Essays on Human Capital

Inaugural-Dissertation

zur Erlangung des Grades eines Doktors
der Wirtschafts- und Gesellschaftswissenschaften

durch die

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

vorgelegt von

Simon Julian Heiler

aus Rüsselsheim

2026

Dekan:	Prof. Dr. Martin Böse
Erstreferent:	Prof. Dr. Moritz Kuhn
Zweitreferent:	Prof. Dr. Thomas Hintermaier

Tag der mündlichen Prüfung:	27. April 2026

Acknowledgements

My sincere thanks go to my supervisors Moritz Kuhn and Pavel Brendler. Their advice and experience have been tremendously beneficial for my education and for this thesis.

I would like to thank my colleagues at the Bonn Graduate School of Economics, in particular Andrea Amelio and Martin Kornejew, for their brilliance and their kindness in allowing me to draw from it, for their curiosity and genuine interest in my work, and for their friendship.

Moreover, I would like to extend my gratitude to the organizers and participants of the Bonn Macro Internal Seminar and the MEF Seminar, as well as faculty, staff, and colleagues at the Institute for Macroeconomics and Econometrics at the University of Bonn. I am grateful for every comment, suggestion, critique, and advice I have received over the years.

Financial support by the Bonn Graduate School of Economics and the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through RTG 2281 "The Macroeconomics of Inequality" and through CRC TR 224 (Project A03) is gratefully acknowledged.

Finally, I would like to thank my family. They have borne many of the burdens that pursuing a PhD entails. Their unwavering support is what ultimately led this thesis to its completion.

Contents

List of Figures

List of Tables

Introduction

Human capital is one of the building blocks of economic activity and is central to vast areas of the economic literature. The concept of human capital has been popularized in the economic literature by Mincer (1958) and Becker (1964). Today, our understanding of human capital is "the knowledge, skills, competencies and attributes embodied in individuals that facilitate the creation of personal, social and economic well-being" (OECD, 2001). However, since Lucas (1988), we understand that while human capital is embodied in individuals, it has important aggregate implications that shape economic development. The perspective that this thesis takes on human capital entails both aspects: human capital consists of the knowledge, skills, competencies, and experiences embodied in individuals that these individuals can utilize as a resource in economic processes, while recognizing that the collective consequences of individuals' decisions to acquire human capital create distinct aggregate economic effects. In this dual role, human capital is at the heart of many of the key forces and mechanisms of a modern economy, as it shapes decisions, sets incentives, and determines outcomes in myriad ways.

This thesis comprises three essays that examine particular ways in which human capital affects economic processes: the accumulation of human capital as a motive for individual choices, the availability of human capital as a determinant of industry composition, and the distribution of human capital as a measure of efficiency in the economy.

In Chapter 1, "Worker Heterogeneity and Optimal Unemployment Insurance", I analyze how policy design interacts with intrinsic motives for human capital accumulation in determining the search behavior of unemployed workers. I demonstrate that understanding the human capital accumulation channel allows for the design of targeted unemployment insurance policies that do not require conditioning on worker characteristics, but instead build on existing forces shaping individual decisions. Using US data, I find that fully conditional policies generate welfare gains of 0.35% of consumption, while simple unconditional policies can achieve up to 0.2% of consumption, or 60% of the gains of the more complex policy.

Chapter 2, "Demographic Change and Technological Choice", analyzes how demographic processes affect the relative availability of human capital as an input factor for production and thus shape decisions of firms to create and employ technology. Synthesizing demography-induced capital abundance à la Krueger and Ludwig (2007) with directed technological change à la Acemoglu and Restrepo (2021), I show how relative scarcity in human capital as an aggregate resource can shape the technological configuration of an economy across sectors of production. I document a general increase in investment per worker in Germany, that is most pronounced for investment in technology in the production sectors, and develop a model that is capable of replicating these patterns.

Chapter 3, "Labor Market Mismatch", assesses how the distribution of idle human capital across geography and occupation relative to the distribution of vacant positions affects efficiency in the labor market. Extending the work of Şahin, Song, Topa, and Violante (2014), I develop methods to measure different aspects of the alignment of searching workers and vacant positions and to quantify the distributional implications of policies that aim to alleviate any misalignment. I construct a novel long-term dataset on vacancies and searching workers in Germany and find mismatch that is economically meaningful. However, I also show that the distributional consequences of trying to reduce this mismatch likely far exceed the benefits.

While this thesis advances our understanding of how different aspects of human capital shape economic activity, each essay also points to important avenues for further research. The analysis in Chapter 1 highlights individual motives driving job search behavior, but does not account for aggregate implications of UI policy, thus abstracting from important channels that could amplify or dampen the effects. Chapter 2 introduces a novel link between demography and technology, yet provides no guidance on how the policymaker should account for this mechanism. Chapter 3 provides an approximation of the cost of implementing an allocation of searchers and vacancies that features less mismatch, but falls short of a proper welfare analysis of any such policy.

Addressing these questions is reserved for future work. I now begin the investigation of human capital with an analysis of how individual human capital accumulation motives interact with unemployment insurance policy design.

References

Acemoglu, Daron, and Pascual Restrepo. 2021. "Demographics and Automation." *Review of Economic Studies* 89 (1): 1–44. [2]

Becker, Gary S. 1964. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. 1st. Chicago: University of Chicago Press. [1]

Krueger, Dirk, and Alexander Ludwig. 2007. "On the consequences of demographic change for rates of returns to capital, and the distribution of wealth and welfare." *Journal of Monetary Economics* 54 (1): 49–87. [2]

Lucas, Robert E. 1988. "On the mechanics of economic development." *Journal of Monetary Economics* 22 (1): 3–42. [1]

Mincer, Jacob. 1958. "Investment in Human Capital and Personal Income Distribution." *Journal of Political Economy* 66 (4): 281–302. [1]

OECD. 2001. *The Well-being of Nations: The Role of Human and Social Capital*. Paris: OECD Publishing. [1]

Şahin, Ayşegül, Joseph Song, Giorgio Topa, and Giovanni L. Violante. 2014. "Mismatch Unemployment." *American Economic Review* 104 (11): 3529–64. [2]

Chapter 1

Worker Heterogeneity and Optimal Unemployment Insurance *

1.1 Introduction

Unemployment insurance (UI) is an important component of most countries' social security systems. At the core of any UI system is a classic insurance-incentive trade-off: more generous UI benefits reduce search effort of the unemployed. This trade-off varies with the workers' characteristics. It has been proposed in the literature to use conditional UI benefits to account for differences in the trade-off that workers face at different stages of their lives (Michelacci and Ruffo, 2015) or at different positions in the wealth distribution (Rendahl, 2012). In practice, however, conditional benefits are difficult, if not impossible, to implement.

This essay analyzes optimal UI policies when workers differ by age and ability. I find that an optimal age-and-ability-dependent policy generates welfare gains equivalent to a consumption increase by 0.35 percentage points. The essay's key finding is that a substantial fraction of these gains can be achieved through a simple, non-conditional policy consisting of a replacement rate on pre-unemployment earnings, a benefit floor and a benefit cap, much like the current US system.

To quantify the effects, I use an OLG model with endogenous human capital accumulation and exogenous job separations calibrated to the US economy. Human capital consists of a permanent ability type and experience, which is accumulated during employment (learning-by-doing) and depreciates during spells of unemployment. The policymaker maximizes expected lifetime utility of a worker at model entry. Optimal age-and-type-dependent replacement rates decrease with age and with ability. Rates for high-productivity workers start at ca. 45% and drop sharply to effectively zero around age 40. Rates for low-productivity workers start

* This chapter is based on Heiler, S., "Worker Heterogeneity and Optimal Unemployment Insurance: The Surprising Power of the Floor", CRC TR 224 Discussion Paper No. 545, 2024.

out even higher, at initially close to 100%, and decrease steadily over the life cycle to ca. 40% at the end of the working life. The optimal non-conditional policy closely replicates these patterns in the first half of the working life, but provides higher UI benefits in the second half of the working life than optimal conditional rates. As mentioned above, the welfare gains from implementing optimal age-and-type-dependent replacement rates correspond to a 0.35 percentage point increase in consumption in all states and periods. The welfare gains from implementing the optimal rate, floor, and cap policy correspond to a 0.2 percentage point increase, or roughly 60% of the gains from the fully conditional policy.

These results are driven by endogenous human capital accumulation. Returns from working are two-fold: employed workers receive wages that can be consumed or saved and accumulate experience, which yields higher future wages. UI replaces lost income, but cannot compensate for foregone experience. Workers with highly productive human capital technologies thus have strong incentives to invest in experience in the first periods of their working life. Towards the end of their working life, they have accumulated significant savings and returns to additional experience are low. The optimal replacement rate for these workers is therefore initially high and then drops significantly with age. Low ability workers, on the other hand, are less productive and experience more and longer unemployment spells. As a consequence, they accumulate less human capital and assets and are less able to self-insure against unemployment throughout their working lives. The optimal replacement rate for low ability workers therefore features less variation over the life cycle than for high ability workers.

The non-conditional UI policy can replicate these patterns by introducing a non-linearity in the effective replacement rate. Low income workers are more likely to be affected by the benefit floor, high income workers are more likely to be affected by the benefit cap. As income correlates strongly with age and ability, younger and lower ability workers are more likely to benefit from the floor and, when they do, the effects on UI benefits are larger. Vice versa, older and higher ability workers' benefits are limited more frequently by the cap, reducing their effective replacement rate.

The essay proceeds as follows: Section 1.2 presents the related literature, Section 1.3 presents empirical findings. Section 1.4 develops the model for policy analysis. The baseline calibration, the fit of the model to the data, and optimal conditional and non-conditional policies are presented in Section 1.5. This includes a discussion of the welfare effects. Section 1.6 concludes.

1.2 Related Literature and Contribution

This essay relates to the vast body of literature on optimal unemployment insurance in the tradition of Shavell and Weiss (1979) and Hopenhayn and Nicolini

(1997). At the core of this literature is the moral hazard problem caused by the unobservability of search effort. The trade-off then is between providing public insurance where private markets are incomplete and (not) distorting the incentives to actively search for work when unemployed. While moral hazard is socially undesirable (as it implements inefficient search effort levels), providing insurance is socially desirable (as it partly resolves market incompleteness).

An important recent finding of this literature is that UI serves a double-role: it provides insurance against income risk, and it provides liquidity to otherwise liquidity-constrained households (see e.g. Shimer and Werning, 2008). This, in turn, implies that longer unemployment duration in response to more generous UI policy is not unambiguously undesirable. When households are liquidity constrained, they search with higher than optimal effort. An increase in average unemployment duration after an increase in UI benefits could thus be caused by households being able to afford to search with optimal effort. Related to this, Chetty (2008) makes two observations: (i) the liquidity effect appears to be dominant for constrained households and (ii) the severity of the moral hazard problem in UI correlates with worker age. The findings indicate that the generosity of the UI system should account for a worker's age and level of savings.

A different strand of the literature focuses on the human capital channel. The importance of the channel has first been pointed out by Brown and Kaufold (1988). They analyze the interaction of UI policy with human capital investment decisions and identify a trade-off between providing more insurance and providing incentives to invest in human capital to self-insure.

Michelacci and Ruffo (2015) then combine the observation that response to UI correlates with age with the insight that the human capital channel is a key driver of this connection. They analyze optimal age-dependent UI policy using a structural model with an explicit age-structure and endogenous human capital accumulation through learning-by-doing. They find that raising UI replacement rates for younger workers and reducing replacement rates for older workers is welfare-enhancing. Their analysis, however, abstracts from any heterogeneity across workers other than age.

Focusing on differences in worker ability, but abstracting from age, Setty and Yedid-Levi (2020) find that the UI system can redistribute resources from high skill to low skill workers. The key channel for this are differential equilibrium unemployment frequency and duration of low skill workers compared to high skill workers. If all workers pay into the insurance system in proportion to their earnings, this means that low skill workers benefit more from the system than high skill workers. If, in addition, the system features a cap on UI benefits, which is more likely to bind for workers with higher skill levels, redistribution from high to low skill workers further increases. They abstract from endogenous labor supply choice and endogenous educational choice or skill investment. Their framework, thus, does not feature feedback effects from higher UI benefit levels to labor supply or human capital investment.

The combination of worker heterogeneity with respect to age and productivity thus represents a gap in the UI literature. This essay contributes to closing this gap. The contribution of this essay is two-fold: First, I analyze the potential welfare gains from setting UI replacement rates when explicitly conditioning on age and productivity is feasible; second, I demonstrate that a sizeable share of the potential welfare gains can be obtained through an implementable (i.e. real-life) policy, inspired by the system currently in place in the US.

1.3 Empirical Evidence

To demonstrate the role of age and idiosyncratic productivity in labor market policy, I attempt to answer two separate questions: First, do relevant differences by age and productivity exist in relevant labor market statistics? Second, does existing policy already account for these differences by differential treatment? These questions will be addressed in the following sections.

Note that this exercise aims at isolating the combined effect of age and idiosyncratic productivity on important determinants of labor market prospects and policies. I therefore attempt to control for observable differences other than age and productivity, meaning that the results do not necessarily coincide with population averages.

1.3.1 Data Sources

For the analysis, I use data from the Current Population Survey (CPS). The CPS is a monthly household survey conducted jointly by the Bureau of Labor Statistics and the US Census Bureau. The CPS uses a rotating panel design where households are interviewed for four consecutive months, excluded for eight months, then interviewed for a final four months. The survey covers approximately 60,000 households monthly from the civilian noninstitutional population aged 16 and older. The sample employs a multistage stratified probability design with state-level representativeness. This analysis uses data from the CPS Basic Monthly Survey (U.S. Census Bureau, 1989–2019), the Annual Social and Economic Supplement (ASEC, U.S. Census Bureau, 1989–2018), and the Job Tenure Supplement (U.S. Census Bureau, 2002–2018)[1], covering the period 1989–2019.[2] The Basic Monthly Survey contains information on an individual's current labor force status, as well as important socio-economic characteristics. Given the structure of

---

1. "Job Tenure and Occupational Mobility Supplement" or "Job Tenure/Occupation Mobility and Training Supplement" in earlier publications.
2. 2002–2019 for the Job Tenure Supplement.

the survey, the data allows for observing changes in the labor force status from one month to the next for about 75% of the sample. The ASEC and the Job Tenure Supplement are yearly supplementary surveys that contain additional information on earnings and tenure.

I combine the CPS data with information on UI program design compiled by the Employment and Training Administration (ETA). The ETA publishes semi-annual summaries of state UI laws for all US states (U.S. Department of Labor, Employment and Training Administration, 1989–2019). These summaries include, among other statistics, the parameters for the computation of weekly benefit amounts as well as upper and lower bounds for weekly benefits.

Finally, I use data on wealth and income from the Survey of Consumer Finances (SCF) extracts (Board of Governors of the Federal Reserve System, 1989–2019). I pool observations from the SCF waves from 1989 to 2019.

1.3.2 Evidence

The goal of this section is to document features of the US labor market and current UI system with a focus on differences by age and education. I highlight key features on risk, ability to mitigate risk, incentives, and UI generosity. Throughout this analysis, idiosyncratic productivity is approximated by the educational attainment of an individual. For the sake of clarity in presentation and sufficiency of sample sizes, I group all observations into three groups: high school dropouts (*low*), high school graduates (*medium*), and college graduates (*high*). Additional details on the analyses and further results can be found in Appendix 1.A.

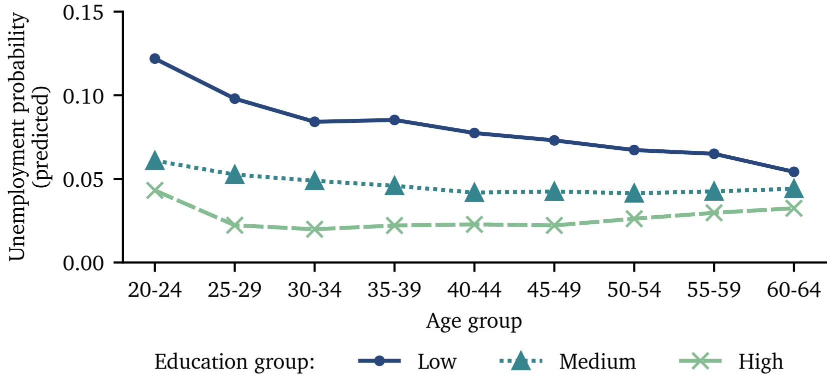
Unemployment risk. The first observation is that the likelihood to be affected by unemployment is not uniform across workers. To demonstrate this, I compute the average probability for a worker to be unemployed by age group and educational attainment using information on individuals' labor force status from the CPS basic monthly surveys. Unemployment probabilities are obtained in two steps: First, I predict individual probabilities using a probit model including an interaction term for age group and education group. Then, I compute conditional effects for each age-education combination.[3] The resulting age-profiles of average probabilities are depicted in Figure 1.1.

It is a well-documented fact that unemployment risk decreases both with age and with education individually.[4] When controlling for both dimensions simultaneously, this relationship only partly holds true. While unemployment rates are

3. For further details on the estimation procedure and a decomposition of the unemployment probabilities into transition rates between employment and unemployment, see Appendix 1.A.

4. On the former, see e.g. Shimer (1999); on the latter, see e.g. Mincer (1991).

Figure 1.1. Unemployment probabilities by age group and education



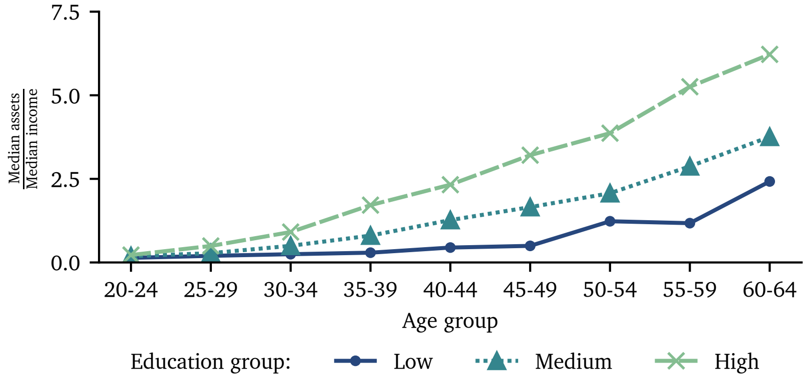
Notes: Average predicted unemployment probabilities by age group and educational attainment (male CPS sample, 1989–2019). Education groups are high-school dropouts (low), high-school graduates (medium), and college graduates (high). Predicted pobabilities obtained using (1.A.1); For details, see Appendix 1.A.
Source: CPS Basic Monthly (U.S. Census Bureau, 1989–2019), own computation.

decreasing by education for all age groups, they are no longer monotonously decreasing over the life cycle for all education groups. To be precise, the rates of high school dropouts and high school graduates without college education are falling over the life cycle, yet the rates for the group with a college degree exhibit a slight u-shape.

Asset holdings. Households not only differ in their risk of being affected by unemployment, but also in their ability to insure themselves against this risk. The ability to self-insure against income loss from unemployment is determined in part by the savings of an individual or household. The assets-over-yearly-income-ratio measures how many years of income a household can substitute by decumulating assets. The CPS does not feature information on household asset holdings. To examine the ability to self-insure by age end education, I compute median assets over median yearly income by age and educational attainment from the SCF extracts.

As can be seen in Figure 1.2, assets increase faster over the life cycle than income for all education groups, resulting in increasing age profiles of the assets-over-income-ratio. Moreover, households in which the reference person has attained more formal education hold significantly more assets relative to income for all ages. Finally, note that the profile is mostly flat for high school dropouts where the median household does not have relevant buffer savings up until age 50, indicating that this group cannot effectively self-insure for most of the life cycle. The evidence indicates significant differences in the ability to self-insure, both with respect to age and education.

Figure 1.2. Median assets over median income by age group and education



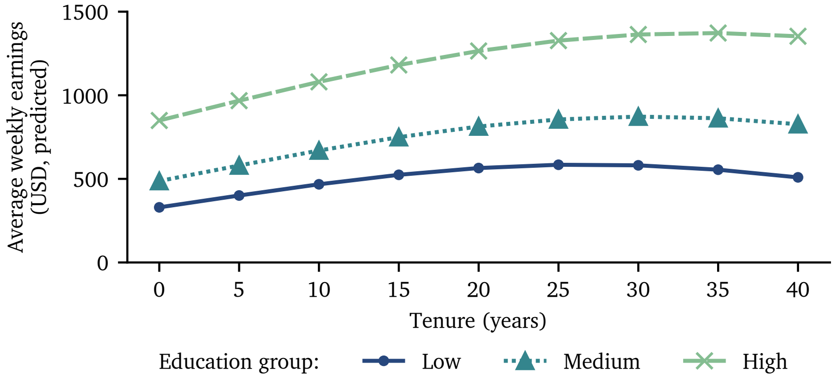
Notes: Median assets over median yearly income by age group and educational attainment (pooled data, 1989–2019). Education groups are high-school dropouts (low), high-school graduates (medium), and college graduates (high).
Source: SCF extracts (Board of Governors of the Federal Reserve System, 1989–2019), own computation.

Returns to education and experience. As mentioned before, understanding workers' incentives to search for employment is key in designing UI policy. The incentives to search are both immediate (generating income for consumption today) and long-term (accumulating experience for higher income and more consumption in the future). The relative strength of these incentives is closely linked to the value of experience in the labor market, reflected in higher earnings. The positive correlation between both education and experience on earnings is well-documented, at least since Mincer (1958). The evidence on the interaction between education and experience with respect to income is less clear. In an early study, Psacharopoulos and Layard (1979) find steeper experience-earnings profiles for higher education groups.

To assess the relationship between work experience and earnings, I use the CPS Job Tenure Supplement. I regress weekly labor earnings on tenure interacted with education, controlling for age, age squared, race, marital status, citizenship status and time and state fixed effects. To demonstrate average marginal effects of tenure by education, I predict average weekly earnings by education and years of experience. I find substantial effects of tenure on earnings that differ considerably across education groups. Figure 1.3 depicts the results.

Earnings are increasing in education and in experience. Returns to experience, i.e. wage increases from additional experience, are positive but decreasing for all education groups. In other words, returns are high for workers with little experience and low for more experienced workers. Finally, while absolute increases are larger for workers with more formal education, relative returns are comparable across groups. The wage-experience profiles indicate that motives to invest in

Figure 1.3. Average weekly earnings by years of experience and education



Notes: Average predicted weekly earnings by job tenure and educational attainment (male CPS sample, 2002–2018). Education groups are high-school dropouts (low), high-school graduates (medium), and college graduates (high). Predicted earnings obtained using 1.A.2; For details, see Appendix 1.A.
Source: CPS Job Tenure Supplements (U.S. Census Bureau, 2002–2018), own computation.

human capital accumulation depend strongly on the current level of experience for all education groups. They are strongest if the current level is low, which is typically the case at the beginning of a worker's employment history.

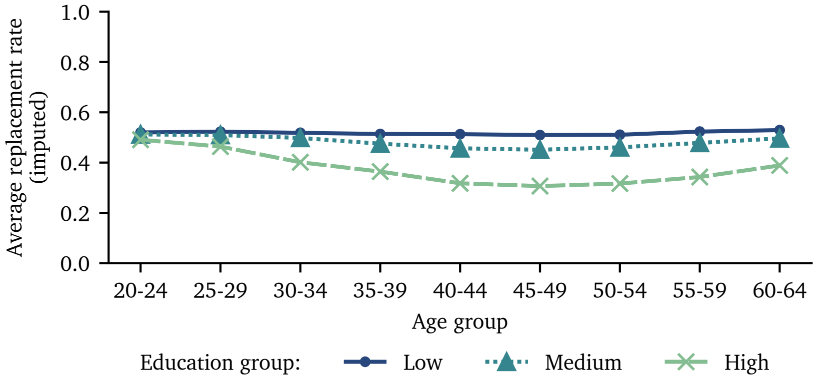
UI replacement rates. Finally, I analyze if and how the current UI system differentiates between workers based on their age or their educational attainment. Although the existing policy does not target groups of workers by these dimensions explicitly, the evidence suggests that UI generosity differs across groups.

The CPS does not include precise information on UI benefits, pre-unemployment wages or on UI replacement rates. UI benefits are therefore imputed in a two-step procedure: first, pre-unemployment wages are imputed for all unemployed individuals in the sample; in a second step, benefits are then imputed based on imputed pre-unemployment wages, following the procedure first employed by Cullen and Gruber (2000).[5] Effective replacement rates result from imputed earnings and benefits. Average statistics by age and education are then again obtained from individual quantities by regressing on observables and predicting average conditional effects. Figure 1.4 depicts average effective replacement rates by age group and education group.

To explain the shape of the profiles, recall one of the key features of the current system: the benefit cap. Naturally, individuals belonging to a group with higher average earnings are more likely to be affected by this cap than individuals belong-

5. For details on the procedure and further details regarding life cycle profiles of imputed earnings and benefits, see Appendix 1.A

Figure 1.4. Average effective UI replacement rate by age group and education



Notes: Average effective UI replacement rate by age group and educational attainment. Education groups are high-school dropouts (low), high-school graduates (medium), and college graduates (high). Effective replacement rates are obtained from average predicted pre-unemployment earnings and UI benefits using (1.A.3) (pre-unemployment earnings are imputed, UI benefits are then derived using the methodology of Cullen and Gruber, 2000); For details, see Appendix 1.A.
Source: CPS Basic Monthly (U.S. Census Bureau, 1989–2019), CPS ASEC (U.S. Census Bureau, 1989–2018), ETA UI policy statistics (U.S. Department of Labor, Employment and Training Administration, 1989–2019), own computation.

ing to a group with lower average earnings. This has two important implications: (i) differences in benefits between education groups are smaller than differences in earnings, i.e. age profiles are closer together, and (ii) the years during which an individual is likely affected by the cap differ across education groups (age-profiles exhibit different shapes over the life cycle). In other words, while average benefits for the low education group follow the profile of average earnings, the age-profile for the high education group is only slightly higher (much less than for earnings) and mostly flat. In combination, this translates into substantial differences in the age-profiles of replacement rates. Because the low education group is mostly unaffected by benefit caps, the replacement rate is almost constant over the life cycle for this group. In contrast, more highly educated individuals are more likely to be affected even in young ages, yielding a lower average replacement rate.[6] Moreover, with growing average earnings over the life cycle, they become even more likely to be affected, further lowering average replacement rates over the life cycle. In sum, there are substantial differences with respect to education between the life cycle profiles of effective replacement rates. Given the simple structure of the system, the rich heterogeneity in effective replacement rates is quite remarkable and clearly designates such systems as a candidate class for policy analysis.

6. For further details on the imputed share of workers affected by the benefit bounds, see Appendix 1.A.

As shown, the forces determining the household response to UI policy vary systematically with age and with productivity. Moreover, existing US policy already differentiates along these dimensions. The obvious next questions are (i) whether this differential treatment is beneficial in terms of welfare and (ii) which policies are optimal. These questions require a quantitative model and are addressed next.

1.4 Model

The model is a life cycle model with endogenous human capital accumulation, idiosyncratic labor risk and idiosyncratic labor productivity. The model is partial in the sense that it abstracts from the productive sector. As a consequence, there are no general equilibrium feedback effects of e.g. aggregate labor supply on wages or of aggregate asset holdings on interest rates.

Framework. There is a continuum of workers with total mass normalized to unity. Before entering the model, each worker draws a discrete type $k \in K$ that captures permanent differences in ability across workers. The type probabilities, and therefore the shares of workers of a given group in the total population, are constant over time. Workers live for a total of $\bar{n}_w + \bar{n}_r$ periods, the first $\bar{n}_w$ of which they are active in the labor market and the last $\bar{n}_r$ of which they are retired. All workers enter the model without having a job.

Preferences. Workers receive utility from consumption c and leisure l in every period of their lives and maximize expected discounted lifetime utility. The flow utility from consumption and leisure utility are assumed to be additively separable. For ease of notation, let $u(c)$ denote the utility component from consumption and $\psi(l)$ denote the utility component from leisure, i.e.

$$U(c,l) = u(c) + \alpha\psi(l) \tag{1.1}$$

where $\alpha \geq 0$ denotes the weight of leisure utility. Both components are characterized by a CRRA specification and the utility from consuming one unit of leisure is normalized to zero:

$$u(c) = \frac{c^{1-\sigma^c}}{1-\sigma^c} \qquad \psi(l) = \frac{l^{1-\sigma^l} - 1}{1-\sigma^l}$$

Preferences are homogeneous, i.e. all workers have the same utility specification and a common time discount factor β.

Financial markets. When workers enter the model, they start with initial assets $a_{k,0}$. Financial markets are incomplete, in that the only available savings device is a riskless bond that pays a constant interest rate r, satisfying $\beta = \frac{1}{1+r}$. Workers are allowed to borrow up to the borrowing limit $\underline{a}$[7].

Labor markets. Labor markets are fragmented: employers can observe worker types and workers of a given type can only work in their respective labor market. As a consequence, job separation rates and job finding rates are independent between markets. Within each labor market, the setup is identical. The timing in each labor market is characterized by two phases: search phase and consumption phase.

In the search phase, workers without a job can find employment by investing time in search. All workers are endowed with one unit of time. The search technology function $\zeta_k(s)$ captures how search effort translates into the probability to find employment. The search technology function is linear and truncated between zero and one

$$\zeta_k(s) \;=\; \max\{0, \min\{\gamma_k s + \mu_k\,, 1\}\} \tag{1.2}$$

The parameters of the search technology function are type-dependent, i.e. allow for differences in the efficacy of search for different types. Successful job search leads to a job in the same period. Time not invested in search is enjoyed as leisure. Workers that are already matched to a job do not search and consume their entire time endowment as leisure, which, given the normalization of the leisure utility function, yields leisure utility of zero.

In the second phase, all employed workers receive pre-tax wages $\bar{\omega}h$, where $\bar{\omega}$ is the average wage level and h is the worker's level of accumulated human capital. In every period of employment, workers face the risk of becoming unemployed with exogenous probability $\delta_{k,n}$. All workers choose how much out of their available resources to consume and how much to save.

Human capital accumulation. Next to the permanent ability of a worker, human capital is modelled as experience in the labor market. Upon entering the model, workers start with initial human capital level $h_{k,0}$. The human capital production technology is captured by a standard learning-by-doing (LBD) specification with participation (see, e.g. Blundell and Macurdy, 1999)

$$h'_k(h,e) \;=\; \big(1-\kappa_k\big)h + H(h,e) \tag{1.3}$$

where h' is the human capital level at the beginning of the next period, e is an indicator of whether the worker is employed and $0 \leq \kappa_k \leq 1$ is the type-specific human capital depreciation rate. In line with the extensive empirical literature on LBD production functions,[8] the function $H(\cdot)$ is assumed to be increasing and

7. I assume a fixed and exogenous borrowing limit that is identical across worker types. While borrowing limits could plausibly be type-dependent, I maintain this assumption for simplicity since the borrowing constraint plays a secondary role in the analysis.

8. For an early work, see Shaw (1989); more recently, see Imai and Keane (2004)

concave in h. For tractability, I assume a specific functional form. I take the specification from Blandin (2018) and adjust for participation and inelastic labor supply

$$H(h,e) \;=\; \mathbb{1}_{\{e=1\}}\lambda_k h^{\phi_k}$$

where λ_k captures the worker's type-specific learning ability and $0 \leq \phi_k < 1$ captures the type-specific curvature of new human capital with respect to current human capital.[9]

Government policy. The economy features three government programs: unemployment insurance, social security (pensions) and a general tax and transfer system. The UI system is financed by a proportional tax on labor income, τ^{UI}, and pays out benefits b^{UI} to unemployed workers. The social security program is also financed by a proportional labor income tax, τ^{SS}, and pays out pension benefits b^{SS} to retired workers. Pension benefits do not depend on type, age, assets, human capital, or earnings history. Finally, the government collects a general income tax, τ^{I}, on labor and capital income to finance lump-sum transfers T. Lump-sum transfers are paid to all workers in all periods and all states.

This essay focuses on the UI program. The social security program and the general tax and transfer system are included to ensure that the model replicates the US economy reasonably well. In essence, the social security program in the model ensures that the level of private savings for retirement is realistic and the general tax and transfers system aims at replicating redistribution across types that is prevalent in the US economy.

The UI program is modelled after the US UI system, in which benefits depend on pre-unemployment earnings. In this model, earnings are a function of experience h. Thus, the relevant quantity for the computation of UI benefits is the level of experience in the period prior to the unemployment spell. Let $\tilde{h}$ denote this quantity. In line with the US system, I assume that unemployed workers are eligible for UI benefits for a maximum of $\bar{m}$ periods. During eligibility, benefits are computed as a fraction of pre-unemployment earnings, potentially subject to minimum and maximum benefit amounts. Denote the UI benefit replacement rate by ρ, and the minimum and maximum amounts by $\underline{b}$ and $\bar{b}$, respectively. In the

9. LBD is one of two main specifications for human capital production in the literature. The alternative is dedicated skill investment in the tradition of Ben-Porath (1967) (BP). Both specifications have in common that workers with low current human capital and workers with highly productive human capital technologies have the strongest incentives to invest in human capital. Under BP, unemployed workers face the trade-off between investing in (or maintaining) human capital, searching for employment, and leisure. Consequently, the results should not depend on the specification of the human capital technology, at least qualitatively. Solving the model under BP is much more involved than under LBD, which is why the latter is chosen as technology.

general formulation, UI benefits are then

$$\tilde{b}_k^{UI}\left(n,\tilde{h}\right) \;=\; \max\left\{\underline{b}\,,\, min\left\{\rho_{k,n}\bar{\omega}\tilde{h}\,,\,\bar{b}\right\}\right\} \tag{1.4}$$

The replacement rate $\rho_{k,n}$ is restricted to be in $[0,1]$ and the benefit bounds are in absolute dollar values and assumed to be non-negative with $\underline{b} < \bar{b}$.

Note that the pre-unemployment experience level is linked to the current level of experience through human capital depreciation during unemployment: experience depreciates at a constant rate, and no additional experience accumulates during unemployment, hence

$$\tilde{h} \;=\; \left(1-\kappa^k\right)^{-m} h$$

Thus, unemployment duration enters the benefit computation in two ways: first, pre-unemployment experience is backed out from current experience and unemployment duration during eligibility, and second, if duration exceeds the maximum number of eligible periods, benefits are reduced to zero. Consequently, UI benefits can be expressed as a function of current experience:

$$b_k^{UI}(n,m,h) \;=\; \begin{cases} \tilde{b}_k^{UI}\left(n,\left(1+\kappa_k\right)^{-m} h\right) & \text{if } m \;\le\; \bar{m} \\ 0 & \text{if } m \;>\; \bar{m} \end{cases} \tag{1.5}$$

Household problem. At the core of the model is the worker optimization problem. There are two sequential decisions to take in every period, following the two phases in the labor market. In the first phase, workers without a job decide how much time to invest in job search and how much leisure to enjoy. In the second phase, all workers decide how much to consume out of their available resources and how much to save.

