kuhn and Hans-Martin v. Gaudecker for invaluable advice, guidance, and encouragement throughout this project. I also thank Christian Bayer, Michael Peters, Joseph Altonji, Kerstin Holzheu and Brigitte Hochmuth as well as participants of various seminars for their helpful feedback and comments. Support by the Deutsche Forschungsgemeinschaft (DFG) through CRC TR 224 (Project C01) is gratefully acknowledged. The usual disclaimer applies.

productive. Together, housing market frictions and region-specific skill accumulation are sufficient to lock workers into their initial regions, with only minor direct moving costs. In light of this interaction between local housing and labor markets, I proceed to conduct policy experiments that show that housing construction in high-wage regions can increase mobility, GDP, and welfare, even in a region where no new houses are constructed.

In the first part of the paper, I leverage high-quality administrative data from Germany and document three key facts. Firstly, I show that worker mobility is indeed low in Germany. At the age of 55, about half of all workers have spent their entire career in the very same commuting zone and about 70% have never moved further than to a commuting zone neighboring their very first one.

Secondly, over the life-cycle, the difference between expensive high- and cheap low-wage regions is not a difference in wage levels but in wage growth. Making use of the large correlation between the wage and the rent level, I split the sample into 3 using rent terciles. Taking regional prices into account, workers who begin their career in the lowest rent tercile actually earn about 6% higher real wages than their peers who start in the highest tercile. However, because of greater wage growth, workers from the highest tercile earn more on average in the middle part and especially at the end of their careers. At age 55, workers in the highest tercile earn about 10% higher real wages than their counterparts in the lowest tercile. I show that these growth patterns persist even when controlling for individual and age by occupation fixed effects in a dynamic diff-in-diff regression, casting doubt over the view that pure worker selection causes the wage differences.

Thirdly, occupations are spatially concentrated. Even after taking different sizes of regional workforces into account, each occupation in the data exhibits strictly convex Lorenz curves across commuting zones. The resulting different regional occupational compositions explain a sizable part of the regional variation in wage levels and wage growth. Depending on the decade, the composition of occupations explains about 60-80% of variation in the wage level and about 45-55% of variation in wage growth over 25 years at the county level.

In the second part of the paper, I propose a two-region general-equilibrium life-cycle model with local labor and rent markets to explain low worker mobility in light of my empirical findings. Workers draw utility from a final consumption good and housing, and can save but not borrow. Workers can choose to relocate between the regions at a resource and a utility moving cost every year. Human capital is two-dimensional, consisting of two separate skills, and is accumulated stochastically through experience. In line with different occupation compositions in each region, one of the two skills is more productive and can be learned faster, and in each region, a separate intermediate output good is produced. One of the regions has faster wage growth on average because its dominant skill can be accumulated faster.

I calibrate the model to the German data and investigate which forces lead to low worker mobility. I find that workers born in the poorer region with less wage growth face a dilemma. In principle, they would like to move to the region with high wage growth to maximize their expected career outcomes. The strongest force that keeps them from moving there is the high rent price in that region. By contrast, older workers well into their careers have accumulated wealth and can afford high rents more easily, but are kept back by their human capital. On average, it reflects their local labor market and is less applicable in the other region. Taking the dynamic interaction between locally adapted human capital and regional housing costs into account, the implied moving costs—measured as compensating differentials between moving and staying—are much lower than typically estimated in the literature, averaging about €39,000. They are especially small for young workers—only about €8,000 for 20-year-olds—and increase with age as human capital becomes more region-specific. I show that estimating a Mincer regression, which assumes full portability of experience across regions, strongly overestimates potential wage gains of moving to the high wage region.

Finally, I use the model to study housing policy as a lever for regional mobility. I conduct a policy experiment in which the government finances the construction of an additional 3% of housing in the high-rent region with greater wage growth. I find that this policy succeeds in inducing more mobility towards the high-wage region through the resulting rent price reduction. A secondary effect is that wages change in both regions. The larger influx of labor supply leads to lower wages in the region with greater growth and to a wage increase in the low-wage region. The policy is welfare-enhancing for workers in both regions, however. In the high-wage region, the welfare gains from cheaper housing outweigh the welfare loss from the wage decrease. Workers in the low-wage region value the wage increase at home, but crucially, the increased opportunity from easier mobility to the growth region as well. Overall consumption equivalent variation for newborn workers is 0.80%. Because more workers benefit from the faster skill accumulation in the rich region, I register a 0.04% increase in final output.

The remainder of the paper is structured as follows: Section 2.2 discusses the data I employ for the empirical analysis. Section 2.3 presents the empirical results. Section 2.4 discusses model assumptions and calibration. Section 2.5 explains low mobility through the lens of the model. The policy experiments are described in Section 2.6. Section 2.7 concludes.

Related Literature My work relates to several branches of the literature. Going back to Sjaastad (1962), the idea that moving can be seen as an investment has been present in the literature for a long time. In an important recent addition to this view Bilal and Rossi-Hansberg (2021), who stress the short-term cost and long-term pay-offs of location choice, and that moving to a cheap location is unlike financial assets, not subject to borrowing constraints. I document a clear application of long-term investment through location choice in light of regional human capital.

Equally important, there is the literature on moving costs and migration. An early seminal contribution is Roback (1982), who stresses the role of local amenities. In a more recent influential study Kennan and Walker (2011) estimate large moving costs but also substantial effects of expected income in different regions. Other papers replicate large moving cost estimates in other settings (e.g., Bryan and Morten, 2019; Tombe and Zhu, 2019; Ransom, 2022). Central to the mechanism in my model, Bayer and Juessen (2012) show that moving costs are overestimated if dynamic selection is not taken into account. Also related to my work, Kaplan and Schulhofer-Wohl (2017a) use a model with occupations and regions to explain a drop in interstate mobility by better information and less dispersed wages between states within occupations. Other papers focus on spatial frictions such as impeded ability to compete (Schmutz and Sidibé, 2018) or increased search costs (Heise and Porzio, 2022; Diaz, Janez, and Wellschmied, 2023). Perhaps closest to my work, Diaz, Janez, and Wellschmied (2023) stress the lifecycle nature of moving and higher rents because of wealthy late-career workers in high-wage regions. Diaz, Janez, and Wellschmied (2023) focus on search frictions, however. While their focus is on search frictions, my paper instead highlights how regional occupation structures and the accumulation of region-specific human capital shape workers' mobility decisions and regional wage dynamics. Moreover, unlike Diaz, Janez, and Wellschmied (2023), my model allows for borrowing constraints and captures general-equilibrium feedback through the production of regional intermediate goods.

I add to this literature in a number of ways. I explicitly take the accumulation of regional human capital because of regional occupation structures over the lifecycle into account. Additionally, I allow workers to accumulate wealth in the presence of borrowing constraints. Finally, I stress the general-equilibrium effects of moving large numbers of workers between regions that produce different intermediate goods.

Further, there is research on local living costs. Moretti (2013) finds that the nominal wage difference between skill groups is, in part, mitigated by high local living costs. Card, Rothstein, and Yi (2025) find that local wage-premia are offset by local living costs. Relatedly, my work connects to the literature on regional housing supply and its implications. Hsieh and Moretti (2019) find that in a Rosen-Roback style model, restrictions on new housing supply impede GDP growth in the US. Other papers, too, find evidence for regulations driving up housing prices (e.g., Glaeser, Gyourko, and Saks, 2005; Glaeser and Ward, 2009; Saiz, 2010; Herkenhoff, Ohanian, and Prescott, 2017; Glaeser and Gyourko, 2018). Building on these findings, I take inelastic housing supply as given and add the point that the resulting high living costs have career implications by making young workers miss the window of moving towards opportunity.

A key aspect of my theory is the specificity of human capital. Going back to Shaw (1984) and Shaw (1987), this idea has long been present in the literature.

More recently, Kambourov and Manovskii (2009) show that there are sizable returns to occupational experience. Sullivan (2010), too, finds that both occupation and industry-specific human capital are key determinants of wages. Another strand of the literature stresses the importance of different tasks performed in different occupations for the accumulation of human capital (e.g., Yamaguchi, 2012; Böhm, Gaudecker, and Schran, 2024).

The point that workers experience larger wage growth and learn more or more quickly in some places than in others, which is key to my theory, is well accepted in the literature. Glaeser and Maré (2001), find that workers in cities have higher wage growth and that a wage premium stays with workers who leave cities, suggesting faster human capital accumulation. Roca and Puga (2016), too, find that workers in large cities accumulate more valuable experience. Dauth et al. (2022) stress that workers in cities have better matching opportunities.

Finally, there is an influential branch of the literature that stresses selection and sorting of workers and firms into locations. Regarding firm selections Bilal (2023) explains regional differences in job stability and unemployment through selection of firms by productivity. Lindenlaub, Oh, and Peters (2022) show that firm sorting has an important influence on wage dispersion. Lhullier (2025) explains between and within-city wage inequality through agglomeration of productive firms. Concerning selection from the worker side, the literature is focused on selection of high-versus low-skilled workers. In this spirit, Borjas, Bronars, and Trejo (1992) stresses the role of differential returns to being high-skilled in different locations. Similarly, Hunt and Mueller (2004) find that workers with high skills migrate to areas with high returns. Dahl (2002) confirms the importance of comparative advantages of high and low education types in Roy-type regional selection models. Behrens, Duranton, and Robert-Nicoud (2014) explain the correlation between skills and city population through sorting by talent and productivity. Diamond (2016) attributes increases in sorting to labor demand changes and changes in local amenities. I add to this literature by stressing that skill is multidimensional and that it evolves with experience that depends on the region where it is gained. As a result, in my model, comparative advantage evolves with regional labor market experience.

2.2 Data

I use the *Sample of Integrated Labor Market Histories* (SIAB) provided by the Institute for Employment Research (IAB) at the German Federal Employment Agency. The SIAB is a 2% random sample of the complete labor market histories of all workers covered by social security in Germany between 1975 and 2017. Approximately 80% of the German workforce is covered, excluding only the self-employed, civil servants, and the military. The data contain detailed information on daily wages and socio-demographics, including age, gender, and education. A key feature of the

SIAB is its regional detail: it reports the county of each workplace, enabling me to track workers' locations over time. Because counties are relatively small administrative units, I measure geographic moves at the level of commuting zones defined by the Federal Office for Building and Regional Planning (BBSR), which better reflect local labor markets. An additional key strength of these data is that they record occupations in 120 levels, based on the German *KldB* classification.

For my analysis, I restrict the sample to male workers born between 1950 and 1964. Female labor force participation was relatively low in these cohorts, and East German observations are only available from 1993; I therefore drop all workers with any East German observations. To focus on long-term career and mobility outcomes, I further exclude workers with very short employment histories, keeping only those whose first observation occurs at age 30 or earlier and whose last observation occurs at age 45 or later.

Although the SIAB is recorded at a daily frequency, I construct an annual panel by retaining only the spell active on June 1 of each year. One limitation of the data is that wages are top-coded at the social security contribution ceiling; about 12% of wage observations are thus right-censored. Following Böhm, Gaudecker, and Schran (2024) and Dustmann, Ludsteck, and Schönberg (2009), I impute these censored wages with a standard procedure. The imputation procedure is described in detail in Appendix 2.A.1.

In addition to the microdata described above, my analysis requires information on local price levels. To measure the county-level price level, I use a regional consumer price index (CPI) published by the BBSR, based on regional prices in 2007. When measuring wage growth and in the calibration, I deflate all monetary variables to 1995 prices using the national CPI from Jordà et al. (2017). Regional price differences are accounted for using the BBSR regional CPI.

I also require information on local rents. For this, I draw on the rent index from the *Indicators and Maps on Urban and Spatial Development* (INKAR) database of the BBSR. These data measure the average offered rent per square meter for apartments in medium to good residential areas and are available from 2010 onward. Because the rent index is only available from 2010, I use 2010 values to rank counties in earlier years. To validate this, I make use of the high correlation between rents and building land prices. Building land prices are available at the county level since 1995, and for even earlier years for major cities. Appendix 2.A.2 shows that building land prices are extremely persistent between 1985 and 2010, and that after 2010, rents and building land prices are highly correlated. This provides confidence that the relative ranking of local rents is stable over time. Further, in my analysis, I aggregate the rent index into three categories (low, medium, high) to classify counties by relative cost using county rent terciles. As a robustness check, I use the government's regional rent categories for calculating housing allowances for low-income households (*Wohngeld*). These categories are based on rents for relatively low-cost apartments. Rerunning my main empirical analysis with these data does not change my empirical conclusions, as reported in Appendix 2.C.

To calibrate the model, I also use the *Socio-Economic Panel* (SOEP). To estimate initial wealth, I restrict the data to individuals observed between the ages of 16 and 20. I make use of each individual's reported asset values from the personal wealth module (financial assets, business assets, real estate assets, insurance and pension assets, and other assets such as valuable possessions). To estimate household size, I estimate the modified OECD equivalent scale on households with exactly one adult male worker born between 1940 and 1964. Details of this procedure are discussed in Appendix 2.E.2. I also use regional data from the federal statistics office (Destatis) and the statistics offices of the German states on the regional total square meters of housing and the regional number of births. Details on the calibration are discussed in Section 2.4.3.

2.3 Empirical Facts

2.3.1 Local Wages and Worker Mobility

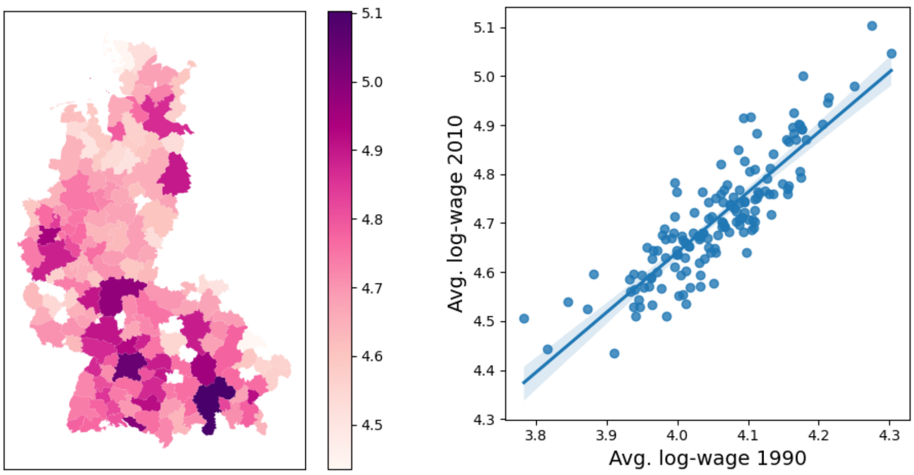
Differences between regional wage levels in Germany are large and persistent. The left panel of Figure 2.1 shows a map of West Germany with average log daily wages at the commuting zone level in 1990. The difference between the highest and the lowest level is about 0.6 log-points (over 60%)[1]. The right panel of Figure 2.1 shows the extremely high level of persistence of local wage levels in a scatter plot. The correlation between commuting zone average levels over the 20-year gap from 1990 to 2010 is 0.868.

Despite these large differences in local wage levels, worker mobility during the same time was low. Figure 2.2 shows the fraction of workers who have never left the commuting zone they initially started their career in against age. By age 55, around half of all workers have never left the commuting zone in which they began their career. The dashed line in Figure 2.2 reveals that as many as 70% of workers have never worked in a commuting zone that does not neighbor their initial commuting zone. This means that a large fraction of careers in Germany played out in only one commuting zone or close by. In particular, this fact is not simply driven by workers who spend their entire careers in high-wage areas. As shown in Figure 2.B.2 in the appendix, the numbers hardly change when reducing the sample to workers who started their career in commuting zones of the lowest wage tercile.

Given the large differences in the local wage level, this is surprising. Should the large and persistent wage level differences not lead to workers moving to high

1. Figure 2.1 in the appendix shows that a large part of the variation survives residualization by age and education.

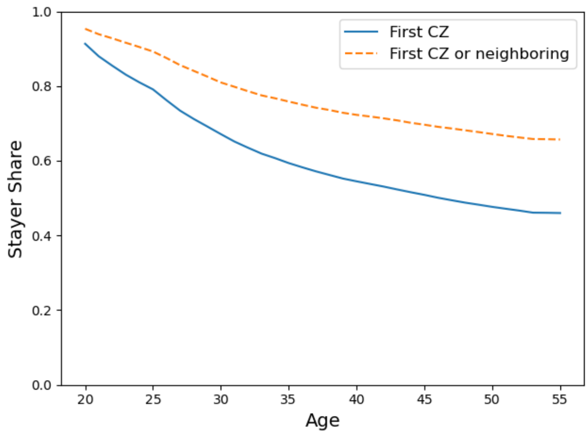
Figure 2.1. Local Wage Level and Persistence



(a) Avg. Log Wage 2010 *(b) Wage Level Persistence*

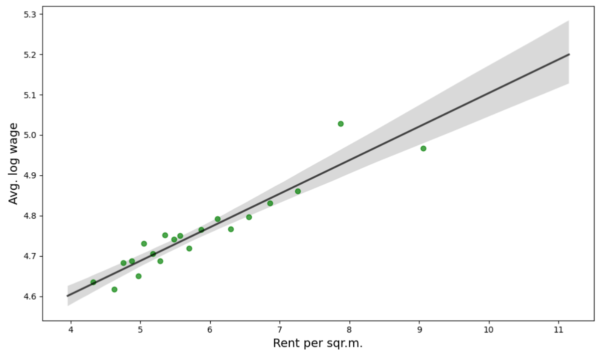
Notes: The left panel maps commuting zones in West Germany, colored by the average log daily wage in 2010 estimated from the SIAB (darker shades indicate higher values). The right panel shows a scatter of average log daily wages in 2010 (vertical axis) on average log daily wages in 1990 (horizontal axis) with an OLS fit. Each point corresponds to one commuting zone. The sample is restricted to full-time workers.

Figure 2.2. Fraction of Never-Movers



Notes: The solid line plots the share of workers in the SIAB sample who have never left their initial commuting zone, by age. The dashed line plots the share of workers who have either remained in their initial commuting zone or moved only to an adjacent commuting zone.

Figure 2.3. Rent vs. Wage-Level



Notes: Binned Scatter of average log-wages (SIAB) against average rents per square meter (INKAR) at the county level with 20 bins. The black line shows an OLS fit from regressing average log wages on average rents per square meter.

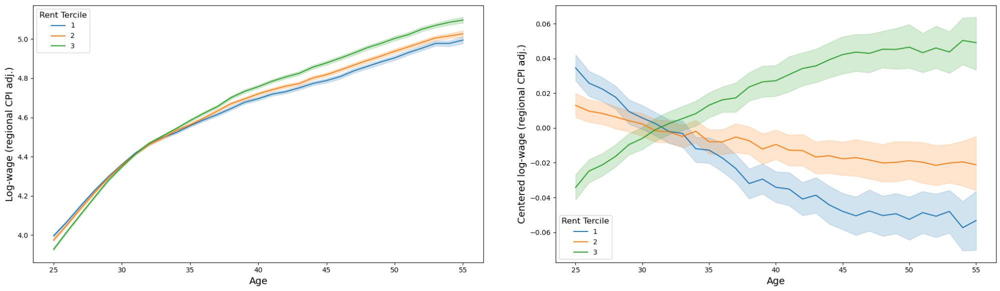
wage regions, leading to high worker mobility? A first candidate answer is that local living costs completely nullify the wage gains. Indeed, Figure 2.3 shows that a large correlation between the local wage and the local rent level exists. The natural question then becomes whether working in a high-wage/high-rent area even pays off.

To get to the bottom of this, I turn my attention to wage profiles earned over the lifecycle. First, I divide counties into terciles using the 2010 INKAR rent index. Then, I divide workers into three groups by the rent tercile that their very first workplace falls into. I then track these workers over their careers and calculate average log wages of full-time[2] workers at each age, using the BBSR CPI to adjust for regional price differences. The left panel of Figure 2.4 shows the result for workers born between 1960 and 1964. The right panel shows the same log-wage profiles with the national average log-wage subtracted at each age. This visualizes the log-wage difference between the three rent-terciles at each age. For example, at age 25, the blue line showing the demeaned profile of the lowest tercile is at 0.03 and the green line showing the highest tercile is at -0.03, indicating that workers who started in the lowest tercile earn about 6% higher real wages on average. Results for the remaining birth years are similar and shown in Appendix 2.B.

Figure 2.4 reveals two key facts. Firstly, after price differences have been accounted for, young workers do not earn higher average real wages in the high-rent/high-wage areas than their counterparts in the low-rent/low-wage areas. As discussed, in fact, at age 25, real wages are about 6% higher in the cheapest bin than in the most expensive bin. Secondly, looking only at the average wage level of different areas hides the fact that over the lifecycle, the difference between re-

2. Since the SIAB reports whether a worker works part-time but not the number of hours this refers to, I elect to exclude part-time wages from the results presented in Figures 2.4.

Figure 2.4. Log Wage Lifecycle Profiles



(a) Plain *(b) Mean Subtracted*

Notes: The left panel shows average log wages by age for workers in full-time employment who began their careers in counties belonging to the first, second, and third rent tercile, respectively. Rents are measured using INKAR data from 2010. The green line corresponds to the highest, the blue line to the bottom, and the orange line to the middle rent tercile. Wages are regionally CPI-adjusted using the BBSR CPI. The right panel shows the same wage profiles after subtracting the overall mean at each age. The sample includes workers born between 1960 and 1964.

gions is not in the wage level but in wage *growth*. While real wages are higher in the cheap bin for workers under 30, wages in the expensive regions grow faster. At 55, the wage gap between the most and least expensive bins has grown to more than 10% on average. In particular, integrating the differences up over the whole career, it is clear that being in the high-wage/high-rent area does pay off. The late career real wage growth outweighs the lower real wage entry level in terms of real earnings. However, it is also clear that young workers in high-wage areas will tend to be poorer because of high prices. From the point of view of a young worker, a career in a high-wage/high-rent area can be viewed as a long-term investment.

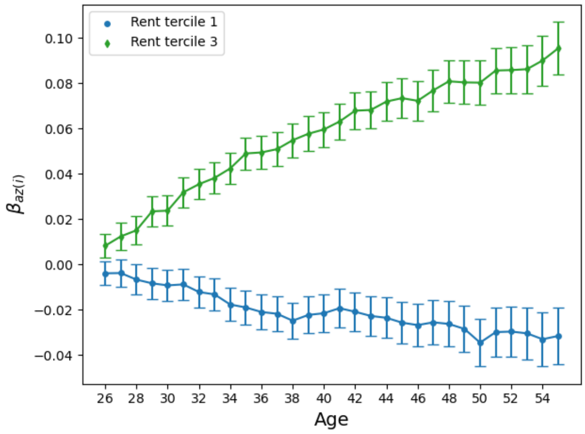
The patterns presented in Figure 2.4 are striking. In principle, however, they could be an artifact of worker selection — for instance, differences in productivity types or education. I therefore estimate the following regression:

$$\log w_{ij} = \sum_{\substack{k=1,3 \\ a=26,\ldots,55}} \mathbb{1}\{a = j\}\mathbb{1}\{\mathrm{z(i)} = k\}\beta_{jz(i)} + \alpha_i + \delta_j + \gamma_{\mathrm{educ}(i)j} + \mu_{\mathrm{pt.time},ij} + \eta_{t(ij)} + e_{ij} \tag{2.1}$$

w_{ij} is the wage of worker i at age j. $z(i)$ is the rent bin that worker i started his career in. α_i, δ_j and $\eta_{t(ij)}$ are worker, age and year fixed effects. $\mu_{\mathrm{pt.\text{-}time},ij}$ is a part-time effect and $\gamma_{\mathrm{educ}(i)j}$ is an age by education level fixed effect. $\beta_{jz(i)}$ is the coefficient of interest. It measures the effect of having started a career in rent bin $z(i)$ on the log-wage at age j. Note that age 25 and the workers who started their careers in the middle rent tercile are excluded as the baseline.

Equation 2.1 controls for two key factors. Firstly, it controls for worker fixed effects. If wage differences were to be explained by sorting of plain productivity

Figure 2.5. Regression Results



Notes: Regression coefficients $\beta_j z$ for age $j = 26, \ldots, 55$ and rent bins $z = 1, 3$ from estimating Equation 2.1. The green line corresponds to workers who started their careers in the highest rent tercile, the blue line corresponds to workers who started their careers in the bottom tercile. Age 25 and the middle rent bin $z = 2$ are excluded as the baseline level. β_{jz} measures the effect of having started a career in rent bin $z(i)$ on the log-wage at age j. Standard errors are clustered at the worker level.

types, this fixed effect should explain the majority of wage differences. Secondly, it controls for age-by-education fixed effects. If the pattern of differential was to be explained by different education types who have selected into their education group because of their ability to learn and who have faster wage growth precisely because of this better learning ability, this effect should explain the growth differences.

Figure 2.5 reports the estimated coefficients of interest and standard errors. Neither of the discussed controls explains the difference in wage growth over the career. The difference in wage effects between the top and the bottom rent bin amounts to about 3% at age 26 and continues to grow steadily with age. The difference between the effects of having started the career in the highest vis-à-vis the lowest rent bin on wages at 55 amounts to about 12% - an effect of similar magnitude compared with the raw means presented in Figure 2.4. This suggests that instead of sorting of plain types or of differently educated types who can learn faster, another mechanism must be at play, producing the observed patterns.

Next, I turn my attention to the type of work performed in the different regions - the workers' occupations.

2.3.2 Regional Occupations

Occupations in Germany are spatially very concentrated. Figure 2.6 shows Lorenz curves based on the share of all workers working in a given occupation in each commuting zone on the left. Each red line represents one of the 120 *KldB88* occupations reported in the SIAB. The black line represents the hypothetical Lorenz curve for an occupation with equally many workers in each commuting zone. All the Lorenz curves are strictly convex, indicating strong spatial concentration of occupations. At the extreme, the bottom line (corresponding to nautical and aeronautical workers) crosses the point $(0.9, 0.247)$, indicating that as many as 75.3% of all nautical/aeronautical workers work in only 10% of commuting zones. Other occupations represented by lines at the very bottom include insurance professionals, chemists, chemistry lab assistants, publicists, and data specialists. Occupations exhibiting the least spatial concentration include carpenters, scaffolders, plastics-processing technicians, and bricklayers.[3]

Part of the extreme concentration displayed by the red Lorenz curves in the left part of Figure 2.6 is driven by the different sizes of commuting zones. The right part of Figure 2.6 shows Lorenz curves in blue that take this into account. Each commuting zone is assigned a section on the x-axis that is proportional to the number of workers contained in the SIAB who work there. Then, commuting zones are sorted by the number of workers per capita in a given occupation. The y-axis still shows accumulated worker shares. The black line now represents the hypothetical Lorenz curve for an occupation with equally many workers per capita in each commuting zone. As expected, the resulting Lorenz curves are closer to the 45-degree line than the uncorrected counterparts. However, all of them are still strictly convex, and occupations still exhibit a large degree of spatial concentration.

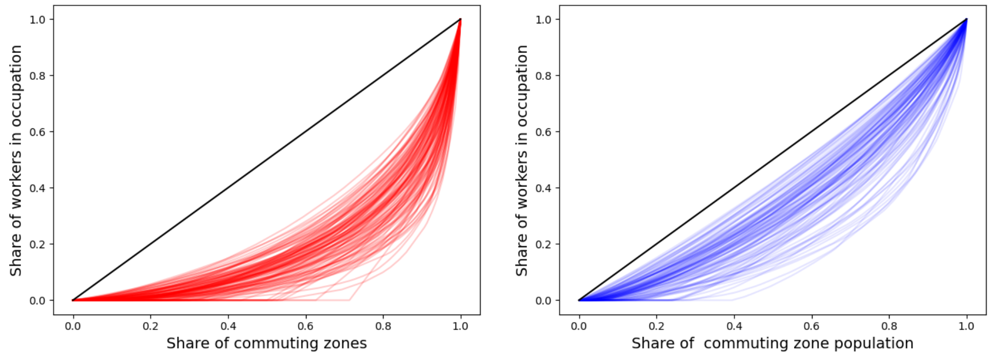
The main takeaway from Figure 2.6 is that workers in different regions are faced with different occupation compositions. This naturally raises the question if these regional occupation structures explain the regional differences in wage levels and wage growth shown in the previous section. To investigate this, I revisit the log-wage profiles shown in Figure 2.4 and conduct a simple counterfactual exercise that eliminates any regional wage differences *within*-occupation. This isolates the role of occupational composition. If the occupation composition were identical across rent bins, the log-wage profiles should then coincide exactly.

I proceed as follows. As a first step, I calculate average log wages by occupation and age across all regions in West Germany. Each worker is then assigned this national average wage according to their occupation and age, regardless of their region. I then recompute the lifecycle wage profiles of the three regional rent bins using these counterfactual wages.

Figure 2.7 compares the observed average log-wage profiles already presented in Figure 2.4 to the resulting counterfactual log-wage profiles. The solid lines are the empirical wage profiles already shown in Figure 2.4 and the dashed lines are the counterfactual log wage profiles without intra-occupational wage dispersion. The

3. As demonstrated by e.g Dauth, Fuchs, and Otto (2018), industries are regionally concentrated too. Since industries and occupations are closely connected and the model I propose in Section 2.4 does not explicitly differentiate between the two notions, and the SIAB is better suited to speak to occupations, I restrict my analysis to occupations.

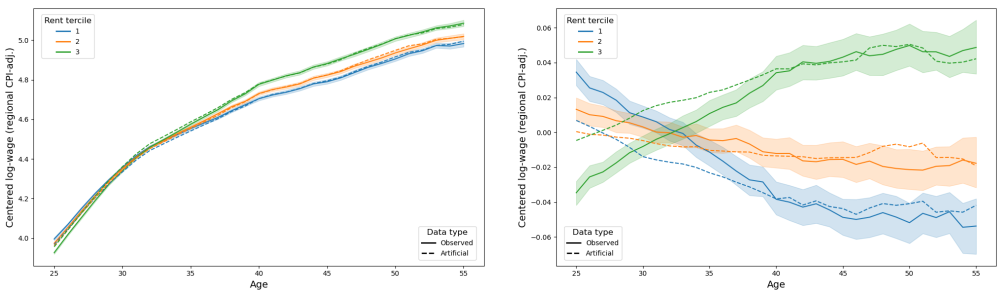
Figure 2.6. Occupation Lorenz Curves



(a) CZ-level *(b) CZ-size-corrected*

Notes: Lorenz curves based on shares of all workers working in an occupation who work in each of the BBSR commuting zones. On the left, the commuting zones on the x-axis are sorted from the lowest employment share to the highest employment share separately for each occupation. Each red line represents one of the 120 *KldB88* occupations reported in the SIAB. The black solid line represents the 45-degree line, i.e., the Lorenz curve for an occupation that has equally many workers in each commuting zone. On the right, commuting zones on the x-axis are sorted by employment per capita estimated from the SIAB. Each commuting zone is assigned a segment on the x-axis that is proportional to its size in the SIAB. The black solid line represents the 45-degree line, i.e., the Lorenz curve for an occupation that has equally many workers per capita in each commuting zone. The sample is restricted to the year 2000.

Figure 2.7. Observed vs. Counterfactual Log Wage Profiles



(a) Plain *(b) Mean Subtracted*

Notes: The solid lines show the average log wage profiles for workers in full-time employment who began their careers in counties belonging to the first, second, and third rent tercile, respectively. The dashed lines show counterfactual log wage profiles that results from assigning every worker the national average wage, conditional on occupation and age. The green line corresponds to the highest, the blue line to the bottom, and the orange line to the middle rent tercile. Wages are regionally CPI-adjusted to the level of Bonn using the BBSR CPI. The right panel shows the same wage profiles after subtracting the overall mean at each age. Rents are measured using INKAR data from 2010. The sample includes workers born between 1960 and 1964.

left panel shows average log-wage profiles, while the right panel displays the same profiles after subtracting the national mean at each age. The counterfactual profiles almost exactly replicate the empirical ones. In particular, they closely replicate the differences in wage growth as demonstrated by the right panel.

As discussed, if the occupation composition in the three rent bins were the same, all the log wage profiles would be the same too. In the right panel, this would mean that all three dashed lines should be exactly at 0 at every age. Any difference from 0 must come from the occupation composition behind the three dashed lines. The near-perfect match between empirical and counterfactual profiles implies that the wage differences between workers in the three rent bins can be almost fully attributed to differences in occupational composition, rather than to within-occupation regional wage differentials.

To complement the findings presented in Figure 2.7 with a simple quantitative measure, I next report how much of the between-county variation in wages can be statistically explained by differences in occupational composition. Specifically, I estimate two straightforward regressions—one for wage levels and one for wage growth—to summarize the explanatory power of regional occupation shares in terms of R^2 values.

I first regress average county log wages on the county's occupational employment shares:

$$\overline{\log w_c} = \sum_{\omega \in \Omega} \beta_\omega s_{c\omega} + e_c \tag{2.2}$$

where $s_{c\omega}$ denotes the employment share of occupation ω in county c and Ω is the set of 28 two-digit occupations.[4]

Panel A of Table 2.1 shows that occupation shares explain more than 60% of the between-county variation in wage levels, reaching 80% in the 2000–2010 decade. Thus, most cross-regional wage variation can be attributed to differences in occupational composition.

In Section 2.3.1, I showed, however, that over the lifecycle, the main difference between high- and low-wage regions is not in the entry level, but in wage growth.[5] To examine whether occupational composition also predicts regional wage *growth*, I estimate a second regression where the dependent variable is the average log-wage

4. I work at the county × 2-digit occupation level (rather than BBSR commuting zones × 120 SIAB occupations) to avoid an overparameterized regression. There are 28 2-digit occupations and 269 counties as opposed to 120 occupations and 151 commuting zones. The occupation classification in my SIAB version does not allow differentiation between 2-digit occupations on 3 occasions, leaving only 28 2-digit occupations instead of the full 31 given by the standard KldB classification.

5. All wage levels are adjusted to the price level of 1995 using the CPI from the Macro History Database (Jordà, Schularick, and Taylor, 2017).

Table 2.1. R^2 for Occupation-share Regressions.

Panel A: Wage level by decade		**Panel B: Wage growth by cohort**	
Years	R^2	Cohort	R^2
1980–1990	0.632	1950–1954	0.444
1990–2000	0.708	1955–1959	0.467
2000–2010	0.809	1960–1964	0.543

Notes: Panel A reports R^2-values from estimating Equation 2.2. The outcome variable is the county-level average log wage. Panel B reports R^2-values from estimating Equation 2.3. The outcome variable is the difference between the average log wage at 50 and at 25. Wages and occupation shares are measured at the county and 2-digit occupation level.

difference between ages 50 and 25:

$$\overline{\log w_c^{50}} - \overline{\log w_c^{25}} = \sum_{\omega \in \Omega} \gamma_\omega \tilde{s}_{c\omega} + u_c \tag{2.3}$$

I estimate Equation 2.3 separately by cohort. The occupation shares $\tilde{s}_{c\omega}$ are estimated early in each cohort's career - to be precise, 1975-1980 for the 1950-1954 cohort, 1980-1985 for the 1955-1959 cohort, and 1985-1990 for the 1960-1964 cohort. The resulting R^2 values are reported in Panel B of Table 2.1. Across all three cohorts, the R^2 values are close to 50% of variation explained. This result is particularly striking because wage growth is measured over 25 years, and occupation shares are measured only around the time at which workers begin their careers. Even so, a sizeable part of wage growth variation can be explained by the occupation composition young workers face in a given county.

2.4 Model

2.4.1 Model Motivation

In Section 2.3, I document two key facts. First, using the strong correlation between rent and wage levels, I show that the difference between regions with high wages and high rents and regions with low wages and low rents is not merely a difference in wage levels, but one in wage growth. This is consistent with workers in the high-wage areas accumulating human capital more quickly. Second, the regional occupation composition explains large parts of wage levels and growth, and the different observed wage profiles are wage profiles of workers in different occupations.

In this section, I lay out a general-equilibrium model that reflects these observations and makes use of them to endogenously explain the low levels of worker mobility. The model features two regions with regional labor and rent markets. Regional occupation structure is represented in two ways.

First, in the model, workers accumulate two different skills through experience. Production in each region emphasizes one of the skills, making the accumulation of this skill more likely and more productive. As a result, on average, the human capital of a worker will be more useful in the region in which it was accumulated, making moving less and less attractive as careers progress. This is in line with the literature that documents the accumulation of occupation-specific human capital. Second, I assume that each regional labor market produces a separate intermediate good, reflecting specialization arising from regional occupational structure. These intermediate goods are then aggregated into a final consumption good.

Using the calibrated model, I then proceed to quantify different mechanisms that lead to low mobility. I then conduct a policy exercise in which additional housing is constructed in the region with better career opportunities.

2.4.2 Model Assumptions

Regions and Production The production function is set up as in the canonical model for production with different skills in Acemoglu and Autor (2011), adding spatial segregation. There are two locations called A and B. In each region, there is a regional representative firm that produces a tradable intermediate good using labor as input. The production function of such a regional firm is

$$F_z(L_z) = L_z$$

where z denotes the region and L_z is the total amount of effective labor supplied to the firm in region z.

The two intermediate goods produced in each region are aggregated to a final consumption good by a national aggregation firm with technology

$$Y = \left(F_A(L_A)^{\frac{\sigma-1}{\sigma}} + F_B(L_B)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}$$

where σ is the elasticity of substitution between the two intermediate goods. Labor markets are spatially segmented and competitive. The wage for one unit of labor supplied in region A and the wage for one unit of labor supply $\overline{w}_B$ are given by

$$\overline{w}_A = \frac{\partial Y}{\partial L_A} = \left[1 + \left(\frac{L_B}{L_A}\right)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{1}{\sigma-1}} \qquad \overline{w}_B = \frac{\partial Y}{\partial L_B} = \left[1 + \left(\frac{L_A}{L_B}\right)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{1}{\sigma-1}}.$$

and the wage in Region B takes on the same form. Note that the wage in both regions depends on the ratio of labor supplied in both regions. This is a key consequence of the two regions producing separate intermediate goods. Since both of these goods are needed to produce the final consumption good, wages are not only a function of local labor supply[6], but instead depend on labor supply in the other region too. As a result, worker re-location from one region to the other will - all else equal - raise wages in the region that workers leave and decrease them in the other region.

Each region is endowed with an exogenous stock of housing, denoted by $\overline{\eta_z}$, that is owned by absentee landlords and rented out in spot markets at the endogenous rent price p_z.

Workers - Basic Assumptions Workers are born and draw their birthplace z_0 and an initial endowment of wealth e_0. The worker is born in region A with probability p_A, and the distribution of initial wealth is denoted by F_e. Time is discrete and at annual frequency. Workers work for 40 years and live for 60, spending the last 20 years in retirement. Age is denoted by index j.

Each period before retirement, workers can decide whether to move to the other region at the end of the period. If they decide to move, they pay a resource moving cost κ and incur a utility moving cost $k(j)$, where the latter is increasing with age towards the end of the lifecycle. $k(j)$ captures the emotional and social cost of relocation—leaving behind familiar surroundings, friends, and family. The cost rises late in the life cycle as workers become more settled and socially embedded in their region. Each period before retirement, workers draw an *iid* taste shock for each region ζ_{zj} from an Extreme Value Type 1 distribution.

Workers can invest in an asset a at an exogenous rate R that is determined in international financial markets.

Workers - Labor Supply Workers work full-time for the local firm in the region where they live. The number of units of effective labor they supply to the local firm is a function of idiosyncratic productivity ψ_j and two skill levels - h_{aj} and h_{bj}. Idiosyncratic productivity ψ_j follows an $AR(1)$ process in log-space:

$$\log \psi_j = \rho_\psi \log \psi_{j-1} + \varepsilon_j, \quad \varepsilon_j \sim \mathcal{N}(0, \sigma_\psi)$$

Both skill levels can take N discrete values on the grid $\mathcal{H} = \{h_1, \ldots, h_N\}$ with $h_1 < \ldots, < h_N$. Given productivity and skill levels, the effective units of labor supplied inelastically by each worker are determined by region-specific functions:

$$\begin{aligned} \ell_A(h_a, h_b, \psi) &= h_a^{\alpha_A} \cdot h_b^{1-\alpha_A} \cdot \psi \\ \ell_B(h_a, h_b, \psi) &= h_a^{\alpha_B} \cdot h_b^{1-\alpha_B} \cdot \psi \end{aligned}$$

6. Note that earnings only depending on local TFP and labor supply is the most prevalent assumption in the spatial labor literature - see e.g. Bilal (2023), Hsieh and Moretti (2019), Diaz, Janez, and Wellschmied (2023).

A worker's labor earnings, given skill levels and productivity, are again region-specific and given by

$$\begin{aligned} w_A(h_a, h_b, \psi) &= \ell_A(h_a, h_b, \psi) \cdot \overline{w}_A \\ w_B(h_a, h_b, \psi) &= \ell_B(h_a, h_b, \psi) \cdot \overline{w}_B \end{aligned}$$

In each region, one of the two skills is more productive at producing the local output good. I assume that h_a is the dominant skill in region A and that h_b is the dominant skill in region B Formally, this means that $\alpha_A > \frac{1}{2}$ and that $\alpha_B < \frac{1}{2}$.

All workers at age $j = 1$ start their careers with skill-levels $h_a = 1$ and $h_b = 1$. Through experience, workers stochastically move up the skill grid. Every year that workers work, there is a probability of the skill index of either or both skills moving to the next grid point if they have not yet reached the h_N. For $i - 1, \dots, N - 1$ this probability takes the form

$$P(h_a' = h_{i+1} \mid h_a = h_i, z, j) = \begin{cases} \alpha_A \cdot \overline{p_a} \cdot \delta_a^{j-1}, & z = A, \\ \alpha_B \cdot \overline{p_a} \cdot \delta_a^{j-1}, & z = B, \end{cases}$$

$$P(h_b' = h_{i+1} \mid h_a = h_i, z, j) = \begin{cases} (1 - \alpha_A) \cdot \overline{p_b} \cdot \delta_b^{j-1}, & z = A, \\ (1 - \alpha_B) \cdot \overline{p_b} \cdot \delta_b^{j-1}, & z = B. \end{cases}$$

The interpretation is straightforward. First, the probability exponentially decays with age j. This means that the older workers become, the slower they learn and the lower the probability of making the next step. Second, the exponential decay is different for the two skills. This means that one of the skills can be learned more quickly, or alternatively, that the probability of a meaningful skill increase is higher simply because there is more to learn. Third, the probability is region-specific in that it is scaled by the Cobb-Douglas exponent from the labor supply function. This means that each skill can be learned more quickly in the region where it is dominant, capturing that skills are learned faster if they are used more through learning by doing, and if there are more people who have this skill, from which it is possible to learn. The Cobb-Douglas exponent captures both - how important the skill is for making the local good - and the exposure to experience, which scales the skill-inherent progress probabilities.

Workers - Value Functions Workers draw utility from consumption of the consumption good c and housing units η. The flow utility function is Stone-Geary and given by

$$u_j(c, \eta) = \frac{\nu_j}{1 - \lambda} \left(\left(\frac{c}{\nu_j} \right)^{\gamma} \cdot \left(\left(\frac{\eta}{\nu_j} \right) - \nu_j \underline{\eta} \right)^{1-\gamma} \right)^{1-\lambda}.$$

A couple of points warrant discussion. First, there is a subsistence level of housing consumption $\underline{\eta}$, capturing that a worker cannot survive without a minimum amount of shelter. Another realistic effect of this assumption is that poorer workers spend a greater share of their expenditure on housing. After covering expenses for $\overline{\eta}$, workers spend constant shares γ and $1-\gamma$ on consumption and housing, respectively.

Second, there are household size weights $\{\nu_j\}_{j=0}^{T}$. Using the SOEP, I estimate the share of household income in families with one adult male in the cohorts I based my empirical results on to be about 80% (See Appendix 2.E.1 for details). I therefore assume that the labor income earned by any one worker is, in fact, the household income of this worker's household. Consequently, the age-dependent scaling parameter ν_j captures the size of a worker's household over the lifecycle and the fact that the larger a household, the larger the marginal utility of consumption and housing is. Finally, λ is the intertemporal elasticity of substitution.

Each working-age individual solves a dynamic location and consumption–savings problem. At every age j, a worker with current state (z, h_a, h_b, a, ψ) decides (i) whether to remain in the current region or move to the other region, and (ii) how much to consume, save, and rent housing for that period. The value function of a working-age worker is therefore

$$V(j,z,h_a,h_b,a,\psi) \;=\; \max_{z'\in\{A,B\}} \left\{V_{z'}(j,z,h_a,h_b,a,\psi) - \mathbb{1}\{z \neq z'\}\, k(j)\, \nu_j + \zeta_{z'}\right\} \quad (2.4)$$

where z' denotes the chosen region for the next period, $V_{z'}(\cdot)$ is the conditional value function given that choice, $k(j)\,\nu_j$ is the age- and household-size–specific utility cost of moving, and $\zeta_{z'}$ is a region-specific idiosyncratic preference shock drawn from an Extreme Value Type I distribution. This structure yields a standard logit decision rule for regional mobility.

Conditional on choosing region z', the worker optimizes over consumption c, next-period assets a', and housing η:

$$V_{z'}(j,z,h_a,h_b,a,\psi) \;=\; \max_{a',c,\eta} \left\{u_j(c,\eta) + \beta\, \mathbb{E}_j\left[V(j+1,z',h'_a,h'_b,a',\psi')\right]\right\}$$

$$\begin{aligned} s.t. \quad \mathbb{1}\{z \neq z'\}\kappa\,\nu_j + p_z\eta + c + a' \;&=\; w_1(\omega(z), h_{\omega(z)}, \psi)(1-\tau) + Ra \\ a' \;&\geq\; 0 \end{aligned}$$

The term $\kappa\,\nu_j$ represents the resource cost of moving, paid only when the worker relocates ($z' \neq z$), and scaled by household size ν_j. Both moving costs are multiplied by ν_j to reflect that larger households face higher financial and utility costs when moving. Expectations are taken over next-period skill levels, idiosyncratic productivity, and regional taste-shocks, conditional on all information available at age j. Workers pay a linear tax τ on labor income.

At the beginning of the lifecycle, there is an initial period $j=0$. Workers are born with an initial productivity draw ψ_0, wealth endowment e_0, and birthplace z_0. Before entering the first working period ($j=1$), they decide in which region to start their career. This decision is subject to the usual idiosyncratic taste shocks ζ_A and

ζ_B, and takes into account moving costs relative to their birthplace. The value of a newborn worker is

$$V(0, z_0, e_0, \psi_0) = \max_{z' \in \{A,B\}} \left\{ V^0_{z'}(z_0, e_0, \psi_0) - \mathbb{1}\{z_0 \neq z'\}\, k(0)\, \nu_0 + \zeta_{z'} \right\}, \tag{2.5}$$

where z' denotes the region chosen for the start of working life. The conditional value function $V^0_{z'}(\cdot)$ is the continuation value from entering region z' at age $j = 1$:

$$\begin{aligned} V^0_{z'}(z_0, e_0, \psi_0) &= \max_{a'} \left\{ \beta\, \mathbb{E}_0 \left[V(1, z', h_{a,1}, h_{b,1}, a', \psi_1) \right] \right\} \\ &\text{s.t.} \quad \mathbb{1}\{z_0 \neq z'\} \kappa\, \nu_0 + a' = e_0, \quad a' \geq 0. \end{aligned}$$

Workers do not make consumption or housing choices in this initial period because it is assumed that they still live with their parents and that this choice is made for them.

After the working phase, workers retire in their last workplace and can no longer move. They receive unemployment benefits $b_i = s_b \cdot w_{i40}$, where w_{i40} is the total labor income, worker i earned in the last working-age period and $s_b \in (0, 1)$[7]. The value function of a retired worker for $j < 60$ is

$$\begin{aligned} V^{\text{ret}}(j, z, b, a) &= \max_{c, \eta, a' \geq 0} \left\{ u_j(c, \eta) + \beta V^{\text{ret}}(j+1, z, b, a') \right\} \\ s.t. \quad c &= Ra + b - a' - p_z \eta \\ a' &\geq 0 \end{aligned} \tag{2.6}$$

In the final period, workers consume all their wealth, i.e., the value function in the final period of a worker's life is

$$\begin{aligned} V^{\text{final}}(z, b, a) &= \max_{c, \eta} \{ u_{60}(c, \eta) \} \\ s.t. \quad c &= Ra + b - p_z \eta. \end{aligned} \tag{2.7}$$

7. This means that b is a direct function of the skill levels, productivity, and location in the last period. As such, the value function of the retired worker could be expressed as a function of the same arguments as a working phase value function. To save notation, (h_a, h_b, ψ) are condensed into the state b, however.

Equilibrium Individual state variables consist of age j, location z, skill levels h_a and h_b, productivity ψ and assets a. The individual state vector is denoted by $x := (j, z, h_a, h_b, \psi, a) \in \mathbb{X}$ where $\mathbb{X}$ is the state space. Let $\mathbb{A} = \{x \in \mathbb{X} \mid z = A\}$ and $\mathbb{B} = \{x \in \mathbb{X} \mid z = B\}$. Let μ_A denote the cumulative distribution of states contained in $\mathbb{A}$ and μ_B denote the cumulative distribution of states contained in $\mathbb{B}$. A competitive steady state equilibrium is a collection of value functions $\{V(x), V^0(x), V^{\text{ret}}(x)\}$[8], worker decision rules $\{z'(x), c(x), \eta(x), a'(x)\}$, equilibrium rent prices p_A, p_B and unit wages $\overline{w}_A, \overline{w}_B$ such that

(1) Given rent prices p_A, p_B and unit wages $\overline{w}_A, \overline{w}_B$, the decision rules $\{z'(x), c(x), \eta(x), a'(x)\}$ solve the workers' decisions by solving the recursive maximization problems (2.4)-(2.7) with the value functions $\{V(x), V^0(x), V^{\text{ret}}(x)\}$.

(2) Given the household decisions $\{z'(x), c(x), \eta(x), a'(x)\}$

$$
\begin{aligned}
L_A &= \int_{\mathbb{A}} \ell_A(h_a, h_b, \psi) d\mu_A(a) \\
L_B &= \int_{\mathbb{B}} \ell_B(h_a, h_b, \psi) d\mu_B(b)
\end{aligned}
$$

and the unit wages $\overline{w}_A, \overline{w}_B$ maximize the final consumption good firm's profits given $\frac{L_A}{L_B}$.

(3) Rent Markets clear, i.e.

$$
\begin{aligned}
\int_{\mathbb{A}} \eta(a) d\mu_A(a) &= \overline{\eta}_A \\
\int_{\mathbb{B}} \eta(b) d\mu_B(b) &= \overline{\eta}_B.
\end{aligned}
$$

The equilibrium is solved for using numerical methods. See Appendix 2.D for details.

2.4.3 Model Calibration

I calibrate the model to the cohort born between 1960 and 1964. All prices and wages are expressed in 1995 values. Regional price differences are captured endogenously through the model's rent prices and the implied regional CPI. These differences are left untargeted; their fit is discussed in Section 2.4.4.

Since the model has only two regions, I define the data analogue regions as follows: I define all counties contained in the lower two rent bins (rent tercile 1 and

8. Note that I use $V^{\text{ret}}(j, z, b, a) = V^{\text{ret}}(j, z, h_a, h_b, \psi, a) = V^{ret}(x)$ where h_a, h_b and ψ refer to the skill levels and productivity in the last working age period of a worker.

2 over counties) from Section 2.3 to be region A and the highest rent bin (tercile 3) to be region B. With this definition, about 56% of observations of the cohort in question in the SIAB are from Region A and the remaining 44% are from B. I interpret model age $j = 1, \dots, 40$ to represent real-world age $20, \dots, 59$. In the following, I explain the calibration of parameters.

To translate real-world to model prices, I use the average full-time wage in the SIAB and the average wage in the model.

Utility Parameters The intertemporal elasticity of substitution λ is set to the standard value 2 found, e.g., in Kaplan and Schulhofer-Wohl (2017b). The elasticity of substitution between consumption and housing γ targets 22% of expenditure share for housing by renters reported by Schlattmann (2024), taking into account the subsistence level of housing $\underline{\eta}$. $\underline{\eta}$ itself is set to the model equivalent of $9m^2$ since this is the minimum legal personal living space in most German states.

The discount factor β is set to 0.98.

Finally, the household size weights $\{\nu_j\}_{j=0}^{60}$ are estimated using the modified OECD equivalence scale and SOEP data. I fit a cubic polynomial to the household sizes of 25-70 year old male workers, who are the only adult male in their household born between 1940 and 1964. Before 25 and after 70, I impose that the resulting weight is the same as the nearest fitted value, since there are only relatively few observations of very young and very old households. The estimated values are shown in Figure 2.E.2 in Appendix 2.E.

Production Parameters A key parameter is the elasticity of substitution between the intermediate goods produced in each region, as it determines how strong general equilibrium effects of labor relocation will act on local wage levels. Since the production function of the aggregation firm can be interpreted as a representation of trade of sectoral intermediate goods between regions, I choose $\sigma = 4$ in line with the interval of values estimated by the trade literature for Armington elasticities (e.g., Feenstra et al., 2018; Bajzik et al., 2020).

Region Parameters Regional housing supply is estimated using data from the regional branch of the German Federal Statistics Office. These data report the total square meters of residential buildings at the county level. The total number of square meters in the analogues of regions A and B, denoted by $\overline{\eta}_A^{\text{data}}$ and $\overline{\eta}_B^{\text{data}}$, is obtained by summation across counties. Model housing supplies are then given by

$$\overline{\eta}_A = \overline{\eta}_A^{\text{data}} \cdot \phi, \quad \overline{\eta}_B = \overline{\eta}_B^{\text{data}} \cdot \phi,$$

where ϕ converts total square meters into efficiency units of housing. It is chosen such that the model matches the average rent price per square meter in West Germany in 1995 of €5.22, according to the statistical yearbooks from Destatis.