Given the separability of the flow utility, the normalization of the leisure utility function and the assumptions on labor supply and leisure consumption, the problem can be captured by three states through which the workers transition: *employed*, *unemployed*, and *searching*. The worker maximization problem can then be represented by a sequence of three value functions for each type k and age n, corresponding to the three states a worker can be in within a given period. Denote by $c_k^e(n,h,a,a')$ the consumption level of an employed worker with age n, human capital h, current assets a, and next periods assets a'. Let m denote unemployment duration, i.e. the count of consecutive periods (including the current period) in which the worker entered the *unemployed* state during the consumption phase. Then, $c_k^u(n,m,h,a,a')$ denotes the consumption level of an unemployed worker with age n, unemployment duration m, human capital h, current assets a, and next periods assets a'. Moreover, denote by $V_k^e(n,h,a)$ the expected present value of utility over the remaining life cycle for an employed worker of type k, of age n, with human capital level h, and asset holdings a, assuming the worker behaves optimally in all subsequent periods. Denote by $V_k^u(n,m,h,a)$ and $V_k^s(n,m,h,a)$ the same quantity for unemployed workers and searching workers, respectively, with an unemployment duration of m periods. The value of being employed is then given by the sum of utility from consumption in the current period and expected discounted continuation value of either directly transitioning to *employed* or being separated from the job and transitioning to the initial *searching* state, given that assets are chosen optimally:

$$\begin{aligned} V_k^e(n,h,a) = \max_{a' \geq \underline{a}} \Big\{ & u\left(c_k^e\left(n,h,a,a'\right)\right) \\ & + \beta\left(1-\delta_{k,n}\right) V_k^e\left(n+1, h'(h,1), a'\right) \\ & + \beta \delta_{k,n} V_k^s\left(n+1, 0, h'(h,1), a'\right) \Big\} \end{aligned} \tag{1.6}$$

The value of being in the *searching* state is given by the sum of leisure utility and the expected utility from either being *employed* or *unemployed* in the same period, depending on whether the job search was successful or not, given that search effort is chosen optimally:

$$\begin{aligned} V_k^s(n,m,h,a) = \max_{s \in [0,1]} \Big\{ & \alpha \psi(1-s) \\ & + \zeta_k(s) V_k^e(n,h,a) \\ & + \left[1-\zeta_k(s)\right] V_k^u(n,m+1,h,a) \Big\} \end{aligned} \tag{1.7}$$

The value of being in the *unemployed* state is given by

$$V_k^u(n,m,h,a) = \max_{a' \geq \underline{a}} \left\{ u\left(c_k^u\left(n,m,h,a,a'\right)\right) + \beta V_k^s\left(n+1, h'(h,0), m, a'\right) \right\} \tag{1.8}$$

which is the sum of consumption under unemployment and the expected discounted value from transitioning to *searching*, again given that assets are chosen optimally.

The consumption levels in the states of the consumption phase can be obtained from the budget constraint: For employed workers, this yields

$$c_k^e\left(n,h,a,a'\right) = \left(1-\tau^{UI}-\tau^{SS}-\tau^I\right)\bar{\omega} h + \left[1+\left(1-\tau^I\right)r\right]a + T - a' \tag{1.9}$$

For unemployed workers, the budget constraint implies that

$$c_k^u\left(n,m,h,a,a'\right) = b_k^{UI}(n,m,h) + \left[1+(1-\tau^I)r\right]a + T - a' \tag{1.10}$$

where $b_k^{UI}(n,m,h)$ is the UI benefit function imposed by the policymaker.

Retired agents do not participate in the labor market anymore, hence only the consumption phase remains, where workers decide how much out of retirement pension income and accumulated capital they consume. In the absence of survival risk or any other risk during retirement, retired workers perfectly smooth consumption over the entire retirement period. In other words, they consume their retirement income plus the annuity value of their savings, $c^r(a) = b^{SS} + T + \frac{(1-\tau^I)r[1+(1-\tau^I)r]^{\bar{n}_r}}{[1+(1-\tau^I)r]^{\bar{n}_r}-1}a$, and enjoy one unit of leisure in every period. Consumption during retirement thus only depends on the level of transfers and the asset level upon retirement, but neither on the workers acquired human capital nor on the state in the first period of retirement (i.e. the state to which the worker transitions from the last period of the working age). The value of retiring with asset level a for a worker of type k is independent of the employment status, i.e. $V_k^e(\bar{n}_w + 1, h, a) = V_k^u(\bar{n}_w + 1, h, a)$, and is given by

$$V_k^e\left(\bar{n}_w + 1, h, a\right) = \frac{1-\beta^{\bar{n}_r}}{1-\beta} U\left(c^r(a), 1\right) = \frac{1-\beta^{\bar{n}_r}}{1-\beta} u\left(c^r(a)\right) \quad \forall k \in K \tag{1.11}$$

All value functions can be solved for by backwards induction. Making extensive use of envelope conditions on optimal choices of consumption and search effort yields a set of three recursive first order conditions. Together with the terminal condition $c_k(\bar{n}_w + \bar{n}_r + 1, h, a) = 0 \; \forall \, k \in K$, value functions at model entry can be obtained iteratively. Recall that workers enter the model without a job (i.e. in the *searching* state) and with initial human capital $h_{k,0}$ and initial assets $a_{k,0}$. Thus, after the type of a worker has materialized, the expected discounted value of lifetime utility for that worker is given by $V_k^s(0, 0, h_{0,k}, a_{0,k})$. Before drawing the type, a worker has expected discounted value of lifetime utility of

$$\bar{V}^0 = \sum_{k \in K} \chi_k V_k^s\left(0, 0, h_{k,0}, a_{k,0}\right) \tag{1.12}$$

For further details on the solution algorithm, see Appendix 1.B.

Government problem. In this framework, the government problem consists of choosing the optimal UI system. For this purpose, the social security tax τ^{SS} and the general income tax τ^I are assumed to be exogenous to the government. The levels of retirement pensions b^{SS} and lump-sum transfers T are then determined by budget balance. All three government programs are assumed to be self-financing, i.e. the government faces three budget constraints. The budget constraints can be expressed via the net budget position of the government for a given program. Let B_k^P denote net fiscal contribution by program and worker type, i.e. the net present value of all contributions from and payments to workers of type k over their entire life cycle under program P. Note that positive values indicate net contributions and negative values indicate net transfers received. The government discounts future payments using the pre-tax interest rate r. The net budget positions for unemployment insurance, social security and general tax and transfers, respectively, are then given by

$$\begin{aligned} B_k^{UI} &= \sum_{n=0}^{\bar{n}_w} (1+r)^{-n} \int \tau^{UI} \bar{\omega} h \chi_k^e(n, dh) \\ &- \sum_{n=0}^{\bar{n}_w} (1+r)^{-n} \sum_{m=1}^{\bar{m}} \int b_k^{UI}(n, m, h) \chi_k^u(n, m, dh) \end{aligned} \tag{1.13}$$

$$\begin{aligned} B_k^{SS} &= \sum_{n=0}^{\bar{n}_w} (1+r)^{-n} \int \tau^{SS} \bar{\omega} h \chi_k^e(n, dh) \\ &- \sum_{n=\bar{n}_w+1}^{\bar{n}_w+\bar{n}_r} (1+r)^{-n} b^{SS} \chi_k \end{aligned} \tag{1.14}$$

$$\begin{aligned} B_k^{I} &= \sum_{n=0}^{\bar{n}_w} (1+r)^{-n} \int \tau^{I} \bar{\omega} h \chi_k^e(n, dh) + \sum_{n=0}^{\bar{n}_w+\bar{n}_r} (1+r)^{-n} \int \tau^I r a \chi_k(n, da) \\ &- \sum_{n=0}^{\bar{n}_w+\bar{n}_r} (1+r)^{-n} T \chi_k \end{aligned} \tag{1.15}$$

where χ_k are the type weights, $\chi_k^e(n, dh)$ is the measures of employed workers of age n with human capital level h, $\chi_k^u(n, m, dh)$ is the measure of unemployed workers of age n with unemployment duration of m periods and human capital h, and $\chi_k(n, da)$ is the measure of all workers of age n and with assets a.

For the main analysis, I assume that the government budgets need to be balanced in aggregates only, i.e. I allow for transfers across types within a given program. This is a slight deviation from the strand of literature assuming actuarially fair policies (Hopenhayn and Nicolini, 1997; Shimer and Werning, 2007, 2008). By allowing for transfers across types, the policies are only actuarially fair before the type of the agent is drawn. In other words, before entering the model, the worker expects zero net transfers in present value. Once the type is revealed, some workers are net payers and some workers are net receivers. The assumption is motivated by how UI systems are implemented in reality. While typically not the main objective of the system, most UI policies feature some degree of redistribution, through differential unemployment probabilities by types and often through caps on benefit amounts. For the assessment of the welfare implications of UI policy, it is therefore necessary to account for the distributional effects as well. Naturally, effects from more efficient search behavior and effects from redistribution blend in such a treatment.

The aim of the analysis is to measure potential welfare improvements that stem from targeting specific groups of workers. As outlined above, the government is bound to functional form (1.4) for the UI benefit function. The core analysis now consists of restricting the parameter space that is available to the policymaker and solving the corresponding optimization problem. Formally, let Θ denote the parameter space that the policymaker can choose from and θ denote a choice from this set. The optimization problem of the policymaker can then be stated as

$$\max_{\theta \in \Theta} \left\{ \bar{V}^0 \right\} \tag{1.16}$$

$$\text{s.t.} \sum_{k \in K} B_k^{UI} = 0 \tag{1.17}$$

$$\sum_{k \in K} B_k^{SS} = 0 \tag{1.18}$$

$$\sum_{k \in K} B_k^{I} = 0 \tag{1.19}$$

Constraint (1.17) captures the direct effects on the government budget of implementing a given UI program. Constraints (1.18) and (1.19) mechanically adjust transfers to account for spillover effects of changes in labor supply induced by UI policy.

This completes the model description. With all components in place, I now turn to model calibration and results.

1.5 Results

1.5.1 Calibration and Model Fit

The model is calibrated to male individuals in the US and parameters are obtained by matching selected model moments to their empirical counterparts. The complete calibration is summarized in Table 1.C.1.

Timing conventions. One model period corresponds to one quarter and workers are assumed to enter the model at age 20. Workers are active in the labor market for 45 years (corresponding to $\bar{n}_w = 180$ periods), then retire deterministically at age 65, are retired for 20 more years (corresponding to $\bar{n}_r = 80$ periods) and exit the model deterministically at age 85.

Human capital technology. The human capital production technology parameters are calibrated via indirect inference using CPS ASEC data on wages of male workers aged 20 to 64 between 1990 and 2010. In the model, wages depend not on age, but on accumulated human capital, which in turn depends on a worker's human capital production technology and employment history. Initial human capital levels and the parameters of the human capital production functions, $\{h_k^0, \lambda_k, \phi_k, \kappa_k\}_{k \in K}$, are chosen such that average equilibrium wages in the model match empirical wage profiles.[10] Throughout the analysis, the average wage level is normalized to $\bar{\omega} = 1$. Calibrated initial human capital levels are

10. For details on the construction of target wage profiles, see Appendix 1.C.1.

$h^0_{low} = 0.7$, $h^0_{medium} = 0.9$, and $h^0_{high} = 1.1$. The learning ability parameters are set to $\lambda_{low} = 0.03$, $\lambda_{medium} = 0.04$, and $\lambda_{high} = 0.06$, and the curvature parameter of the human capital production function is set to $\phi_k = 0.1$ for all types. Experience depreciates at rate $\delta^h_k = 0.025$ for all types. This indicates that *high* types start at a higher initial level and accumulate human capital more efficiently, yet all types are affected by loss of experience to the same degree.

Separation probabilities. Separation rates are directly calibrated to observed separation probabilities. As mentioned in Section 1.3, the CPS contains quarter-to-quarter changes in labor force status for about one quarter of the sample. I use these observations to compute average 3-months transition probabilities from employment to unemployment. For this, I first estimate a probit regression for the transition probability, controlling for age, education, race, marital status, and time effects. I then predict individual transition probabilities using the estimated model. The predicted transition probabilities are then used to compute average predicted transition probabilities by age group and education group (see Appendix 1.A for more details). To obtain separation probabilities for all model ages, I construct a spline function through these averages.[11]

Preferences. Risk aversion parameters for consumption and leisure are set to $\sigma^c = \sigma^l = 2.0$, the weight of leisure utility is set to $\alpha = 1.0$. This is within the range of the literature for specifications with separable consumption and leisure utility (see e.g. Guvenen, Kuruscu, and Ozkan, 2013, Heathcote, Storesletten, and Violante, 2017). The discount factor is set to $\beta = 0.99$ which matches annual interest rates of approximately 4%.

Search technology. As search effort is not observed in the data used for calibration, the parameters of the search technology are calibrated indirectly by matching unemployment rates by age and type. With separation rates fixed at predicted rates, unemployment rates in the model are determined by job finding rates only. Note that preferences on leisure and search technology are not individually identified through job finding rates. After choosing a value for the preference parameter σ^l, the parameters of the search technology are selected by matching empirical unemployment rates.[12] I allow for the search technology to depend on type. The calibrated values for the slope parameter $\{\gamma_k\}_{k\in K}$ are $\{1.00, 1.01, 1.09\}$ for low, medium, and high productivity workers, respectively. The calibrated values for the intercept parameter $\{\mu_k\}_{k\in K}$ are $\{0.14, 0.12, 0.08\}$. These values imply that additional search effort raises the job finding probability more for more productive

11. The knots of the spline correspond to the average age in the respective age group.

12. The calibration targets are the unemployment probabilities depicted in Figure 1.1. For details on the computation of these probabilities, see Appendix 1.A.

workers, yet low productivity workers have a higher probability to find employment without searching at all.

Initial assets and borrowing limit. Supported by the empirical evidence presented in Section 1.3, all workers are assumed to enter the model with zero assets, i.e. $a_k^0 = 0 \, \forall k \in K$. The borrowing limit is taken from Michelacci and Ruffo (2015), where it is calibrated to the 2007 Survey of Consumer Finances (SCF). The calibration target is the fifth percentile of the net worth distribution of workers under 35, divided by average quarterly total income, which amounts to −0.61 in the data. The constraint is thus set to −0.61 times the mean quarterly total income in the economy, which corresponds to $\underline{a} = -1.12$ in the base calibration.

Policy Parameters. Government policy consists of the UI benefit function, pension benefits, lump-sum transfers, and the respective tax rates. In the base calibration, the benefit function $b_k^{UI,base}(n, m, h)$ is given by a common and constant UI replacement rate on pre-unemployment earnings which is exogenously set to $\bar{\rho} = 0.5$. This is well in line with the empirical findings in Section 1.3 and comparable studies in the literature (see e.g. Chetty, 2008). The maximum UI benefit duration is set to $\bar{m} = 4$ quarters.[13] The UI income tax rate is chosen endogenously to keep the government budget condition (1.13) balanced. The equilibrium UI tax rate in the base calibration is $\tau^{UI} = 0.014$.

The labor income tax financing the social security system is set to $\tau^{SS} = 0.05$. The US Social Security Administration (SSA) reports tax rates for employers ranging between 5.015% and 5.3% over the calibration period (Office of the Chief Actuary, 2020). The equilibrium pension benefits (determined via budget constraint (1.14)) are $b^{SS} = 0.67$. This implies a ratio of retirement pensions over mean quarterly labor income of roughly 0.4, which is well in line with the data as reported by OECD (2007).

Finally, the general income tax rate is set to $\tau^I = 0.15$. The congressional budget office reports average income taxes to range between 7.2% and 11.8% over the calibration period with an average of 9.3% (Congressional Budget Office, 2020). The corresponding level of equilibrium lump-sum transfers implied by budget condition (1.15) is $T = 0.2$.

Figure 1.5 depicts the model fit for key variables by worker type. Panel 1.5a shows simulated relative wages (normalized by the average wage at model entry) against the empirical profiles of wages by type (normalized by the average wage at

13. Note that statutory benefit duration in US states is less than one year (typically 26 weeks). However, during the observation period, several federal or state programs have temporarily extended the duration of UI benefit beyond the statutory limit (up to 99 weeks during the Great Recession). The assumption represents an attempt to account for these extensions. A more restrictive assumption would likely dampen the results, as shorter benefit duration implies that fewer workers are covered under the policy.

age 20). The model matches the empirical profiles closely, both in terms of shape and level. The model fails to replicate the decrease in empirical wage curves for workers close to retirement. This could hint at several potential shortcomings of the model. For once, the assumption on inelastic labor supply might be too strong: if the downward-sloping wage profile close to retirement is generated by workers choosing to work fewer hours, the model is incapable of replicating this feature of the data. Alternatively, human capital depreciation, in particular late in the employment biography, could be stronger than modelled. The close fit of wage profiles over the majority of the working-age nevertheless supports the modelling assumptions.

Panel 1.5b shows simulated unemployment rates versus their empirical counterpart. The empirical targets correspond to the unemployment probabilities presented in Section 1.3. The simulated unemployment rates match the data very well. Since separation rates are matched exactly, unemployment probabilities are pinned down by the endogenous job finding rates. These are in turn determined by the combination of leisure utility and search technology. The good fit between simulated and empirical unemployment probabilities thus supports the assumptions made for leisure utility (preference homogeneity) and search technology (linear technology).

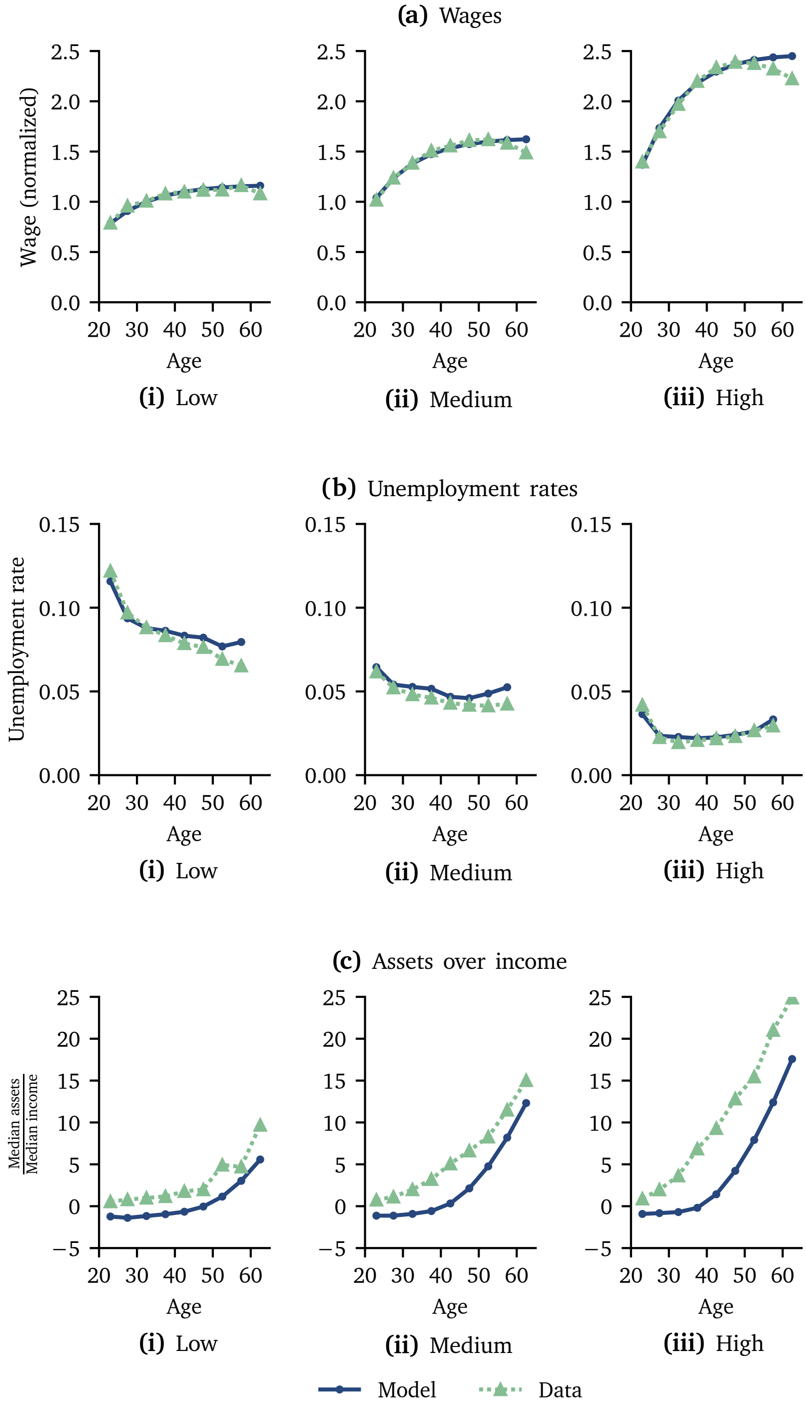
Panel 1.5c shows simulated assets-over-quarterly-income ratios versus their empirical counterpart. Note that assets are not targeted in the calibration. Overall, the model generates life cycle profiles of relative asset holdings that fit the empirical data remarkably well. The overall good fit of levels and profiles indicates that the model realistically replicates the ability to self-insure by age and productivity.

The model underestimates asset holdings of young agents across all education groups. In particular, workers of all types borrow against future income (up to the borrowing constraint) early in life and only accumulate wealth at later ages. Consequently, the simulated relative asset holdings of medium and, especially, high types are lower than in the data. While this difference could be reduced by imposing a stricter borrowing limit (a zero borrowing limit would likely bring the model closer to the data) such a restriction would be counterfactual given observed borrowing behavior. I maintain the calibration of Michelacci and Ruffo (2015) for two complementary reasons. First, it ensures consistency with the existing literature. Second, it represents the conservative choice: since the model generates less variation in the ability to self-insure against employment loss across types than observed in the data, any predictions regarding the variation in optimal replacement rates across types will be understated rather than overstated.

1.5.2 Policy Experiments

I now consider a series of policy experiments to assess the welfare implications of optimal UI policy. The policy experiments boil down to placing different restric-

Figure 1.5. Fit of model moments vs. observed data



Notes: Comparison of model and data moments; Panel 1.5a depicts wages (data: normalized such that the average wage at model entry/ age 20 is one, for details see Appendix 1.C); Panel 1.5b depicts unemployment rates (data: average predicted unemployment probabilities); Panel 1.5c depicts the ratio of median assets to median quarterly income by age group and educational attainment. Education groups are high-school dropouts (low), high-school graduates (medium), and college graduates (high).
Source: CPS Basic Monthly (U.S. Census Bureau, 1989–2019), CPS Tenure Supplements (U.S. Census Bureau, 2002–2018), SCF extracts (Board of Governors of the Federal Reserve System, 1989–2019), own computation.

tions on Θ in (1.16) and solving for optimal parameters. The first restriction I analyze is replicating a common assumption from the literature: in exercise (A), the instrument available to the policymaker is a simple replacement rate, that is common across types and constant by age:

$$b_k^{UI,A}\left(n,\tilde{h}\right) = \bar{\rho}\bar{\omega}\tilde{h} \;\; \forall \; n \;\; \forall \; k \in K$$

For the optimization problem (1.16), the parameter vector is $\theta^A = \{\bar{\rho}\}$ and the parameter space is $\Theta^A = [0,1]$.

The work of Michelacci and Ruffo (2015) and the empirical evidence presented in Section 1.3 highlight the importance of a worker's age and type in the tradeoff underlying the UI design problem. Consequently, I assess two restrictions that are targeting these dimensions. Under the second restriction, the policymaker can condition the policy on age, but not on type; under the third restriction, the policymaker can condition on both. Formally, in exercise (B), replacement rates are age-dependent, implying that

$$b_k^{UI,B}\left(n,\tilde{h}\right) = \rho_n\bar{\omega}\tilde{h} \;\; \forall \; k \in K$$

with $\theta^B = \left\{\{\rho_n\}_{n=0}^{\bar{n}_w}\right\}$ and $\Theta^B = [0,1]^{\bar{n}}$. Similarly, in exercise (C), replacement rates are age-and-type-dependent, implying that

$$b_k^{UI,C}\left(n,\tilde{h}\right) = \rho_{k,n}\bar{\omega}\tilde{h}$$

with $\theta^C = \left\{\left\{\{\rho_{k,n}\}_{n=0}^{\bar{n}_w}\right\}_{k\in K}\right\}$ and $\Theta^C = [0,1]^{\bar{n}K}$.

The final restriction I analyze is motivated by the current US UI system. In exercise (D), the available instruments are again a common and constant replacement rate, coupled with a minimum and maximum benefit amount. Formally, the benefit function is given by

$$b_k^{UI,D}\left(n,\tilde{h}\right) = \min\left\{\underline{b}, \max\left\{\bar{\rho}\bar{\omega}\tilde{h}, \bar{b}\right\}\right\} \;\; \forall \; n \;\; \forall \; k \in K$$

implying $\theta^D = \left\{\bar{\rho}, \underline{b}, \bar{b}\right\}$ and $\Theta^D = [0,1] \times \mathbb{R}_+^2$.

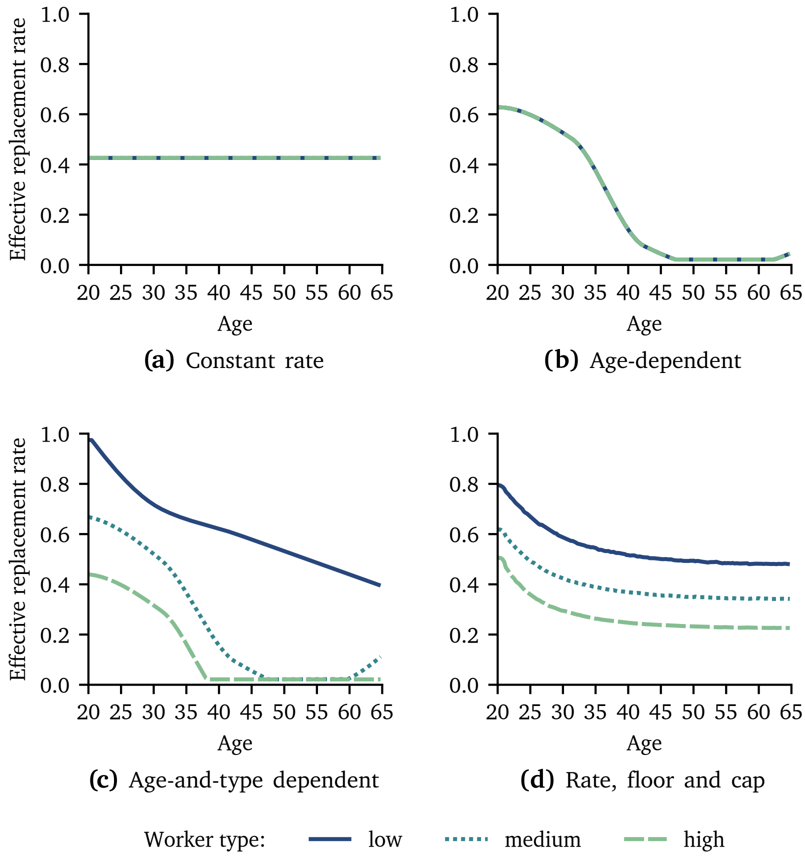
The logic of this structure is the following: Exercise (A) is a replication of previous work on optimal UI with similar modelling choices (e.g. Chetty, 2008). Exercises (B) and (C) are purely hypothetical: explicitly conditioning the policy on age is politically infeasible, and conditioning the policy on worker type is, at best, highly complex (due to observation and measurement issues) and potentially practically infeasible. The exercises are informative on the value of being able to target UI policy on these characteristics. In this sense, they yield an upper bound on the welfare gains that could be generated if it were feasible to directly condition on worker characteristics. The final exercise (D) then asks how much of these potential gains can be achieved if full use is made of the existing system. Note that this system does not feature any conditionality on (observable or unobservable) worker characteristics. As mentioned before, while both floor and cap are featured in the US UI system, only the cap has bite in the current implementation. Realizing the welfare gains under this policy regime would not require major policy reforms, but rather changing the parameters of the system that is already in place. The comparison across exercises (A)-(D) allows for a quantification of both the value of targeting information and the effectiveness of practical policy tools in achieving welfare improvements.

Figure 1.6 depicts the resulting life cycle profiles of effective replacement rates. For the exercises (A), (B), and (C), replacement rates are directly set by the policymaker. In exercise (D), effective replacement rates result from binding lower and upper bounds on benefits. When restricted to a single replacement rate for all workers, the policymaker implements $\bar{\rho} = 0.44$, which is close to the benchmark calibration (and, hence, the current implementation in the US). The optimal age-dependent policy exhibits the key features Michelacci and Ruffo (2015) found in their analysis on average workers: high replacement rates for young workers and a steady decline over the life cycle. Even when explicitly accounting for low-type workers, optimal age-dependent replacement rates drop to (close to) zero starting at age 45.

The picture changes significantly when the policymaker can target worker types specifically: while the replacement rate for all workers still decrease over the life cycle, low-type workers now receive substantial UI benefits even at older ages. The profiles of optimal age-and-type-dependent replacement rates are depicted in Figure 1.6b. The falling life cycle profile prevails, yet optimal replacement rates are higher throughout the life cycle for lower productivity the workers. Note that the policies exhibit two key features: first, for all types, replacement rates are falling over the life cycle, indicating a targeting to those workers that have the least ability to self insure (redistribution from old to young); second, on average, high-type workers receive less public insurance than medium-type and low-type workers, reflecting that income loss due to unemployment is particularly hurting low-type workers (redistribution from high to low).

As already demonstrated in Section 1.3, the simple structure consisting of a fixed and common replacement rate, a benefit floor and a benefit cap can nonetheless generate rich heterogeneity in effective replacement rates with respect to age and type (see Figure 1.6c). With parameters chosen optimally, this policy generates life cycle profiles that feature remarkable qualitative similarities to the profiles generated by age-and-type-dependent policies: higher replacement rates with a moderate decrease over the life cycle for low types, and lower replacement rates with a more pronounced decrease over the life cycle for high types. What is noteworthy about this, is that these patterns arise from a much smaller set of parameters and do not involve any conditionality. The optimal replacement rate is 0.22, and the benefit floor and cap are set at approximately 55% and 70% of the

Figure 1.6. Effective replacement rates under optimal policies



Notes: Effective replacement rates under optimized policy parameters by age and worker type. Policy choices are: (A) common and constant replacement rate (1.6a), (B) age-dependent replacement rates, common across types (1.6b), (C) age-and-type dependent replacement rates (1.6c), and (D) common and constant replacement rate, benefit floor, and benefit cap (1.6d). Policy parameters obtained from solving (1.16). The complete calibrations are summarized in tables 1.C.2–1.C.5.
Source: Own computation.

Table 1.1. Consumption equivalents of optimal policies

	Consumption equivalent			
Policy	Low	Medium	High	Average
(A) Common and constant rate	-0.06	0.03	0.13	0.04
(B) Age-dependent rate	0.04	0.24	0.40	0.25
(C) Age-and-type-dependent rate	2.16	0.21	-0.34	0.35
(D) Rate, floor, and cap	1.07	0.14	-0.16	0.20

Notes: Consumption equivalents are calculated using equation 1.B.19. The reference calibration for all equivalents is the *baseline* calibration. For details on the computation of consumption equivalents, see Appendix 1.B.2. Average consumption equivalents are obtained as weighted average over types.
Source: Own computation.

average wage level, respectively. The key difference to the current parametrization of the US UI system is a substantial increase in the benefit floor.

1.5.3 Welfare Analysis

To quantify the effects of the presented policies, I now turn to a simple welfare analysis using consumption equivalents. To construct consumption equivalents in an economy with heterogeneous workers, I first compute consumption equivalents for all types separately and then calculate averages, weighted by population shares.[14] The type-specific equivalents measure how consumption in the baseline setup would have to change – proportionally by the same factor in all periods and states – in order to make the worker at entry as well off as under the alternative setup, while keeping all other quantities constant (including utility from leisure). The average measure cannot be interpreted in the same way, yet is still indicative of the relative size of the welfare effects of the proposed policies.

Table 1.1 summarizes the results of the welfare comparison. First, note that the consumption equivalents of the optimal common and constant replacement rate are small. This is in line with the existing results that the US UI system is, on average, close to optimal. Type-specific consumption equivalents under (A) are positive for medium and high-type workers and negative for low-type workers, indicating that, on average, low-type worker would prefer a more generous UI system, while medium-type and high-type workers favor less public insurance.

The average consumption equivalent corresponding to optimally setting age-dependent replacement rates is 0.25 percentage points.[15] The type-specific equivalents indicate that the policy creates a welfare gain for all types, but high-type

14. For details on the computation of type-specific consumption equivalents, see Appendix 1.B.

workers benefit most. The intuition is that high-type workers value both the increased generosity for young workers and the effective abolishing of public insurance for older workers.

Additionally conditioning on worker types increases welfare gains to 0.35 percentage points of consumption in all states and periods, or by about 40% compared to age-dependent rates. Contrary to conditioning on age only, welfare gains are now largest for low-type workers and even turn negative for high types. The reason is that by targeting types, the policymaker can redistribute not only from old to young, but also from high-type workers to low-type workers.

Finally, the average welfare gains from imposing an optimal combination of common and constant replacement rate, benefit floor and benefit cap are equal to an increase in consumption of 0.2 percentage points This amounts to 80% of the gains from conditioning the replacement rate on age only and about 60% of the gains from conditioning the replacement rate on both age and type. Thus, the majority of the welfare gains from the hypothetical targeted policies can be achieved under a simple, untargeted policy. Note that type-specific welfare effects are roughly half the effects of the age-and-type-dependent policy, indicating that rate, floor, and cap work in a structurally similar way, yet are less effective.

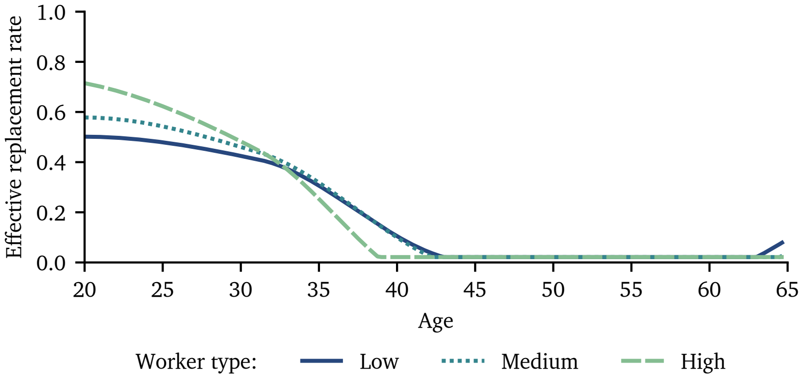
As mentioned before, part of these welfare gains stem from changes in redistribution across types. To disentangle the effects, I now turn to a decomposition of the welfare effect from conditioning on age and type. As mentioned before, when unemployment risks differ across workers, income-tax financed UI systems feature some degree of redistribution by design. These distributional effects are not the primary goal of the policy and if the policymaker wants to redistribute resources, other instruments are better suited. In most cases, changing UI policy also changes the degree of redistribution embedded in the system. To assess alternative policies, it is therefore important to differentiate the welfare effects from resolving the friction (here: unobservable search effort) and redistribution. For this, I compute optimal age-and-type-dependent replacement rates, holding the UI budget by type constant (i.e. at the level of the baseline calibration). In other words, I compute optimal age-and-type-dependent replacement rates in an economy with the same net transfers (in present value) across types as the baseline economy. Formally, denoting the exercise by (C'), the optimization problem becomes

$$\max_{\theta \in \Theta^C} \left\{ \bar{V}^0 \right\} \tag{1.20}$$

$$\text{s.t.} \quad B_k^P = B_k^{P,base} \quad \forall k \in K \quad \forall P \in \{UI, SS, I\}$$

15. The CEs found in this exercise are thus smaller than the welfare effects of age-dependent replacement rates found in Michelacci and Ruffo (2015), which amount to 0.8 percentage points of consumption.

Figure 1.7. Optimal age-and-type-dependent replacement rates under fixed budgets



Notes: Effective replacement rates by age and worker type under optimized policy parameters for exercise (C'). Policy parameters are age-and-type-dependent replacement rates, the government program budgets by type are fixed at baseline levels. Policy parameters obtained from solving (1.20). The complete calibration is summarized in Table 1.C.6.
Source: Own computation.

where $B_k^{P,base}$ are the net fiscal positions by type and program under the baseline calibration.

Figure 1.7 depicts the age profiles of replacement rates from problem (1.20). Optimal replacement rates follow a declining life cycle profile for all worker types. High-type workers exhibit the steepest decline because they experience the largest age-related differences in their ability to self-insure. By contrast, optimal replacement rates decline more gradually for low-type workers, who retain limited self-insurance capacity throughout their working lives.