The probability of being born in region A, denoted by $P(A)$, is estimated at 55% using county-level birth data from Regionalstatistik.

Endowment Distribution I use wealth data from the SOEP for individuals under 20 to estimate the initial wealth distribution F_e. Since 59% of under-20-year-olds report zero wealth, I assume that workers draw an initial endowment of 0 with probability 0.59, and with probability 0.41 draw a positive endowment from a log-normal distribution:

$$\log e_0 \mid e_0 > 0 \sim \mathcal{N}(\mu_e, \sigma_e).$$

The parameters μ_e and σ_e are estimated by maximum likelihood.

Idiosyncratic Productivity The continuous $AR(1)$ process for idiosyncratic productivity is approximated by using Rouwenhorst discretization with five points. The auto-regressive coefficient ρ_ψ is estimated from the SIAB by estimating

$$\log w_{ij} = \alpha_{z(ij)} + \gamma_{\text{edu}(i)} + \delta_{\omega(i)} + t_{ij}^{\omega}\beta_1 + t_{ij}^{z}\beta_2 + j\beta_3 + j^2\beta_4 + u_{ij}$$

where z denotes the sample analogue region of A and B, $\alpha_{z(ij)}$ is a region fixed effect, $\delta_{\omega(i)}$ is an occupation fixed-effect, t_{ij}^{ω} and t_{ij}^{z} are occupational and regional tenure, and $\gamma_{\text{edu}(i)}$ is an education effect. The occupation and region effects proxy the regional wage levels $\overline{w}_A$ and $\overline{w}_B$. The tenure variables, as well as the age profile, capture the part of the wage that the model explains through human capital accumulation of the two skills.

Next, I estimate ρ_ψ as the persistence parameter of the residual component by fitting

$$\hat{u}_{ij} = \rho_\psi \hat{u}_{i,j-1} + \varepsilon_{ij},$$

where $\hat{u}_{ij}$ denotes the empirical residuals from the first-stage regression. The variance of the innovation σ_ψ^2 is then chosen such that the variance of log wages in the model matches the variance of the empirical log wages.

Moving Costs There are two types of moving costs: the resource cost κ and the utility cost $k(j)$. The resource cost κ is fixed at the model equivalent of €5,000. This value is intended to represent the total out-of-pocket expenses associated with a long-distance move within Germany, including transportation, deposits, and partial furnishing.[9]

The functional form of $k(j)$ is chosen to be approximately constant for most of a worker's career and to rise only toward the end of working life, starting around $j = 26$ (age 45 in the data). This captures that older workers become increasingly reluctant to move as they approach retirement—for example, because they have stronger social and family ties or are more attached to their local environment.

9. Publicly available estimates for long-distance moves (e.g., between major cities such as Duisburg and Munich) suggest direct costs of about €2,000–2,500 for small flats, excluding deposits and new furniture (Check24, price comparison platform).

Formally, I specify

$$k(j) = k \cdot \text{scale}(j),$$

where scale(j) is a smooth S-shaped function that remains near one until about age 45 and then increases. The level parameter k is calibrated internally. Details on the exact functional form of scale(j) and the calibrated age profile of $k(j)$ are provided in Appendix 2.E.3.

Human Capital The human capital grid is defined as $\mathscr{H} = \{h_1, \ldots, h_{10}\}$ with $N = 10$ grid points. To anchor its range, I use the empirical 10th and 90th percentiles of log wages in the SIAB, denoted by d_1 and d_{10}. The first point in $\mathscr{H}$ is normalized to one, $h_1 = 1$. The highest grid point h_{10} is pinned down by the empirical 90–10 wage ratio, i.e.

$$\frac{h_{10}}{h_1} = \frac{d_{10}}{d_1}.$$

The remaining points are spaced equidistantly in log space between h_1 and h_{10}.

Retirement Benefits / Taxes I choose the replacement rate s_b such that, on average, workers receive 40% of the average wage they earned as a retirement benefit b. The linear labor income tax τ is adjusted to exactly balance the government budget, i.e., such that the sum over all retirement benefits equals the sum of all collected taxes.

Internal Calibration The above-described procedure leaves 8 parameters that need to be calibrated internally: The skill accumulation probability decay parameters $\bar{p}_a, \bar{p}_b, \delta_a$ and δ_b, the labor skill substitution parameters α_A and α_B, the baseline utility moving cost k, and the scale of the taste shock σ_ζ. These parameters are chosen to match the life cycle wage profiles of both workers who start their career in A and workers who start their career in B, and the fractions of workers who have never left their first region, again both for workers from A and B, all simultaneously.

Table 2.2 gives an overview of all calibrated parameters.

Table 2.2. Model Parameters

Parameter	Description	Value	Source/Target
Set			
σ	Armington elast.	4	Trade Lit.
λ	IES	2	standard value
β	Discount factor	0.98	standard value
R	Interest Rate	1.03	standard value
$\underline{\eta}$	Subsistence housing	model eq. of 9 m^2	German Regulation
κ	Moving cost	€5,000	Set
Directly Estimated			
$\{v_j\}_{j=0}^{60}$	HH weights	cf. Figure 2.E.2	SOEP
$P(A)$	Birth prob.	0.55	Destatis
p_e	Wealth mass	0.41	SOEP
(μ_e, σ_e)	Wealth shape	$(\log 0.136, 1.544)$	SOEP
$\bar{\eta}_A, \bar{\eta}_B$	Housing supply	$(4.482 \cdot 10^8, 2.940 \cdot 10^8)$	Destatis
ρ_ψ	AR(1) coeff.	0.77	SIAB
$\mathscr{H} = \{h_1, \dots, h_{10}\}$	Skill Grid	$\{1, \dots, 5.687\}$	SIAB
Internal			
γ	cons.-hs. elast.	0.820	22% hs. exp. share
σ_ψ^2	Shock var.	0.38	$\mathrm{Var}(\log w)$
$\bar{p}_a, \bar{p}_b$	skill acc. prob.	(0.75, 0.83)	LC wages/stayer-share
δ_a, δ_b	skill acc. prob. decay	(0.482, 0.672)	LC wages/stayer-share
α_A, α_B	Skill subst.	(0.783, 0.262)	LC wages/stayer-share
k	Utility moving cost	0.180	LC wages/stayer-share
σ_ζ	Taste scale	0.2	LC wages/stayer-share
ϕ	Housing efficiency	1.335	Avg. rent price p.sq.m
s_b	Replacement rate of b	0.339	40% of avg. wage
τ	Tax rate	0.2025	Balanced budget

2.4.4 Model Fit

The left panel of Figure 2.8 shows average log-wage profiles of workers who begin their careers in region A and region B in the model, compared to the same profiles calculated from the SIAB. Both shape and career wage growth in the model fit the data extremely well. Note that neither the model nor the empirical profiles have been adjusted for the regional price level.

The right panel, by contrast, shows wage profiles that have been regionally CPI-adjusted. For the empirical data, I use the regional *BBSR* CPI, as in Section 2.3.1. For the simulated data, I construct an analogous model-internal regional CPI as

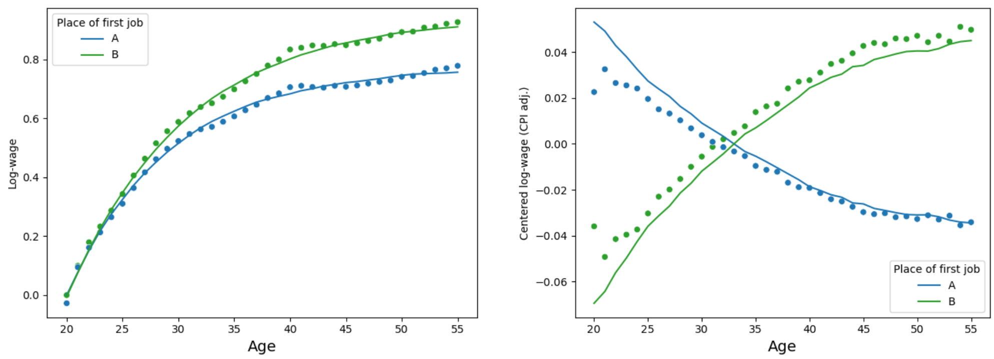
$$CPI(A) = \frac{p_A \cdot q_\eta + 1 \cdot q_c}{p_B \cdot q_\eta + 1 \cdot q_c}, \quad CPI(B) = 1$$

where q_η and q_c are the average quantities of housing and consumption consumed in the whole economy, across both regions. I use economy-wide averages since the *BBSR* index also uses the overall German consumption basket, which is used by Destatis. To better illustrate differences between the regions over the lifecycle, the national average log wage is subtracted from each profile in Panel (b).

Note that only the unadjusted wages shown in Panel (a) are targeted directly. The regional real-wage differences in Panel (b), which are implied by the model's endogenous rent and price structure, nonetheless align closely with their empirical counterparts.

Figure 2.9 shows the fraction of workers who have never left their initial region, by age and region of labor market entry. These empirical profiles are used as cali-

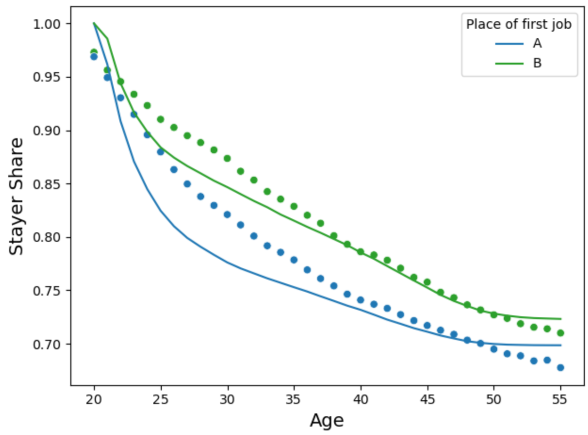
Figure 2.8. Model Fit: Wages



(a) Plain *(b) Mean Subtracted, Regional CPI-adj.*

Notes: The left panel shows average log-wage lifecycles of workers who started their career in A vs B in the model vs the data. The right panel shows the same profiles but with the mean subtracted at every age and adjusted using the regional CPI. For the empirical data, the BBSR CPI is used, the model CPI is calculated using endogenous rent prices and the consumption good as the numeraire. The solid lines represent simulated model outcomes. The dots represent SIAB data estimates.

Figure 2.9. Fraction of Never-Movers



Notes: The solid lines plot the share of workers in the model simulation who have never left their initial region, by their place of labor market entry. The dots plot the same shares in the *SIAB* using the empirical analog of Regions *A* and *B*.

bration targets. The green line represents workers who entered the labor market in region B, the blue line represents workers who entered in region A. The dots show the same statistic in the SIAB using the empirical analog for A and B. The green dots, for example, show the fraction of workers who started their career in a county of the highest rent tercile and have never worked in a county in the lower two terciles. Even though workers move slightly too often in the first 5 years of their lives, the model calibration generates a very realistic mobility level over the lifecycle. In particular, the level difference between workers who began their careers in A vs B is very close to the empirical difference.

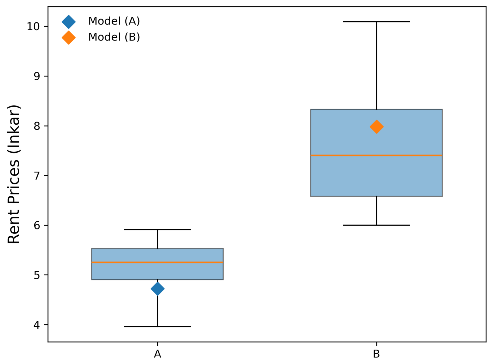
Table 2.3 reports model outcome values of targets that are not already shown in either Panel (a) of Figure 2.8 or Figure 2.9. All targets are met within reasonable tolerance.

Table 2.3. Model Fit

Moment	**Model**	**Target**
Variance of log-wage	0.50	0.52
Housing expenditure share	22%	22%
Avg. per sq. m. rent	€5.22	€5.22
b-replacement rate of avg. wage	41%	40%

Rent prices in the two regions are untargeted. Per-square-meter prices in the model amount to €3.95 in region A and €6.66 in region B. To put these values into perspective, Figure 2.10 compares them to the distribution of average county rent prices faced by workers in the SIAB in 2010, as measured by the INKAR index. The

Figure 2.10. Rent Price Fit



Notes: The boxplots show the distribution of average per-square-meter rents faced by workers in 2010 in counties corresponding to the model's regions *A* and *B*, based on the 2010 INKAR rent index. The boxes represent the interquartile range (25th–75th percentile) of county-level average rents, and the whiskers indicate the full observed range. The orange lines mark the medians, and the diamonds show the model's rent levels, rescaled to 2010 using a national rent deflator (FRED), solely for comparability of price levels.

model rents are reasonably close to the empirical medians and fall well within the support of observed rents.

2.5 Worker Mobility in the Model

2.5.1 Explaining Low Worker Mobility

As demonstrated in Section 2.4.4, workers in region B experience substantially stronger wage growth on average. Why, then, do workers not move toward opportunity more frequently? In this section, I quantify different channels that limit mobility.

In the model, four forces limit workers' mobility across regions. Most directly, there are two types of moving costs: the resource cost and the utility cost. Both act as direct frictions, discouraging relocation by reducing the net benefits of moving. In addition, two further mechanisms become important depending on a worker's career stage.

First, there is the high rent price in region B. This holds back young workers in particular. Young workers have neither accumulated assets nor human capital and earn low wages. This means that they have less to spend on consumption and therefore their housing consumption is relatively close to the subsistence level. As a result, the marginal utility from the larger housing they can afford in region A is particularly high, making moving to region B less attractive.

Second, there is human capital mismatch. This, by contrast, affects mostly older workers. Older workers have, on average, accumulated human capital that is more valuable in the region they have spent the most time in. If they move to the other region, their labor earnings drop because their skills are less productive at producing the local output good. On top of that, the older workers, the lower the probabilities of advancing the skill grids become. After a move, older workers are therefore not only mismatched but additionally have lower probabilities of adjusting their skills to the local labor market.

To quantify these effects, at different stages in life, I simulate the moving costs in terms of money, taking all monetary and non-monetary considerations into account. I implement this by calculating how much money each worker in the model would need to be compensated with, in order to choose to move to the other region, abstracting from the *iid* extreme-value taste-shock.[10] The importance of the different channels can then be measured by simulating how the moving costs change if this channel is shut down. To achieve this, both types of moving cost parameters can be set to zero. Note that this includes moving costs for any future moves, given a state. To eliminate the rent price channel, the rent price in region B is set to the level of region A. The remaining residual after all three frictions are removed reflects the contribution of human capital mismatch and the underlying differences in skill accumulation probabilities between regions. Note that when shutting down any of the cost channels, compensating differentials can become negative, meaning that workers would require a payment to stay in their current region.

As a first step, I simulate moving costs for different age groups, presented in Table 2.4. Overall, average moving costs are about €39, 000. However, for workers in their 20s, moving costs are only just under €16, 000 on average. After that, they increase with age. For workers approaching retirement, the drastic increase in moving costs can be explained by the exogenous rise in utility moving costs $k(j)$. For workers below 45, however, this happens endogenously, through the discussed channels.

I proceed by conducting the discussed decomposition. First, I turn my attention to young workers. The left panel of Table 2.5 shows average moving costs by region for 20-year-old workers. Starting with young workers, the baseline in row (1) shows that moving costs are around €8,000 in both regions. When the resource cost κ is set to zero (row 2), moving costs fall sharply, to about €2,000 in both regions—confirming that out-of-pocket expenses are particularly burdensome for young workers who have little wealth. Eliminating the utility cost $k(j)$ (row 3) yields

10. In the presence of taste shock, it is impossible to pay workers high enough transfers to make them move to the other region with certainty. Instead, I calculate the minimal payment necessary to move workers to the other region with a probability greater than $\frac{1}{2}$, essentially calculating moving costs in the absence of contemporary taste shocks.

Since it is infeasible to calculate these numbers exactly for each state without extreme use of computational time, I approximate the number by selecting the smallest sufficient payment on an exponentially spaced grid between €−37,500 and €1.5 million with 1,000 points. Workers take prices as given, and no general equilibrium effects are taken into account.

Table 2.4. Average Moving Costs by Age Group

	Overall	20–29	30–39	40–49	50–59
Mean cost (€)	39,000	15,900	21,600	25,900	90,500

Notes: The table reports average simulated moving costs by age group, expressed in 1995 euros. Moving costs measure the monetary compensation each worker in the model would require to be indifferent between staying and moving to the other region, taking both monetary and non-monetary factors into account. Calculations abstract from idiosyncratic taste shocks and hold equilibrium prices fixed.

only a modest reduction in moving costs, suggesting that psychic or convenience costs matter less early in life.

Setting rents in region B equal to those in A (row 4) has by far the strongest effect. Without the regional rent difference, moving costs for workers from A drop to €—13,000, on average, indicating that if it wasn't for the high rents in region B, young workers in A had a strong incentive to move. By contrast, workers from B would then require about €14,600 in compensation to move to A—a much higher amount than in the baseline. This highlights that high rents in region B are the key factor deterring young workers in A from moving toward opportunity. At the same time, these high rents make staying in B relatively unattractive for young workers there, offsetting their better wage prospects and resulting in similar baseline moving costs across regions.

When all moving costs and rent differences are eliminated (row 6), workers from A have an even higher willingness to pay to move to B - almost €19,000 on average. Workers from B, by contrast, still require a compensation of €4,500 because moving away slows down their human capital accumulation until they can move back.

In sum, high rent prices in region B are the dominant barrier preventing young workers in A from moving toward better career opportunities. In contrast, workers in B face little incentive to move, as lower rents elsewhere do not compensate for the loss in wage growth potential.

Next, I turn my attention to older workers. The right panel of Table 2.5 shows average moving costs by region for 45-year-old workers. Baseline moving costs are much higher at 45 than at 20—about €24,000 in region A and €27,000 in region B. When the resource cost κ is removed (row 2), moving costs fall by roughly one quarter, indicating that out-of-pocket expenses remain relevant but are no longer the main driver. Eliminating the utility cost $k(j)$ (row 3) reduces costs by a similar magnitude.

Equalizing rents between regions (row 4) leads to a modest reduction in moving costs for workers in A but a strong increase for workers in B. At age 45, workers have already accumulated wealth and typically rent larger amounts of housing. As

Table 2.5. Decomposition of Moving Costs by Region and Age (in €)

	Age 20		Age 45	
	Region A	Region B	Region A	Region B
(1) Baseline (all frictions)	8,400	8,100	24,400	27,400
(2) No resource cost	2,100	1,900	18,200	21,200
(3) No utility cost	6,700	6,600	15,400	16,500
(4) Equal rent prices ($p_B = p_A$)	-13,000	14,600	20,300	32,700
(5) No resource or utility cost	500	400	9,200	10,300
(6) No resource or utility costs and $p_B = p_A$	-18,900	4,500	5,500	13,600

Notes: Entries show the average monetary compensation (in euros) required to make a worker move from their current region to the other. Columns report averages for regions *A* and *B* at ages 20 and 45. Each counterfactual removes one or more frictions: (2) resource cost κ, (3) utility cost $k(j)$, (4) rent price differential, (5) both costs simultaneously, and (6) all aforementioned frictions combined.

a result, the marginal utility from additional housing space is smaller than for young workers, and higher rents in region *B* no longer constitute a major barrier to moving.

Setting both the resource and utility costs to zero (row 5) cuts moving costs by about two-thirds relative to the baseline, yet a considerable residual remains. Even when all frictions—including rent differences—are removed (row 6), moving costs remain at roughly €5,500 in *A* and €13,600 in *B*. These residual costs capture the effect of human capital mismatch: older workers have accumulated skills that are region-specific, and moving implies a loss in productivity and slower future accumulation in the new local labor market.

It is now clear why workers move so little over the lifecycle. Workers in region *A* face a dilemma. While they are still young, they would, in principle, like to move to region *B* to benefit from faster human capital accumulation. But with little wealth and low entry-level wages, high rents make moving to *B* unattractive. Their only option is to remain in *A* and save until, with higher wealth levels, they can more easily afford the high rents.

Workers in region *B*, by contrast, have little incentive to move to region *A* in search of career opportunities. They endure their early years with limited wealth and high rents but are later compensated through faster wage growth and stronger career progression.

In sum, mobility is not low in the model because moving is extremely costly, but because relatively minor costs at the beginning of young workers' careers lock them into a path. High rents prevent young workers from moving when it would be most beneficial. As they age, the accumulation of region-specific human capital locks workers into their regions.

Table 2.6. Average Log Wage Differences: Mincer Prediction vs. Model Counterfactual

	Region A		Region B	
	Mincer Prediction	Model Counterfactual	Mincer Prediction	Model Counterfactual
Mean	0.104	-0.275	-0.229	-0.455
Median	0.113	-0.310	-0.233	-0.473

Notes: The table reports the mean and median log wage differences between simulated model wages and predicted wages in the other region from the Mincer regression and from the structural model counterfactual, separately by region. Positive values indicate that predicted wages exceed observed wages.

2.5.2 Potential Wages

The moving costs calculated in the previous section are quite low, compared to some literature estimates. To illustrate why, I compare potential earnings in the other region in the model with predictions a basic Mincer regression produces when fitted to the model simulations. After all, Mincer regressions are a standard tool in the literature to estimate potential earnings. Through the lens of the model, however, a basic Mincer Regression struggles to produce accurate predictions for potential earnings in the other region.

The most standard Mincer regression with a regional fixed effect takes the form

$$\log \text{wage}_{ij} = \delta + \beta_1 \text{schooling}_i + \beta_2 \text{experience}_{ij} + \beta_3 \text{experience}_{ij}^2 + e_{ij},$$

where $\delta_{z(ij)}$ is the region fixed effect. In the model, schooling does not feature explicitly. However, workers enter the economy with different levels of persistent productivity ψ_{i0}. I therefore use workers' initial productivity to proxy education and let it enter as a fixed effect. Experience is given by age. I therefore estimate

$$\log \text{wage}_{ij} = \alpha_{\psi_{i0}} + \delta_{z(ij)} + \beta_1 \cdot j + \beta_2 \cdot j^2 + e_{ij}. \tag{2.8}$$

Table 2.6 shows the mean and median difference between a worker's current log-wage and the log-wage predicted for this worker in the other region by the fitted model given in Equation 2.8 for 45-year-old workers. For comparison, the mean and median of the actual model counterfactual, given a worker's state, are included as well. The model counterfactual takes into account regional prices by using the model-internal regional CPI discussed in Section 2.4.3 to adjust wages.

In the case of workers from region A, the Mincer Regression overestimates potential earnings by a large margin. On average, the regression predicts a sizable wage of over 10%, whereas the actual counterfactual implies a real wage of 0.275 log-points. This is because naively estimating Equation 2.8 fails to acknowledge that general experience does not capture the level of the skill required in region B. Additionally, it does not take the high rents into account.

Likewise, the regression model underestimates the wage loss of workers from region B, because it fails to take the regional nature of human capital into account.[11]

An econometrician facing the simulated model data who assumes that experience can be fully utilized after a move to another region may estimate a model similar to Equation 2.8 and conclude that there are large unrealized gains from moving to region B. Such an interpretation would suggest that high moving costs are necessary to explain why workers from region A remain in their home region. Taking regional human capital and living costs into account, however, these large incentives do not exist through the lens of the model, where regional human capital and high living costs are the main drivers of low mobility. In this sense, the model's implications echo the findings of Kambourov and Manovskii (2009), who show that human capital is highly occupation-specific. Acknowledging the regional concentration of occupations, therefore, becomes crucial in explaining low mobility.

2.6 Housing Supply Policy

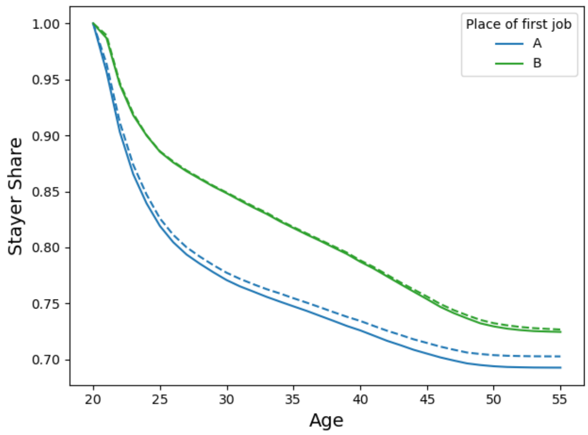
In Section 2.5.1, I discussed two main findings. First, moving typically only makes economic sense for young workers, since they have not yet accumulated much region-specific human capital and therefore face little skill mismatch after a move. Second, the main factor preventing young workers in region A from moving to region B is the high rent level. In light of the limited housing supply, older workers who occupy large housing units do not internalize the effect of their housing consumption on the ability of young workers to move to region B. This section presents a policy experiment aimed specifically at mitigating this effect by expanding the housing stock in region B. In particular, I let the government increase the housing stock in region B through public investment.

According to data from the Federal Statistical Office (Destatis),[12] constructing one square meter of housing in 1995 cost €1,254, and a square meter of undeveloped building land cost €58.02. I therefore use the model equivalent of €1,312.02 as the price for building an additional efficiency unit of housing. The construction project is assumed to be fully financed through government debt. Interest payments at the exogenous model rate R are covered by adjustments to the labor income tax τ. The government remains the owner of the newly constructed housing and receives rental income from renting it out at the market rent price.

11. Note that this wage loss does not capture that a given level of wealth is worth more in real terms in Region A than in Region B, mitigating any wage losses when moving to Region A.

12. Specifically, I take the building price from the Destatis report "Bauen und Wohnen, Baugenehmigungen, Baukosten" (Construction and Housing – Building Permits and Construction Costs) and the land price from "Fachserie 17 / Reihe 5 Preise Kaufwerte für Bauland" (Series 17 / Series 5: Prices – Purchase Prices for Building Land).

Figure 2.11. Fraction of Never-Movers (Counterfactual)



Notes: The solid lines plot the share of workers in the model simulation who have never left their initial region, by their place of labor market entry after a 3% increase in the housing stock in region *B*. Dashed lines show the same fractions in the baseline economy.

I study a construction project that increases the housing stock in region B by 3%. The magnitude is motivated by large-scale real-world construction projects. For example, an ongoing project in the city of Mannheim[13] is projected to provide housing for about 10,000 people—a roughly 3% increase in the city's population. This policy directly lowers the rent price in region B, which falls to €6.51 compared to €6.66 in the baseline economy. As a result, moving to region B becomes much more attractive, especially for young workers. Figure 2.11 shows the fraction of workers who have never moved by origin region. The baseline economy is depicted by dashed lines, and the post-policy economy by solid lines. Clearly, mobility out of region A increases. Consequently, rents in region A adjust slightly downward as well—to €3.93 compared to €3.95 in the baseline—reflecting lower local housing demand.

In addition to the rent effects, the policy also influences wage levels. Since the two regions produce separate intermediate goods, reallocating labor from region A to region B raises the price of labor in region A and lowers it in region B. This wage adjustment is amplified by the fact that more workers in region B now accumulate human capital more rapidly. As a result, wages per efficiency unit rise by 0.18% in region A and fall by 0.18% in region B.

The faster human capital accumulation of a larger number of workers in region B raises the overall effective labor supply in the economy, which increases by about 0.05%. Aggregate output rises as well, but only modestly, since more labor is now devoted to producing the intermediate good from region B, which exhibits di-

13. The project is frequently called the Franklin-Quartier project in the press coverage, as it is located in the Franklin district of Mannheim.

Table 2.7. CEV after Housing Construction

Decomposition	Veil of Ignorance	Conditional on A	Conditional on B
Full Policy	0.80%	0.76%	0.85%
Δp_A	0.07%	0.12%	0.03%
$\Delta \bar{w}_A$	0.15%	0.23%	0.04%
Δp_B	0.41%	0.17%	0.71%
$\Delta \bar{w}_B$	-0.12%	-0.05%	-0.22%
$\Delta \tau$	0.29%	0.29%	0.29%

Notes: Consumption equivalent variation (CEV) of newborn workers with respect to a counterfactual economy where the government increases the housing stock in region B by 3%. The decomposition in rows 2–6 isolates the contribution of each general-equilibrium price change by resetting one variable at a time to its baseline level and recomputing the resulting CEV.

minishing returns in the final output production. Overall, output in the new steady state is 0.03% higher than in the baseline.[14]

Next, I analyze how these changes translate into welfare effects. Table 2.7 reports the consumption equivalent variation (CEV) of newborn workers relative to the baseline. Behind the veil of ignorance, newborn workers experience a sizeable welfare gain of 0.8%. Conditional on being born in region B, the welfare gain rises slightly to 0.85%, as these workers directly benefit from the larger housing supply. The CEV for workers born in region A, at 0.76%, is still substantial.

The decomposition in Table 2.7 reveals that the rent reduction in region B is the dominant welfare channel. Its effect is particularly strong for workers born in B, amounting to 0.41% CEV. This benefit is partially offset by the negative wage effect in region B of about -0.12%. For workers from region A, the changes in local rents and wages both have positive effects (0.12% and 0.23%, respectively). Importantly, however, the lower rent in region B contributes even more—around 0.7% CEV—reflecting the welfare gain from being able to move more easily to the high-opportunity region.

Additionally, the construction project is fiscally profitable. Government rental income from the new housing exceeds the interest payments on the construction debt, resulting in a further welfare gain of 0.29% CEV through lower taxes.

Finally, Table 2.8 presents welfare effects by endowment tercile. The gains are largest for workers in the lowest tercile, particularly those born in region B. Since poorer households spend a larger share of their income on housing, they benefit disproportionately from lower rents. Although the policy raises welfare for all groups, it is most beneficial to poorer workers.

For comparison, I also consider an alternative policy aimed at reducing housing demand rather than expanding supply. Specifically, I examine a downsizing incen-

14. GDP is measured as total output Y, excluding construction activity.

Table 2.8. CEV by Endowment

Endowment Tercile	Veil of Ignorance	Conditional on A	Conditional on B
1	0.80%	0.79%	0.77%
2	0.76%	0.75%	0.75%
3	0.85%	0.84%	0.79%

Notes: Consumption equivalent variation of newborn workers by endowment tercile with respect to a counterfactual economy where the government increases the housing stock in region B by 3%.

tive similar to programs discussed in the public debate, where retired households receive transfers if they move to smaller dwellings. In the model, however, this policy proves largely ineffective. It induces only marginal reductions in housing consumption among retirees, with negligible effects on rents, wages, or mobility. Consequently, the welfare effect is slightly negative, as the fiscal cost of the subsidy outweighs its benefits. The details of the experiment are described in Appendix 2.G.2.

2.7 Conclusion

In this study, I revisit the question of why workers are not more mobile despite large regional wage differences. Using high-quality administrative data from Germany, I document that regional wage differences manifest themselves not in higher entry-level wages but in larger wage growth over the lifecycle. I show that occupations are regionally concentrated and that large parts of the variation in local wage levels and growth can be accounted for by regional occupation compositions.

Motivated by these findings, I propose a two-region general equilibrium model with local labor and rent markets. In the model, different occupation compositions are represented by different skills that are differentially productive in the two regions and are learned stochastically through experience, as well as different intermediate output goods being made in each region. In one of the regions, workers specialize in a skill that can be learned faster, and experience greater wage growth over their careers. Workers can move between the regions every year, draw utility from consumption and housing, and are subject to a borrowing constraint. I calibrate the model to the data, and I find that for young workers, the high rent prices in the high-wage region are a major reason not to move there. Older workers have, on average, accumulated human capital that is more productive in their current region and have no incentives to move. I find that average moving costs, measured by the amount of compensation needed to move workers, are only about €39,000.

I conduct a policy experiment in which the government pays for the construction of 10% more housing in the high-wage region. I find that this policy leads to more mobility into the high-wage region. In particular, it is welfare-improving for workers from both regions. Workers in the high-wage region value the decreasing rent prices.

Workers from the low-wage region experience wage increases and moderate rent decreases. On top of that, they benefit from increased possibilities to move to the high-wage region for a better career. A similar policy, incentivizing retired workers to downsize their housing consumption, fails to increase welfare.

For future research, there are several ways forward. For once, extending the model to include home-owners is interesting, since home-owners are even more reluctant to move and, in particular, dislike falling house prices that could be induced after construction projects. Another interesting idea could be opening the model economy to international trade. This could alter general equilibrium effects because of labor re-allocation, since international demand for one of the intermediate output goods might be greater. Finally, introducing a notion of differential talent to speak to early life skill mismatch could be interesting to introduce even stronger inequality of opportunity. It could also lead to larger GDP responses after mobility-inducing policies. I view this as fruitful ground for future research.

2. Appendix

2.A Data

2.A.1 Wage Imputation

This appendix describes the procedure used to impute wages that are top-coded at the social security contribution ceiling. I follow Böhm, Gaudecker, and Schran (2024), in turn following Dustmann, Ludsteck, and Schönberg (2009), and implement a Tobit model for (log) daily wages with right-censoring separately for each year. Let $y_i^* = x_i'\beta + \varepsilon_i$ denote the latent log wage with $\varepsilon_i \sim \mathcal{N}(0, \sigma_g^2)$, where g indexes an education × age-group cell and a three-level bin of the regional rent distribution. Because I later split the sample by the same three rent bins, I perform the imputation separately for these bins. This ensures that the upper-tail fit is not distorted by systematic regional cost differences that correlate with the censoring threshold and wage structure. Importantly, however, my results are almost unchanged if I instead pool regions and omit the rent-bin split at the imputation stage.

The observed outcome is $y_i = \min\{y_i^*, c_t\}$, where c_t is the time-t contribution ceiling mapped to daily real wages. I estimate (β, σ_g) by maximum likelihood using all observations, weighting each record by employment spell length. The covariate vector x_i includes age, the individual's censoring rate, the average wage from other years (to capture persistent individual effects), and education. Allowing the wage variance to differ by education × age-group cell follows Dustmann, Ludsteck, and Schönberg (2009) and improves fit for the upper tail.

I then replace censored observations with simulated draws from a truncated normal distribution, using the Tobit prediction as the mean and the estimated Tobit variance as the dispersion parameter.

2.A.2 INKAR Rent Data

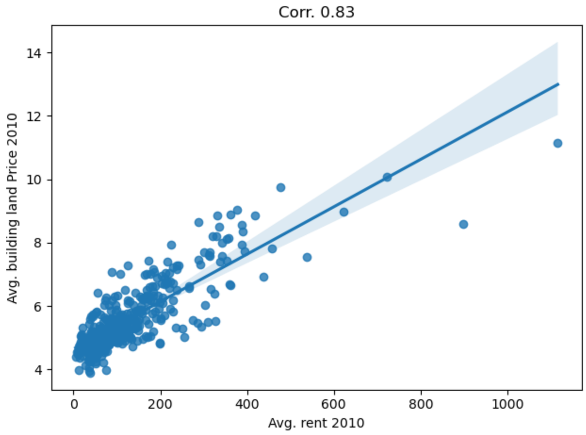
The INKAR rent index is only published from 2010 onwards. To assert that it can still be used to measure rent-terciles in earlier years, I turn to building land prices.

Figure 2.A.1 shows the correlation of rents as measured by the INKAR rent index with average building land prices at the county level in a scatter plot. Building land prices have been obtained from the Regional Statistics Database (Regionalstatistik), a database published by the statistics offices of the German Federal States. The correlation is 83%.

Local building land prices can be traced back much further than local rents. At the county level, they are reported from 1995 onwards. Before that, they are published for Germany's major cities in the annual report of the Federal Statistics Office (Destatis). Panel (a) of Figure 2.A.2 shows the auto-correlation of building land prices at the county level between 1995 and 2010. Panel (b) shows the same correlation for the major cities reported in the report of 1980. In both cases, auto-correlation is about 90%.

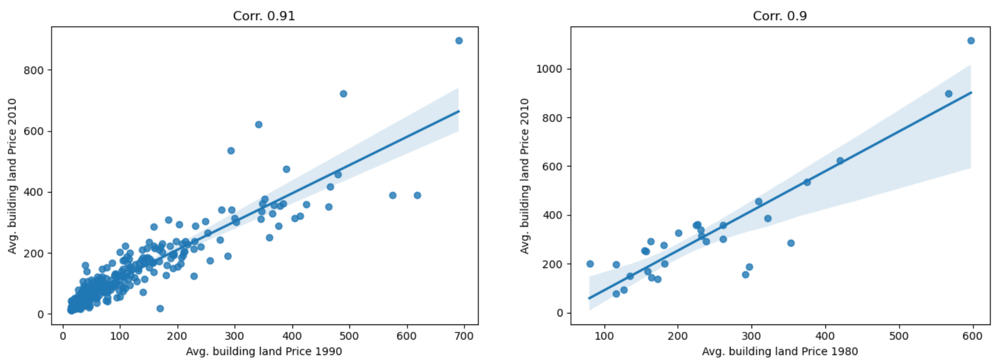
This provides reassurance that rents are equally persistent over time and the grouping into terciles works. Additionally, I conduct robustness checks with alternative data based on a German rent subsidy scheme reported in Appendix 2.C.

Figure 2.A.1. Correlation of County Rents and Building Land Prices in 2010



Notes: Scatter plot of average offered rent per m^2 (INKAR, 2010) against building land prices from the Regional Statistics Database (Regionalstatistik), maintained jointly by the Federal Statistical Office of Germany (Destatis) and the statistical offices of the federal states. Each point is a county. The fitted line shows the OLS regression of log rent on log building land price.

Figure 2.A.2. Persistence of Building Land Prices Over Time



(a) 1995–2010 *(b) 1980–2010*

Notes: Each panel plots the building land price in the initial year on the horizontal axis and the price in 2010 on the vertical axis for all counties. The fitted line is OLS with slope reported in the panel. The prices after 1995 are from the Regional Statistics Database (Regionalstatistik), maintained jointly by the Federal Statistical Office of Germany (Destatis) and the statistical offices of the federal states. The 1980 prices come from the statistical Yearbook (Statistisches Jahrbuch) of the Federal Statistical Office of Germany (Destatis) and include only selected cities.

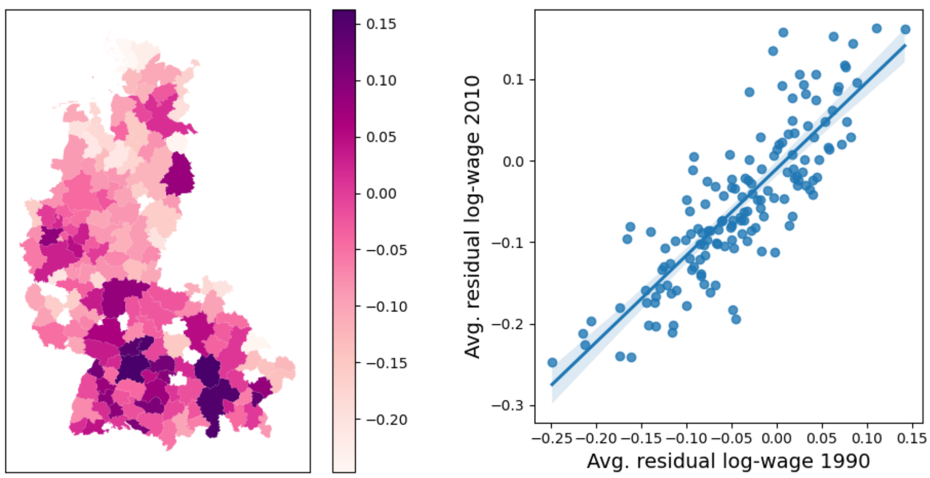
2.B Additional Empirical Results

Figure 2.B.1 shows a map of average log wages that have been residualized by regressing on age and six levels of education in the left panel. The right panel shows the auto-correlation between 1990 and 2010 of the same average residuals.

Figure 2.B.2 shows the share of workers who never moved to any other commuting zone and the share of workers who never moved further away from their first commuting zone than to a neighboring commuting zone for workers who begin their careers in commuting zones of the lowest wage tercile. Commuting Zone average wages are used to calculate these terciles.

Figure 2.B.3 and Figure 2.B.4 show the results for the 1960-1964 cohorts in Figure 2.4 in the main text for the 1950-1954 and the 1955-1959 cohorts, respectively.

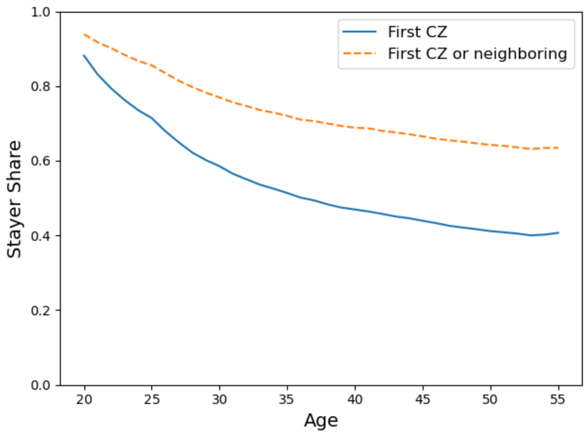
Figure 2.B.1. Local Wage Level and Persistence - Residualized



(a) Avg. Log Wage 2010 *(b) Wage Level Persistence*

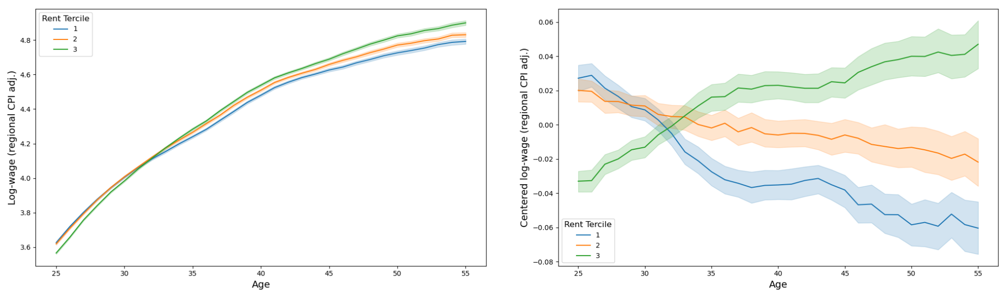
Notes: The left panel maps commuting zones in West Germany, shaded by the average residual log daily wage in 2010. Residual wages are obtained from regressions of log daily wages on age and education fixed effects (six education levels) estimated using the SIAB. Darker shades indicate higher values. The right panel plots the average residual log daily wage in 2010 (vertical axis) against the average log daily wage in 1990 (horizontal axis) with an OLS fit. Each point represents one commuting zone. The sample is restricted to full-time workers.

Figure 2.B.2. Fraction of Never-Movers



Notes: The sample has been reduced to workers who begin their career in commuting zones of the lowest income tercile. The solid line plots the share of workers in the SIAB sample who have never left their initial commuting zone, by age. The dashed line plots the share of workers who have either remained in their initial commuting zone or moved only to an adjacent commuting zone.

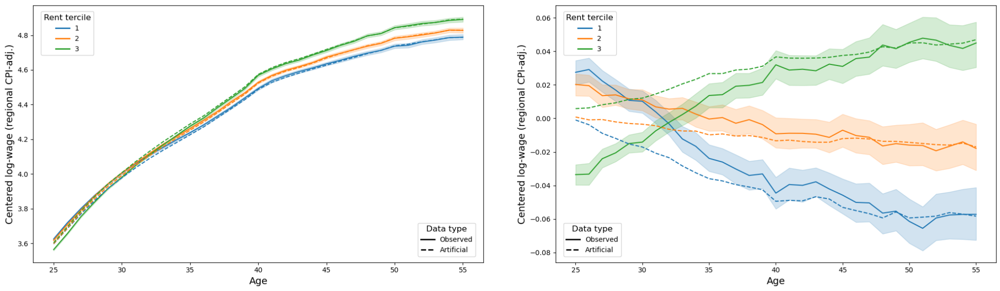
Figure 2.B.3. Log Wage Lifecycle Profiles



(a) Plain *(b) Mean Subtracted*

Notes: The left panel shows average log wages by age for workers who began their careers in counties belonging to the first, second, and third rent tercile, respectively. Rents are measured using INKAR data from 2010. The green line corresponds to the highest, the blue line to the bottom, and the orange line to the middle rent tercile. Wages are regionally CPI-adjusted to the level of Bonn using the BBSR CPI. The right panel shows the same wage profiles after subtracting the overall mean at each age. The confidence bands shown are at the 95% level. The sample includes workers born between 1950 and 1954.

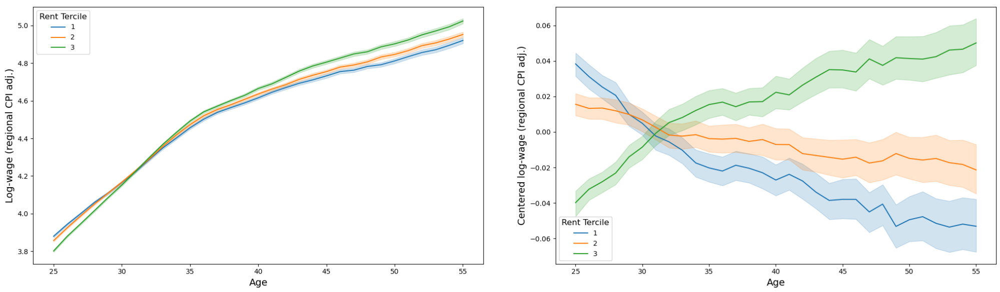
Figure 2.B.5. Observed vs. Counterfactual Log Wage Profiles



(a) Plain *(b) Mean Subtracted*

Notes: The solid lines show the average log wage profiles for workers in full-time employment who began their careers in counties belonging to the first, second, and third rent tercile, respectively. The dashed lines show counterfactual log wage profile that results from assigning every worker the national average wage, conditional on occupation and age. The green line corresponds to the highest, the blue line to the bottom, and the orange line to the middle rent tercile. Wages are regionally CPI-adjusted to the level of Bonn using the BBSR CPI. The right panel shows the same wage profiles after subtracting the overall mean at each age. Rents are measured using INKAR data from 2010. The sample includes workers born between 1950 and 1954.

Figure 2.B.4. Log Wage Lifecycle Profiles



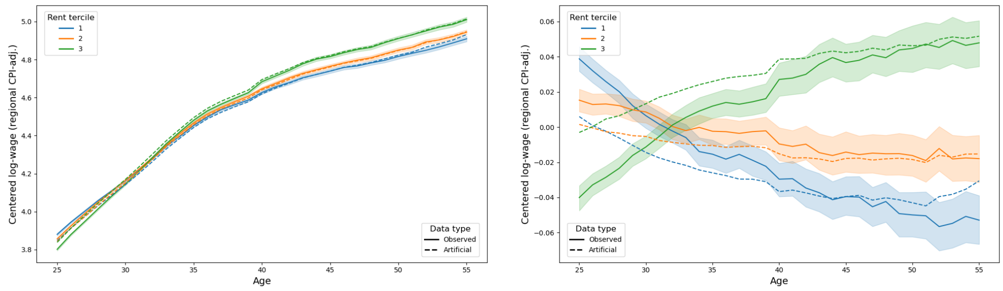
(a) Plain *(b) Mean Subtracted*

Notes: The left panel shows average log wages by age for workers who began their careers in counties belonging to the first, second, and third rent tercile, respectively. Rents are measured using INKAR data from 2010. Wages are regionally CPI-adjusted to the level of Bonn using the BBSR CPI. The right panel shows the same wage profiles after subtracting the overall mean at each age. The sample includes workers born between 1955 and 1959.

Figures 2.B.5 and 2.B.6 show the results reported in Figure 2.7 in the main text for workers born 1950-1954 and 1955-1959, respectively.

Figure 2.B.7 shows an interesting additional result. The fraction of workers who change their occupation when they move to another commuting zone is quite high - about 50%. This highlights the importance of regional occupational composition.

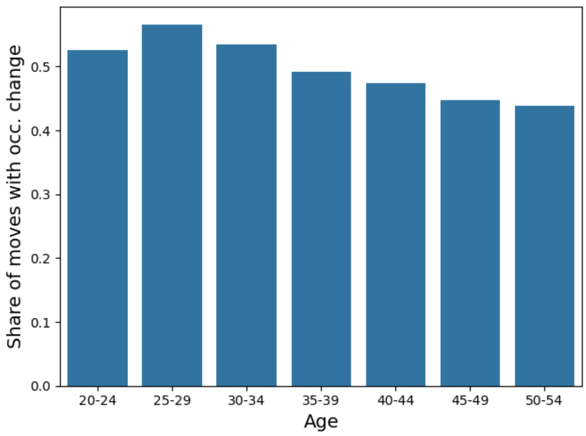
Figure 2.B.6. Observed vs. Counterfactual Log Wage Profiles



(a) Plain *(b) Mean Subtracted*

Notes: The solid lines show the average log wage profiles for workers in full-time employment who began their careers in counties belonging to the first, second, and third rent tercile, respectively. The dashed lines show counterfactual log wage profile that results from assigning every worker the national average wage, conditional on occupation and age. The green line corresponds to the highest, the blue line to the bottom, and the orange line to the middle rent tercile. Wages are regionally CPI-adjusted to the level of Bonn using the BBSR CPI. The right panel shows the same wage profiles after subtracting the overall mean at each age. Rents are measured using INKAR data from 2010. The sample includes workers born between 1955 and 1959.

Figure 2.B.7. Occupation Change Shares Among Movers



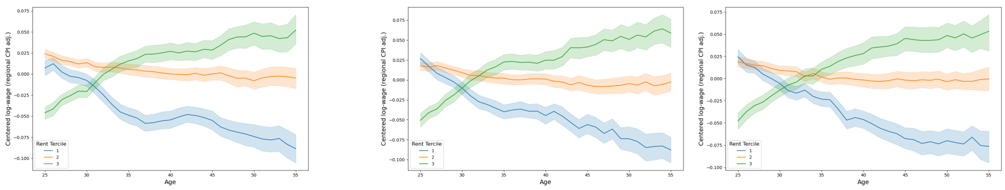
Notes: Fractions of workers who change their occupation among all workers who move to a different commuting zone by age group. Sample pools birth years 1950-1964.

2.C Alternative Rent Data (Wohngeld)

Figure 2.C.1 shows the same results shown in the right panels of Figures 2.4, 2.B.3, and 2.B.4 but with a different way of splitting the sample. In Germany, low-income households are eligible for a rent subsidy. Crucially, it depends on the local rent price of non-expensive housing. I obtained these categories from the Federal Statis-

tics Office with the caveat that some counties are missing, leading me to use the INKAR data for the main specifications shown in the main text. To this end, all German counties are grouped into 6 categories. I group counties into three bins of two categories each, and reproduce the log-wage profiles I obtained using rent terciles based on the INKAR rent index. The obtained results are extremely similar, lending confidence that the results are indeed correct.

Figure 2.C.1. Log Wage Lifecycle Profiles - Rent Subsidy Category

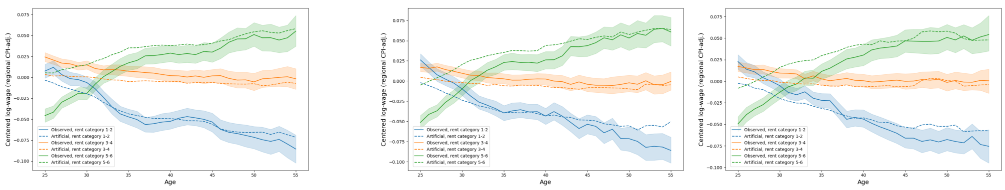


(a) 1950-1954 *(b) 1955-1959* *(c) 1960-1964*

Notes: Log wages by age for workers who began their careers in counties belonging to the three rent bins. Rents are measured using the categories of the subsidiy scheme *Wohngeld* that groups counties into 6 bins ranging from cheap to expensive. The green line corresponds to the highest, the blue line to the bottom, and the orange line to the middle 2 rent bins. Wages are regionally CPI-adjusted to the level of Bonn using the BBSR CPI. All panels show the same wage profiles after subtracting the overall mean at each age. The confidence bands shown are at the 95% level. Each panel shows results for different 5-year cohorts.

Figure 2.C.2 reproduces the counterfactual occupation-weighted wage profiles shown in Figure 2.7 in the main text. Results are extremely similar again.

Figure 2.C.2. Observed vs. Counterfactual Log Wage Profiles - Rent Subsidy Category



(a) 1950-1954 *(b) 1955-1959* *(c) 1960-1964*

Notes: Log wages by age for workers who began their careers in counties belonging to three rent bins. Rents are measured using the categories of the subsidiy scheme *Wohngeld* that groups counties into 6 bins ranging from cheap to expensive. The dashed lines show counterfactual log wage profile that results from assigning every worker the national average wage, conditional on occupation and age. The green line corresponds to the highest, the blue line to the bottom, and the orange line to the middle 2 rent bins. Wages are regionally CPI-adjusted to the level of Bonn using the BBSR CPI. All panels show the same wage profiles after subtracting the overall mean at each age. The confidence bands shown are at the 95% level. Each panel shows results for different 5-year cohorts.

2.D Solving the Model

2.D.1 Solving the Model with EGM

The worker problem is solved by means of backwards induction. This can often be done efficiently using an endogenous grid method (EGM) as in Carroll (2006). However, since the location choice in the model is discrete, I cannot use the standard EGM. Instead, I rely on the endogenous grid method for discrete-continuous dynamic choice models proposed by Iskhakov et al. (2017).

In the last period of the lifecycle, first-order conditions can be used to analytically solve for the worker decisions. In the periods between the working phase and the final period, the inverse Euler equation of retired workers is

$$c = \left[R\beta c'^{-\lambda} \left(\frac{\nu'}{\nu} \right)^{\lambda} \right]^{-1/\lambda}$$

where primes indicate state in the next period. Worker decisions can then be found using standard EGM methods.