The corresponding consumption equivalents are summarized in Table 1.2. There are two key observations: First, implementing optimal age-profiles by type generates significant welfare increases, even when redistribution across types is unchanged; about 90% of the average welfare gain of exercise (C) can be achieved with budgets fixed to baseline levels. Second, the additional average welfare gains from allowing the extent of redistribution via the UI system to change are limited and stem mainly from redistributing from high-type workers to low-type workers. The individual redistribution effects are sufficiently strong to completely flip the result for the high-types: the consumption equivalent turns from +0.3 percentage points with fixed budgets to -0.3 percentage points with a combined budget.

In summary, optimal replacement rates are decreasing over the life cycle, and more so for high-type workers vs. low-type workers, even when the budget by type is fixed.

Table 1.2. Decomposition of consumption equivalents of optimal policies

	Consumption equivalent			
Policy	Low	Medium	High	Average
(C) Age-and-type-dependent rate	2.16	0.21	-0.34	0.35
(C') Age-and-type-dependent rate (fixed budget)	0.35	0.30	0.32	0.32

Notes: Consumption equivalents are calculated using equation 1.B.19. The reference scenario for all equivalents is the *baseline* calibration. Average consumption equivalents are obtained as weighted average over types.
Source: Own computation.

1.6 Conclusion

Differences in idiosyncratic productivity, translating into differences in labor market risks and opportunities over the life cycle, generate rich heterogeneity in the decision context in which the labor supply choice is taken. Using a life cycle model with permanent productivity types and endogenous human capital accumulation, I find that UI policies that account for these differences can generate sizeable welfare gains.

I replicate the finding of Michelacci and Ruffo (2015) that conditioning UI policy on age can generate sizeable welfare gains and that optimal replacement rates fall with age. In my calibration, implementing optimal age-dependent replacement rates is equivalent to an increase of consumption of 0.25 percentage points in all states and periods. Moreover, optimally conditioning UI replacement rates on age and type generates welfare gains of 0.35 percentage points of consumption. Approximately 90%, or about 0.32 percentage points of consumption, come from inducing more efficient search behavior, and the remainder stems from an increase in redistribution across types relative to baseline.

As implementing policies explicitly conditioning on age and idiosyncratic productivity is difficult, if not impossible, these are hypothetical considerations. However, simple and robustly implementable policies exist and can generate a substantial share of the welfare gains of these conditional policies. The current US UI system proves to be one such policy, though the current parameterization leaves potential for improvement. A substantial increase of the benefit floor to around 55% of the average wage level, coupled with a UI replacement rate of 22% of pre-unemployment earnings, and a benefit cap at approximately 70% of the average wage level, yields welfare gains that are equivalent to raising consumption in all states and periods by about 0.25%. This amounts to ca. 80% of the welfare gains from conditioning replacement rates on age and to about 60% of the gains from a fully age-and-type-dependent policy.

The analysis is focused on different options for the UI benefit function. This focus represents one of the major limitations of the analysis. A range of studies have shown that extending the policies to include the financing side of the policy can vastly increase the potential welfare gains (see e.g. Huggett and Parra, 2010 or Michelacci and Ruffo, 2015). A natural extension of this work is thus to include tax policies in the analysis.

References

Ben-Porath, Yoram. 1967. "The Production of Human Capital and the Life Cycle of Earnings." *Journal of Political Economy* 75 (4): 352–65. [16]

Blandin, Adam. 2018. "Learning by Doing and Ben-Porath: Life-cycle Predictions and Policy Implications." *Journal of Economic Dynamics and Control* 90: 220–35. [16]

Blundell, Richard, and Thomas Macurdy. 1999. "Chapter 27 - Labor Supply: A Review of Alternative Approaches." In. Edited by Orley C. Ashenfelter and David Card. Vol. 3, Handbook of Labor Economics. Elsevier, 1559–695. [15]

Brown, Eleanor, and Howard Kaufold. 1988. "Human Capital Accumulation and the Optimal Level of Unemployment Insurance Provision." *Journal of Labor Economics* 6 (4): 493–514. [7]

Chetty, Raj. 2008. "Moral Hazard versus Liquidity and Optimal Unemployment Insurance." *Journal of Political Economy* 116 (2): 173–234. [7, 23, 26, 37]

Congressional Budget Office. 2020. "The Distribution of Household Income, 2017." Washington, DC: Congress of the United States. [23]

Cullen, Julie Berry, and Jonathan Gruber. 2000. "Does Unemployment Insurance Crowd Out Spousal Labor Supply?" *Journal of Labor Economics* 18 (3): 546–72. [12, 13, 37–39]

Guvenen, Fatih, Burhanettin Kuruscu, and Serdar Ozkan. 2013. "Taxation of Human Capital and Wage Inequality: A Cross-Country Analysis." *Review of Economic Studies* 81 (2): 818–50. [22]

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante. 2017. "Optimal Tax Progressivity: An Analytical Framework*." *Quarterly Journal of Economics* 132 (4): 1693–754. [22]

Hopenhayn, Hugo A., and Juan Pablo Nicolini. 1997. "Optimal Unemployment Insurance." *Journal of Political Economy* 105 (2): 412–38. [6, 20]

Huggett, Mark, and Juan Carlos Parra. 2010. "How Well Does the U.S. Social Insurance System Provide Social Insurance?" *Journal of Political Economy* 118 (1): 76–112. [33]

Imai, Susumu, and Michael P. Keane. 2004. "Intertemporal Labor Supply and Human Capital Accumulation*." *International Economic Review* 45 (2): 601–41. [15]

Michelacci, Claudio, and Hernán Ruffo. 2015. "Optimal Life Cycle Unemployment Insurance." *American Economic Review* 105 (2): 816–59. [5, 7, 23, 24, 26, 27, 30, 32, 33, 37]

Mincer, Jacob. 1958. "Investment in Human Capital and Personal Income Distribution." *Journal of Political Economy* 66 (4): 281–302. [11]

Mincer, Jacob. 1991. "Education and Unemployment." *NBER Working Paper*, (3838): Also published in Studies in Human Capital, Elgar Publishers, 1993. [9]

Miranda, Mario J., and Paul L. Fackler. 2004. *Applied Computational Economics and Finance*. Vol. 1, MIT Press Books 0262633094. The MIT Press. [56]

OECD. 2007. *Pensions at a Glance 2007*, 204. [23]

Office of the Chief Actuary. 2020. "Social Security Tax Rates." Washington, DC: Social Security Agency. [23]

Psacharopoulos, George, and Richard Layard. 1979. "Human Capital and Earnings: British Evidence and a Critique." *Review of Economic Studies* 46 (3): 485–503. [11]

Rendahl, Pontus. 2012. "Asset-based Unemployment Insurance." *International Economic Review* 53 (3): 743–70. [5]

Setty, Ofer, and Yaniv Yedid-Levi. 2020. "On the Provision of Unemployment Insurance when Workers are Ex-Ante Heterogeneous." *Journal of the European Economic Association*, (05): jvaa015. [7]

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

Shaw, Kathryn L. 1989. "Life-Cycle Labor Supply with Human Capital Accumulation." *International Economic Review* 30 (2): 431–56. [15]

Shimer, Robert. 1999. "Why Is the U.S. Unemployment Rate So Much Lower?" In *NBER Macroeconomics Annual 1998*. Edited by Ben S. Bernanke and Julio J. Rotemberg. Vol. 13, Cambridge, MA: MIT Press. [9]

Shimer, Robert, and Iván Werning. 2007. "Reservation Wages and Unemployment Insurance." *Quarterly Journal of Economics* 122 (3): 1145–85. [20]

Shimer, Robert, and Iván Werning. 2008. "Liquidity and Insurance for the Unemployed." *American Economic Review* 98 (5): 1922–42. [7, 20]

U.S. Census Bureau. 1989–2018. *Current Population Survey, Annual Social and Economic Supplement, 1989-2018*. Annual household survey microdata. U.S. Census Bureau. URL: https://www.census.gov/programs-surveys/cps.html. March CPS with detailed income, employment, and demographic data. [8, 13, 38, 39]

U.S. Census Bureau. 1989–2019. *Current Population Survey, Basic Monthly Files, 1989-2019*. Survey microdata. U.S. Census Bureau. URL: https://www.census.gov/programs-surveys/cps.html. January 1989 - December 2019, monthly microdata. [8, 10, 13, 25, 36, 38, 39]

U.S. Census Bureau. 2002–2018. *Current Population Survey, Tenure Supplement, 2002-2018*. Survey data. U.S. Census Bureau. URL: https://www.census.gov/programs-surveys/cps.html. [8, 12, 25]

U.S. Department of Labor, Employment and Training Administration. 1989–2019. "Significant Provisions of State Unemployment Insurance Laws." Semiannual report. U.S. Department of Labor, Employment, and Training Administration. January 1989 through July 2019 editions, valid through December 2019. [9, 13, 37–39]

Board of Governors of the Federal Reserve System. 1989–2019. *Survey of Consumer Finances, Extract Data, 1989-2019*. Household survey microdata. Board of Governors of the Federal Reserve System. URL: https://www.federalreserve.gov/econres/scfindex.htm. Triennial cross-sectional survey, waves 1989, 1992, 1995, 1998, 2001, 2004, 2007, 2010, 2013, 2016, 2019. [9, 11, 25]

Appendix 1.A Empirical Analysis

This Appendix describes the procedures used for estimation and imputation of data in more details and presents some additional results.

Labor market statistics. An individual's labor force status is reported as a categorical variable in the CPS basic monthly surveys. For my analysis, I only differentiate between three categories: *employed* (combining "employed - at work" and "employed - absent"), *unemployed* (combining "unemployed - looking" and "unemployed - on layoff"), and *not in labor force* (combining "retired", "disabled", "unavailable", and "other"). As mentioned before, for part of the sample, the labor market status in the month following the current observation is available, allowing to observe status changes. The outcome variables are therefore defined as the individual having labor market status *unemployed* in the observation period (for unemployment probabilities), the individual having labor market status *employed* in the observation period and *unemployed* in the following period (for employed-to-unemployed transitions), and the individual having labor market status *unemployed* in the observation period and *employed* in the following period. Only observations for which the labor market status is observed for both the current and the following period are considered for the estimation of transition probabilities.

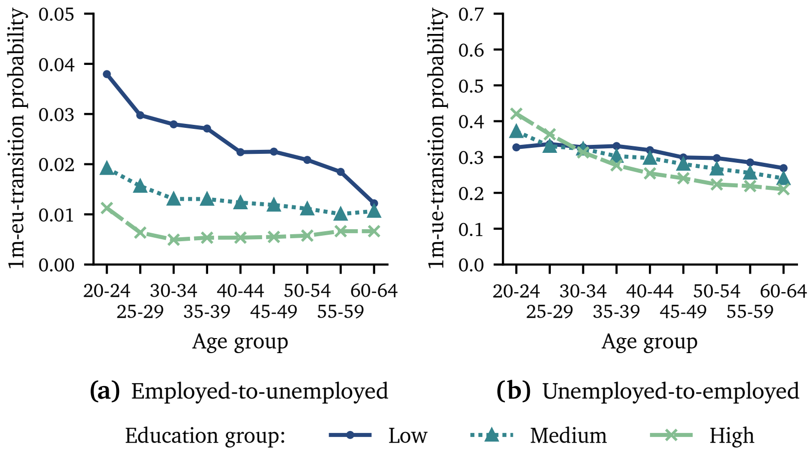
To predict individual unemployment and transition probabilities, I estimate probit models with the interaction of age group and education group as covariate of interest, controlling for observables of the individual, as well as location and time. To be precise, I estimate the following regression equation

$$\Phi^{-1}\left(P\left(Y_i = 1\right)\right) = \alpha + \sum_{j=1}^{J} \sum_{k \in K} \beta_k^j \mathbb{1}_{\{\text{AgeGroup}_i = j\}} \mathbb{1}_{\{\text{Education}_i = k\}} + \gamma \boldsymbol{X}_i + \varepsilon_i \quad (1.A.1)$$

where Y_i is the labor market status variable of interest, $\boldsymbol{X}_i$ contains marital status and race, as well as state, month, and year fixed effects. The coefficient estimates β_k^j then measure the combined effect of age group and education on the probability of a given status event occurring.[16] The fixed effects ensure that differences across states (capturing geographic location and jurisdiction), years (capturing long term trends), and months (capturing yearly cyclicality in the data) are controlled for. Using the estimated model, I compute predicted unemployment probabilities for each age-education combination. These predictions represent the expected probability of the labor market status event occurring for the average individual in the sample, where "average" refers to the empirical distribution of

16. Note that one age-education combination serves as the reference category and is absorbed in the intercept α.

Figure 1.A.1. 1-month transition probabilities



Notes: Average predicted transition probability between labor force statuses by age group and educational attainment (male CPS sample, 1989–2018). Education groups are high-school dropouts (low), high-school graduates (medium), and college graduates (high). Predicted probabilities are obtained using (1.A.1) for transitions from *employed* to *unemployed* (left panel) and from *unemployed* to *employed* (right panel).
Source: CPS Basic Monthly (U.S. Census Bureau, 1989–2019), own computation.

control variables (state, year, month, marital status, and race) observed in the data. The resulting average unemployment probability is depicted in Figure 1.1, and the average probabilities to transition between labor market statuses are depicted in Figure 1.A.1.

As can be seen in Figure 1.A.1a, the employed to unemployed transition probabilities exhibit a similar pattern as unemployment rates: a strict ordering by education for all age groups and a general decrease over the life cycle. Average unemployed to employed transition probabilities, however, present a different picture (see Figure 1.A.1). While there is a general common downward trend over the life cycle, there is no strict ordering by education across age groups: the transition probability for college graduates is highest across education groups for young workers and lowest for older workers.

Earnings by tenure. To obtain earnings by tenure and education group, I estimate the following regression equation:

$$Y_i = \alpha + \sum_{k \in K} \beta_k^1 \mathbb{1}_{\{\text{Education}_i = k\}} \text{Tenure}_i + \sum_{k \in K} \beta_k^2 \mathbb{1}_{\{\text{Education}_i = k\}} \left(\text{Tenure}_i\right)^2 + \gamma \boldsymbol{X}_i + \varepsilon_i \tag{1.A.2}$$

where the dependent variable Y_i is the log of weekly earnings (measured in 1990 USD), and $\boldsymbol{X}_i$ is a vector of controls including age, age squared, marital status, race, as well as year and state fixed effects. As before, average predicted earnings for each tenure-education combination are computed using the estimated model. The resulting average earnings are depicted in Figure 1.3.

Pre-unemployment earnings and UI benefits. Given the limitations in the CPS data, unemployment benefits are computed in several steps. Since pre-unemployment earnings are not reported in the CPS, I impute earnings in the period prior to the unemployment spell.[17] For this, I run a conventional wage regression for each year and for each state using wage information from the CPS ASEC:

$$Y_i = \alpha + \beta^1 \text{Age}_i + \beta^2 \left(\text{Age}_i\right)^2 + \gamma \text{Education}_i + \delta \text{Married}_i + \theta \text{Race}_i + \varepsilon_i$$

where the dependent variable Y_i is the log of reported "*usual weekly earnings before deductions*", the independent variables are a quadratic polynomial of age, dummy variables for white and black individuals, a dummy variable for married individuals, as well as four dummies for educational attainment ("*Less than a High School Diploma*", "*High School graduates, no College*", "*Some College or Associate Degree*", and "*Bachelor's Degree and higher*"). Using the estimated coefficients, individual pre-unemployment earnings are then imputed for all unemployed male individuals aged between 16 and 64 years in the respective CPS basic monthly surveys.

In the second step, unemployment benefits are imputed using the approach of Cullen and Gruber (2000). The characteristics of the UI system are extracted from the ETA summaries (U.S. Department of Labor, Employment and Training Administration, 1989–2019). In the current setup, I simulate base benefits and, for the states in which they apply, additional benefits from dependent allowances. For the number of dependents, I currently only use information on unemployed / non-working spouses, as information of dependent children is not available for all years in the sample in the CPS. In most US states, weekly benefit amounts are a function of wages in a base period prior to unemployment (usually the five quarters prior to the unemployment spell). A precise calculation of potential UI benefits would thus require knowledge about the recent earning history of the individual. As the CPS basic monthly survey features a very limited panel dimension, base period earnings cannot be observed directly and need to be approximated. I use multiples of imputed pre-unemployment weekly earnings as proxy for base period earnings. Moreover, I currently do not simulate qualification requirements for unemployment benefits that exist in some states. Both simplifying assumptions are in line with the literature (see e.g. Michelacci and Ruffo, 2015, and Chetty,

17. Note that the structure of the CPS allows, in principle, for observing pre-unemployment earnings for individuals with one to three or with twelve to fifteen months of unemployment. I do not use this information, but instead impute pre-unemployment earnings for two reasons: the data structure is potentially subject to selection bias, and the data quality of the observations for which pre-unemployment earnings are directly available is unclear.

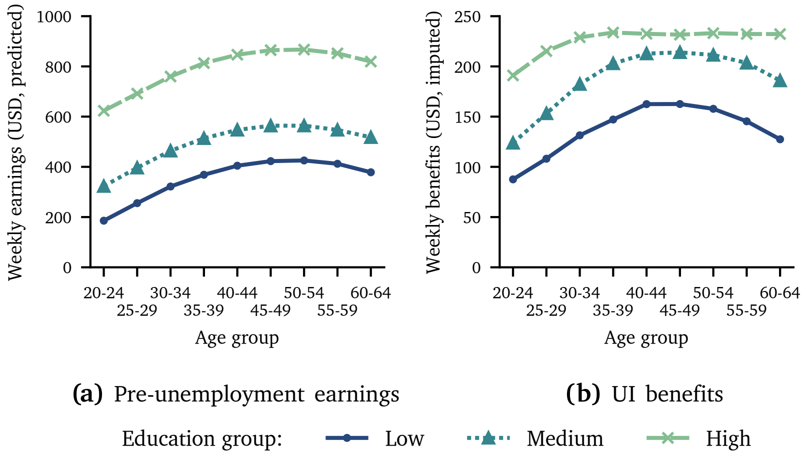
2008). The validity of this approach has also been demonstrated in the original study by Cullen and Gruber (2000): approximating base period earnings with earnings in the output quarter yields a correlation between potential benefits of 0.90, and additionally abstracting from qualification rules lowers the correlation between approximated and actual potential benefits to 0.88.

Once pre-unemployment earnings and benefits have been imputed, I estimate the following regression equation to obtain population-average age profiles by education group:

$$\tilde{Y}_i = \alpha + \sum_{j=1}^{J} \sum_{k \in K} \beta_k^j \mathbb{1}_{\{\text{AgeGroup}_i = j\}} \mathbb{1}_{\{\text{Education}_i = k\}} + \gamma \boldsymbol{X}_i + \delta^s + \theta_y^t + \theta_m^t + \varepsilon_i \quad (1.A.3)$$

where $\tilde{Y}_i$ are an individuals imputed pre-unemployment earnings and UI benefits, respectively, $\boldsymbol{X}_i$ contains marital status and race, δ^s are state fixed effects, and θ_y^t and θ_m^t are year and month fixed effects, respectively. The average predicted pre-unemployment earnings and predicted benefits are then again obtained from the estimated models and depicted in Figure 1.A.2.

Figure 1.A.2. Average earnings and benefits by age and education

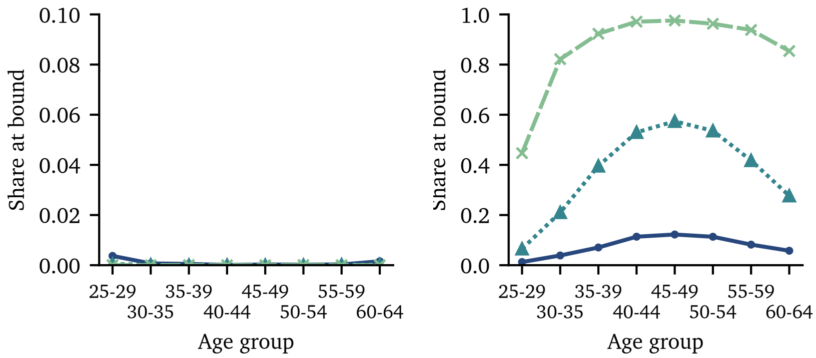


Notes: Average predicted weekly pre-unemployment earnings (left panel) and average predicted weekly UI benefits (right panel) by age group and educational attainment. Earnings and benefits are measured in 1990 USD. Education groups are high-school dropouts (low), high-school graduates (medium), and college graduates (high). Predicted earnings and benefits are obtained using (1.A.3).
Source: CPS Basic Monthly (U.S. Census Bureau, 1989–2019), CPS ASEC (U.S. Census Bureau, 1989–2018), ETA UI policy statistics (U.S. Department of Labor, Employment and Training Administration, 1989–2019), own computation.

The flattening of the benefit profiles is driven by the benefit cap: higher average earnings imply that a larger fraction of workers are affected by the cap. The higher proportion of workers at the upper bound then translates into lower effective replacement rates.

The imputation procedure allows to compute the share of individuals affected by the UI bounds, which is depicted in Figure 1.A.3. As can be seen, the floor is largely ineffective: the share of individuals affected is virtually zero for all but low education workers at the beginning of their working life and close to retirement. As expected from the benefit profiles, the upper bound plays a significant role for many individuals, especially for high education workers. The share of recipients at the bound exhibits an inverse u-shape over the life cycle, following the shape of (pre-unemployment) earnings. For intermediate age groups, almost all high education recipients are at the bound. This is the main driver behind the heterogeneity in effective replacement rates discussed in Section 1.3.

Figure 1.A.3. Share of UI benefit recipients affected by bounds



(a) Benefit floor **(b)** Benefit cap

Education group: Low Medium High

Notes: Average share of UI benefit recipients affected by benefit floor (left panel) and benefit cap (right panel) by age group and educational attainment (full CPS sample, 1989–2018). Education groups are high-school dropouts (low), high-school graduates (medium), and college graduates (high). Benefits are imputed from imputed pre-unemployment earnings using the methodology of Cullen and Gruber (2000).
Source: CPS Basic Monthly (U.S. Census Bureau, 1989–2019), CPS ASEC (U.S. Census Bureau, 1989–2018), ETA UI policy statistics (U.S. Department of Labor, Employment and Training Administration, 1989–2019), own computation.

Appendix 1.B Proofs and Derivations

1.B.1 Solution of the Baseline Economy

Solving the household problem. The first order conditions for the consumption choices are obtained from the value functions in Section 1.4. The following derivations focus on interior solutions, thus omitting the Kuhn-Tucker multipliers for the constraints on parameters. In case any of the constraints are binding, the solutions can be directly obtained from the constraints.

Let $a_k^{e*}(n,h,a)$ and $a_k^{u*}(n,m,h,a)$ be the asset choice that solves the maximization problem in a given period for employed workers and unemployed workers, respectively and let

$$\begin{aligned} c_k^{e*}(n,h,a) &= c_k^e\left(n,h,a,a_k^{e*}(n,h,a)\right) \\ c_k^{u*}(n,m,h,a) &= c_k^u\left(n,m,h,a,a_k^{u*}(n,m,h,a)\right) \end{aligned}$$

denote the corresponding optimal consumption choices. For interior solutions, $a_k^{e*}(n,h,a)$ satisfies

$$\begin{aligned} \frac{du\left(c_k^e\left(n,h,a,a_k^{e*}(n,h,a)\right)\right)}{dc} &= \beta\left(1-\delta_{k,n}\right)\frac{\partial V_k^e\left(n+1,h_k'(h,1),a_k^{e*}(n,h,a)\right)}{\partial a} \\ &\quad + \beta\delta_{k,n}\frac{\partial V_k^s\left(n+1,0,h_k'(h,1),a_k^{e*}(n,h,a)\right)}{\partial a} \end{aligned} \tag{1.B.1}$$

The maximum operator in the value function (1.6) can be dropped by substituting the solution to the maximization problem. The value function is then given by

$$\begin{aligned} V_k^e(n,h,a) &= u\left(c_k^e\left(n,h,a,a_k^{e*}(n,h,a)\right)\right) \\ &\quad + \beta\left(1-\delta_{k,n}\right)V_k^e\left(n+1,h_k'(h,1),a_k^{e*}(n,h,a)\right) \\ &\quad + \beta\delta_{k,n}V_k^s\left(n+1,0,h_k'(h,1),a_k^{e*}(n,h,a)\right) \end{aligned} \tag{1.B.2}$$

Taking derivatives w.r.t. a yields

$$
\begin{aligned}
\frac{\partial V_k^e(n,h,a)}{\partial a} &= \frac{\partial c_k^e\left(n,h,a,a_k^{e*}(n,h,a)\right)}{\partial a}\frac{du\left(c_k^e\left(n,h,a,a_k^{e*}(n,h,a)\right)\right)}{dc} \\
&+ \frac{\partial c_k^e\left(n,h,a,a_k^{e*}(n,h,a)\right)}{\partial a'}\frac{\partial a_k^{e*}(n,h,a)}{\partial a} \\
&\quad \frac{du\left(c_k^e\left(n,h,a,a_k^{e*}(n,h,a)\right)\right)}{dc} \\
&+ \beta(1-\delta_{k,n})\frac{\partial a_k^{e*}(n,h,a)}{\partial a}\frac{\partial V_k^e\left(n+1,h_k'(h,1),a_k^{e*}(n,h,a)\right)}{\partial a'} \\
&+ \beta\delta_{k,n}\frac{\partial a_k^{e*}(n,h,a)}{\partial a}\frac{\partial V_k^s\left(n+1,0,h_k'(h,1),a_k^{e*}(n,h,a)\right)}{\partial a'} \\
&= \left[1+(1-\tau^I)r\right]\frac{du\left(c_k^e\left(n,h,a,a_k^{e*}(n,h,a)\right)\right)}{dc} \\
&- \frac{\partial a_k^{e*}(n,h,a)}{\partial a}\frac{du\left(c_k^e\left(n,h,a,a_k^{e*}(n,h,a)\right)\right)}{dc} \\
&+ \frac{\partial a_k^{e*}(n,h,a)}{\partial a}\beta(1-\delta_{k,n})\frac{\partial V_k^e\left(n+1,h_k'(h,1),a_k^{e*}(n,h,a)\right)}{\partial a'} \\
&+ \frac{\partial a_k^{e*}(n,h,a)}{\partial a}\beta\delta_{k,n}\frac{\partial V_k^s\left(n+1,0,h_k'(h,1),a_k^{e*}(n,h,a)\right)}{\partial a'} \\
&= \left[1+(1-\tau^I)r\right]\frac{du\left(c_k^{e*}(n,h,a)\right)}{dc}
\end{aligned}
\tag{1.B.3}
$$

where the last step uses the envelope condition (1.B.1). Analogously, again making use of the respective envelope conditions on optimal asset policies, the partial derivative of the value function for unemployed workers is

$$
\frac{\partial V_k^u(n,m,h,a)}{\partial a} = \left[1+(1-\tau^I)r\right]\frac{du\left(c_k^{u*}(n,m,h,a)\right)}{dc} \tag{1.B.4}
$$

Now, let $s_k^*(n,m,h,a)$ solve the maximization problem in equation (1.7). This implies that, for interior solutions, $s_k^*(n,m,h,a)$ satisfies

$$
\frac{d\psi\left(1-s_k^*(n,m,h,a)\right)}{d(1-s)} = \alpha^{-1}\frac{d\zeta_k\left(s_k^*(n,m,h,a)\right)}{ds}\left[V_k^e(n,h,a)-V_k^u(n,m+1,h,a)\right] \tag{1.B.5}
$$

Again substituting the solution for the maximum operator, the value function for searching workers (1.7) becomes

$$
\begin{aligned}
V_k^s(n,m,h,a) &= \alpha\psi\left(1-s_k^*(n,m,h,a)\right) \\
&+ \zeta_k\left(s_k^*(n,m,h,a)\right)V_k^e(n,h,a) \\
&+ \left[1-\zeta_k\left(s_k^*(n,m,h,a)\right)\right]V_k^u(n,m+1,h,a)
\end{aligned}
\tag{1.B.6}
$$

Taking derivatives w.r.t. current asset holdings a and employing the envelope condition (1.B.5) yields

$$\begin{aligned}\frac{\partial V_k^s(n,m,h,a)}{\partial a} &= \zeta_k\left(s_k^*(n,m,h,a)\right)\frac{\partial V_k^e(n,h,a)}{\partial a}\\ &\quad+\left[1-\zeta_k\left(s_k^*(n,m,h,a)\right)\right]\frac{\partial V_k^u(n,m+1,h,a)}{\partial a}\end{aligned} \tag{1.B.7}$$

With these expressions for the derivatives of the value functions, it is straightforward to compute the first order conditions for interior solutions to the consumption/ savings choice. Imposing the optimality conditions defining $a_k^{e*}(\cdot)$, $a_k^{u*}(\cdot)$ and $s_k^*(\cdot)$ for all periods, the first-order condition for employed workers becomes

$$\begin{aligned}\frac{du\left(c_k^{e*}(n,h,a)\right)}{dc} &= \beta\left(1-\delta_{k,n}\right)\frac{\partial V_k^e\left(n+1,h_k'(h,1),a_k^{e*}(n,h,a)\right)}{\partial a}\\ &\quad+\beta\delta_{k,n}\frac{\partial V_k^s\left(n+1,0,h_k'(h,1),a_k^{e*}(n,h,a)\right)}{\partial a}\end{aligned}$$

Substituting (1.B.3), (1.B.4), and (1.B.7) yields

$$\begin{aligned}\frac{du\left(c_k^{e*}(n,h,a)\right)}{dc} &= \beta\left[1+(1-\tau^I)r\right]\Bigg\{\\ &\left(1-\delta_{k,n}\right)\frac{du\left(c_k^{e*}\left(n+1,h_k'(h,1),a_k^{e*}(n,h,a)\right)\right)}{dc}\\ &+\delta_{k,n}\Bigg[\zeta_k\left(s_k^*\left(n+1,0,h_k'(h,1),a_k^{e*}(n,h,a)\right)\right)\\ &\qquad\frac{du\left(c_k^{e*}\left(n+1,h_k'(h,1),a_k^{e*}(n,h,a)\right)\right)}{dc}\\ &\qquad+\left[1-\zeta_k\left(s_k^*\left(n+1,0,h_k'(h,1),a_k^{e*}(n,h,a)\right)\right)\right]\\ &\qquad\frac{du\left(c_k^{u*}\left(n+1,1,h_k'(h,1),a_k^{e*}(n,h,a)\right)\right)}{dc}\Bigg]\Bigg\}\end{aligned} \tag{1.B.8}$$

Applying the functional form assumed for utility, i.e. substituting $\frac{du(c)}{dc}=c^{-\sigma^c}$ and $\left(\frac{du}{dc}\right)^{-1}(y)=y^{-\frac{1}{\sigma^c}}$, then yields

$$
\begin{aligned}
c_k^{e*}(n,h,a) \;=\; & \left\{\beta\left[1+(1-\tau^I)r\right]\right\}^{-\frac{1}{\sigma^c}} \Bigg\{ \\
& \left(1-\delta_{k,n}\right) c_k^{e*}\left(n+1, h_k'(h,1), a_k^{e*}(n,h,a)\right)^{-\sigma^c} \\
& + \delta_{k,n}\Bigg[\zeta_k\left(s_k^*\left(n+1,0,h_k'(h,1),a_k^{e*}(n,h,a)\right)\right) \\
& \quad c_k^{e*}\left(n+1, h_k'(h,1), a_k^{e*}(n,h,a)\right)^{-\sigma^c} \\
& \quad + \left[1-\zeta_k\left(s_k^*\left(n+1,0,h_k'(h,1),a_k^{e*}(n,h,a)\right)\right)\right] \\
& \quad c_k^{u*}\left(n+1,1,h_k'(h,1),a_k^{e*}(n,h,a),1\right)^{-\sigma^c}\Bigg]\Bigg\}^{-\frac{1}{\sigma^c}}
\end{aligned}
\tag{1.B.9}
$$

The first-order condition for consumption of unemployed workers is derived analogously. It is given by

$$
\begin{aligned}
c_k^{u*}(n,m,h,a) \;=\; & \left\{\beta(1+(1-\tau^I)r)\right\}^{-\frac{1}{\sigma^c}} \Bigg\{ \\
& \zeta_k\left(s_k^*\left(n+1,m,h_k'(h,0),a_k^{u*}(n,m,h,a)\right)\right) \\
& \quad c_k^{e*}\left(n+1,h_k'(h,0),a_k^{u*}(n,m,h,a)\right)^{-\sigma^c} \\
& + \left[1-\zeta_k\left(s_k^*\left(n+1,m,h_k'(h,0),a_k^{u*}(n,m,h,a)\right)\right)\right] \\
& \quad c_k^{u*}\left(n+1,m+1h_k'(h,0),a_k^{u*}(n,h,a)\right)^{-\sigma^c}\Bigg\}^{-\frac{1}{\sigma^c}}
\end{aligned}
\tag{1.B.10}
$$

Finally, solving the envelope condition (1.B.5) for search effort s_k^* yields the first-order condition

$$
\begin{aligned}
s_k^*(n,m,h,a) \;&=\; 1-\left(\frac{d\psi}{d(1-s)}\right)^{-1}(z) \\
\text{with } z \;&=\; \alpha^{-1}\frac{d\zeta_k\left(s_k^*(n,m,h,a)\right)}{ds}\left[V_k^e(n,h,a)-V_k^u(n,m+1,h,a)\right]
\end{aligned}
$$

where $\left(\frac{d\psi}{d(1-s)}\right)^{-1}(\cdot)$ is the inverse of the derivative of the leisure utility function w.r.t. leisure. Thus, the first-order condition for search effort is

$$
s_k^*(n,m,h,a) \;=\; 1-\alpha^{\frac{1}{\sigma^l}}\left\{\frac{d\zeta_k\left(s_k^*(n,m,h,a)\right)}{ds}\left[V_k^e(n,h,a)-V_k^u(n,m+1,h,a)\right]\right\}^{-\frac{1}{\sigma^l}}
\tag{1.B.11}
$$

With linear search technology, $\frac{d\zeta_k(s_k^*(n,h,a))}{ds}$ is a constant and the FOC (1.B.11) provides a closed form solution for optimal search effort.

With all first order conditions in place, it is now straightforward to solve the household optimization problem by backwards induction. As mentioned before, retired households optimally consume their retirement pension income plus the annuity of their asset holdings. Value functions, consumption/ saving policy functions and search effort policy functions for working age households can be obtained by iterating over the first-order conditions (1.B.9), (1.B.9), and (1.B.11) for all working age periods.