During the working phase, the discrete choice needs to be handled. Every period, for every possible state, the worker's problem is solved twice, taking the location choice (i.e., z') as given. Using the worker's inverse Euler equation

$$c = \Bigg[R\beta \int \Bigg(\sum_{h'_a \in \mathcal{H}} \sum_{h'_b \in \mathcal{H}} \sum_{z'' \in \{A,B\}} P(z'' \mid \psi', h'_a, h'_b, z', \psi') P(h'_a \mid h_a) P(h'_b \mid h_b) \\ c'(z'', h'_a, h'_b, \psi')^{-\lambda} \left(\frac{p_z}{p_{z'}} \right)^{(1-\gamma)(1-\lambda)} \left(\frac{\nu'}{\nu} \right)^{\lambda} \Bigg) \mathrm{d}F(\psi' \mid \psi) \Bigg]^{-1/\lambda}.$$

and given a grid of consumption in the next period, implied through next period's already determined consumption policy function by an exogenous grid of assets in the next period, an endogenous grid of consumption in the current period can be calculated. Note that variables with a prime denote state variables in the next period, and z'' denotes the location choice in the period after the next period. Given the functional form of u_j, grid points corresponding to the endogenous contemporary consumption grid of contemporary housing consumption η and subsequently the saving decision, as well as implied conditional value functions V_A and V_B, can be calculated. The borrowing constraint is handled by adding a point with $a = 0$ and implied consumption decisions into the respective grids where workers would like to borrow if they could.

Given that the taste shocks ζ_j are *iid* and *ETV*1 distributed, the probabilities of choosing A or B can be calculated in closed form at every grid point with the famous logit probability formula (Domencich and McFadden (1975)).[15] Policy functions are

15. As discussed by Iskhakov et al. (2017), this procedure can produce non-monotonous points in the conditional value grids that need to be eliminated through an upper envelope.

then interpolated linearly and passed on to the period before the contemporary period.

In the 0-period after birth, the same logit formula can be applied to obtain choice probabilities for each region, taking initial productivity and endowment as given.

This backward induction procedure yields value functions and workers' decisions for given prices $\overline{w}_A, \overline{w}_B, p_A$ and p_B. Since local housing stocks are fixed and $\overline{w}_A$ and $\overline{w}_B$ are closed form function of $\frac{L_A}{L_B}$ solving the model then is a matter of finding $\frac{L_A}{L_B}$ such that the implied wages are profit maximizing for the aggregation firm and result from the workers decision problem given these same wages and rent prices that clear rent markets at the same time.

Aggregate variables $L_A, L_B, \eta_A^d, \eta_B^d$, where the latter two denote local aggregate housing demand, are obtained by simulating 100,000 individuals using the calculated worker decisions.

The fixed-point problem of finding $\frac{L_A}{L_B}, p_A$ and p_B is then solved by updating initial guesses in nested loops. $\frac{L_A}{L_B}$ is updated using Anderson acceleration, and the rent prices are updated with a log-linear updating rule.

2.D.2 Solving the Model under Downsizing

Under the policy discussed in section 2.G.2, EGM is no longer applicable for retired workers in region B because there can be corner solutions. Workers who would otherwise have consumed slightly more than $71.02m^2$ may find it optimal to consume exactly the threshold. Similarly, with the transfer, workers who would have rented housing just below this threshold may now find the threshold optimal. I therefore use a standard grid-search method where for every asset grid point and every level of consumption on a grid, the value of choosing freely above the threshold, below the threshold, and the corner, as well as the resulting savings decision solution, are compared. Policy functions are then interpolated on the same asset grid used in the EGM method described in the previous section. For non-retired workers and workers retired in A, the same EGM algorithm as before is employed.

2.E Calibration

2.E.1 Male Income Shares

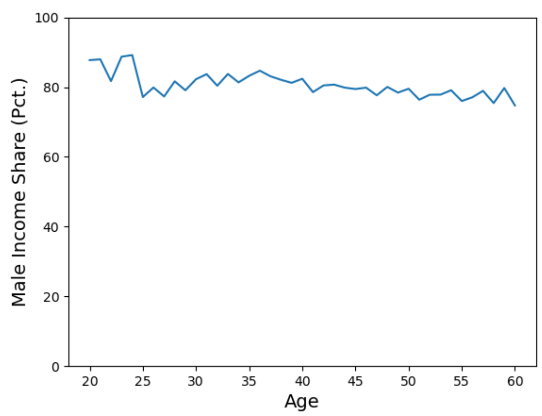
To document how the within-household income distribution evolves with age, I use the German Socio-Economic Panel (SOEP) to construct the age profile of the male income share shown in Figure 2.E.1. I merge the individual-level labor income records with demographic and household identifiers. All monetary amounts are deflated using the German consumer price index from Jordà, Schularick, and Taylor (2017).

For each household and survey year, I compute individual ages as the difference between survey year and year of birth, dropping implausible values where the implied age is negative. To identify the household head, I restrict attention to males with positive earnings and select, within each household, the oldest such male. As in my main empirical analysis, I restrict the analysis to households whose head was born between 1950 and 1964. I then restrict the sample to households with exactly one adult male earner and a head aged 60 or below.

For each remaining household-year, I aggregate real wages by gender and compute the within-household share of male earnings.

Figure 2.E.1 shows the resulting profile. The male income share is close to 100 percent for younger heads and gradually declines with age as female earnings and joint retirement patterns become more prevalent. Only households with a single working-age male are included, so that the pattern captures variation in the relative contribution of female partners rather than compositional shifts in household structure.

Figure 2.E.1. Average Male Income Share by Age of Household Head



Notes: The figure plots the average share of male earnings in total household labor income by the age of the male household head. The sample is restricted to SOEP households with exactly one adult male earner born between 1950 and 1964 and with positive household labor income.

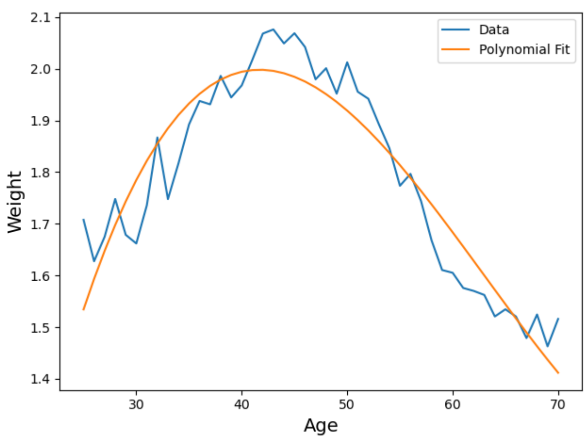
2.E.2 Household Size Weights

Household consumption is scaled using the modified OECD equivalence scale. To construct the underlying household information, I use the German Socio-Economic Panel (SOEP). I merge the individual-level labor income records with household identifiers and survey years and adjust wages using the consumer price index to obtain real values. For each household, I identify the male head as the oldest working-age male with positive earnings, and restrict the sample to cohorts born between 1940 and 1964.

For these households, I calculate the number of adults and children and then assign an equivalence weight following the OECD method, where additional adults and children increase household needs by fixed fractions relative to the first adult. I then average these weights by the age of the household head, which provides an age profile of equivalence scales.

Because the raw averages are somewhat irregular, I fit a smooth polynomial curve to the values between ages 25 and 70. For ages below 25 and above 70, the curve is extended by holding the fitted values constant at the boundaries. The resulting schedule gives a smooth and age-dependent measure of household equivalence weights that I feed into the model. Figure 2.E.2 shows the resulting weights.

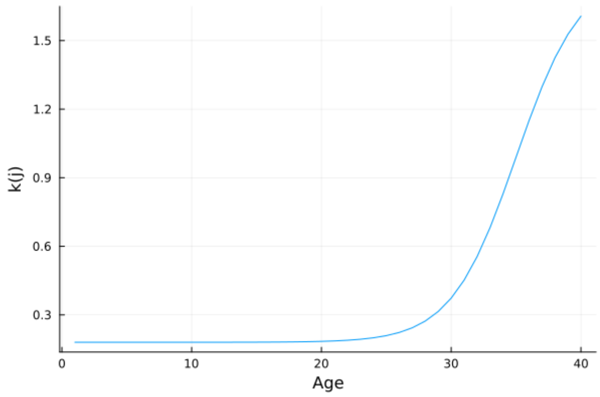
Figure 2.E.2. Modified OECD Equivalence Weights



Notes: The blue line shows the modified OECD equivalent weights of households in the SOEP with exactly one male adult born between 1940 and 1964 against age. The orange line is the cubic polynomial fit that is fed into the model.

2.E.3 Utility Moving Cost

Figure 2.E.3. Utility Moving Costs



Notes: Calibrated values of utility moving costs $k(j)$ by age.

The functional form of $k(j)$ is set to

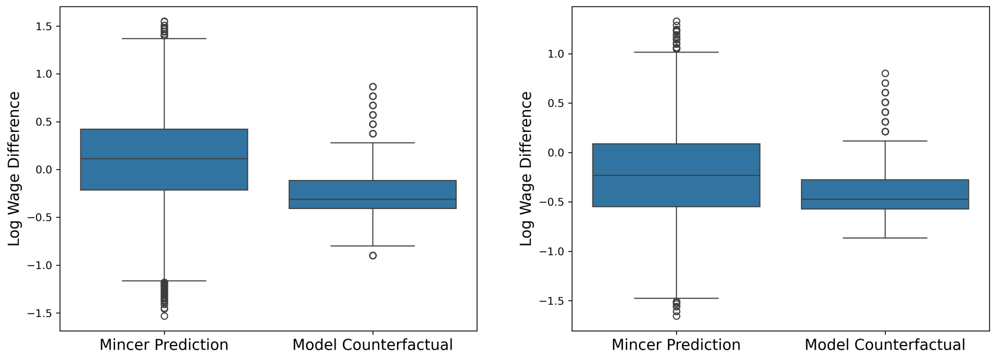
$$k(j) \;=\; k\Big[\,1 + (S-1)\,\sigma\big(2\beta\,(j-\mu)\big)\Big], \qquad s(j) \;=\; \sigma\big(2\beta\,(j-\mu)\big).$$

where $\mu = 35$, $\beta = 0.2$ and $S = 10$. The resulting calibrated values are shown in Figure 2.E.3.

2.F Additional Model Results

Figure 2.F.1 shows the distribution of predicted and model counterfactual wage changes from working in the other region discussed in section 2.5.2.

Figure 2.F.1. Predicted Wage Changes



(a) Workers in Region A *(b) Workers in Region B*

Notes: Distribution of log wage differences between observed wages and model-based predictions. The left boxplot in each panel refers to the Mincer regression, while the right boxplot refers to the model counterfactual. In each boxplot, the box represents the interquartile range, the line inside the box the median, and the whiskers extend to 1.5 times the interquartile range. Points beyond the whiskers indicate outliers.

2.G Additional Policy Results

2.G.1 CEV of Older Workers

The construction policy presented in section 2.6 is beneficial for newborn workers as shown in the main text. However, it is beneficial for older workers too.

At age 45, CEV for workers in region A with median productivity, skill levels, and savings from the construction policy is 0.61%. The number is smaller than the 0.79% CEV for newborn workers. This is because workers at 45 have no more incentives to move to region B because of human capital mismatch and no longer benefit from greater moving opportunities. They do still benefit from the wage increase and the rent price decrease in region A.

By contrast, at 0.84%, median 45-year-old workers in region B have about the same CEV as newborn workers. This is because they stille directly benefit from the rent price reduction.

2.G.2 Downsizing Policy

Housing construction effectively increases welfare and makes moving to the region with better career opportunities easier, as discussed in the previous section. An alternative policy that is often discussed in public debates is incentivizing households to downsize their housing consumption in order to make room for more people. In the German state of Baden-Württemberg, for example, municipalities can receive a bonus payment of €3000 per household that moves to a flat that is at least 15 square meters smaller than the previous flat. The municipalities may then choose to pay this bonus out to the household directly.

I implement a stylized version of this policy for comparison with the construction project. In the model, there are two reasons that lead to high rents in region B. One is small housing supply as targeted by the construction policy. The other is old workers having accumulated a lot of wealth and consuming large quantities of housing. In the model, in region B, workers under the age of 30 rent a median flat of $58.7m^2$, whereas the median flat of retired workers is as large as $87.05m^2$.

I test the following policy. Workers who are retired and choose to live in a flat that is $15m^2$ smaller than the pre-policy median (i.e. $72.05m^2$) receive a lump sum transfer of €250. The net present value of receiving this transfer for every year of retirement is close to €3000. The technical details of the implementation are discussed in Appendix 2.D.2.

The policy is not effective at all. The average housing consumption of retired workers in region B is reduced by only about $0.5m^2$. As a result, none of the general-equilibrium objects change measurably. Table 2.G.1 shows the welfare effects on newborn workers. The effect is small and negative - the policy has no meaningful impact on aggregates but still leads to tax increases.

Table 2.G.1. CEV after Downsizing Policy

	Veil of ignorance	Conditional on A	Conditional on B
CEV	-0.19%	-0.19%	-0.19%

Notes: Consumption equivalent variation of newborn workers with respect to a counterfactual economy where the government pays lump sum transfers of €250 to retired workers who take up less than 72.05m^2 of housing.

Compared to housing construction, the downsizing policy has two main disadvantages. First, unless incentives are huge (and therefore extremely expensive), it only affects retired workers at the margin, and the housing space that is freed up is therefore limited. Second, houses that are constructed once can be used by subsequent generations, warranting financing the construction through government debt. Incentivizing old households to live in smaller flats needs to be paid for continuously and is therefore expensive.

References

Acemoglu, Daron, and David Autor. 2011. "Chapter 12 - Skills, Tasks and Technologies: Implications for Employment and Earnings," edited by David Card and Orley Ashenfelter, 4: 1043–171. Handbook of Labor Economics. Elsevier. [64]

Bajzik, Josef, Tomas Havranek, Zuzana Irsova, and Jiri Schwarz. 2020. "Estimating the Armington elasticity: The importance of study design and publication bias." *Journal of International Economics* 127: 103383. [70]

Bayer, Christian, and Falko Juessen. 2012. "On the dynamics of interstate migration: Migration costs and self-selection." *Review of Economic Dynamics* 15 (3): 377–401. [52]

Behrens, Kristian, Gilles Duranton, and Frédéric Robert-Nicoud. 2014. "Productive Cities: Sorting, Selection, and Agglomeration." *Journal of Political Economy* 122 (3): 507–53. [53]

Bilal, Adrien. 2023. "The Geography of Unemployment." *Quarterly Journal of Economics* 138 (3): 1507–76. [53, 65]

Bilal, Adrien, and Esteban Rossi-Hansberg. 2021. "Location as an Asset." *Econometrica* 89 (5): 2459–95. [51]

Böhm, Michael J., Hans-Martin von Gaudecker, and Felix Schran. 2024. "Occupation Growth, Skill Prices, and Wage Inequality." *Journal of Labor Economics* 42 (1): 201–43. [53, 54, 86]

Borjas, George J., Stephen G. Bronars, and Stephen J. Trejo. 1992. "Self-selection and internal migration in the United States." *Journal of Urban Economics* 32 (2): 159–85. [53]

Bryan, Gharad, and Melanie Morten. 2019. "The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia." *Journal of Political Economy* 127 (5): 2229–68. [52]

Card, David, Jesse Rothstein, and Moises Yi. 2025. "Location, Location, Location." *American Economic Journal: Applied Economics* 17 (1): 297–336. [52]

Carroll, Christopher D. 2006. "The method of endogenous gridpoints for solving dynamic stochastic optimization problems." *Economics Letters* 91 (3): 312–20. [93]

Dahl, Gordon B. 2002. "Mobility and the Return to Education: Testing a Roy Model with Multiple Markets." *Econometrica* 70 (6): 2367–420. [53]

Dauth, Wolfgang, Sebastian Findeisen, Enrico Moretti, and Jens Suedekum. 2022. "Matching in cities." *Journal of the European Economic Association* 20 (4): 1478–521. [53]

Dauth, Wolfgang, Michaela Fuchs, and Anne Otto. 2018. "Long-run processes of geographical concentration and dispersion: Evidence from Germany." *Papers in Regional Science* 97 (3): 569–93. [60]

Diamond, Rebecca. 2016. "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000." *American Economic Review* 106 (3): 479–524. [53]

Diaz, Antonia, Alvaro Janez, and Felix Wellschmied. 2023. "Geographic Mobility over the Life-Cycle." IZA Discussion Papers 15896. Institute of Labor Economics (IZA). [52, 65]

Domencich, Thomas A., and Daniel L. McFadden. 1975. *Urban Travel Demand: A Behavioral Analysis.* 93: xv + 215. Contributions to Economic Analysis. Amsterdam / New York: North-Holland Publishing Company / American Elsevier. [93]

Dustmann, Christian, Johannes Ludsteck, and Uta Schönberg. 2009. "Revisiting the German Wage Structure." *Quarterly Journal of Economics* 124 (2): 843–81. [54, 86]

Feenstra, Robert C., Philip Luck, Maurice Obstfeld, and Katheryn N. Russ. 2018. "In Search of the Armington Elasticity." *Review of Economics and Statistics* 100 (1): 135–50. [70]

Glaeser, Edward, and Joseph Gyourko. 2018. "The Economic Implications of Housing Supply." *Journal of Economic Perspectives* 32 (1): 3–30. [52]

Glaeser, Edward L., Joseph Gyourko, and Raven Saks. 2005. "Why Is Manhattan So Expensive? Regulation and the Rise in Housing Prices." *Journal of Law and Economics* 48 (2): 331–69. [52]

Glaeser, Edward L., and David C. Maré. 2001. "Cities and Skills." *Journal of Labor Economics* 19 (2): 316–42. [53]

Glaeser, Edward L., and Bryce A. Ward. 2009. "The causes and consequences of land use regulation: Evidence from Greater Boston." *Journal of Urban Economics* 65 (3): 265–78. [52]

Heise, Sebastian, and Tommaso Porzio. 2022. "Labor Misallocation Across Firms and Regions." Working Paper, Working Paper Series 30298. National Bureau of Economic Research. [52]

Herkenhoff, Kyle F, Lee E Ohanian, and Edward C Prescott. 2017. "Tarnishing the Golden and Empire States: Land-Use Restrictions and the U.S. Economic Slowdown." Working Paper, Working Paper Series 23790. National Bureau of Economic Research. [52]

Hsieh, Chang-Tai, and Enrico Moretti. 2019. "Housing Constraints and Spatial Misallocation." *American Economic Journal: Macroeconomics* 11 (2): 1–39. [52, 65]

Hunt, Gary L., and Richard E. Mueller. 2004. "North American Migration: Returns to Skill, Border Effects, and Mobility Costs." *Review of Economics and Statistics* 86 (4): 988–1007. [53]

Iskhakov, Fedor, Thomas H. Jørgensen, John Rust, and Bertel Schjerning. 2017. "The endogenous grid method for discrete-continuous dynamic choice models with (or without) taste shocks." *Quantitative Economics* 8 (2): 317–65. [93]

Jordà, Òscar, Moritz Schularick, and Alan M. Taylor. 2017. "Macrofinancial History and the New Business Cycle Facts." *NBER Macroeconomics Annual* 31: 213–63. [62, 95]

Kambourov, Gueorgui, and Iourii Manovskii. 2009. "OCCUPATIONAL SPECIFICITY OF HUMAN CAPITAL." *International Economic Review* 50 (1): 63–115. [53, 81]

Kaplan, Greg, and Sam Schulhofer-Wohl. 2017a. "UNDERSTANDING THE LONG-RUN DECLINE IN INTERSTATE MIGRATION." *International Economic Review* 58 (1): 57–94. [52]

Kaplan, Greg, and Sam Schulhofer-Wohl. 2017b. "UNDERSTANDING THE LONG-RUN DECLINE IN INTERSTATE MIGRATION." *International Economic Review* 58 (1): 57–94. [70]

Kennan, John, and James R Walker. 2011. "The effect of expected income on individual migration decisions." *Econometrica* 79 (1): 211–51. [52]

Lhullier, Hugo. 2025. "The Local Root of Wage Inequality." *University of Chicago, Becker Friedman Institute for Economics Working Paper,* nos. 2025-78. [53]

Lindenlaub, Ilse, Ryungha Oh, and Michael Peters. 2022. "Firm sorting and spatial inequality." Working paper. National Bureau of Economic Research. [53]

Moretti, Enrico. 2013. "Real Wage Inequality." *American Economic Journal: Applied Economics* 5 (1): 65–103. [52]

Ransom, Tyler. 2022. "Labor Market Frictions and Moving Costs of the Employed and Unemployed." *Journal of Human Resources* 57 (S): S137–S166. [52]

Roback, Jennifer. 1982. "Wages, Rents, and the Quality of Life." *Journal of Political Economy* 90 (6): 1257–78. [52]

Roca, Jorge De La, and Diego Puga. 2016. "Learning by Working in Big Cities." *Review of Economic Studies* 84 (1): 106–42. [53]

Saiz, Albert. 2010. "The Geographic Determinants of Housing Supply*." *Quarterly Journal of Economics* 125 (3): 1253–96. [52]

Schlattmann, Lennard. 2024. "Spatial Redistribution of Carbon Taxes." *JMP.* [70]

Schmutz, Benoît, and Modibo Sidibé. 2018. "Frictional Labour Mobility." *Review of Economic Studies* 86 (4): 1779–826. [52]

Shaw, Kathryn L. 1984. "A Formulation of the Earnings Function Using the Concept of Occupational Investment." *Journal of Human Resources* 19 (3): 319–40. [52]

Shaw, Kathryn L. 1987. "Occupational Change, Employer Change, and the Transferability of Skills." *Southern Economic Journal* 53 (3): 702–19. [52]

Sjaastad, Larry A. 1962. "The Costs and Returns of Human Migration." *Journal of Political Economy* 70 (5): 80–93. [51]

Sullivan, Paul. 2010. "Empirical evidence on occupation and industry specific human capital." *Labour Economics* 17 (3): 567–80. [53]

Tombe, Trevor, and Xiaodong Zhu. 2019. "Trade, Migration, and Productivity: A Quantitative Analysis of China." *American Economic Review* 109 (5): 1843–72. [52]

Yamaguchi, Shintaro. 2012. "Tasks and Heterogeneous Human Capital." *Journal of Labor Economics* 30 (1): 1–53. [53]

Chapter 3

Carrots or Sticks? Short-Time Work vs. Lay-Off Taxes*

Joint with Gero Stiepelmann

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 hand, there is the short-time work system that rewards the retention of endangered workers through subsidizing and enabling hours reductions[1].

This raises the natural question, which of the two policy instruments 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 been shown to have the problem of introducing new inefficiencies into the economy through the distortion of working hours (e.g., Burdett and Wright, 1989; Stiepelmann, 2026). However, these re-

* We thank Michael Peters and Martin Souchier for invaluable feedback and guidance during the development of this project. We gratefully acknowledge financial support by the Bonn Graduate School of Economics and the Yale Graduate School of Arts and Sciences. Gero Stieplemann gratefully acknowledges financial support by the Institute of Macroeconomics and Econometrics Bonn and Johannes Weber gratefully acknowledges financial support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 (Project C01). The views expressed in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Board of Governors, the Federal Reserve Bank of Cleveland, or the Federal Reserve System.

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.

sults crucially hinge on the absence of financial constraints for firms. This assumption is widespread in the search and matching literature, but clearly 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 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 27% of firms in the economy or more are financially constrained.

The backbone of our model is a canonical Diamond-Mortensen-Pissarides type search and matching model (Mortensen and Pissarides, 1993, 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.

The 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 separate after productivity falls below the separation threshold. To replicate the U.S. experience-rated unemployment insurance system, we assume that lay-off taxes are backloaded to the next period, thereby avoiding firm bankruptcies.

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. This implies that for sufficiently low productivity realizations, firms make ex-post losses. However, if the firm is financially constrained, it cannot make ex-post losses and loses its ability to fully insure workers. Low enough productivity shocks make the borrowing constraint binding, and the 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. Even though firms would like to keep the worker employed to avoid future lay-off taxes, workers quit unitarily when income falls enough, rendering the lay-off tax ineffective[2].

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 workers with income insurance, where constrained firms can no longer provide it. We proceed to show that short-time work cannot simply be replaced by providing loans to financially constrained firms. This is because where short-time work increases the joint surplus of firms and workers by providing insurance to workers, loans merely shift funds across periods and therefore have no bearing on a worker's wage and her decision to quit. Unlike short-time work, loans therefore do not prevent inefficient separations.

2. In the U.S., workers who quit following a substantial deterioration in job conditions remain eligible for unemployment insurance benefits.

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 of firms face financial constraints. Calibrating the model to the U.S. economy and targeting a 45% unemployment insurance replacement rate, we find that short-time work is the superior policy instrument if 27% of firms in the economy or more are financially constrained, which is a realistic figure for the U.S. economy. If the planner is allowed to optimize unemployment insurance benefits jointly with either short-time work or lay-off taxes, this threshold even drops to 24%.

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., Baily, 1978; Shavell and Weiss, 1979). 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) show 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 lay-offs 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 studies focus on the potential advantages of short-time work. Balleer et al. (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 recent advancements, 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 (2026) 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 this branch of the literature by enhancing the understanding of the interplay between short-time work and unemployment insurance by highlighting the pivotal role of 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 responses. Unlike Wasmer and Weil (2004) and Iliopulos, Langot, and Sopraseuth (2019), who model hiring costs as working capital so that constraints primarily affect vacancy posting, in our setting financial constraints operate through separations, as constrained firms are less able to sustain unprofitable matches.

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

$$y(\epsilon,h) \; = \; \epsilon \cdot h^{\alpha} - (\mu_{\epsilon} - \epsilon) \cdot k_f$$

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 $(\mu_{\epsilon} - \epsilon) \cdot k_f$ 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 denote the probability that firms and workers separate. 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 s whenever their productivity changes. The government budget constraint is

$$n^s \cdot s \; = \; (1-n) \cdot b - n^s \cdot \rho \cdot F$$

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,

$$y(\epsilon, h(\epsilon)) - w^c(\epsilon) \geq -\beta \cdot \left[\lambda \cdot \bar{J} + (1-\lambda) \cdot J^c(\epsilon)\right]$$

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

$$J^u(\epsilon) = y(\epsilon, h(\epsilon)) - w^u(\epsilon) + \beta \cdot \left[\lambda \cdot \bar{J} + (1-\lambda) J^u(\epsilon)\right]$$

and the value of a constrained firm *after* productivity realizes to ϵ that is

$$J^c(\epsilon) = y(\epsilon, h(\epsilon)) - w^c(\epsilon) + \beta \cdot \left[\lambda \cdot \bar{J} + (1-\lambda) \cdot J^c(\epsilon)\right]$$

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

$$\bar{J} = -s + (1-p)\left(\int_{\epsilon_s^u}^{\infty} J^u(\epsilon)\, dG(\epsilon)\right) + p\left(\int_{\epsilon_s^c}^{\infty} J^c(\epsilon)\, dG(\epsilon)\right) \quad - \rho \cdot \beta \cdot F$$

Firms separate from a worker once productivity drops below the separation threshold ϵ_s^u or ϵ_s^c for the unconstrained and constrained case, respectively. We denote the probability of separation when a new shock arrives as ρ. When separation occurs, the firm has to pay a lay-off tax F. In accordance with the experience rating system in the US, we assume that lay-off taxes are backloaded, i.e., they are paid with a one-period delay. This ensures that financial constraints do not bind at the time of separation. The lay-off tax is therefore discounted by β.[3]

Firms can freely enter the labor market and post vacancies at cost k_v. The mass of vacancies v must therefore solve

$$\frac{k_v}{q} = \beta \cdot \bar{J}$$

3. In principle, a firm could remain financially constrained in the subsequent period and thus be unable to pay the lay-off tax, potentially leading to bankruptcy. We abstract from this possibility and assume that all firms are able to meet their tax obligations when due.

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

$$V^u(\epsilon) = u(w^u(\epsilon) - \phi(h(\epsilon))) + \beta \cdot \left[\lambda \bar{V} + (1-\lambda) V^u(\epsilon)\right]$$

The value of a worker at a constrained firm *after* productivity realizes to ϵ is

$$V^c(\epsilon) = u(w^c(\epsilon) - \phi(h(\epsilon))) + \beta \cdot \left[\lambda \bar{V} + (1-\lambda) V^c(\epsilon)\right]$$

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

$$U = u(b) + \beta \cdot \left[f \cdot \bar{V} + (1-f) \cdot U\right]$$

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:

$$V^u(\epsilon) \leq U, \quad V^c(\epsilon) \leq U$$

Formally, the bargaining outcome solves

$$\max_{\mathcal{X}} \quad \bar{J}^{1-\eta} \left(\bar{V} - U\right)^{\eta}$$

subject to

$$\begin{aligned} V^u(\epsilon) \geq U, \quad V^c(\epsilon) &\geq U && \text{(commitment problem)} \\ y(\epsilon, h^c(\epsilon)) - w^c(\epsilon) &\geq -\beta\left[\lambda \bar{J} + (1-\lambda) J^c(\epsilon)\right] && \text{(financial constraint)} \end{aligned}$$

where the choice set is given by

$$\mathcal{X} = \left\{w^u(\epsilon), w^c(\epsilon), h^u(\epsilon), h^c(\epsilon), \epsilon_s^u, \epsilon_s^c\right\}.$$

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. Appendix 3.C.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:

$$c^w(\epsilon) = z(\epsilon) + \lambda \cdot \beta \cdot \bar{J}$$

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

$$\alpha \cdot \epsilon \cdot h(\epsilon)^{\alpha-1} = h(\epsilon)^{\psi}$$

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:

$$z(\epsilon_s^u) + [1 - \beta \cdot (1-\lambda)] \cdot \beta \cdot F + \frac{u(c^u) - u(b)}{u'(c^w)} - c^w + \frac{\lambda - \eta \cdot f}{1-\eta} \cdot \frac{k_v}{q} = 0$$

$$\frac{u(c^w(\epsilon_s^c)) - u(b)}{u'(c^w)} + (\lambda - f) \cdot \frac{\eta}{1-\eta} \cdot \frac{k_v}{q} = 0$$

Note that F enters the separation condition for unconstrained firms only. In the unconstrained case, firms fully 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 ($V^u(\epsilon_s^u) > U$). The lay-off tax increases the surplus and therefore lowers the separation threshold ϵ_s^u. However, a sufficiently low productivity realization in a financially constrained firm is passed directly to the worker's income. For low enough productivity levels, the continuation value of employment falls below the worker's outside option, such that $V^c(\epsilon_s^c) \leq U$, and the worker quits unilaterally. Although the firm would prefer to retain the worker in order to avoid paying the backloaded lay-off tax, it lacks the financial resources to offer a wage that satisfies the worker's participation constraint. Consequently, the lay-off tax does not affect the separation decision.

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 $\rho^c = G(\epsilon_s^c)$. 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 unemployed 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.C 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 $s_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon))$ 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 s_{stw}. Choosing an hours threshold D is a realistic implementation of what short-time work rules mandate in e.g., Germany.[4] 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, the government chooses the policy parameters D, s_{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

$$\begin{aligned} n^s \cdot s \; = \; & \frac{1-p}{\lambda} \cdot n^s \cdot \int_{\epsilon_s^u}^{\max\{\epsilon_{\text{stw}}, \epsilon_s^u\}} s_{\text{stw}} \cdot (\bar{h} - h_{\text{stw}}(\epsilon))\, dG(\epsilon) \\ & + \frac{p}{\lambda} \cdot n^s \cdot \int_{\epsilon_s^c}^{\max\{\epsilon_{\text{stw}}, \epsilon_s^c\}} s_{\text{stw}} \cdot (\bar{h} - h_{\text{stw}}(\epsilon))\, dG(\epsilon) \\ & + (1-n) \cdot b \end{aligned}$$

where n^s is the mass of matches that receive a new productivity shock. ϵ_s^u and ϵ_s^c 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

$$y(\epsilon, h(\epsilon)) - w^c(\epsilon) \; \geq \; -\beta \cdot \left[\lambda \cdot \bar{J} + (1-\lambda) \cdot J^c(\epsilon)\right]$$

and with short-time work

$$y(\epsilon, h_{\text{stw}}(\epsilon)) - w^c_{\text{stw}}(\epsilon) \; \geq \; -\beta \cdot \left[\lambda \cdot \bar{J} + (1-\lambda) \cdot J^c_{\text{stw}}(\epsilon)\right]$$

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 $w^c(\epsilon)$ 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

4. 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.

$h^u(\epsilon)$ 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

$$J^u(\epsilon) \;=\; y(\epsilon, h(\epsilon)) - w^u(\epsilon) + \beta \cdot \left[\lambda \bar{J} + (1-\lambda) J^u(\epsilon)\right]$$

and the value of an unconstrained firm *with* short-time work *after* productivity realizes is

$$J^u_{\text{stw}}(\epsilon) \;=\; y(\epsilon, h_{\text{stw}}(\epsilon)) - w^u_{\text{stw}}(\epsilon) + \beta \cdot \left[\lambda \bar{J} + (1-\lambda) J^u_{\text{stw}}(\epsilon)\right].$$

The value of firm *without* short-time work *after* productivity realizes to ϵ that is constrained is

$$J^c(\epsilon) \;=\; y(\epsilon, h(\epsilon)) - w^c(\epsilon) + \beta \cdot \left[\lambda \bar{J} + (1-\lambda) J^c(\epsilon)\right]$$

and the value of a constrained firm *with* short-time work *after* productivity realizes is

$$J^c_{\text{stw}}(\epsilon) \;=\; y(\epsilon, h_{\text{stw}}(\epsilon)) - w^c_{\text{stw}}(\epsilon) + \beta \cdot \left[\lambda \bar{J} + (1-\lambda) J^c_{\text{stw}}(\epsilon)\right]$$

where $\bar{J}$ is the expected firm value *before* the shock has realized. It is given by

$$\begin{aligned}\bar{J} \;=\; -s + (1-p)&\left(\int_{\max\{\xi^u_s, \epsilon_{\text{stw}}\}}^{\infty} J^u(\epsilon)\, dG(\epsilon) + \int_{\epsilon^u_s}^{\epsilon_{\max\{\epsilon^u_s, \text{stw}\}}} J^u_{\text{stw}}(\epsilon)\, dG(\epsilon)\right)\\ +\, p&\left(\int_{\max\{\xi^c_s, \epsilon_{\text{stw}}\}}^{\infty} J^c(\epsilon)\, dG(\epsilon) + \int_{\epsilon^c_s}^{\max\{\epsilon^c_s, \epsilon_{\text{stw}}\}} J^c_{\text{stw}}(\epsilon)\, dG(\epsilon)\right)\end{aligned}$$

The max operators in the integral bounds 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

$$\frac{k_v}{q} \;=\; \beta \cdot \bar{J}$$

Workers The value of a worker at an unconstrained firm *without* short-time work *after* productivity realizes to ϵ is

$$V^u(\epsilon) \;=\; u(w^u(\epsilon) - \phi(h(\epsilon))) + \beta \cdot \left[\lambda \bar{V} + (1-\lambda) V^u(\epsilon)\right]$$

and the value of a worker at an unconstrained firm *with* short-time work *after* productivity realizes is

$$V^u_{\text{stw}}(\epsilon) \;=\; u(w^u_{\text{stw}}(\epsilon) + s_{\text{stw}} \cdot (\bar{h} - h_{\text{stw}}(\epsilon)) - \phi(h_{\text{stw}}(\epsilon))) + \beta \cdot \left[\lambda \bar{V} + (1-\lambda) V^u_{\text{stw}}(\epsilon)\right]$$

The value of worker *without* short-time work *after* productivity realizes to ϵ who works at a constrained firm is

$$V^c(\epsilon) \;=\; u(w^c(\epsilon) - \phi(h(\epsilon))) + \beta \cdot \left[\lambda \bar{V} + (1-\lambda) V^c(\epsilon)\right]$$

and the value of a worker *with* short-time work *after* productivity realizes who works at a constrained firm is

$$V^c_{\text{stw}}(\epsilon) \;=\; u(w^c_{\text{stw}}(\epsilon) + s_{\text{stw}} \cdot (\bar{h} - h_{\text{stw}}(\epsilon)) - \phi(h^c_{\text{stw}}(\epsilon))) + \beta \cdot \left[\lambda \bar{V} + (1-\lambda) V^c_{\text{stw}}(\epsilon)\right]$$

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.

$$\begin{aligned} V^u(\epsilon) \leq U, \quad V^u_{\text{stw}}(\epsilon) \leq U \\ V^c(\epsilon) \leq U, \quad V^c_{\text{stw}}(\epsilon) \leq U \end{aligned}$$

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 solves

$$\max_{\mathcal{X}} \quad \bar{J}^{1-\eta}\left(\bar{V} - U\right)^{\eta}$$

Subject to

$$\begin{aligned} V^u(\epsilon) \geq U, V^c(\epsilon) \geq U, V^u_{\text{stw}}(\epsilon) \geq U, V^c_{\text{stw}}(\epsilon) \geq U \quad & \text{(commitment problem)} \\ y(\epsilon, h^c(\epsilon)) - w^c(\epsilon) \geq -\beta\left[\lambda\bar{J} + (1-\lambda)J^c(\epsilon)\right] \quad & \text{(financial constraint)} \\ y(\epsilon, h^c_{\text{stw}}(\epsilon)) - w^c_{\text{stw}}(\epsilon) \geq -\beta\left[\lambda\bar{J} + (1-\lambda)J^c_{\text{stw}}(\epsilon)\right] \quad & \text{(financial constraint, STW)} \end{aligned}$$

where the choice set is given by

$$\mathcal{X} = \{w^u(\epsilon), w^c(\epsilon), w^u_{\text{stw}}(\epsilon), w^c_{\text{stw}}(\epsilon), h^u(\epsilon), h^c(\epsilon), h^u_{\text{stw}}(\epsilon), h^c_{\text{stw}}(\epsilon), \epsilon^u_s, \epsilon^c_s, \xi^u_s, \xi^c_s\}.$$

Here, η denotes the bargaining power of workers. Firms and workers take both the worker's participation constraints and the firm's financial constraints into account when writing contracts. Appendix 3.D.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:

$$u'(c^w(\epsilon)) = u'(c^w_{\text{stw}}(\epsilon)) = u'(c^w)$$

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:

$$\begin{aligned} c^w(\epsilon) &= z(\epsilon) + \lambda \cdot \frac{k_v}{q} \\ c^w_{\text{stw}}(\epsilon) &= z_{\text{stw}}(\epsilon) + s_{\text{stw}} \cdot (\bar{h} - h_{\text{stw}}(\epsilon)) + \lambda \cdot \frac{k_v}{q}. \end{aligned}$$

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

$$\begin{aligned} \alpha \cdot \epsilon \cdot h(\epsilon)^{\alpha-1} &= h(\epsilon)^{\psi} \\ \alpha \cdot \epsilon \cdot h_{\text{stw}}(\epsilon)^{\alpha-1} &= h_{\text{stw}}(\epsilon)^{\psi} + s_{\text{stw}}. \end{aligned}$$

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 s_{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:

$$\begin{aligned} z(\xi^u_s) + \frac{u(c^w) - u(b)}{u'(c^w)} - c^w + \frac{\lambda - \eta f}{1-\eta} \cdot \frac{k_v}{q} &= 0 \\ z_{\text{stw}}(\epsilon^u_s) + s_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon^u_s)) + \frac{u(c^w) - u(b)}{u'(c^w)} - c^w + \frac{\lambda - \eta f}{1-\eta} \cdot \frac{k_v}{q} &= 0 \\ \frac{u(c(\xi^u_s)) - u(b)}{u'(c^w)} + (\lambda - f)\frac{\eta}{1-\eta} \cdot \frac{k_v}{q} &= 0 \\ \frac{u(c_{\text{stw}}(\epsilon^c_s)) - u(b)}{u'(c^w)} + (\lambda - f)\frac{\eta}{1-\eta} \cdot \frac{k_v}{q} &= 0 \end{aligned}$$

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 short-time work, 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.

In constrained 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. Short-time work increases the worker's available income in such cases, thereby increasing the worker's incentive to remain with the firm. In essence, short-time work operates by reducing the wage burden on the firm required to sustain this consumption level. In effect, short-time work lowers the cost of providing income insurance, thereby reducing the firm's incentive to initiate separations.

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:

$$\begin{aligned} n^u(p) \;&=\; \frac{1-p}{\lambda}\cdot(1-\rho^u)\cdot n^s \\ n^c(p) \;&=\; \frac{p}{\lambda}\cdot(1-\rho^c)\cdot n^s \end{aligned}$$

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.D 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.C and 3.D in the appendix. To make the theoretical results more concise, we henceforth assume that $\beta \to 1$ for the remainder of the paper.

To state the welfare function formally, 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 (2026) shows that $\Omega \geq 0$, $\frac{\partial \Omega}{\partial s_{\text{stw}}} \geq 0$ and $\frac{\partial \Omega}{\partial \epsilon_{\text{stw}}} \geq 0$.

The welfare function is given by

$$\begin{aligned} W(\pi) &= n^u \cdot u(c^w) + n^c \cdot u^c + (1-n)\cdot u(b) \\ &\quad + \nu^f \cdot u((n\cdot(z-\Omega) - n^u\cdot c^w - n^c \cdot e^c - \tau(b) - \theta\cdot(1-n)\cdot k_v)/\nu^f) \end{aligned} \tag{3.1}$$

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 $\pi = (b, \epsilon_{\text{stw}}, s_{\text{stw}})$ 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.E 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 a 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*

$$\begin{aligned}\underbrace{(1-n)\cdot\left(\frac{u'(b)-u'(c^w)}{u'(c^w)}\right)}_{\text{Income insurance to unemployed, MIUB}} &= \underbrace{\left(-\frac{df}{db}\right)^{ge}\cdot u\cdot L_v(b)}_{\text{Less vacancy posting, MLV}} + \underbrace{\left(\frac{d\epsilon_s^u}{db}\right)^{ge}\cdot\frac{\partial n^u(p)}{\partial \epsilon_s^u}\cdot L_s^u(b,F)}_{\text{More separation in unconstrained firms, }MLS^u} \\ &+ \underbrace{\left(\frac{d\xi_s^c}{db}\right)^{ge}\cdot\frac{\partial n^c(p)}{\partial \epsilon_s^c}\cdot L_{s,\xi}^c(b)}_{\text{More separations in constrained firms, }MLS^c} \\ &+ \underbrace{MIE_b(b)}_{\text{Less income insurance to employed}}\end{aligned}$$

where L_v denotes the social value of hiring an additional worker, L_s^u and L_s^c denote the social value of the marginal match in an unconstrained and constrained firm,

$$\begin{aligned} L_v(b) &= \left(\frac{\eta - \gamma}{(1-\gamma)(1-\eta)} \cdot \bar{J} + \frac{1}{f} \cdot b \right) \\ L_s^u(b,F) &= \lambda \cdot \left(\frac{1}{f} \cdot b - \beta \cdot F \right) \\ L_s^c(b) &= \frac{\lambda}{f} \cdot b \end{aligned}$$

and MIE_b denotes the social costs from the insurance loss in unconstrained firms when UI benefits are set more generously.

$$MIE_b(b) = n^c(p) \cdot \frac{1}{1-\rho^c} \cdot \left(-\left(\frac{d\theta}{db} \right)^{ge} \right) \cdot \left(\int_{\epsilon_s^c}^{\epsilon^p} \lambda \cdot \frac{\gamma}{f} \cdot \frac{u'(c(\epsilon)) - u'(c^w)}{u'(c^w)} \, dG(\epsilon) \right) \cdot k_v$$

PROOF: See Section 3.F.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 separate earlier. Additionally, 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, MIE^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[5] 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 the firm cannot sustain a wage that the worker is willing to work for, it cannot prevent the worker from quitting. Thus, the firm has no way of avoiding paying lay-off taxes, and lay-off taxes have no bearing on the separation decision.

As a result, lay-off taxes can only mitigate the negative fiscal externalities of UI by combating separations in *unconstrained* firms. This means that when $p = 0$, lay-off taxes can induce the efficient number of separations in the 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 backloaded lay-off tax F is determined by its FOC:*

$$F = \frac{1}{\beta} \cdot \frac{1}{f} \cdot b + BE_{lt}(b) \tag{3.2}$$

PROOF: See Section 3.F.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

5. Note that ϵ^p in IE_θ denotes the productivity level at which financial constraints become binding for a constrained firm.

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.F in the appendix.

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. Jung and Kuester (2015) show that a combination of a vacancy subsidy and a lay-off tax can decentralize the planner's solution in a search-and-matching model, as the vacancy subsidy additionally operates on the vacancy-posting margin.

3.3.3 The Optimal Short-Time Work Regime

Under the short-time work regime the Ramsey planner chooses s_{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 $(s_{stw}, \epsilon_{stw})$ must fulfill the following first-order condition:*

$$
\begin{aligned}
\underbrace{(1-n)\cdot\left(\frac{u'(b)-u'(c^w)}{u'(c^w)}\right)}_{\textit{Income insurance to unemployed, MUIB}} &= \underbrace{\left(-\frac{df}{db}\right)^{ge}\cdot u\cdot L_v(b)}_{\textit{Less vacancy posting, MLV}} \\
&+ \underbrace{\left(-\frac{d\epsilon_s^u}{db}\right)^{ge}\cdot\frac{\partial n^u(p)}{\partial\epsilon_s^u}\cdot L_s^u(b,s_{stw})}_{\textit{More separations in unconstr. firms, } MLS^u} \\
&+ \underbrace{\mathbb{1}(\epsilon_{stw}\geq\epsilon_s^c)\cdot\left(-\frac{d\epsilon_s^c}{db}\right)^{ge}\cdot\frac{\partial n^c(p)}{\partial\epsilon_s^c}\cdot L_s^c(b,s_{stw})}_{\textit{More separations in constr. firms with STW, } MLS_\epsilon^c} \\
&+ \underbrace{\mathbb{1}(\epsilon_{stw}\leq\xi_s^c)\cdot\left(-\frac{d\xi_s^c}{db}\right)^{ge}\cdot\frac{\partial n^c(p)}{\partial\xi_s^c}\cdot L_{s,\xi}^c(b)}_{\textit{More separations in constr. firms without STW, } MLS_\xi^c} \\
&+ \underbrace{MIE_b^c}_{\textit{Less income insurance to employed workers}}
\end{aligned}
$$

where L_v denotes the social value of hiring an additional worker

$$L_v = \left(\frac{\eta-\gamma}{(1-\gamma)(1-\eta)}\cdot\bar{J}+\frac{1}{f}\cdot b\right)$$

L_s^u, L_s^c and $L_{s,\xi}^c$ denote the social value of a marginal match in a constrained firm, in an unconstrained firm with access to STW, and an unconstrained firm without access to STW,

$$L_s^i = \left(\frac{\lambda}{f}\cdot b - s_{stw}\cdot\left(\bar{h}-h_{stw}(\xi_s^i)\right)\right) \quad \text{with} \quad i\in\{u,c\}, \quad L_{s,\xi}^c = \frac{\lambda}{f}\cdot b$$

and MIE_b^c denotes the total insurance loss from marginally increasing UI benefits in constrained firms

$$
\begin{aligned}
MIE_b^c = n^c(p)\cdot\left(-\frac{d\theta}{db}\right)^{ge}\cdot\frac{1}{1-\rho^c}\cdot\Bigg(&\int_{\max\{\epsilon_{stw},\xi_s^c\}}^{\epsilon^p}\lambda\cdot\frac{\gamma}{f}\cdot\frac{u'(c(\epsilon))-u'(c^w)}{u'(c^w)}\,dG(\epsilon) \\
&+\int_{\epsilon_s^c}^{\max\{\epsilon_{stw},\epsilon_s^c\}}\lambda\cdot\frac{\gamma}{f}\cdot\frac{u'(c_{stw}(\epsilon))-u'(c^w)}{u'(c^w)}\,dG(\epsilon)\Bigg)\cdot k_v
\end{aligned}
$$

PROOF: See Section 3.G.6 in the appendix.

As in the lay-off tax regime, the marginal social benefit of more generous unemployment benefits (*MUIB*) must equal the associated marginal welfare losses. These losses arise from three sources: fewer vacancies being posted (*MLV*), increased separations in both unconstrained and constrained firms (*MLS^u* and *MLS^c*), and a loss of insurance in constrained firms (MIE^c). Importantly, constrained firms feature two distinct separation margins. Matches may separate with short-time work (STW) support at the threshold ϵ_s^c, or without STW support at ξ_s^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 the lay-off tax could only counteract the adverse effect of unemployment benefits on separations in unconstrained firms. By contrast, the short-time work parameters $(s_{\text{stw}}, \epsilon_{\text{stw}})$ 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, counteracting inefficiently high separations and inefficiently low insurance provision.

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 from separations (MLS) terms, but also enhances income insurance by acting on the MIE_b^c 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_{stw} is set even more 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 in both unconstrained and constrained firms and the adverse effects of hours distortions. This is reflected in how optimal short-time work is set.

We begin by examining how the short-time work eligibility threshold ϵ_{stw} should be determined. The following proposition establishes that there are three candidates for the optimal eligibility threshold. The optimal threshold can be found by comparing the welfare levels induced by these three candidates. As discussed in section 3.2.3, the optimal eligibility condition in terms of working hours D is uniquely implied by the optimal ϵ_{stw}.

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.*[6] *The optimal eligibility threshold $\epsilon^{\star}_{stw}$ is one of three candidates pinned down by the following conditions:*

Candidate 1

Suppose it is optimal to set $\epsilon_{stw} \geq \xi^{c}_{s}$ (letting all firms in need access short-time work). If

$$MWL_{\epsilon_{stw}} + BE_{stw,2} \quad > \quad MIE^{c}_{\epsilon_{stw}}, \tag{3.3}$$

then the optimal threshold is given by $\epsilon_{stw} = \xi^{c}_{s}$.

Otherwise, the optimal threshold solves

$$MWL_{\epsilon_{stw}} + BE_{stw,2} \quad = \quad MIE^{c}_{\epsilon_{stw}}. \tag{3.4}$$

Here, $MWL_{\epsilon_{stw}} = n \cdot \frac{\partial \Omega}{\partial \epsilon_{stw}}$ denotes the total welfare loss from marginally increasing STW benefits, and $MIE^{c}_{\epsilon_{stw}}$ denotes the marginal welfare gain from providing income insurance:

$$MIE^{c}_{\epsilon_{stw}} = n^{s} \cdot \frac{p}{\lambda} \cdot g(\epsilon_{stw}) \cdot \frac{\dfrac{u(c_{stw}(\epsilon_{stw})) - u(c(\epsilon_{stw}))}{c_{stw}(\epsilon_{stw}) - c(\epsilon_{stw})} - u'(c^{w})}{u'(c^{w})} \cdot [c_{stw}(\epsilon_{stw}) - c(\epsilon_{stw})].$$

Candidate 2

Suppose it is optimal to set $\epsilon_{stw} < \epsilon^{c}_{s}$ with $\xi^{u}_{s} < \epsilon^{c}_{s}$ (full exclusion of constrained firms from STW). Then the optimal threshold is given by $\epsilon_{stw} = \xi^{u}_{s}$.

Candidate 3

Let $I = \left[\max\{\epsilon^{c}_{s}, \xi^{u}_{s}\}, \xi^{c}_{s}\right]$. Suppose it is optimal to set $\epsilon_{stw} \in I$ (partial exclusion of constrained firms in need of STW). If

$$\underbrace{L^{c}_{s,\epsilon}(\epsilon_{stw})}_{\text{Social value of preserved match}} \quad > \quad \underbrace{\frac{MWL^{u}_{\epsilon_{stw}}}{MMS^{c}_{\epsilon_{stw}}}}_{\substack{\text{Social cost of hours distortion}\\ \text{in unconstrained firms}\\ \text{per match retained}}} + BE_{stw,2} \quad \forall \epsilon \in I, \tag{3.5}$$

6. 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 in Section 3.G.5 in the appendix.

then the optimal threshold is given by $\epsilon_{stw} = \xi_s^c$.

If

$$L_{s,\epsilon}^c(\epsilon_{stw}) \quad < \quad \frac{MWL_{\epsilon_{stw}}^u}{MMS_{\epsilon_{stw}}^c} + BE_{stw,2} \quad \forall \epsilon \in I, \tag{3.6}$$

then the optimal threshold is given by $\epsilon_{stw} = \max\{\epsilon_s^c, \xi_s^u\}$.

Otherwise, the optimal threshold is given by ϵ_{stw} *that solves*

$$L_{s,\epsilon}^c(\epsilon_{stw}) \quad = \quad \frac{MWL_{\epsilon_{stw}}^u}{MMS_{\epsilon_{stw}}^c} + BE_{stw,2} \tag{3.7}$$

Here, $L_s^c(\epsilon_{stw})$ *denotes the social value of an unconstrained match retained by marginally loosening the participation constraint.* $MMS_{\epsilon_{stw}}^c = n^s \cdot \frac{p}{\lambda} \cdot g(\epsilon_{stw})$ *denotes the number of constrained matches retained with a looser eligibility condition.* $MWL_{s_{stw}}^u = n^s \cdot \frac{1-p}{\lambda} \cdot \frac{\partial \Omega}{\partial \epsilon_{stw}}$ *denotes the additional welfare cost of working hours distortions in unconstrained firms when loosening the eligibility condition. Finally,* $BE_{stw,2}$ *captures general equilibrium effects through the generalized Nash bargaining process.*

PROOF: See Section 3.G.5 in the appendix.

To explain the economic logic behind Proposition 4 and why it can only characterize candidates we first discuss the basic trade-off the planner faces. For this, one intuitive additional result is needed, namely that constrained firms will always separate at even higher productivity levels than their unconstrained counterparts as stated by the following lemma:

Lemma 1. *Constrained firms separate for at higher productivity levels than unconstrained firms:*

$$\epsilon_s^c > \epsilon_s^u \quad \textit{and} \quad \xi_s^c > \xi_s^u$$

PROOF: See Section 3.H in the appendix.

Remember that any mass of firms on short-time work introduces socially suboptimal 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 access to short-time work.

When choosing ϵ_{stw}, the planner therefore faces a trade-off. The threshold can be set leniently enough to allow constrained matches that would otherwise separate access to short-time work—that is, $\epsilon_{\text{stw}} \geq \xi_s^c$. However, in this case, a mass of unconstrained matches with productivity $\epsilon \in \left[\xi_s^u, \epsilon_{\text{stw}}\right]$ will also gain access to short-time work, even though they would not separate in the absence of the program, thereby introducing socially costly distortions in working hours in unconstrained firms.

Alternatively, to avoid these distortions, the planner could choose to ignore constrained firms and set the eligibility threshold to $\epsilon_{\text{stw}} = \xi_s^u$. At this threshold, unconstrained firms separate without access to short-time work. In this case, constrained firms with $\epsilon \in \left[\max\{\epsilon_s^c, \xi_s^u\}, \xi_s^c\right]$ are excluded from the STW system even though they are in greater need of support than unconstrained firms. Since which of these is the lesser evil depends on the exact parameters, Proposition 4 can only characterize candidates for the optimal eligibility threshold.

With this in mind, we now turn to explaining the rationale behind each of the three candidates for the optimal eligibility threshold. One possibility is that the eligibility condition should be set so loosely that all firms in need of support can access short-time work. In this case, candidate 1 is optimal. To provide access to short-time work to all firms in need, $\epsilon_{\text{stw}} = \xi_s^c$ 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.

Proposition 4 characterizes the conditions under which either of these cases is optimal. Our numerical results presented in Section 3.5, indeed, show that quantitatively, the distortion effect dominates and that $\epsilon_{\text{stw}} = \xi_s^c$ will be optimal.

A second possibility is to set the eligibility threshold to include only unconstrained firms that are actually in need of short-time work, thereby excluding all constrained firms with a productivity level between ξ_s^u and ξ_s^c from the system. In this case, candidate 2 is optimal. The optimal threshold is given by $\epsilon_{\text{stw}} = \xi_s^u$. As argued by Stiepelmann (2026) $\epsilon_{\text{stw}} < \xi_s^u$ 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.

A third possibility is that the eligibility criterion for short-time work (STW) should be set sufficiently strict that some financially constrained firms are excluded, even though STW is necessary to retain their workers, but not strict enough to exclude all financially constrained firms. In this case, candidate 3 is optimal. In this scenario, all unconstrained firms that require STW for worker retention can participate in short-time work, along with some that could survive without it. The Ramsey planner must balance the social gains from protecting constrained firms through a looser eligibility threshold against the additional hours distortions imposed on unconstrained firms where STW is not needed to retain workers. This trade-off is captured by Inequalities 3.5 and 3.6 and depends on the value of p.

A higher p increases the share of workers in financially constrained firms. When p is large, the social value of retaining these workers dominates the costs of distorting hours in unconstrained firms, as the marginal cost of distortion MWL^u_{stw} declines with p while the number of constrained firms retained MMS^c_{stw} rises. In this case, the eligibility threshold is set sufficiently loose to allow all unconstrained firms for which STW is needed to retain workers to participate, i.e., $\epsilon_{\text{stw}} = \xi^c_s$. Conversely, when p is small, hours distortion costs dominate, and the planner effectively ignores STW, setting $\epsilon_{\text{stw}} = \xi^u_s$. For intermediate values of p, the optimal threshold might balance the social benefits of worker retention against the social costs of hours distortions in unconstrained firms. Our numerical results in section 3.5 show that either the social value of worker retention or the social cost of working hours distortion dominate for almost all values of p.

Next, we turn to the optimal level of STW benefits s_{stw}, characterized by the following proposition:

Proposition 5. *The optimal short-time work subsidy s_{stw} is pinned down by the first order condition*

$$\bar{s}_{stw} = \underbrace{\frac{T_{UI}}{T_{STW}}}_{\text{Fiscal Ext.}} \cdot b + \underbrace{\frac{MIE^c_{s_{stw}}}{MMS_{s_{stw}}}}_{\text{Insurance}} - \underbrace{\frac{MWL_{s_{stw}}}{MMS_{s_{stw}}}}_{\text{Distortion}} + \underbrace{BE_{stw,3}}_{\text{Bargaining Effect}}$$

Here, $T_{UI} = \frac{1}{f}$ denotes the expected duration that a worker spends on the UI system, $T_{STW} = \frac{1}{\lambda}$ denotes the expected duration that a workers spends on the STW system, $MIE^c_{s_{stw}}$ denotes the marginal welfare gain from increasing insurance provided to workers in unconstrained firms:

$$MIE^c_{s_{stw}} = \frac{n^c(p)}{1-\rho^c} \cdot \int_{\epsilon^c_s}^{\max\{\epsilon^c_s, \epsilon_{stw}\}} \left(\frac{u'(c_{stw}(\epsilon)) - u'(c^w)}{u'(c^w)} \right) (\bar{h} - h_{stw}(\epsilon))\, dG(\epsilon)$$

Further, $MWL_{\epsilon_{stw}} = n \cdot \frac{\partial \Omega}{\partial \epsilon_{stw}}$ denotes the marginal welfare loss from hours distortion from increasing STW benefits. $MMS_{s_{stw}} = MMS^c_{s_{stw}} + MMS^u_{s_{stw}}$ denotes the marginal number of matches saved with an increase of STW benefits, where

$$\begin{aligned} MMS^u_{s_{stw}} &= n^s \cdot \frac{1-p}{\lambda} \cdot g(\epsilon^u_s) \cdot \frac{\partial \epsilon^u_s}{\partial s_{stw}} \\ MMS^c_{s_{stw}} &= \mathbb{1}\{\epsilon_{stw} \geq \epsilon^c_s\} \cdot n^s \cdot \frac{p}{\lambda} \cdot g(\epsilon^c_s) \cdot \frac{\partial \epsilon^c_s}{\partial s_{stw}} \end{aligned}$$

denote the marginal number of matches saved in constrained and unconstrained firms. The net-transfer of firms and workers is weighted by how many matches of constrained and unconstrained firms can be retained with an increase in STW benefits:

$$\bar{s}_{stw} = \frac{MMS^u_{s_{stw}}}{MMS_{s_{stw}}} \cdot s_{stw} \cdot \left(\bar{h} - h_{stw}(\epsilon^u_s)\right) + \frac{MMS^c_{s_{stw}}}{MMS_{s_{stw}}} \cdot s_{stw} \cdot \left(\bar{h} - h_{stw}(\epsilon^c_s)\right)$$

PROOF: See Section 3.G.4 in the appendix.

Proposition 5 highlights both the forces that make more generous short-time work attractive to the planner and the associated social costs. The most important motive is the fiscal externality correction arising from the unemployment insurance (UI) system. Short-time work helps internalize the fiscal externality that UI imposes on separation decisions in both constrained and unconstrained firms. In unconstrained firms, STW reduces inefficient separations by lowering the cost of insuring workers when the match is hit by a low-productivity shock. In financially constrained firms, STW partly compensates workers for the associated income loss and thereby reduces their incentive to quit.

The optimal level of STW benefits depends on the expected duration of time T_{stw} spent in the STW and UI systems. If the time a match spends in STW T_{stw} is relatively long, the Ramsey planner sets lower benefits in order to encourage reallocation toward more productive matches. Conversely, if workers tend to spend a long time in the UI system, higher STW benefits become desirable, as preserving the match and waiting for the firm's recovery may restore productivity.

The second term reflects the insurance motive for short-time work. It shows that short-time work can provide income insurance to workers in financially constrained firms. Here, $MIE^{c}_{s_{\text{stw}}}$ denotes the marginal insurance effect—the total welfare gain from increasing income insurance through higher STW benefits. As in the standard optimal UI framework, the planner would ideally like to fully insure workers against income shocks.

However, the Ramsey planner faces a trade-off: increasing short-time work benefits improves insurance but simultaneously weakens separation incentives. This can be a problem because with full insurance, labor does not get reallocated even for very low levels of productivity, which is inefficient. The optimal level of short-time work benefits therefore balances the insurance motive against the behavioral distortion arising from inefficient worker retention. This trade-off is reflected in the formula by the fact that $MIE^{c}_{s_{\text{stw}}}$ is scaled by the marginal number of matches preserved when short-time work benefits increase $MMS_{s_{\text{stw}}}$. If short-time work is highly effective at preventing separations, providing insurance through short-time work becomes more distortionary, since higher benefits lead to a larger increase in inefficient retention. In this case, the optimal level of short-time work benefits is lower.

The main reason why the planner does not want to set short-time work benefits too high is that short-time work distorts firms' working hours choices, generating a marginal welfare loss captured by the MWL_{stw} term. Since the planner trades off the employment stabilization benefits of worker retention against the efficiency costs from distorted hours, the marginal welfare loss MWL_{stw} is scaled by the marginal number of matches preserved, $MMS_{s_{\text{stw}}}$.

It is noteworthy how the Ramsey planner sets the *average net* subsidy $\bar{s}_{\text{stw}}$ optimally and then adjusts the actual benefits s_{stw} to achieve this optimal level. Since constrained firms have higher separation thresholds ($\epsilon_s^c > \epsilon_s^u$), it will hold that financially constrained firms reduce working hours less and get a smaller net subsidy: $s_{\text{stw}} \cdot \left(\bar{h} - h_{\text{stw}}(\epsilon_s^c)\right) < s_{\text{stw}} \cdot \left(\bar{h} - h_{\text{stw}}(\epsilon_s^u)\right)$. The net subsidies are weighted according to how many firms the Ramsey planner can save with an increase of short-time work benefits ($MMS^i_{s_{\text{stw}}}$). Therefore, optimal short-time work benefits depend on the share of financially constrained firms. As constrained firms receive a smaller net subsidy, an increase in their share requires higher benefits to compensate for the smaller reduction in working hours. At the same time, a higher fraction of constrained firms raises the marginal insurance gain $MIE^c_{s_{\text{stw}}}$, further increasing optimal benefits.

3.3.4 Why Loans Cannot Replace Short-Time Work

The theoretical analysis shows that short-time work does particularly well when a lot of firms are financially constrained. A common critique of using short-time work to address inefficiencies arising from financial constraints is that such constraints could instead be alleviated by providing loans to firms or workers. However, the financial friction we consider cannot be resolved through the provision of loans.

In our model, the financial friction arises from the assumption that financially constrained firms cannot write fully state-contingent contracts that insure them against low-productivity realizations. Therefore, in low-productivity states, firms do not have the resources to pay workers enough to sustain the match. However, a Ramsey planner would prefer to retain some of these matches, as workers may inefficiently quit in order to receive UI benefits. Providing loans does not resolve this inefficiency because loans merely shift funds across periods and therefore do not alter the fundamental surplus generated by the match and thus the financial constraint.

To see this, assume the government offers a loan to the firm. For simplicity, assume that the government loans a sum of L to the worker, who has to be repaid if the firm gets productive in the next period and that does not have to be repaid if it is not productive. We assume that the loan is fairly priced so that

$$L = \beta \cdot \lambda \cdot R \cdot L + \beta \cdot 0 \quad \Longleftrightarrow \quad R = \frac{1}{\beta \cdot \lambda}.$$

Consider the impact of the loan on the financial constraint:

$$y(\epsilon, h(\epsilon)) + L - w^c(\epsilon) \geq -\beta \cdot \left[\lambda \cdot \left(\bar{J} - R \cdot L\right) + (1-\lambda) \cdot J^c(\epsilon)\right]$$

Inserting the fairly priced loan into the formula shows that loans do not alter the financial constraint:

$$y(\epsilon, h(\epsilon)) + L - w^c(\epsilon) \geq L - \beta \cdot \left[\lambda \cdot \bar{J} + (1-\lambda) \cdot J^c(\epsilon)\right]$$

The loan proves ineffective. The reason for this is that loans do not alter the fundamental surplus of firms and workers. This is because the constraint faced by the firm is not a simple liquidity constraint. To sustain employment, the joint surplus of firms and workers must increase. Short-time work achieves this by compensating workers for lost income when hours are reduced, thereby raising the joint surplus of the match.

3.3.5 Comparing Lay-Off Taxes and Short-Time Work

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. In the following sections, we turn to quantitatively assessing the relative performance of the two policy regimes for different values of p.

3.4 Calibration

We calibrate the model to the U.S. economy at monthly frequency. We calibrate for *CRRA* utility, i.e. $u(c) = \frac{c^{1-\phi}}{1-\phi}$. For ϕ, we set the standard value of 2. We set the labor elasticity of the production function α to 0.65 following Christoffel and Linzert (2010). The elasticity of matching with respect to unemployment γ is set to the standard value of 0.65. As customary, we set the worker bargaining power η to 0.65 as well to fulfill the Hosios condition (Hosios, 1990). Note that 0.65 is well within the range of estimates reported by Petrongolo and Pissarides (2001). The disutility of labor parameter ψ is set to 1.5 to target a Frisch elasticity of labor supply of 0.65 as in Domeij and Floden (2006).

We set the shock arrival probability λ to 0.07, so that a firm will remain at any drawn productivity level for $1/0.07 = 14.29$ months on average. This choice is informed by the estimate of yearly plant-level autocorrelation of productivity of around 0.4 by Abraham and White (2006).[7]

The location of the productivity shock distribution μ is set to −0.80 to normalize the average wage w to 1. Since the effective dispersion of productivity shocks is governed by the cost shock parameter c_f, which is calibrated internally, we set σ to 0.12 following Krause and Lubik (2007).

The remaining parameters are calibrated internally. The calibration targets are shown in Table 3.1. All targets are met exactly.

Table 3.1. Model calibration targets

Parameter	Description	Value
Worker flows		
q	Vacancy-filling rate	0.3381
f	Job-finding rate	0.41
ρ^{eff}	Effective monthly separation rate	0.03
Policy parameters		
b^{rep}	UI replacement rate	0.45
F^{rep}	Lay-off tax coverage of UI payments	0.5

Notes: $b^{\text{rep}} \equiv \frac{b}{w}$ denotes the unemployment insurance (UI) replacement rate, i.e. benefits b relative to the average wage w (normalized to 1 in the calibration). The lay-off tax coverage target is reported as $\frac{f}{b}F^{\text{cal}}$, i.e. what fraction of expected UI payments to a laid-off worker is covered by the lay-off tax F.

7. In the model, the autocorrelation of productivity is given by $1-\lambda$. Therefore, $\lambda = 0.07$ implies a monthly autocorrelation of 0.93, which implies a yearly autocorrelation of 0.42.

For the purpose of calibration, we set p to 0.5. We target a UI replacement rate $b^{\text{rep}} = \frac{b}{w}$ of 0.45 as reported by Engen and Gruber (2001). Following Jung and Kuester (2015), we target a lay-off tax coverage F^{rep} of 0.5, which means that the lay-off tax covers 50% of the expected UI payments transferred to a worker after a lay-off.[8] Given these policy parameter targets to inform the calibration of b and F, we calibrate the remaining parameters (matching efficiency $\bar{m}$, cost shock scale c_f, and vacancy posting cost k_v) to match standard worker flow targets. Note that since ρ is the probability of separation conditional on a new shock arriving, the effective separation rate ρ^{eff} is different from ρ. It is given by $\rho^{\text{eff}} = \frac{n \cdot \lambda \cdot \rho + (1-n) \cdot f \cdot \rho}{n_0}$, where n_0 is the mass of matches at the beginning of a period. It includes the mass of already existing matches n and new matches $(1-n) \cdot f$. Existing matches draw a new shock with probability λ, while all new matches draw a productivity shock and, like every other match, will separate with probability ρ once productivity realizes.

Table 3.2 summarizes the calibrated and set parameters of the model.

Table 3.2. Model Parameters: Calibrated and Set Values

Parameter	Description	Value
Calibrated Parameters		
$\bar{m}$	Matching efficiency parameter	0.383
c_f	Cost shock parameter	0.407
k_v	Vacancy posting cost	0.086
b	Unemployment benefit level	0.450
$\bar{h}$	Regular working hours level	0.953
F^{cal}	Lay-off tax	0.549
Set Parameters		
ϕ	Coefficient of relative risk aversion	2
α	Labor elasticity in production function	0.65
γ	Elasticity of matching w.r.t. unemployment	0.65
η	Worker bargaining power	0.65
ψ	Disutility of labor parameter	1.5
λ	Shock arrival probability	0.07
μ	Location of productivity shocks	-0.80
σ	Scale of productivity shocks	0.12
p	Fraction of financially constrained firms	0.5

8. Since the expected duration of unemployment is $1/f$, the expected UI payments per unemployed worker are given by $\frac{b}{f}$, implying that $F^{\text{rep}} = \frac{F \cdot f}{b}$.

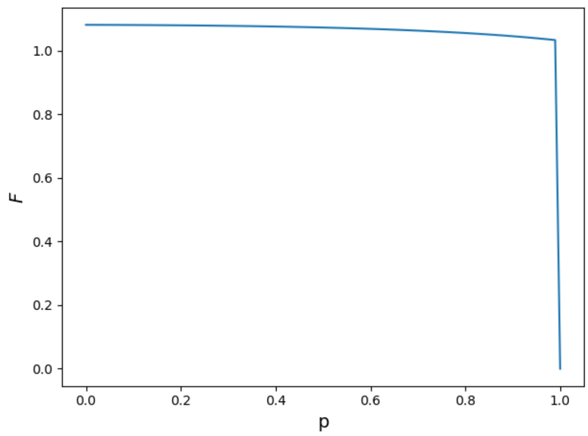
3.5 Quantitative Results

3.5.1 Taking UI as given

As a first benchmark, we compute the optimal lay-off tax and short-time work parameters chosen by a Ramsey planner who takes the calibrated UI system (i.e. $b = 0.45$) as given. To calculate the Ramsey optimal policy parameters, we solve the Ramsey planner's maximization as described in Section 3.3.1 numerically. The details are described in Section 3.J in the appendix.

Figure 3.1 shows the results for the optimal lay-off tax F across $p \in [0, 1]$. The result is what FOC pinning down the optimal lay-off tax (Equation 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 in $BE_{\mathrm{lt}}(b)$. At $p = 1$, lay-off taxes lose their bite completely as they can no longer target any inefficiencies. 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.

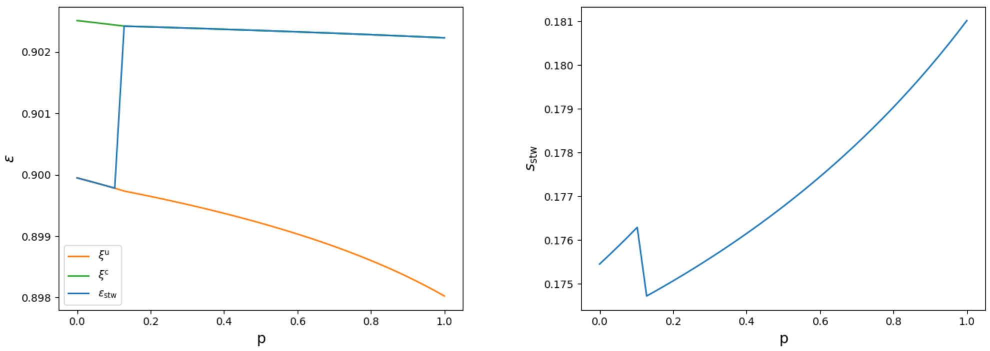
Figure 3.1. Optimal Lay-Off Tax



Notes: Optimal lay-off tax F set by the Ramsey planner across $p \in [0, 1]$. The calibrated UI parameter of $b = 0.45$ is taken as given.

Figure 3.2 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 3.2, the benefits from protecting constrained firms eventually outweigh the added distortion from protecting excess unconstrained firms. At this point ($p = 0.128$), the eligibility condition jumps to the threshold ξ_s^c at which constrained firms without short-time work separate.

Figure 3.2. Optimal Short-time Work



(a) Optimal ε_{stw} *(b) Optimal s_{stw}*

Notes: Optimal eligibility threshold ε_{stw} and optimal short-time work benefits s_{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. The calibrated UI parameter of $b = 0.45$ is taken as given.

The optimal generosity parameter s_{stw} is increasing in p except at the kink at $p = 0.128$. 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 s_{stw} has to rise in p as long as ϵ_{stw} does not change much. Additionally, the planner can provide more insurance with higher s_{stw} which becomes more valuable as p increases, also contributing to the upward slope. 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 s_{stw} as the more lenient eligibility threshold is introduced, explaining the kink.

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 3.2 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 slopes are different for the two policy regimes: as p increases, lay-off taxes gradually lose effectiveness, leading to a faster decline of welfare. The two curves have different slopes and intersect at $p = 0.275$. The interpretation is simple: If more than 27% 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 3.3 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

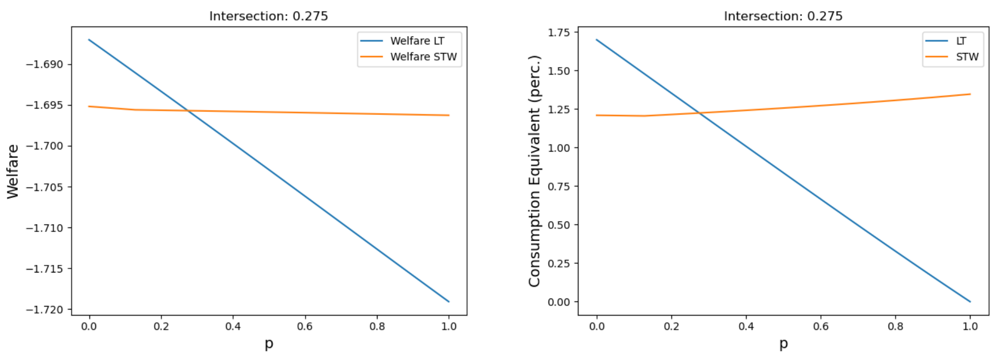
$$u(c_0 \cdot (1 + \Delta^{\text{policy}})) \;=\; u(c_{\text{policy}})$$

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 (with the calibrated UI parameter $b = 0.45$), 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 1.25-1.75%. Secondly, as p increases, short-time work becomes not only preferable to lay-off taxes, but improvements in terms of consumption equivalent variation even increase. This stems from the effect that short-time work helps firms provide insurance, which becomes increasingly important for higher p.

Aggregate outcomes of the economies associated with both policy regimes are shown in Figure 3.B.1 in Appendix 3.B.

Figure 3.3. Welfare Comparison



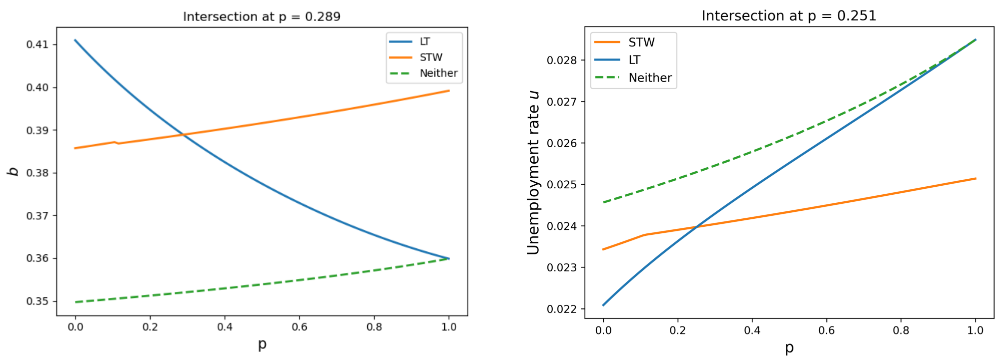
(a) Welfare *(b) Consumption Equivalent*

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. The calibrated UI parameter of $b = 0.45$ is taken as given. 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, but with the calibrated UI parameter in terms of consumption equivalent variation.

3.5.2 Full Optimization

We now turn to the full optimization problem, where the Ramsey planner optimally chooses either lay-off tax or short-time work parameters simultaneously with the unemployment benefit. As a first benchmark, we compute how the Ramsey planner sets the optimal unemployment benefit b absent both lay-off taxes and short-time work. Figure 3.4 shows the result as a green dashed line. Without lay-off taxes that could counteract the fiscal externality of the unemployment benefit, the optimal unemployment benefit b is set to a lower level compared to the calibrated value of 0.45, ranging between 0.35 and 0.36. It slightly increases in p, indicating that the planner values higher insurance more than lowering unemployment when faced with more separations and therefore higher unemployment because of more constrained firms.

Figure 3.4. Optimal Unemployment Benefits



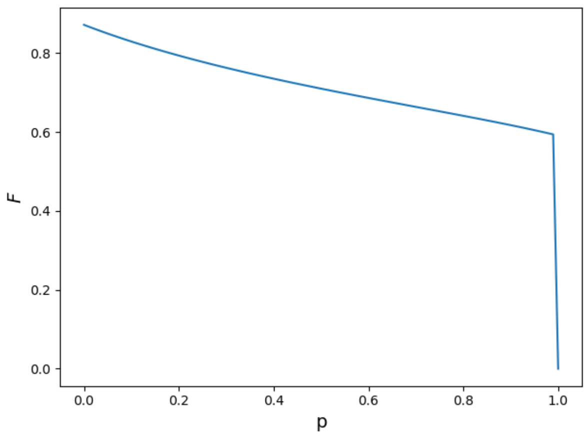
(a) Optimal Unemployment Benefits b *(b) Unemployment Rate u*

Notes: The optimal unemployment benefit b set by the Ramsey planner across $p \in [0, 1]$ without either lay-off taxes or short-time work (dashed green line), while simultaneously setting the lay-off tax (blue line) and while simultaneously setting short-time work (orange line) is shown in plot (a) on the left. The resulting unemployment rate u across p in the three cases is shown in plot (b) on the right.

Having established how unemployment benefits are set in the absence of supplementary policies, we now turn to the problem of jointly optimizing the unemployment benefit b and the lay-off tax F. The first thing to note is that b is set considerably higher than without supplementary policies for $p < 1$ as shown by the blue line in Figure 3.4. This is because the planner now has the lay-off tax at her disposal to counteract the fiscal externality of the unemployment benefit, which allows her to set a higher unemployment benefit and thus provide more insurance to workers. The second thing to note is that b sharply decreases with p until it reaches the same level as without supplementary policies at $p = 1$. The reason behind this downward slope is that with more constrained firms, the lay-off tax that can only act on unconstrained matches becomes less and less effective at counteracting the fiscal externality of the unemployment benefit. This forces the planner to set a lower unemployment benefit to reduce the externality. Indeed, as the right panel of Figure 3.4 shows, the unemployment rates associated with the regime with no supplementary policies and the regime with the optimal lay-off tax converge as p increases until the two curves meet at $p = 1$ as well.

A secondary effect of this is that the optimal lay-off tax F shown in Figure 3.5 also decreases in p as the planner sets lower and lower unemployment benefits and thus has to counteract a smaller and smaller fiscal externality in unconstrained matches. At $p = 1$, the lay-off tax loses its bite completely, and the planner sets the same unemployment benefit as without the option to set a lay-off tax.

Figure 3.5. Optimal Lay-Off Tax

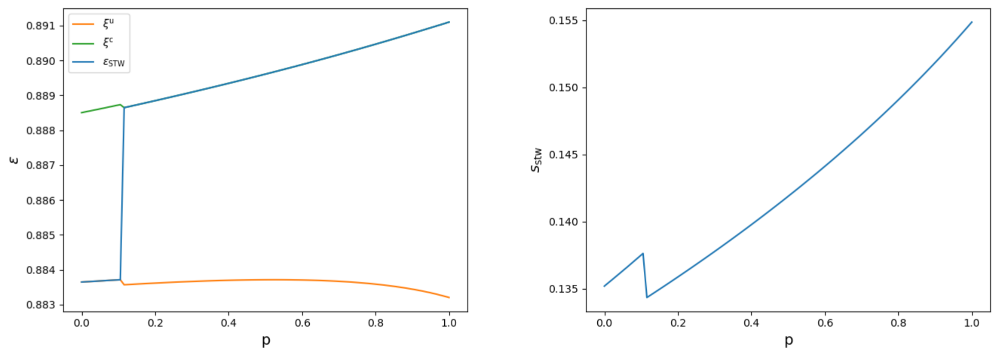


Notes: The optimal lay-off tax F across $p \in [0, 1]$ when the Ramsey planner optimally chooses the optimal lay-off tax F and the optimal unemployment benefit b simultaneously.

Next, we turn to the joint optimization of the unemployment benefit b and the short-time work parameters ϵ_{stw} and s_{stw}. Like with the lay-off tax, the optimal unemployment benefit b with short-time work shown in Figure 3.4 as an orange line is set at a higher level than without supplementary policies. Unlike with the lay-off tax, however, the optimal unemployment benefit b increases with p. The reason for this pattern is that with more constrained firms, more matches are at risk of separation, leading to higher welfare gains from providing insurance through the unemployment benefit. Since short-time work remains effective at counteracting the fiscal externality of the unemployment benefit even with more constrained firms, the planner is not kept from increasing b by increasing the number of separations as with the lay-off tax. As the right panel of Figure 3.4 shows, the unemployment rate under the short-time work regime remains much lower for higher p than in the other regimes as a result.

This dynamic also explains how the optimal short-time work parameters ϵ_{stw} and s_{stw} shown in Figure 3.6 are set differently than in the case without optimizing

Figure 3.6. Optimal Short-time Work



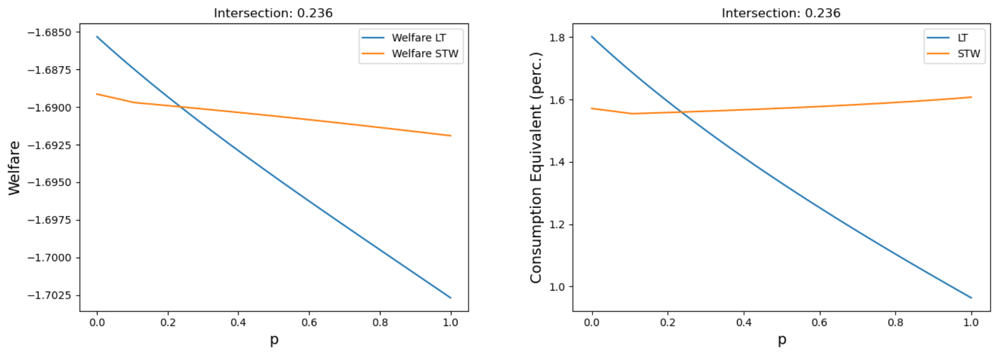
(a) Optimal ε_{stw} *(b) Optimal s_{stw}*

Notes: Optimal eligibility threshold ε_{stw} and optimal short-time work benefits s_{stw} set by the Ramsey planner across $p \in [0, 1]$ while setting the optimal unemployment benefit b simultaneously. 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.

b discussed in the previous section. The qualitative patterns of the optimal eligibility thresholds persist: the planner still initially sets $\epsilon_{\text{stw}} = \xi_s^u$ before transitioning to $\epsilon_{\text{stw}} = \xi_s^c$ as p increases for the same reasons described in Section 3.5.1. However, there are quantitative differences. Both ξ_s^u and with it the eligibility threshold, now increase with p, reflecting that higher unemployment benefits lead to higher separation thresholds and thus firms are in need of support sooner. Meanwhile, the optimal short-time work generosity parameter s_{stw} increases far more steeply in p than in the case where b is held fixed. This sharp increase arises because the higher unemployment benefits amplify the fiscal externality that the planner must counteract, necessitating a much higher short-time work generosity to keep the average net subsidy at the optimal level.

Finally, we compare the welfare of the optimal lay-off tax and the optimal short-time work when the Ramsey planner optimally sets the unemployment benefit b simultaneously. Figure 3.7 shows the results. As expected, the welfare of both optimal policies is higher than without optimizing b. The improvement over an economy with calibrated UI parameters but without either lay-off taxes or short-time work in terms of consumption-equivalent variation is substantial especially for higher values of p. The improvement over the benchmark economy with only the calibrated unemployment insurance and neither short-time work nor lay-off taxes is around 1.6% with short-time work instead of around 1.25% and is close to 1.8% with lay-off taxes instead of around 1.75% for $p = 0$. In particular, in the case of full optimization, short-time work becomes the superior policy instrument at an even lower share of financially constrained firms of $p = 0.236$ compared to $p = 0.275$ when b is held fixed.

Figure 3.7. Welfare Comparison



(a) Welfare *(b) Consumption Equivalent*

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, while the Ramsey planner optimally chooses the optimal unemployment benefit b simultaneously. Plot (b) on the right shows improvements of optimal short-time work and optimal lay-off taxes over the economy without either of short-time work or lay-off taxes, as well as without setting b optimally, but with the calibrated UI parameter in terms of consumption equivalent variation.

3.6 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 27% or more of firms are financially constrained, short-time work dominates lay-off taxes in welfare terms. If the level of unemployment benefits is optimized simultaneously, this threshold falls to 24%.

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 believe this offers a promising direction for future work.

3. Appendix

3.A Table of Model Parameters, Variables and Functions

Table 3.A.1. Model Parameters, Variables and Functions Part 1

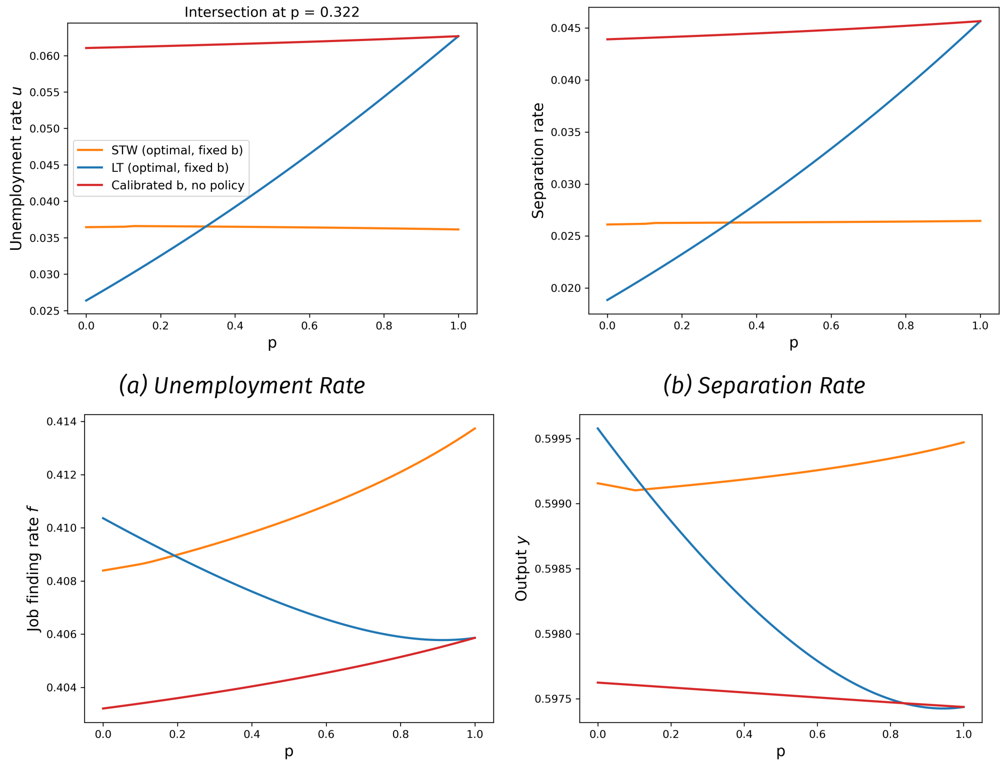
Symbol	**Description**
n	Mass of employed workers
$u = 1 - n$	Mass of unemployed workers
v	Mass of posted vacancies
$\theta = \frac{v}{1-n}$	Labor market tightness
f	Job-finding rate
q	Vacancy-filling rate
ρ	Separation rate
ρ^u, ρ^c	Separation rates (unconstrained/ constrained)
ρ^F	Separation with lay-off 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
$\bar{m}, \gamma$	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
s	Lump-sum tax per shock
s_{stw}	STW subsidy per hour gap
k_v	Cost of posting a vacancy
η	Bargaining power of workers
ϕ	Risk aversion

Table 3.A.2. Model Parameters, Variables and Functions Part 2

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

3.B Aggregate Outcomes

Figure 3.B.1. Aggregate Outcomes



(a) Unemployment Rate

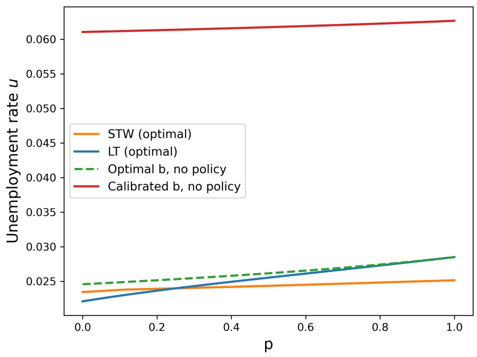
(b) Separation Rate

(c) Job Finding Rate

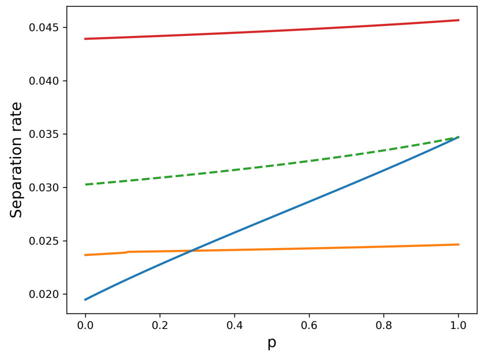
(d) Output

Notes: Aggregate outcomes for the different policy regimes taking the calibrated unemployment benefit level of $b = 0.45$ as given against the share of constrained firms, p. The red lines are associated with a regime with no lay-off tax and no short-time work, the blue lines with a regime with optimal UI benefits, and the orange lines with optimal short-time work. **(a)** Unemployment rate, **(b)** separation rate, **(c)** job finding rate and **(d)** output.

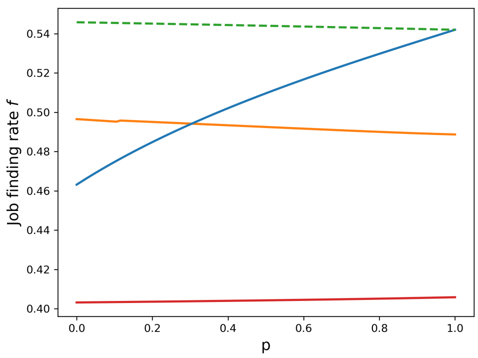
Figure 3.B.2. Aggregate Outcomes - Full Optimization



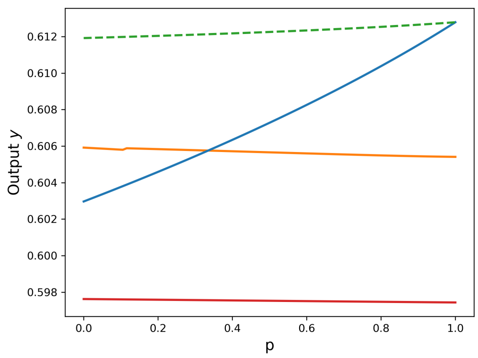
(a) Unemployment Rate



(b) Separation Rate



(c) Job Finding Rate



(d) Output

Notes: Aggregate outcomes for the different policy regimes (calibrated UI, optimal UI, optimal UI with optimal LT and optimal UI with optimal STW) as a function of the share of constrained firms, p. **(a)** Unemployment rate, **(b)** separation rate, **(c)** job finding rate and **(d)** output.

3.C Equilibrium Lay-off Tax

3.C.1 Bargaining

Nash-Bargaining Problem:

$$\max_{w^u(\epsilon),w^c(\epsilon),h^u(\epsilon),h^c(\epsilon),\epsilon_s^u,\epsilon_s^c} \quad \bar{J}^{(1-\eta)}\left(\bar{V}-U\right)^{\eta}$$

subject to

(1) Financial constraint of financially constrained firms:

$$y(\epsilon,h(\epsilon)) - w^c(\epsilon) + \beta \cdot \left[\lambda \cdot \bar{J} + (1-\lambda)\cdot J^c(\epsilon)\right] \geq 0$$

(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

$$z(\epsilon) - c^c(\epsilon) + \beta \cdot \left[\lambda \cdot \bar{J} + (1-\lambda)\cdot J^c(\epsilon)\right] > 0$$

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:

$$u(c^i(\epsilon)) - u(b) + (\lambda - f)\cdot\beta\cdot(\bar{V}-U) \geq 0$$

Let $\epsilon_s^{i,\mathrm{cp}}$ denote the threshold at which the worker wants to separate unitarily:

$$u(c^i(\epsilon_s^{i,\mathrm{cp}})) - u(b) + (\lambda - f)\cdot\beta\cdot(\bar{V}-U) = 0$$

We can rewrite the value functions of the firm and the surplus of the worker as:

$$\begin{aligned}
\bar{J} \;=\; & -\tau \\
& + (1-p)\cdot\int_{\max\{\epsilon_s^{u,B},\epsilon_s^{u,\mathrm{cp}}\}}^{\infty} \frac{z(\epsilon) - c^u(\epsilon) + \lambda\cdot\beta\cdot\bar{J}}{1-\beta\cdot(1-\lambda)} dG(\epsilon) \\
& + (1-p)\cdot G(\max\{\epsilon_s^{u,B},\epsilon_s^{u,\mathrm{cp}}\})\cdot\beta\cdot F \\
& + p\cdot\int_{\max\{\epsilon_s^{c,B},\epsilon_s^{c,\mathrm{cp}}\}}^{\infty} \frac{z(\epsilon) - c^c(\epsilon) + \lambda\cdot\beta\cdot\bar{J}}{1-\beta\cdot(1-\lambda)} dG(\epsilon) \\
& + p\cdot G(\max\{\epsilon_s^{c,B},\epsilon_s^{c,\mathrm{cp}}\})\cdot\beta\cdot F \\
(\bar{V}-U) \;=\; & (1-p)\cdot\int_{\max\{\epsilon_s^{u,B},\epsilon_s^{u,\mathrm{cp}}\}}^{\infty} \frac{u(c^u(\epsilon)) - u(b) + \lambda\cdot\beta\cdot(\bar{V}-U)}{1-\beta\cdot(1-\lambda)} dG(\epsilon) \\
& + p\cdot\int_{\max\{\epsilon_s^{c,B},\epsilon_s^{c,\mathrm{cp}}\}}^{\infty} \frac{u(c^c(\epsilon)) - u(b) + \lambda\cdot\beta\cdot(\bar{V}-U)}{1-\beta\cdot(1-\lambda)} dG(\epsilon)
\end{aligned}$$

Note that $\epsilon_s^{i,B}$ denotes the separation threshold in the bargaining outcome. Separations occurs either when the productivity falls below the bargained separation threshold or when the commitment constrained of the worker is violated $\epsilon < \max\{\epsilon_s^{c,B}, \epsilon_s^{c,\text{cp}}\}$. Finally, note that it is equivalent to maximize over $w^i(\epsilon)$ and $c^i(\epsilon)$ as

$$c^i(z) = w^i(\epsilon) - \phi(h^i(\epsilon)).$$

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

$$\begin{aligned}\mathscr{B} = {} & \left(\bar{J}\right)^{\eta} \cdot \left(\bar{V} - U\right)^{1-\eta} \\ & - \int_0^{\infty} \mathscr{M}(\epsilon)\left[z(\epsilon) - c^c(\epsilon) + \beta \cdot \left[\lambda \cdot \bar{J} + (\lambda - f) \cdot J^c(\epsilon)\right]\right] d\epsilon\end{aligned}$$

With Kuhn-Tucker conditions for a maximum:

$$\begin{aligned}(I) \quad & \mathscr{M}(\epsilon) \le 0 \\ (II) \quad & \mathscr{M}(\epsilon) \cdot \left[z(\epsilon) - c^c(\epsilon) + \beta \cdot \left[\lambda \cdot \bar{J} + (1-\lambda) \cdot J^c(\epsilon)\right]\right] = 0\end{aligned}$$

$$\begin{aligned}\frac{\partial \mathscr{B}}{\partial c^u(\epsilon)} &= -(1-\eta) \cdot (1-p) \cdot g(\epsilon) \cdot \left(\frac{\bar{V} - U}{\bar{J}}\right)^{\eta} \\ &\quad + \eta \cdot (1-p) \cdot g(\epsilon) \cdot u'\left(c^u(\epsilon)\right) \cdot \left(\frac{\bar{J}}{\bar{V} - U}\right)^{1-\eta} \\ &= 0\end{aligned}$$

$$\Longleftrightarrow \quad \frac{1-\eta}{\eta} \cdot \frac{\bar{V} - U}{\bar{J}} = u'\left(c^u(\epsilon)\right) \quad \Rightarrow \quad u'\left(c^u(\epsilon)\right) = u'(c^w)$$

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:

$$J = (1-\eta) \cdot \bar{S}, \quad \frac{\bar{V} - U}{u'(c^u)} = \eta \cdot \bar{S}$$

$$\frac{\partial \mathscr{B}}{\partial c^c(\epsilon)} = -(1-p)\cdot g(\epsilon)\cdot(1-\eta)\cdot\left(\frac{\bar{V}-U}{\bar{S}}\right)^{\eta} + (1-p)\cdot g(\epsilon)\cdot u'\left(c^u(\epsilon)\right)\cdot U\cdot\left(\frac{J}{\bar{V}-U}\right)^{1-\eta} + \mathscr{M}(\epsilon) \stackrel{!}{=} 0$$

Insert the surplus splitting rule:

$$\Rightarrow \quad -(1-p)\cdot g(\epsilon)\cdot\left(\frac{\eta}{1-\eta}\cdot u'(c^w)\right)^{\eta} + (1-p)\cdot g(\epsilon)\cdot u'\left(c^c(\epsilon)\right)\cdot\eta + \mathscr{M}(\epsilon)\cdot\left(u'(c^w)\cdot\frac{1-\eta}{\eta}\right)^{1-\eta} = 0$$

$$\Leftrightarrow \ (1-p)\cdot g(\epsilon)\cdot\eta\cdot\left(u'(c^c(\epsilon))-u'(c^w)\right) = -\mathscr{M}(\epsilon)\cdot\left(u'(c^u)\cdot\frac{\eta}{1-\eta}\right)^{1-\eta}$$

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

$$z(\epsilon) - c^w + \beta\cdot\left[\lambda\cdot\bar{J} + (1-\lambda)\cdot J^c(\epsilon)\right] \geq 0$$

then the constraint is non-binding and $\mathscr{M}(\epsilon) = 0$. ✓

If

$$z(\epsilon) - c^w + \beta\cdot\left[\lambda\cdot\bar{J} + (1-\lambda)\cdot J^c(\epsilon)\right] < 0$$

then the constraint is binding and $\mathscr{M}(\epsilon) < 0$. ✓
Thus, it is optimal for the firm to pay as much as it can:

$$J^c(\epsilon) = z(\epsilon) - c^c(\epsilon) + \beta\cdot\left[\lambda\cdot\bar{J} + (1-\lambda)\cdot J^c(\epsilon)\right] = 0$$

$$\Leftrightarrow \quad c^c(\epsilon) = z(\epsilon) + \lambda\cdot\beta\cdot\bar{J}$$

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:

$$
\begin{aligned}
\bar{S} =\ & -\tau \\
& + (1-p) \cdot \int_{\max\{\epsilon_s^{u,B}, \epsilon_s^{u,\text{cp}}\}}^{\infty} \frac{z(\epsilon) + \frac{u(c^w)-u(b)}{u'(c^w)} - c^w + (1-\eta \cdot f) \cdot \beta \cdot \bar{S}}{1-\beta \cdot (1-\lambda)} dG(\epsilon) \\
& - (1-p) \cdot G(\max\{\epsilon_s^{u,B}, \epsilon_s^{u,\text{cp}}\}) \cdot \beta \cdot F \\
& + p \cdot \int_{\epsilon^p}^{\infty} \frac{z(\epsilon) + \frac{u(c^w)-u(b)}{u'(c^w)} - c^w + (1-\eta \cdot f) \cdot \beta \cdot \bar{S}}{1-\beta \cdot (1-\lambda)} dG(\epsilon) \\
& + p \cdot \int_{\max\{\epsilon_s^{c,B}, \epsilon_s^{c,\text{cp}}\}}^{\epsilon_p} \frac{z(\epsilon) + \frac{u(c^c(\epsilon))-u(b)}{u'(c^w)} - c^c(\epsilon) + (1-\eta \cdot f) \cdot \beta \cdot \bar{S}}{1-\beta \cdot (1-\lambda)} dG(\epsilon) \\
& - (1-p) \cdot G(\max\{\epsilon_s^{c,B}, \epsilon_s^{c,\text{cp}}\}) \cdot \beta \cdot F
\end{aligned}
$$

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 $\epsilon_s^{u,B} > \epsilon_s^{u,cp}$

$$
\begin{aligned}
\frac{\partial \mathscr{B}}{\partial \epsilon_s^{u,b}} =\ & -(1-p) \cdot g(\epsilon_s^{u,b}) \cdot \Bigg[\frac{1}{1-\beta \cdot (1-\lambda)} \cdot \Big(z(\epsilon_s^{u,b}) + \frac{u(c^w)-u(b)}{u'(c^w)} \\
& - c^w + (\lambda - \eta f) \cdot \beta \cdot \bar{S} \Big) + \beta \cdot F \Bigg] = 0
\end{aligned}
$$

$$
\iff \quad z(\epsilon_s^{u,b}) + \big(1-\beta \cdot (1-\lambda)\big) \cdot \beta \cdot F + \frac{u(c^w)-u(b)}{u'(c^w)} - c^w + (\lambda - \eta f) \cdot \beta \cdot \bar{S} = 0
$$

Insert free-entry Condition:

$$
z(\epsilon_s^{u,b}) + \big(1-\beta \cdot (1-\lambda)\big) \cdot \beta \cdot F + \frac{u(c^w)-u(b)}{u'(c^w)} - c^w + \frac{\lambda - \eta f}{1-\eta} \cdot \left(\frac{k_v}{q} \right) = 0
$$

Note that workers do not have a commitment problem $\epsilon_s^{u,B} > \epsilon_s^{u,cp}$, as workers would never quit under the insurance constraint of the firm:

$$V(\epsilon) - U \;=\; u(c^w) - u(b) + (\lambda - f) \cdot \beta \cdot (\bar{V} - U) \;>\; 0$$

Workers in unconstrained firms always have a positive surplus!
Suppose that $\epsilon_s^{c,B} \geq \epsilon_s^{c,cp}$

$$\frac{\partial \mathscr{B}}{\partial \epsilon_s^{c,B}} \;=\; -(1-p) \cdot g(\epsilon_s^{c,B}) \cdot \left[\frac{1}{1 - \beta \cdot (1-\lambda)} \cdot \left(z(\epsilon_s^{c,B}) + \frac{u\left(c^c(\epsilon_s^{c,B})\right) - u(b)}{u'(c^w)} \right. \right.$$
$$\left. \left. - c^c(\epsilon_s^{c,B}) + (\lambda - \eta f) \cdot \beta \cdot \bar{S} \right) + F \right] \;=\; 0$$

$$\iff \quad z(\epsilon_s^{c,B}) + \left(1 - \beta \cdot (1-\lambda)\right) \cdot \beta \cdot F + \frac{u(c^c(\epsilon_s^{c,B})) - u(b)}{u'(c^w)} - c^c(\epsilon_s^{c,B}) + (\lambda - \eta f) \cdot \beta \cdot \bar{S} \;=\; 0$$

Insert free-entry condition:

$$z(\epsilon_s^{c,B}) + \left(1 - \beta \cdot (1-\lambda)\right) \cdot \beta \cdot F + \frac{u(c^c(\epsilon_s^{c,B})) - u(b)}{u'(c^w)} - c^c(\epsilon_s^{cp}) + \frac{\lambda - \eta f}{1-\eta} \cdot \frac{k_v}{q} \;=\; 0$$

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 paying future lay-off taxes, firms would like workers to stay within the firm for negative surplus values. Thus, the worker decides to leave the firm unilaterally $\epsilon_s^{c,B} < \epsilon_s^{c,cp}$.

3.C.2 Job Creation and Wage Equation

The notation of this section follows Appendix 3.E for lay-off taxes.

(I) Value of an unconstrained firm after the idiosyncratic productivity shock has realized:

$$\begin{aligned} J^u(\epsilon) \;&=\; z(\epsilon) - c^w + \beta \cdot \left[\lambda \bar{J} + (1-\lambda) \cdot J^u(\epsilon)\right] \\ \Leftrightarrow \quad J^u(\epsilon) \;&=\; \frac{1}{1 - \beta \cdot (1-\lambda)} \left[z(\epsilon) - c^w + \lambda \cdot \beta \cdot \bar{J}\right] \end{aligned}$$

(II) Value of a constrained firm without binding constraints after the idiosyncratic productivity shock has realized:

$$\begin{aligned} J^c(\epsilon) \;&=\; z(\epsilon) - c^w + \beta \cdot \left[\lambda \bar{J} + (1-\lambda) \cdot J^c(\epsilon)\right] \\ \Leftrightarrow \quad J^c(\epsilon) \;&=\; \frac{1}{1 - \beta \cdot (1-\lambda)} \left(z(\epsilon) - c^w + \lambda \cdot \beta \cdot \bar{J}\right) \end{aligned}$$

(III) Value of a constrained firm with binding constraints after the idiosyncratic productivity shock has realized:

$$\begin{aligned} J^c(\epsilon) &= z(\epsilon) - c^w(\epsilon) + \beta \cdot \left[\lambda \bar{J} + (1-\lambda) \cdot J^c(\epsilon)\right] \\ \Leftrightarrow \quad J^c(\epsilon) &= \frac{1}{1-\beta \cdot (1-\lambda)} \left(z(\epsilon) - c^w(\epsilon) + \lambda \cdot \beta \cdot \bar{J}\right) \end{aligned}$$

(IV) Value of a firm before the idiosyncratic productivity shock has realized:

$$\bar{J} = -\tau + (1-p) \cdot \int_{\epsilon_s^u}^{\infty} J^u(\epsilon)\, dG(\epsilon) + p \cdot \int_{\epsilon_s^c}^{\infty} J^c(\epsilon)\, dG(\epsilon) - \rho \cdot F$$

Inserting (I)-(III) into (IV) gives:

$$\begin{aligned} \Leftrightarrow \quad \bar{J} = -\tau + \frac{1}{1-\beta \cdot (1-\lambda)} \Big[& (1-\rho) \cdot z - (1-p) \cdot (1-\rho^u) \cdot c^w - p \cdot (1-\rho^c) \cdot e^c \\ & + (1-p) \cdot \lambda \cdot \beta \cdot \bar{J} \Big] - (1-p) \cdot \rho^u \cdot F \end{aligned}$$

Here z denotes the average production (without distortions) and e^c denotes the average cost of a constrained worker for a firm. Both are defined in section 3.E of the appendix.