Solving for equilibrium tax policies. Once the policy functions have been derived, the government budget constraints can be checked by computing expected present values of net cost to the government at model entry. Since there are separate government programs that are individually required to have a balanced budget, this computation needs to be conducted for all programs separately. For this, denote the present value of net cost to the government under program P of an employed worker of type k, age n, human capital level h, and asset holdings a by $C_k^{P,e}(n,h,a)$. Let $C_k^{P,u}(n,m,h,a)$ denote the same quantity for unemployed workers with m periods of unemployment. Given the optimal policies derived above, these present value cost functions can then be expressed recursively for all working age periods for employed workers by

$$
\begin{aligned}
C_k^{P,e}(n,h,a) \;=\; & \pi_k^{P,e}(n,h,a) \\
& + \frac{1}{1+r}\left(1-\delta_{k,n}\right) C_k^{P,e}\left(n+1, h_k'(h,1), a_k^{e*}(n,h,a)\right) \\
& + \frac{1}{1+r}\delta_{k,n}\Big[\zeta_k\left(s_k^*\left(n+1,0,h_k'(h,1),a_k^{e*}(n,h,a)\right)\right) \\
& \qquad C_k^{P,e}\left(n+1,h_k'(h,1),a_k^{e*}(n,h,a)\right) \\
& \qquad + \left[1-\zeta_k\left(s_k^*\left(n+1,0,h_k'(h,1),a_k^{e*}(n,h,a)\right)\right)\right] \\
& \qquad C_k^{P,u}\left(n+1,1,h_k'(h,1),a_k^{e*}(n,h,a)\right)\Big]
\end{aligned}
\tag{1.B.12}
$$

where $\pi_k^{P,e}(n,h,a)$ is the net per-period cost to the government under program P of an employed worker of type k. For unemployed workers, present value cost functions are given by

$$
\begin{aligned}
C_k^{P,u}(n,m,h,a) \;=\; & \pi_k^{P,u}(n,m,h,a) \\
& + \frac{1}{1+r}\Big[\zeta_k\left(s_k^*\left(n+1,m,h_k'(h,0),a_k^{u*}(n,h,a)\right)\right) \\
& \qquad C_k^{P,e}\left(n+1,h_k'(h,0),a_k^{u*}(n,h,a)\right) \\
& \qquad + \left[1-\zeta_k\left(s_k^*\left(n+1,m,h_k'(h,0),a_k^{u*}(n,h,a)\right)\right)\right] \\
& \qquad C_k^{P,u}\left(n+1,m+1,h_k'(h,0),a_k^{u*}(n,h,a)\right)\Big]
\end{aligned}
\tag{1.B.13}
$$

where $\pi_k^{P,u}(n,m,h,a)$ is the net per-period cost to the government under program P of an unemployed worker of type k with unemployment duration m.

The expected discounted cost under program P for workers at model entry, i.e. in the searching stage at period zero, are given by

$$\begin{aligned} C_k^{P,0} &= \zeta_k\left(s_k^*\left(0,0,h_k^0,a_k^0\right)\right)C_k^{P,e}\left(0,h_k^0,a_k^0\right) \\ &\quad + \left[1-\zeta_k\left(s_k^*\left(0,0,h_k^0,a_k^0\right)\right)\right]C_k^{P,u}\left(0,1,h_k^0,a_k^0\right) \end{aligned} \tag{1.B.14}$$

The government budget of program P is satisfied if the expected cost at model entry (i.e. prior to drawing the type) is zero:

$$\sum_{k\in K}\chi_k C_k^{P,0} = 0 \qquad \forall P \in \{UI,SS,I\} \tag{1.B.15}$$

The net per-period cost functions for the individual programs are

$$\pi_k^{UI,e}(n,h,a) = -\tau^{UI}\bar{\omega}h \qquad \text{and} \qquad \pi_k^{UI,u}(n,m,h,a) = b_k^{UI}(n,m,h)$$

$$\pi_k^{SS,e}(n,h,a) = \pi_k^{SS,u}(n,m,h,a) = \begin{cases} -\tau^{SS}\bar{\omega}h & \text{if } n \le \bar{n}_w \\ b_k^{SS} & \text{if } n > \bar{n}_w \end{cases}$$

$$\pi_k^{I,e}(n,h,a) = \pi_k^{I,u}(n,m,h,a) = \begin{cases} T-\tau^I\left(\bar{\omega}h+ra\right) & \text{if } n \le \bar{n}_w \\ T-\tau^I ra & \text{if } n > \bar{n}_w \end{cases}$$

Note that, during retirement, workers of all types receive pension benefits b^{SS} for all remaining periods of life in addition to the lump-sum transfers T that all agents receive in all periods. Retired workers only pay income tax on their asset income. Cost from pensions and lump-sum transfers are constant streams and can be summarized in net present value terms (discounted to the first period of retirement) by the discount factor $\beta^{r,c}$ given by

$$\beta^{r,c} = \frac{(1+r)^{\bar{n}_r}-1}{r(1+r)^{\bar{n}_r-1}}$$

Revenues from taxing asset income represent interest payments on an amortized loan and can be summarized in net present value terms (again discounted to the first period of retirement) by the discount factor $\beta^{r,a}$ given by

$$\beta^{r,a} = \frac{\left(1-\tau^I\right)\left[1+\left(1-\tau^I\right)r\right]^{\bar{n}_r} + (1+r)^{\bar{n}_r}\left\{\tau^I\left[1+(1-\tau^I)r\right]^{\bar{n}_r}-1\right\}}{(1+r)^{\bar{n}_r-1}\left\{\left[1+(1-\tau^I)r\right]^{\bar{n}_r}-1\right\}}$$

The present value of net cost to the government for retired workers, discounted to the first period of retirement, is then given by

$$\begin{aligned} C_k^{UI,e}\left(\bar{n}_w+1,h,a\right) &= C_k^{UI,u}\left(\bar{n}_w+1,m,h,a\right) = 0 \\ C_k^{SS,e}\left(\bar{n}_w+1,h,a\right) &= C_k^{SS,u}\left(\bar{n}_w+1,m,h,a\right) = \beta^{r,c}b^{SS} \\ C_k^{I,e}\left(\bar{n}_w+1,h,a\right) &= C_k^{I,u}\left(\bar{n}_w+1,m,h,a\right) = \beta^{r,c}T-\beta^{r,a}a \end{aligned} \tag{1.B.16}$$

With expressions (1.B.12)–(1.B.16), period zero cost functions for employed and unemployed workers can be obtained by backwards induction. For a given UI policy choice, the model is solved by (i) guessing equilibrium parameters of the tax functions, (ii) deriving optimal policies and present value cost functions, (iii) checking the government budget constraint and, if necessary, adjusting the guess for the income tax parameters until a pre-specified tolerance for the budget condition is met.

1.B.2 Consumption Equivalents

As mentioned before, consumption equivalents in this framework are defined as the percentage change in consumption in every state and period required to make workers as well off under the baseline economy as under an alternative calibration, keeping leisure utility constant. The exercise therefore consists in (i) isolating the utility from consumption only in the baseline scenario and (ii) computing the required proportional change in consumption to equate total utility in both scenarios.

To formalize the concept, consider a given set of policy functions for consumption and search effort. Recall that $V_k^e(n,h,a)$ denotes the present value of total utility for an employed worker of type k, age n, human capital level h, and asset holdings a given these policy functions. Denote by $U_k^e(n,h,a)$ and $L_k^e(n,h,a)$ the present values of consumption utility and leisure utility, respectively, for the same worker and the same set of policies. By definition,

$$V_k^e(n,h,a) \;=\; U_k^e(n,h,a) + L_k^e(n,h,a)$$

Analogously, let $U_k^u(n,m,h,a)$, $L_k^u(n,m,h,a)$, $U_k^s(n,m,h,a)$, and $L_k^s(n,m,h,a)$ be the present values of consumption utility and leisure utility for unemployed and searching workers with an unemployment spell of m periods. Now, consider workers that exert the same search effort as before, but consume η times the consumption level of the above allocation in all states and periods. Denote the present value at entry of consumption utility for this worker by $\tilde{U}_k^s(\eta)$. Note that, by definition, $\tilde{U}_k^s(1) = U_k^s\left(0,0,h_k^0,a_k^0\right)$. Also note that, due to the CRRA functional form, the adjusted consumption utility in any period is given by

$$u(\eta c) \;=\; \frac{(\eta c)^{1-\sigma^c}}{1-\sigma^c} \;=\; \eta^{1-\sigma^c}\frac{c^{1-\sigma^c}}{1-\sigma^c} \;=\; \eta^{1-\sigma^c}u(c)$$

Since all utility terms in the present value of consumption utility exhibit this property, the multiplier can be drawn out of the expectation:

$$\tilde{U}_k^s(\eta) \;=\; \eta^{1-\sigma^c}\tilde{U}_k^s(1) \;=\; \eta^{1-\sigma^c}U_k^s\left(0,0,h_k^0,a_k^0\right)$$

Now, let $V_k^{s,base}\left(0,0,h_k^0,a_k^0\right)$, $U_k^{s,base}\left(0,0,h_k^0,a_k^0\right)$, and $L_k^{s,base}\left(0,0,h_k^0,a_k^0\right)$ be the present values at entry of a searching worker of type k with initial human capital

h_k^0 and initial wealth a_k^0 for the quantities total utility, consumption utility, and leisure utility, respectively. Let $\tilde{U}_k^{s,base}(\eta)$ denote the expected discounted value of utility from consuming η times the consumption under the baseline scenario. Let $V_k^{s,A}(0,0,h_k^0,a_k^0)$ be the net present value of total utility at entry of some alternative scenario (A). The consumption multiplier η^A that corresponds to this alternative scenario (A) is then defined by the condition

$$\begin{aligned} V_k^{s,A}\left(0,0,h_k^0,a_k^0\right) &= \tilde{U}_k^{s,base}\left(\eta^A\right) + L_k^{s,base}\left(0,0,h_k^0,a_k^0\right) \\ &= \left(\eta^A\right)^{1-\sigma^c} U_k^{s,base}\left(0,h_k^0,a_k^0\right) + L_k^{s,base}\left(0,h_k^0,a_k^0\right) \end{aligned} \tag{1.B.17}$$

The welfare gain of the alternative scenario relative to the baseline scenario is captured by the difference in expected present value at entry. Formally,

$$\begin{aligned} \Delta V_k^{A,base} &= V_k^{s,A}\left(0,0,h_k^0,a_k^0\right) - V_k^{s,base}\left(0,0,h_k^0,a_k^0\right) \\ &= \tilde{U}_k^{s,base}\left(\eta^A\right) + L_k^{s,base}\left(0,0,h_k^0,a_k^0\right) - V_k^{s,base}\left(0,0,h_k^0,a_k^0\right) \\ &= \left(\eta^A\right)^{1-\sigma^c} U_k^{s,base}\left(0,0,h_k^0,a_k^0\right) + L_k^{s,base}\left(0,0,h_k^0,a_k^0\right) - V_k^{s,base}\left(0,0,h_k^0,a_k^0\right) \\ &= \left(\eta^A\right)^{1-\sigma^c} U_k^{s,base}\left(0,0,h_k^0,a_k^0\right) - U_k^{s,base}\left(0,0,h_k^0,a_k^0\right) \\ &= \left[\left(\eta^A\right)^{1-\sigma^c} - 1\right] U_k^{s,base}\left(0,0,h_k^0,a_k^0\right) \end{aligned} \tag{1.B.18}$$

Denote by $\eta\left(\Delta V_k^{A,base}\right)$ the multiplier that solves (1.B.18) for a given alternative scenario (A) and a given baseline scenario. The consumption equivalent for workers of type k of the welfare change from implementing alternative scenario (A) instead of the baseline scenario is defined as the percentage change in consumption in all states and periods in the baseline scenario that is required to make the worker as well off as under the alternative scenario. Formally, with $CE\left(\Delta V_k^{A,base}\right) = \eta\left(\Delta V_k^{A,base}\right) - 1$,

$$CE\left(\Delta V_k^{A,base}\right) = \left[\frac{\Delta V_k^{A,base}}{U_k^{s,base}\left(0,0,h_k^0,a_k^0\right)} + 1\right]^{\frac{1}{1-\sigma^c}} - 1 \tag{1.B.19}$$

Thus, the only quantities required to compute consumption equivalents are the difference in present value at model entry between the alternative scenario and the baseline scenario and the present value at entry of consumption utility in the baseline scenario.

Solving for consumption utility in a base scenario. To compute consumption utility in the base scenario, consider the set of optimal policies

$$\left\{\left\{c_k^{e*}(n,h,a), \left\{c_k^{u*}(n,m,h,a)\right\}_{m=1}^{\bar{m}}, \left\{s_k^{*}(n,m,h,a)\right\}_{m=0}^{\bar{m}}\right\}_{n=1}^{\bar{n}_w}\right\}_{k\in K} \tag{1.B.20}$$

Consumption policies and the respective budget constraints can then be used to derive optimal savings policies

$$\left\{\left\{a_k^{e*}(n,h,a),\left\{a_k^{u*}(n,m,h,a)\right\}_{m=1}^{\bar{m}}\right\}_{n=0}^{\bar{n}_w}\right\}_{k\in K} \tag{1.B.21}$$

With the complete set of optimal policies from the baseline scenario, the present value of consumption utility by state, type, age, human capital level, asset level, and unemployment duration can be expressed recursively by the following set of equations for employed workers

$$\begin{aligned}
U_k^e(n,h,a) \;=\;& u\left(c_k^{e*}(n,h,a)\right) \\
&+\beta\left(1-\delta_{k,n}\right)U_k^e\left(n+1,h_k'(h,1),a_k^{e*}(n,h,a)\right) \\
&+\beta\delta_{k,n}\Big[\zeta_k\left(s_k^*\left(n+1,0,h_k'(h,1),a_k^{e*}(n,h,a)\right)\right) \\
&\qquad U_k^e\left(n+1,h_k'(h,1),a_k^{e*}(n,h,a)\right) \\
&\quad+\left[1-\zeta_k\left(s_k^*\left(n+1,0,h_k'(h,1),a_k^{e*}(n,h,a)\right)\right)\right] \\
&\qquad U_k^u\left(n+1,1,h_k'(h,1),a_k^{e*}(n,h,a)\right)\Big]
\end{aligned} \tag{1.B.22}$$

and for unemployed workers

$$\begin{aligned}
U_k^u(n,m,h,a) \;=\;& u\left(c_k^{u*}(n,m,h,a)\right) \\
&+\beta\zeta_k\left(s_k^*\left(n+1,m,h_k'(h,0),a_k^{u*}(n,h,a)\right)\right) \\
&\qquad U_k^e\left(n+1,h_k'(h,0),a_k^{u*}(n,h,a)\right) \\
&+\beta\left[1-\zeta_k\left(s_k^*\left(n+1,m,h_k'(h,0),a_k^{u*}(n,h,a)\right)\right)\right] \\
&\qquad U_k^u\left(n+1,m+1,h_k'(h,0),a_k^{u*}(n,h,a)\right)
\end{aligned} \tag{1.B.23}$$

Consumption in retirement is constant and consists of government transfers and the annuity income of asset holdings at the beginning of retirement

$$c^r(a) \;=\; b^{SS}+T+\frac{\left(1-\tau^I\right)r\left[1+\left(1-\tau^I\right)r\right]^{\bar{n}_r}}{\left[1+(1-\tau^I)\,r\right]^{\bar{n}r}-1}a \tag{1.B.24}$$

During retirement, utility is entirely generated through consumption. Therefore,

$$U_k^e\left(\bar{n}_w+1,h,a\right) \;=\; U_k^u\left(\bar{n}_w+1,m,h,a\right) \;=\; \frac{1-\beta^{\bar{n}_r}}{1-\beta}u\left(c_k^r(h,a)\right) \tag{1.B.25}$$

Starting from equation (1.B.25) and solving backwards using equations (1.B.22) and (1.B.23) then yields $U_k^{s,base}\left(0,0,h_k^0,a_k^0\right)$.

Appendix 1.C Model Calibration

1.C.1 Construction of Calibration Targets

Relative wages. The construction of wage targets for the calibration follows a multistep procedure. First, I construct population-wide age profiles of relative wages. For this, I estimate the following regression equation

$$\log(\text{earnings}_i) \;=\; \beta^0 + \sum_{j=2}^{J} \beta^j \mathbb{1}_{\{\text{AgeGroup}_i = j\}} + \gamma \boldsymbol{X}_i + \delta^s + \theta^t + \varepsilon_i$$

where the base age group ($j = 1$) is '20' (years of age), and subsequent age groups are '21 to 24', seven 5-year brackets and '60 and older'. The vector $\boldsymbol{X}_i$ contains information on educational attainment, race, and marital status, and δ^s and θ^t are state and year fixed effects, respectively. Population-wide average relative wages (relative to the wage at age 20) are then obtained as $\bar{\omega}^j = \exp\left(\hat{\beta}^j\right)$.

Next, I estimate relative average wages (relative to the base category) by age group and education group. Since college education may not yet be completed at younger ages, I only use data from age 25 onward. Specifically, I estimate

$$\log(\text{earnings}_i) \;=\; \tilde{\beta}_0 + \sum_{j=3}^{J}\sum_{k=2}^{K} \beta_k^j \mathbb{1}_{\{\text{AgeGroup}_i = j\}} \times \mathbb{1}_{\{\text{Education}_i = k\}} + \tilde{\gamma}\tilde{\boldsymbol{X}}_i + \tilde{\delta}^s + \tilde{\theta}^t + \tilde{\varepsilon}_i$$

where age groups are as above (dropping the first two categories) and the base category is low education workers aged between 25 and 29. The vector $\tilde{\boldsymbol{X}}_i$ now only contains information on race and marital status, and $\tilde{\delta}^s$ and $\tilde{\theta}^t$ are again state and year fixed effects. Average wages by age group and education type (relative to the base category) are then obtained as $\tilde{\omega}_k^j = \exp\left(\hat{\beta}_k^j\right)$ with $\tilde{\omega}_{low}^{\text{'25 to 29'}} = 1.0$. The missing relative wages for age groups '20' and '21 to 24' are then computed as

$$\tilde{\omega}_k^j \;=\; \tilde{\omega}_k^{\text{'25 to 29'}} \frac{\bar{\omega}^j}{\bar{\omega}^{\text{'25 to 29'}}} \quad \text{for } j \in \{\text{'20'}, \text{'21 to 24'}\}$$

The underlying assumption is that wage growth from groups '20' and '21 to 24' to group '25 to 29' is the same across education types and equal to the population-average growth rate for these groups.

Finally, relative wages by age group and education are rescaled such that the average relative wage at age 20 equals one:

$$\omega_k^j \;=\; \frac{\tilde{\omega}_k^j}{\sum_{k \in K} \chi_k \tilde{\omega}_k^{\text{'20'}}}$$

where χ_k represents the population share of education group k. This treatment ensures that relative wages for all education groups are computed consistently on the basis of reliable data.

1.C.2 Calibration Tables

Table 1.C.1. Baseline calibration

Parameter	Definition	Value		
		Low	Medium	High
$\bar{n}_w$	Working periods		180	
$\bar{n}_r$	Retirement periods		80	
β	Discount factor		0.99	
σ^c	Risk aversion coefficient for consumption		2.0	
σ^l	Risk aversion coefficient for leisure		2.0	
α	Weight of leisure utility		1.0	
χ_k	Type share of population	0.11	0.58	0.31
$\bar{\omega}$	Average wage level		1.00	
$h'(h,w)$	Human capital production technologies			
h_k^k	Initial human capital level	0.70	0.90	1.10
λ_k	Learning ability parameter	0.03	0.04	0.06
ϕ_k	Human capital curvature parameter	0.10	0.10	0.10
δ_k^h	Human capital depreciation rate	0.03	0.03	0.03
$\zeta_k(s)$	Job search technologies			
γ_k	Search technology slope parameter	1.00	1.01	1.09
μ_k	Search technology intercept	0.14	0.12	0.08
$\delta_{k,n}$	Separation rate by age (years)			
	$n = 10$ (22.5)	0.079	0.038	0.021
	$n = 30$ (27.5)	0.063	0.033	0.013
	$n = 50$ (32.5)	0.058	0.030	0.012
	$n = 70$ (37.5)	0.055	0.028	0.012
	$n = 90$ (42.5)	0.050	0.026	0.013
	$n = 110$ (47.5)	0.048	0.025	0.013
	$n = 130$ (52.5)	0.043	0.024	0.014
	$n = 150$ (57.5)	0.039	0.024	0.016
	$n = 170$ (62.5)	0.034	0.023	0.017
b^{SS}	Retirement pensions	0.673	0.673	0.673
T	Lumpsum transfers	0.203	0.203	0.203
$\underline{a}$	Borrowing constraint		-1.12	
τ^{UI}	Unemployment insurance tax rate	0.014	0.014	0.014
τ^{SS}	Social security tax rate	0.050	0.050	0.050
τ^I	General income tax rate	0.150	0.150	0.150
$\bar{m}$	Maximum benefit duration in periods		4	
$\rho_{k,n}$	UI replacement rate by age (years)			
	$n = 0$ (20)	0.50	0.50	0.50
	$n = 45$ (31.25)	0.50	0.50	0.50
	$n = 90$ (42.5)	0.50	0.50	0.50
	$n = 135$ (53.75)	0.50	0.50	0.50
	$n = 180$ (65)	0.50	0.50	0.50
$\underline{b}$	UI floor		0.00	
$\bar{b}$	UI cap		inf	

Notes: Model parameters in the baseline setup. For details on the calibration, see Section 1.5.1. The functions $\delta_{k,n}$ and $\rho_{k,n}$ are cubic splines through values in the table.

Table 1.C.2. Exercise (A): optimal common and constant replacement rate

Parameter	Definition	Value		
		Low	Medium	High
$\bar{n}_w$	Working periods		180	
$\bar{n}_r$	Retirement periods		80	
β	Discount factor		0.99	
σ^c	Risk aversion coefficient for consumption		2.0	
σ^l	Risk aversion coefficient for leisure		2.0	
α	Weight of leisure utility		1.0	
χ_k	Type share of population	0.11	0.58	0.31
$\bar{\omega}$	Average wage level		1.00	
$h'(h,w)$	Human capital production technologies			
h_k^k	Initial human capital level	0.70	0.90	1.10
λ_k	Learning ability parameter	0.03	0.04	0.06
ϕ_k	Human capital curvature parameter	0.10	0.10	0.10
δ_k^h	Human capital depreciation rate	0.03	0.03	0.03
$\zeta_k(s)$	Job search technologies			
γ_k	Search technology slope parameter	1.00	1.01	1.09
μ_k	Search technology intercept	0.14	0.12	0.08
$\delta_{k,n}$	Separation rate by age (years)			
	$n = 10$ (22.5)	0.079	0.038	0.021
	$n = 30$ (27.5)	0.063	0.033	0.013
	$n = 50$ (32.5)	0.058	0.030	0.012
	$n = 70$ (37.5)	0.055	0.028	0.012
	$n = 90$ (42.5)	0.050	0.026	0.013
	$n = 110$ (47.5)	0.048	0.025	0.013
	$n = 130$ (52.5)	0.043	0.024	0.014
	$n = 150$ (57.5)	0.039	0.024	0.016
	$n = 170$ (62.5)	0.034	0.023	0.017
b^{SS}	Retirement pensions	0.673	0.673	0.673
T	Lumpsum transfers	0.203	0.203	0.203
$\underline{a}$	Borrowing constraint		-1.12	
τ^{UI}	Unemployment insurance tax rate	0.014	0.014	0.014
τ^{SS}	Social security tax rate	0.050	0.050	0.050
τ^I	General income tax rate	0.150	0.150	0.150
$\bar{m}$	Maximum benefit duration in periods		4	
$\rho_{k,n}$	UI replacement rate by age (years)			
	$n = 0$ (20)	0.44	0.44	0.44
	$n = 45$ (31.25)	0.44	0.44	0.44
	$n = 90$ (42.5)	0.44	0.44	0.44
	$n = 135$ (53.75)	0.44	0.44	0.44
	$n = 180$ (65)	0.44	0.44	0.44
$\underline{b}$	UI floor		0.00	
$\bar{b}$	UI cap		inf	

Notes: Model parameters in the calibration with optimimized policy parameters under combined government program budgets (exercise A). Policy parameter is the UI replacement rate (common across types and constant over worker age). Policy parameters are obtained by solving (1.16) with parameter space Θ^A; remaining parameters as in baseline calibration. The functions $\delta_{k,n}$ and $\rho_{k,n}$ are cubic splines through values in the table.

Table 1.C.3. Exercise (B): optimal age-dependent replacement rate

Parameter	Definition	Value		
		Low	Medium	High
$\bar{n}_w$	Working periods		180	
$\bar{n}_r$	Retirement periods		80	
β	Discount factor		0.99	
σ^c	Risk aversion coefficient for consumption		2.0	
σ^l	Risk aversion coefficient for leisure		2.0	
α	Weight of leisure utility		1.0	
χ_k	Type share of population	0.11	0.58	0.31
$\bar{\omega}$	Average wage level		1.00	
$h'(h,w)$	Human capital production technologies			
h_k^k	Initial human capital level	0.70	0.90	1.10
λ_k	Learning ability parameter	0.03	0.04	0.06
ϕ_k	Human capital curvature parameter	0.10	0.10	0.10
δ_k^h	Human capital depreciation rate	0.03	0.03	0.03
$\zeta_k(s)$	Job search technologies			
γ_k	Search technology slope parameter	1.00	1.01	1.09
μ_k	Search technology intercept	0.14	0.12	0.08
$\delta_{k,n}$	Separation rate by age (years)			
	$n = 10$ (22.5)	0.079	0.038	0.021
	$n = 30$ (27.5)	0.063	0.033	0.013
	$n = 50$ (32.5)	0.058	0.030	0.012
	$n = 70$ (37.5)	0.055	0.028	0.012
	$n = 90$ (42.5)	0.050	0.026	0.013
	$n = 110$ (47.5)	0.048	0.025	0.013
	$n = 130$ (52.5)	0.043	0.024	0.014
	$n = 150$ (57.5)	0.039	0.024	0.016
	$n = 170$ (62.5)	0.034	0.023	0.017
b^{SS}	Retirement pensions	0.674	0.674	0.674
T	Lumpsum transfers	0.204	0.204	0.204
$\underline{a}$	Borrowing constraint		-1.12	
τ^{UI}	Unemployment insurance tax rate	0.011	0.011	0.011
τ^{SS}	Social security tax rate	0.050	0.050	0.050
τ^I	General income tax rate	0.150	0.150	0.150
$\bar{m}$	Maximum benefit duration in periods		4	
$\rho_{k,n}$	UI replacement rate by age (years)			
	$n = 0$ (20)	0.64	0.64	0.64
	$n = 45$ (31.25)	0.52	0.52	0.52
	$n = 90$ (42.5)	0.08	0.08	0.08
	$n = 135$ (53.75)	-0.01	-0.01	-0.01
	$n = 180$ (65)	0.05	0.05	0.05
$\underline{b}$	UI floor		0.00	
$\bar{b}$	UI cap		inf	

Notes: Model parameters in the calibration with optimimized policy parameters under combined government program budgets (exercise B). Policy parameters are replacement rates that are common across types and varying over worker age. Policy parameters are obtained by solving (1.16) with parameter space Θ^B; remaining parameters as in baseline calibration. The functions $\delta_{k,n}$ and $\rho_{k,n}$ are cubic splines through values in the table.

Table 1.C.4. Exercise (C): optimal age-and-type-dependent replacement rate

		Value		
Parameter	Definition	Low	Medium	High
$\bar{n}_w$	Working periods		180	
$\bar{n}_r$	Retirement periods		80	
β	Discount factor		0.99	
σ^c	Risk aversion coefficient for consumption		2.0	
σ^l	Risk aversion coefficient for leisure		2.0	
α	Weight of leisure utility		1.0	
χ_k	Type share of population	0.11	0.58	0.31
$\bar{\omega}$	Average wage level		1.00	
$h'(h,w)$	Human capital production technologies			
h_k^k	Initial human capital level	0.70	0.90	1.10
λ_k	Learning ability parameter	0.03	0.04	0.06
ϕ_k	Human capital curvature parameter	0.10	0.10	0.10
δ_k^h	Human capital depreciation rate	0.03	0.03	0.03
$\zeta_k(s)$	Job search technologies			
γ_k	Search technology slope parameter	1.00	1.01	1.09
μ_k	Search technology intercept	0.14	0.12	0.08
$\delta_{k,n}$	Separation rate by age (years)			
	$n=10$ (22.5)	0.079	0.038	0.021
	$n=30$ (27.5)	0.063	0.033	0.013
	$n=50$ (32.5)	0.058	0.030	0.012
	$n=70$ (37.5)	0.055	0.028	0.012
	$n=90$ (42.5)	0.050	0.026	0.013
	$n=110$ (47.5)	0.048	0.025	0.013
	$n=130$ (52.5)	0.043	0.024	0.014
	$n=150$ (57.5)	0.039	0.024	0.016
	$n=170$ (62.5)	0.034	0.023	0.017
b^{SS}	Retirement pensions	0.674	0.674	0.674
T	Lumpsum transfers	0.204	0.204	0.204
$\underline{a}$	Borrowing constraint		-1.12	
τ^{UI}	Unemployment insurance tax rate	0.012	0.012	0.012
τ^{SS}	Social security tax rate	0.050	0.050	0.050
τ^{I}	General income tax rate	0.150	0.150	0.150
$\bar{m}$	Maximum benefit duration in periods		4	
$\rho_{k,n}$	UI replacement rate by age (years)			
	$n=0$ (20)	1.02	0.69	0.45
	$n=45$ (31.25)	0.72	0.51	0.30
	$n=90$ (42.5)	0.62	0.09	-0.09
	$n=135$ (53.75)	0.51	-0.02	-0.02
	$n=180$ (65)	0.41	0.11	0.02
$\underline{b}$	UI floor		0.00	
$\bar{b}$	UI cap		inf	

Notes: Model parameters in the calibration with optimimized policy parameters under combined government program budgets (exercise C). Policy parameters are replacement rates that are varying across types and worker age. Policy parameters are obtained by solving (1.16) with parameter space Θ^C; remaining parameters as in baseline calibration. The functions $\delta_{k,n}$ and $\rho_{k,n}$ are cubic splines through values in the table.

Table 1.C.5. Exercise (D): optimal replacement rate, benefit floor and benefit cap

		Value		
Parameter	Definition	Low	Medium	High
$\bar{n}_w$	Working periods		180	
$\bar{n}_r$	Retirement periods		80	
β	Discount factor		0.99	
σ^c	Risk aversion coefficient for consumption		2.0	
σ^l	Risk aversion coefficient for leisure		2.0	
α	Weight of leisure utility		1.0	
χ_k	Type share of population	0.11	0.58	0.31
$\bar{\omega}$	Average wage level		1.00	
$h'(h,w)$	Human capital production technologies			
h_k^k	Initial human capital level	0.70	0.90	1.10
λ_k	Learning ability parameter	0.03	0.04	0.06
ϕ_k	Human capital curvature parameter	0.10	0.10	0.10
δ_k^h	Human capital depreciation rate	0.03	0.03	0.03
$\zeta_k(s)$	Job search technologies			
γ_k	Search technology slope parameter	1.00	1.01	1.09
μ_k	Search technology intercept	0.14	0.12	0.08
$\delta_{k,n}$	Separation rate by age (years)			
	$n = 10$ (22.5)	0.079	0.038	0.021
	$n = 30$ (27.5)	0.063	0.033	0.013
	$n = 50$ (32.5)	0.058	0.030	0.012
	$n = 70$ (37.5)	0.055	0.028	0.012
	$n = 90$ (42.5)	0.050	0.026	0.013
	$n = 110$ (47.5)	0.048	0.025	0.013
	$n = 130$ (52.5)	0.043	0.024	0.014
	$n = 150$ (57.5)	0.039	0.024	0.016
	$n = 170$ (62.5)	0.034	0.023	0.017
b^{SS}	Retirement pensions	0.673	0.673	0.673
T	Lumpsum transfers	0.203	0.203	0.203
$\underline{a}$	Borrowing constraint		-1.12	
τ^{UI}	Unemployment insurance tax rate	0.013	0.013	0.013
τ^{SS}	Social security tax rate	0.050	0.050	0.050
τ^I	General income tax rate	0.150	0.150	0.150
$\bar{m}$	Maximum benefit duration in periods		4	
$\rho_{k,n}$	UI replacement rate by age (years)			
	$n = 0$ (20)	0.22	0.22	0.22
	$n = 45$ (31.25)	0.22	0.22	0.22
	$n = 90$ (42.5)	0.22	0.22	0.22
	$n = 135$ (53.75)	0.22	0.22	0.22
	$n = 180$ (65)	0.22	0.22	0.22
$\underline{b}$	UI floor		0.56	
$\bar{b}$	UI cap		0.69	

Notes: Model parameters in the calibration with optimimized policy parameters under combined government program budgets (exercise D). Policy parameters are replacement rate, benefit floor, and benefit cap that that are common across types and constant over age. Policy parameters are obtained by solving (1.16) with parameter space Θ^D; remaining parameters as in baseline calibration. The functions $\delta_{k,n}$ and $\rho_{k,n}$ are cubic splines through values in the table.

Table 1.C.6. Exercise (C'): optimal age-and-type-dependent rates, fixed budgets

Parameter	Definition	Value		
		Low	Medium	High
$\bar{n}_w$	Working periods		180	
$\bar{n}_r$	Retirement periods		80	
β	Discount factor		0.99	
σ^c	Risk aversion coefficient for consumption		2.0	
σ^l	Risk aversion coefficient for leisure		2.0	
α	Weight of leisure utility		1.0	
χ_k	Type share of population	0.11	0.58	0.31
$\bar{\omega}$	Average wage level		1.00	
$h'(h,w)$	Human capital production technologies			
h_k^k	Initial human capital level	0.70	0.90	1.10
λ_k	Learning ability parameter	0.03	0.04	0.06
ϕ_k	Human capital curvature parameter	0.10	0.10	0.10
δ_k^h	Human capital depreciation rate	0.03	0.03	0.03
$\zeta_k(s)$	Job search technologies			
γ_k	Search technology slope parameter	1.00	1.01	1.09
μ_k	Search technology intercept	0.14	0.12	0.08
$\delta_{k,n}$	Separation rate by age (years)			
	$n = 10$ (22.5)	0.079	0.038	0.021
	$n = 30$ (27.5)	0.063	0.033	0.013
	$n = 50$ (32.5)	0.058	0.030	0.012
	$n = 70$ (37.5)	0.055	0.028	0.012
	$n = 90$ (42.5)	0.050	0.026	0.013
	$n = 110$ (47.5)	0.048	0.025	0.013
	$n = 130$ (52.5)	0.043	0.024	0.014
	$n = 150$ (57.5)	0.039	0.024	0.016
	$n = 170$ (62.5)	0.034	0.023	0.017
b^{SS}	Retirement pensions	0.677	0.675	0.673
T	Lumpsum transfers	0.205	0.204	0.203
$\underline{a}$	Borrowing constraint		-1.12	
τ^{UI}	Unemployment insurance tax rate	0.004	0.010	0.013
τ^{SS}	Social security tax rate	0.050	0.050	0.050
τ^I	General income tax rate	0.150	0.150	0.150
$\bar{m}$	Maximum benefit duration in periods		4	
$\rho_{k,n}$	UI replacement rate by age (years)			
	$n = 0$ (20)	0.51	0.59	0.73
	$n = 45$ (31.25)	0.42	0.45	0.45
	$n = 90$ (42.5)	0.03	0.01	-0.10
	$n = 135$ (53.75)	-0.15	-0.23	-0.14
	$n = 180$ (65)	0.09	0.03	0.02
$\underline{b}$	UI floor		0.00	
$\bar{b}$	UI cap		inf	

Notes: Model parameters in the calibration with optimimized policy parameters under fixed government budget by worker type (exercise C'). Program budgets are obtained from baseline calibration. Policy parameters are replacement by age and worker type and are obtained by solving (1.20); remaining parameters as in baseline calibration. The functions $\delta_{k,n}$ and $\rho_{k,n}$ are cubic splines through values in the table.

Appendix 1.D Computational Implementation

Model solution. As indicated in Section 1.4 and Appendix 1.B, the model is solved numerically by backwards induction. For tractability, unemployment duration (not UI eligibility) is capped at some maximum duration $M > \bar{m}$. Since workers do not receive benefits on unemployment spells longer than $\bar{m}$, the only implication is that, from duration M onward, human capital does not further depreciate. This allows for solving a finite number of value functions for workers in the *unemployed* and *searching* states. The treatment is innocuous since, under reasonably calibrated parameters, no worker experiences unemployment spells with a duration exceeding M.[18] Separation probabilities $\delta_{k,n}$ and UI replacement rates $\rho_{k,n}$ are implemented as cubic Hermite splines. Knots are placed at 5-year age intervals starting from age 22.5 years for separation probabilities, and at 11.25-year intervals starting from age 20 for replacement rates. The function values at the knots are the calibrated parameters.

Model calibration. The model is calibrated using indirect inference, i.e. by minimizing the distance between model and data moments. The distance measure is the average root mean square error (RMSE) across types. The average RMSE is minimized using quasi-Newton methods.[19] The gradient of the objective function w.r.t. the parameter vector is computed at each step using the two-sided finite differences method. For details on the numerical methods, see Miranda and Fackler (2004).

18. For most calibrations, workers can "force" transitioning to employment (the domain of $\zeta_k(s)$ includes probability 1), and leisure utility is bounded above by zero while consumption utility satisfies the lower Inada condition. Once UI benefit eligibility expires, all current consumption needs to be financed from asset holdings, which are bounded (since initial assets are finite and accumulation is bounded by finite wages). Consequently, as unemployment duration increases and asset holdings approach zero, all affected workers will choose to ensure transitioning to employment.

19. I use the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm to compute the direction of the update step in combination with a backtracking line search algorithm to compute the optimal step length.

Chapter 2

Demographic Change and Technological Choice

2.1 Introduction

As most developed economies, Germany is in the midst of a demographic transformation. For at least three decades, birth rates have been decreasing and life expectancy has been increasing. As a result, the population is aging and is expected to continue doing so for at least three more decades. This transition is slow-moving in nature, but with baby boomers approaching retirement, a tipping point has been reached, introducing a fundamental economic shift: labor becomes increasingly scarce while capital grows relatively abundant.

The relative availability of basic economic input factors (capital and labor) is a key determinant of industry structure. For one thing, firms choose the mix of input factors with which they can produce output at the lowest cost given their current production technology. For another, firms actively decide which technologies to invest in given relative prices. The mechanism I examine here is simple: demography-driven changes in the relative price of input factors triggers a response from firms; depending on the characteristics of the industry they operate in, they can invest in labor-saving technology or engage in wage competition. The aim of this essay is two-fold: first, I document how changing demographics have affected the availability of input factors and how firms have responded; second, I develop a model to simulate these adaptations going forward.

Towards the first aim, I use a novel application of the German Federal Employment Agency's labor scarcity methodology extended over a much longer time horizon. I document a structural break around 2010 when demographic pressures began to bite. Before 2010, labor scarcity remained low and roughly constant across both production and services. Since then, it has risen significantly in both sectors, triggering markedly different firm responses. Generally speaking, both real wages and investment have been growing. Diving deeper into investment though reveals interesting patterns: In production sectors, software, and research and development (R&D) investment has exploded while machinery investment has grown modestly. In services, both categories remain largely flat. This divergence suggests that production firms are racing to automate as labor becomes scarce, while service firms (constrained by limited substitution possibilities) cannot easily replace workers with machines.

Towards the second aim, I develop a multi-sector life cycle model with endogenous technology choice that combines two established frameworks. I extend the Dvorkin and Monge-Naranjo (2019) worker sector-choice model with explicit age structure and savings, and integrate it with a multi-sector version of the Acemoglu (2010) production framework featuring dynamic technology choice by monopolistic innovators. I derive conditions for the existence of a stationary solution and apply the model to simulate the German economy between 2000 and 2020.

Qualitatively, the model can replicate all patterns observed in the data: Real wages are increasing and more so for services, investment in machines and technology is increasing for production and flat for services. Quantitatively, the model does not explain the entire effect: observed changes in investment for machines and structures are captured, and approximately one-fourth of the change in technology investment is explained through demographic forces alone. When coupled with declining technology costs, the mechanism accounts for up to one-third of observed technology investment changes. While falling short of explaining the full magnitude of observed patterns, these results demonstrate that the demographic channel is both qualitatively and quantitatively meaningful. In forward-looking simulations based on projected population dynamics, I find that demographic forces alone generate 0.2% annual growth in machine productivity between 2000 and 2040, suggesting that current trends will intensify as aging accelerates. Thus, Germany's experience could provide a preview of the challenges that many other developed economies will soon face.

The essay proceeds as follows: Section 2.2 discusses the contribution to the literature and Section 2.3 presents the empirical evidence, both descriptive and from regression analyses; Section 2.4 then introduces the theoretical model and derives key equilibrium characteristics; Section 2.5 demonstrates the core mechanism before simulating technological change in the German economy based on observed demographic changes; Section 2.6 concludes.

2.2 Related Literature and Contribution

This essay synthesizes and extends three distinct strands of literature to analyze how demographic change drives differential technological responses across sectors.

The first strand analyzes the economic implications of demographic change. Much of the literature focuses on social security or related policies that are immediately affected by changes in the demographic composition of an economy. De Nardi, İmrohoroğlu, and Sargent (1999) pioneer the use of demographic projections to assess implications for the Social Security and Medicare systems arising from population aging. Specifically focusing on the availability of input factors to production, Krueger and Ludwig (2007) show that aging reduces rates of return to capital in the US by roughly 0.9 percentage points between 2005 and 2050, while wages are increasing by roughly 4% over the same period. Börsch-Supan, Ludwig, and Winter (2006) then demonstrate that the demographic transition that most developed economies are undergoing affects the availability of capital not only on the national level, but can also induce changes in international flows of capital.

The second strand of literature is concerned with the substitutability of human and physical capital. In most cases, this literature is focused on the use of robots in manufacturing industries and addresses whether increased use of robotics is associated with loss of manufacturing jobs. Examples are Dauth, Findeisen, Südekum, and Woessner (2017) for the German economy, and Graetz and Michaels (2018), who provide cross-country evidence on robot adoption patterns and total factor productivity. Autor and Dorn (2013) specifically account for differences across economic sectors and document how differences in automation potential lead to polarization in employment and wages. Linking demographic change and innovation, Acemoglu and Restrepo (2021) demonstrate that aging leads to greater automation because it creates shortages of middle-aged workers who specialize in manual production tasks. Their mechanism operates through workforce composition, i.e. the relative scarcity of workers aged 36–55 who are most susceptible to robot substitution. Finally, in a theoretical treatment of labor scarcity and substitution of labor and machines, Acemoglu (2010) derives conditions for when labor scarcity induces innovation vs. when innovation is hampered by labor scarcity.

The third strand examines how workers are allocated to sectors of the economy. The particular contribution I build on from this literature is that of Dvorkin and Monge-Naranjo (2019), who highlight life cycle human capital accumulation as a key determinant of sectoral choice by workers. They model how relative wages affect workers' sector choices over their careers, and how sectoral composition, in turn, determines the aggregate growth rate of the economy.

I combine insights from these three literatures to analyze economy-wide effects of demographics on technology adoption. Unlike Acemoglu and Restrepo (2021), who focus on age composition within the workforce, I examine how demographic change affects the fundamental scarcity of labor relative to capital as factors of production. While their mechanism operates through worker heterogeneity by age, the mechanism examined here operates through factor market equilibrium: aging simultaneously reduces the total supply of labor (as boomers retire) and increases the relative supply of capital (as older workers tend to hold more capital), creating economy-wide scarcity that firms address differently depending on their technological constraints.

This approach captures macroeconomic implications that age-composition mechanisms miss. Specifically, I show how capital abundance from demographic transition, combined with labor scarcity, drives differential sectoral responses based on technological possibilities for factor substitution. Production sectors with high capital-labor substitutability respond by investing heavily in labor-saving technologies, while service sectors with limited substitution possibilities compete primarily through wages (in line with the theoretical predictions of Acemoglu, 2010).

My empirical contribution complements the theoretical innovation by extending the German Federal Employment Agency's labor scarcity methodology over a much longer time horizon, revealing the structural break around 2010 when demographic pressures intensified. This timing coincides with the acceleration of differential investment patterns across sectors, providing crucial identification for the proposed mechanism.

2.3 Empirical Evidence

In this section, I present empirical evidence for the mechanism outlined above. I first present data sources and descriptive evidence on the demographic transformation of the German economy as well as the development of key statistics related to the relative supply and demand of capital and labor during this transition. I conclude with a regression analysis that assesses how different industries respond to the challenges posed by the relative scarcity of labor arising from the changing demographics.

2.3.1 Data Sources

Demographic data is taken from the UN World Population Prospects 2022 (WPP). From the WPP, I use information on observed and projected population size and conditional survival probabilities by age between 1950 and 2100 (1950–2021: observed data; 2022–2100: medium projection scenario).

For labor market data, I draw extensively on the German Sample of Integrated Labour Market Biographies (SIAB) from the Bundesagentur für Arbeit (Federal Employment Agency of Germany, BA). I use the SIAB Regional File (or SIAB-R), Version 7519 v1 (Frodermann, Ganzer, Schmucker, and Berge, 2021a). The dataset contains information on age, employment status, occupation, and earnings. I use observations from the SIAB between January 1998 and December 2019.

Information on wages, investment, and employment is obtained from the Volkswirtschaftliche Gesamtrechnung (German National Accounts, VGR) and the Volkswirtschaftliche Gesamtrechnung der Länder (National Accounts of the German Federal States, VGRdL). Data from the VGR includes gross investment by investment category, as well as wages and employment by economic sector. Investment categories are 'structures', 'machines' and 'other investment', where other investment is defined as "[...] intellectual property (software and databases, research and development, copyrights, exploration drilling), livestock and crops". For sectors other than agriculture and mining, other investment therefore mainly consists of software, databases, and R&D, which I will also refer to as 'technology' throughout the analysis. The VGRdL data contains data on investment, wages, and employment by German federal states and by economic activity. In the VGRdL dataset, the investment categories 'machines' and 'other investment' are only reported as one combined category. From both VGR and VGRdL, I use data from 2000 to 2022.

Finally, information on interest rates is obtained from the Bundesbank (German Central Bank, BB). I use overnight real interest rates on bank deposits between 2000 and 2022.

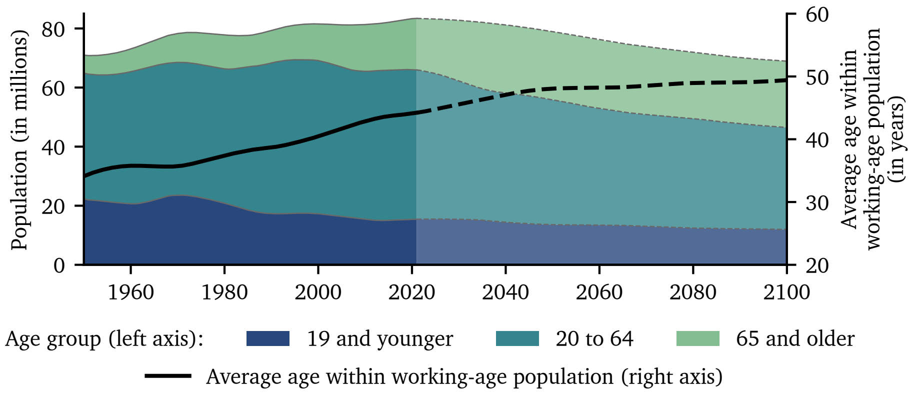
Where applicable, economic activities are categorized according to the German Classification of Economic Activities issue 2008 (WZ08) in all datasets. Throughout the analysis, economic activities are aggregated to two broad industry domains: 'production' (WZ08 sections A-F) and 'non-government services' (WZ08 sections G–N). Government and other public activities (WZ08 sectors O–U) are excluded from the analyses. Occupational data is reported according to the Classification of Occupations (KldB) of the German Statistical Office. Data is coded according to KldB issue 1988 (until 2011) and issue 2010 (from 2012 onward). Geographic observation units are labor market regions (derived from IAB labor market regions) and German federal states. To ensure consistency across dataset, some regions and occupation groups are merged. For further details regarding all data sources, data handling, and data classifications, see Appendix 2.A.

2.3.2 Descriptive Statistics

According to the projections by the UN, the total population is expected to peak at ca. 83 million in the 2020s and then decrease by almost 20% until 2100. Figure 2.1 depicts the observed and projected further demographic transition of Germany from 1950 to 2100. More importantly for the labor market, the working-age population has been increasing steadily until the 1990s, exceeding 50 million for the first time in 1988, and remained relatively constant until the early 2020s, before dropping below 50 million again in 2023. From there, it is projected to continue to decrease to ca. 43 million, or by ca. 15%, by 2050.

This surge and decline in the working age population has been accompanied by a general aging of both the population as a whole and the working-age population. The average age within the working-age population, i.e. the population between 20 and 64, was relatively stable at ca. 38 years until 1980 and has since been increasing to ca. 44 years in the late 2010s. The increase is expected to continue relatively linearly until 2050, where it will be around 48 years of age. Interestingly, while the overall demographic transition is projected to continue well after 2050,

Figure 2.1. Demographics of the German population 1950–2100

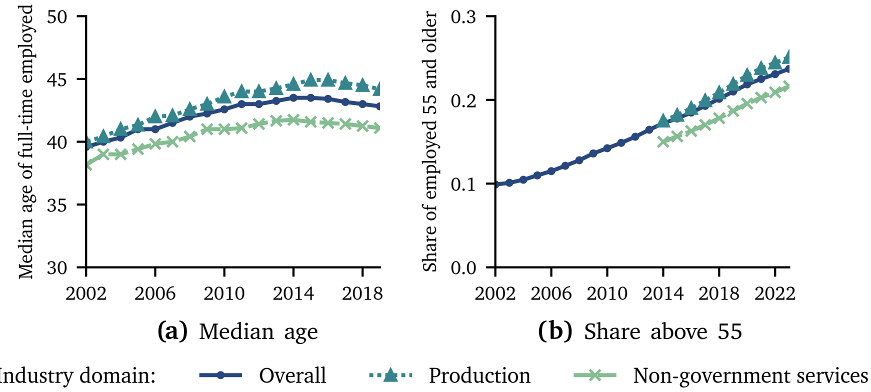


Notes: Population by broad age groups (left axis) and average worker age (right axis); average worker age is defined as the average age within the population group aged 20 to 64. Data from 2022 onward are projections (medium projection scenario).
Source: UNWPP (United Nations, Department of Economic and Social Affairs, Population Division, 2022), own computation.

the average age of a working-age individual is projected to remain relatively stable at slightly less than 50 years. In sum, the average person within working-age is projected to be 10 years older in 2050 than in 1980. The WPP data is informative about size and composition of the working-age population, which represents the total potential labor supply of the economy. Thus, the economy has shifted from a regime of growing potential labor supply (until 1990) to stable potential labor supply (from 1990 to 2020) and then decreasing potential labor supply (from 2020 to 2050). Although changes in labor force participation may affect how potential labor supply translates into actual labor supply, the changes in potential labor supply are likely to affect the actual availability of labor.

Some effects of the demographic transition can already be observed in the active labor force. Figure 2.2 depicts median age of full-time employed workers and the share of full-time employed workers above the age of 55 by industry cluster. Between 1998 and 2015, median age has increased from 38.7 (37.1) to 44.9 (41.6) in production (non-government services) where it has remained relatively constant until 2019. Median worker age is correlated with proximity to retirement and, therefore, informative about the pressure to replace workers that leave the work force due to demographics. A more direct measure of the demographic pressure to replace retiring workers is the share of the work force that is close to the statutory retirement age. The share of workers aged 55 or older (so roughly within 10 years of statutory retirement) has been stable at ca. 10% until the mid-2000s. From then on, it has increased steadily to almost 25% by 2023. The overall trend holds within industry domains with slightly higher shares in production than in services. Both indicators provide evidence that the aging of the

Figure 2.2. Demographics of the labor force by industry domain



Notes: Median age of full-time employed (SIAB data, overall and by industry domain, left panel), and share of full-time employed aged 55 and older (BA data, overall and by industry domain, right panel; BA data by industry only available from 2014 onward). Industry domains are aggregated from WZ08 industry sections (1-letter codes); 'Production' entails sections A–F, 'non-government services' entails sections G–N.
Source: BA (Statistik der Bundesagentur für Arbeit, 2014–2024, 2024d), SIAB (Frodermann et al., 2021a), own computation.

Baby Boomer cohorts has translated into higher median ages and higher shares of workers approaching retirement, both overall and within industry domains.

The central premise of this essay is that demographic change exerts immediate effects on labor supply, whereas its influence on labor demand is more mediated or indirect. To evaluate this assumption, I seek to quantify changes in labor supply attributable to demographic shifts. However, since labor supply and demand are endogenously determined, isolating a pure measure of labor supply change is inherently challenging. Moreover, in the absence of market frictions, standard economic theory suggests that prices adjust to equate supply and demand, leaving no theoretical room for persistent labor shortages or scarcity.

Despite this, understanding how shifts in relative supply affect market participants' behavior remains crucial, particularly under real-world conditions where frictions, rigidities, and institutional constraints exist. To capture such imbalances, I draw on the methodology developed by the BA, which constructs a composite index of labor scarcity. This index combines six sub-indicators, each designed to capture a distinct dimension of situations in which labor supply falls short of demand. I adapt this methodology to a novel dataset in order to generate a time-varying measure of occupational-level labor scarcity. Full details of the index construction are provided in Appendix 2.A.2.

To derive measures of labor scarcity at the industry level, I aggregate scarcity measures defined over a different segmentation dimension using employment weights. Specifically, I apply the following approximation:

$$z_{it} \equiv \sum_k w_{ikt} z_{ikt} \approx \sum_k \bar{w}_{ik} \bar{z}_{kt} \equiv \tilde{z}_{it} \tag{2.1}$$

where i indexes industries, k denotes observation units at which scarcity is measured (e.g., regions, occupations, or region-occupation pairs), and t is the year. The variable z_{ikt} captures labor scarcity for unit k within industry i at time t, and w_{ikt} denotes that unit's share in total employment in industry i at time t. The first equivalence is an exact decomposition of industry-level scarcity, while the second equivalence defines the approximate measure I employ. The approximation relies on time-averaged employment shares $\bar{w}_{ik}$ and a scarcity measure that is observation-unit specific (though not industry-specific), $\bar{z}_{kt}$.

I adopt this approach for two main reasons. First, the methodology developed by the BA was explicitly designed to measure scarcity at the occupational level and does not directly extend to industries. Second, several sub-indicators required to construct the scarcity index are only available at the segmentation level of the observation units (e.g., occupations or regions), making direct computation at the industry level infeasible.

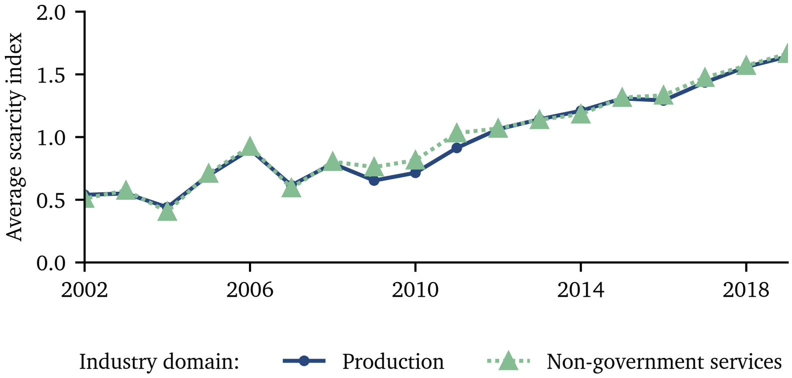
In the main specification, the observation units k correspond to region-occupation pairs. Regions are labor market regions, and occupations are defined at the level of KldB main occupation groups (KldB 1988 level-2-codes/ KldB 2010 2-digit codes).[1]

Figure 2.3 depicts average labor scarcity by industry domain. Note that both level and change of labor scarcity are very similar for both domains: after an initial period with moderate and relatively stable average scarcity around 0.5 (2002–2009), the scarcity index is increasing almost linearly to more than 1.5 in 2019. The evidence on labor scarcity thus suggests that both industry domains are affected similarly. Finally, note that the distinct increase in labor scarcity occurs later than the increase in median worker age depicted in Figure 2.2a.

The mechanism suggested in this essay operates through changes in factor prices. Figure 2.4 exhibits the development of real interest rates (left panel) and real wages by industry cluster (right panel). Real interest rates exhibit a decreasing trend over the observation period. Wages again exhibit a pattern of relative stability for the 2002–2009 period, followed by an upward trend for the 2009–2019 period. Note that, for the 2009–2019 period, wage growth is more pronounced for services than for production. Increasing substitution of labor and machines in production and increasing (relative) wages in services would represent incentives for more workers to seek employment in the services sector. This is supported by the data: relative employment in production over services has steadily declined over the observation period from approximately 3:4 in 2000 to approximately 3:5 in 2022 (see Figure 2.4b).

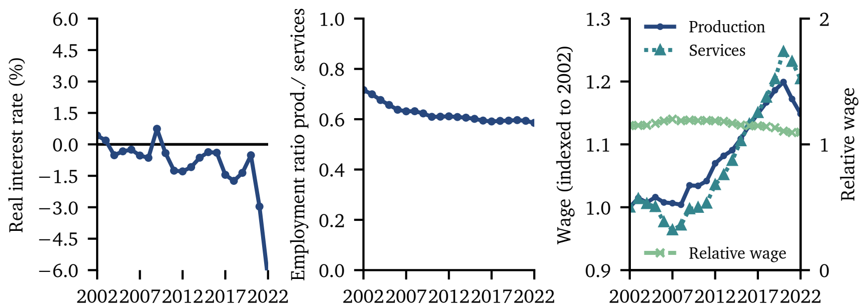
1. The results are robust to alternative segmentations of observation units, see Appendix 2.A.4 for details.

Figure 2.3. Average labor scarcity by industry domain



Notes: Average labor scarcity index by industry domain. Data by industry is aggregated from scarcity by labor market region and occupation main groups and average employment shares by labor market region and occupation main group within industries using equation (2.1). Labor market regions are coarsened IAB labor market regions (see Table 2.A.2), occupation main groups are coarsened from KldB 1988 level-2-codes (2000–2011, see Appendix 2.A.3) and KldB 2010 2-digit codes (2012–2019). Industry domains are aggregated from WZ08 industry sections (1-letter codes); 'Production' entails sections A–F, 'non-government services' entails sections G–N.
Source: SIAB (Frodermann et al., 2021a), own computation.

Figure 2.4. Real interest rate, real wages, and relative employment by industry domain



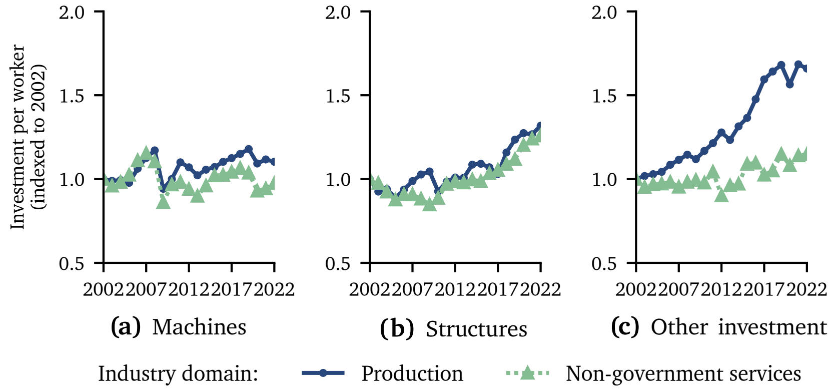
(a) Real interest rate **(b)** Employment ratio **(c)** Real wages

Notes: Real interest rates (left panel), relative employment across industry domains (middle panel), and real hourly wages by industry domain (indexed to 2002 and relative, right panel). Relative employment is defined as total count of workers in 'production' divided by total count of workers in 'non-government services'. Industry domains are aggregated from WZ08 industry sections (1-letter codes); 'Production' entails sections A–F, 'non-government services' entails sections G–N.
Source: BB (Deutsche Bundesbank, 2025), VGRdL (Arbeitskreis "Volkswirtschaftliche Gesamtrechnung der Länder", 2025a), own computation.

Finally, investment can potentially mitigate the twin-challenge of decreased potential labor supply and aging of working-age population. In particular, by investing in equipment and, more importantly, in technology, that is in more and better machines, manual labor can potentially be substituted by machines. Indeed, total investment per worker has increased by almost 20% between 1991 and 2024.

There are, however, important differences across investment classes and industries. Figure 2.5 depicts real investment per worker, separated in investment in struc-

Figure 2.5. Investment per worker by type and industry domain



Notes: Real investment by type within industry domain divided by total workers per industry domain, indexed to 2002. Industry domains are aggregated from WZ08 industry sections (1-letter codes); 'Production' entails sections A–F, 'non-government services' entails sections G–N.
Source: VGR (Statistisches Bundesamt (Destatis), Genesis Online, 2025a,b), own computation.

tures, investment in machines, and other investment, by industry cluster between 2002 and 2022 (indexed to 2002). Investment in machines has remained relatively constant, with investment in production sectors slightly increasing. Investment in structures has moderately increased in both sectors. Importantly, other investment (i.e. investment in technology) has increased for both industry clusters, but substantially stronger for production than for services. The increase in investment is particularly pronounced for the period 2009–2019, coinciding with the strong increase in labor scarcity described above. This indicates that the response to labor scarcity follows different strategies in the two industry domains: in production, firms attempt to substitute labor with more and better machines, while in services, firms engage in wage competition in order to attract more labor.

Note that, so far, all evidence has been descriptive. In the next section, I present further econometric evidence on the relationship between demographics, labor scarcity and investment and wages.

2.3.3 Regression Analysis

This subsection pursues to main goals: First, I want to establish the relationship between changes in the demographic composition of the labor force and the measures of labor scarcity presented above; second, I want to provide more rigorous evidence on how firms respond to labor scarcity by industry.

Towards the first goal, I regress the scarcity measure on different proxies for the demographic transition. Here, the key challenge lies in disentangling any direct connection between scarcity and demographics from general trends. As shown above, demographic variables related to population aging and the scarcity measures exhibit a clear time trend. To isolate the effects of demographics on labor scarcity from aggregate trends in both, I exploit variation in the particular exposure of observation units to demographic pressures. Specifically, I estimate the following regression equation:

$$z_{jlt} \;=\; \beta(x_{jlt} - \bar{x}_t) + \gamma_j + \delta_l + \varepsilon_{jlt} \qquad (2.2)$$

where j indexes regions, l indexes occupation groups, and t is the year. The dependent variable z_{jlt} is the scarcity index computed from the sub-indicators for observation cell jl, and x_{jlt} and $\bar{x}_t$ are the demographic variables observed in that cell and nation-wide, respectively. Region and occupation fixed effects are captured by γ_j and δ_l. Using deviations from nation-wide averages preserves variation for region-occupation-pairs that age differently from the national trend. This approach exploits the fact that firms in regions and occupations with differential exposure to demographic pressures should respond differently to these idiosyncratic pressures, holding aggregate trends affecting all units similarly constant.

Moreover, to reduce concerns regarding reverse causality or simultaneity, I use lagged variables for the demographic indicators. In the main specification, I use the 5-year-lag of the share of full-time employed workers within five years of statutory retirement age as the demographic measure, and pairs of labor market regions and occupation main groups as observation units. In other words, I use variation in how much more a region-occupation unit was exposed to the pressure to replace retiring employees (relative to the national average) five years ago to identify differences in labor scarcity observed today.[2] Since the classification of occupation-level has changed in 2012 from KldB 1988 to KldB 2010, I construct two separate datasets on which I estimate equation (2.2). The first dataset contains data from 2000 to 2011 and observation units are combinations of labor market regions and KldB 1988 level-2-codes, the second dataset contains data from 2012 to 2019 and units are labor market regions and KldB 2-digit codes. Table 2.1 summarizes the results. The coefficient estimate for the first dataset (2000–2011) is insignificant, and the coefficient estimate for the second dataset (2012–2019) is positive and statistically significant at the 1% confidence level. These findings align well with the broader demographic trends presented above. Between the 1980s and the early 2000s, the share of employed workers above 55 years of age was stable at ca. 10%. In the mid-2000s, the large Baby Boomer cohorts reached age 55 and this share started to increase to almost 25% in 2023. This led to a phase of demographic pressure buildup, as many more workers were

2. For robustness checks using alternative datasets, demographic variables, and segmentations of observation units, see Appendix 2.A.4.

Table 2.1. Effects of workforce aging on labor scarcity

	KldB 1988 (2000-2011)	KldB 2010 (2012-2019)
$lag_5(Share\ 5\ years\ to\ retirement)$	−0.91	1.63^{**}
	(0.52)	(0.60)
R^2	0.23	0.40
Adj. R^2	0.22	0.39
Num. obs.	5,032	2,994

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Notes: Coefficient estimates for β of (2.2). Observation units are combinations of labor market regions and occupation main groups. Labor market regions are coarsened IAB labor market regions (see Table 2.A.2). Occupation main groups are coarsened from KldB 1988 level-2-codes (first column) and KldB 2010 2-digit codes (second column); For KldB 1988, some level-2-codes are combined for consistency with SIAB data (see Appendix 2.A.3). Estimation based on KldB 1988 uses data from 2000–2011, estimation based on KldB 2010 uses data from 2012–2019, estimation based on federal states uses data from 2000–2019.
Source: SIAB (Frodermann et al., 2021a), own computation.

in the pre-retirement phase but had not retired yet. From ca. 2010 onwards, the pressure translated into measurable scarcity as retirement became more prevalent among these cohorts (either through early retirement options or after reaching the statutory retirement age). The evidence thus is consistent with the premise that exposure to demographic pressure leads to higher labor scarcity, in particular for the aggregate increase observed after 2009. Note that the primary purpose of this exercise is to establish a relation between changes in the demographic composition of the work force and labor scarcity. I make no claims about the exact mechanism. A quantitative analysis of the relationship is beyond the scope of this essay.

Having established a link between demographics and labor scarcity as measured here, I now turn to the relationship of labor scarcity and firms' responses to it. For this, I use data on investment, wages, employment, and hours worked by industry and German federal state from the VGRdL. While this data adds the geographical dimension (German federal states), the information on investment is less granular: 'investment in machines' and 'other investment' is reported jointly. Consequently, the data does not allow differentiating the effects of labor scarcity on these investment categories.

Again, average labor scarcity by industry domain and state is computed as employment-weighted average of scarcity measures:

$$z_{ijt} \;\equiv\; \sum_{l} w_{ijkt} z_{ijkt} \;\approx\; \sum_{k} \bar{w}_{ijk} \bar{z}_{jkt} \;=\; \tilde{z}_{ijt} \tag{2.3}$$

where i is the industry domain, j is the state, k is the observation unit of labor scarcity and t is the year of the observation. In the main specification, observation units are occupation main groups (KldB 1988 level 2 codes / KldB 2010 2-digit codes).

Labor supply and, thus, labor scarcity is endogenous. So are wages and investment levels. The identification challenge now lies in identifying variations in labor scarcity that are exogenous to firms' decisions regarding investments and wages. I address this challenge in multiple ways: First, the construction of the average scarcity measure by industry (though in part motivated by data limitations) separates predetermined employment shares from occupation-region scarcity shifts, creating plausibly exogenous variation in scarcity. Second, the inclusion of comprehensive fixed effects (state, industry, and time) absorbs many potential confounders while variation from the three-way interaction remains.[3] Specifically, I estimate the following regression equation:

$$y_{ijt} \;=\; \beta_1(\tilde{z}_{ijt} * Production_i) + \beta_2(\tilde{z}_{ijt} * Services_i) + \gamma_i + \delta_j + \sigma_t + \varepsilon_{ijt} \tag{2.4}$$

3. Further robustness checks can be found in Appendix 2.A.4.

where $\tilde{z}_{ijt}$ is the average labor scarcity index for industry domain i and state j at time t as computed by (2.3), γ_i, δ_j, and σ_t are fixed effects for industry, state, and time, respectively.

Note that the aggregation method used for the scarcity measure is structurally similar to a shift-share (Bartik) design, where industry-level exposure to occupation-level labor scarcity shocks is computed using pre-determined occupational employment shares. This approach allows for a transparent and interpretable measure of how changes in occupation-level scarcity propagate to industries, conditional on their workforce composition. As in shift-share designs, the validity of this approach relies on the assumption that occupation-level scarcity shocks are exogenous to industry-specific unobservables. In particular, the key assumptions are exogeneity of shifts (i.e. national industry trends are uncorrelated with local unobserved factors affecting the outcome), relevance of the instrument (i.e. local industry composition creates meaningful differential exposure to national trends) and the exclusion restriction (i.e. the instrument only affects outcomes through the endogenous variable of interest). In this specific setup with time-averaged employment shares, shares are not strictly pre-determined, and, hence, potentially influenced by investment/ wages, yet the averaging limits this effect.

Table 2.2 summarizes the results. Coefficients are positive and significant for investment in machines and other investment in production industries and for hourly wages in service industries. This is evidence that the changes described above (production firms invest in more and better machines upon labor scarcity, while service firms engage in wage competition) are in fact (partially) driven by an increase in labor scarcity, which are in turn (at least partially) driven by demographic change.

2.4 Model

This section presents the model. The model consists of several components: the demographic process governs model entry and exit of workers; the worker model describes how workers choose which economic sector to work in and how human capital is accumulated; the production model describes how human capital, physical capital and technology are combined to produce output, as well as how technology is created; financial markets close the model as they transform savings by workers into investments in physical capital. Each component is presented in detail, including key results and statements on the properties of the model.

The focus of this essay is to assess the connection between the demographics of an economy and the use of technology in the production process within the economy. The core mechanism is the change in the relative availability of capital and labor that corresponds to a change in the demographic process. To

Table 2.2. Effects of labor scarcity on investment and wages

	$\log\left(\frac{\text{Investment}}{\text{Worker}}\right)$		$\log\left(\frac{\text{Wage}}{\text{Hour}}\right)$
	Machines & other	Structures	
Scarcity * Production	0.08^{***}	-0.02	0.01
	(0.02)	(0.03)	(0.01)
Scarcity * Non-govt. services	0.04	0.03	0.02^{**}
	(0.02)	(0.03)	(0.01)
R^2	0.75	0.92	0.96
Adj. R^2	0.73	0.92	0.96
Num. obs.	640	640	640

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Notes: Coefficient estimates for β_1 and β_2 of (2.4). Average labor scarcity by industry domain and state is computed from scarcity by occupation main group and average employment shares of occupation main groups within industry-state-combinations using (2.3). Occupation main groups are coarsened from KldB 1988 level-2-codes (2000–2011, see Appendix 2.A.3) and KldB 2010 2-digit codes (2012–2019). Industry domains are aggregated from WZ08 industry sections (1-letter codes); 'Production' entails sections A–F, 'non-government services' entails sections G–N.
Source: SIAB (Frodermann et al., 2021a), VGRdL (Arbeitskreis "Volkswirtschaftliche Gesamtrechnung der Länder", 2025a,b), own computation.

isolate effects of relative, not absolute, availability, it is convenient to eliminate differences resulting from changes in the total size of the economy. This can be achieved by normalizing the model with the total size of the population. With additional assumptions, the normalized model becomes time-invariant, enabling the solution and comparison of the economy's steady states. For each component of the model, I first present the general case (i.e. in absolute terms) and then present the normalized version of the model (i.e. in per-capita terms) including the assumptions required for the normalization and for stationarity of the solution. The proofs of all formal statements, as well as additional derivations, can be found in Appendix 2.B.

2.4.1 Demographics

Time is discrete. In the general case, a continuum of workers of mass $\phi_{t,0}$ enters the model at the beginning of every period t. Workers are active in the labor force for S^w periods, after which they retire for another S^r periods, before they leave the economy. Agents are subject to survival risk: in each period t, a fraction $(1-\zeta_{t,s})$ of agents of age s dies. In other words, $\zeta_t = \{\zeta_{t,s}\}_{s=1}^{S^w+S^r}$ denotes the conditional survival probabilities for agents of age s in period t. Workers have full information about their survival prospects when they enter the model, i.e. they know the relevant vector of survival probabilities $\{\zeta_{t,0}, \zeta_{t+1,1}, \dots, \zeta_{t+S^w+S^r,S^w+S^r}\}$.