(V) Value of an unconstrained worker after the idiosyncratic productivity shock has realized:

$$\begin{aligned} V^u(\epsilon) &= u(c^w) + \beta \cdot \left[\lambda \bar{V} + (1-\lambda) \cdot V^u(\epsilon)\right] \\ \Leftrightarrow \quad V^u(\epsilon) &= \frac{1}{1-\beta \cdot (1-\lambda)} \left(u(c^w) + \lambda \cdot \beta \cdot \bar{V}\right) \end{aligned}$$

(VI) Value of a constrained worker without binding constraints after the idiosyncratic productivity shock has realized:

$$\begin{aligned} V^c(\epsilon) &= u(c^w) + \beta \cdot \left[\lambda \bar{V} + (1-\lambda) \cdot V^c(\epsilon)\right] \\ \Leftrightarrow \quad V^c(\epsilon) &= \frac{1}{1-\beta \cdot (1-\lambda)} \left(u(c^w) + \lambda \cdot \beta \cdot \bar{V}\right) \end{aligned}$$

(VII) Value of a constrained worker with binding constraints after the idiosyncratic productivity shock has realized:

$$\begin{aligned} V^c(\epsilon) &= u(c^w(\epsilon)) + \beta \cdot \left[\lambda \bar{V} + (1-\lambda) \cdot V^c(\epsilon)\right] \\ \Leftrightarrow \quad V^c(\epsilon) &= \frac{1}{1-\beta \cdot (1-\lambda)} \left(u(c^w(\epsilon)) + \lambda \cdot \beta \cdot \bar{V}\right) \end{aligned}$$

(VIII) Value of a worker before the idiosyncratic productivity shock has realized:

$$\bar{V} = (1-p) \cdot \int_{\epsilon_s^u}^{\infty} V^u(\epsilon)\, dG(\epsilon) + p \cdot \int_{\xi_s^c}^{\infty} V^c(\epsilon)\, dG(\epsilon) + \rho \cdot U$$

Inserting (V)-(VII) into (VIII) gives:

$$\Leftrightarrow \quad \bar{V} = \frac{1}{1-\beta\cdot(1-\lambda)}\Big[(1-p)(1-\rho^u)\cdot u(c^w) + p\cdot(1-\rho^c)\cdot u^c + \beta\cdot\big[(1-\rho)\cdot\lambda\cdot\beta\cdot\bar{V} + \rho\cdot U\big]\Big]$$

u^c is defined as the average utility of a worker in a constrained firm. Its definition can be found in appendix 3.E.

(IX) Unemployment:

$$U = u(b) + \beta\cdot\big[f\cdot\bar{V} + (1-f)\cdot U\big]$$

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:

$$\begin{aligned}
& V^i(\epsilon) - U = u(c^i(\epsilon)) - u(b) + (\lambda - f)\cdot V^i(\epsilon) + (\lambda - f)\cdot\beta\cdot\bar{V} - (\lambda - f)\cdot\beta\cdot U \\
\Leftrightarrow \quad & V^i(\epsilon) - U = u(c^i(\epsilon)) - u(b) + (\lambda - f)\cdot\beta\cdot(V^i(\epsilon) - U) + (\lambda - f)\cdot\beta\cdot(\bar{V} - U) \\
\Leftrightarrow \quad & V^i(\epsilon) - U = \frac{1}{1-\beta\cdot(1-\lambda)}\Big(u(c^i(\epsilon)) - u(b) + (\lambda - f)\cdot\beta\cdot(\bar{V} - U)\Big)
\end{aligned}$$

B) Next, we calculate the expected surplus of the worker, before the shocks have been realized:

$$\begin{aligned}
\bar{V} - U &= (1-\rho)\cdot\int_{\epsilon_s^u}^{\infty}\Big(V^u(\epsilon) - U\Big)\,dG(\epsilon) + p\cdot\int_{\epsilon_s^c}^{\infty}\Big(V^c(\epsilon) - U\Big)\,dG(\epsilon) \\
&= \frac{1}{1-\beta\cdot(1-\lambda)}\Bigg[(1-\rho)\cdot(1-\rho^u)\cdot\Big(u(c^u) - u(b)\Big) \\
&\quad + p\cdot(1-\rho^c)\cdot\Big(u(c^c) - u(b)\Big) \\
&\quad + (1-\rho)\cdot(\lambda - f)\cdot\beta\cdot(\bar{V} - U)\Bigg]
\end{aligned}$$

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:

$$\begin{aligned}
S &= \bar{J} + \frac{\bar{V} - U}{u'(c^u)} \\
&= -\tau + \frac{1}{1-\beta\cdot(1-\lambda)}\Bigg[(1-\rho)\cdot z \\
&\quad + (1-\rho)\cdot(1-\rho^u)\cdot\left(\frac{u(c^u) - u(b)}{u'(c^u)} - c^u\right) \\
&\quad + p\cdot(1-\rho^c)\cdot\left(\frac{u(c^c) - u(b)}{u'(c^u)} - e^c\right) \\
&\quad + (1-\rho)\cdot\lambda\cdot\beta\cdot\bar{J} + (1-\rho)\cdot(\lambda - f)\cdot\beta\cdot\frac{\bar{V} - U}{u'(c^u)}\Bigg] - \rho\cdot\beta\cdot F
\end{aligned}$$

Using the surplus splitting rule, we can rewrite the equation above as:

$$\begin{aligned} S \; = \; & -\tau + \frac{1}{1-\beta \cdot (1-\lambda)} \Bigg[(1-\rho) \cdot z \\ & + (1-\rho) \cdot (1-\rho^u) \cdot \left(\frac{u(c^u) - u(b)}{u'(c^u)} - c^u \right) \\ & + p \cdot (1-\rho^c) \cdot \left(\frac{u(c^c) - u(b)}{u'(c^u)} - e^c \right) \\ & + (1-\rho) \cdot (\lambda - \eta f) \cdot \beta \cdot S \Bigg] - \rho \cdot \beta \cdot F \end{aligned}$$

Next, we want to replace the tax in the surplus equation. We can do this by using the budget constraint of the government:

$$\begin{aligned} n^s \cdot \tau \; &= \; -n^s \cdot \rho \cdot F + (1-n) \cdot b \\ \Leftrightarrow \qquad \tau \; &= \; -\rho \cdot F + \frac{1-n}{n^s} \cdot b \\ \text{Insert } n^s \; &= \; \frac{\lambda}{1-\rho} \cdot n \\ \Rightarrow \qquad \tau \; &= \; \rho \cdot F + \frac{1}{\lambda} \cdot (1-\rho) \cdot \frac{1-n}{n} \cdot b \end{aligned}$$

Inserting it into the equation for the joint surplus gives:

$$\begin{aligned} S \; = \; & \frac{1}{1-\beta \cdot (1-\lambda)} \Bigg[(1-\rho) \cdot \left(z - \frac{1-n}{n} \cdot b \right) + (1-p) \cdot (1-\rho^u) \cdot \left(\frac{u(c^w) - u(b)}{u'(c^w)} - c^w \right) \\ & + p \cdot (1-\rho^c) \cdot \left(\frac{u(c^c) - u(b)}{u'(c^w)} - e^c \right) + (1-\rho) \cdot (\lambda - \eta f) \cdot \beta \cdot S \Bigg] \\ & - \frac{1}{\lambda} \cdot (1-\rho) \cdot \frac{1-n}{n} \cdot b + \rho \cdot (1-\beta) \cdot F \end{aligned}$$

Next, insert the free-entry condition: $\frac{k_v}{q} = \beta \cdot \bar{J}$

$$\begin{aligned} \frac{1}{1-\eta} \cdot \frac{k_v}{q} \; = \; & \frac{\beta}{1-\beta \cdot (1-\lambda)} \cdot \Bigg[(1-\rho) \cdot z \\ & + (1-p) \cdot (1-\rho^u) \cdot \left(\frac{u(c^w) - u(b)}{u'(c^w)} - c^w \right) \\ & + p \cdot (1-\rho^c) \cdot \left(\frac{u(c^c) - u(b)}{u'(c^w)} - e^c \right) \\ & + (1-\rho) \cdot \left(\frac{\lambda - \eta f}{1-\eta} \cdot \frac{k_v}{q} \right) \Bigg] \\ & - \frac{\beta}{\lambda} \cdot (1-\rho) \cdot \frac{1-n}{n} \cdot b + \rho \cdot \beta \cdot (1-\beta) \cdot F \end{aligned}$$

The subsequent derivation of the wage equation is completely analogous to the derivation under STW regime described in Section 3.D.2. The wage equation thus reads:

$$
\begin{aligned}
(1-p)\cdot(1-\rho^u)\cdot(1-\eta)\cdot\left(\frac{u(c^w)-u(b)}{u'(c^w)}+\eta\cdot c^w\right) =& \\
&\eta\cdot[(1-\rho)\cdot z+(1-\rho)\cdot\theta\cdot k_v] \\
&-\frac{1-\beta\cdot(1-\lambda)}{\lambda}\cdot(1-\rho)\cdot\frac{1-n}{n}\cdot b \\
&+\frac{1-\beta\cdot(1-\lambda)}{\lambda}\cdot\rho\cdot\beta\cdot(1-\beta)\cdot F \\
&-p\cdot(1-\rho^c)\cdot(1-\eta)\cdot\left(\frac{u(c^c)-u(b)}{u'(c^w)}+\eta\cdot e^c\right)
\end{aligned}
$$

3.D Equilibrium STW

3.D.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:

$$
\begin{aligned}
y(\epsilon,h(\epsilon))-w^c(\epsilon) \;&\geq\; \beta\cdot\left[\lambda\cdot\bar{J}+(1-\lambda)\cdot J^c(\epsilon)\right] \\
y(\epsilon,h_{\text{stw}}(\epsilon))-w^c_{\text{stw}}(\epsilon) \;&\geq\; \beta\cdot\left[\lambda\cdot\bar{J}+(1-\lambda)\cdot J^c_{\text{stw}}(\epsilon)\right]
\end{aligned}
$$

(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:

$$
u(c^j_i(\epsilon))-u(b)+(\lambda-f)\cdot\beta\cdot(\bar{V}-U) \;\geq\; 0
$$

Let $\epsilon^{j,\text{cp}}_s$ denote the threshold at which the worker wants to separate from a firm with STW support:

$$
u(c^j_i(\epsilon^{j,\text{stw},\text{cp}}_s))-u(b)+(\lambda-f)\cdot\beta\cdot(\bar{V}-U) \;=\; 0
$$

Likewise, we can determine the threshold at which workers leave firms without STW support, $\xi_s^{j,\mathrm{cp}}$, as:

$$u(c_i^j(\xi_s^{j,\mathrm{cp}})) - u(b) + (\lambda - f)\cdot\beta\cdot(\bar{V} - U) \;=\; 0$$

We can rewrite the value functions of the firm and the surplus of the worker as:

$$\begin{aligned}
\bar{J} \;=\; -\tau + (1-p)\cdot\Bigg[\int_{\max\{\xi_s^{u,B},\,\xi_s^{u,\mathrm{cp}},\,\epsilon_{\mathrm{stw}}\}}^{\infty} \frac{1}{1-\beta\cdot(1-\lambda)}\Big(z(\epsilon) - c^u(\epsilon) + \lambda\beta\bar{J}\Big)\,dG(\epsilon)\\
+\int_{\max\{\epsilon_s^{u,B},\,\epsilon_s^{u,\mathrm{cp}}\}}^{\max\{\epsilon_s^{u,B},\epsilon_s^{u,cp},\epsilon_{\mathrm{stw}}\}} \frac{1}{1-\beta\cdot(1-\lambda)}\Big(z_{\mathrm{stw}}(\epsilon) - c^u_{\mathrm{stw}}(\epsilon) + \lambda\beta\bar{J}\Big)\,dG(\epsilon)\Bigg]\\
+\,p\cdot\Bigg[\int_{\max\{\xi_s^{c,B},\,\xi_s^{c,\mathrm{cp}},\,\epsilon_{\mathrm{stw}}\}}^{\infty} \frac{1}{1-\beta\cdot(1-\lambda)}\Big(z(\epsilon) - c^c(\epsilon) + \lambda\beta\bar{J}\Big)\,dG(\epsilon)\\
+\int_{\max\{\epsilon_s^{c,B},\,\epsilon_s^{c,\mathrm{cp}}\}}^{\max\{\epsilon_s^{c,B},\epsilon_s^{c,cp},\epsilon_{\mathrm{stw}}\}} \frac{1}{1-\beta\cdot(1-\lambda)}\Big(z_{\mathrm{stw}}(\epsilon) - c^c_{\mathrm{stw}}(\epsilon) + \lambda\beta\bar{J}\Big)\,dG(\epsilon)\Bigg]
\end{aligned}$$

$$\begin{aligned}
&(\bar{V} - U)\\
&= (1-p)\cdot\Bigg[\int_{\max\{\xi_s^{u,B},\,\xi_s^{u,\mathrm{cp}},\,\epsilon_{\mathrm{stw}}\}}^{\infty} \frac{1}{1-\beta\cdot(1-\lambda)}\Big(u(c^u(\epsilon)) - u(b) + \lambda\beta(\bar{V}-U)\Big)\,dG(\epsilon)\\
&+\int_{\max\{\epsilon_s^{u,B},\,\epsilon_s^{u,\mathrm{cp}}\}}^{\max\{\epsilon_s^{u,B},\epsilon_s^{u,cp},\epsilon_{\mathrm{stw}}\}} \frac{1}{1-\beta\cdot(1-\lambda)}\Big(u(c^u_{\mathrm{stw}}(\epsilon)) + \tau_{\mathrm{stw}}(\bar{h} - h_{\mathrm{stw}}(\epsilon)) - u(b) + \lambda\beta(\bar{V}-U)\Big)\,dG(\epsilon)\Bigg]\\
&+p\cdot\Bigg[\int_{\max\{\xi_s^{c,B},\,\xi_s^{c,\mathrm{cp}},\,\epsilon_{\mathrm{stw}}\}}^{\infty} \frac{1}{1-\beta\cdot(1-\lambda)}\Big(u(c^c(\epsilon)) - u(b) + \lambda\beta(\bar{V}-U)\Big)\,dG(\epsilon)\\
&+\int_{\max\{\epsilon_s^{c,B},\,\epsilon_s^{c,\mathrm{cp}}\}}^{\max\{\epsilon_s^{c,B},\epsilon_s^{c,cp},\epsilon_{\mathrm{stw}}\}} \frac{1}{1-\beta\cdot(1-\lambda)}\Big(u(c^c_{\mathrm{stw}}(\epsilon)) + \tau_{\mathrm{stw}}(\bar{h} - h_{\mathrm{stw}}(\epsilon)) - u(b) + \lambda\beta(\bar{V}-U)\Big)\,dG(\epsilon)\Bigg]
\end{aligned}$$

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.

$$\max_{\substack{c^u(\epsilon),\,c^c(\epsilon),\,c^u_{\mathrm{stw}}(\epsilon),\,c^c_{\mathrm{stw}}(\epsilon),\\ h^u(\epsilon),\,h^c(\epsilon),\,h^u_{\mathrm{stw}}(\epsilon),\,h^c_{\mathrm{stw}}(\epsilon),\\ \xi_s^{u,B},\,\xi_s^{c,B},\,\epsilon_s^{u,B},\,\epsilon_s^{c,B}}} \big(\bar{J}\big)^{1-\eta}\cdot\big(\bar{V}-U\big)^{\eta}$$

Financial constraints of financially constrained firms

$$
\begin{aligned}
y(\epsilon, h(\epsilon)) - w^c(\epsilon) \;&\geq\; \beta \cdot \left[\lambda \cdot \bar{J} + (1-\lambda) \cdot J^c(\epsilon)\right] \\
y(\epsilon, h_{\text{stw}}(\epsilon)) - w^c_{\text{stw}}(\epsilon) \;&\geq\; \beta \cdot \left[\lambda \cdot \bar{J} + (1-\lambda) \cdot J^c_{\text{stw}}(\epsilon)\right]
\end{aligned}
$$

Set up Kuhn-Tucker Conditions:

$$
\begin{aligned}
&\left(\bar{J}\right)^{\eta} \cdot \left(\bar{V} - U\right)^{1-\eta} - \int_0^{\infty} \mathscr{M}(\epsilon)\left[z(\epsilon) - c^c(\epsilon) + \beta \cdot \left[\lambda \cdot \bar{J} + (1-\lambda) \cdot J^c(\epsilon)\right]\right] dG(\epsilon) \\
&\quad - \int_0^{\infty} \mathscr{M}_{\text{stw}}(\epsilon)\left[z_{\text{stw}}(\epsilon) - c^c_{\text{stw}}(\epsilon) + \beta \cdot \left[\lambda \cdot \bar{J} + (1-\lambda) \cdot J^c_{\text{stw}}(\epsilon)\right]\right] dG(\epsilon)
\end{aligned}
$$

With Kuhn-Tucker conditions for a maximum:

(A)(i) $\mathscr{M}(\epsilon) \leq 0$

(ii) $\mathscr{M}(\epsilon)\left[z(\epsilon) - c^c(\epsilon) + \beta \cdot \left[\lambda \cdot \bar{J} + (1-\lambda) \cdot J^c(\epsilon)\right]\right] = 0$

(B)(i) $\mathscr{M}_{\text{stw}}(\epsilon) \leq 0$

(ii) $\mathscr{M}_{\text{stw}}(\epsilon)\left[z_{\text{stw}}(\epsilon) - c^c_{\text{stw}}(\epsilon) + \beta \cdot \left[\lambda \cdot \bar{J} + (1-\lambda) \cdot J^c_{\text{stw}}(\epsilon)\right]\right] = 0$

FOC for $c^u(\epsilon)$:

$$
\begin{aligned}
\frac{\partial \mathscr{B}}{\partial c^u(\epsilon)} \;&=\; -(1-\eta) \cdot (1-p) \cdot g(\epsilon)\left(\frac{\bar{V} - U}{\bar{J}}\right)^{\eta} \\
&\quad + \eta \cdot (1-p) \cdot g(\epsilon) \cdot u'(c^u(\epsilon))\left(\frac{\bar{J}}{\bar{V} - U}\right)^{1-\eta} \;=\; 0
\end{aligned}
$$

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 $c^u_{\text{stw}}(\epsilon)$:

$$
\begin{aligned}
\frac{\partial \mathscr{B}}{\partial c^u_{\text{stw}}(\epsilon)} \;&=\; -(1-\eta)(1-p) \cdot g(\epsilon)\left(\frac{\bar{V} - U}{\bar{J}}\right)^{\eta} \\
&\quad + \eta \cdot (1-p) \cdot g(\epsilon) \cdot u'\left(c^u_{\text{stw}}(\epsilon)\right) \cdot \eta \cdot \left(\frac{\bar{J}}{\bar{V} - U}\right)^{1-\eta} \;=\; 0
\end{aligned}
$$

Unconstrained firms want to insure workers against idiosyncratic productivity shocks, also on STW:

$$\begin{aligned} \iff \quad \frac{1-\eta}{\eta} \cdot \frac{\bar{V}-U}{\bar{J}} &= u'\left(c^u(\epsilon)\right) \\ \Rightarrow \quad u'\left(c^u_{\text{stw}}(\epsilon)\right) &= u'\left(c^u_{\text{stw}}(\epsilon')\right) = u'(c^w) \end{aligned}$$

FOC for $c^c(\epsilon)$:

$$\begin{aligned} \frac{\partial \mathscr{B}}{\partial c^c(\epsilon)} &= -(1-p) \cdot g(\epsilon) \cdot (1-\eta) \cdot \left(\frac{\bar{V}-U}{\bar{J}}\right)^{\eta} \\ &\quad + (1-p) \cdot g(\epsilon) \cdot u'(c^c(\epsilon)) \cdot \eta \cdot \left(\frac{\bar{J}}{\bar{V}-U}\right)^{1-\eta} + \mathscr{M}(\epsilon) \\ &\stackrel{!}{=} 0 \end{aligned}$$

Insert the surplus splitting rule:

$$\begin{aligned} &- (1-p) \cdot g(\epsilon) \cdot \left(\frac{\eta}{1-\eta} \cdot u'(c^w)\right)^{\eta} \\ &+ (1-p) \cdot g(\epsilon) \cdot u'(c^c(\epsilon)) \cdot \eta \cdot \left(\frac{1-\eta}{\eta} \cdot \frac{1}{u'(c^w)}\right) + \mathscr{M}(\epsilon) \stackrel{!}{=} 0 \end{aligned}$$

$$\iff \quad (1-p) \cdot g(\epsilon) \cdot \eta \cdot \left(u'(c^c(\epsilon)) - u'(c^w)\right) = -\mathscr{M}(\epsilon) \cdot \left(u'(c^w) \cdot \frac{\eta}{1-\eta}\right)^{1-\eta}$$

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

$$z(\epsilon) - c^w + \beta \cdot \left[\lambda \cdot \bar{J} + (1-\lambda) \cdot J^c(\epsilon)\right] \geq 0,$$

then the constraint is non-binding and $\mathscr{M}(\epsilon) = 0$. ✓

If

$$z(\epsilon) - c^w + \beta \cdot \left[\lambda \cdot \bar{J} + (1-\lambda) \cdot J^c(\epsilon)\right] < 0,$$

then the constraint is binding and $\mathscr{M}(\epsilon) < 0$. ✓

Thus, it is optimal for the firm to pay as much as it can:

$$J^c(\epsilon) = z(\epsilon) - c^c(\epsilon) + \beta \cdot \left[\lambda \cdot \bar{J} + (1-\lambda) \cdot J^c(\epsilon)\right] = 0$$

$$\iff \quad c^c(\epsilon) = z(\epsilon) + \lambda \cdot \beta \cdot \bar{J}$$

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

$$c_{\text{stw}}(\epsilon) = \epsilon_{\text{stw}}(\epsilon) + \tau_{\text{stw}} \cdot (\bar{h} - h_{\text{stw}}(\epsilon)) + \lambda \cdot \beta \cdot \bar{J} \quad \text{when it is binding.}$$

Using the surplus splitting rule, we can simplify the Nash-Bargaining problem to:

$$\max_{h^u(\epsilon), h^c(\epsilon), h^u_{\text{stw}}(\epsilon), h^c_{\text{stw}}(\epsilon),\, \xi^u_s, \xi^c_s, \xi^{u^c}_s, \xi^{c,c}_s} (1-\eta)^{1-\eta} \cdot (u'(c^w) \cdot \eta)^{\eta} \cdot \bar{S}$$

That is, we just have to maximize the joint surplus:

$$\begin{aligned}
\bar{S} = &-\tau \\
&+ (1-p) \cdot \int_{\max\{\xi^{u,B}_s, \xi^{u,cp}_s, \epsilon_{\text{stw}}\}}^{\infty} \frac{1}{1-\beta \cdot (1-\lambda)} \Big(z(\epsilon) + \frac{u(c^w) - u(b)}{u'(c^w)} - c^w + (1 - \eta \cdot f) \cdot \beta \cdot \bar{S} \Big) dG(\epsilon) \\
&+ (1-p) \cdot \int_{\max\{\epsilon^{u,B}_s, \epsilon^{u,cp}_s\}}^{\max\{\epsilon^{u,B}_s, \epsilon^{u,cp}_s, \epsilon_{\text{stw}}\}} \frac{1}{1-\beta \cdot (1-\lambda)} \Big(z(\epsilon) + \frac{u(c^w) - u(b)}{u'(c^w)} - c^w \\
&\qquad + \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon)) + (1 - \eta \cdot f) \cdot \beta \cdot \bar{S} \Big) dG(\epsilon) \\
&+ p \cdot \int_{\epsilon^p}^{\infty} \frac{1}{1-\beta \cdot (1-\lambda)} \Big(z(\epsilon) + \frac{u(c^w) - u(b)}{u'(c^w)} - c^w + (1 - \eta \cdot f) \cdot \beta \cdot \bar{S} \Big) dG(\epsilon) \\
&+ p \cdot \int_{\max\{\xi^c_s, \xi^{c,cp}_s, \epsilon_{\text{stw}}\}}^{\epsilon^p} \frac{1}{1-\beta \cdot (1-\lambda)} \Big(z(\epsilon) + \frac{u(c^c(\epsilon)) - u(b)}{u'(c^w)} - c^c(\epsilon) + (1 - \eta \cdot f) \cdot \beta \cdot \bar{S} \Big) dG(\epsilon) \\
&+ p \cdot \int_{\max\{\epsilon^c_s, \epsilon^{c,cp}_s\}}^{\max\{\epsilon^c_s, \epsilon^{c,cp}_s, \epsilon_{\text{stw}}\}} \frac{1}{1-\beta \cdot (1-\lambda)} \Big(z_{\text{stw}}(\epsilon) + \frac{u(c^c_{\text{stw}}(\epsilon)) - u(b)}{u'(c^w)} \\
&\qquad - c^c(\epsilon) + (1 - \eta \cdot f) \cdot \beta \cdot \bar{S} \Big) dG(\epsilon)
\end{aligned}$$

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:

$$\begin{aligned}
\frac{\partial \mathscr{B}}{\partial h^u(\epsilon)} &= (1-p) \cdot g(\epsilon) \cdot \left(\frac{\partial y(\epsilon, h^u(\epsilon))}{\partial h^u(\epsilon)} - \phi'(h^u(\epsilon)) \right) = 0 \\
\Leftrightarrow \quad \frac{\partial y(\epsilon, h^u(\epsilon))}{\partial h^u(\epsilon)} &= \phi'(h^u(\epsilon))
\end{aligned}$$

The FOC for working hours in constrained firms *outside* STW is:

$$\frac{\partial \mathcal{B}}{\partial h^c(\epsilon)} = p \cdot g(\epsilon) \cdot \left(\frac{\partial y(\epsilon, h^c(\epsilon))}{\partial h^c(\epsilon)} - \phi'(h^c(\epsilon)) \right) \cdot \frac{u'(c^c(\epsilon))}{u'(c^w)} = 0$$
$$\Leftrightarrow \quad \frac{\partial y(\epsilon, h^c(\epsilon))}{\partial h^c(\epsilon)} = \phi'(h^c(\epsilon))$$

The FOC for working hours in unconstrained firms *on* STW is:

$$\frac{\partial \mathcal{B}}{\partial h^u_{\text{stw}}(\epsilon)} = (1-p) \cdot g(\epsilon) \cdot \left(\frac{\partial y(\epsilon, h^u_{\text{stw}}(\epsilon))}{\partial h^u_{\text{stw}}(\epsilon)} - \phi'(h^u_{\text{stw}}(\epsilon)) - \tau_{\text{stw}} \right) = 0$$
$$\Leftrightarrow \quad \frac{\partial y(\epsilon, h^u_{\text{stw}}(\epsilon))}{\partial h^u_{\text{stw}}(\epsilon)} = \phi'(h^u_{\text{stw}}(\epsilon)) + \tau_{\text{stw}}$$

The FOC for working hours in constrained firms *on* STW is:

$$\frac{\partial \mathcal{B}}{\partial h^c(\epsilon)} = p \cdot g(\epsilon) \cdot \left(\frac{\partial y(\epsilon, h^c(\epsilon))}{\partial h^c(\epsilon)} - \phi'(h^c(\epsilon)) - \tau_{\text{stw}} \right) \cdot \frac{u'(c^c(\epsilon))}{u'(c^w)} = 0$$
$$\Leftrightarrow \quad \frac{\partial y(\epsilon, h^c(\epsilon))}{\partial h^c(\epsilon)} = \phi'(h^c(\epsilon)) + \tau_{\text{stw}}$$

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:

$$\frac{\partial \mathcal{B}}{\partial \xi_s^{u,B}} = -(1-p) \cdot g(\xi_s^{u,B}) \cdot \frac{1}{1-\beta \cdot (1-\lambda)} \left(z(\xi_s^{u,B}) + \frac{u(c^w) - u(b)}{u'(c^w)} - c^w + (\lambda - \eta f) \cdot \beta \cdot \bar{S} \right) = 0$$

$$\Longleftrightarrow \quad z(\xi_s^{u,B}) + \frac{u(c^w) - u(b)}{u'(c^w)} - c^w + (\lambda - \eta f) \cdot \beta \cdot \bar{S} = 0$$

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:

$$V(\epsilon) - U = u(c^w) - u(b) + (\lambda - f) \cdot \beta \cdot (\bar{V} - U) > 0$$

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:

$$\frac{\partial \mathcal{B}}{\partial \xi_s^{c,B}} = -(1-p)\cdot g(\xi_s^{c,B})\cdot\Bigg[\frac{1}{1-\beta\cdot(1-\lambda)}\cdot\Big(z(\xi_s^{c,B})+\frac{u\left(c^c(\xi_s^{c,B})\right)-u(b)}{u'(c^w)}$$
$$-c^c(\xi_s^{c,B})+(\lambda-\eta f)\cdot\beta\cdot\bar{S}\Big)\Bigg] = 0$$

$$\iff \quad z(\xi_s^{c,B})+\frac{u(c^c(\xi_s^{c,B}))-u(b)}{u'(c^w)}-c^c(\xi_s^{c,B})+(\lambda-\eta f)\cdot\beta\cdot\bar{S} = 0$$

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:

$$\frac{\partial \mathcal{B}}{\partial \epsilon_s^{u,B}} = -(1-p)\cdot g(\epsilon_s^{u,B})\cdot\Bigg[\frac{1}{1-\beta\cdot(1-\lambda)}\cdot\Big(z_{\text{stw}}(\epsilon_s^{u,B})+\frac{u(c^w)-u(b)}{u'(c^w)}-c^w$$
$$+\tau_{\text{stw}}\cdot(\bar{h}-h_{\text{stw}}(\epsilon_s^{u,B}))+(\lambda-\eta f)\cdot\beta\cdot\bar{S}\Big)\Bigg] = 0$$

$$\iff \quad z_{\text{stw}}(\epsilon_s^{u,B})+\frac{u(c^w)-u(b)}{u'(c^w)}-c^w+\tau_{\text{stw}}(\bar{h}-h_{\text{stw}}(\epsilon_s^{u,B}))+(\lambda-\eta f)\cdot\beta\cdot\bar{S} = 0$$

Insert free-entry Condition:

$$z_{\text{stw}}(\epsilon_s^{u,B})+\frac{u(c^w)-u(b)}{u'(c^w)}-c^w+\frac{\lambda-\eta f}{1-\eta}\cdot\frac{k_v}{q} = 0$$

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:

$$V(\epsilon)-U = u(c^w)-u(b)+(\lambda-f)\cdot\beta\cdot(\bar{V}-U) > 0$$

It always gets positive surplus!
Suppose that

$$\max\{\epsilon_s^{c,B}, \epsilon_s^{c,\mathrm{cp}}, \epsilon_{\mathrm{stw}}\} = \epsilon_s^{c,B}$$

Then the FOC for the separation threshold of the constrained firm with STW support is:

$$\begin{aligned}\frac{\partial \mathscr{B}}{\partial \epsilon_s^{c,B}} &= -(1-p)\cdot g(\epsilon_s^{c,B})\cdot\Bigg[\frac{1}{1-\beta\cdot(1-\lambda)}\cdot\Bigg(z_{\mathrm{stw}}(\epsilon_s^{c,B}) + \frac{u\left(c^c_{\mathrm{stw}}(\epsilon_s^{c,B})\right)-u(b)}{u'(c^w)}\\ &\quad - c^c_{\mathrm{stw}}(\epsilon_s^{c,B}) + (\lambda-\eta f)\cdot\beta\cdot\bar{S}\Bigg)\Bigg] = 0\end{aligned}$$

$$\iff \quad z(\epsilon_s^{c,B}) + \frac{u(c^c_{\mathrm{stw}}(\epsilon_s^{c,B}))-u(b)}{u'(c^w)} - c^c_{\mathrm{stw}}(\epsilon_s^{c,B}) + (\lambda-\eta f)\cdot\beta\cdot\bar{S} = 0$$

Insert free-entry condition:

$$z_{\mathrm{stw}}(\epsilon_s^{c,B}) + \frac{u(c^c_{\mathrm{stw}}(\epsilon_s^{c,B}))-u(b)}{u'(c^w)} - c^c_{\mathrm{stw}}(\epsilon_s^{c,B}) + \frac{\lambda-\eta f}{1-\eta}\cdot\frac{k_v}{q} = 0$$

The participation constraint of workers can be written with free-entry condition as:

$$z_{\mathrm{stw}}(\epsilon_s^{c,\mathrm{cp}}) + \frac{u(c^c_{\mathrm{stw}}(\epsilon_s^{c,\mathrm{cp}}))-u(b)}{u'(c^w)} - c^c_{\mathrm{stw}}(\epsilon_s^{c,\mathrm{cp}}) + \frac{\lambda-\eta f}{1-\eta}\cdot\frac{k_v}{q} = 0$$

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.D.2 Job Creation and Wage Equation

This section derives the job creation and the equation for STW. It follows its notation part 3.E of the appendix.

(I) Value of an unconstrained firm outside STW after the idiosyncratic productivity shock has realized:

$$\begin{aligned} J^u(\epsilon) &= z(\epsilon) - c^w + \beta\cdot\left[\lambda\bar{J} + (1-\lambda)J^u(\epsilon)\right]\\ \Leftrightarrow \quad J^u(\epsilon) &= \frac{1}{1-\beta\cdot(1-\lambda)}\left(z(\epsilon) - c^w + \lambda\beta\bar{J}\right)\end{aligned}$$

(II) Value of an unconstrained firm on STW after the idiosyncratic productivity shock has realized:

$$\begin{aligned} J^u_{\mathrm{stw}}(\epsilon) &= z_{\mathrm{stw}}(\epsilon) + \tau_{\mathrm{stw}}(\bar{h} - h_{\mathrm{stw}}(\epsilon)) - c^w + \beta\left[\lambda\bar{J} + (1-\lambda)J^u_{\mathrm{stw}}(\epsilon)\right]\\ \Leftrightarrow \quad J^u_{\mathrm{stw}}(\epsilon) &= \frac{1}{1-\beta\cdot(1-\lambda)}\left(z_{\mathrm{stw}}(\epsilon) - c^w + \lambda\beta\bar{J}\right)\end{aligned}$$

(III) Value of a constrained firm outside STW with non-binding constraints after the idiosyncratic productivity shock has realized:

$$\begin{aligned} J^c(\epsilon) &= z(\epsilon) - c^w + \beta \cdot \left[\lambda \bar{J} + (1-\lambda) J^c(\epsilon)\right] \\ \Leftrightarrow \quad J^c(\epsilon) &= \frac{1}{1-\beta \cdot (1-\lambda)} \left(z(\epsilon) - c^w + \lambda\beta\bar{J}\right) \end{aligned}$$

(IV) Value of a constrained firm outside STW with binding constraints after the idiosyncratic productivity shock has realized:

$$\begin{aligned} J^c(\epsilon) &= z(\epsilon) - c^w(\epsilon) + \beta \left[\lambda \bar{J} + (1-\lambda) J^c(\epsilon)\right] \\ \Leftrightarrow \quad J^c(\epsilon) &= \frac{1}{1-\beta \cdot (1-\lambda)} \left(z(\epsilon) - c^w(\epsilon) + \lambda\beta\bar{J}\right) \end{aligned}$$

(V) Value of a constrained firm on STW after the idiosyncratic productivity shock has realized:

$$\begin{aligned} J^c_{\text{stw}}(\epsilon) &= z_{\text{stw}}(\epsilon) - c^w_{\text{stw}}(\epsilon) + \beta \cdot \left[\lambda J + (1-\lambda) J^c_{\text{stw}}(\epsilon)\right] \\ \Leftrightarrow \quad J^c_{\text{stw}}(\epsilon) &= \frac{1}{1-\beta \cdot (1-\lambda)} \left(z_{\text{stw}}(\epsilon) - c^w_{\text{stw}}(\epsilon) + \lambda\beta J\right) \end{aligned}$$

(VI) Value of a firm before the idiosyncratic productivity shock has realized:

$$\begin{aligned} \bar{J} = {} & -\tau + (1-p) \cdot \left(\int_{\epsilon_{\text{stw}}}^{\infty} J^u(\epsilon)\, dG(\epsilon) + \int_{\epsilon_s^u}^{\epsilon_{\text{stw}}} J^u_{\text{stw}}(\epsilon)\, dG(\epsilon) \right) \\ & + p \cdot \left(\int_{\max\{\epsilon_{\text{stw}}, \xi_s^c\}}^{\infty} J^c(\epsilon)\, dG(\epsilon) + \int_{\epsilon_s^c}^{\max\{\epsilon_{\text{stw}}, \epsilon_s^c\}} J^c_{\text{stw}}(\epsilon)\, dG(\epsilon) \right) \end{aligned}$$

Inserting (I) - (IV) into (V) gives:

$$\begin{aligned} \bar{J} = {} & -\tau \\ & + \frac{1}{1-\beta \cdot (1-\lambda)} \Bigg[(1-\rho) \cdot (z - \Omega) \\ & - (1-p) \cdot (1-\rho^u) \cdot \left(c^w - \int_{\epsilon_s^u}^{\epsilon_{\text{stw}}} \tau_{\text{stw}} (\bar{h} - h_{\text{stw}}(\epsilon))\, dG(\epsilon) \right) \\ & - p \cdot (1-\rho^c) \cdot \left(e^c - \int_{\epsilon_s^c}^{\max\{\epsilon_{\text{stw}}, \epsilon_s^c\}} \tau_{\text{stw}} (\bar{h} - h_{\text{stw}}(\epsilon))\, dG(\epsilon) \right) \\ & + (1-\rho) \cdot \lambda \cdot \beta \cdot \bar{J} \Bigg] \end{aligned}$$

Here z denotes the average production (without distortions) and ec denotes the average cost of a constrained worker for a firm. Both are defined in section D of the appendix.

(VII) Value of an unconstrained worker outside STW after the idiosyncratic productivity shock has realized:

$$\begin{aligned} V^u(\epsilon) &= u(c^w) + \beta \cdot \left[\lambda \bar{V} + (1-\lambda)\, V^u(\epsilon)\right] \\ \Leftrightarrow \quad V^u(\epsilon) &= \frac{1}{1-\beta\cdot(1-\lambda)}\left(u(c^w) + \lambda\beta\bar{V}\right) \end{aligned}$$

(VIII) Value of an unconstrained worker on STW after the idiosyncratic productivity shock has realized:

$$\begin{aligned} V^u_{\text{stw}}(\epsilon) &= u(c^w) + \beta \cdot \left[\lambda \bar{V} + (1-\lambda)\, V^u_{\text{stw}}(\epsilon)\right] \\ \Leftrightarrow \quad V^u_{\text{stw}}(\epsilon) &= \frac{1}{1-\beta\cdot(1-\lambda)}\left(u(c^w) + \lambda\beta\bar{V}\right) \end{aligned}$$

(IX) Value of a constrained worker outside STW with non-binding constraints after the idiosyncratic productivity shock has realized:

$$\begin{aligned} V^c(\epsilon) &= u(c^w) + \beta \cdot \left[\lambda \bar{V} + (1-\lambda)\, V^c(\epsilon)\right] \\ \Leftrightarrow \quad V^c(\epsilon) &= \frac{1}{1-\beta\cdot(1-\lambda)}\left(u(c^w) + \lambda\beta\bar{V}\right) \end{aligned}$$

(X) Value of a constrained worker outside STW with binding constraints after the idiosyncratic productivity shock has realized:

$$\begin{aligned} V^c(\epsilon) &= u(c^w(\epsilon)) + \beta \cdot \left[\lambda \bar{V} + (1-\lambda)\, V^c(\epsilon)\right] \\ \Leftrightarrow \quad V^c(\epsilon) &= \frac{1}{1-\beta\cdot(1-\lambda)}\left(u(c^w(\epsilon)) + \lambda\beta\bar{V}\right) \end{aligned}$$

(XI) Value of a constrained worker on STW after the idiosyncratic productivity shock has realized:

$$\begin{aligned} V^c_{\text{stw}}(\epsilon) &= u(c_{\text{stw}}(\epsilon)) + \beta \cdot \left[\lambda \bar{V} + (1-\lambda)\, V^c_{\text{stw}}(\epsilon)\right] \\ \Leftrightarrow \quad V^c_{\text{stw}}(\epsilon) &= \frac{1}{1-\beta\cdot(1-\lambda)}\left(u(c_{\text{stw}}(\epsilon)) + \lambda\beta\bar{V}\right) \end{aligned}$$

(XII) Value of a worker before the idiosyncratic productivity shock has realized:

$$\begin{aligned} \bar{V} = (1-p)\cdot\Bigg(&\int_{\epsilon_{\text{stw}}}^{\infty} V^u(\epsilon)\, dG(\epsilon) \\ &+ \int_{\epsilon_s^u}^{\epsilon_{\text{stw}}} V^u_{\text{stw}}(\epsilon)\, dG(\epsilon)\Bigg) \\ &+ p\cdot\Bigg(\int_{\epsilon_p}^{\infty} V^c(\epsilon)\, dG(\epsilon) \\ &+ \int_{\max\{\xi_s^c,\epsilon_{\text{stw}}\}}^{\epsilon_p} V^c(\epsilon)\, dG(\epsilon) \\ &+ \int_{\epsilon_s^c}^{\max\{\epsilon_s^c,\epsilon_{\text{stw}}\}} V^c_{\text{stw}}(\epsilon)\, dG(\epsilon)\Bigg) \\ &+ \rho\cdot U \end{aligned}$$

Inserting (VII) - (XI) into (XII) gives:

$$\begin{aligned}\bar{V} \;=\; & \frac{1}{1-\beta\cdot(1-\lambda)}\Big[(1-p)(1-\rho^u)\cdot u(c^w) + p\cdot(1-\rho^c)\cdot u^c \\ & + (1-\rho)\cdot\lambda\cdot\beta\cdot\bar{V} + \rho\cdot U\Big]\end{aligned}$$

Here u^c denotes the average utility of a constrained worker for a firm. Both are defined in section 3.E of the appendix.

(XIII) Unemployment:

$$U \;=\; u(b) + \beta\cdot\big[f\cdot\bar{V} + (1-f)\cdot U\big]$$

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:

$$\begin{aligned} & V_i^j(\epsilon) - U \;=\; u(c_i^j(\epsilon)) - u(b) + (\lambda - f)\cdot V_i^j(\epsilon) + (\lambda - f)\cdot\beta\cdot\bar{V} - (\lambda - f)\cdot\beta\cdot U \\ \Leftrightarrow\quad & V_i^j(\epsilon) - U \;=\; u(c_i^j(\epsilon)) - u(b) + (\lambda - f)\cdot\beta\cdot(V_i^j(\epsilon) - U) + (\lambda - f)\cdot\beta\cdot(\bar{V} - U) \\ \Leftrightarrow\quad & V_i^j(\epsilon) - U \;=\; u(c_i^j(\epsilon)) - u(b) + (\lambda - f)\cdot(V_i^j(\epsilon) - U) + (\lambda - f)\cdot\beta\cdot(\bar{V} - U) \\ \Leftrightarrow\quad & V_i^j(\epsilon) - U \;=\; \frac{1}{1-\beta\cdot(1-\lambda)}\Big(u(c_i^j(\epsilon)) - u(b) + (\lambda - f)\cdot\beta\cdot(\bar{V} - U)\Big)\end{aligned}$$

for $i \in \{\text{stw}, \text{no stw}\}$ and $j \in \{u, c\}$.

(B) The surplus of the worker before shock realizations are known:

$$\begin{aligned}\bar{V} - U \;=\; & (1-p)\cdot\left(\int_{\epsilon_{\text{stw}}}^{\infty} V^u(\epsilon) - U\, dG(\epsilon) + \int_{\epsilon_s^u}^{\epsilon_{\text{stw}}} V_{\text{stw}}^u(\epsilon) - U\, dG(\epsilon)\right) \\ & + p\cdot\left(\int_{\epsilon_p}^{\infty} V(\epsilon) - U\, dG(\epsilon) + \int_{\max\{\xi_s^c, \epsilon_{\text{stw}}\}}^{\epsilon_p} V^c(\epsilon) - U\, dG(\epsilon)\right. \\ & \left. + \int_{\epsilon_s^c}^{\max\{\epsilon_s^c, \epsilon_{\text{stw}}\}} V_{\text{stw}}^c(\epsilon) - U\, dG(\epsilon)\right)\end{aligned}$$

Inserting A into B gives:

$$\begin{aligned}\bar{V} - U \;=\; & \frac{1}{1-\beta\cdot(1-\lambda)}\Big[(1-p)(1-\rho^u)\cdot(u(c^w) - u(b)) \\ & + p\cdot(1-\rho^c)\cdot(u^c - u(b)) \\ & + (\lambda - f)\cdot\beta\cdot(\bar{V} - U)\Big]\end{aligned}$$

(C) Finally, we can calculate the joint surplus before shock realizations are known:

$$\begin{aligned}
S &= J + \frac{\bar{V} - U}{u'(c^w)} \\
&= -\tau + \frac{1}{1 - \beta \cdot (1-\lambda)} \Bigg[(1-\rho)(1-\rho^u) \cdot \Bigg(\frac{u(c^w) - u(b)}{u'(c^w)} - c^w \\
&\quad + \frac{1}{1-\rho^u} \int_{\epsilon_s^u}^{\epsilon_{\text{stw}}} \tau_{\text{stw}} \cdot (\bar{h} - h_{\text{stw}}(\epsilon))\, dG(\epsilon) \Bigg) \\
&\quad + p(1-\rho^c) \cdot \Bigg(\frac{u^c - u(b)}{u'(c^w)} - e^c \\
&\quad + \frac{1}{1-\rho^c} \int_{\epsilon_s^c}^{\max\{\epsilon_{\text{stw}}, \epsilon_s^c\}} \tau_{\text{stw}} (\bar{h} - h_{\text{stw}}(\epsilon))\, dG(\epsilon) \Bigg) \\
&\quad + (1-\rho)\lambda\beta J + (1-\rho)(\lambda - f) \cdot \beta \cdot \frac{\bar{V} - U}{u'(c^w)} \Bigg]
\end{aligned}$$

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:

$$
\begin{aligned}
S &= J + \frac{\bar{V} - U}{u'(c^w)} = \eta S + (1-\eta) S \\
&= -\tau + \frac{1}{1-\beta \cdot (1-\lambda)} \Bigg[(1-\rho)\left(z - \Omega - \frac{1-n}{n} b\right) \\
&\quad + (1-p)(1-\rho^u)\left(\frac{u(c^w) - u(b)}{u'(c^w)} - c^w\right) \\
&\quad + (1-p)\int_{\epsilon_s^u}^{\epsilon_{\text{stw}}} \tau_{\text{stw}} \cdot (\bar{h} - h_{\text{stw}}(\epsilon))\, dG(\epsilon) \\
&\quad + p(1-\rho^c)\left(\frac{u^c - u(b)}{u'(c^w)} - e^c\right) \\
&\quad + p\int_{\epsilon_s^c}^{\max\{\epsilon_{\text{stw}}, \epsilon_s^c\}} \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon))\, dG(\epsilon) \Bigg] \\
\Leftrightarrow \quad S &= -\tau + \frac{1}{1-\beta \cdot (1-\lambda)} \Bigg[(1-\rho)\left(z - \Omega - \frac{1-n}{n} b\right) \\
&\quad + (1-p)(1-\rho^u)\left(\frac{u(c^w) - u(b)}{u'(c^w)} - c^w\right) \\
&\quad + (1-p)\int_{\epsilon_s^u}^{\epsilon_{\text{stw}}} (\bar{h} - h_{\text{stw}}(\epsilon))\, dG(\epsilon) \\
&\quad + p(1-\rho^c)\left(\frac{u^c - u(b)}{u'(c^w)} - e^c\right) \\
&\quad + p\int_{\epsilon_s^c}^{\max\{\epsilon_{\text{stw}}, \epsilon_s^c\}} \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon))\, dG(\epsilon) \\
&\quad + (1-\rho)(\lambda - \eta \cdot f) \cdot \beta \cdot S \Bigg]
\end{aligned}
$$

Next, we want to replace the tax. To do this, we need to know the budget constraint of the government:

$$
\begin{aligned}
n^s \cdot \tau &= \frac{n}{1-\rho} \cdot \Bigg((1-p)\int_{\epsilon_s^u}^{\epsilon_{\text{stw}}} \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon))\, dG(\epsilon) \\
&\quad + p\int_{\epsilon_s^c}^{\max\{\epsilon_{\text{stw}}, \epsilon_s^c\}} \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon))\, dG(\epsilon) \Bigg) \\
&\quad + (1-n) \cdot b
\end{aligned}
$$

Inserting the employment equation

$$
n = (1-\rho) \cdot \frac{n^s}{\lambda}
$$

gives:

$$\begin{aligned}\tau = &\frac{1}{\lambda}\cdot\left((1-p)\int_{\epsilon_s^u}^{\infty}\tau_{\text{stw}}(\bar{h}-h_{\text{stw}}(\epsilon))\,dG(\epsilon)\right.\\&\left.+p\int_{\epsilon_s^c}^{\max\{\epsilon_{\text{stw}},\epsilon_s^c\}}\tau_{\text{stw}}(\bar{h}-h_{\text{stw}}(\epsilon))\,dG(\epsilon)\right)\\&+\frac{(1-\rho)}{n^s}\cdot b\end{aligned}$$

Inserting the expression for the number of firms that receive a shock

$$n^s = \frac{n}{1-\rho}\cdot\lambda$$

gives:

$$\begin{aligned}\tau = &\frac{1}{\lambda}\cdot\left((1-p)\int_{\epsilon_s^u}^{\infty}\tau_{\text{stw}}(\bar{h}-h_{\text{stw}}(\epsilon))\,dG(\epsilon)\right.\\&\left.+p\int_{\epsilon_s^c}^{\max\{\epsilon_{\text{stw}},\epsilon_s^c\}}\tau_{\text{stw}}(\bar{h}-h_{\text{stw}}(\epsilon))\,dG(\epsilon)\right)\\&+\frac{1}{\lambda}\cdot(1-\rho)\cdot\left(\frac{1-n}{n}\cdot b\right)\end{aligned}$$

Inserting the tax into the surplus equation gives:

$$\begin{aligned}S = &\frac{1}{1-\beta\cdot(1-\lambda)}\Bigg[(1-\rho)\,(z-\Omega)\\&+(1-p)(1-\rho^u)\left(\frac{u(c^w)-u(b)}{u'(c^w)}-c^w\right)\\&+p(1-\rho^c)\left(\frac{u^c-u(b)}{u'(c^w)}-e^c\right)\\&+(1-\rho)(\lambda-f)\cdot\beta\cdot S\Bigg]\\&-\left(\frac{1}{\lambda}-\frac{1}{1-\beta\cdot(1-\lambda)}\right)\cdot(1-p)\int_{\epsilon_s^u}^{\epsilon_{\text{stw}}}(\bar{h}-h_{\text{stw}}(\epsilon))\,dG(\epsilon)\\&-\left(\frac{1}{\lambda}-\frac{1}{1-\beta\cdot(1-\lambda)}\right)\cdot p\int_{\epsilon_s^c}^{\max\{\epsilon_{\text{stw}},\epsilon_s^c\}}\tau_{\text{stw}}(\bar{h}-h_{\text{stw}}(\epsilon))\,dG(\epsilon)\\&-\frac{1}{\lambda}\cdot(1-\rho)\cdot\frac{1-n}{n}\cdot b\end{aligned}$$

Inserting the free-entry condition

$$\frac{k_v}{q} = \beta\cdot J$$

gives:

$$\frac{1}{1-\eta}\cdot\frac{k_v}{q} = \frac{\beta}{1-\beta\cdot(1-\lambda)}\Bigg[(1-\rho)\left(z-\Omega-\frac{1-n}{n}\cdot b\right) + (1-p)(1-\rho^u)\left(\frac{u(c^w)-u(b)}{u'(c^w)}-c^w\right) + p(1-\rho^c)\left(\frac{u^c-u(b)}{u'(c^w)}-e^c\right) + (1-\rho)\cdot\frac{\lambda-\eta\cdot f}{1-\eta}\cdot\frac{k_v}{q}\Bigg]$$

This is the job-creation condition used in the Ramsey problem for STW in Appendix 3.E.

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:

$$\bar{J} = \frac{1}{1-\beta\cdot(1-\lambda)}\Big[(1-\rho)(z-\Omega)-p(1-\rho^u)e^w-(1-p)(1-\rho^c)e+(1-\rho)\lambda\beta\bar{J}\Big] - \left(\frac{1}{\lambda}-\frac{1}{1-\beta\cdot(1-\lambda)}\right)\cdot(1-p)\int_{\epsilon_s^u}^{\epsilon_{\text{stw}}}\tau_{\text{stw}}(\bar{h}-h_{\text{stw}}(\epsilon))\,dG(\epsilon) - \left(\frac{1}{\lambda}-\frac{1}{1-\beta\cdot(1-\lambda)}\right)\cdot p\int_{\epsilon_s^c}^{\max\{\epsilon_{\text{stw}},\epsilon_s^c\}}\tau_{\text{stw}}(\bar{h}-h_{\text{stw}}(\epsilon))\,dG(\epsilon) - \frac{1}{\lambda}\cdot(1-\rho)\cdot\frac{1-n}{n}\cdot b$$

Likewise, we can calculate the surplus of a worker:

$$\bar{V}-U = \frac{1}{1-\beta\cdot(1-\lambda)}\Big[(1-p)(1-\rho^u)\big(u(c^w)-u(b)\big)+p(1-\rho^c)\big(u^c-u(b)\big) + (1-\rho)(\lambda-f)\beta(\bar{V}-U)\Big]$$

Next, we can insert the value of a worker for a firm and the surplus of a worker into the surplus splitting rule:

$$
\begin{aligned}
\frac{\eta}{1-\beta\cdot(1-\lambda)}\cdot\Big[&(1-\rho)(z-\Omega)-p(1-\rho^{u})\cdot c^{w}-(1-p)(1-\rho^{c})\cdot e^{c}\\
&+(1-\rho)\cdot\lambda\cdot\beta\cdot\bar{J}\Big]\\
&-\left(\frac{1}{\lambda}-\frac{1}{1-\beta\cdot(1-\lambda)}\right)\cdot(1-p)\int_{\epsilon_s^u}^{\epsilon_{\mathrm{stw}}}\tau_{\mathrm{stw}}(\bar{h}-h_{\mathrm{stw}}(\epsilon))\,dG(\epsilon)\\
&-\left(\frac{1}{\lambda}-\frac{1}{1-\beta\cdot(1-\lambda)}\right)\cdot p\int_{\epsilon_s^c}^{\max\{\epsilon_{\mathrm{stw}},\epsilon_s^c\}}\tau_{\mathrm{stw}}(\bar{h}-h_{\mathrm{stw}}(\epsilon))\,dG(\epsilon)\\
&-\frac{1}{\lambda}\cdot(1-\rho)\cdot\frac{1-n}{n}\cdot b\\
=&\ \frac{1-\eta}{1-\beta\cdot(1-\lambda)}\cdot\frac{1}{u'(c^{w})}\cdot\Big[(1-p)(1-\rho^{u})\cdot(u(c^{w})-u(b))\\
&+p(1-\rho^{c})\cdot(u^{c}-u(b))\\
&+(1-\rho)(\lambda-f)\beta(\bar{V}-U)\Big]
\end{aligned}
$$

Insert the surplus splitting rule again: $\eta\cdot\bar{J}=(1-\eta)\cdot\frac{\bar{V}-U}{u'(c^{w})}$

$$
\begin{aligned}
\frac{\eta}{1-\beta\cdot(1-\lambda)}\cdot\Big[&(1-\rho)(z-\Omega)-p(1-\rho^{u})\cdot c^{w}-(1-p)(1-\rho^{c})\cdot e^{c}\\
&+(1-\rho)\cdot\lambda\cdot\beta\cdot\bar{J}\Big]\\
&-\left(\frac{1}{\lambda}-\frac{1}{1-\beta\cdot(1-\lambda)}\right)\cdot(1-p)\int_{\epsilon_s^u}^{\epsilon_{\mathrm{stw}}}\tau_{\mathrm{stw}}(\bar{h}-h_{\mathrm{stw}}(\epsilon))\,dG(\epsilon)\\
&-\left(\frac{1}{\lambda}-\frac{1}{1-\beta\cdot(1-\lambda)}\right)\cdot p\int_{\epsilon_s^c}^{\max\{\epsilon_{\mathrm{stw}},\epsilon_s^c\}}\tau_{\mathrm{stw}}(\bar{h}-h_{\mathrm{stw}}(\epsilon))\,dG(\epsilon)\\
&-\frac{1}{\lambda}\cdot(1-\rho)\cdot\frac{1-n}{n}\cdot b\\
=&\ \frac{1-\eta}{1-\beta\cdot(1-\lambda)}\cdot\frac{1}{u'(c^{w})}\cdot\Big[(1-p)(1-\rho^{u})\cdot(u(c^{w})-u(b))\\
&+p(1-\rho^{c})\cdot(u^{c}-u(b))\\
&+(1-\rho)(\lambda-f)\cdot u'(c^{w})\cdot\frac{\eta}{1-\eta}\cdot\beta\cdot\bar{J}\Big]
\end{aligned}
$$

$$\Leftrightarrow \quad \frac{\eta}{1-\beta\cdot(1-\lambda)}\Big[(1-\rho)(z-\Omega)-p(1-\rho^u)\cdot c^w-(1-p)(1-\rho^c)\cdot e^c$$
$$+(1-\rho)\cdot\lambda\cdot f\cdot\beta\cdot\bar{J}\Big]$$
$$-\left(\frac{1}{\lambda}-\frac{1}{1-\beta\cdot(1-\lambda)}\right)\cdot(1-p)\int_{\epsilon_s^u}^{\epsilon_{\text{stw}}}\tau_{\text{stw}}(\bar{h}-h_{\text{stw}}(\epsilon))\,dG(\epsilon)$$
$$-\left(\frac{1}{\lambda}-\frac{1}{1-\beta\cdot(1-\lambda)}\right)\cdot p\int_{\epsilon_s^c}^{\max\{\epsilon_{\text{stw}},\epsilon_s^c\}}\tau_{\text{stw}}(\bar{h}-h_{\text{stw}}(\epsilon))\,dG(\epsilon)$$
$$-\frac{1}{\lambda}\cdot(1-\rho)\cdot\frac{1-n}{n}\cdot b$$
$$=\frac{1-\eta}{1-\beta\cdot(1-\lambda)}\cdot\frac{1}{u'(c^w)}\Big[(1-p)(1-\rho^u)(u(c^w)-u(b))$$
$$+p(1-\rho^c)(u^c-u(b))\Big]$$

$$\Leftrightarrow \quad \eta\cdot\Big[(1-\rho)(z-\Omega)+(1-\rho)\cdot f\cdot\beta\cdot\bar{J}\Big]$$
$$-\left(\frac{1-\beta\cdot(1-\lambda)}{\lambda}-1\right)\cdot(1-p)\int_{\epsilon_s^u}^{\epsilon_{\text{stw}}}\tau_{\text{stw}}(\bar{h}-h_{\text{stw}}(\epsilon))\,dG(\epsilon)$$
$$-\left(\frac{1-\beta\cdot(1-\lambda)}{\lambda}-1\right)\cdot p\int_{\epsilon_s^c}^{\max\{\epsilon_{\text{stw}},\epsilon_s^c\}}\tau_{\text{stw}}(\bar{h}-h_{\text{stw}}(\epsilon))\,dG(\epsilon)$$
$$-\frac{1-\beta\cdot(1-\lambda)}{\lambda}\cdot(1-\rho)\cdot\frac{1-n}{n}\cdot b$$
$$=(1-p)(1-\rho^u)(1-\eta)\cdot\left(\frac{u(c^w)-u(b)}{u'(c^w)}+\eta\cdot c^w\right)$$
$$+p(1-\rho^c)(1-\eta)\cdot\left(\frac{u^c-u(b)}{u'(c^w)}+\eta\cdot e^c\right)$$

Inserting the free-entry condition again, we get the wage equation from Appendix 3.E for STW.

$$(1-p)(1-\rho^u)(1-\eta)\cdot\frac{u(c^w)-u(b)}{u'(c^w)}+\eta\cdot c^w$$
$$-\left(\frac{1-\beta\cdot(1-\lambda)}{\lambda}-1\right)\cdot(1-p)\int_{\epsilon_s^u}^{\epsilon_{\text{stw}}}\tau_{\text{stw}}(\bar{h}-h_{\text{stw}}(\epsilon))\,dG(\epsilon)$$
$$-\left(\frac{1-\beta\cdot(1-\lambda)}{\lambda}-1\right)\cdot p\int_{\epsilon_s^c}^{\max\{\epsilon_{\text{stw}},\epsilon_s^c\}}\tau_{\text{stw}}(\bar{h}-h_{\text{stw}}(\epsilon))\,dG(\epsilon)$$
$$-\frac{1-\beta\cdot(1-\lambda)}{\lambda}\cdot(1-\rho)\cdot\frac{1-n}{n}\cdot b$$
$$=\eta\cdot\Big[(1-\rho)(z-\Omega)-(1-\rho)\cdot\frac{1-n}{n}\cdot b+(1-\rho)\cdot\theta\cdot k_v\Big]$$
$$-p(1-\rho^c)(1-\eta)\cdot\left(\frac{u^c-u(b)}{u'(c^w)}+\eta\cdot e^c\right)$$

3.E The full Ramsey Problems

3.E.1 Lay-off Tax

For $\beta \to 1$ the Ramsey planner's problem reads:

$$\max_{b,\tau_{\text{stw}},\epsilon_{\text{stw}}} n^u \cdot u(c^w) + n^c \cdot u^c + (1-n) \cdot u(b) \\ + v^f \cdot u((n \cdot z - n^u \cdot c^w - n^c \cdot e^c - \tau(b) - \theta \cdot (1-n) \cdot k_v)/v^f)$$

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:

$$\begin{aligned}WE =&(1-p)\cdot(1-\rho^u)\cdot\left(\eta\cdot c^w + (1-\eta)\cdot\frac{u(c^w)-u(b)}{u'(c^w)}\right)\\&-\eta\cdot(1-\rho)\cdot\left(z - \frac{1-n}{n}\cdot b + \theta\cdot\frac{k_v}{q}\right)\\&+p\cdot(1-\rho^c)\cdot\left[(1-\eta)\cdot\frac{u^c-u(b)}{u'(c^w)} + \eta\cdot e^c\right] = 0\end{aligned}$$

(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.E.2 Short-Time Work

For $\beta \to 1$ the Ramsey planner's problem reads:

$$\max_{b,\tau_{\text{stw}},\epsilon_{\text{stw}}} n^u \cdot u(c^w) + n^c \cdot u^c + (1-n) \cdot u(b)$$
$$+ \nu^f \cdot u((n \cdot z - n^u \cdot c^w - n^c \cdot e^c - \tau(b) - \theta \cdot (1-n) \cdot k_v)/\nu^f)$$

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:

$$u^c = \frac{1}{1-\rho^c}\Bigg((1-G(\epsilon^p)) \cdot u(c^w) + \int_{\max\{\epsilon_{\text{stw}},\xi_s^c\}}^{\epsilon_p} u(c^w(\epsilon))\, dG(\epsilon)$$
$$+ \int_{\epsilon_s^c}^{\max\{\epsilon_{\text{stw}},\epsilon_s^c\}} u(c^w_{\text{stw}}(\epsilon)) dG(\epsilon)\Bigg)$$

(IX) Average cost of a constrained worker for a firm:

$$e^c = \frac{1}{1-\rho^c}\Bigg[(1-G(\epsilon^p)) \cdot c^w + \int_{\max\{\epsilon_{\text{stw}},\xi_s^c\}}^{\epsilon_p} c^w(\epsilon)\, dG(\epsilon)$$
$$+ \int_{\epsilon_s^c}^{\max\{\epsilon_{\text{stw}},\epsilon_s^c\}} \big(c^w_{\text{stw}}(\epsilon) - \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon))\big)\, dG(\epsilon)\Bigg]$$

(X) Average production (without distortions):

$$z = \frac{1}{1-\rho}\left((1-p)\int_{\xi_s^u}^{\infty} z(\epsilon)\,dG(\epsilon) + p\int_{\max\{\xi_s^c,\epsilon_{\text{stw}}\}}^{\infty} z(\epsilon)\,dG(\epsilon) + p\cdot\int_{\epsilon_s^c}^{\max\{\epsilon_s^c,\epsilon_{\text{stw}}\}} z(\epsilon)\,dG(\epsilon)\right)$$

(XI) Average distortion of working hours:

$$\Omega = \frac{1}{1-\rho}\left((1-p)\int_{\epsilon_s^u}^{\epsilon_{\text{stw}}} \Omega(\epsilon)\,dG(\epsilon) + p\int_{\epsilon_s^c}^{\max\{\epsilon_{\text{stw}},\epsilon_s^c\}} \Omega(\epsilon)\,dG(\epsilon)\right)$$
$$\text{with}\quad \Omega(\epsilon) = \bar{z}(\epsilon) - z_{\text{stw}}(\epsilon)$$

(XII) Job-creation condition:

$$\frac{1}{1-\eta}\cdot\frac{k_v}{q} = \frac{1}{\lambda}\Bigg[(1-\rho)(z-\Omega-\frac{1-n}{n}b) + p(1-\rho^u)\left(\frac{u(c^w)-u(b)}{u'(c^w)} - c^w\right) + (1-p)(1-\rho^c)\left(\frac{u^c-u(b)}{u'(c^w)} - e^c\right) + (1-\rho)\cdot\frac{\lambda-\eta\cdot f}{1-\eta}\cdot\frac{k_v}{q}\Bigg]$$

(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:

$$z_{\text{stw}}(\epsilon_s^u) + \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon_s^u)) + \frac{u(c^w)-u(b)}{u'(c^w)} - c^w + \frac{\lambda-\eta\cdot f}{1-\eta}\cdot\frac{k_v}{q} = 0$$

(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:

$$\frac{u(c^w_{\text{stw}}(\epsilon_s^c)))}{u'(c^w)} + (\lambda - f) \cdot \frac{\eta}{1-\eta} \cdot \frac{k_v}{q} = 0$$

(XIX) Consumption, constrained worker outside STW

$$c^w(\epsilon) = z(\epsilon) + \lambda \cdot \frac{k_v}{q}$$

(XX) Consumption, constrained worker on STW

$$c^w_{\text{stw}}(\epsilon) = z_{\text{stw}}(\epsilon) + \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon)) + \lambda \cdot \frac{k_v}{q}$$

3.F 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.F.1 Employment (Lay-Off Tax)

$$
\begin{aligned}
\frac{\partial \mathscr{L}}{\partial n^s} &= -\lambda_{n^s} + \frac{1-p}{\lambda} \cdot (1-\rho^u) \cdot \left(u(c^w) - u'(c^w) \cdot c^u\right) \\
&\quad + \frac{p}{\lambda} \cdot (1-\rho^c) \cdot \left(u(c^w) - u'(c^w) \cdot e^c\right) \\
&\quad + \frac{1-p}{\lambda} \cdot (1-\rho^u) \cdot \lambda_{n^u} + \frac{p}{\lambda} \cdot (1-\rho^c) \cdot \lambda_{n^c} \;=\; 0
\end{aligned}
$$

$$
\frac{\partial \mathscr{L}}{\partial n^u} = -\lambda_{n^u} + \lambda_n \;=\; 0
$$

$$
\frac{\partial \mathscr{L}}{\partial n^c} = -\lambda_{n^c} + \lambda_n \;=\; 0
$$

$$
\begin{aligned}
\frac{\partial \mathscr{L}}{\partial n} &= -\lambda_n + \lambda_{n^s} \cdot (\lambda - q(\theta) \cdot \theta) + u'(c^w) \cdot [z + b + \theta \cdot c] - u(b) \\
&\quad + \frac{1-\rho}{\lambda} \cdot \frac{b}{n^2} \cdot \lambda_\theta + \eta(1-\rho) \cdot \frac{b}{n^2} \cdot \lambda_c \;=\; 0
\end{aligned}
$$

$$
\begin{aligned}
\Leftrightarrow \quad \frac{\lambda_n}{u'(c^w)} &= (\lambda - q(\theta) \cdot \theta) \cdot \frac{\lambda_{n^s}}{u'(c^w)} + [z + b + \theta \cdot c] - \frac{u(b)}{u'(c^u)} \\
&\quad + \frac{1-\rho}{\lambda} \cdot \frac{b}{n^2} \cdot \frac{\lambda_\theta}{u'(c^w)} + \frac{\lambda \cdot b}{n^2} \cdot \frac{\lambda_c}{u'(c^w)} + \eta \cdot \frac{b}{n^2} \cdot \frac{\lambda_c}{u'(c^w)}
\end{aligned}
$$

3.F.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:

$$
\begin{aligned}
\frac{\partial \mathscr{L}}{\partial \theta} &= -k_v(1-n)u'(c^w) + IE_\theta \cdot k_v + (1-\gamma)q(\theta)(1-n)\lambda_{n^s} \\
&\quad - \frac{\gamma}{1-\eta} k_v \lambda_\theta + \frac{\lambda\gamma - f\eta}{(1-\eta)f}\left(\frac{1-\rho}{\lambda}\lambda_\theta - \lambda_{\epsilon_s^u}\right) \\
&\quad - \left(\lambda \cdot \frac{\gamma}{f} u'(c(\epsilon_s^u)) + \frac{\lambda\gamma - f}{(1-\eta)f}\eta\right)\lambda_{\epsilon_s^c} \\
&\quad - \lambda_c \cdot \frac{\partial WE}{\partial \theta}
\end{aligned}
$$

$$
\Leftrightarrow \quad \frac{\lambda_{n^s}}{u'(c^w)} = \frac{1+\chi}{1-\eta} \cdot \frac{k_v}{q}
$$

with

$$
\begin{aligned}
\chi = {} & \frac{1}{u \cdot u'(c^w)} \cdot \left[\frac{1}{f} \cdot \frac{1}{1-\eta} \cdot \left(\gamma \lambda_\theta - (\lambda\gamma - f\eta) \left(\frac{1-\rho}{\lambda} \lambda_\theta - \lambda_{\epsilon_s^u} \right) \right) \right. \\
& + \left(\lambda \cdot \frac{\gamma}{f} \cdot u'(c^w(\epsilon_s^u)) + \frac{\lambda\gamma - f}{(1-\eta)f} \cdot \eta \right) \lambda_{\epsilon_s^c} \\
& \left. + \frac{1}{k_v} \cdot \left(\lambda_c \cdot \frac{\partial WE}{\partial \theta} - IE_\theta \right) \right]
\end{aligned}
$$

Using the FOCs for employment and labor market tightness gives:

$$
\begin{aligned}
u'(c^w) \cdot \frac{1+\chi}{1-\gamma} \cdot \frac{k_v}{q} = {} & \frac{1-p}{\lambda} \cdot (1-\rho^u) \cdot \left(u(c^w) - u(b) - u'(c^w) \cdot c^w \right) \\
& + \frac{p}{\lambda} \cdot (1-\rho^c) \cdot \left(u^c - u(b) - u'(c^w) \cdot e^c \right) \\
& + u'(c^w) \cdot \frac{1-\rho}{\lambda} \cdot \left(z + b + \frac{\lambda - \gamma f + \chi(\lambda - f)}{1-\gamma} \cdot \frac{k_v}{q} \right) \\
& + \frac{1-\rho}{\lambda} \cdot \frac{\lambda_\theta}{n^s} \cdot \frac{b}{n} + \frac{1-\rho}{n^s} \cdot \frac{b}{n} \cdot \lambda_c
\end{aligned}
$$

Rearranging gives the *Optimal Job Creation Condition:*

$$
\begin{aligned}
\frac{1+\chi}{1-\gamma} \cdot \frac{k_v}{q} = {} & \frac{1-\rho}{\lambda} \cdot (z+b) + \frac{1-\rho}{\lambda} \cdot \frac{\lambda_\theta}{n^s \cdot u'(c^w)} \cdot \frac{b}{n} + \frac{1-\rho}{\lambda} \cdot \frac{\lambda \cdot \lambda_c}{n^s \cdot u'(c^w)} \cdot \frac{b}{n} \\
& + \frac{1-p}{\lambda} \cdot (1-\rho^u) \cdot \left(\frac{u(c^w) - u(b)}{u'(c^w)} - c^w \right) \\
& + \frac{p}{\lambda} \cdot (1-\rho^c) \cdot \left(\frac{u^c - u(b)}{u'(c^w)} - e^c \right) \\
& + \frac{1-\rho}{\lambda} \cdot \frac{\lambda - \gamma f + \chi \cdot (\lambda - f)}{1-\gamma} \cdot \frac{k_v}{q}
\end{aligned}
$$

Subtracting the decentralized job-creation condition from the optimal gives:

$$
\begin{aligned}
\left(\chi - \frac{\eta - \gamma}{1-\eta} \right) \cdot \frac{1}{1-\gamma} \cdot \frac{k_v}{q} = {} & \left(1 + \frac{\lambda_\theta + \eta + \lambda \cdot \lambda_c}{n^s \cdot u'(c^w)} \right) \cdot \frac{1-\rho}{\lambda} \cdot \frac{b}{n} \\
& + \frac{(1-\rho) \cdot (\lambda - f)}{\lambda} \cdot \left(\chi - \frac{\eta - \gamma}{1-\gamma} \right) \cdot \frac{k_v}{q}
\end{aligned}
$$

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 Optimal Separation Condition (Unconstrained Firms)

FOC for the separation threshold of unconstrained firms:

$$
\begin{aligned}
-\frac{\partial \mathscr{L}}{\partial \epsilon_s^u} = {} & \frac{n}{1-\rho} \cdot (1-p) \cdot g(\epsilon_s^u) \cdot u'(c^w) \cdot \left(z(\epsilon_s^u) + \frac{u(c^w)}{u'(c^w)} - c^w - z \right) \\
& + \lambda_\theta \cdot \frac{1-p}{\lambda} \cdot g(\epsilon_s^u) \cdot \left[z(\epsilon_s^u) + \frac{u(c^w) - u(b)}{u'(c^w)} - c^w - \frac{1-n}{n} \cdot b + \frac{\lambda - \eta f}{1-\eta} \cdot \frac{k_v}{q} \right] \\
& + \lambda_c \cdot \frac{\partial WE}{\partial \epsilon_s^u} + \frac{1-p}{\lambda} \cdot g(\epsilon_s^u) \cdot n^s \cdot \lambda_{n^u} + \lambda_{\epsilon_s^u} \cdot \frac{\partial S_{\text{lt}}^u(\epsilon_s^u)}{\partial \epsilon_s^u} = 0
\end{aligned}
$$

Insert for λ_{n^u} and subtract decentralized separation condition of unconstrained firms:

$$
\begin{aligned}
& g(\epsilon_s^u) \cdot \frac{1-p}{\lambda} \cdot n^s \cdot \left[z(\epsilon_s^u) + \frac{u(c^w)}{u'(c^w)} - c^w + \frac{\lambda_n}{u'(c^w)} \right] \\
& - \frac{\lambda_\theta}{u'(c^w)} \cdot \frac{1-p}{\lambda} \cdot g(\epsilon_s^u) \cdot \left[\lambda F + \frac{1-n}{n} \cdot b \right] \\
& + \frac{\lambda_c}{u'(c^w)} \cdot \frac{\partial WE}{\partial \epsilon_s^u} + \frac{\lambda_{\epsilon_s^u}}{u'(c^w)} \cdot \frac{\partial S_{\text{lt}}^u(\epsilon_s^u)}{\partial \epsilon_s^u} = 0
\end{aligned}
$$

Insert for λ_n:

$$
\begin{aligned}
& g(\epsilon_s^u) \cdot \frac{1-p}{\lambda} \cdot n^s \left[z(\epsilon_s^u) + \frac{u(c^w)}{u'(c^w)} - c^w - z \right] \\
& + g(\epsilon_s^u) \cdot \frac{1-p}{\lambda} \cdot n^s \left[(\lambda - q(\theta) \cdot \theta) \cdot \frac{\lambda_{n^s}}{u'(c^w)} + (z + b + \theta \cdot c) - \frac{u(b)}{u'(c^w)} \right. \\
& \qquad \left. + \frac{1-\rho}{\lambda} \cdot \frac{b}{n^2} \cdot \frac{\lambda_\theta}{u'(c^w)} + \eta \cdot \frac{b}{n^2} \cdot \frac{\lambda_c}{u'(c^w)} \right] \\
& + \frac{\lambda_\theta}{u'(c^w)} \cdot \frac{1-p}{\lambda} \cdot g(\epsilon_s^u) \left[\lambda F + \frac{1-n}{n} \cdot b \right] \\
& + \frac{\lambda_c}{u'(c^w)} \cdot \frac{\partial WE}{\partial \epsilon_s^u} + \frac{\lambda_{\epsilon_s^u}}{u'(c^w)} \cdot \frac{\partial S_{\text{stw}}^u(\epsilon_s^u)}{\partial \epsilon_s^u} = 0
\end{aligned}
$$

$$
\begin{aligned}
\iff \quad & z(\epsilon_s^u) + \frac{u(c^w) - u(b)}{u'(c^w)} - c^w + b + \big(\lambda - q(\theta) \cdot \theta\big) \cdot \frac{\lambda_{n^s}}{u'(c^w)} + \theta \cdot c \\
& - \frac{\lambda_\theta}{n^s \cdot u'(c^w)} \left(\lambda F + \frac{1-n}{n} \cdot b - \frac{b}{n} \right) \\
& + \frac{\lambda_c}{n^s \cdot u'(c^w)} \left(\eta \cdot \frac{b}{n} + \frac{\lambda}{1-p} \cdot \frac{1}{g(\epsilon_s^u)} \cdot \frac{\partial WE}{\partial \epsilon_s^u} \right) \\
& + \frac{\lambda_{\epsilon_s^u}}{n^s \cdot u'(c^w)} \cdot \frac{\lambda}{1-p} \cdot \frac{1}{g(\epsilon_s^u)} \cdot \frac{\partial S_{\text{lt}}^u(\epsilon_s^u)}{\partial \epsilon_s^u} = 0
\end{aligned}
$$

Inserting λ_n^s gives the *Optimal Separation Condition*:

$$z(\epsilon_s^u) + \frac{u(c^w) - u(b)}{u'(c^w)} - c^w + b + \frac{\lambda - \gamma \cdot f + \chi \cdot (\lambda - f)}{1-\gamma} \cdot \frac{k_v}{q}$$
$$- \frac{\lambda_\theta}{n^s \cdot u'(c^w)} \cdot (\lambda \cdot F - b) + \frac{\lambda_c}{n^s \cdot u'(c^w)} \cdot \left(\eta \cdot \frac{b}{n} + \frac{\lambda}{1-p} \cdot \frac{1}{g(\epsilon_s^u)} \cdot \frac{\partial WE}{\partial \epsilon_s^u}\right) = 0$$

Subtracting the decentralized separation condition and rearranging gives:

$$\begin{aligned} \lambda \cdot F = {} & b + (\lambda - \theta q(\theta)) \cdot \left(\chi - \frac{\eta - \gamma}{1-\gamma}\right) \cdot \frac{k_v}{q} \\ & + \frac{\lambda_{\epsilon_s^u}}{n^s \cdot u'(c^w)} \cdot \frac{\lambda}{1-p} \cdot \frac{1}{g(\epsilon_s^u)} \cdot \frac{\partial S_{\text{lt}}^u(\epsilon_s^u)}{\partial \epsilon_s^u} \\ & - \frac{\lambda_\theta}{n^s \cdot u'(c^w)} \cdot (\lambda F - b) \\ & + \frac{\lambda_c}{n^s \cdot u'(c^w)} \left(\eta \cdot \frac{b}{n} + \frac{\lambda}{1-p} \cdot \frac{1}{g(\epsilon_s^u)} \cdot \frac{\partial WE}{\partial \epsilon_s^u}\right) \end{aligned}$$

$$\begin{aligned} \iff \quad 0 = {} & \left(1 + \frac{\lambda_\theta}{u'(c^w) \cdot n^s}\right) \cdot \left(b + (\lambda - f) \cdot \frac{\frac{1-\rho}{\lambda}}{\rho + (1-\rho) \cdot \frac{f}{\lambda}} \cdot \frac{b}{n} - \lambda \cdot F\right) \\ & + \frac{\lambda_c}{n^s \cdot u'(c^w)} \left(\eta \cdot \frac{b}{n} + \frac{\lambda}{1-p} \cdot \frac{1}{g(\epsilon_s^u)} \cdot \frac{\partial WE}{\partial \epsilon_s^u}\right) \\ & + \frac{\lambda}{1-p} \cdot \frac{1}{g(\epsilon_s^u)} \cdot \frac{\lambda_{\epsilon_s^u}}{n^s \cdot u'(c^w)} \cdot \frac{\partial S_{\text{lt}}^u(\epsilon_s^u)}{\partial \epsilon_s^u} \end{aligned}$$

$$\begin{aligned} \iff \quad 0 = {} & \left(1 + \frac{\lambda_\theta}{u'(c^w) \cdot n^s}\right) \cdot \left(b + (\lambda - f) \cdot \frac{b}{f} - \lambda F\right) \\ & + \frac{\lambda}{1-p} \cdot \frac{1}{g(\epsilon_s^u)} \cdot \frac{\lambda_{\epsilon_s^u}}{n^s \cdot u'(c^w)} \frac{\partial S_{\text{stw}}^u(\epsilon_s^u)}{\partial \epsilon_s^u} \\ & + \frac{\lambda_c}{n^s \cdot u'(c^w)} \cdot \left(\eta \cdot \frac{b}{n} + \frac{\lambda}{1-p} \cdot \frac{1}{g(\epsilon_s^u)} \cdot \frac{\partial WE}{\partial \epsilon_s^u}\right) \end{aligned}$$

$$\begin{aligned} \iff \quad 0 = {} & \left(1 + \frac{\lambda_\theta}{u'(c^w) \cdot n^s}\right) \cdot \left(\frac{\lambda}{f} \cdot b - \lambda F\right) \\ & + \frac{\lambda}{1-p} \cdot \frac{1}{g(\epsilon_s^u)} \cdot \frac{\lambda_{\epsilon_s^u}}{n^s \cdot u'(c^w)} \cdot \frac{\partial S_{\text{lt}}^u(\epsilon_s^u)}{\partial \epsilon_s^u} \\ & + \frac{\lambda_c}{n^s \cdot u'(c^w)} \cdot \left(\eta \cdot \frac{b}{n} + \frac{\lambda}{1-p} \cdot \frac{1}{g(\epsilon_s^u)} \cdot \frac{\partial WE}{\partial \epsilon_s^u}\right) \end{aligned}$$

Rearranging gives the Lagrange multiplier for unconstrained firms:

$$
\begin{aligned}
-\frac{\lambda_{\epsilon_s^u}}{n^s \cdot u'(c^w)} = \frac{1-p}{\lambda} \cdot \frac{g(\epsilon_s^u)}{\frac{\partial S^u_{\text{stw}}(\epsilon_s^u)}{\partial \epsilon_s^u}} \cdot \Bigg(\left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right) \cdot \left(\frac{\lambda}{f} \cdot b - \lambda F\right) \\
+ \frac{\lambda_c}{n^s \cdot u'(c^w)} \cdot \left(\eta \cdot \frac{b}{n} + \frac{1}{g(\epsilon_s^u)} \cdot \frac{\lambda}{1-p} \cdot \frac{1}{g(\epsilon_s^u)} \cdot \frac{\partial WE}{\partial \epsilon_s^u}\right)\Bigg)
\end{aligned}
$$

3.F.4 Optimal Separation Condition (Constrained Firms)

The FOC for the separation threshold in the constrained firm is:

$$
\begin{aligned}
-\frac{\partial \mathscr{L}}{\partial \epsilon_s^c} = & \frac{n}{1-\rho} \cdot p \cdot g(\epsilon_s^c) \cdot u'(c^w) \left(z(\epsilon_s^c) + \frac{u(c^w(\epsilon_s^c)) - u(c^w)}{u'(c^w)} - c^w(\epsilon_s^c) - z \right) \\
& + \lambda_\theta \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^c) \left[z(\epsilon_s^c) + \frac{c^w(\epsilon_s^c) - u(b)}{u'(c^w)} - c^w(\epsilon_s^c) - \frac{1-n}{n} \cdot b + \frac{\lambda - \eta f}{1-\eta} \cdot \frac{k_v}{q} \right] \\
& + \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot n^s \cdot \lambda_{n^c} \\
& + \lambda_c \cdot \frac{\partial WE}{\partial \epsilon_s^c} \\
& + \lambda_{\epsilon_s^c} \cdot \frac{\partial S^c_{\text{lt}}(\epsilon_s^c)}{\partial \epsilon_s^c} = 0
\end{aligned}
$$

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*

$$
\begin{aligned}
-\frac{\partial \mathcal{L}}{\partial \epsilon_s^c} = &\; z(\epsilon_s^c) + \frac{u(c^w(\epsilon_s^c))}{u'(c^w)} - c^w(\epsilon_s^c) + b + \frac{\lambda - \gamma \cdot f + \chi \cdot (\lambda - f)}{1-\gamma} \cdot \frac{k_v}{q} \\
&+ \frac{\lambda_\theta}{n^s \cdot u'(c^w)} \cdot b \\
&+ \frac{\lambda_c}{n^s \cdot u'(c^w)} \left(\eta \cdot \frac{b}{n} + \frac{1}{g(\epsilon_s^c)} \cdot \frac{\lambda}{p} \cdot \frac{\partial WE}{\partial \epsilon_s^c} \right) \\
&+ \frac{\lambda}{p} \cdot \frac{1}{n^s \cdot g(\epsilon_s^c)} \cdot \frac{\lambda_{\epsilon_s^c}}{u'(c^w)} \cdot \frac{\partial S_{\mathrm{lt}}^c(\epsilon_s^c)}{\partial \epsilon_s^c} = 0
\end{aligned}
$$

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:

$$
\begin{aligned}
-\frac{\lambda_{\epsilon_s^c}}{n^s \cdot u'(c^w)} = &\; \frac{p}{\lambda} \cdot \frac{g(\epsilon_s^c)}{\frac{\partial S_{\mathrm{stw}}^c(\epsilon_s^c)}{\partial \epsilon_s^c}} \cdot \left(\left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)} \right) \cdot \frac{\lambda}{f} \cdot b \right. \\
&\left. + \frac{\lambda_c}{n^s \cdot u'(c^w)} \cdot \left(\eta \cdot \frac{b}{n} + \frac{1}{g(\epsilon_s^c)} \cdot \frac{\lambda}{p} \cdot \frac{\partial WE}{\partial \epsilon_s^c} \right) \right)
\end{aligned}
$$

3.F.5 The Optimal Lay-Off Tax

We start from

$$
\frac{\partial \mathcal{L}}{\partial F} = -\lambda_{\epsilon_s^u} = 0
$$

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

$$
\begin{aligned}
-\frac{\lambda_{\epsilon_s^u}}{n^s \cdot u'(c^w)} = &\; \frac{p}{\lambda} \cdot \frac{g(\epsilon_s^u)}{\frac{\partial S_{\mathrm{lt}}^u(\epsilon_s^u)}{\partial \epsilon_s^u}} \cdot \left(\left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)} \right) \cdot \left(\frac{\lambda}{f} \cdot b - \lambda \cdot F \right) \right. \\
&\left. + \frac{\lambda_c}{n^s \cdot u'(c^w)} \left(\eta \cdot \frac{b}{n} + \frac{\lambda}{g(\epsilon_s^u)} \cdot \frac{\lambda}{1-p} \cdot \frac{\partial WE}{\partial \epsilon_s^u} \right) \right)
\end{aligned}
$$

This leads to:

$$
\begin{aligned}
F &= \frac{1}{f} \cdot b + BE_{lt} \\
BE_{lt} &= \frac{1}{\lambda} \frac{1}{\left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right)} \cdot \frac{\lambda_c}{n^s \cdot u'(c^w)} \cdot \left(\eta \cdot \frac{b}{n} + \frac{1}{g(\epsilon_s^u)} \cdot \frac{\lambda}{1-p} \cdot \frac{\partial WE}{\partial \epsilon_s^u}\right)
\end{aligned}
$$

3.F.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.

$$
\begin{aligned}
(1-\eta) \cdot \left(u'(b) - u'(c^w)\right) &= \lambda_\theta \cdot (1-\rho) \cdot \left(\frac{1-n}{n} + \frac{u'(b)}{u'(c^w)}\right) \\
&\quad + \left(\lambda_{\epsilon_s^c} + \lambda_{\xi_s^c}\right) \cdot \left(-\frac{u'(b)}{u'(c^w)}\right) \\
&\quad + \lambda_c \cdot (1-\rho) \cdot \left(\eta \cdot \frac{1-n}{n} + (1-\eta) \cdot \frac{u'(b)}{u'(c^w)}\right)
\end{aligned}
$$

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:

$$
\begin{aligned}
-\frac{\lambda_{\epsilon_s^u}}{n^s \cdot u'(c^w)} &= \frac{1-p}{\lambda} \cdot g(\epsilon_s^u) \cdot \frac{1}{\frac{\partial S_{\text{stw}}^u(\epsilon_s^u)}{\partial \epsilon_s^u}} \cdot \left[\left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right) \cdot \left(\frac{\lambda}{f} \cdot b - \lambda \cdot F\right)\right. \\
&\quad \left. + \frac{\lambda_c}{n^s \cdot u'(c^w)} \cdot \left(\eta \cdot \frac{b}{n} + \frac{1}{g(\epsilon_s^u)} \cdot \frac{\lambda}{1-p} \cdot \frac{\partial WE}{\partial \epsilon_s^u}\right)\right]
\end{aligned}
$$

$$
\begin{aligned}
-\frac{\lambda_{\epsilon_s^c}}{n^s \cdot u'(c^w)} &= \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot \frac{1}{\frac{\partial S_{\text{stw}}^c(\epsilon_s^c)}{\partial \epsilon_s^c}} \cdot \left[\left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right) \cdot \frac{\lambda}{f} \cdot b\right. \\
&\quad \left. + \frac{\lambda_c}{n^s \cdot u'(c^w)} \cdot \left(\eta \cdot \frac{b}{n} + \frac{1}{g(\epsilon_s^c)} \cdot \frac{\lambda}{p} \cdot \frac{\partial WE}{\partial \epsilon_s^c}\right)\right]
\end{aligned}
$$

First, let us find an expression for λ_c:

$$
\begin{aligned}
\frac{\partial \mathscr{L}}{\partial c^w} = &-\Bigg((1-p)\cdot(1-\rho^u)\cdot\frac{u(c^w)-u(b)}{u'(c^w)}\cdot\frac{u''(c^w)}{u'(c^w)} \\
&+p\cdot(1-\rho^c)\cdot\frac{u^c-u(b)}{u'(c^w)}\cdot\frac{u''(c^w)}{u'(c^w)}\Bigg)\cdot\lambda_\theta \\
&-(1-\eta)\cdot\Bigg((1-p)\cdot(1-\rho^u)\cdot\left(1-(1-\eta)\cdot\frac{u(c^w)-u(b)}{u'(c^w)}\cdot\frac{u''(c^w)}{u'(c^w)}\right) \\
&\quad -p\cdot(1-\rho^c)\cdot(1-\eta)\cdot\frac{u^c-u(b)}{u'(c^w)}\cdot\frac{u''(c^w)}{u'(c^w)}\Bigg)\cdot\lambda_c \\
&+\frac{u(c^w)-u(b)}{u'(c^w)}\cdot\frac{u''(c^w)}{u'(c^w)}\cdot\lambda_{\epsilon_s^u} \\
&+\frac{u(c_{\text{stw}}(\epsilon_s^c))-u(b)}{u'(c^w)}\cdot\frac{u''(c^w)}{u'(c^w)}\cdot\lambda_{\epsilon_s^c} = 0
\end{aligned}
$$

Next we can solve for the Lagrange multiplier of the wage equation:

$$
\begin{aligned}
\lambda_c = -\frac{1}{\left(\frac{\partial WE}{\partial c^w}\right)^{\text{total}}}\Bigg[&\frac{\partial \epsilon_s^u}{\partial c^w}\cdot\frac{1-p}{\lambda}\cdot g(\epsilon_s^u)\cdot n^s\cdot u'(c^w)\cdot\left(\frac{\lambda}{f}\cdot b-\lambda\cdot F\right) \\
&+\frac{\partial \epsilon_s^c}{\partial c^w}\cdot\frac{p}{\lambda}\cdot g(\epsilon_s^c)\cdot n^s\cdot u'(c^w)\cdot\frac{\lambda}{f}\cdot b \\
&+\left(-\frac{\partial S^{\text{total}}}{\partial c^w}\right)\cdot\lambda_\theta\Bigg]
\end{aligned}
$$

Insert λ_c into the Lagrange multipliers for the separation conditions. To do that, let us rewrite the Lagrange multipliers of the separation conditions:

$$
\begin{aligned}
\lambda_{\epsilon_s^u} = -\frac{1}{\frac{\partial S_{\text{stw}}^u(\epsilon_s^u)}{\partial \epsilon_s^u}}\cdot&\Bigg(n^s\cdot u'(c^w)\cdot\frac{1-p}{\lambda}\cdot g(\epsilon_s^u)\cdot\left(1+\frac{\lambda_\theta}{n^s\cdot u'(c^w)}\right)\cdot\left(\frac{\lambda}{f}\cdot b-\lambda\cdot F\right) \\
&+\lambda_c\cdot\left(\frac{\partial n}{\partial \epsilon_s^u}\cdot\left(-\frac{\partial WE}{\partial n}\right)+\frac{\partial WE}{\partial \epsilon_s^u}\right)\Bigg) \\
\lambda_{\epsilon_s^c} = -\frac{1}{\frac{\partial S_{\text{stw}}^c(\epsilon_s^c)}{\partial \epsilon_s^c}}\cdot&\Bigg(n^s\cdot u'(c^w)\cdot\frac{p}{\lambda}\cdot g(\epsilon_s^c)\cdot\left(1+\frac{\lambda_\theta}{n^s\cdot u'(c^w)}\right)\cdot\frac{\lambda}{f}\cdot b \\
&+\lambda_c\cdot\left(\frac{\partial n}{\partial \epsilon_s^c}\cdot\left(-\frac{\partial WE}{\partial n}\right)+\frac{\partial WE}{\partial \epsilon_s^c}\right)\Bigg)
\end{aligned}
$$

Inserting λ_c gives:

$$\begin{aligned}
\lambda_{\epsilon_s^u} = & -\frac{1}{\frac{\partial S_{\text{stw}}^u(\epsilon_s^u)}{\partial \epsilon_s^u}} \cdot \left(n^s \cdot u'(c^w) \cdot \frac{1-p}{\lambda} \cdot g(\epsilon_s^u) \cdot \left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right) \cdot \left(\frac{\lambda}{f} \cdot b - \lambda \cdot F\right) \right. \\
& + \left(\frac{\partial n}{\partial \epsilon_s^u} \cdot \left(-\frac{\partial c^w}{\partial n}\right)^{\text{total}} + \left(\frac{\partial c^w}{\partial \epsilon_s^u}\right)^{\text{total}} \right) \\
& \cdot \left[\frac{\partial \epsilon_s^u}{\partial c^w} \cdot \frac{1-p}{\lambda} \cdot g(\epsilon_s^u) \cdot n^s \cdot u'(c^w) \cdot \left(\frac{\lambda}{f} \cdot b - \lambda \cdot F\right) \right. \\
& \left.\left. + \frac{\partial \epsilon_s^c}{\partial c^w} \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot n^s \cdot u'(c^w) \cdot \frac{\lambda}{f} \cdot b + \left(-\frac{\partial S^{\text{total}}}{\partial c^w}\right) \cdot \lambda_\theta \right] \right)
\end{aligned}$$

$$\begin{aligned}
\lambda_{\epsilon_s^c} = & \frac{1}{\frac{\partial S_{\text{stw}}^c(\epsilon_s^c)}{\partial \epsilon_s^c}} \cdot \left(n^s \cdot u'(c^w) \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot \left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right) \cdot \frac{\lambda}{f} \cdot b \right. \\
& + \left(\left(\frac{\partial n}{\partial \epsilon_s^c}\right) \cdot \left(-\frac{\partial c^w}{\partial n}\right)^{\text{total}} + \left(\frac{\partial c^w}{\partial \epsilon_s^c}\right)^{\text{total}} \right) \\
& \cdot \left[\frac{\partial \epsilon_s^u}{\partial c^w} \cdot \frac{1-p}{\lambda} \cdot g(\epsilon_s^u) \cdot n^s \cdot u'(c^w) \cdot \left(\frac{\lambda}{f} \cdot b - \lambda \cdot F\right) \right. \\
& \left.\left. + \frac{\partial \epsilon_s^c}{\partial c^w} \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot n^s \cdot u'(c^w) \cdot \frac{\lambda}{f} \cdot b + \left(-\frac{\partial S^{\text{total}}}{\partial c^w}\right) \cdot \lambda_\theta \right] \right)
\end{aligned}$$

Insert λ_c into the Lagrange multipliers for the separation conditions. To do that, let us rewrite the Lagrange multipliers of the separation conditions:

$$\begin{aligned}
\lambda_{\epsilon_s^u} = & -\frac{1}{\frac{\partial S_{\text{stw}}^u(\epsilon_s^u)}{\partial \epsilon_s^u}} \cdot \left(-n^s \cdot u'(c^w) \cdot \frac{1-p}{\lambda} \cdot g(\epsilon_s^u) \cdot \left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right) \right. \\
& \left. \cdot \left(\frac{\lambda}{f} \cdot b - \lambda \cdot F\right) + \lambda_c \cdot \left(\frac{\partial n}{\partial \epsilon_s^u} \cdot \left(-\frac{\partial WE}{\partial n}\right) + \frac{\partial WE}{\partial \epsilon_s^u} \right) \right) \\
\lambda_{\epsilon_s^c} = & -\frac{1}{\frac{\partial S_{\text{stw}}^c(\epsilon_s^c)}{\partial \epsilon_s^c}} \cdot \left(n^s \cdot u'(c^w) \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot \left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right) \cdot \frac{\lambda}{f} \cdot b \right. \\
& \left. + \lambda_c \cdot \left(\frac{\partial n}{\partial \epsilon_s^c} \cdot \left(-\frac{\partial WE}{\partial n}\right) + \frac{\partial WE}{\partial \epsilon_s^c} \right) \right)
\end{aligned}$$

Inserting λ_c gives:

$$
\begin{aligned}
\lambda_{\epsilon_s^u} = & -\frac{1}{\frac{\partial S^u_{\text{stw}}(\epsilon_s^u)}{\partial \epsilon_s^u}} \cdot \Bigg(n^s \cdot u'(c^w) \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^u) \cdot \left(\lambda + \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right) \cdot \left(\frac{\lambda}{f} \cdot b - \lambda \cdot F\right) \\
& + \left(\frac{\partial n}{\partial \epsilon_s^u} \cdot \left(-\frac{\partial c^w}{\partial n}\right)^{\text{total}} + \left(\frac{\partial c^w}{\partial \epsilon_s^u}\right)^{\text{total}} \right) \\
& \cdot \Bigg[\frac{\partial \epsilon_s^u}{\partial c^w} \cdot \frac{1-p}{\lambda} \cdot g(\epsilon_s^u) \cdot n^s \cdot u'(c^w) \cdot \left(\frac{\lambda}{f} \cdot b - \lambda \cdot F\right) \\
& + \frac{\partial \epsilon_s^c}{\partial c^w} \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot n^s \cdot u'(c^w) \cdot \frac{\lambda}{f} \cdot b + \left(-\frac{\partial S^{\text{total}}}{\partial c^w}\right) \cdot \lambda_\theta \Bigg] \Bigg)
\end{aligned}
$$

$$
\begin{aligned}
\lambda_{\epsilon_s^c} = & -\frac{1}{\frac{\partial S^c_{\text{stw}}(\epsilon_s^c)}{\partial \epsilon_s^c}} \cdot \Bigg(n^s \cdot u'(c^w) \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot \left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right) \cdot \frac{\lambda}{f} \cdot b \\
& + \left(\left(\frac{\partial n}{\partial \epsilon_s^c}\right) \cdot \left(-\frac{\partial c^w}{\partial n}\right)^{\text{total}} + \left(\frac{\partial c^w}{\partial \epsilon_s^c}\right)^{\text{total}} \right) \\
& \cdot \Bigg[\frac{\partial \epsilon_s^u}{\partial c^w} \cdot \frac{1-p}{\lambda} \cdot g(\epsilon_s^u) \cdot n^s \cdot u'(c^w) \cdot \left(\frac{\lambda}{f} \cdot b - \lambda \cdot F\right) \\
& + \frac{\partial \epsilon_s^c}{\partial c^w} \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot n^s \cdot u'(c^w) \cdot \frac{\lambda}{f} \cdot b + \left(-\frac{\partial S^{\text{total}}}{\partial c^w}\right) \cdot \lambda_\theta \Bigg] \Bigg)
\end{aligned}
$$

Now we have everything to calculate λ_θ from the two equations:

$$
(1) \quad \left(\chi - \frac{\eta - \gamma}{1-\eta}\right) \cdot \frac{1}{1-\gamma} \cdot \frac{k_v}{q} = \left(1 + \frac{\lambda_\theta + \eta \cdot \lambda \cdot \lambda_c}{n^s \cdot u'(c^w)}\right) \cdot \frac{b}{f}
$$

$$
\begin{aligned}
(2) \quad \chi = & \frac{1}{1-\eta} \cdot \frac{1}{u \cdot f} \cdot \frac{1}{u'(c^w)} \cdot \left[\gamma - (\lambda\gamma - f\eta) \cdot \left(\frac{1}{\lambda}(1-\rho)\lambda_\theta - \lambda_{\epsilon_s^u}\right)\right] \\
& + \frac{1}{u} \cdot \frac{1}{u'(c^w)} \cdot \left(\lambda \cdot \frac{\gamma}{f} \cdot u'(c^w) + \left(\frac{1}{1-\eta} \cdot \frac{\lambda\gamma - f}{f}\right) \cdot \eta\right) \cdot \lambda_{\epsilon_s^c} \\
& + \frac{1}{u} \cdot \frac{1}{u'(c^w)} \cdot \lambda_c \cdot \frac{\partial WE}{\partial \theta} \cdot \frac{1}{k_v} \\
& - \frac{1}{u} \cdot \frac{1}{u'(c^w)} \cdot IE_\theta
\end{aligned}
$$

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

$$\begin{aligned}M \cdot \lambda_\theta \;&=\; u'(c^w) \cdot \frac{\partial f}{\partial \theta} \cdot u \cdot \left(\frac{\eta - \gamma}{(1-\gamma)(1-\eta)} \cdot \frac{k_v}{q} + \left(1 + \frac{\lambda_\theta}{n^s u'(c^w)}\right) \cdot \frac{b}{f}\right)\\ &\quad + u'(c^w) \cdot n^s \cdot \left(-\frac{\partial \epsilon_s^u}{\partial \theta}\right)^{\text{total}} \cdot g(\epsilon_s^u) \cdot \frac{1-p}{\lambda}\left(\frac{\lambda}{f}b - \lambda F\right)\\ &\quad + u'(c^w) \cdot n^s \cdot \left(-\frac{\partial \epsilon_s^c}{\partial \theta}\right)^{\text{total}} \cdot g(\epsilon_s^c) \cdot \frac{p}{\lambda} \cdot \frac{\lambda}{f}b\\ &\quad + u'(c^w) \cdot n^s \cdot \tilde{IE}_\theta\end{aligned}$$

with

$$
\begin{aligned}
M = & \left[\gamma - (\lambda\gamma - f \cdot \eta) \cdot \frac{1}{\lambda} \cdot (1-\rho)\right] \cdot \frac{1}{1-\eta} \cdot \frac{k_v}{f} \\
& + \left(-\frac{\partial f}{\partial \theta} \cdot u \cdot \frac{b}{f}\right) \\
& + \left(\frac{\partial \epsilon_s^u}{\partial \theta}\right)^{\text{total}} \cdot g(\epsilon_s^u) \cdot \frac{1-p}{\lambda} \cdot \left(\frac{\lambda}{f} b - \lambda \cdot F\right) \\
& + \left(\frac{\partial \epsilon_s^c}{\partial \theta}\right)^{\text{total}} \cdot g(\epsilon_s^c) \cdot \frac{p}{\lambda} \cdot \frac{\lambda}{f} b \\
& + \left(\frac{\partial c^w}{\partial \theta}\right)^{\text{total}} \cdot \left(-\frac{\partial S}{\partial c^w}\right)^{\text{total}} \\
& + \tilde{IE}_\theta
\end{aligned}
$$

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:

$$
\begin{aligned}
\lambda_\theta = & \ u'(c^w) \cdot \left(\frac{\partial f}{\partial S}\right)^{\text{ge}} \cdot u \cdot \left(\frac{\eta - \gamma}{(1-\gamma)(1-\eta)} \cdot \frac{k_v}{q} + \frac{\lambda}{f} \cdot b\right) \\
& + u'(c^w) \cdot \left(-\frac{\partial \epsilon_s^u}{\partial S}\right)^{\text{ge}} \cdot \frac{\partial n^u}{\partial \epsilon_s^u} \cdot \left(\frac{\lambda}{f} b - \lambda F\right) \\
& + u'(c^w) \cdot \left(-\frac{\partial \epsilon_s^c}{\partial S}\right)^{\text{ge}} \cdot \frac{\partial n^c}{\partial \epsilon_s^c} \cdot \frac{\lambda}{f} b \\
& + u'(c^w) \cdot n^c \cdot \left(\frac{\partial \theta}{\partial S}\right)^{\text{ge}} \cdot \hat{IE}_\theta
\end{aligned}
$$

Note that

$$
\left(\frac{\partial f}{\partial S}\right)^{\text{ge}} = \frac{1}{M} \cdot \frac{\partial f}{\partial S}
$$

$$
\left(\frac{\partial \epsilon_s^u}{\partial S}\right)^{\text{ge}} = \frac{1}{M} \cdot \left(\frac{\partial \epsilon_s^u}{\partial S}\right)^{\text{total}}
$$

$$
\left(\frac{\partial \epsilon_s^c}{\partial S}\right)^{\text{ge}} = \frac{1}{M} \cdot \left(\frac{\partial \epsilon_s^c}{\partial S}\right)^{\text{total}}
$$