In any period t, the mass of workers of age s is denoted by $\phi_{t,s}$ and given by

$$\phi_{t,s} = \phi_{t-s,0} \prod_{m=0}^{s-1} \zeta_{t-s+m,m} = \phi_{t-s,0}\tilde{\zeta}_{t,s-1}$$

where $\tilde{\zeta}_{t,s-1} = \prod_{m=1}^{s-1} \zeta_{t-s+m,m}$ denotes the unconditional probability of a worker of age s of surviving from model entry in period $t-s$ to the current period t. The total population size in period t, denoted by ϕ_t, is

$$\phi_t = \sum_{s=0}^{S^w+S^r} \phi_{t,s} = \sum_{s=0}^{S^w+S^r} \phi_{t-s,0}\tilde{\zeta}_{t,s-1}$$

For the normalization, let $\tilde{\phi}_{t,s}$ denote the share of workers of age s in period t and assume that:

(D1) *The mass of workers entering the model $\phi_{t,0}$ grows at constant rate $G^n = (1+n)$, and the conditional survival rates are constant over time, $\zeta_{t,s} = \zeta_s \forall s, t$.*

Constant conditional survival probabilities imply that unconditional survival probabilities are constant as well. With the mass of newborn workers growing at constant rate $(1+n)$, the mass of workers of age s in period t can then be expressed as

$$\phi_{t,s} = \phi_{t-s,0}\tilde{\zeta}_{t,s-1} = \phi_{t,0}(1+n)^{-s}\tilde{\zeta}_{s-1}$$

and the total mass of workers in period t is

$$\phi_t \;=\; \sum_{s=0}^{S^w+S^r} \phi_{t,s} \;=\; \phi_{t,0} \sum_{s=0}^{S^w+S^r} (1+n)^{-s} \tilde{\zeta}_{s-1} \;=\; \phi_{t,0}\bar{\phi}$$

where $\bar{\phi} = \sum_{s=0}^{S^w+S^r} (1+n)^{-s}\tilde{\zeta}_{s-1}$. Dividing the mass of workers of age s by the total mass of workers in period t yields population shares

$$\tilde{\phi}_{t,s} \;=\; \frac{\phi_{t,s}}{\phi_t} \;=\; \frac{\phi_{t,0}(1+n)^s \tilde{\zeta}_{s-1}}{\phi_{t,0}\bar{\phi}} \;=\; \bar{\phi}^{-1}(1+n)^s \tilde{\zeta}_{s-1} \;=\; \tilde{\phi}_s$$

which are time-invariant.

2.4.2 Workers

The worker model builds on the worker model in Dvorkin and Monge-Naranjo (2019). The economy is composed of J separate economic sectors. Workers choose in which sector to work every period and accumulate human capital throughout their working life. Human capital consists of an absolute advantage and a relative advantage in the labor market, which are separable. The absolute advantage is captured by a worker's level of human capital h which can accumulate and depreciate over time subject to the worker's employment history. The relative advantage is captured by a worker's vector of labor market opportunities ε which is assumed to be stochastic and renewed every period. Moving between sectors of the economy is subject to a utility cost χ that captures aspects such as having to relocate or adapting to a new environment. Human capital accumulation is captured by the transfer parameter τ that captures how much human capital is transferred to the next period, depending on the sector choice of the worker.

New workers enter the model unattached to any sector and with initial human capital h_0. They draw their initial labor market opportunities $\varepsilon_{t,0}$ and choose the sector j in which they begin their working life in the following period. They incur the initial non-pecuniary cost $\chi^j_{t,0}$ and accumulate human capital according to the human capital accumulation function

$$h_1 \;=\; h_0 \tau^j_{t,0} \varepsilon^j_0 \tag{2.5}$$

During the working age ($s = 1, \ldots, S^w - 1$), workers enter a period connected to some sector $j = 1, \ldots, J$. At the beginning of the period, each worker draws a vector of labor market opportunities $\varepsilon_{t,s} = \{\varepsilon^j_{t,s}\}^J_{j=1}$. Workers then work full-time in their current sector and receive real wages $\omega^j_t h \varepsilon^j_{t,s}$. Workers save a constant fraction τ^p of their income for retirement and consume the remaining labor income. Then, they choose the sector for the next period. Moving from one sector to another entails certain costs or benefits. There are direct utility (or non-pecuniary) costs, as well as indirect costs and benefits through imperfect transferability and differences in accumulation of human capital across sectors. The direct cost are captured by the $J \times J$ matrix $\mathcal{X}_{t,s}$, where $\chi_{t,s}^{j,l}$ captures the cost for an agent of age s of moving from occupation j in period t to occupation l in period $t+1$. Note that, in principle, the non-pecuniary costs can depend on both worker age and time. The indirect cost are captured by the $J \times J$ matrix $\mathcal{T}_{t,s}$, where $\tau_{t,s}^{j,l}$ captures the change in the absolute value of human capital for an agent of age s of moving from occupation j in period t to occupation l in period $t+1$. Here, $\tau_{t,s}^{j,l} < 1$ corresponds to a depreciation of human capital, e.g. due to imperfect transferability of skills, and $\tau_{t,s}^{j,l} > 1$ corresponds to accumulation of human capital, e.g. through learning-by-doing. Again, human capital transferability can, in principle, depend on both age and time.

In the last period before retirement (i.e. at age S^w), workers enter the period attached to some sector, work, and consume their net wages, but do not choose a new sector. In retirement, workers receive retirement payments b, which represent returns on their previous savings that have been invested in the financial markets. As workers only save during working age, retired workers are effectively hand-to-mouth. Thus, the present value of retirement, discounted to the last period of the working age and as of period t, can be summarized as

$$\Omega_t(b_t) = u(b_t) \sum_{m=1}^{S^r} \beta^m \Pi_{n=0}^{m-1} \zeta_{t+m,S^w+m} \tag{2.6}$$

Let $V_{t,s}(h_t, j, \varepsilon_t)$ denote the value function of a worker of age s in period t with current human capital h_t, who is currently attached to sector j, and has drawn labor market opportunities ε_t. The worker problem can then be expressed in recursive form as

$$\begin{aligned} V_{t,0}(h_t^0, \varepsilon_t) &= \beta \max_l \left\{ \chi_{t,0}^l \mathbb{E}_{\varepsilon_{t+1}} \left[V_{t+1,1}(h_t^0 \tau_{t,0}^l \varepsilon_t^l, l, \varepsilon_{t+1}) \right] \right\} \\ V_{t,s}(h_t, j, \varepsilon_t) &= u\left(\left(1 - \tau^p\right) \omega_t^j h_t \varepsilon_t^j \right) + \beta \zeta_{t,s} \max_l \left\{ \chi_{t,s}^{j,l} \mathbb{E}_{\varepsilon_{t+1}} \left[V_{t+1,s+1}(h_t \tau_{t,s}^{j,l} \varepsilon_t^l, l, \varepsilon_{t+1}) \right] \right\} \\ V_{t,S^w}(h_t, j, \varepsilon_t) &= u\left(\left(1 - \tau^p\right) \omega_t^j h_t \varepsilon_t^j \right) + \Omega_t(b_t) \end{aligned} \tag{2.7}$$

As outlined above, the human capital of workers is assumed to be separable in their absolute advantage h, which makes them more productive in all sectors, and their relative advantage ε, which determines their relative productivity across sectors. With some additional assumptions on workers' consumption utility and the structure of retirement income, the absolute advantage h can be factored out of the sector choice problem, as summarized in proposition 1.

Proposition 1 (Worker Problem Factorization). *Assume that:*

(W1) *Consumption utility is CRRA with parameter $\gamma > 0, \gamma \neq 1$, i.e.*

$$u(c) = \frac{c^{1-\gamma}}{1-\gamma}$$

(W2) *Retirement payments are proportional to the absolute value of human capital at retirement, i.e.*

$$b_t \propto h_t^{S^w} \quad \Leftrightarrow \quad \exists \bar{b}_t \ s.\ th.\ b_t = h_t^{S^w} \bar{b}_t$$

Then, the absolute value of human capital can be factored out of the worker problem and (2.7) is equivalent to

$$\begin{aligned} v_{t,0}(\varepsilon_t) &= \beta \max_l \left\{ \chi_{t,0}^l \mathbb{E}_{\varepsilon_{t+1}} \left[v_{t+1,1}(l, \varepsilon_{t+1}) \right] \right\} \\ v_{t,s}(j, \varepsilon_t) &= \frac{\left[(1-\tau^p) \omega_t^j \varepsilon_t^j \right]^{1-\gamma}}{1-\gamma} + \beta \zeta_{t,s} \max_l \left\{ \chi_{t,s}^{j,l} (\tau_{t,s}^{j,l} \varepsilon_t^l)^{1-\gamma} \mathbb{E}_{\varepsilon_{t+1}} \left[v_{t+1,s+1}(l, \varepsilon_{t+1}) \right] \right\} \\ v_{t,S^w}(j, \varepsilon_t) &= \frac{\left[(1-\tau^p) \omega_t^j \varepsilon_t^j \right]^{1-\gamma}}{1-\gamma} + \Omega_t(\bar{b}_t) \end{aligned} \tag{2.8}$$

The proof of Proposition 1 can be found in Appendix 2.B.

Continuation values, and therefore sector choices, are closely linked to workers' expectations on their future labor market opportunities. Analogous to Dvorkin and Monge-Naranjo (2019), I assume that the labor market opportunities are distributed according to a type-II extreme value distribution (or Fréchet distribution) for all occupations. This assumption yields closed form solutions for the conditional expectations of the normalized value functions $v_{t,s}^j$ as well as for the probabilities for a worker of age s to switch from occupation j to occupation l from period t to $t+1$. These results are summarized in proposition 2.

Proposition 2 (Worker Problem Solution). *Denote by $\mathscr{M}_{t,0}$ the initial sector choice vector in period t and by $\mathscr{M}_{t,s}$ the sector choice matrix of workers of age s in period t. Moreover, let $v_{t,s}^j = \mathbb{E}_{\varepsilon_t}\left[v_{t,s}\left(j, \varepsilon_t^j\right)\right]$ and $\Upsilon_t^j = \mathbb{E}_{\varepsilon_t}\left[\left(\varepsilon_t^j\right)^{1-\gamma}\right]$ and assume that:*

(W3) *All real wages ω_t^j are strictly positive.*

(W4) *For all ages s, all periods t, and all sectors j, the shocks ε_t^j are independently distributed from a Fréchet distribution with scale parameter λ^j and curvature parameter κ.*

Then,

(1) the expected value of being attached to occupation j at age $0 < s < S^w$ is

(a) if $0 < \gamma < 1$:

$$\begin{aligned} v_{t,s}^j &= \Upsilon_t^j \frac{\left[(1-\tau^p)\, \omega_t^j \right]^{1-\gamma}}{1-\gamma} \\ &\quad + \beta \zeta_{t,s} \Gamma\left(1 - \frac{1-\gamma}{\kappa}\right) \left[\sum_{l=1}^{J} \left(\chi_{t,s}^{j,l} v_{t+1,s+1}^l \right)^{\frac{\kappa}{1-\gamma}} \left(\tau_{t,s}^{j,l} \lambda^j \right)^{\kappa} \right]^{\frac{1-\gamma}{\kappa}} \end{aligned}$$

(b) if $\gamma > 1$*:*

$$v^j_{t,s} = \Upsilon^j_t \frac{\left[(1-\tau^p)\,\omega^j_t\right]^{1-\gamma}}{1-\gamma} - \beta\zeta_{t,s}\Gamma\left(1-\frac{1-\gamma}{\kappa}\right)\left[\sum_{l=1}^{J}\left(-\chi^{j,l}_{t,s}v^l_{t+1,s+1}\right)^{\frac{\kappa}{1-\gamma}}\left(\tau^{j,l}_{t,s}\lambda^j\right)^{\kappa}\right]^{\frac{1-\gamma}{\kappa}}$$

and the expected value at age S^w *is*

$$v^j_{t,S^w} = \Upsilon^j_t \frac{[(1-\tau^p)\omega^j_t]^{1-\gamma}}{1-\gamma} + \Omega_t(\bar{b}_t)$$

(2) *the elements of* $\mathscr{M}_{t,0}$ *and* $\mathscr{M}_{t,s}$*, i.e. the share of new workers choosing sector l and the share of workers of age s switching from occupation j to occupation l in period t, are given by*

(a) if $0 < \gamma < 1$*:*

$$\mu^l_{t,0} = \frac{\left[\lambda^l\tau^l_{t,0}\left(\chi^l_{t,0}v^l_{t+1,1}\right)^{\frac{1}{1-\gamma}}\right]^{\kappa}}{\sum_{k=1}^{J}\left[\lambda^k\tau^k_{t,0}\left(\chi^k_{t,0}v^k_{t+1,1}\right)^{\frac{1}{1-\gamma}}\right]^{\kappa}}$$

and

$$\mu^{j,l}_{t,s} = \frac{\left[\lambda^l\tau^{j,l}_{t,s}\left(\chi^{j,l}_{t,s}v^l_{t+1,s+1}\right)^{\frac{1}{1-\gamma}}\right]^{\kappa}}{\sum_{k=1}^{J}\left[\lambda^k\tau^{j,k}_{t,s}\left(\chi^{j,k}_{t,s}v^k_{t+1,s+1}\right)^{\frac{1}{1-\gamma}}\right]^{\kappa}}$$

respectively, and

(b) if $\gamma > 1$*:*

$$\mu^l_{t,0} = \frac{\left[\lambda^l\tau^l_{t,0}\left(-\chi^l_{t,0}v^l_{t+1,1}\right)^{\frac{1}{1-\gamma}}\right]^{\kappa}}{\sum_{k=1}^{J}\left[\lambda^k\tau^k_{t,0}\left(-\chi^k_{t,0}v^k_{t+1,1}\right)^{\frac{1}{1-\gamma}}\right]^{\kappa}}$$

and

$$\mu^{j,l}_{t,s} = \frac{\left[\lambda^l\tau^{j,l}_{t,s}\left(-\chi^{j,l}_{t,s}v^l_{t+1,s+1}\right)^{\frac{1}{1-\gamma}}\right]^{\kappa}}{\sum_{k=1}^{J}\left[\lambda^k\tau^{j,k}_{t,s}\left(-\chi^{j,k}_{t,s}v^k_{t+1,s+1}\right)^{\frac{1}{1-\gamma}}\right]^{\kappa}}$$

respectively.

The proof of Proposition 2 can be found in Appendix 2.B. Note that expected value functions and sector choice probabilities are fully determined by wages, the parameters governing the demographic process (specifically, survival probabilities $\zeta_{t,s}$), and the parameters governing the transfer of human capital ($\mathscr{T}_{t,s}$ and $\mathscr{X}_{t,s}$). In fact, a proportional change in all wages cancels out of the sector choice probabilities, meaning that only relative wages matter for workers' individual decisions.

Aggregating over all workers, individual choice probabilities yield proportions of workers choosing a specific sector. The distribution of active workers in the economy by age and sector is therefore determined by two components only: the demographic process, entailing the sequence of masses of newborn workers and conditional survival probabilities $\{\phi_{t,0}, \zeta_t\}_{t=0}^{\infty}$, and the sequences of initial sector choice vectors and sector choice matrices by age, $\{\mathscr{M}_{t,0}, \{\mathscr{M}_{t,s}\}_{s=1}^{S^w+S^r}\}_{t=0}^{\infty}$. Lemma 3 summarizes this result, the proof can again be found in Appendix 2.B.

Lemma 3 (Distribution of workers across sectors). *Let $\phi_{t,s}^j$ denote the mass of workers of age s attached to occupation j in period t. The mass of workers by age and sector in period t is then given by the matrix*

$$\Phi_t = \begin{bmatrix} \phi_{t,1}^1 & \cdots & \phi_{t,1}^J \\ \vdots & \ddots & \vdots \\ \phi_{t,S^w} & \cdots & \phi_{t,S^w}^J \end{bmatrix} = \begin{bmatrix} \phi_{t-1,0}\mathscr{M}_{t-1,0} \\ \phi_{t-2,0}\zeta_{t-1,1}\mathscr{M}_{t-2,0}\mathscr{M}_{t-1,1} \\ \vdots \\ \phi_{t-S^w,0}\tilde{\zeta}_{t,S^w-1}\mathscr{M}_{t-S^w,0}\prod_{m=1}^{S^w-1}\mathscr{M}_{t-S^w+m,m} \end{bmatrix}$$

Similarly, the aggregate distribution of human capital by age and sector can be derived in closed form. To see this, first note that the human capital of a worker of a given age and attached to a given sector is closely related to sector choice probabilities: Workers always choose the sector that yields the highest expected continuation value (which accounts for current wages and expected accumulation of human capital). The average labor market opportunity for a given sector within the group of workers that have chosen this sector for the next period thus is given by the conditional expectation of the labor market opportunity shock. The distributional assumptions made above yield closed form expressions for this conditional expectation. Lemma 4 summarizes this result.

Lemma 4 (Expected value of human capital). *Let $\bar{\varepsilon}_{t,0}^l$ denote the expected value of the labor market opportunity in sector l of newborn workers in period t that have chosen l for period $t+1$. Let $\bar{\varepsilon}_{t,s}^{j,l}$ denote the expected value of the labor market opportunity in sector l of workers of age s that are attached to sector j in period t and have chosen l for period $t+1$. Then, $\bar{\varepsilon}_{t,0}^l$ and $\bar{\varepsilon}_{t,s}^{j,l}$ are given by*

$$\bar{\varepsilon}_{t,0}^l = \mathbb{E}_\varepsilon\left[\varepsilon^l | l = \arg\max_k \left\{\chi_{t,0}^k(\tau_{t,0}^k\varepsilon^k)^{1-\gamma}v_{t+1,1}^k\right\}\right] = \Gamma\left(1-\frac{1}{\kappa}\right)\lambda^l\left(\mu_{t,0}^l\right)^{-\frac{1}{\kappa}}$$

and

$$\bar{\varepsilon}^{j,l}_{t,s} = \mathbb{E}_\varepsilon\left[\varepsilon^l | l = \arg\max_k \left\{\chi^{j,k}_{t,s}(\tau^{j,k}_{t,s}\varepsilon^k)^{1-\gamma} v^k_{t+1,s+1}\right\}\right] = \Gamma\left(1-\frac{1}{\kappa}\right)\lambda^l\left(\mu^{j,l}_{t,s}\right)^{-\frac{1}{\kappa}}$$

respectively.

Again, choice probabilities and model parameters fully determine the expected values of labor market opportunities. These same shocks are the key determinants of human capital accumulation between periods. Aggregating over all workers then yields average human capital by age and sector from expected values of the individual workers' problem. Lemma 5 summarizes this result.

Lemma 5 (Human Capital Distribution). *Let $h^j_{t,s}$ denote the sum of human capital of workers of age s attached to sector j in period t. Let $\tilde{\Gamma} = \Gamma\left(1-\frac{1}{\kappa}\right)\lambda$ and $\tilde{\mathcal{M}}_{t,s} = \mathcal{T}_{t,s} \circ \left(\mathcal{M}_{t,s}\right)^{1-\frac{1}{\kappa}}$, where $\circ$ denotes the element-wise (or Hadamard) product and $(\cdot)^x$ denotes the element-wise power. The total human capital in the economy is then given by the matrix*

$$H_t = \begin{bmatrix} h^1_{t,1} & \cdots & h^J_{t,1} \\ \vdots & \ddots & \vdots \\ h^1_{t,S^w} & \cdots & h^J_{t,S^w} \end{bmatrix}$$
$$= \begin{bmatrix} \phi_{t-1,0} h^0_{t-1} \tilde{\Gamma} \circ \tilde{\mathcal{M}}_{t-1,0} \\ \phi_{t-2,0} h^0_{t-2} \zeta_{t-1,1} \left(\tilde{\Gamma}\right)^2 \circ \left[\tilde{\mathcal{M}}_{t-2,0} \tilde{\mathcal{M}}_{t-1,1}\right] \\ \vdots \\ \phi_{t-S^w,0} h^0_{t-S^w} \zeta_{t,S^w-1} \left(\tilde{\Gamma}\right)^{S^w} \circ \left[\tilde{\mathcal{M}}_{t-S^w,0} \prod_{m=1}^{S^w-1} \tilde{\mathcal{M}}_{t-S^w+m,m}\right] \end{bmatrix}$$

The labor supply in sector j in a given period t is the total human capital of all workers attached to that sector in that period,

$$H^{supply,j}_t = \sum_{s=1}^{S^w} h^j_{t,s} \tag{2.9}$$

which corresponds to the sum over column j of matrix H_t.

Lemma 5 captures how the demographic process and model parameters shape the total supply of labor in the economy. As outlined above, the focus of the exercise is to identify consequences of changes in relative availability of capital and labor. It is therefore convenient to again normalize these distributions with the total size of the population. The necessary assumptions for the normalization and the resulting distributions of workers and human capital are summarized in Proposition 6.

Proposition 6 (Normalized distributions of workers and human capital). *Let $\tilde{h}_s^j$ denote the normalized human capital of workers of age s attached to sector j and ε' denote next periods' labor market opportunities. Assume that (D1) holds and that:*

(W5) *Human capital transferability and non-pecuniary switching cost parameter are constant over time ($\mathcal{T}_{t,s} = \mathcal{T}_s$ and $\mathcal{X}_{t,s} = \mathcal{X}_s$ $\forall t, s$).*

(W6) *Initial human capital per worker is constant over time ($h_t^0 = h^0$ $\forall t$).*

(W7) *Real wages and retirement payments are constant over time ($\omega_t = \omega$ $\forall t$ and $b_t = b$ $\forall t$).*

Then, the worker problem (2.8) is time-invariant with

$$\begin{aligned} v_0(\varepsilon) &= \beta \max_l \left\{ \chi_0^l \mathbb{E}_{\varepsilon'} \left[v_1(l, \varepsilon') \right] \right\} \\ v_s(j, \varepsilon) &= \frac{\left[(1-\tau^p)\, \omega^j \varepsilon^j \right]^{1-\gamma}}{1-\gamma} + \beta \zeta_s \max_l \left\{ \chi_s^{j,l} \left(\tau_s^{j,l} \varepsilon^l \right)^{1-\gamma} \mathbb{E}_{\varepsilon'} \left[v_{s+1}\left(l, \varepsilon'\right) \right] \right\} \\ v_{S^w}(j, \varepsilon) &= \frac{\left[(1-\tau^p)\, \omega^j \varepsilon^j \right]^{1-\gamma}}{1-\gamma} + \Omega(\bar{b}) \end{aligned} \tag{2.10}$$

Moreover, initial sector choice vectors and sector choice probabilities by age are constant over time

$$\mathcal{M}_{t,0} = \mathcal{M}_0 \quad \forall t \quad \text{and} \quad \mathcal{M}_{t,s} = \mathcal{M}_s \quad \forall t, s$$

Further, the normalized distribution of workers (i.e. the share of workers by age and sector) is constant over time and given by

$$\tilde{\Phi} = \begin{bmatrix} \tilde{\phi}_1^1 & \dots & \tilde{\phi}_1^J \\ \vdots & \ddots & \vdots \\ \tilde{\phi}_{S^w} & \dots & \tilde{\phi}_{S^w}^J \end{bmatrix} = \begin{bmatrix} \tilde{\phi}_1 \mathcal{M}_0 \\ \tilde{\phi}_2 \mathcal{M}_0 \mathcal{M}_1 \\ \vdots \\ \tilde{\phi}_{S^w} \mathcal{M}_0 \prod_{s=1}^{S^w-1} \mathcal{M}_s \end{bmatrix}$$

and the normalized distribution of human capital (i.e. per-capita human capital by age and sector) is constant over time and given by

$$\tilde{H} = \begin{bmatrix} \tilde{h}_1^1 & \dots & \tilde{h}_1^J \\ \vdots & \ddots & \vdots \\ \tilde{h}_{S^w} & \dots & \tilde{h}_{S^w}^J \end{bmatrix} = h^0 \begin{bmatrix} \tilde{\phi}_1 \tilde{\Gamma} \circ \tilde{\mathcal{M}}_0 \\ \tilde{\phi}_2 \left(\tilde{\Gamma}\right)^2 \circ \left[\tilde{\mathcal{M}}_0 \tilde{\mathcal{M}}_1\right] \\ \vdots \\ \tilde{\phi}_{S^w} \left(\tilde{\Gamma}\right)^{S^w} \circ \left[\tilde{\mathcal{M}}_0 \prod_{s=1}^{S^w-1} \tilde{\mathcal{M}}_s\right] \end{bmatrix}$$

with $\tilde{\mathcal{M}}_0 = \mathcal{T}_0 \circ \left(\mathcal{M}_0\right)^{1-\frac{1}{\kappa}}$ and $\tilde{\mathcal{M}}_s = \mathcal{T}_s \circ \left(\mathcal{M}_s\right)^{1-\frac{1}{\kappa}}$, where $\circ$ denotes the element-wise (or Hadamard) product and $(\cdot)^x$ denotes the element-wise power.

The proof of Proposition 6 can be found in Appendix 2.B. A direct consequence of Proposition 6 is that the normalized (i.e. per-capita) labor supply in sector j is also constant over time and given by

$$\tilde{H}^{supply,j} = \sum_{s=1}^{S^w} \tilde{h}_s^j \tag{2.11}$$

which corresponds to the sum over column j of matrix $\tilde{H}$. Moreover, average earnings per worker in the economy are closely linked to the distributions of workers and human capital derived above. Let $\bar{\omega}_s^j$ denote the average earnings of a worker of age s working in sector j and $\bar{\omega}_s$ the average earnings of a worker of age s (i.e. across all sectors). Average earnings per worker are then given by

$$\bar{\omega}_s^j = \frac{\Gamma\left(1-\frac{1}{\kappa}\right)\lambda^j \tilde{h}_s^j \omega^j}{\tilde{\phi}_s^j} \quad \text{and} \quad \bar{\omega}_s = \sum_{j=1}^{J} \frac{\tilde{\phi}_s^j}{\tilde{\phi}_s} \bar{\omega}_s^j \tag{2.12}$$

In summary, for a given constant demographic process (i.e. a process satisfying (D1)), and a given vector of constant real wages, workers' sector choice probabilities are time invariant and fully determined by model parameters. Thus, the distributions of workers and human capital are time invariant and so are labor supply and average earnings per worker. In the next section, I turn to the production side of the economy.

2.4.3 Production Model

2.4.3.1 Environment

The production model extends the model developed by Acemoglu (2010). The production side mirrors the J sectors of the worker model. Each sector entails atomistic intermediate goods producers and a single technology monopolist that provides the technology used within the sectors. In addition, atomistic final goods producers produce the single consumption good of the economy. Intermediate goods and the final good are produced competitively. The production structure can be thought of as being organized in layers: first, sectoral producers combine human capital H^j, machines M^j, structures K^j, and technology $T^j(\theta^j)$ to produce intermediate goods Q^j; then, the final consumption good Y is produced by combining intermediate goods. Technology is at the core of the mechanism and entails two aspects: the technology level θ, which requires investment in research and development, and the quantity of technology used in production $T(\theta)$. The technology level represents the stock of patents, while the quantity of technology used in production represents the amount of licenses to utilize these patents in production bought by intermediate goods producers.[4] The technology level includes the

4. Throughout this section, 'technology level' and 'stock of patents' will be used to refer to θ and 'technology quantity' and 'licenses' will be used to refer to $T(\theta)$.

patents for any technology that helps in the automation of processes. In this analysis, the technology level is assumed to augment the machines used in the production process and thereby indirectly creates opportunities to substitute human capital with machines. Different technology levels can then be interpreted as degrees of sophistication of the machines involved: For call center operations, technology might range from manual switchboards to automatic switchboards, to computer-aided call-centers, and finally fully integrated systems using artificial intelligence (AI); For machining, technology might range from non-programmed (manual) machining to hard-wired programmed logic controllers, to programmable logic controllers, to networked controllers and finally to fully integrated tools (internet of things, IoT). In practical terms, the technology spectrum ranges from purely manual operation (human operators connecting calls or manually operating machines) to fully automated systems with minimal human involvement (chatbots with speech-to-text and text-to-speech functionalities or fully automated processing lines). The monopolists decide which level of technology is available for production and the intermediate goods producers decide how much of this technology is used in the production process.

The intermediate goods producer hires labor (measured in effective units of labor H_t^j) at nominal wage w_t^j, rents machines M_t^j and structures K_t^j at nominal rates $r_t^{M,j}$ and $r_t^{K,j}$, respectively, and purchases licenses $T_t^j(\theta_t^j)$ for using technology of level θ_t^j at nominal price $p_t^{T,j}$. The intermediate good Q^j is a Cobb-Douglas combination of an intermediate input G^j, which is in turn produced from labor, machines, and structures, and licenses. The intermediate good is sold to final goods producers at nominal price $p_t^{Q^j}$. Formally, let the production function of the representative sectoral producer be given by

$$Q_t^j = \varphi^j \left[G^j \left(H_t^j, M_t^j, K_t^j, \theta_t^j \right) \right]^\alpha \left[T_t^j(\theta_t^j) \right]^{1-\alpha} \tag{2.13}$$

with

$$G^j \left(H_t^j, M_t^j, K_t^j, \theta_t^j \right) = \left[(1-\eta) \left(\theta_t^j M_t^j \right)^{\frac{\sigma_j - 1}{\sigma_j}} + \eta \left(H_t^j \right)^{\frac{\sigma_j - 1}{\sigma_j}} \right]^{\frac{\sigma_j \nu}{\sigma_j - 1}} \left(K_t^j \right)^{1-\nu} \tag{2.14}$$

Note that the intermediate input production function G^j exhibits constant returns to scale in labor, machines, and structures and, for $\nu < 1$, decreasing returns to scale in labor and machines. The intermediate goods producers' optimization problem is

$$\begin{aligned} \max_{\{H_t^j, M_t^j, K_t^j, T_t^j(\theta_t^j)\}} \quad & p_t^{Q^j} Q_t^j - w_t^j H_t^j - r_t^{M,j} M_t^j - r_t^{K,j} K_t^j - p_t^{T,j} T_t^j(\theta_t^j) \\ \text{s.t.} \quad & (2.13),\ (2.14) \end{aligned} \tag{2.15}$$

The final goods producers purchase intermediate goods at their respective prices $p_t^{Q^j}$ to produce and sell the consumption goods at price p_t^Y, using a CES aggregator of sectoral goods[5], i.e.

$$Y_t = \left[\sum_{j}^{J} \varrho^j (Q_t^j)^{\frac{1-\psi}{\psi}} \right]^{\frac{\psi}{1-\psi}} \tag{2.16}$$

The optimization problem of final goods producers is

$$\max_{\{Q_t^j\}_{j=1}^J} \left\{ p_t^Y Y_t - \sum_{j=1} p_t^{Q^j} Q_t^j \right\} \quad \text{s.t. } (2.16) \tag{2.17}$$

The technology monopolists invest in new patents and market the existing technology to intermediate goods producers. In each period, the (pre-) existing technology of level θ_t^j (i.e. the stock of patents at the beginning of the period) is used to create licenses to use this technology $T_t^j\left(\theta_t^j\right)$ at marginal cost x_t^j, which is then sold at price $p_t^{T,j}$. The monopolists maximize the total sum of discounted utility from consumption. In every period t, the net profits from selling patents $T_t^j\left(\theta_t^j\right)$, denoted by $\Pi_t^{\theta,j}$, can either be consumed, $C_t^{M,j}$, or invested in research, $I_t^{\theta,j}$, to produce new patents and increase the technology level in period $t+1$. Consumption yields utility $U_t^M\left(C_t^{M,j}\right)$. Creating new patents incurs a cost to the monopolist, which consists of two components: a direct component, denoted by $C_t(\cdot)$ that represents the effort to produce new patents, and a cost externality $A(\cdot)$ that captures the relationship between the current technology level and the cost to produce new patents. For now, I only assume that the total cost of creating new patents is the same across all sectors and is given by the product of direct cost and externality. The direct cost is non-negative and monotonically increasing in the amount of new patents created and the externality is strictly positive for all current technology levels. Existing technology depreciates at a constant rate δ^θ (i.e. patents expire after a fixed period of time). Denote the amount of patents created in period t by $\Delta\theta_t^j$. By definition,

$$\Delta\theta_t^j = \theta_{t+1}^j - \left(1 - \delta^\theta\right)\theta_t^j$$

i.e. the amount of newly created patents in period t is the difference between the stock of patents next period and the non-depreciated part of the current stock of technology. The investment required to implement a given level of technology next period is then given by the corresponding total cost of creating the required amount of new patents:

$$I_t^{\theta,j} = A\left(\theta^j\right) C_t\left(\Delta\theta_t^j\right)$$

5. Note that this includes perfect substitutes, Cobb-Douglas, and perfect complements with $\rho \to 1$, $\rho \to 0$, and $\rho \to -\infty$, respectively

Using this relationship, patents created in period t can be defined as a function of investment and the current technology level. Denote the patent creation function by $F_t\left(I_t^{\theta,j}, \theta^j\right)$. Then,

$$I_t^{\theta,j} = A\left(\theta^j\right) C_t\left(F_t\left(I_t^{\theta,j}, \theta_t^j\right)\right) \quad \Leftrightarrow \quad F_t\left(I_t^{\theta,j}, \theta_t^j\right) = C_t^{-1}\left(\left[A(\theta^j)\right]^{-1} I_t^{\theta,j}\right)$$

With some additional technical conditions on $A(\cdot)$ and $C_t(\cdot)$, the technology creation function $F_t(\cdot)$ is well-defined and is continuous and monotonically increasing in $I_t^{\theta,j}$.[6]

The optimization problem of the technology monopolists therefore consist of two choices: how much to invest in future technology levels (new patents) and how much to charge for the use of existing technology (licenses). The optimization problem can be stated in recursive form as

$$\begin{aligned} V_t^j\left(\theta_t^j\right) &= \max_{\{p_t^{T,j}, I_t^{\theta,j}\}} U_t^M\left(\Pi_t^j\left(\theta_t^j, p_t^{T,j}\right) - I_t^{\theta,j}\right) + \beta V_{t+1}^j\left(\theta_{t+1}^j\right) \\ \text{s.t.} \quad \Pi_t^j\left(\theta_t^j, p_t^{T,j}\right) &= \left(p_t^{T,j} - x_t^j\right) T_t^j\left(\theta_t^j\right) \\ \theta_{t+1}^j &= \left(1 - \delta^\theta\right)\theta_t^j + F_t\left(I_t^{\theta,j}, \theta_t^j\right) \\ p_t^{T,j} &\in \mathbb{R}_0^+, \; I_t^{\theta,j} \in \left[0, \Pi_t^j\left(\theta_t^j, p_t^{T,j}\right)\right] \end{aligned} \tag{2.18}$$

In the next section, I present some general properties of the production environment, before again normalizing the production model by the total size of the population.

2.4.3.2 Properties of the Production Model

The production structure is composed of mostly standard components: human capital and machines are bundled in a CES production function with machine-augmenting technology, these bundles are then combined with structures in a Cobb-Douglas production function to produce intermediate inputs G^j; the intermediate inputs are, in turn, combined with technology licenses T in another Cobb-Douglas function to produce intermediate goods, which are finally combined to the consumption good in a CES aggregator. The use of standard components allows for the derivation of unitary cost and demand functions, building on the known results for CES and Cobb-Douglas production functions. This is summarized in Proposition 7, the proof of which can again be found in Appendix 2.B.

6. The conditions are that both functions are continuous, that domain and range of $C(\cdot)$ are the non-negative real numbers and that $C(\cdot)$ and $A(\cdot)$ are bounded over closed sets. For more details see the proof of Proposition 8 in Appendix 2.B.

Proposition 7 (Unitary cost and demand functions). *Suppose the production structure is given by equations (2.16), (2.13) and (2.14) and intermediate and final goods are produced competitively. Then, the unit demand functions, i.e. the quantities used to produce one unit of the final consumption good are, for all sectors j,*

$$Q_t^j = \left(\varrho^j \frac{p_t^Y}{p_t^{Q^j}}\right)^{\psi} Y_t \tag{2.19}$$

$$G_t^j = \alpha \frac{p_t^{Q^j}}{p_t^{G^j}} Q_t^j \tag{2.20}$$

$$H_t^j = \nu \left(\eta \frac{p_t^{B^j}}{w^j}\right)^{\sigma_j} \frac{p_t^{G^j}}{p_t^{B^j}} G_t^j \tag{2.21}$$

$$M_t^j = \nu \left((1-\eta) \frac{\theta_t^j p_t^{B^j}}{r_t^{M,j}}\right)^{\sigma_j} \frac{p_t^{G^j}}{\theta_t^j p_t^{B^j}} G^j \tag{2.22}$$

$$K_t^j = (1-\nu) \frac{p_t^{G^j}}{r_t^{K,j}} G_t^j \tag{2.23}$$

Moreover, the equilibrium prices of the final consumption good, intermediate goods, intermediate inputs, and intermediate bundles are equal to the minimal unitary cost of producing the respective goods, i.e. $p_t^Y = c_t^Y$ *and* $\{p_t^{Q^j}, p_t^{G^j}, p_t^{B^j}\}_{j=1}^J = \{c_t^{Q^j}, c_t^{G^j}, c_t^{B^j}\}_{j=1}^J$, *where minimal unitary cost are given by*

$$c_t^Y = \begin{cases} \prod_{j=1}^J \left(\frac{c_t^{Q^j}}{\varrho^j}\right)^{\varrho^j} & \text{if } \psi = 1 \\ \left[\sum_j^J (\varrho^j)^{\psi} (c_t^{Q^j})^{1-\psi}\right]^{\frac{1}{1-\psi}} & \text{if } \psi \neq 1 \end{cases} \tag{2.24}$$

$$c_t^{Q^j} = (\varphi^j)^{-1} \left(\frac{c_t^{G^j}}{\alpha}\right)^{\alpha} \left(\frac{x_t^j}{1-\alpha}\right)^{1-\alpha} \tag{2.25}$$

$$c_t^{G^j} = \left(\frac{c_t^{B^j}}{\nu}\right)^{\nu} \left(\frac{r_t^{K,j}}{1-\nu}\right)^{1-\nu} \tag{2.26}$$

$$c_t^{B^j} = \begin{cases} \left(\frac{r_t^{M,j}}{(1-\eta)\theta_t^j}\right)^{1-\eta} \left(\frac{w_t^j}{\eta}\right)^{\eta} & \text{if } \sigma_j = 1 \\ \left[(1-\eta)^{\sigma_j} \left(\frac{r_t^{M,j}}{\theta_t^j}\right)^{1-\sigma_j} + \eta^{\sigma_j} \left(w_t^j\right)^{1-\sigma_j}\right]^{\frac{1}{1-\sigma_j}} & \text{if } \sigma_j \neq 1 \end{cases} \tag{2.27}$$

The non-standard components of this production environment are level and quantity of technology and the role of monopolists in providing it. The technology monopolists solve their optimization problem subject to the demand for technology licenses $T^j(\theta^j)$ by intermediate goods producers. First, note that demand for licenses follows directly from the Cobb-Douglas specification of intermediate goods production. Taking derivatives of (2.15) w.r.t. T_t^j and rearranging yields

$$T_t^j\left(\theta_t^j\right) = \left((1-\alpha)\varphi^j \frac{p_t^{Q^j}}{x_t^j}\right)^{\frac{1}{\alpha}} G^j\left(H_t^j, M_t^j, K_t^j, \theta_t^j\right) \tag{2.28}$$

This has two implications: first, the intermediate goods production function can be simplified by substituting (2.28) to

$$Q_t^j = \left(\varphi^j\right)^{\frac{1}{\alpha}} \left((1-\alpha)\frac{p_t^{Q^j}}{x_t^j}\right)^{\frac{1-\alpha}{\alpha}} G^j\left(H_t^j, M_t^j, K_t^j, \theta_t^j\right) \tag{2.29}$$

Thus, as $G^j(\cdot)$ exhibits constant returns to scale in $\left\{H_t^j, M_t^j, K_t^j\right\}$, so does $Q^j(\cdot)$ for all sectors j. Combining this result across sectors, the final goods production function exhibits constant returns to scale in $\left\{H_t^j, M_t^j, K_t^j\right\}_{j=1}^J$. In other words, if all input factors in all sectors are multiplied by a constant, total output is also multiplied by that same constant.

Second, demand for technology licenses itself exhibits constant returns to scale in input factors. Both of these properties allow to re-write the production model in terms of per-capita quantities.

2.4.3.3 Normalization

I apply two normalizations to the production model: As with the worker model, I now normalize the production side of the model in per-capita terms. Moreover, I normalize all prices with the price of the final consumption good, obtaining real prices. As detailed above, intermediate goods production and final goods production exhibit constant returns to scale in $\{H_t^j, M_t^j, K_t^j\}_{j=1}^J$, so normalizing goods production is straightforward. Let $\tilde{X}_t^j = \phi_t^{-1} X_t^j$ for $\left\{H_t^j, M_t^j, K_t^j, G_t^j, Q_t^j\right\}_{j=1}^J$ and $\tilde{Y}_t = \phi_t^{-1} Y_t$ for all periods t. Unitary cost functions only depend on factor prices, hence equations (2.24)–(2.27) expressed in real prices become

$$\pi_t^Y = 1 \tag{2.30}$$

$$\pi_t^{Q^j} = (\varphi^j)^{-1} \left(\frac{\pi_t^{G^j}}{\alpha}\right)^{\alpha} \left(\frac{\pi_t^{T,j}}{1-\alpha}\right)^{1-\alpha} \tag{2.31}$$

$$\pi_t^{G^j} = \left(\frac{\pi_t^{B^j}}{\nu}\right)^{\nu} \left(\frac{\rho_t^{K,j}}{1-\nu}\right)^{1-\nu} \tag{2.32}$$

$$\pi_t^{B^j} = \begin{cases} \left(\frac{\rho_t^{M,j}}{(1-\eta)\theta_t^j}\right)^{1-\eta} \left(\frac{\omega_t^j}{\eta}\right)^{\eta} & \text{if } \sigma_j = 1 \\ \left[\left(1-\eta\right)^{\sigma_j} \left(\frac{\rho_t^{M,j}}{\theta_t^j}\right)^{1-\sigma_j} + \eta^{\sigma_j} (\omega_t^j)^{1-\sigma_j}\right]^{\frac{1}{1-\sigma_j}} & \text{if } \sigma_j \neq 1 \end{cases} \tag{2.33}$$

where $\rho_t^{K,j} = \frac{r_t^{K,j}}{p_t^Y}$, $\rho_t^{M,j} = \frac{r_t^{M,j}}{p_t^Y}$, $\omega_t^j = \frac{w_t^j}{p_t^Y}$, and $\pi_t^{T,j} = \frac{p_t^{T,j}}{p_t^Y}$.

Per-capita unit demand functions (i.e. input quantities required to produce one unit of output per capita) are

$$\tilde{Q}_t^j = \left(\varrho^j \frac{1}{\pi_t^{Q^j}}\right)^{\psi} \tilde{Y}_t \tag{2.34}$$

$$\tilde{G}_t^j = \alpha \frac{\pi_t^{Q^j}}{\pi_t^{G^j}} \tilde{Q}_t^j \tag{2.35}$$

$$\tilde{H}_t^j = \nu \left(\eta \frac{\pi_t^{B^j}}{w^j}\right)^{\sigma_j} \frac{\pi_t^{G^j}}{\pi_t^{B^j}} \tilde{G}_t^j \tag{2.36}$$

$$\tilde{M}_t^j = \nu \left((1-\eta) \frac{\theta_t^j \pi_t^{B^j}}{\rho_t^{M,j}}\right)^{\sigma_j} \frac{\pi_t^{G^j}}{\theta_t^j \pi_t^{B^j}} \tilde{G}^j \tag{2.37}$$

$$\tilde{K}_t^j = (1-\nu) \frac{\pi_t^{G^j}}{\rho_t^{K,j}} \tilde{G}_t^j \tag{2.38}$$

The normalization of the technology monopolists' problem requires additional assumptions on monopolists' utility and on the cost of creating new patents. The necessary assumptions and the resulting normalized optimization problem are summarized in Proposition 8.

Proposition 8 (Normalization of the technology monopolists' problem). *Assume that:*

(P1) *The monopolists' utility is linear in per-capita profits, i.e.*

$$U_t^M\left(C_t^{M,j}\right) = \frac{C_t^{M,j}}{\phi_t} = \frac{\Pi^j\left(\theta^j, p_t^{T,j}\right) - I_t^{\theta,j}}{\phi_t} = \tilde{\Pi}^j\left(\theta^j, p_t^{T,j}\right) - \tilde{I}_t^{\theta,j}$$

and the monopolists' discount factor satisfies $\beta_t = \frac{1}{1+\rho_t}$.

(P2) *The per-capita direct cost of inventing new patents* $\tilde{C}(\cdot)$ *is constant over time, i.e.*

$$C_t\left(\Delta\theta_t^j\right) = \phi_t \tilde{C}\left(\Delta\theta_t^j\right)$$

and is twice continuously differentiable, non-negative, strictly increasing, strictly convex, has range $\mathbb{R}_0^+$, *is bounded over closed sets, and satisfies* $\tilde{C}(0) = 0$.

(P3) *The technology externality* $A(\cdot)$ *is twice continuously differentiable, strictly positive, non-increasing, convex, bounded, and satisfies*

$$-\left[A\left(\theta^j\right)\right]^{-1} \frac{dA\left(\theta^j\right)}{d\theta^j} \leq 1 \quad \forall \theta^j \in \Theta \ \forall j$$

and

$$-\theta^j\left[A\left(\theta^j\right)\right]^{-1}\frac{dA\left(\theta^j\right)}{d\theta^j} \;\leq\; \min\left\{\frac{(1-\nu)}{2},\;\frac{1}{2\sigma_j}\right\} \quad \forall\theta^j \;\in\; \Theta \;\;\forall j$$

and, if it is strictly convex, satisfies

$$\left[A\left(\theta^j\right)\left(\frac{d^2A\left(\theta^j\right)}{\left(\theta^j\right)^2}\right)\right]^{-1}\left[\frac{dA\left(\theta^j\right)}{d\theta^j}\right]^2 \;\leq\; \frac{1}{2} \quad \forall\theta^j \;\in\; \Theta \;\;\forall j$$

(P4) *If the technology externality is not constant, the two functions satisfy*

$$\left[A\left(\theta^j\right)\left(\frac{d^2A\left(\theta^j\right)}{\left(\theta^j\right)^2}\right)\right]^{-1}\left[\frac{dA\left(\theta^j\right)}{d\theta^j}\right]^2 \;\leq\; \tilde{C}\left(\Delta\theta^j\right)\left(\frac{d^2\tilde{C}\left(\Delta\theta^j\right)}{d\left(\Delta\theta^j\right)^2}\right)\left[\frac{d\tilde{C}\left(\Delta\theta^j\right)}{d\Delta\theta^j}\right]^{-2}$$

$$\forall(\theta^j,\theta^{j\prime}) \;\in\; X^j \;\;\forall j$$

Then, for $0<\sigma_j<1$, *if real input prices and per-capita stocks of human capital, machines, and structures are strictly positive and constant over time, i.e. if*

$$\left\{\omega_t^j,\rho_t^{M,j},\rho_t^{K,j},\xi_t^j,\tilde{H}_t^j,\tilde{M}_t^j,\tilde{K}_t^j\right\}_{j=1}^J \;=\; \left\{\omega^j,\rho^{M,j},\rho^{K,j},\xi^j,\tilde{H}^j,\tilde{M}^j,\tilde{K}^j\right\}_{j=1}^J \;>\; 0$$

(1) the monopolist optimization problem is stationary in per-capita terms, i.e. (2.18) is equivalent to

$$\begin{aligned} V^j\left(\theta^j\right) \;&=\; \max_{\tilde{I}^{\theta,j}}\left\{\tilde{\Pi}^{j*}(\theta^j)-\tilde{I}^{\theta,j}+\frac{1}{1+\rho}V^j\left(\theta^{j\prime}\right)\right\} \\ \text{s.t.}\quad \tilde{\Pi}^{j*}(\theta^j) \;&=\; (1-\alpha)\,\pi^{G^j}G^j\left(\tilde{H}^j,\tilde{M}^j,\tilde{K}^j,\theta^j\right) \\ \theta^{j\prime} \;&=\; \left(1-\delta^\theta\right)\theta^j+\tilde{F}\left(\tilde{I}^{\theta,j},\theta^j\right) \\ \tilde{I}^{\theta,j} \;&\in\; \left[0,\tilde{\Pi}^{j*}\left(\theta^j\right)\right] \end{aligned} \tag{2.39}$$

where $\tilde{F}\left(\tilde{I}^{\theta,j},\theta^j\right)=\tilde{C}^{-1}\left(\left[A\left(\theta^j\right)\right]^{-1}\tilde{I}^{\theta,j}\right)$,

(2) the optimal real price of licenses is a constant markup over real marginal cost

$$\pi^{T,j}=\frac{\xi^j}{1-\alpha}$$

where $\xi^j=\frac{x^j}{c_t^Y}$ *is the real marginal cost of producing one license for technology in sector j.*

(3) a solution to (2.39) exists, is unique, strictly increasing, and continuously differentiable, and the investment policy function $\tilde{I}^{\theta,j*}(\theta^j)$ *is single-valued and continuous and satisfies*

$$\frac{\partial\tilde{F}^j\left(\tilde{I}^{\theta,j*}\left(\theta^j\right),\theta^j\right)}{\partial\tilde{I}^{\theta,j}}\left(\left.\frac{dV^j(z)}{dz}\right|_{z=\left(1-\delta^\theta\right)\theta^j+\tilde{F}^j\left(\tilde{I}^{\theta,j*}\left(\theta^j\right),\theta^j\right)}\right) \;=\; 1+\rho$$

The proof of Proposition 8 can be found in Appendix 2.B. The basic intuition behind the proposition is that a solution is guaranteed if the total cost of creating new patents is convex. In the absence of a cost externality (i.e. if $A(\cdot)$ is constant), convexity of $C(\cdot)$ suffices. If the cost externality is not constant, it needs to be decreasing in θ^j (to ensure that monotonicity of profits is preserved in the objective function) and sufficiently convex (to ensure that total cost is non-negative). If this is the case, i.e. if $A(\cdot)$ is decreasing and convex, the convexity of the indirect cost function must be small (absolute and relative to the direct cost function) as to preserve convexity of the total cost function (and, thus, concavity of the constraint set).

For illustration, suppose for now that the both cost functions are power functions, i.e.

$$A\left(\theta^j\right) = \left(\theta^j + 1\right)^{-\varsigma_e} \quad \text{and} \quad \tilde{C}\left(\Delta\theta^j\right) = \iota\left(\Delta\theta^j\right)^{\varsigma_d}$$

Then, the conditions regarding the cost of creating new patents become simple conditions on the exponents of the cost functions: assumptions (P2)–(P4) are satisfied if

$$\varsigma_e \in \left(0, \frac{1-\nu}{2}\right] \quad \text{and} \quad \varsigma_d \geq \varsigma_e + 1$$

In other words, if the cost externality is mild (ς_e is small) and the direct cost function is sufficiently convex to outweigh the externality, a solution to the monopolists' problem exists and has the aforementioned characteristics.

Returning to the general case, I now show that there is a stationary technology level, i.e. a stock of patents that is constant over time.

Corollary 9 (Stationary technology levels). *Assume that assumptions (P1)–(P4) hold. Then, for $0 < \sigma_j < 1$, if real input prices and per-capita stocks of human capital, machines, and structures are constant over time, a stationary technology level exists, i.e.*

$$\forall j \; : \; \exists \theta^{j*} s.th. \;\; \left(1 - \delta^{\theta}\right)\theta^{j*} + \tilde{F}^j\left(\tilde{I}^{\theta,j*}\left(\theta^{j*}\right), \theta^{j*}\right) = \theta^{j*}$$

The stationary technology level is characterized by

$$\frac{d\tilde{\Pi}^{j*}\left(\theta^{j*}\right)}{d\theta^j} = \left(\rho + \delta^{\theta}\right)A\left(\theta^{j*}\right)\frac{d\tilde{C}\left(\delta^{\theta}\theta^{j*}\right)}{d\Delta\theta^j} + \frac{dA\left(\theta^{j*}\right)}{d\theta^j}\tilde{C}\left(\delta^{\theta}\theta^{j*}\right) \tag{2.40}$$

The proof of Corollary 9 can be found in Appendix 2.B. Note that strict concavity of monopolists' profits[7] imply that the LHS of condition (2.40) is strictly decreasing. This, in turn, means that if the RHS of (2.40) is non-decreasing, the

7. See the proof of Proposition 8 for further details on the properties of the monopolists' profit functions.

stationary technology level must be unique. Taking derivatives of the RHS w.r.t. θ^j yields

$$\begin{aligned}\frac{dRHS\left(\theta^j\right)}{d\theta^j} &= \delta^\theta\left(\rho+\delta^\theta\right)A\left(\theta^j\right)\frac{d^2\tilde{C}\left(\delta^\theta\theta^j\right)}{d\left(\Delta\theta^j\right)^2} \\ &\quad+\left(\rho+2\delta^\theta\right)\frac{dA\left(\theta^j\right)}{d\theta^j}\frac{d\tilde{C}\left(\delta^\theta\theta^j\right)}{d\Delta\theta^j} \\ &\quad+\frac{d^2A\left(\theta^j\right)}{d\left(\theta^j\right)^2}\tilde{C}\left(\delta^\theta\theta^j\right)\end{aligned}$$

In the absence of a cost externality (i.e. with $\frac{dA(\theta^j)}{d\theta^j}=0$), the RHS is strictly increasing and the stationary technology level is unique. With $\frac{dA(\theta^j)}{d\theta^j}<0$, monotonicity of the RHS depends on the functional forms of $A(\cdot)$ and $\tilde{C}(\cdot)$ and the parameters of the model.

Suppose as before that both cost functions are power functions. Uniqueness of stationary technology levels is the guaranteed by the conditions $\varsigma_e \in \left(0, \frac{1-\nu}{2}\right]$ and $\varsigma_d \geq \varsigma_e + 1$ derived above.[8] This results completes the production side of the model. To close the model, I now turn to a description of the financial sector of the economy.

2.4.4 Investment

The financial markets act as intermediaries between workers and firms for the supply of physical capital. They collect savings from workers during working ages, invest in machines and structures, and then rent out these machines and structures to firms. The proceeds from renting out physical capital are then returned to retired workers.

Recall that workers save a fixed fraction τ^p of their earnings and that pension payments are proportional to the absolute value of human capital at retirement. Aggregating over all workers within a given cohort then yields total savings of all workers of a given age or total payments to retirees of said age, respectively (depending on whether the given cohort is working or has already retired). The total supply of funds (or total balance in the financial markets) from the cohort that is of a given age in a given period then consists of all accumulated past savings and accrued interest minus all past retirement payments. Moreover, by aggregating over all workers within a cohort, individual contributions (and, thus, earnings) can be replaced with average contributions (earnings). To be precise, consider the cohort that is of age s in period t. In the first period after model entry of this cohort, $t-s+1$, the workers have then been of model age

8. See Appendix 2.B for details.

1 and have saved a total amount $\phi_{t-s+1,1}\tau^p\bar{\omega}_{t-s+1,1}$. These savings have accumulated interest with compound interest factor $\prod_{m=1}^{s-1}\left(1+\rho_{t-s+m}\right)$ until period t. Similarly, the contributions from period $t-s+2$, in which the workers have been of model age 2, have accumulated interest with compound interest factor $\prod_{m=1}^{s-2}\left(1+\rho_{t-s+m}\right)$. If the cohort is older than S^w, the workers have saved until period $t-s+S^w$ and, starting from period $t-s+S^w+1$, received payments from the financial markets. The total payout received at age S^w+1 (i.e. in period $t-s+S^w+1$) was $-\phi_{t-s+S^w+1,S^w+1}\bar{b}_{t-s+S^w}\omega_{t-s+S^w,S^w}$ and has since accumulated interest with compound interest factor $\prod_{m=1}^{s-(S^w+1)}\left(1+\rho_{t-s+m}\right)$ Summing over all periods in which the workers of age s have been active in the economy yields the balance of funds of a given cohort:

$$B_{t,s} = \phi_{t-s,0}\sum_{m=1}^{s}\left[\left(\prod_{l=1}^{m-1}\zeta_{t-s+l,l}\right)\left(\prod_{l=m}^{s-1}\left(1+\rho_{t-s+l}\right)\right)\right.$$
$$\left.\left[\mathbb{1}_{\{m\leq S^w\}}\tau^p\bar{\omega}_{t-s+m,m} - \mathbb{1}_{\{m>S^w\}}\bar{b}_{t-s+S^w}\bar{\omega}_{t-s+S^w,S^w}\right]\right]$$

The total supply of funds in the economy is the given by summing over all cohorts, i.e. $B_t = \sum_{s=1}^{S^w+S^r} B_{t,s}$.

All available funds are invested across sectors, either in machines, denoted by $I_t^{M,j}$, or in structures, denoted by $I_t^{K,j}$. Further, the existing stocks of machines and structures depreciate at rates δ^M and δ^K, respectively. The stocks of physical assets thus evolve according to

$$M_{t+1}^j = \left(1-\delta^M\right)M_t^j + I_t^{M,j} \quad \text{and} \quad K_{t+1}^j = \left(1-\delta^K\right)K_t^j + I_t^{K,j} \quad \forall j,t \tag{2.41}$$

The real gross returns in the financial markets are the real rental rates on the existing stock of physical capital, i.e. $\rho_t^{M,j}M_t^j$ and $\rho_t^{K,j}K_t^j$. Financial markets are assumed to be efficient, i.e. there is no opportunity for arbitrage across asset classes. In other words, the net returns on all assets are equal to the real interest rate

$$\rho_t^{M,j} - \delta^M = \rho_t^{K,j} - \delta^K = \rho_t \quad \forall j,t \tag{2.42}$$

Note that this implies that rental rates for machines and structures are identical across sectors. Any change in the real interest rate (or, more generally, in the relative availability of capital) thus affects all asset classes and all sectors in the same way. Using the no-arbitrage condition (2.42), total net investment in period t can be written as

$$\sum_{j=1}^{J}\left(I_t^{M,j} - \delta^M M_t^j + I_t^{K,j} - \delta^K K_t^j\right) = \rho_t B_t + \sum_{s=1}^{S^w}\phi_{t,s}\tau^p\bar{\omega}_{t,s}$$
$$- \sum_{s=S^w+1}^{S^w+S^r}\phi_{t,s}\bar{b}_{t-s+S^w}\bar{\omega}_{t-s+S^w,S^w}$$

where the first summand on the right hand side is the net returns on all assets, the second summand is the total inflow of savings from workers, and the third summand is the total outflow of annuity payments to retirees. Note that this is an accounting identity: all inflows into the financial market (either as savings from workers or rental payments from firms) are either paid out as retirement payments to workers or invested in new physical capital.

If real interest rates and the retirement factor are constant over time and some conditions for the worker problem are satisfied, normalizing the investment sector of the economy is straightforward. In particular, suppose that $\rho_t = \rho \ \forall t$ and $\bar{b}_t = \bar{b} \ \forall t$ and that assumptions (W1)– (W7) hold. Equation (2.42) becomes

$$\rho^{M,j} - \delta^M = \rho^{K,j} - \delta^K = \rho \quad \forall j \tag{2.43}$$

and the total per-capita supply of physical capital is then constant over time and given by

$$\begin{aligned}\tilde{B} = \bar{\phi}^{-1} \sum_{s=1}^{S^w+S^r} \Bigg\{ (1+n)^{-s} \sum_{m=1}^{s} \Bigg[\Bigg(\prod_{l=1}^{m-1} \zeta_l \Bigg) (1+\rho)^{s-m} \\ \Big[\mathbb{1}_{\{m \leq S^w\}} \tau^p \bar{\omega}_m - \mathbb{1}_{\{m > S^w\}} \bar{b} \bar{\omega}_{S^w} \Big] \Bigg] \Bigg\}\end{aligned} \tag{2.44}$$

Total net per-capita investments is constant as well and given by

$$\sum_{j=1}^{J} \left(\tilde{I}_t^{M,j} - \delta^M \tilde{M}_t^j + \tilde{I}_t^{K,j} - \delta^K \tilde{K}_t^j \right) = \rho \tilde{B} + \tau^p \sum_{s=1}^{S^w} \tilde{\phi}_s \bar{\omega}_s - \bar{b} \sum_{s=S^w+1}^{S^w+S^r} \tilde{\phi}_s \bar{\omega}_{S^w} \tag{2.45}$$

Closing the model requires an assumption on how the pension factor $\bar{b}$ is determined. Note that this modeling choice has implications on the redistribution of resources across cohorts and, consequently, on worker welfare. On the other hand, this choice determines net investments in new physical capital. Stationary stocks of physical capital require zero net investment, as implied by the following assumption:

(I1) *The retirement factor $\bar{b}$ satisfies*

$$\rho \tilde{B} = \bar{b} \sum_{s=S^w+1}^{S^w+S^r} \tilde{\phi}_s \bar{\omega}_{S^w} - \tau^p \sum_{s=1}^{S^w} \tilde{\phi}_s \bar{\omega}_s \tag{2.46}$$

Assumption (I1) implies financial markets pool the current proceeds from renting out capital with the current contributions of all workers and then pay out everything to all retirees.[9] It is per se unclear, why the capital markets would offer precisely this contract to workers. The financial markets in this model are purely mechanical and motivating the assumption would require a richer model. For an alternative interpretation of this modelling choice, suppose that, instead of decentralized financial markets, the intermediation of capital supply and capital demand is carried out by a centralized pension system. If the pension system is tasked with providing pensions to retirees and replacing the depreciated part of the capital stock while collecting contributions from workers and rental rates from firms, Assumption (I1) is equivalent to a per-period budget balance requirement on the pension system. As mentioned before, this assumption naturally has implications for redistribution across cohorts. The scope of this essay is to assess a particular channel that connects demographic change with technology choices in the setup of the production process and with aggregate labor shares across sectors. While important, welfare consequences on cohorts of workers implied by these mechanics are beyond the scope and left for future research. The assumptions made are therefore innocuous for the purpose of this analysis.

Finally, note that, with zero net investment, aggregate stocks of machines and structures satisfy

$$\tilde{I}^{M,j} = \delta^M \tilde{M}^j \quad \text{and} \quad \tilde{I}^{K,j} = \delta^K \tilde{K}^j \qquad \forall j \tag{2.47}$$

2.4.5 Equilibrium

With all components of the model in place I now turn to defining the equilibrium and a discussion of equilibrium conditions and properties. First, I define the equilibrium for the general case.

Definition 10 (General equilibrium)**.** Given, an initial population of workers and their human capital $\{\Phi_0, H_0\}$, an exogenous demographic process $\{\phi_{t,0}, \{\zeta_{t,s}\}_{s=0}^{S^w+S^r}\}_{t=0}^{\infty}$, an exogenous sequence of real marginal cost of producing technology licenses $\{\{\xi_t^j\}_{j=1}^J\}_{t=0}^{\infty}$, and initial stocks of machines and structures $\{M_0^j, K_0^j\}_{j=1}^J$ a **general equilibrium** is a collection of

(i) a system of real prices $\{\rho_t, \{\omega_t^j, \rho_t^{M,j}, \rho_t^{K,j}, \pi_t^j\}_{j=1}^J\}_{t=0}^{\infty}$,

(ii) aggregate stocks of human capital, machines, and structures, $\{\{H_t^j, M_t^j, K_t^j\}_{j=1}^J\}_{t=0}^{\infty}$, aggregate levels of investment, $\{\{I_t^{M,j}, I_t^{K,j}\}_{j=1}^J\}_{t=0}^{\infty}$, and aggregate output, $\{Y_t\}_{t=0}^{\infty}$,

(iii) distributions of workers and human capital $\{\Phi_t, H_t\}_{t=0}^{\infty}$,

9. Note that the constraint needs to hold only in the aggregate. Individual benefits still depend on individual human capital upon retirement, thus reflecting the individual workers' employment histories and past contributions.

(iv) technology investment policies $\{\{I_t^{\theta,j*}(\theta^j)\}_{j=1}^J\}_{t=0}^\infty$ and technology levels $\{\{\theta_t^j\}_{j=1}^J\}_{t=0}^\infty$,

(v) retirement factors $\{\bar{b}_t\}_{t=0}^\infty$, and

(vi) expected worker values and sector choice probabilities $\{\mathscr{M}_{t,0}, \{\{v_{t,s}^j\}_{j=1}^J, \mathscr{M}_{t,s}\}_{s=1}^{S^w}\}_{t=0}^\infty$

such that, in every period t,

(a) given wages and initial human capital, worker value functions and sector choices solve the individual workers' optimization problem, i.e. $\{\mathscr{M}_{t,0}, \{\{v_{t,s}^j\}_{j=1}^J, \mathscr{M}_{t,s}\}_{s=1}^{S^w}\}_{t=0}^\infty$ solve the system of equations (2.8),

(b) given prices, aggregate stocks of human and physical capital $\{\{H_t^j, M_t^j, K_t^j\}_{j=1}^J\}_{t=0}^\infty$ and aggregate output $\{Y_t\}_{t=0}^\infty$ satisfy (2.20)–(2.23),

(c) given prices and levels of inputs, technology pricing $\{\{\pi_t^{T,j}\}_{j=1}^J\}_{t=0}^\infty$ and technology investment policy functions $\{\{I_t^{\theta,j*}(\theta_t^j)\}_{j=1}^J\}_{t=0}^\infty$ solve (2.18),

(d) given worker choices, the human capital demand by sector is consistent with the distribution of human capital H_t,

(e) real rental rates satisfy (2.42) and capital markets clear, i.e. $\sum_{j=1}^J (M_t^j + K_t^j) = B_t$, and

(f) the aggregate stocks of physical capital are consistent with net investment, i.e. (2.41) holds.

Next, consider the normalized version of the model. As shown before, all components of the model can be represented in per-capita terms. Further, under the appropriate restrictions on parameters, all components of the model are time-invariant in per-capita terms. An equilibrium can thus be described as a collection of constant variables, instead of infinite sequences as above. The equilibrium definition can be refined for this special case:

Definition 11 (Balanced-growth equilibrium)**.** Given a time-invariant exogenous demographic process $\{n, \{\zeta_s\}_{s=0}^{S^w+S^r}\}$ and an exogenous constant real marginal cost of producing technology licenses $\{\xi^j\}_{j=1}^J$, a **balanced-growth equilibrium** is a collection of

(i) a constant real interest rate, constant real wages, and constant real rental rates for machines and structures $\{\rho, \{\omega^j, \rho^{M,j}, \rho^{K,j}\}_{j=1}^J\}$,

(ii) constant per-capita-levels of human capital, machines, structures, investment, and output $\{\{\tilde{H}^j, \tilde{M}^j, \tilde{K}^j, \tilde{I}^{M,j}, \tilde{I}^{K,j}\}_{j=1}^J, \tilde{Y}\}$,

(iii) constant per-capita distributions of workers and human capital $\{\tilde{\Phi}, \tilde{H}\}$,

(iv) time-invariant per-capita technology investment policies $\{\tilde{I}^{\theta,j*}(\theta^j)\}_{j=1}^J$ and constant technology levels $\{\theta^{j*}\}_{j=1}^J$,

(v) a constant retirement factor $\bar{b}$, and

(vi) time-invariant expected worker values and sector choice probabilities $\{\mathscr{M}_0, \{v_s^j\}_{j=1}^J, \mathscr{M}_s\}_{s=0}^{S^w}$

such that

(a) given wages and initial human capital, worker value functions and sector choices solve the individual workers' optimization problem, i.e. $\{\mathscr{M}_0, \{\{v_s^j\}_{j=1}^J, \mathscr{M}_s\}_{s=1}^{S^w}\}$ solve the system of equations (2.10),

(b) given input prices, $\{\tilde{H}^j, \tilde{M}^j, \tilde{K}^j\}_{j=1}^J$ and $\tilde{Y}$ satisfy (2.34)–(2.38) (subject to (2.30)–(2.33)),

(c) given prices and per-capita levels of inputs, technology investment policy functions $\tilde{I}^{\theta,j*}(\theta^j)$ solve (2.39),

(d) technology levels satisfy the stationarity condition (2.40),

(e) given worker choices, the per-capita human capital demand by sector is consistent with the per-capita distribution of human capital $\tilde{H}$,

(f) real rental rates satisfy (2.43) and capital markets clear, i.e. $\sum_{j=1}^J(\tilde{M}^j + \tilde{K}^j) = \tilde{B}$, and

(g) the aggregate stocks of physical capital are consistent with net investment, i.e. (2.47) holds.

Building on the results derived above, it is straightforward to show that a balanced-growth equilibrium exists if assumptions (W5) and (W6) for the worker problem, assumptions (P1)–(P4) for the production problem, and Assumption (I1) for the investment model hold. To begin, suppose that real input prices and per-capita stocks of human capital, machines and structures are constant over time. Then, using Proposition 8 and Corollary 9, the monopolists' problem is stationary, has a unique solution and a stationary technology level exists. Now, if technology levels are constant over time (together with real input prices and per-capita stocks of human capital, machines and structures), so are intermediate and final goods prices. Moreover, with constant prices for intermediate and final goods, factor demand is constant (in per-capita terms).

Proposition 6 then states that constant real wages imply constant per-capita supply of human capital. As shown in Section 2.4.4, constant real wages, a constant real interest rate, and a constant retirement payment factor imply constant per-capita supply of physical capital and constant total net investment. Together with the no-arbitrage condition, this implies that net investment in all assets is constant over time. Moreover, net investment is zero by Assumption (I1). Thus, stocks of machines and structures are also constant over time, validating the initial supposition.

As shown before, with power cost functions for the creation of new patents, stationary technology levels, and therefore the balanced-growth equilibrium, are unique. This result closes the model section.

2.5 Results

2.5.1 Parametrization

For illustrating the mechanism, I first solve a simple stylized parametrization of the model. The simple model consists of two sectors that are supposed to represent the production domain (sector A) and the service domain (sector B) of the economy. The only difference between the two sectors is how easily labor and machines can be substituted within the sector. In sector A ('production'), labor and machines are relatively more substitutable, while in sector B ('services'), labor and machines are less substitutable. This is captured by the elasticity of substitution between labor and machines σ_j: in sector A, labor and machines are weak complements ($\sigma_A = 0.8$, corresponding to a substitution parameter of −0.25); in sector B, labor and machines are strong complements ($\sigma_B = 0.2$, corresponding to a substitution parameter of −4.0). Other than that, the setup is entirely symmetrical.

A model period corresponds to a year. Workers are active in the labor market for 45 years, after which they are retired for at most 35 years. The population growth rate in the baseline parametrization is zero and survival probabilities are the empirical survival frequencies by age observed in the year 2000 from the WPP.

Input weights in the production process are set to match factor shares. Target factor shares are 0.6 for labor and 0.4 for capital, which is further split into structures (0.23), machines (0.1), and technology (0.07). The split within capital matches average shares of investment across these categories as observed in the VGR data. The corresponding model parameters are $\alpha = 0.93$ and $\nu = 0.75$. To preserve as much symmetry as possible, the weight of human capital in intermediate bundle production is identical across sectors and set to $\eta = 0.65$. The elasticity of substitution for intermediate goods in final goods production is set to $\psi = 0.2$, corresponding to a substitution parameter of −4.0, which is standard in the literature for cross-sector elasticities.

For the cost of technology production (patent creation and licensing), I employ an ad-hoc parametrization: the marginal cost of creating technology licenses is set to $\xi = 0.05$ (which is in the same order of magnitude as rental rates for machines and structures); the cost of patent creation is assumed to follow power cost functions with scale parameter $\iota = 0.005$, and exponents $\varsigma_d = 1.2$ (direct cost) and $\varsigma_e = 0.125$ (cost externality). Thus, the externality of existing patents on the cost of creating new patents is small (recall that $\varsigma_e = 0$ corresponds to no externality). The convexity of the direct cost function is stronger than linear, yet weaker than quadratic.