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:

$$\begin{aligned}\lambda_{\epsilon_s^u} =& -\frac{u'(c^w)}{\left(\frac{\partial S^u_{\text{stw}}(\epsilon_s^u)}{\partial \epsilon_s^u}\right)^{\text{total}}} \Bigg[\left(\frac{\partial n^u}{\partial \epsilon_s^u}\right)^{\text{ge}} \cdot \left(\frac{\lambda}{f} b - \lambda F\right) + \left(\frac{\partial \epsilon_s^c}{\partial \epsilon_s^u}\right)^{\text{ge}} \cdot \frac{\partial n^c}{\partial \epsilon_s^c} \cdot \frac{\lambda}{f} b\\ &+ \frac{\partial c^w}{\partial \epsilon_s^u} \cdot \left(-\frac{\partial f}{\partial c^w}\right)^{\text{ge}} \cdot u \cdot \left(\frac{\eta-\gamma}{(1-\gamma)(1-\eta)} \cdot \frac{k_v}{q} + \frac{b}{f}\right)\\ &+ n^c \cdot \frac{\partial c^w}{\partial \epsilon_s^u} \cdot \left(-\frac{\partial \theta}{\partial c^w}\right)^{\text{ge}} \cdot \hat{IE}_\theta \Bigg]\end{aligned}$$

$$\begin{aligned}\lambda_{\epsilon_s^c} =& -\frac{u'(c^w)}{\left(\frac{\partial S^c_{\text{stw}}(\epsilon_s^c)}{\partial \epsilon_s^c}\right)^{\text{total}}} \Bigg[\frac{\partial \epsilon_s^u}{\partial \epsilon_s^c} \cdot \left(\frac{\partial n^u}{\partial \epsilon_s^u}\right)^{\text{ge}} \cdot \left(\frac{\lambda}{f} b - \lambda F\right)\\ &+ \left(\frac{\partial n^c}{\partial \epsilon_s^c}\right)^{\text{ge}} \cdot \frac{\lambda}{f} b + \frac{\partial c^w}{\partial \epsilon_s^c} \cdot \left(-\frac{\partial f}{\partial c^w}\right)^{\text{ge}} \cdot u \cdot \left(\frac{\eta-\gamma}{(1-\gamma)(1-\eta)} \cdot \frac{k_v}{q} + \frac{b}{f}\right)\\ &+ n^c \cdot \frac{\partial c^w}{\partial \epsilon_s^c} \cdot \left(-\frac{\partial \theta}{\partial c^w}\right)^{\text{ge}} \cdot \hat{IE}_\theta \Bigg]\end{aligned}$$

For convenience, the optimal UI benefits are replicated here:

$$\begin{aligned}(1-n)\cdot\big(u'(b)-u'(c^w)\big) \;=\; &\lambda_\theta\cdot(1-\rho)\cdot\left(\frac{1-n}{n}+\frac{u'(b)}{u'(c^w)}\right)\cdot\left(-\frac{u'(b)}{u'(c^w)}\right)\\ &+(\lambda_{\epsilon_s^u}+\lambda_{\epsilon_s^c})\cdot\left(-\frac{u'(b)}{u'(c^w)}\right)\\ &+\lambda_c\cdot(1-\rho)\cdot\left(\eta\cdot\frac{1-n}{n}+(1-\eta)\cdot\frac{u'(b)}{u'(c^w)}\right)\end{aligned}$$

Inserting the Lagrange multiplier gives:

$$\begin{aligned}&(1-n)\cdot\big(u'(b)-u'(c^w)\big)\\ &\quad= u'(c^w)\cdot\left[\frac{\partial c^w}{\partial\epsilon_s^u}\cdot\left(-\frac{\partial f}{\partial c^w}\right)^{\text{ge}}\cdot u\cdot\left(\frac{\eta-\gamma}{(1-\gamma)(1-\eta)}\cdot\frac{k_v}{q}+\frac{b}{f}\right)\right]\\ &\quad+u'(c^w)\cdot\left[\left(\frac{\partial\epsilon_s^u}{\partial b}\right)^{\text{ge}}\cdot\frac{\partial n^u}{\partial\epsilon_s^u}\cdot\left(\frac{\lambda}{f}b-\lambda F\right)\right]\\ &\quad+u'(c^w)\cdot\left[\left(\frac{\partial\epsilon_s^c}{\partial b}\right)^{\text{ge}}\cdot\frac{\partial n^c}{\partial\epsilon_s^c}\cdot\frac{\lambda}{f}b\right]\\ &\quad+u'(c^w)\cdot\left[n^c\cdot\left(-\frac{\partial\theta}{\partial b}\right)^{\text{ge}}\cdot\hat{I\!E}_\theta\right].\end{aligned}$$

With:

$$\begin{aligned}\left(\frac{\partial\epsilon_s^u}{\partial b}\right)^{\text{ge}}\cdot\frac{\partial n^u}{\partial\epsilon_s^u} = &\underbrace{\left(\frac{\partial\epsilon_s^u}{\partial b}\right)^{\text{ge}}\cdot\frac{\partial n^u}{\partial\epsilon_s^u}}_{\text{direct effects}}\\ &+\underbrace{\left[\frac{\partial S}{\partial b}\cdot\left(\frac{\partial\epsilon_s^u}{\partial S}\right)^{\text{ge}}+\frac{\partial\epsilon_s^c}{\partial b}\cdot\left(\frac{\partial\epsilon_s^u}{\partial\epsilon_s^c}\right)^{\text{ge}}+\frac{\partial c^w}{\partial b}\cdot\left(\frac{\partial\epsilon_s^u}{\partial c^w}\right)^{\text{ge}}\right]\cdot\frac{\partial n^u}{\partial\epsilon_s^u}}_{\text{indirect effect}}\end{aligned}$$

$$\begin{aligned}\left(\frac{\partial\epsilon_s^c}{\partial b}\right)^{\text{ge}}\cdot\frac{\partial n^c}{\partial\epsilon_s^c} = &\underbrace{\left(\frac{\partial\epsilon_s^c}{\partial b}\right)^{\text{ge}}\cdot\frac{\partial n^c}{\partial\epsilon_s^c}}_{\text{direct effects}}\\ &+\underbrace{\left[\frac{\partial S}{\partial b}\cdot\left(\frac{\partial\epsilon_s^c}{\partial S}\right)^{\text{ge}}+\frac{\partial\epsilon_s^u}{\partial b}\cdot\left(\frac{\partial\epsilon_s^c}{\partial\epsilon_s^u}\right)^{\text{ge}}+\frac{\partial c^w}{\partial b}\cdot\left(\frac{\partial\epsilon_s^c}{\partial c^w}\right)^{\text{ge}}\right]\cdot\frac{\partial n^c}{\partial\epsilon_s^c}}_{\text{indirect effect}}\end{aligned}$$

$$\begin{aligned}\left(\frac{\partial f}{\partial b}\right)^{\text{ge}} = &\underbrace{\left(\frac{\partial S}{\partial b}\cdot\left(-\frac{\partial f}{\partial S}\right)^{\text{ge}}\right)}_{\text{direct effect}}\\ &+\underbrace{\left[\frac{\partial\epsilon_s^u}{\partial b}\cdot\frac{\partial c^w}{\partial\epsilon_s^u}+\frac{\partial\epsilon_s^c}{\partial b}\cdot\frac{\partial c^w}{\partial\epsilon_s^c}+\frac{\partial c^w}{\partial b}\right]\cdot\left(\frac{\partial f}{\partial c^w}\right)^{\text{ge}}}_{\text{indirect effect}}\end{aligned}$$

$$\left(\frac{\partial\theta}{\partial b}\right)^{\text{ge}} = \underbrace{\frac{\partial S}{\partial b}\cdot\left(\frac{\partial\theta}{\partial S}\right)^{\text{ge}}}_{\text{direct effect}} + \underbrace{\left[\frac{\partial\epsilon_s^u}{\partial b}\cdot\frac{\partial c^w}{\partial\epsilon_s^u}+\frac{\partial\epsilon_s^c}{\partial b}\cdot\frac{\partial c^w}{\partial\epsilon_s^c}+\frac{\partial c^w}{\partial b}\right]\cdot\left(\frac{\partial\theta}{\partial c^w}\right)^{\text{ge}}}_{\text{indirect effect}}$$

3.F.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:

$$(1-\eta)\cdot\big(u'(b)-u'(c^w)\big) = \lambda_\theta\cdot(1-\rho)\cdot\left(\frac{1-n}{n}+\frac{u'(b)}{u'(c^w)}\right)+\lambda_{\epsilon_s^u}\cdot\left(-\frac{u'(b)}{u'(c^w)}\right) + \lambda_c\cdot(1-\rho)\cdot\left(\eta\cdot\frac{1-n}{n}+(1-\eta)\cdot\frac{u'(b)}{u'(c^w)}\right)$$

The Lagrange multiplier for the separation condition of unconstrained firms is equal to

$$-\frac{\lambda_{\epsilon_s^c}}{n^s\cdot u'(c^w)} = \frac{p}{\lambda}\cdot g(\epsilon_s^c)\cdot\frac{1}{\frac{\partial S^c_{\text{stw}}(\epsilon_s^c)}{\partial\epsilon_s^c}}\cdot\left[\left(1+\frac{\lambda_\theta}{n^s\cdot u'(c^w)}\right)\cdot\frac{\lambda}{f}\cdot b + \frac{\lambda_c}{n^s\cdot u'(c^w)}\cdot\left(\eta\cdot\frac{b}{n}+\frac{1}{g(\epsilon_s^c)}\cdot\frac{\lambda}{p}\cdot\frac{\partial WE}{\partial\epsilon_s^c}\right)\right]$$

First, let us find an expression for λ_c:

$$\begin{aligned}\frac{\partial\mathscr{L}}{\partial c^w} = &-\Bigg((1-p)\cdot(1-\rho^u)\cdot\frac{u(c^w)-u(b)}{u'(c^w)}\cdot\frac{u''(c^w)}{u'(c^w)}\\ &+p\cdot(1-\rho^c)\cdot\frac{u^c-u(b)}{u'(c^w)}\cdot\frac{u''(c^w)}{u'(c^w)}\Bigg)\cdot\lambda_\theta\\ &-(1-\eta)\cdot\Bigg((1-p)\cdot(1-\rho^u)\cdot\left(1-(1-\eta)\cdot\frac{u(c^w)-u(b)}{u'(c^w)}\cdot\frac{u''(c^w)}{u'(c^w)}\right)\\ &-p\cdot(1-\rho^c)\cdot(1-\eta)\cdot\frac{u^c-u(b)}{u'(c^w)}\cdot\frac{u''(c^w)}{u'(c^w)}\Bigg)\cdot\lambda_c\\ &+\frac{u(c_{\text{stw}}(\epsilon_s^c))-u(b)}{u'(c^w)}\cdot\frac{u''(c^w)}{u'(c^w)}\cdot\lambda_{\epsilon_s^c} = 0\end{aligned}$$

Insert Lagrange multipliers for separation condition:

$$
\begin{aligned}
&-\Bigg((1-p)\cdot(1-\rho^u)\cdot\frac{u(c^w)-u(b)}{u'(c^w)}\cdot\frac{u''(c^w)}{u'(c^w)}\\
&\quad+p\cdot(1-\rho^c)\cdot\frac{u^c-u(b)}{u'(c^w)}\cdot\frac{u''(c^w)}{u'(c^w)}\Bigg)\cdot\lambda_\theta\\
&-(1-\eta)\cdot(1-p)\cdot(1-\rho^u)\cdot\left(1-(1-\eta)\cdot\frac{u(c^w)-u(b)}{u'(c^w)}\cdot\frac{u''(c^w)}{u'(c^w)}\right)\\
&\quad\cdot p\cdot(1-\rho^c)\cdot(1-\eta)\cdot\frac{u^c-u(b)}{u'(c^w)}\cdot\frac{u''(c^w)}{u'(c^w)}\cdot\lambda_c\\
&+\frac{\partial\epsilon_s^c}{\partial c^w}\cdot\frac{p}{\lambda}\cdot g(\epsilon_s^c)\cdot\left(\left(n^s\cdot u'(c^w)+\lambda_\theta\right)\cdot\left(\frac{\lambda}{f}\cdot b\right)\right.\\
&\quad\left.+\lambda_c\cdot\left(\eta\cdot\frac{b}{n}+\frac{1}{g(\epsilon_s^c)}\cdot\frac{\lambda}{p}\cdot\frac{\partial WE}{\partial\epsilon_s^c}\right)\right)\;=\;0
\end{aligned}
$$

Rearranging gives:

$$
\begin{aligned}
0\;=&\;\lambda_\theta\cdot\frac{\partial S^{\text{total}}}{\partial c^w}\\
&+\lambda_c\cdot\frac{\partial WE^{\text{total}}}{\partial c^w}\\
&+\frac{\partial\epsilon_s^c}{\partial c^w}\cdot\frac{p}{\lambda}\cdot g(\epsilon_s^c)\cdot n^s\cdot u'(c^w)\cdot\frac{\lambda}{f}b
\end{aligned}
$$

with

$$
\begin{aligned}
\frac{\partial S^{\text{total}}}{\partial c^w}\;=&\;\Bigg[p(1-\rho^u)\frac{u(c^w)-u(b)}{u'(c^w)}\frac{u''(c^w)}{u'(c^w)}\\
&-(1-p)(1-\rho^c)\frac{u^c-u(b)}{u'(c^w)}\frac{u''(c^w)}{u'(c^w)}\\
&+\frac{\partial\epsilon_s^c}{\partial\theta}\cdot\frac{p}{\lambda}\cdot g(\epsilon_s^c)\cdot\frac{\lambda}{f}b\Bigg]
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{\partial WE^{\text{total}}}{\partial c^w}\;=&\;\Bigg[-(1-\eta)(1-p\rho^u)\left(1-(1-\eta)\frac{u(c^w)-u(b)}{u'(c^w)}\frac{u''(c^w)}{u'(c^w)}\right)\\
&+(1-(1-p)\rho^c)(1-\eta)\frac{u^c-u(b)}{u'(c^w)}\frac{u''(c^w)}{u'(c^w)}\\
&+\frac{\partial\epsilon_s^c}{\partial c^w}\cdot\frac{p}{\lambda}\cdot g(\epsilon_s^c)\cdot\lambda_c\left(\eta\frac{b}{n}+\frac{1}{g(\epsilon_s^c)}\cdot\frac{\lambda}{p}\cdot\frac{\partial WE}{\partial\epsilon_s^c}\right)\Bigg]
\end{aligned}
$$

Next we can solve for the Lagrange multiplier of the wage equation:

$$\lambda_c = -\frac{1}{\left(\frac{\partial WE}{\partial c^w}\right)^{\text{total}}}\left[\frac{\partial \epsilon_s^c}{\partial c^w}\cdot\frac{p}{\lambda}\cdot g(\epsilon_s^c)\cdot n^s\cdot u'(c^w)\cdot\frac{\lambda}{f}\cdot b\right.$$
$$\left.+\left(-\frac{\partial S^{\text{total}}}{\partial c^w}\right)\cdot\lambda_\theta\right]$$

Now we have everything to calculate λ_θ from the two equations:

$$(1)\quad \left(\chi-\frac{\eta-\gamma}{1-\eta}\right)\cdot\frac{1}{1-\gamma}\cdot\frac{k_v}{q} = \left(1+\frac{\lambda_\theta+\eta\cdot\lambda\cdot\lambda_c}{n^s\cdot u'(c^w)}\right)\cdot\frac{b}{f}$$

$$\begin{aligned}(2)\quad \chi =\ & \frac{1}{1-\eta}\cdot\frac{1}{u\cdot f}\cdot\frac{1}{u'(c^w)}\cdot\left[\gamma-(\lambda\gamma-f\eta)\cdot\left(\frac{1}{\lambda}(1-\rho)\lambda_\theta\right)\right]\\ &+\frac{1}{u}\cdot\frac{1}{u'(c^w)}\cdot\left(\lambda\cdot\frac{\gamma}{f}\cdot u'(c^w)+\left(\frac{1}{1-\eta}\cdot\frac{\lambda\gamma-f}{f}\right)\cdot\eta\right)\cdot\lambda_{\epsilon_s^c}\\ &+\frac{1}{u}\cdot\frac{1}{u'(c^w)}\cdot\lambda_c\cdot\frac{\partial WE}{\partial\theta}\cdot\frac{1}{k_v}\\ &-\frac{1}{u}\cdot\frac{1}{u'(c^w)}\cdot IE_\theta\end{aligned}$$

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:

$$\left[\gamma-(\lambda\gamma-f\eta)\cdot\frac{1}{\lambda}\cdot(1-\rho)\right]\cdot\frac{1}{1-\eta}\cdot\frac{k_v}{f}\cdot\lambda_\theta$$

$$= u'(c^w)\cdot\left(\frac{\partial f}{\partial\theta}\right)\cdot u\cdot\left(\frac{\eta-\gamma}{(1-\gamma)(1-\eta)}\cdot\frac{k_v}{q}+\left(1+\frac{\lambda_\theta+\eta\cdot\lambda\cdot\lambda_c}{n^s\cdot u'(c^w)}\right)\cdot\frac{b}{f}\right)$$

$$+\left(\lambda\cdot\frac{\gamma}{f}\cdot u'(c^w)+\left(\frac{1}{1-\eta}\cdot\frac{\lambda\gamma-f}{f}\right)\cdot\eta\right)\cdot k_v\cdot(-\lambda_{\epsilon_s^c})$$

$$+\frac{\partial WE}{\partial\theta}\cdot(-\lambda_c)\quad+IE_\theta$$

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:

$$\lambda_c = -\frac{u'(c^w)}{\left(\frac{\partial WE}{\partial c^w}\right)^{\text{total}}}\left[\left(\frac{\partial \epsilon_s^u}{\partial c^w}\right)^{\text{ge}} \cdot \frac{\partial n^u}{\partial \epsilon_s^u} \cdot \left(\frac{\lambda}{f}b - \lambda F\right) + \left(\frac{\partial \epsilon_s^c}{\partial c^w}\right)^{\text{ge}} \cdot \frac{\partial n^c}{\partial \epsilon_s^c} \cdot \frac{\lambda}{f}b \right.$$
$$\left. + \left(-\frac{\partial f}{\partial c^w}\right)^{\text{ge}} \cdot u \cdot \left(\frac{\eta-\gamma}{(1-\gamma)(1-\eta)} \cdot \frac{k_v}{q} + \frac{b}{f}\right) + n^c \cdot \left(-\frac{\partial \theta}{\partial c^w}\right)^{\text{ge}} \cdot \hat{IE}_\theta \right]$$

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:

$$(1-n) \cdot \left(u'(b) - u'(c^w)\right) = \lambda_\theta \cdot (1-\rho) \cdot \left(\frac{1-n}{n} + \frac{u'(b)}{u'(c^w)}\right) \cdot \left(-\frac{u'(b)}{u'(c^w)}\right)$$
$$+ \lambda_{\epsilon_s^c} \cdot \left(-\frac{u'(b)}{u'(c^w)}\right)$$
$$+ \lambda_c \cdot (1-\rho) \cdot \left(\eta \cdot \frac{1-n}{n} + (1-\eta) \cdot \frac{u'(b)}{u'(c^w)}\right)$$

Inserting the Lagrange multiplier gives

$$(1-n) \cdot \left(u'(b) - u'(c^w)\right) = u'(c^w) \cdot \left[\left(-\frac{\partial f}{\partial b}\right)^{\text{ge}} \cdot u \cdot \left(\frac{\eta-\gamma}{(1-\gamma)(1-\eta)} \cdot \frac{k_v}{q} + \frac{b}{f}\right) \right.$$
$$+ \left(\frac{\partial \epsilon_s^c}{\partial b}\right)^{\text{ge}} \cdot \frac{\partial n^c}{\partial \epsilon_s^c} \cdot \frac{\lambda}{f}b$$
$$\left. + n^c \cdot \left(-\frac{\partial \theta}{\partial b}\right)^{\text{ge}} \cdot \hat{IE}_\theta \right]$$

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

$$\begin{aligned}(1-n)\cdot\frac{\big(u'(b)-u'(c^w)\big)}{u'(c^w)} &= \left(-\frac{\partial f}{\partial b}\right)^{\text{ge}}\cdot u\cdot L_v \\ &\quad + \left(\frac{\partial\epsilon_s^c}{\partial b}\right)^{\text{ge}}\cdot\frac{\partial n^c}{\partial\epsilon_s^c}\cdot L_s^c + n^c\cdot\left(-\frac{\partial\theta}{\partial b}\right)^{\text{ge}}\cdot\hat{IE}_\theta\end{aligned}$$

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.

3.G 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.G.1 Employment

$$
\begin{aligned}
\frac{\partial \mathscr{L}}{\partial n^s} &= -\lambda_{n^s} + \frac{1-p}{\lambda}(1-\rho^u)\cdot\big(u(c^w) - u'(c^w)\cdot c^w\big) \\
&\quad + \frac{p}{\lambda}(1-\rho^c)\cdot\big(u^c - u'(c^w)\cdot e^c\big) \\
&\quad + \frac{1-p}{\lambda}\cdot(1-\rho^c)\cdot\lambda_{n^u} + \frac{p}{\lambda}(1-\rho^c)\cdot\lambda_{n^c} = 0 \\
\frac{\partial \mathscr{L}}{\partial n^u} &= -\lambda_{n^u} + \lambda_n = 0 \\
\frac{\partial \mathscr{L}}{\partial n^c} &= -\lambda_{n^c} + \lambda_n = 0 \\
\frac{\partial \mathscr{L}}{\partial n} &= -\lambda_n + \lambda_{n^s}\cdot(\lambda - q(\theta)\cdot\theta) \\
&\quad + u'(c^w)\cdot[(z-\Omega) + b + \theta\cdot k_v] - u(b) \\
&\quad + \frac{(1-\rho)}{\lambda}\cdot\frac{b}{n^2}\cdot\lambda_\theta + \eta(1-\rho)\cdot\frac{b}{n^2}\cdot\lambda_c \\
\Leftrightarrow \qquad \frac{\lambda_n}{u'(c^w)} &= (\lambda - q(\theta)\cdot\theta) + [z - \Omega + b + \theta\cdot k_v] - \frac{u(b)}{u'(c^w)} \\
&\quad + \frac{1-\rho}{\lambda}\cdot\frac{b}{n^2}\cdot\frac{\lambda_\theta}{u'(c^w)} + \eta\cdot(1-\rho)\cdot\frac{b}{n^2}\cdot\frac{\lambda_c}{u'(c^w)}
\end{aligned}
$$

3.G.2 Job Creation

Before we start, let us define the Insurance effect:

$$
IE_\theta = n^s\cdot u'(c^w)\cdot\left(1 + \frac{\lambda_\theta}{n^s\cdot u'(c^w)}\right)\cdot \tilde{IE}_\theta
$$

$$
\tilde{IE}_\theta = \frac{p}{\lambda}\left(\int_{\max\{\xi_s^c,\epsilon_{\text{stw}}\}}^{\epsilon_p} \frac{\lambda}{f}\cdot\gamma\cdot\frac{u'(c(\epsilon)) - u'(c^w)}{u'(c^w)}\,dG(\epsilon)\right.
$$

$$
\left.+\int_{\epsilon_s^c}^{\max\{\epsilon_{\text{stw}},\epsilon_s^c\}} \frac{\lambda}{f}\cdot\gamma\cdot\frac{u'(c_{\text{stw}}(\epsilon)) - u'(c^w)}{u'(c^w)}\,dG(\epsilon)\right)\cdot k_v
$$

The FOC for the labor market tightness can be written as:

$$
\begin{aligned}
\frac{\partial \mathscr{L}}{\partial \theta} = & -k_v(1-n)\cdot u'(c^w) + IE_\theta + (1-\gamma)\cdot q(\theta)\cdot(1-n)\cdot\lambda_{n^s} \\
& - \frac{1}{1-\eta}\cdot\gamma\cdot k_v + \frac{\lambda_\theta}{f} + \frac{1}{1-\eta}\cdot\frac{\lambda\gamma - f\eta}{f}\cdot\left(\frac{1}{\lambda}\cdot(1-\rho)\cdot\lambda_\theta - \lambda_{\epsilon_s^u}\right) \\
& - \left(\lambda\cdot\frac{\gamma}{f}\cdot u'(c^w_{\text{stw}}(\epsilon_s^c)) + \left(\frac{1}{1-\eta}\cdot\frac{\lambda\gamma - f}{f}\right)\eta\right)\cdot\lambda_{\epsilon_s^c} \\
& - \left(\lambda\cdot\frac{\gamma}{f}\cdot u'(c^w(\xi_s^c)) + \frac{1}{\lambda}\cdot\left(\frac{1}{1-\eta}\cdot\frac{\lambda\gamma - f}{f}\right)\eta\right)\cdot\lambda_{\xi_s^c} \\
& - \lambda_c\cdot\frac{\partial WE}{\partial\theta}
\end{aligned}
$$

$$
\iff \qquad \frac{\lambda_{n^s}}{u'(c^w)} = \frac{1+\chi}{1-\gamma}\cdot\frac{k_v}{q}
$$

$$
\begin{aligned}
\text{with}\quad \chi = & \frac{1}{(1-n)\cdot u'(c^w)}\cdot\frac{1}{f} \\
\cdot\ & \frac{1}{1-\eta}\cdot\left(\gamma\cdot\lambda_\theta - (\lambda\cdot\gamma - f\cdot\eta)\cdot\left(\frac{1}{\lambda}\cdot(1-\rho)\cdot\lambda_\theta - \lambda_{\epsilon_s^u}\right)\right) \\
& + \frac{1}{(1-n)\cdot u'(c^w)}\left(\lambda\cdot\frac{\gamma}{f}u'(c^w_{\text{stw}}(\epsilon_s^c)) + \left(\frac{1}{1-\eta}\cdot\frac{\lambda\gamma - f}{f}\right)\eta\right)\cdot\lambda_{\epsilon_s^c} \\
& + \frac{1}{(1-n)\cdot u'(c^w)}\left(\lambda\cdot\frac{\gamma}{f}u'(c^w(\xi_s^c)) + \left(\frac{1}{1-\eta}\cdot\frac{\lambda\gamma - f}{f}\right)\eta\right)\cdot\lambda_{\xi_s^c} \\
& + \frac{\lambda_c}{(1-n)\cdot u'(c^w)}\cdot\frac{\partial WE}{\partial\theta}/k_v \\
& - \frac{1}{(1-n)\cdot u'(c^w)}\cdot IE_\theta/k_v
\end{aligned}
$$

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.G.3 Lagrange Multipliers for Separation Condition

Optimal separation condition unconstrained firms

$$\begin{aligned}
& -\frac{\partial\mathscr{L}}{\partial\epsilon_s^u} \\
& = \frac{n}{1-\rho}\cdot(1-p)\cdot g(\epsilon_s^u)\cdot u'(c^w)\cdot\left(z(\epsilon_s^u)-\Omega(\epsilon_s^u)+\frac{u(c^w)}{u'(c^w)}-c^w-(z+\Omega)\right) \\
& + \lambda_\theta\cdot\frac{1-p}{\lambda}\cdot g(\epsilon_s^u)\cdot\left[z(\epsilon_s^u)-\Omega(\epsilon_s^u)+\frac{u(c^w)-u(b)}{u'(c^w)}-c^w\right. \\
& \left.-\frac{1-n}{n}\cdot b+\frac{\lambda-\eta f}{1-\eta}\cdot\frac{k_v}{q}\right] \\
& + \lambda_c\cdot\frac{\partial WE}{\partial\epsilon_s^u}+\frac{1-p}{\lambda}\cdot g(\epsilon_s^u)\cdot n^s\cdot\lambda_{n^u}+\lambda_{\epsilon_s^u}\cdot\frac{\partial S^u_{\text{stw}}(\epsilon_s^u)}{\partial\epsilon_s^u} \stackrel{!}{=} 0
\end{aligned}$$

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}

$$\begin{aligned}
\iff\quad & z(\epsilon_s^u)-\Omega(\epsilon_s^u)+\frac{u(c^w)-u(b)}{u'(c^w)}-c^w+b \\
& +\frac{\lambda-\gamma f+\chi(\lambda-f)}{1-\gamma}\cdot\frac{k_v}{q} \\
& -\frac{\lambda_\theta}{n^s\cdot u'(c^w)}\big(\tau_{\text{stw}}(\bar{h}-h_{\text{stw}}(\epsilon_s^u))-b\big) \\
& +\lambda\cdot\frac{\lambda_c}{n^s\cdot u'(c^w)}\left(\eta\cdot\frac{b}{n}+\frac{\lambda}{1-p}\cdot\frac{1}{g(\epsilon_s^u)}\cdot\frac{\partial WE}{\partial\epsilon_s^u}\right) \\
& +\frac{\lambda_{\epsilon_s^u}}{n^s\cdot u'(c^w)}\cdot\frac{\lambda}{1-p}\cdot\frac{1}{g(\epsilon_s^u)}\cdot\frac{\partial S_{\text{stw}}^u(\epsilon_s^u)}{\partial\epsilon_s^u}=0
\end{aligned}$$

Subtract decentralized separation condition:

$$\begin{aligned}
\tau_{\text{stw}}(\bar{h}-h_{\text{stw}}(\epsilon_s^u)) \;=\; & b+\big(\lambda-\theta q(\theta)\big)\cdot\left(\chi-\frac{\eta-\gamma}{1-\eta}\right)\cdot\frac{1}{1-\gamma}\cdot\frac{k_v}{q}\\
&+\frac{\lambda_{\epsilon_s^u}}{n^s\cdot u'(c^w)}\cdot\frac{\lambda}{1-p}\cdot\frac{1}{g(\epsilon_s^u)}\cdot\frac{\partial S_{\text{stw}}^u(\epsilon_s^u)}{\partial\epsilon_s^u}\\
&-\frac{\lambda_\theta\cdot\theta}{n^s\cdot u'(c^w)}\cdot\big(\tau_{\text{stw}}(\bar{h}-h_{\text{stw}}(\epsilon_s^u))-b\big)\\
&+\frac{\lambda_c}{n^s\cdot u'(c^w)}\left(\eta\cdot\frac{b}{n}+\frac{\lambda}{1-p}\cdot\frac{1}{g(\epsilon_s^u)}\cdot\frac{\partial WE}{\partial\epsilon_s^u}\right)
\end{aligned}$$

Hence

$$\begin{aligned}
0 \;=\; & \left(1+\frac{\lambda_\theta}{u'(c^w)\cdot n^s}\right)\\
&\cdot\left(b+(\lambda-f)\cdot\frac{(1-\rho/\lambda)}{(\rho+(1-\rho)\cdot f/\lambda)}\cdot\frac{b}{n}-\tau_{\text{stw}}(\bar{h}-h_{\text{stw}}(\epsilon_s^u))\right)\\
&+\frac{\lambda}{1-p}\cdot\frac{1}{g(\epsilon_s^u)}\cdot\frac{\lambda_{\epsilon_s^u}}{n^s\cdot u'(c^w)}\cdot\frac{\partial S_{\text{stw}}^u(\epsilon_s^u)}{\partial\epsilon_s^u}\\
&+\frac{\lambda_c}{n^s\cdot u'(c^w)}\left(\eta\cdot\frac{b}{n}+\frac{\lambda}{1-p}\cdot\frac{1}{g(\epsilon_s^u)}\cdot\frac{\partial WE}{\partial\epsilon_s^u}\right)\\
\iff\quad 0 \;=\; & \left(1+\frac{\lambda_\theta}{u'(c^w)\cdot n^s}\right)\cdot\left(\frac{\lambda}{f}\cdot b-\tau_{\text{stw}}(\bar{h}-h_{\text{stw}}(\epsilon_s^u))\right)\\
&+\frac{\lambda}{1-p}\cdot\frac{1}{g(\epsilon_s^u)}\cdot\frac{\lambda_{\epsilon_s^u}}{n^s\cdot u'(c^w)}\cdot\frac{\partial S_{\text{stw}}^u(\epsilon_s^u)}{\partial\epsilon_s^u}\\
&+\frac{\lambda_c}{n^s\cdot u'(c^w)}\left(\eta\cdot\frac{b}{n}+\frac{\lambda}{1-p}\cdot\frac{1}{g(\epsilon_s^u)}\cdot\frac{\partial WE}{\partial\epsilon_s^u}\right)
\end{aligned}$$

And finally:

$$\begin{aligned}
-\frac{\lambda_{\epsilon_s^u}}{n^s\cdot u'(c^w)} \;=\; & \frac{(1-p)\cdot g(\epsilon_s^u)}{\lambda}\cdot\frac{1}{\frac{\partial S_{\text{stw}}^u(\epsilon_s^u)}{\partial\epsilon_s^u}}\cdot\left(\left(1+\frac{\lambda_\theta}{n^s\cdot u'(c^w)}\right)\right.\\
&\cdot\left(\frac{\lambda}{f}\cdot b-\tau_{\text{stw}}\cdot\big(\bar{h}-h_{\text{stw}}(\epsilon_s^u)\big)\right)\\
&\left.+\frac{\lambda_c}{n^s\cdot u'(c^w)}\cdot\left(\eta\cdot\frac{b}{n}+\frac{\lambda}{1-p}\cdot\frac{1}{g(\epsilon_s^u)}\cdot\frac{\partial WE}{\partial\epsilon_s^u}\right)\right)
\end{aligned}$$

Optimal separation condition constrained firms

$$
\begin{aligned}
-\frac{\partial \mathscr{L}}{\partial \epsilon_s^c} = \Bigg[& \frac{n}{1-\rho} \cdot p \cdot g(\epsilon_s^c) \cdot u'(c^w) \Big(z(\epsilon_s^c) - \Omega(\epsilon_s^c) \\
& + \frac{u(c_{\text{stw}}(\epsilon_s^c))}{u'(c^w)} - c_{\text{stw}}(\epsilon_s^c) - (z+\Omega) \Big) \\
& + \lambda_\theta \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^c) \Bigg[z(\epsilon_s^c) - \Omega(\epsilon_s^c) + \frac{u(c_{\text{stw}}(\epsilon_s^c)) - u(b)}{u'(c^w)} - c^w(\epsilon_s^c) \\
& - \frac{1-n}{n} \cdot b + \frac{\lambda - \eta f}{1-\eta} \cdot \frac{k_v}{q} \Bigg] \\
& - \lambda_c \cdot \frac{\partial WE}{\partial \epsilon_s^c} + \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot n^s \cdot \lambda_{n^c} \Bigg] \cdot \mathbb{1}(\epsilon_{\text{stw}} \geq \epsilon_s^c) + \lambda^c_{\epsilon_s^c} \cdot \frac{\partial S^u_{\text{stw}}(\epsilon_s^c)}{\partial \epsilon_s^c} = 0
\end{aligned}
$$

Subtract decentralized separation condition (constrained firms):

$$
\begin{aligned}
-\frac{\partial \mathscr{L}}{\partial \epsilon_s^c} = \Bigg[& \frac{n}{1-\rho} \cdot p \cdot g(\epsilon_s^c) \Bigg(z_{\text{stw}}(\epsilon_s^c) - \Omega(\epsilon_s^c) + \frac{u(c_{\text{stw}}(\epsilon_s^c)) - u(c^w)}{u'(c^w)} - (z+\Omega) \Bigg) \\
& + \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot n^s \cdot \frac{\lambda_{n^c}}{u'(c^w)} \\
& - \frac{\lambda_\theta}{u'(c^w)} \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^c) \Bigg[\tau_{\text{stw}} \cdot (\bar{h} - h_{\text{stw}}(\epsilon_s^c)) + \frac{1-n}{n} \cdot b \Bigg] \\
& + \frac{\lambda_c}{u'(c^w)} \cdot \frac{\partial WE}{\partial \epsilon_s^c} \Bigg] \cdot \mathbb{1}(\epsilon_{\text{stw}} \geq \epsilon_s^c) + \frac{\lambda^c_{\epsilon_s^c}}{u'(c^w)} \cdot \frac{\partial S^c_{\text{stw}}(\epsilon_s^c)}{\partial \epsilon_s^c} = 0
\end{aligned}
$$

Insert λ_{n^c} and λ_n

$$
\begin{aligned}
\iff -\frac{\partial \mathscr{L}}{\partial \epsilon_s^c} = \Bigg[& \frac{p}{\lambda} \cdot n^s g(\epsilon_s^c) \Bigg(z_{\text{stw}}(\epsilon_s^c) - \Omega(\epsilon_s^c) + \frac{u(c_{\text{stw}}(\epsilon_s^c))}{u'(c^w)} - u(c^w) - (z+\Omega) \Bigg) \\
& + \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot n^s \cdot \Bigg((\lambda - q(\theta) \cdot \theta) \cdot \frac{\lambda_{n^s}}{u'(c^w)} \\
& + [z - \Omega + b + \theta \cdot c] - \frac{u(b)}{u'(c^w)} \\
& \quad + \frac{1-\rho}{\lambda} \cdot \frac{b}{n^2} \cdot \lambda_\theta \cdot \frac{1}{u'(c^w)} + \eta \cdot \frac{b}{n^2} \cdot \frac{\lambda_c}{u'(c^w)} \Bigg) \Bigg] \\
& - \frac{\lambda_\theta}{u'(c^w)} \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot \Bigg[\tau_{\text{stw}} \cdot (\bar{h} - h_{\text{stw}}(\epsilon_s^c)) + \frac{1-n}{n} \cdot b \Bigg] \\
& + \frac{\lambda_c}{u'(c^w)} \cdot \frac{\partial WE}{\partial \epsilon_s^c} \cdot \mathbb{1}(\epsilon_{\text{stw}} \geq \epsilon_s^c) + \frac{\lambda^c_{\epsilon_s^c}}{u'(c^w)} \cdot \frac{\partial S^c_{\text{stw}}(\epsilon_s^c)}{\partial \epsilon_s^c} = 0
\end{aligned}
$$

Insert λ_{n^s}

$$
\begin{aligned}
-\frac{\partial \mathscr{L}}{\partial \epsilon_s^c} \;=\; \Bigg[& z_{\text{stw}}(\epsilon_s^c) - \Omega(\epsilon_s^c) + \frac{u(c_{\text{stw}}(\epsilon_s^c))}{u'(c^w)} - c_{\text{stw}}(\epsilon_s^c) + b \\
& + \frac{\lambda - \gamma \cdot f + \chi \cdot (\lambda - f)}{1-\gamma} \cdot \frac{k_v}{q} \\
& - \frac{\lambda_\theta}{n^s \cdot u'(c^w)} \left[\tau_{\text{stw}} \cdot (\bar{h} - h_{\text{stw}}(\epsilon_s^c)) - b \right] \\
& + \frac{\lambda_c}{n^s \cdot u'(c^w)} \left(\eta \cdot \frac{b}{n} + \frac{1}{g(\epsilon_s^c)} \cdot \frac{\lambda}{p} \cdot \frac{\partial WE}{\partial \epsilon_s^c} \right) \Bigg] \cdot \mathbb{1}(\epsilon_{\text{stw}} \;\geq\; \xi_s^c) \\
& + \frac{\lambda}{p} \cdot \frac{1}{n^s \cdot g(\epsilon_s^c)} \cdot \frac{\lambda^c_{\epsilon_s^c}}{u'(c^w)} \cdot \frac{\partial S^c_{\text{stw}}(\epsilon_s^c)}{\partial \epsilon_s^c} \;=\; 0
\end{aligned}
$$

And finally:

$$
\begin{aligned}
-\frac{\lambda_{\epsilon_s^c}}{n^s \cdot u'(c^w)} \;=\; & \frac{p}{\lambda} \cdot \frac{g(\epsilon_s^c)}{\frac{\partial s^c_{\text{stw}}(\epsilon_s^c)}{\partial \epsilon_s^c}} \cdot \Bigg(\left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right) \cdot \left(\frac{\lambda}{f} \cdot b - \tau_{\text{stw}} \cdot (\bar{h} - h_{\text{stw}}(\epsilon_s^c)) \right) \\
& + \frac{\lambda_c}{n^s \cdot u'(c^w)} \left(\eta \cdot \frac{b}{n} + \frac{1}{g(\epsilon_s^c)} \cdot \frac{\lambda}{p} \cdot \frac{\partial WE}{\partial \epsilon_s^c} \right) \Bigg) \cdot \mathbb{1}(\epsilon_{\text{stw}} \;\geq\; \epsilon_s^c)
\end{aligned}
$$

Analogously, it can be derived that

$$
\begin{aligned}
-\frac{\lambda_{\xi_s^c}}{n^s \cdot u'(c^w)} \;=\; & \frac{p}{\lambda} \cdot \frac{g(\xi_s^c)}{\frac{\partial s^c_{\text{stw}}(\xi_s^c)}{\partial \xi^c}} \cdot \Bigg(\left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right) \cdot \left(\frac{\lambda}{f} \cdot b \right) \\
& + \frac{\lambda \cdot \lambda_c}{n^s \cdot u'(c^w)} \left(\eta \cdot \frac{b}{n} + \frac{1}{g(\xi_s^c)} \cdot \frac{\lambda}{p} \cdot \frac{\partial WE}{\partial \xi^c} \right) \Bigg) \cdot \mathbb{1}(\epsilon_{\text{stw}} \;\leq\; \xi_s^c)
\end{aligned}
$$

3.G.4 Optimal STW Benefits

$$
\begin{aligned}
\frac{\partial \mathscr{L}}{\partial \tau_{\text{stw}}} \;=\; & \frac{p}{\lambda} \cdot n^s \cdot \int_{\epsilon_s^c}^{\max\{\epsilon_{\text{stw}}, \epsilon_s^c\}} \left(u'(c_{\text{stw}}(\epsilon)) - u'(c^w) \right) \left(\bar{h} - h_{\text{stw}}(\epsilon) \right) dG(\epsilon) \\
& + \lambda_\theta \cdot \frac{p}{\lambda} \cdot \int_{\epsilon_s^c}^{\max\{\epsilon_{\text{stw}}, \epsilon_s^c\}} \left(\frac{u'(c_{\text{stw}}(\epsilon)) - u'(c^w)}{u'(c^w)} \right) \left(\bar{h} - h_{\text{stw}}(\epsilon) \right) dG(\epsilon) \\
& - \frac{n^s \cdot (1-\rho)}{\lambda} \cdot u'(c^w) \cdot \frac{\partial \Omega}{\partial \tau_{\text{stw}}} - \lambda_\theta \cdot \frac{(1-\rho)}{\lambda} \cdot \frac{\partial \Omega}{\partial \tau_{\text{stw}}} - \lambda_c \cdot \frac{\partial WE}{\partial \tau_{\text{stw}}} \\
& - \lambda_{\epsilon_s^u} \cdot \frac{\partial S^u_{\text{stw}}(\epsilon_s^u)}{\partial \tau_{\text{stw}}} - \lambda_{\epsilon_s^c} \cdot \frac{\partial S^c_{\text{stw}}(\epsilon_s^c)}{\partial \tau_{\text{stw}}} \;=\; 0
\end{aligned}
$$

Insert Lagrange multiplier separation condition:

$$\begin{aligned}
&\left(1+\frac{\lambda_\theta}{n^s\cdot u'(c^w)}\right)\cdot\left(\frac{p}{\lambda}\int_{\epsilon_s^c}^{\max\{\epsilon_s^c,\epsilon_{\text{stw}}\}}\frac{u'(c_{\text{stw}}(\epsilon))-u'(c^w)}{u'(c^w)}\cdot(\bar{h}-h_{\text{stw}}(\epsilon))\,dG(\epsilon)\right.\\
&\qquad\left.-(1-\rho)\cdot\frac{\partial\Omega}{\partial\tau_{\text{stw}}}\right)-\frac{\lambda_c}{n^s\cdot u'(c^w)}\cdot\frac{\partial WE}{\partial\tau_{\text{stw}}}\\
&\quad+\frac{(1-p)}{\lambda}\cdot g(\epsilon_s^u)\cdot\frac{\partial\epsilon_s^u}{\partial\tau_{\text{stw}}}\cdot\left(\left(1+\frac{\lambda_\theta}{n^s\cdot u'(c^w)}\right)\cdot\left(\frac{\lambda}{f}\cdot b-\tau_{\text{stw}}\cdot(\bar{h}-h_{\text{stw}}(\epsilon_s^u))\right)\right.\\
&\qquad\left.+\frac{\lambda_c}{n^s\cdot u'(c^w)}\cdot\left(\eta\cdot\frac{b}{n}+\frac{1}{g(\epsilon_s^u)}\cdot\frac{\lambda}{1-p}\cdot\frac{\partial WE}{\partial\epsilon_s^u}\right)\right)\\
&\quad+\frac{p}{\lambda}\cdot g(\epsilon_s^c)\cdot\frac{\partial\epsilon_s^c}{\partial\tau_{\text{stw}}}\cdot\left(\left(1+\frac{\lambda_\theta}{n^s\cdot u'(c^w)}\right)\cdot\left(\frac{\lambda}{f}\cdot b-\tau_{\text{stw}}\cdot(\bar{h}-h_{\text{stw}}(\epsilon_s^c))\right)\right.\\
&\qquad\left.+\frac{\lambda_c}{n^s\cdot u'(c^w)}\cdot\left(\eta\cdot\frac{b}{n}+\frac{1}{g(\epsilon_s^c)}\cdot\frac{\lambda}{p}\cdot\frac{\partial WE}{\partial\epsilon_s^c}\right)\right)=0
\end{aligned}$$

Finally:

$$\begin{aligned}
\bar{\tau}_{\text{stw}} = &\underbrace{\frac{\lambda}{f}b}_{\text{Fiscal Ext.}}\\
&+\underbrace{\frac{\mathbb{1}\{\epsilon_{\text{stw}}\geq\epsilon_s^c\}}{MMS}\cdot n^s\cdot\frac{p}{\lambda}\int_{\epsilon_s^c}^{\epsilon_{\text{stw}}}\left(\frac{u'(c_{\text{stw}}(\epsilon))-u'(c^w)}{u'(c^w)}\right)(\bar{h}-h_{\text{stw}}(\epsilon))\,dG(\epsilon)}_{\text{Insurance}}\\
&\underbrace{-\frac{n^s}{MMS}\cdot\frac{1}{\lambda}\cdot(1-\rho)\frac{\partial\,\Omega}{\partial\,\tau_{\text{stw}}}}_{\text{Distortion}}+\underbrace{BE_{\text{stw},3}}_{\text{Bargaining Effect}}
\end{aligned}$$

$$\begin{aligned}
\bar{\tau}_{\text{stw}} = &\frac{MMS^u}{MMS}\tau_{\text{stw}}(\bar{h}-h_{\text{stw}}(\epsilon_s^u))\\
&+\frac{MMS^c}{MMS}\tau_{\text{stw}}(\bar{h}-h_{\text{stw}}(\epsilon_s^c))
\end{aligned}$$

where

$$MMS^u = \frac{(1-p)}{\lambda}\cdot g(\epsilon_s^u)\cdot\frac{\partial\epsilon_s^u}{\partial\tau_{\text{stw}}}\cdot n^s,\quad MMS^c = \mathbb{1}(\epsilon_{\text{stw}}\geq\epsilon_s^c)\cdot\frac{p}{\lambda}\cdot g(\epsilon_s^c)\cdot\frac{\partial\epsilon_s^c}{\partial\tau_{\text{stw}}}\cdot n^s,$$

$$MMS = MMS^c + MMS^u$$

and

$$
\begin{aligned}
BE_{\text{stw},3} \;=\; & \frac{1}{\left(1+\frac{\lambda_\theta}{n^s u'(c^w)}\right)} \\
& \times\Bigg[-\frac{\lambda_c}{\varphi\, n^s u'(c^w)}\,\frac{\partial WE}{\partial \tau_{\text{stw}}}+\frac{\varphi^u}{\varphi}\,\frac{\lambda_c}{n^s u'(c^w)}\Big(\eta\,\frac{b}{n}+\frac{1}{g(\epsilon_s^u)}\,\frac{\lambda}{1-p}\,\frac{\partial WE}{\partial \epsilon_s^u}\Big) \\
& \quad+\frac{\varphi^c}{\varphi}\,\frac{\lambda_c}{n^s u'(c^w)}\Big(\eta\,\frac{b}{n}+\frac{1}{g(\epsilon_s^c)}\,\frac{\lambda}{p}\,\frac{\partial WE}{\partial \epsilon_s^c}\Big)\Bigg]
\end{aligned}
$$

3.G.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 $\xi_s^c > \xi_s^u$.

Candidate I:

$$
\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \epsilon_{\text{stw}}} \;=\; & g(\epsilon_{\text{stw}}) \cdot \frac{p}{\lambda} \cdot n^s \cdot \Big(u(c_{\text{stw}}(\epsilon_{\text{stw}})) - u(c(\epsilon_{\text{stw}})) \\
& - u'(c^w) \cdot [c_{\text{stw}}(\epsilon_{\text{stw}}) - c(\epsilon_{\text{stw}})]\Big) \\
& - n \cdot \frac{\partial \Omega}{\partial \epsilon_{\text{stw}}} \\
& + \lambda_\theta \cdot \frac{p}{\lambda} \cdot \Big(u(c_{\text{stw}}(\epsilon_{\text{stw}})) - u(c(\epsilon_{\text{stw}})) \\
& - u'(c^w) \cdot [c_{\text{stw}}(\epsilon_{\text{stw}}) - c(\epsilon_{\text{stw}})]\Big) \\
& - \lambda_\theta \cdot \frac{1-\rho}{\lambda} \cdot \frac{\partial \Omega}{\partial \epsilon_{\text{stw}}} \\
& - \lambda_c \cdot \frac{\partial WE}{\partial \epsilon_{\text{stw}}} \overset{!}{\geq} 0
\end{aligned}
$$

$$
\begin{aligned}
\Leftrightarrow \quad & \left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right) \cdot \left(\frac{1-\rho}{\lambda} \cdot n^s \cdot \frac{\partial \Omega}{\partial \epsilon_{\text{stw}}} + BE\right) \\
& \overset{!}{\geq} \left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right) \cdot \frac{p}{\lambda} \cdot g(\epsilon_{\text{stw}}) \cdot n^s \\
& \cdot \left(\frac{\dfrac{u(c_{\text{stw}}(\epsilon_{\text{stw}})) - u(c(\epsilon_{\text{stw}}))}{c_{\text{stw}}(\epsilon_{\text{stw}}) - c(\epsilon_{\text{stw}})} - u'(c^w)}{u'(c^w)} \right) \\
& \cdot [c_{\text{stw}}(\epsilon_{\text{stw}}) - c(\epsilon_{\text{stw}})]
\end{aligned}
$$

$$
\begin{aligned}
\Leftrightarrow \quad & \left(\frac{1-\rho}{\lambda} \cdot n^s \cdot \frac{\partial \Omega}{\partial \epsilon_{\text{stw}}} + BE\right) \overset{!}{\geq} \frac{p}{\lambda} \cdot g(\epsilon_{\text{stw}}) \cdot n^s \\
& \cdot \left(\frac{\dfrac{u(c_{\text{stw}}(\epsilon_{\text{stw}})) - u(c(\epsilon_{\text{stw}}))}{c_{\text{stw}}(\epsilon_{\text{stw}}) - c(\epsilon_{\text{stw}})} - u'(c^w)}{u'(c^w)} \right) \\
& \cdot [c_{\text{stw}}(\epsilon_{\text{stw}}) - c(\epsilon_{\text{stw}})]
\end{aligned}
$$

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 participation threshold is determined by balancing the additional cost of hours distortions against the benefit of providing extra insurance.

Candidate II:
FOC for the eligibility threshold:

$$
\begin{aligned}
\frac{\partial \mathscr{L}}{\partial \epsilon_{\text{stw}}} &= -n^s \cdot u'(c^w) \cdot \frac{1-p}{\lambda} \cdot (1-\rho^u) \cdot \frac{\partial \Omega_u}{\partial \epsilon_{\text{stw}}} u'(c^w) \\
&\quad - (\lambda_\theta - \eta \cdot \lambda_c) \cdot \frac{1-p}{\lambda} \cdot (1-\rho^u) \cdot \frac{\partial \Omega_u}{\partial \epsilon_{\text{stw}}} \\
&= -\left(n^s \cdot u'(c^w) + \lambda_\theta + \eta \cdot \lambda_c\right) \cdot \frac{1-p}{\lambda} \cdot (1-\rho^u) \cdot \frac{\partial \Omega_u}{\partial \epsilon_{\text{stw}}}
\end{aligned}
$$

Under reasonable parameter values (assuming that the Ramsey planner never wants to use short-time work to impede vacancy posting by destroying the efficiency of the match, capured by the Lagrange multipliers λ_θ and λ_c), we get:

$$
\frac{\partial \mathscr{L}}{\partial \epsilon_{\text{stw}}} < 0 \quad \text{thus it is optimal to set} \quad \epsilon_{\text{stw}} = \epsilon_s^u.
$$

Candidate III:
FOC for the eligibility threshold:

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \epsilon_{\mathrm{stw}}} = & -n^u \cdot \frac{\partial \Omega_u}{\partial \epsilon_{\mathrm{stw}}} \cdot u'(c^w) \\
& + \frac{n}{1-\rho} \cdot p \cdot g(\epsilon_{\mathrm{stw}}) \cdot u'(c^w) \\
& \quad \cdot \left(z_{\mathrm{stw}}(\epsilon_{\mathrm{stw}}) + \frac{u(c_{\mathrm{stw}}(\epsilon_{\mathrm{stw}}))}{u'(c^w)} - c_{\mathrm{stw}}(\epsilon_{\mathrm{stw}}) - (z + \Omega) \right) \\
& + \lambda_\theta \cdot \frac{(1-p)}{\lambda} \cdot (1-\rho^u) \cdot \frac{\partial \Omega_u}{\partial \epsilon_{\mathrm{stw}}} \\
& + \frac{p}{\lambda} \cdot g(\epsilon_{\mathrm{stw}}) \\
& \quad \cdot \Bigg[z(\epsilon_{\mathrm{stw}}) + \frac{u(c_{\mathrm{stw}}(\epsilon_{\mathrm{stw}})) - u(b)}{u'(c^w)} \\
& \qquad - c_{\mathrm{stw}}(\epsilon_{\mathrm{stw}}) - \frac{1-n}{n} \cdot b + \frac{\lambda - \eta \cdot f}{1-\eta} \cdot \frac{k_v}{q} \Bigg] \\
& - \lambda_c \cdot \frac{\partial WE}{\partial \epsilon_{\mathrm{stw}}} + \frac{p}{\lambda} \cdot g(\epsilon_{\mathrm{stw}}) \cdot n^s \cdot \lambda_{n^c} \overset{!}{>} 0.
\end{aligned}$$

Insert n^u

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \epsilon_{\mathrm{stw}}} = & -n^s \cdot (1-\rho^u) \cdot \left(\frac{1-p}{\lambda} \right) \cdot \frac{\partial \Omega_u}{\partial \epsilon_{\mathrm{stw}}} \cdot u'(c^w) \\
& + n^s \cdot \frac{p}{\lambda} \cdot g(\epsilon_{\mathrm{stw}}) \cdot u'(c^w) \cdot \left(z_{\mathrm{stw}}(\epsilon_{\mathrm{stw}}) + \frac{u(c_{\mathrm{stw}}(\epsilon_{\mathrm{stw}}))}{u'(c^w)} \right) \\
& - n^s \cdot \frac{p}{\lambda} \cdot g(\epsilon_{\mathrm{stw}}) \cdot u'(c^w) \cdot \left(c_{\mathrm{stw}}(\epsilon_{\mathrm{stw}}) + z + \Omega \right) \\
& + \lambda_\theta \cdot \frac{g(\epsilon_{\mathrm{stw}})}{\lambda} \cdot (1-p)(1-\rho^u) \cdot \frac{\partial \Omega_u}{\partial \epsilon_{\mathrm{stw}}} \\
& + \frac{p}{\lambda} \cdot g(\epsilon_{\mathrm{stw}}) \cdot \Bigg(z_{\mathrm{stw}}(\epsilon_{\mathrm{stw}}) + \frac{u(c_{\mathrm{stw}}(\epsilon_{\mathrm{stw}})) - u(b)}{u'(c^w)} - c_{\mathrm{stw}}(\epsilon_{\mathrm{stw}}) - \frac{1-n}{n} \cdot b \\
& \qquad + \frac{\lambda - \eta \cdot f}{1-\eta} \cdot \left(\frac{k_v}{q} \right) \Bigg) \\
& - \lambda_c \cdot \frac{\partial WE}{\partial \epsilon_{\mathrm{stw}}} + \frac{p}{\lambda} \cdot g(\epsilon_{\mathrm{stw}}) \cdot n^s \cdot \lambda_{n^c} \overset{!}{>} 0
\end{aligned}$$

$$
\begin{aligned}
\Leftrightarrow & -(n^s \cdot u'(c^w) + \lambda_\theta) \cdot (1-\rho^u) \cdot \left(\frac{1-p}{\lambda}\right) \cdot \frac{\partial \Omega_u}{\partial \epsilon_{\text{stw}}} \\
& + n^s \cdot \frac{p}{\lambda} \cdot g(\epsilon_{\text{stw}}) \cdot u'(c^w) \\
& \cdot \left(z_{\text{stw}}(\epsilon_{\text{stw}}) + \frac{u(c_{\text{stw}}(\epsilon_{\text{stw}}))}{u'(c^w)} - c_{\text{stw}}(\epsilon_{\text{stw}}) - (z+\Omega) + \frac{\lambda_{n^c}}{u'(c^w)} \right) \\
& + \frac{p}{\lambda} \cdot g(\epsilon_{\text{stw}}) \cdot \Bigg[z_{\text{stw}}(\epsilon_{\text{stw}}) + \frac{u(c_{\text{stw}}(\epsilon_{\text{stw}})) - u(b)}{u'(c^w)} - c_{\text{stw}}(\epsilon_{\text{stw}}) \\
& \qquad - \frac{1-n}{n} \cdot b + \frac{\lambda - \eta \cdot f}{1-\eta} \cdot \left(\frac{k_v}{q}\right) \Bigg] \\
& - \lambda_c \cdot \frac{\partial WE}{\partial \epsilon_{\text{stw}}} \overset{!}{>} 0
\end{aligned}
$$

Insert λ_{n^c} and subtract decentralized separation condition for constrained firms:

$$
\begin{aligned}
\Leftrightarrow \quad & -\left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right) \cdot \frac{(1-p)}{\lambda} \cdot (1-\rho^u) \cdot \frac{\partial \Omega_u}{\partial \epsilon_{\text{stw}}} \cdot \frac{1}{g(\epsilon_{\text{stw}})} \\
& + \frac{p}{\lambda} \cdot \Bigg[b + z_{\text{stw}}(\epsilon_{\text{stw}}) + \frac{u(c_{\text{stw}}(\epsilon_{\text{stw}})) - u(b)}{u'(c^w)} - c_{\text{stw}}(\epsilon_{\text{stw}}) \\
& + \frac{\lambda - \eta \cdot f + \chi \cdot (\lambda - f)}{1-\eta} \cdot \left(\frac{k_v}{q}\right) \Bigg] \\
& - \frac{\lambda_\theta}{n^s \cdot u'(c^w)} \cdot \frac{p}{\lambda} \cdot [\tau_{\text{stw}} \cdot (h - h_{\text{stw}}(\epsilon_{\text{stw}})) - b] \\
& - \frac{\lambda_c}{n^s \cdot u'(c^w)} \cdot \left(\frac{\partial WE}{\partial \epsilon_{\text{stw}}} \cdot \frac{1}{g(\epsilon_{\text{stw}})} \cdot \frac{\partial S_{\text{stw}}(\epsilon_{\text{stw}})}{\partial \epsilon_{\text{stw}}} + \frac{p}{\lambda} \cdot \eta \cdot \frac{b}{n} \right) \overset{!}{>} 0
\end{aligned}
$$

Subtract decentralized separation condition:

$$
\begin{aligned}
\Leftrightarrow \quad & -\left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right) \cdot \frac{(1-p)}{\lambda} \cdot (1-\rho^u) \cdot \frac{\partial \Omega_u}{\partial \epsilon_{\text{stw}}} \cdot \frac{1}{g(\epsilon_{\text{stw}})} \\
& + \frac{p}{\lambda} \cdot \Bigg[z_{\text{stw}}(\epsilon_{\text{stw}}) - z_{\text{stw}}(\epsilon_s) + \frac{u(c_{\text{stw}}(\epsilon_{\text{stw}})) - u(c_{\text{stw}}(\epsilon_s^c))}{u'(c^w)} \\
& \quad - \tau_{\text{stw}} \cdot (h - h_{\text{stw}}(\epsilon_s^c)) + b + (1-\rho) \cdot (\lambda - f) \cdot \left(\chi - \frac{\eta - \gamma}{1-\eta}\right) \cdot \frac{1}{1-\gamma} \cdot \frac{k_v}{q} \Bigg] \\
& + \frac{\lambda_\theta}{n^s \cdot u'(c^w)} \cdot \frac{p}{\lambda} \cdot \Bigg[z_{\text{stw}}(\epsilon_{\text{stw}}) - z_{\text{stw}}(\epsilon_s^c) + \frac{u(c_{\text{stw}}(\epsilon_{\text{stw}})) - u(c_{\text{stw}}(\epsilon_s^c))}{u'(c^w)} \\
& - \tau_{\text{stw}} \cdot (h - h_{\text{stw}}(\epsilon_s^c)) + b \Bigg] \\
& - \frac{\lambda_c}{n^s \cdot u'(c^w)} \cdot \left(\frac{\partial WE}{\partial \epsilon_{\text{stw}}} \cdot \frac{1}{g(\epsilon_{\text{stw}})} \cdot \frac{\partial S_{\text{stw}}(\epsilon_{\text{stw}})}{\partial \epsilon_{\text{stw}}} + \frac{p}{\lambda} \cdot \eta \cdot \frac{b}{n} \right) \overset{!}{>} 0
\end{aligned}
$$

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:

$$
\begin{aligned}
BE = {} & g(\epsilon_{\text{stw}}) \cdot \frac{\lambda_c}{u'(c^w)} \cdot \left(\frac{\partial WE}{\partial \epsilon_{\text{stw}}} \cdot \frac{1}{g(\epsilon_{\text{stw}})} \cdot \frac{\partial S_{\text{stw}}(\epsilon_{\text{stw}})}{\partial \epsilon_{\text{stw}}} + \eta \cdot \frac{b}{n} \right) \\
& \Bigg/ \left(1 + \frac{\lambda_\theta}{n^s \cdot u'(c^w)} \right)
\end{aligned}
$$

The Bargaining effect from the proposition is equal $BE/MMS_{\epsilon_{\text{stw}}}$. This concludes the proof of Proposition 4.

3.G.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:

$$
\begin{aligned}
-\frac{\lambda_{\epsilon_s^u}}{n^s\cdot u'(c^w)} = {} & \frac{1-p}{\lambda}\cdot\frac{g(\epsilon_s^u)}{\frac{\partial S_{\text{stw}}^u(\epsilon_s^u)}{\partial\epsilon_s^u}}\\
& \cdot\left(\left(1+\frac{\lambda_\theta}{n^s\cdot u'(c^w)}\right)\cdot\left(\frac{\lambda}{f}\cdot b-\tau_{\text{stw}}\cdot\left(\bar{h}-h_{\text{stw}}(\epsilon_s^u)\right)\right)\right.\\
& \left.+\frac{\lambda_c}{n^s\cdot u'(c^w)}\cdot\left(\eta\cdot\frac{b}{n}+\frac{1}{g(\epsilon_s^u)}\cdot\frac{\lambda}{1-p}\cdot\frac{\partial WE}{\partial\epsilon_s^u}\right)\right)
\end{aligned}
$$

$$
\begin{aligned}
-\frac{\lambda_{\xi_s^c}}{n^s\cdot u'(c^w)} = {} & \frac{p}{\lambda}\cdot\frac{g(\xi_s^c)}{\frac{\partial S_{\text{stw}}^c(\xi_s^c)}{\partial\xi_s^c}}\cdot\left(\left(1+\frac{\lambda_\theta}{n^s\cdot u'(c^w)}\right)\cdot\frac{\lambda}{f}\cdot b\right.\\
& \left.+\frac{\lambda_c}{n^s\cdot u'(c^w)}\left(\eta\cdot\frac{b}{n}+\frac{1}{g(\xi_s^c)}\cdot\frac{\lambda}{p}\cdot\frac{\partial WE}{\partial\xi_s^c}\right)\right)\cdot\mathbb{1}(\epsilon_{\text{stw}}\le\xi_s^c)
\end{aligned}
$$

$$
\begin{aligned}
-\frac{\lambda_{\epsilon_s^c}}{n^s\cdot u'(c^w)} = {} & \frac{p}{\lambda}\cdot\frac{g(\epsilon_s^c)}{\frac{\partial S_{\text{stw}}^c(\epsilon_s^c)}{\partial\epsilon_s^c}}\cdot\left(\left(1+\frac{\lambda_\theta}{n^s\cdot u'(c^w)}\right)\cdot\left(\frac{\lambda}{f}\cdot b-\tau_{\text{stw}}\cdot\left(\bar{h}-h_{\text{stw}}(\epsilon_s^c)\right)\right)\right.\\
& \left.+\frac{\lambda_c}{n^s\cdot u'(c^w)}\cdot\left(\eta\cdot\frac{b}{n}+\frac{1}{g(\epsilon_s^c)}\cdot\frac{\lambda}{p}\cdot\frac{\partial WE}{\partial\epsilon_s^c}\right)\right)\cdot\mathbb{1}(\epsilon_{\text{stw}}\ge\epsilon_s^c)
\end{aligned}
$$

First, let us find an expression for λ_c:

$$
\begin{aligned}
\frac{\partial \mathscr{L}}{\partial c^w} = & -\Bigg[(1-p)\cdot(1-\rho^u)\cdot\frac{u(c^w)-u(b)}{u'(c^w)}\cdot\frac{u''(c^w)}{u'(c^w)} \\
& \quad + p\cdot(1-\rho^c)\cdot\frac{u^c-u(b)}{u'(c^w)}\cdot\frac{u''(c^w)}{u'(c^w)}\Bigg]\cdot\lambda_\theta \\
& -(1-\eta)\cdot(1-p)\cdot(1-\rho^u)\cdot\left(1-(1-\eta)\cdot\frac{u(c^w)-u(b)}{u'(c^w)}\cdot\frac{u''(c^w)}{u'(c^w)}\right) \\
& -(1-p)\cdot(1-\rho^c)\cdot(1-\eta)\cdot\frac{u^c-u(b)}{u'(c^w)}\cdot\frac{u''(c^w)}{u'(c^w)}\cdot\lambda_c \\
& +\frac{u(c^w)-u(b)}{u'(c^w)}\cdot\frac{u''(c^w)}{u'(c^w)}\cdot\lambda_{\epsilon_s^u} \\
& +\frac{u(c(\xi_s^c))-u(b)}{u'(c^w)}\cdot\frac{u''(c^w)}{u'(c^w)}\cdot\lambda_{\xi_s^c} \\
& +\frac{u(c_{\text{stw}}(\epsilon_s^c))-u(b)}{u'(c^w)}\cdot\frac{u''(c^w)}{u'(c^w)}\cdot\lambda_{\epsilon_s^c} = 0
\end{aligned}
$$

Insert Lagrange multipliers for separation conditions

$$
\begin{aligned}
& -\left(p(1-\rho^u)\cdot\frac{u(c^w)-u(b)}{u'(c^w)}\cdot\frac{u''(c^w)}{u'(c^w)}+(1-p)(1-\rho^c)\cdot\frac{u^c-u(b)}{u'(c^w)}\cdot\frac{u''(c^w)}{u'(c^w)}\right)\lambda_\theta \\
& -(1-\eta)\cdot\left((1-p)\cdot(1-\rho^u)\cdot\left(1-(1-\eta)\cdot\frac{u(c^w)-u(b)}{u'(c^w)}\cdot\frac{u''(c^w)}{u'(c^w)}\right)\right. \\
& \left. -p\cdot(1-\rho^c)\cdot(1-\eta)\cdot\frac{u^c-u(b)}{u'(c^w)}\cdot\frac{u''(c^w)}{u'(c^w)}\right)\lambda_c \\
& +\frac{\partial\epsilon_s^u}{\partial c^w}\cdot\frac{1-p}{\lambda}\cdot g(\epsilon_s^u)\left(n^s\cdot u'(c^w)+\lambda_\theta\right)\cdot\left(\frac{\lambda}{f}\cdot b-\tau_{\text{stw}}\cdot(\bar{h}-h_{\text{stw}}(\epsilon_s^u))\right) \\
& +\lambda_c\cdot\left(\eta\cdot\frac{b}{n}+\frac{1}{g(\epsilon_s^u)}\cdot\frac{\lambda}{1-p}\cdot\frac{\partial WE}{\partial\epsilon_s^u}\right) \\
& +\frac{\partial\xi_s^c}{\partial c^w}\cdot\frac{p}{\lambda}\cdot g(\xi_s^c)\left(n^s\cdot u'(c^w)+\lambda_\theta\right)\cdot\left(\frac{\lambda}{f}\cdot b+\lambda_c\cdot\left(\eta\cdot\frac{b}{n}-\frac{1}{g(\xi_s^c)}\cdot\frac{\lambda}{p}\cdot\frac{\partial WE}{\partial\xi_s^c}\right)\right) \\
& \cdot\mathbb{1}(\epsilon_{\text{stw}}\leq\xi_s^c) \\
& +\frac{\partial\xi_s^c}{\partial c^w}\cdot\frac{p}{\lambda}\cdot g(\xi_s^c)\left(n^s\cdot u'(c^w)+\lambda_\theta\right)\cdot\left(\frac{\lambda}{f}\cdot b-\tau_{\text{stw}}\cdot(\bar{h}-h_{\text{stw}}(\xi_s^c))\right) \\
& +\lambda_c\cdot\left(\eta\cdot\frac{b}{n}+\frac{1}{g(\xi_s^c)}\cdot\frac{\lambda}{p}\cdot\frac{\partial WE}{\partial\xi_s^c}\right)\cdot\mathbb{1}(\epsilon_{\text{stw}}\geq\epsilon_s^c) = 0
\end{aligned}
$$

Rearranging yields

$$
\begin{aligned}
0 \ = &\lambda_\theta \cdot \frac{\partial S^{\text{total}}}{\partial c^w} + \lambda_c \cdot \frac{\partial WE^{\text{total}}}{\partial c^w} \\
&+ \frac{\partial \epsilon_s^u}{\partial c^w} \cdot \frac{1-p}{\lambda} \cdot g(\epsilon_s^u) \cdot n^s \cdot u'(c^w) \cdot \left(\frac{\lambda}{f} \cdot b - \tau_{\text{stw}} \cdot (\bar{h} - h_{\text{stw}}(\epsilon_s^u)) \right) \\
&+ \frac{\partial \xi_s^c}{\partial c^w} \cdot \frac{p}{\lambda} \cdot g(\xi_s^c) \cdot n^s \cdot u'(c^w) \cdot \frac{\lambda}{f} \cdot b \cdot \mathbb{1}(\epsilon_{\text{stw}} \ \leq \ \xi_s^c) \\
&+ \frac{\partial \epsilon_s^c}{\partial c^w} \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot n^s \cdot u'(c^w) \cdot \left(\frac{\lambda}{f} \cdot b - \tau_{\text{stw}} \cdot (\bar{h} - h_{\text{stw}}(\epsilon_s^c)) \right) \cdot \mathbb{1}(\epsilon_{\text{stw}} \ \geq \ \epsilon_s^c)
\end{aligned}
$$

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:

$$
\begin{aligned}
\lambda_c \ = \ &\frac{-1}{\left(\frac{\partial WE}{\partial c^w}\right)^{\text{total}}} \Bigg[\frac{\partial \epsilon_s^u}{\partial c^w} \cdot \frac{1-p}{\lambda} \cdot g(\epsilon_s^u) \cdot n^s \cdot u'(c^w) \left(\frac{\lambda}{f} \cdot b - \tau_{\text{stw}} \cdot \left(\bar{h} - h_{\text{stw}}(\epsilon_s^u) \right) \right) \\
&+ \frac{\partial \xi_s^c}{\partial c^w} \cdot \frac{p}{\lambda} \cdot g(\xi_s^c) \cdot n^s \cdot u'(c^w) \left(\frac{\lambda}{f} \cdot b \right) \cdot \mathbb{1}(\epsilon_{\text{stw}} \ \leq \ \xi_s^c) \\
&+ \frac{\partial \epsilon_s^c}{\partial c^w} \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot n^s \cdot u'(c^w) \left(\frac{\lambda}{f} \cdot b - \tau_{\text{stw}} \cdot \left(\bar{h} - h_{\text{stw}}(\epsilon_s^c) \right) \right) \cdot \mathbb{1}(\epsilon_{\text{stw}} \ \geq \ \epsilon_s^c) \\
&+ \left(-\frac{\partial S_{\text{total}}}{\partial c^w} \right) \cdot \lambda_\theta \Bigg]
\end{aligned}
$$

Insert λ_c into the Lagrange multipliers for the separation conditions. To do that, let us rewrite the Lagrange multipliers of the separation conditions:

$$\begin{aligned}
\lambda_{\epsilon_s^u} = & -\frac{1}{\frac{\partial S^u_{\text{stw}}(\epsilon_s^u)}{\partial \epsilon_s^u}}\left(n^s \cdot u'(c^w) \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^u)\left(1+\frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right)\right. \\
& \cdot \left(\frac{\lambda}{f} \cdot b - \tau_{\text{stw}} \cdot (\bar{h} - h_{\text{stw}}(\epsilon_s^u))\right) \\
& \left. + \lambda_c\left(\frac{\partial n}{\partial \epsilon_s^u} \cdot \left(-\frac{\partial WE}{\partial n}\right) + \frac{\partial WE}{\partial \epsilon_s^u}\right)\right) \\
\lambda_{\xi_s^c} = & \frac{\mathbb{1}(\epsilon_{\text{stw}} \leq \xi_s^c)}{\frac{\partial S^c_{\text{stw}}(\xi_s^c)}{\partial \epsilon_s^c}}\left(n^s \cdot u'(c^w) \cdot \frac{p}{\lambda} \cdot g(\xi_s^c)\left(1+\frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right) \cdot \frac{\lambda}{f} \cdot b\right. \\
& \left. + \lambda_c\left(\frac{\partial n}{\partial \xi_s^c} \cdot \left(-\frac{\partial WE}{\partial n}\right) + \frac{\partial WE}{\partial \xi_s^c}\right)\right. \\
\lambda_{\epsilon_s^c} = & -\frac{\mathbb{1}(\epsilon_{\text{stw}} \geq \epsilon_s^c)}{\frac{\partial S^c_{\text{stw}}(\epsilon_s^c)}{\partial \epsilon_s^c}}\left(n^s \cdot u'(c^w) \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^c)\left(1+\frac{\lambda_\theta}{n^s \cdot u'(c^w)}\right)\right. \\
& \cdot \left(\frac{\lambda}{f} \cdot b - \tau_{\text{stw}} \cdot (\bar{h} - h_{\text{stw}}(\epsilon_s^c))\right) \\
& \left. + \lambda_c\left(\frac{\partial n}{\partial \epsilon_s^c} \cdot \left(-\frac{\partial WE}{\partial n}\right) + \frac{\partial WE}{\partial \epsilon_s^c}\right)\right)
\end{aligned}$$

Inserting λ_c gives:

$$\begin{aligned}
& \lambda_{\epsilon_s^u} = \\
& -\frac{1}{\frac{\partial S^u_{\text{stw}}(\epsilon_s^u)}{\partial \epsilon_s^u}}\left[n^s u'(c^w) \cdot \frac{1-p}{\lambda} \cdot g(\epsilon_s^u)\left(1+\frac{\lambda_\theta}{n^s u'(c^w)}\right)\left(\frac{\lambda}{f} b - \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon_s^u))\right)\right. \\
& + \left(\frac{\partial n}{\partial \epsilon_s^u}\left(-\frac{\partial c^w}{\partial n}\right)^{\text{total}} + \left(\frac{\partial c^w}{\partial \epsilon_s^u}\right)^{\text{total}}\right) \\
& \cdot \left(\frac{\partial \epsilon_s^u}{\partial c^w} \cdot \frac{1-p}{\lambda} \cdot g(\epsilon_s^u) \cdot n^s u'(c^w)\left(\frac{\lambda}{f} b - \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon_s^u))\right)\right. \\
& + \frac{\partial \xi_s^c}{\partial c^w} \cdot \frac{p}{\lambda} \cdot g(\xi_s^c) \cdot n^s u'(c^w) \cdot \frac{\lambda}{f} b \cdot \mathbb{1}(\epsilon_{\text{stw}} \leq \xi_s^c) \\
& + \frac{\partial \epsilon_s^c}{\partial c^w} \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot n^s u'(c^w)\left(\frac{\lambda}{f} b - \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon_s^c))\right) \cdot \mathbb{1}(\epsilon_{\text{stw}} \geq \epsilon_s^c) \\
& \left.\left. + \left(-\frac{\partial S^{\text{total}}}{\partial c^w}\right) \cdot \lambda_\theta\right)\right]
\end{aligned}$$

$$
\begin{aligned}
\lambda_{\xi_s^c} =& \\
& \frac{\mathbb{1}(\epsilon_{\text{stw}} \le \xi_s^c)}{\frac{\partial S^c_{\text{stw}}(\xi_s^c)}{\partial \xi_s^c}} \Bigg[-n^s u'(c^w) \cdot \frac{p}{\lambda} \cdot g(\xi_s^c) \left(1 + \frac{\lambda_\theta}{n^s u'(c^w)}\right) \left(\frac{\lambda}{f} b - \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\xi_s^c))\right) \\
& + \left(\frac{\partial n}{\partial \xi_s^c} \left(-\frac{\partial c^w}{\partial n}\right)^{\text{total}} + \left(\frac{\partial c^w}{\partial \xi_s^c}\right)^{\text{total}} \right) \\
& \cdot \Bigg(\frac{\partial \epsilon_s^u}{\partial c^w} \cdot \frac{1-p}{\lambda} \cdot g(\epsilon_s^u) \cdot n^s u'(c^w) \left(\frac{\lambda}{f} b - \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon_s^u))\right) \\
& + \frac{\partial \xi_s^c}{\partial c^w} \cdot \frac{p}{\lambda} \cdot g(\xi_s^c) \cdot n^s u'(c^w) \cdot \frac{\lambda}{f} b \cdot \mathbb{1}(\epsilon_{\text{stw}} \le \xi_s^c) \\
& + \frac{\partial \epsilon_s^c}{\partial c^w} \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot n^s u'(c^w) \left(\frac{\lambda}{f} b - \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon_s^c))\right) \cdot \mathbb{1}(\epsilon_{\text{stw}} \ge \epsilon_s^c) \\
& + \left(-\frac{\partial S^{\text{total}}}{\partial c^w}\right) \cdot \lambda_\theta \Bigg) \Bigg]
\end{aligned}
$$

$$
\begin{aligned}
\lambda_{\epsilon_s^c} =& \\
& - \frac{\mathbb{1}(\epsilon_{\text{stw}} \ge \epsilon_s^c)}{\frac{\partial S^c_{\text{stw}}(\epsilon_s^c)}{\partial \epsilon_s^c}} \Bigg[n^s u'(c^w) \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^c) \left(1 + \frac{\lambda_\theta}{n^s u'(c^w)}\right) \left(\frac{\lambda}{f} b - \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon_s^c))\right) \\
& + \left(\frac{\partial n}{\partial \epsilon_s^c} \left(-\frac{\partial c^w}{\partial n}\right)^{\text{total}} + \left(\frac{\partial c^w}{\partial \epsilon_s^c}\right)^{\text{total}} \right) \\
& \cdot \Bigg(\frac{\partial \epsilon_s^u}{\partial c^w} \cdot \frac{1-p}{\lambda} \cdot g(\epsilon_s^u) \cdot n^s u'(c^w) \left(\frac{\lambda}{f} b - \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon_s^u))\right) \\
& + \frac{\partial \xi_s^c}{\partial c^w} \cdot \frac{p}{\lambda} \cdot g(\xi_s^c) \cdot n^s u'(c^w) \cdot \frac{\lambda}{f} b \cdot \mathbb{1}(\epsilon_{\text{stw}} \le \xi_s^c) \\
& + \frac{\partial \epsilon_s^c}{\partial c^w} \cdot \frac{p}{\lambda} \cdot g(\epsilon_s^c) \cdot n^s u'(c^w) \left(\frac{\lambda}{f} b - \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon_s^c))\right) \cdot \mathbb{1}(\epsilon_{\text{stw}} \ge \epsilon_s^c) \\
& + \left(-\frac{\partial S^{\text{total}}}{\partial c^w}\right) \cdot \lambda_\theta \Bigg) \Bigg]
\end{aligned}
$$

Now we have everything to calculate λ_θ from two equations:

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

$$
\begin{aligned}
(2)\quad \chi \;=\; & \frac{1}{1-\eta}\cdot\frac{1}{u\cdot f}\cdot\frac{1}{u'(c^w)}\cdot\left[\gamma-(\lambda\gamma-f\eta)\cdot\left(\frac{1}{\lambda}(1-\rho)\lambda_\theta-\lambda_{\epsilon_s^u}\right)\right] \\
& +\frac{1}{u}\cdot\frac{1}{u'(c^w)}\cdot\left(\lambda\cdot\frac{\gamma}{f}\cdot u'(c^w)+\left(\frac{1}{1-\eta}\cdot\frac{\lambda\gamma-f}{f}\right)\cdot\eta\right)\cdot\lambda_{\xi_s^c} \\
& +\frac{1}{u}\cdot\frac{1}{u'(c^w)}\cdot\left(\lambda\cdot\frac{\gamma}{f}\cdot u'(c^w)+\left(\frac{1}{1-\eta}\cdot\frac{\lambda\gamma-f}{f}\right)\cdot\eta\right)\cdot\lambda_{\epsilon_s^c} \\
& +\frac{1}{u}\cdot\frac{1}{u'(c^w)}\cdot\lambda_c\cdot\frac{\partial WE}{\partial\theta}\cdot\frac{1}{k_v} \\
& -\frac{1}{u}\cdot\frac{1}{u'(c^w)}\cdot IE_\theta
\end{aligned}
$$

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 $\lambda_{\epsilon_s^u}, \lambda_{\epsilon_s^c}, \lambda_{\xi_s^c}, IE_\theta$:

$$
\begin{aligned}
&\left[\gamma - (\lambda\gamma - f\cdot\eta)\cdot\frac{1}{\lambda}\cdot(1-\rho)\right]\cdot\frac{1}{1-\eta}\cdot\frac{k_v}{f}\cdot\lambda_\theta \\
&= u'(c^w)\cdot\left(\frac{\partial f}{\partial\theta}\right)\cdot u\cdot\left(\frac{\eta-\gamma}{(1-\gamma)(1-\eta)}\cdot\frac{k_v}{q}+\left(1+\frac{\lambda_\theta}{n^s u'(c^w)}\right)\cdot\frac{b}{f}\right) \\
&\quad + u'(c^w)\cdot n^s\cdot\left(-\frac{\partial\epsilon_s^u}{\partial\theta}\right)^{\text{total}}\cdot g(\epsilon_s^u) \\
&\quad \cdot\frac{1-p}{\lambda}\left(1+\frac{\lambda_\theta}{n^s u'(c^w)}\right)\left(\frac{\lambda}{f}b-\tau_{\text{stw}}(\bar{h}-h_{\text{stw}}(\epsilon_s^u))\right) \\
&\quad + u'(c^w)\cdot n^s\cdot\left(-\frac{\partial\xi_s^c}{\partial\theta}\right)^{\text{total}}\cdot g(\xi_s^c)\cdot\frac{p}{\lambda}\left(1+\frac{\lambda_\theta}{n^s u'(c^w)}\right)\cdot\frac{\lambda}{f}b\cdot\mathbb{1}(\epsilon_{\text{stw}}\ \le\ \xi_s^c) \\
&\quad + u'(c^w)\cdot n^s\cdot\left(-\frac{\partial\epsilon_s^c}{\partial\theta}\right)^{\text{total}}\cdot g(\epsilon_s^c) \\
&\quad \cdot\frac{p}{\lambda}\left(1+\frac{\lambda_\theta}{n^s u'(c^w)}\right)\left(\frac{\lambda}{f}b-\tau_{\text{stw}}(\bar{h}-h_{\text{stw}}(\epsilon_s^c))\right)\cdot\mathbb{1}(\epsilon_{\text{stw}}\ \ge\ \epsilon_s^c) \\
&\quad + \lambda_\theta\cdot\left(-\frac{\partial c^w}{\partial\theta}\right)^{\text{total}}\cdot\left(-\frac{\partial S}{\partial c^w}\right)^{\text{total}} \\
&\quad + n^s\cdot u'(c^w)\cdot\left(1+\frac{\lambda_\theta}{n^s u'(c^w)}\right)\cdot\tilde{IE}_\theta
\end{aligned}
$$

Denote:

$$
\begin{aligned}
\left(-\frac{\partial\epsilon_s^u}{\partial\theta}\right)^{\text{total}} &= \left(-\frac{\partial\epsilon_s^u}{\partial\theta}\right)+\left(\frac{\partial c^w}{\partial\theta}\right)^{\text{total}}\cdot\left(-\frac{\partial\epsilon_s^u}{\partial c^w}\right) \\
&\quad + \left(-\frac{\partial\epsilon_s^u}{\partial\theta}\right)\cdot\left[\left(\frac{\partial n}{\partial\epsilon_s^u}\right)\cdot\left(-\frac{\partial c^w}{\partial n}\right)^{\text{total}}+\left(\frac{\partial c^w}{\partial\epsilon_s^u}\right)^{\text{total}}\right]\cdot\left(\frac{\partial\epsilon_s^u}{\partial c^w}\right) \\
&\quad + f'(\theta)\cdot\frac{\partial n}{\partial f}\cdot\frac{\partial c^w}{\partial n}\cdot\frac{\partial\epsilon_s^u}{\partial c^w} \\
\left(-\frac{\partial\xi_s^c}{\partial\theta}\right)^{\text{total}} &= \Bigg[\left(-\frac{\partial\xi_s^c}{\partial\theta}\right)+\left(\frac{\partial c^w}{\partial\theta}\right)^{\text{total}}\cdot\left(-\frac{\partial\xi_s^c}{\partial c^w}\right) \\
&\quad + \left(-\frac{\partial\xi_s^c}{\partial\theta}\right)\cdot\left[\left(\frac{\partial n}{\partial\xi_s^c}\right)\cdot\left(-\frac{\partial c^w}{\partial n}\right)^{\text{total}}+\left(\frac{\partial c^w}{\partial\xi_s^c}\right)^{\text{total}}\right]\cdot\left(\frac{\partial\xi_s^c}{\partial c^w}\right) \\
&\quad + f'(\theta)\cdot\frac{\partial n}{\partial f}\cdot\frac{\partial c^w}{\partial n}\cdot\frac{\partial\xi_s^c}{\partial c^w}\Bigg]\cdot\mathbb{1}(\epsilon_{\text{stw}}\ \le\ \xi_s^c)
\end{aligned}
$$

$$
\begin{aligned}
\left(-\frac{\partial \epsilon_s^c}{\partial \theta}\right)^{\text{total}} &= \Bigg[\left(-\frac{\partial \epsilon_s^c}{\partial \theta}\right) + \left(\frac{\partial c^w}{\partial \theta}\right)^{\text{total}} \cdot \left(-\frac{\partial \epsilon_s^c}{\partial c^w}\right) \\
&\quad + \left(-\frac{\partial \epsilon_s^c}{\partial \theta}\right) \cdot \left[\left(\frac{\partial n}{\partial \epsilon_s^c}\right) \cdot \left(-\frac{\partial c^w}{\partial n}\right)^{\text{total}} + \left(\frac{\partial c^w}{\partial \epsilon_s^c}\right)^{\text{total}}\right] \cdot \frac{\partial \epsilon_s^c}{\partial c^w} \\
&\quad + f'(\theta) \cdot \frac{\partial n}{\partial f} \cdot \frac{\partial c^w}{\partial n} \cdot \frac{\partial \epsilon_s^c}{\partial c^w}\Bigg] \cdot \mathbb{1}(\epsilon_{\text{stw}} \geq \epsilon_s^c)
\end{aligned}
$$

$$
\begin{aligned}
\left(-\frac{\partial c^w}{\partial \theta}\right)^{\text{total,2}} &= \left(-\frac{\partial c^w}{\partial \theta}\right)^{\text{total}} \\
&\quad + \left(-\frac{\partial \epsilon_s^u}{\partial \theta}\right) \cdot \left[\left(\frac{\partial n}{\partial \epsilon_s^u}\right) \cdot \left(-\frac{\partial c^w}{\partial n}\right)^{\text{total}} + \left(\frac{\partial c^w}{\partial \epsilon_s^u}\right)^{\text{total}}\right] \\
&\quad + \left(-\frac{\partial \xi_s^c}{\partial \theta}\right) \cdot \left[\left(\frac{\partial n}{\partial \xi_s^c}\right) \cdot \left(-\frac{\partial c^w}{\partial n}\right)^{\text{total}} + \left(\frac{\partial c^w}{\partial \xi_s^c}\right)^{\text{total}}\right] \\
&\quad + \left(-\frac{\partial \epsilon_s^c}{\partial \theta}\right) \cdot \left[\left(\frac{\partial n}{\partial \epsilon_s^c}\right) \cdot \left(-\frac{\partial c^w}{\partial n}\right)^{\text{total}} + \left(\frac{\partial c^w}{\partial \epsilon_s^c}\right)^{\text{total}}\right] \\
&\quad + f'(\theta) \cdot \frac{\partial n}{\partial f} \cdot \frac{\partial c^w}{\partial n}
\end{aligned}
$$

This allows us to calculate the Lagrange multiplier for the job-creation condition:

$$
\begin{aligned}
&M \cdot \lambda_\theta = \\
&u'(c^w) \cdot \frac{\partial f}{\partial \theta} \cdot u \cdot \left(\frac{\eta - \gamma}{(1-\gamma)(1-\eta)} \cdot \frac{k_v}{q} + \left(1 + \frac{\lambda_\theta}{n^s u'(c^w)}\right) \cdot \frac{b}{f}\right) \\
&\quad + u'(c^w) \cdot n^s \cdot \left(\frac{\partial \epsilon_s^u}{\partial \theta}\right)^{\text{total}} \cdot g(\epsilon_s^u) \cdot \frac{1-p}{\lambda}\left(\frac{\lambda}{f} b - \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon_s^u))\right) \\
&\quad + u'(c^w) \cdot n^s \cdot \left(-\frac{\partial \xi_s^c}{\partial \theta}\right)^{\text{total}} \cdot g(\xi_s^c) \cdot \frac{p}{\lambda} \cdot \frac{\lambda}{f} b \cdot \mathbb{1}(\epsilon_{\text{stw}} \leq \xi_s^c) \\
&\quad + u'(c^w) \cdot n^s \cdot \left(-\frac{\partial \epsilon_s^c}{\partial \theta}\right)^{\text{total}} \cdot g(\epsilon_s^c) \\
&\quad \cdot \frac{p}{\lambda}\left(\frac{\lambda}{f} b - \tau_{\text{stw}}(\bar{h} - h_{\text{stw}}(\epsilon_s^c))\right) \cdot \mathbb{1}(\epsilon_{\text{stw}} \geq \epsilon_s^c) \\
&\quad + u'(c^w) \cdot n^s \cdot \tilde{IE}_\theta
\end{aligned}
$$

with

$$
\begin{aligned}
M \;=\; & \left[\gamma - (\lambda\gamma - f\cdot\eta)\cdot\frac{1}{\lambda}\cdot(1-\rho)\right]\cdot\frac{1}{1-\eta}\cdot\frac{k_v}{f} \\
& + \left(-\frac{\partial f}{\partial\theta}\cdot u\cdot\frac{b}{f}\right) \\
& + \left(\frac{\partial \epsilon_s^u}{\partial\theta}\right)^{\text{total}}\cdot g(\epsilon_s^u)\cdot\frac{1-p}{\lambda}\left(\frac{\lambda}{f}b - \tau_{\text{stw}}(\bar h - h_{\text{stw}}(\epsilon_s^u))\right) \\
& + \left(\frac{\partial \xi_s^c}{\partial\theta}\right)^{\text{total}}\cdot g(\xi_s^c)\cdot\frac{p}{\lambda}\cdot\frac{\lambda}{f}b\cdot\mathbb{1}(\epsilon_{\text{stw}}\;\le\;\xi_s^c) \\
& + \left(\frac{\partial \epsilon_s^c}{\partial\theta}\right)^{\text{total}}\cdot g(\epsilon_s^c)\cdot\frac{p}{\lambda}\left(\frac{\lambda}{f}b - \tau_{\text{stw}}(\bar h - h_{\text{stw}}(\epsilon_s^c))\right)\cdot\mathbb{1}(\epsilon_{\text{stw}}\;\ge\;\epsilon_s^c) \\
& + \left(\frac{\partial c^w}{\partial\theta}\right)^{\text{total}}\cdot\left(-\frac{\partial S}{\partial c^w}\right)^{\text{total}} \\
& + \tilde{IE}_\theta
\end{aligned}
$$

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:

$$
\begin{aligned}
& M\cdot\lambda_\theta \;= \\
& u'(c^w)\cdot\left(\frac{\partial f}{\partial S}\right)^{\text{ge}}\cdot u\cdot\left(\frac{\eta-\gamma}{(1-\gamma)(1-\eta)}\cdot\frac{k_v}{q}+\frac{\lambda}{f}\cdot b\right) \\
& \quad + u'(c^w)\cdot n^s\cdot\left(\left(-\frac{\partial\epsilon_s^u}{\partial S}\right)^{\text{ge}}\cdot g(\epsilon_s^u)\cdot\frac{1-p}{\lambda}\left(\frac{\lambda}{f}\cdot b - \tau_{\text{stw}}\cdot(\bar h - h_{\text{stw}}(\epsilon_s^u))\right)\right) \\
& \quad + u'(c^w)\cdot n^s\cdot\left(\left(-\frac{\partial\xi_s^c}{\partial S}\right)^{\text{ge}}\cdot g(\xi_s^c)\cdot\frac{p}{\lambda}\cdot\frac{\lambda}{f}\cdot b\cdot\mathbb{1}(\epsilon_{\text{stw}}\;\le\;\xi_s^c)\right) \\
& \quad + u'(c^w)\cdot n^s \\
& \quad \cdot\left(\left(-\frac{\partial\epsilon_s^c}{\partial S}\right)^{\text{ge}}\cdot g(\epsilon_s^c)\cdot\frac{p}{\lambda}\cdot\left(\frac{\lambda}{f}\cdot b - \tau_{\text{stw}}\cdot(\bar h - h_{\text{stw}}(\epsilon_s^c))\right)\cdot\mathbb{1}(\epsilon_{\text{stw}}\;\ge\;\epsilon_s^c)\right) \\
& \quad + u'(c^w)\cdot n^s\cdot\left(\frac{\partial\theta}{\partial S}\right)^{\text{ge}}\cdot\tilde{IE}_\theta
\end{aligned}
$$

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

$$
\begin{aligned}
\left(\frac{\partial \epsilon_s^u}{\partial b}\right)^{\text{ge}}\cdot\frac{\partial n}{\partial \epsilon_s^u} =\; & \underbrace{\left(\frac{\partial \epsilon_s^u}{\partial b}\right)^{\text{ge}}\left(\frac{\partial n}{\partial \epsilon_s^u}\right)}_{\text{direct effects}}\\
& +\underbrace{\left[\frac{\partial S}{\partial b}\left(\frac{\partial \epsilon_s^u}{\partial S}\right)^{\text{ge}}+\frac{\partial \xi_s^c}{\partial b}\left(\frac{\partial c^w}{\partial \xi_s^c}\right)^{\text{ge}}+\frac{\partial \epsilon_s^c}{\partial b}\left(\frac{\partial c^w}{\partial \epsilon_s^c}\right)^{\text{ge}}+\frac{\partial c^w}{\partial b}\left(\frac{\partial \epsilon_s^u}{\partial c^w}\right)^{\text{ge}}\right]\cdot\frac{\partial n}{\partial \epsilon_s^u}}_{\text{indirect effect}}
\end{aligned}
$$

$$\left(\frac{\partial \xi_s^c}{\partial b}\right)^{\text{ge}} \cdot \frac{\partial n}{\partial \xi_s^c} = \underbrace{\left(\frac{\partial \xi_s^c}{\partial b}\right)^{\text{ge}} \left(\frac{\partial n}{\partial \xi_s^c}\right)}_{\text{direct effects}}$$

$$+ \underbrace{\left[\frac{\partial S}{\partial b}\left(\frac{\partial \xi_s^c}{\partial S}\right)^{\text{ge}} + \frac{\partial \epsilon_s^u}{\partial b}\left(\frac{\partial c^w}{\partial \epsilon_s^u}\right)^{\text{ge}} + \frac{\partial \epsilon_s^c}{\partial b}\left(\frac{\partial c^w}{\partial \epsilon_s^c}\right)^{\text{ge}} + \frac{\partial c^w}{\partial b}\left(\frac{\partial \xi_s^c}{\partial c^w}\right)\right] \cdot \frac{\partial n}{\partial \xi_s^c}}_{\text{indirect effect}}$$

$$\left(\frac{\partial \epsilon_s^c}{\partial b}\right)^{\text{ge}} \cdot \frac{\partial n}{\partial \epsilon_s^c} = \underbrace{\left(\frac{\partial \epsilon_s^c}{\partial b}\right)^{\text{ge}} \left(\frac{\partial n}{\partial \epsilon_s^c}\right)}_{\text{direct effects}}$$

$$+ \underbrace{\left[\frac{\partial S}{\partial b}\left(\frac{\partial \epsilon_s^c}{\partial S}\right)^{\text{ge}} + \frac{\partial \epsilon_s^u}{\partial b}\left(\frac{\partial c^w}{\partial \epsilon_s^u}\right)^{\text{ge}} + \frac{\partial \xi_s^c}{\partial b}\left(\frac{\partial c^w}{\partial \xi_s^c}\right)^{\text{ge}} + \frac{\partial c^w}{\partial b}\left(\frac{\partial \epsilon_s^c}{\partial c^w}\right)\right] \cdot \frac{\partial n}{\partial \epsilon_s^c}}_{\text{indirect effect}}$$

$$\left(-\frac{\partial f}{\partial b}\right)^{\text{ge}} = \underbrace{\frac{\partial S}{\partial b}\left(-\frac{\partial f}{\partial S}\right)^{\text{ge}}}_{\text{direct effect}}$$

$$+ \underbrace{\left[\frac{\partial \epsilon_s^u}{\partial b} \cdot \frac{\partial c^w}{\partial \epsilon_s^u} + \frac{\partial \xi_s^c}{\partial b} \cdot \frac{\partial c^w}{\partial \xi_s^c} + \frac{\partial \epsilon_s^c}{\partial b} \cdot \frac{\partial c^w}{\partial \epsilon_s^c} + \frac{\partial c^w}{\partial b}\right] \cdot \left(-\frac{\partial f}{\partial c^w}\right)^{\text{ge}}}_{\text{indirect effect}}$$

$$\left(-\frac{\partial \theta}{\partial b}\right)^{\text{ge}} = \underbrace{\frac{\partial S}{\partial b}\left(-\frac{\partial \theta}{\partial S}\right)^{\text{ge}}}_{\text{direct effect}}$$

$$+ \underbrace{\left[\frac{\partial \epsilon_s^u}{\partial b} \cdot \frac{\partial c^w}{\partial \epsilon_s^u} + \frac{\partial \xi_s^c}{\partial b} \cdot \frac{\partial c^w}{\partial \xi_s^c} + \frac{\partial \epsilon_s^c}{\partial b} \cdot \frac{\partial c^w}{\partial \epsilon_s^c} + \frac{\partial c^w}{\partial b}\right] \cdot \left(-\frac{\partial \theta}{\partial c^w}\right)^{\text{ge}}}_{\text{indirect effect}}$$

3.H 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: $c^w > c_{\text{stw}}(\epsilon)$

$$\begin{aligned}\epsilon_s^c \;>\; \epsilon_s^u \;\Leftrightarrow\;\; & z_{\text{stw}}(\epsilon) + \frac{u(c^w) - u(b)}{u'(c^w)} - c^w + \tau_{\text{stw}}(\bar{h} - h(\epsilon)) \\ & + \frac{\lambda - \eta f}{1-\eta} \cdot \frac{k_v}{q} \\ & \geq\; z_{\text{stw}}(\epsilon) + \frac{u(c_{\text{stw}}(\epsilon)) - u(b)}{u'(c^w)} - c_{\text{stw}}(\epsilon) \\ & + \tau_{\text{stw}}(\bar{h} - h(\epsilon)) + \frac{\lambda - \eta f}{1-\eta} \cdot \frac{k_v}{q}\end{aligned}$$

$$\Leftrightarrow \quad \frac{u(c^w)}{u'(c^w)} - c^w \;>\; \frac{u(c_{\text{stw}}(\epsilon))}{u'(c^w)} - c_{\text{stw}}(\epsilon)$$

$$\Leftrightarrow \quad \frac{u(c^w) - u(c_{\text{stw}}(\epsilon))}{u'(c^w)} \;>\; c^w - c_{\text{stw}}(\epsilon)$$

$$\Leftrightarrow \quad u(c^w) - u(c_{\text{stw}}(\epsilon)) \;>\; (c^w - c_{\text{stw}}(\epsilon)) \cdot u'(c^w)$$

$$\Leftrightarrow \quad \int_{c_{\text{stw}}(\epsilon)}^{c^w} u'(c)\, dc \;>\; \int_{c_{\text{stw}}(\epsilon)}^{c^w} u'(c^w)\, dc$$

$$\Leftrightarrow \quad \int_{c_{\text{stw}}(\epsilon)}^{c^w} \left[u'(c) - u'(c^w)\right] dc \;>\; 0 \qquad \checkmark \quad \text{(due to risk aversion)}$$

3.I Calibration

To calibrate the model, the following system of equations can be solved, taking targets as exogenous. The endogenous model variables used in the system are pinned down at equilibrium level as well. Note that we use $\beta = 1$.

Exogenous:

$$f,\ q, \rho, b_{\text{rep}}, n$$

Endogenous:

$$k_f, \bar{m}, b, \epsilon_s^c, \epsilon_s^u, c^w, \omega$$

(I)

$$\rho \;=\; (1-p)\cdot G(\epsilon_s^u) + p\cdot G(\epsilon_s^c)$$

(II)

$$0 \;=\; z(\epsilon_s^u) + \lambda\cdot F + \frac{u(c^w)-u(b)}{u'(c^w)} - c^w + \frac{\lambda-\eta f}{1-\eta}\cdot\frac{k_v}{q}$$

(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)

$$\omega = \frac{1}{1-\rho} \cdot \Bigg((1-p) \cdot \int_{\epsilon_s^u}^{\infty} \left[c^w + \phi(h(\epsilon))\right] dG(\epsilon) \\ + p \cdot \int_{\epsilon_p}^{\infty} \left[c^w + \phi(h(\epsilon))\right] dG(\epsilon) \\ + p \cdot \int_{\epsilon_s^c}^{\epsilon_p} \left[c(\epsilon) + \phi(h(\epsilon))\right] dG(\epsilon) \Bigg)$$

(VIII)

$$b = b_{\text{rep}} \cdot \omega$$

3.J Optimization

To optimize for policy parameters, we use the following algorithm: First we use the global Covariance Matrix Adaptation Evolution Strategy (CMA-ES) with 8 random starting points. Then we select the best 4 solutions and use a local trust region optimization algorithm to further optimize these candidates. As the local optimization algorithm we rely on BOBYQA (Bound Optimization BY Quadratic Approximation). Finally, we select the best solution among these 4 candidates as the optimal policy parameters. This procedure is conducted independently for each value of p. Note, that since the Ramsey planner must obey the equilibrium conditions, each welfare evaluation of welfare within these optimization algorithms requires solving the model for the given policy parameters. We rely on an evenly spaced grid for p of 96 points between 0 and 1.

References

Abraham, Arpad, and Kirk White. 2006. “The Dynamics of Plant-Level Productivity in U.S. Manufacturing.” Working paper CES-WP-06-20. Center for Economic Studies, U.S. Census Bureau. [135]

Baily, Martin Neil. 1978. “Some aspects of optimal unemployment insurance.” *Journal of public Economics* 10 (3): 379–402. [108]

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. [109]

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. [105, 108, 125]

Braun, Helge, and Björn Brügemann. 2017. “Welfare Effects of Short-Time Compensation.” Tinbergen Institute Discussion Papers 17-010/VI. Tinbergen Institute. [109]

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. [105, 108]

Cahuc, Pierre. 2024. “The Micro and Macro Economics of Short-Time Work.” IZA Discussion Papers 17111. Institute of Labor Economics (IZA). [109]

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). [109]

Cahuc, Pierre, and André Zylberberg. 2008. “Optimum income taxation and layoff taxes.” *Journal of Public Economics* 92 (10-11): 2003–19. [105, 108, 125]

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. [109]

Chetty, Raj. 2006. “A general formula for the optimal level of social insurance.” *Journal of Public Economics* 90 (10-11): 1879–901. [108]

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. [135]

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. [109]

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. [135]

Drechsel, Thomas. 2023. “Earnings-Based Borrowing Constraints and Macroeconomic Fluctuations.” *American Economic Journal: Macroeconomics* 15 (2): 1–34. [109]

Duggan, Mark, Audrey Guo, and Andrew C. Johnston. 2023. “Experience Rating as an Automatic Stabilizer.” *Tax Policy and the Economy* 37: 109–33. [108]

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. [109]

Engen, Eric M., and Jonathan Gruber. 2001. “Unemployment insurance and precautionary saving.” *Journal of Monetary Economics* 47 (3): 545–79. [136]

Fonseca, Julia, and Bernardus Van Doornik. 2022. “Financial development and labor market outcomes: Evidence from Brazil.” *Journal of Financial Economics* 143 (1): 550–68. [109]

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. [109]

Giupponi, Giulia, Camille Landais, and Alice Lapeyre. 2021. “Should We Insure Workers or Jobs During Recessions?” Working paper. London, Centre for Economic Policy Research. [109]

Hopenhayn, Hugo, and Juan Pablo Nicolini. 1997. “Optimal Unemployment Insurance.” *Journal of Political Economy* 105 (2): 412–38. [108]

Hosios, Arthur J. 1990. “On the Efficiency of Matching and Related Models of Search and Unemployment.” *Review of Economic Studies* 57 (2): 279–98. [135]

Iliopulos, Eleni, FranÃ§ois Langot, and Thepthida Sopraseuth. 2019. “Welfare Cost of Fluctuations When Labor Market Search Interacts with Financial Frictions.” *Journal of Money, Credit and Banking* 51 (8): 2207–37. [109]

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. [108]

Jung, Philip, and Keith Kuester. 2015. “Optimal Labor-Market Policy in Recessions.” *American Economic Journal: Macroeconomics* 7 (2): 124–56. [108, 125, 136]

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. [135]

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. [108]

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. [108]

Michau, Jean-Baptiste. 2015. “Optimal labor market policy with search frictions and risk-averse workers.” *Labour Economics* 35 (C): 93–107. [108]

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. [106]

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. [106]

Ottonello, Pablo, and Thomas Winberry. 2020. “FINANCIAL HETEROGENEITY AND THE INVESTMENT CHANNEL OF MONETARY POLICY.” *Econometrica* 88 (6): pp. 2473–2502. [109]

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. [135]

Postel-Vinay, Fabien, and Hélène Turon. 2011. “Severance Packages.” IZA Discussion Papers 6023. Institute of Labor Economics (IZA). [108]

Ratner, David. 2013. “Unemployment insurance experience rating and labor market dynamics.” [108]

Shavell, Steven, and Laurence Weiss. 1979. “The Optimal Payment of Unemployment Insurance Benefits over Time.” *Journal of Political Economy* 87 (6): 1347–62. [108]

Stiepelmann, Gero. 2026. “Optimal Short-Time Work Policy in Recessions.” Working Paper 26-07. Federal Reserve Bank of Cleveland. [105, 109, 121, 130]

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. [115]

Wasmer, Etienne, and Philippe Weil. 2004. "The Macroeconomics of Labor and Credit Market Imperfections." *American Economic Review* 94 (4): 944–63. [109]