Worker preferences are standard: the risk aversion parameter is set to $\gamma = 2$ and the discount rate is 5% per year. The parameters of the distribution of labor market characteristics is taken from Dvorkin and Monge-Naranjo (2019) with $\kappa = 15$ and $\lambda^j = \Gamma\left(\frac{\kappa-1}{\kappa}\right)^{-1} \approx 0.96$. Initial human capital is normalized to 1.

Human capital accumulates on the job (i.e. within the same sector) with 2.5% per year, and switching sectors results in a 15% loss of human capital. There is no loss or gain of human capital from the initial occupation choice ($\tau_0^j = 1\ \forall j$) and there are no non-pecuniary switching cost ($\chi_s^{jl} = 1\ \forall s, j, l$). Workers save 10% of their labor income every period.

Depreciation rates of machines and structures are set to 5% and 2%, respectively, which is standard. The depreciation rate of patents is 5%, corresponding to a patent duration of 20 years, which is the standard duration of a patent under European Patent Law. The complete parametrization is captured in Table 2.3.

2.5.2 Mechanism

This subsection illustrates the core mechanism of the model in more detail. For this, I vary the population growth rate n from -2% to +2% and solve for the balanced growth equilibrium of the corresponding economy. Note that this is a comparison of the steady states of different economies.

First, recall that the population growth rate directly affects the shares of the population that are of working age and of retirement age: if following cohorts are larger (smaller) than previous cohorts, there are relatively more (less) younger people in the population. This relationship is illustrated by the old-age dependency ratio as depicted in Figure 2.6a. The differences in the age structure of the population translates into differences in the relative availability of input factors: labor is only provided by the working-age population and older individuals tend to hold more capital than younger individuals (as they have saved longer). The relative availability of input factors is then reflected in prices. Figures 2.6b and 2.6c depict the relationship between population growth rate and factor prices in the model. With increasing population growth (i.e. decreasing OADR), the share of retirees in the economy is decreasing and real interest rates are increasing, reflecting the reduction in the relative supply of capital. The opposite holds true for labor: higher population growth yields relatively more workers and wages are decreasing with population growth. Note that the decrease in real wages is slightly stronger in sector B ('services'), leading to a slight increase in the relative wage of sector A ('production') over sector B when population growth is higher.

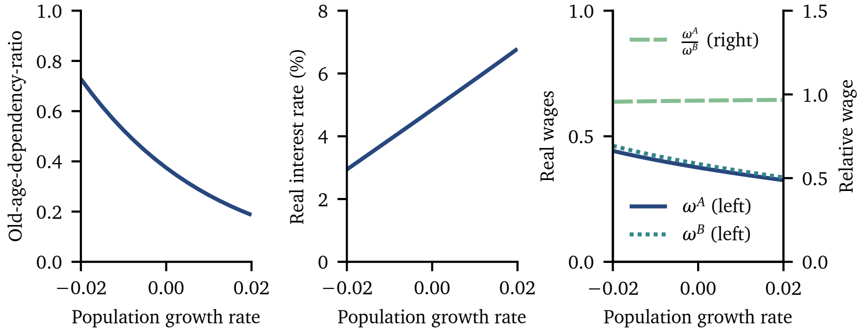
These changes in the prices of input factors affect firms' decision regarding the optimal mix of input factors in producing intermediate goods, which in turn affects the monopolists' decision to provide technology. The first aspect is reflected in changes of the factor shares, as depicted in Figure 2.7. With lower population growth, sector A ('production') employs relatively more physical capital and relatively less human capital. The higher the population growth rate, the more evenly input factors are split (across factors and across sectors). Lower population growth thus causes more specialized utilization of input factors: production firms

Table 2.3. Model parametrization

Parameter	Definition	Value: Sector A	Value: Sector B
General			
J	Number of sectors	2	
S^w	Number of periods working	45	
S^r	Number of periods retired	35	
Demographics			
n	Population growth rate	0.0	
ζ_s	Survival probability at age s	2000 observed	
Production			
α	Weight of intermediate input in intermediate good production	0.93	
ν	Weight of bundles in intermediate input production	0.75	
η	Weight of human capital in intermediate bundle production	0.65	0.65
ψ	Elasticity of substitution between intermediate goods	0.2	
σ_j	Elasticity of substitution between labor and machines	0.8	0.2
φ^j	Intermediate good production scale parameter	1.0	1.0
ϱ^j	Sector weight in final good production	0.5	0.5
ξ^j	Cost of producing licenses	0.05	0.05
ι	Scale parameter of cost of producing patents	0.005	
ς_d	Exponent of direct cost of producing patents	1.2	
ς_e	Exponent of cost externality in producing patents	0.125	
Workers			
β	Discount factor	0.95	
γ	Risk aversion parameter	2.0	
h_0	Initial human capital	1.0	
κ	Curvature parameter of labor market opportunity distribution	15	
λ^j	Scale parameter of labor market opportunity distribution	0.96	0.96
τ_0^j	Initial human capital transfer factor	1.0	1.0
$\tau_s^{j,j}$	Human capital transfer factor - same sector	1.025	1.025
$\tau_s^{j,l}$	Human capital transfer factor - other sector	0.85	0.85
χ_0^j	Initial non-pecuniar transfer cost	1.0	1.0
$\chi_s^{j,j}$	Non-pecuniary transfer cost - same sector	1.0	1.0
$\chi_s^{j,l}$	Non-pecuniary transfer cost - other sector	1.0	1.0
τ^p	Savings rate	0.1	
Investment			
δ^M	Depreciation rate of machines	0.05	
δ^K	Depreciation rate of structures	0.02	
δ^θ	Depreciation rate of technology	0.05	

Notes: Baseline parametrization; unless stated otherwise, these parameters are used throughout the analysis.

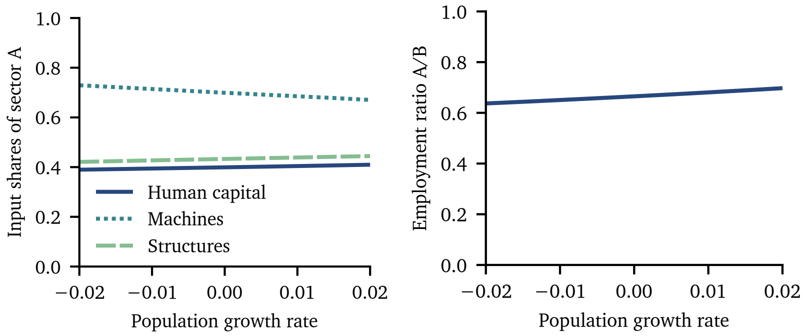
Figure 2.6. OADR and factor prices across population growth rates



(a) OADR **(b)** Real interest rate **(c)** Real wages

Notes: Old-age dependency ratio (left panel), equilibrium real interest rate (middle panel) and real wages by sector (right panel); OADR obtained as $\sum_{s=S^w+1}^{S^w-S^r} \tilde{\phi}_s / \sum_{s=1}^{S^w} \tilde{\phi}_s$ and factor prices obtained by solving for the balanced-growth equilibrium of an economy with the respective population growth rate; all remaining parameters as summarized in Table 2.3.
Source: Own computation.

Figure 2.7. Factor shares across population growth rates



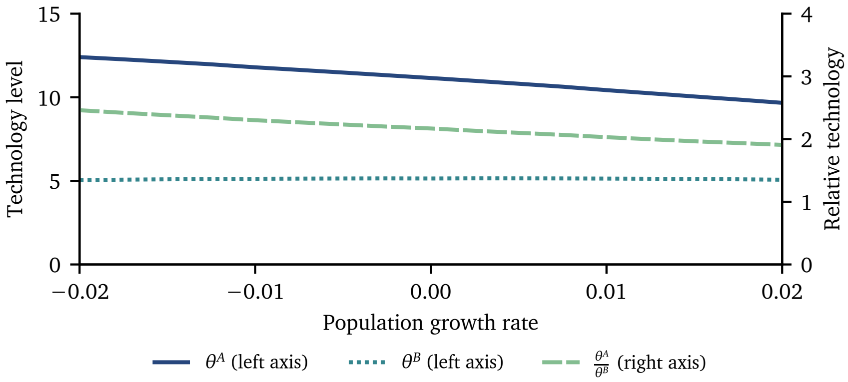
(a) Input shares in production **(b)** Relative employment

Notes: Share of total input factors used in production in sector A (left panel) and ratio of workers attached to sector A over sector B (right panel); factor shares computed as $\tilde{X}^A/(\tilde{X}^A + \tilde{X}^B)$, worker share computed as $\tilde{\phi}^A/\tilde{\phi}^B$ in the balanced-growth equilibrium of the economy with the respective population growth rate; all remaining parameters as summarized in Table 2.3.
Source: Own computation.

more heavily relies on physical capital, whereas services firms use relatively more human capital in the production process.

This specialization on particular input factors is reflected in the use of machine-enhancing technology, as captured by Figure 2.8. The key observations are: Technology levels are decreasing with population growth (relative labor supply vs. capital is increasing with population growth, hence less need for technology). Technology levels are higher in sectors where human capital and machines are more substitutable (it is easier to replace humans with machines there, therefore in-

Figure 2.8. Sectoral technology levels across population growth



Notes: Absolute and relative technology levels by sector; technology levels obtained by solving for the balanced-growth equilibrium of an economy with the respective population growth rate; all remaining parameters as summarized in Table 2.3.
Source: Own computation.

vesting in better machines makes more sense there). The difference in technology levels is decreasing with population growth. In other words, decreasing population growth and the specialization on physical vs. human capital leads to a polarization of technology levels. In this particular parametrization, the technology level of sector B is even decreasing with lower population growth. As the population growth rate decreases, sector A, where machines and human capital are more substitutable, becomes "very high-tech", whereas sector B, which is more complementary, becomes "low tech". Note that, while the more substitutable sector experiences a stronger increase in technology levels upon a reduction in population growth irrespective of a particular parametrization of the model, the direction of the change in technology levels in the more complementary model depends on model parameters.[10]

2.5.3 Simulation

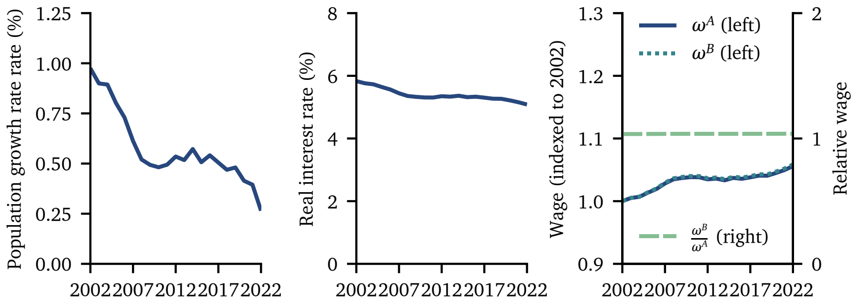
In the main analysis, I simulate the impact of the demographic transition on technology by sector. For this, I first construct an exogenous demographic process that matches the key features of the data. I take observed survival frequencies by year and back out the population growth rate by matching the observed old-age dependency ratio for that year. This way, the resulting economy features the same ratio of retirees to active workers as the data. Figure 2.9a depicts the resulting sequence of population growth rates. I then solve for the stationary equilibrium

10. This relates to the result in Acemoglu (2010) that, upon a labor supply reduction, technology can increase (when technology is *strongly labor saving*) or decrease (when technology is *strongly labor complementary*).

for each year. Note that this represents a sequence of steady states of different economies, thus abstracting from transitional dynamics.

As the mechanism operates via relative prices, I first simulate real prices for physical and human capital. Figures 2.9b and 2.9c depict real interest rate and real wages. The real interest rate falls by ca. 0.75 percentage points and real

Figure 2.9. Simulated population growth rate and factor prices



(a) Pop. growth **(b)** Real interest rate **(c)** Real wages

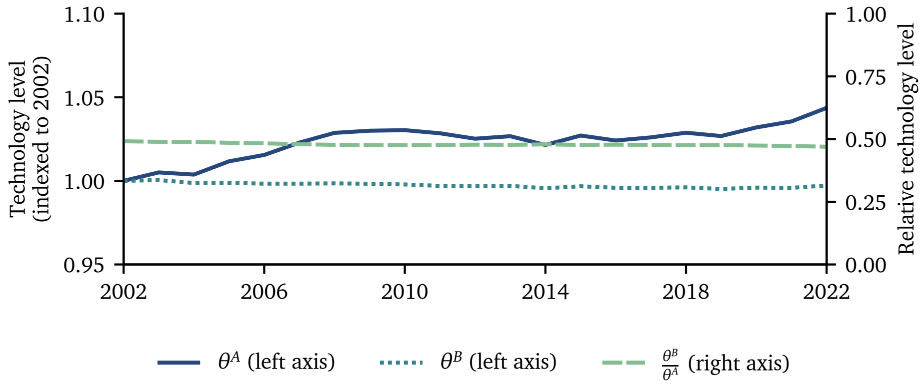
Notes: Simulated population growth rate (left panel), equilibrium real interest rate (middle panel, indexed to 2002) and real wages by sector (right panel, indexed to 2002). Population growth rate fitted to match observed OADR (given observed survival probabilities); all remaining parameters as summarized in Table 2.3. Factor prices obtained by solving for the balanced-growth equilibrium of an economy with the respective population growth rate.
Source: Own computation.

wages increase by ca. 6%, with the increase being slightly larger for services than for production. Note that these patterns are qualitatively in line with the empirical findings presented in Section 2.3, though do not capture the full extent of observed changes. The results are, however, remarkably similar to the effects of demographic change on factor prices computed by Krueger and Ludwig (2007) for the US economy.

Next, I turn to endogenous technology levels. Recall that, as technology is machine-augmenting in this setup, any increase in the technology level can be interpreted as an increase in the productivity of machines. Figure 2.10 depicts the simulated absolute and relative technology levels by sector. The level of technology in the model increases by ca. 5% for production between 2002 and 2022 and remains flat for services. This finding again relates to the results of Acemoglu (2010), who finds that whether labor scarcity encourages or discourages innovation critically depends on whether technology is labor saving or labor complementary. Here, the substitutability of machines and labor in production is sufficiently high for labor scarcity to be driving innovation, whereas this is not the case for services.

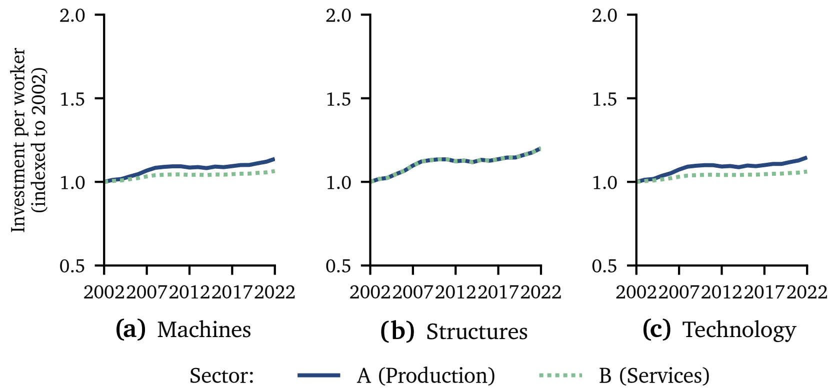
As technology levels are not directly observable, there is no empirical counterpart. Therefore, I analyze the development of investment in technology vis-à-vis physical capital. I compare gross investment per worker with the empirical results

Figure 2.10. Simulated technology levels by sector



Notes: Simulated technology levels by sector (indexed to 2002) and relative. Population growth rate fitted to match observed OADR (given observed survival probabilities); all remaining parameters as summarized in Table 2.3. Technology levels obtained by solving for the balanced-growth equilibrium of an economy with the respective population growth rate.
Source: Own computation.

Figure 2.11. Simulated investment per worker by type and sector



Notes: Gross investment per worker in machines (left panel), structures (middle panel), and patents (right panel), indexed to 2002. Population growth rate fitted to match observed OADR (given observed survival probabilities); all remaining parameters as summarized in Table 2.3. Gross investment per worker is computed according to (2.48).
Source: Own computation.

presented in Section 2.3, where gross investment per worker is computed as

$$\frac{\tilde{I}^{M,j}}{\tilde{\phi}^j} = \frac{\delta^M \tilde{M}^j}{\tilde{\phi}^j} \qquad \frac{\tilde{I}^{K,j}}{\tilde{\phi}^j} = \frac{\delta^K \tilde{K}^j}{\tilde{\phi}^j} \qquad \frac{\tilde{I}^{\theta,j}}{\tilde{\phi}^j} = \frac{A(\theta^{j*})C(\delta^\theta \theta^{j*})}{\tilde{\phi}^j} \tag{2.48}$$

with $\tilde{\phi}^j = \sum_{s=1}^{S^w} \tilde{\phi}_s^j$. The simulated development of investment per worker is depicted in Figure 2.11. The results again qualitatively match the empirical observations presented in Section 2.3: Investment in machines is moderately increasing for the production sector, while it remains relatively flat for services; investment in structures is increasing symmetrically in both sectors; investment in technology is again increasing for production and flat for services, where the increase in production is stronger than for machines. Quantitatively, the model falls short of the full effect observed in the data.

Thus, this simple setup is capable of replicating the key features of the empirical data qualitatively, though can only account for part of the observed changes.

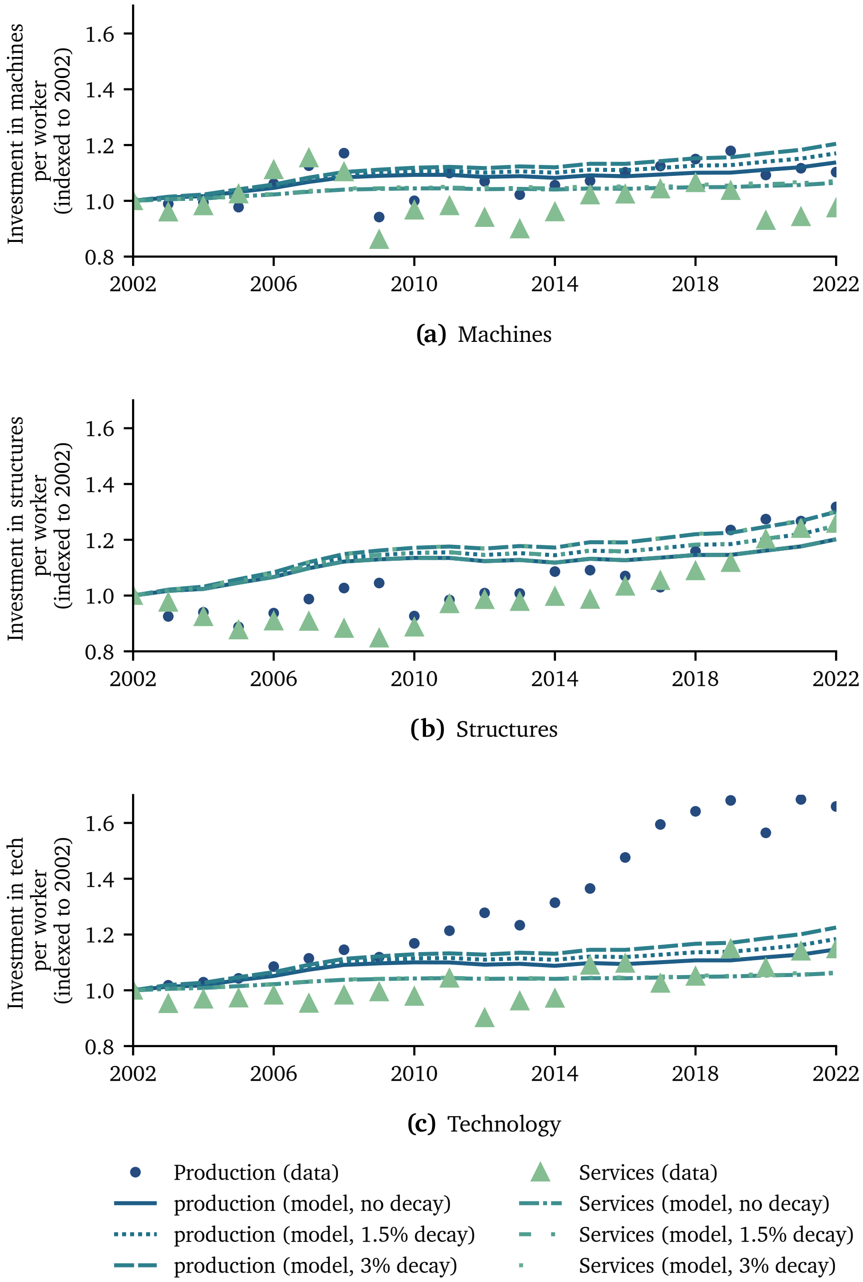
Extension: Cost decay. The analysis so far assumes a constant cost structure: any changes in the marginal cost of producing additional patents stems from the change in technology level and the marginal cost of creating licenses to use these patents in production is constant over time. The evidence suggests to the contrary that the cost of technology has been decreasing substantially over the past decades and exceedingly so. Dvorkin and Monge-Naranjo (2019) estimate an annual decline in the cost of equipment of 1.5%–3.0%.

In this extension, I incorporate decreasing cost of technology in the analysis. I model the falling cost of technology as a decrease in the marginal cost of creating licenses, ξ^j. The cost of creating licenses from patents can be thought of as a publishing or distribution cost. It thus captures e.g. improvements in connectivity of devices and established standards for the distribution of software applications, but not productivity growth in the creation of software.

In addition to the baseline case without cost decay, I assess two scenarios: 1.5% and 3.0% of annual reduction in licensing cost. All scenarios are normalized such that the licensing cost in 2002 is equal to the baseline cost. I repeat the simulation using the same demographic process for all scenarios. Figure 2.12 depicts the results. Investment per worker generally increases across industry domains and across investment categories when licensing cost are decreasing. Investment in machines and in structures are in line with the data for all scenarios. Cost decay contributes to closing the gap between model and data for investment in technology, yet effects in the model still fall significantly short of observed changes. Interestingly, decreases in licensing cost appear to have negligible effects on investment in machines and in technology for services firms.

Finally, the model not only helps explaining observed changes in investment patterns, but also allows for simulating the economy going forward. I extend the analysis by simulating the model not only on observed demographic changes until 2022, but also on projected further changes until 2050. Figure 2.13 depicts the resulting path of technology levels by sector for all three scenarios. In the baseline simulation, the technology level in production is increasing until ca. 2040 at approximately 0.2% p.a. and slightly decreasing again thereafter. Coupled with an annual reduction of technology licensing cost of 3.0%, the technology level (and, thus, productivity of machines) in production is increasing by ca. 0.4% p.a. un-

Figure 2.12. Simulated and observed investment per worker by type and sector



(a) Machines

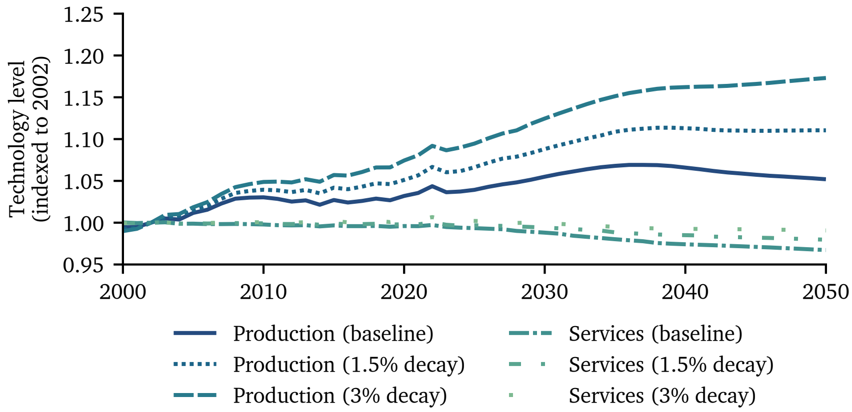
(b) Structures

(c) Technology

Notes: Gross investment per worker in machines (left panel), structures (middle panel), and patents (right panel), indexed to 2002. Survival probabilities are observed frequencies in respective year, population growth rates are adjusted to match observed OADR, marginal cost decreasing by 3% p.a. (2002=baseline); all remaining parameters as summarized in Table 2.3. Gross investment per worker is computed according to (2.48).

Source: VGR (Statistisches Bundesamt (Destatis), Genesis Online, 2025a,b), Own computation.

Figure 2.13. Simulated technology levels by sector under cost decay scenarios



Notes: Simulated technology levels by sector (indexed to 2002) for baseline scenario and with 1.5% and 3.0% annual decay in technology licensing cost. Population growth rate fitted to match observed/ projected OADR (given observed/ projected survival probabilities); all remaining parameters as summarized in Table 2.3. Technology levels obtained by solving for the balanced-growth equilibrium of an economy with the respective population growth rate and marginal cost of creating technology licenses.
Source: Own computation.

til 2040 and continues to increase thereafter. The technology level in services is marginally decreasing over time, even with decaying licensing cost.

The simulation indicates that the proposed mechanism is quantitatively meaningful, i.e. the contribution of the endogenous technology response to changes in the demographic composition of the population is a relevant factor in shaping industry composition.

2.6 Conclusion

Demographic change is drastically altering the population composition of most developed economies. Germany's transition has been underway for three decades and will likely continue for at least three more, making it an ideal laboratory for understanding how demographic forces reshape economic structures.

The presented analysis provides evidence for three fundamental patterns. First, demographic forces reached a tipping point around 2010, when a structural break occurred and demographic pressure began translating into measurable labor scarcity. Second, these pressures affect all industries uniformly—the scarcity is economy-wide rather than sector-specific. Third, firms' responses to these pressures depend critically on their production technology: sectors with high capital-labor substitutability invest heavily in labor-saving technologies, while sectors with limited substitution possibilities engage primarily in wage competition.

To understand these patterns, I develop a tractable multi-sector model featuring population dynamics, worker life cycle sector choices, machine-augmenting technology, and monopolistic innovation. The model's core mechanism demonstrates that aging simultaneously creates capital abundance (consistent with Krueger and Ludwig, 2007) and labor scarcity, triggering technological progress in some but not all sectors (in line with Acemoglu, 2010). The framework successfully replicates observed patterns of investment and wage growth across sectors with different substitutability characteristics.

Quantitatively, simulations based on observed and projected demographic changes yield approximately 0.2% annual productivity growth from demographic forces alone between 2000 and 2050. While this may appear modest, it represents a persistent and accelerating force that compounds over time, suggesting that demographic-driven technological change will become increasingly important as aging intensifies.

The patterns documented here extend well beyond Germany's borders. As developed economies worldwide face similar demographic transitions, understanding how sectoral heterogeneity mediates aggregate demographic effects becomes crucial for both firms and policymakers. Countries experiencing rapid aging should anticipate differential sectoral responses based on their industrial composition and technological characteristics, with production-intensive economies likely seeing greater technology adoption and service-intensive economies facing more pronounced wage pressures.

These findings also highlight the importance of considering demographic forces when analyzing long-term economic growth and technological change. As populations age globally, the interaction between demographic transitions and sectoral production technologies will increasingly shape the direction of innovation and the distribution of economic gains across industries and workers.

References

Acemoglu, Daron. 2010. "When Does Labor Scarcity Encourage Innovation?" *Journal of Political Economy* 118 (6): 1037–78. [58–60, 80, 99, 100, 105]

Acemoglu, Daron, and Pascual Restrepo. 2021. "Demographics and Automation." *Review of Economic Studies* 89 (1): 1–44. [59]

Arbeitskreis "Volkswirtschaftliche Gesamtrechnung der Länder". 2025a. *Arbeitnehmerentgelt, Bruttolöhne und -gehälter in den Ländern der Bundesrepublik Deutschland 1991 bis 2024 - Reihe 1, Ländergebnisse Band 2 (Revision 2024)*. German. Volkswirtschaftliche Gesamtrechnung der Länder. Stuttgart. URL: https://www.vgrdl.de/ (visited on 06/16/2025). [65, 71, 108, 119, 120]

Arbeitskreis "Volkswirtschaftliche Gesamtrechnung der Länder". 2025b. *Bruttoanlageinvestitionen in den Ländern der Bundesrepublik Deutschland 1991 bis 2022 - Reihe 1, Ländergebnisse Band 3 (Revision 2024)*. German. Volkswirtschaftliche Gesamtrechnung der Länder. Stuttgart. URL: https://www.vgrdl.de/ (visited on 06/16/2025). [71, 108, 119, 120]

Autor, David H, and David Dorn. 2013. "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market." *American Economic Review* 103 (5): 1553–97. [59]

Börsch-Supan, Axel, Alexander Ludwig, and Joachim Winter. 2006. "Ageing, Pension Reform and Capital Flows: A Multi-Country Simulation Model." *Economica* 73 (292): 625–58. [59]

Dauth, Wolfgang, Sebastian Findeisen, Jens Südekum, and Nicole Woessner. 2017. "German Robots-The Impact of Industrial Robots on Workers." *CEP Discussion Paper*, (1500): [59]

De Nardi, Mariacristina, Selahattin İmrohoroğlu, and Thomas J. Sargent. 1999. "Projected U.S. Demographics and Social Security." *Review of Economic Dynamics* 2 (3): 575–615. [59]

Deutsche Bundesbank. 2025. *Effective interest rates of German banks – New business – Households' deposits, overnight (Series SUD101 / SUR101)*. Deutsche Bundesbank. URL: https://www.bundesbank.de/en/statistics/money-and-capital-markets/interest-rates-and-yields/mfi-interest-rate-statistics-amounts-outstanding-new-business--651536 (visited on 03/27/2025). [65, 110]

Dvorkin, Maximiliano, and Alexander Monge-Naranjo. 2019. "Occupation Mobility, Human Capital and the Aggregate Consequences of Task-Biased Innovations." Working Papers 2019-064. Human Capital, and Economic Opportunity Working Group. [58, 59, 73, 75, 95, 102, 122, 128]

Frodermann, Corinna, Andreas Ganzer, Alexandra Schmucker, and Philipp Vom Berge. 2021a. *Factually anonymous Version of the Sample of Integrated Labour Market Biographies (SIAB-Regionalfile) – Version 7519 v1*. [60, 63, 65, 68, 71, 108, 114–117, 119, 120]

Frodermann, Corinna, Andreas Ganzer, Alexandra Schmucker, and Philipp Vom Berge. 2021b. "Sample of Integrated Labour Market Biographies Regional File (SIAB-R) 1975-2019." FDZ-Datenreport 05/2021 (en). Nürnberg. [112]

Gartner, Hermann. 2005. "The imputation of wages above the contribution limit with the German IAB employment sample." FDZ-Methodenreport 02/2005 (en). Nürnberg. [109]

Graetz, Georg, and Guy Michaels. 2018. "Robots at Work." *Review of Economics and Statistics* 100 (5): 753–68. [59]

Krueger, Dirk, and Alexander Ludwig. 2007. "On the consequences of demographic change for rates of returns to capital, and the distribution of wealth and welfare." *Journal of Monetary Economics* 54 (1): 49–87. [59, 100, 105]

Statistik der Bundesagentur für Arbeit. 2013–2024. *Regionalreport über Beschäftigte (Quartalszahlen)*. Nürnberg. URL: https://statistik.arbeitsagentur.de/SiteGlobals/Forms/Suche/Einzelheftsuche_Formular.html?topic_f=beschaeftigung-reg-bst-reg (visited on 03/27/2025). [109, 110]

Statistik der Bundesagentur für Arbeit. 2014–2024. *Beschäftigte nach Wirtschaftszweigen (WZ 2008) (Quartalszahlen)*. Nürnberg. URL: https://statistik.arbeitsagentur.de/SiteGlobals/Forms/Suche/Einzelheftsuche_Formular.html?topic_f=beschaeftigung-sozbe-wz-heft (visited on 03/27/2025). [63, 109]

Statistik der Bundesagentur für Arbeit. 2016–2023. *Gemeldete Arbeitsstellen (Monatszahlen)*. Nürnberg. URL: https://statistik.arbeitsagentur.de/SiteGlobals/Forms/Suche/Einzelheftsuche_Formular.html?topic_f=gemeldete-arbeitsstellen (visited on 03/27/2025). [109, 110]

Statistik der Bundesagentur für Arbeit. 2020. "Grundlagen: Methodenbericht – Engpassanalyse – Methodische Weiterentwicklung." Technical Report. Nürnberg. [110]

Statistik der Bundesagentur für Arbeit. 2024a. *Arbeitslose – Kreise und Gemeinden (Monats- und Jahreszahlen)*. Nürnberg. URL: https://statistik.arbeitsagentur.de/Statistikdaten/Detail/Aktuell/iiia4/arbeitslose-kgd-ab1998/arbeitslose-kgd-ab1998-d-0-xlsx.xlsx?__blob=publicationFile&v=1 (visited on 03/27/2025). [109, 110]

Statistik der Bundesagentur für Arbeit. 2024b. *Arbeitsmarktstatistiken zur Bestimmung des Enpassindikators*. Custom dataset. Nürnberg. Order no. 354576. [110, 112, 114]

Statistik der Bundesagentur für Arbeit. 2024c. "Berichte: Blickpunkt Arbeitsmarkt – Fachkräfteengpassanalyse 2023." Technical Report. Nürnberg. [111]

Statistik der Bundesagentur für Arbeit. 2024d. *Beschäftigte nach ausgewählten Merkmalen (Zeitreihe Quartalszahlen)*. Nürnberg. URL: https://statistik.arbeitsagentur.de/Statistikdaten/Detail/Aktuell/iiia6/beschaeftigung-sozbe-zr-ausgewmerkmale-altersgr/zr-ausgewmerkmale-altersgr-d-0-xlsx.xlsx?__blob=publicationFile&v=1 (visited on 03/27/2025). [63, 109]

Statistisches Bundesamt (Destatis), Genesis Online. 2025a. *National accounts - Gross fixed capital formation (nominal/price-adjusted): Germany, years, industries, types of fixed assets (Table 81000-0129)*. URL: https://www-genesis.destatis.de/datenbank/online/statistic/81000/table/81000-0129 (visited on 06/20/2025). [66, 103, 108]

Statistisches Bundesamt (Destatis), Genesis Online. 2025b. *National accounts - Persons in employment: Germany, years, industries (Table 81000-0123)*. URL: https://www-genesis.destatis.de/datenbank/online/statistic/81000/table/81000-0123 (visited on 06/20/2025). [66, 103, 108]

Stokey, Nancy L., Robert E. Lucas, and Edward C. Prescott. 1989. *Recursive Methods in Economic Dynamics*. Harvard University Press. [142, 143]

United Nations, Department of Economic and Social Affairs, Population Division. 2022. "World Population Prospects 2022." Online Edition. [62, 108]

Appendix 2.A Empirical Analysis

This Appendix provides additional details on data sources, methodology, and classifications used for the empirical analysis, as well as additional empirical results.

2.A.1 Data Sources

The World Population Prospects dataset (WPP, United Nations, Department of Economic and Social Affairs, Population Division, 2022) is the United Nations Population Division's official collection of demographic estimates and projections, providing comprehensive population data for all countries and regions from 1950 to 2100. The dataset includes annual population figures by age and sex, fertility rates, mortality indicators, and migration data based on national statistics. The projections incorporate multiple scenarios (medium, high, low, and constant fertility variants) based on probabilistic forecasting models that account for uncertainty in future demographic trends. I use data on population size by age (observed and projected), observed survival frequencies, and projected survival probabilities for Germany from the 2022 issue of the WPP. For all projections, the medium scenario is used.

Data from the German National Accounts (VGR) is provided online through the Genesis Online database by the German Statistical Office (Destatis). Data from two tables hase been used in the analysis: Table 81000-0123 (Gross fixed capital formation (nominal/price-adjusted): Germany, years, industries, types of fixed assets, Statistisches Bundesamt (Destatis), Genesis Online, 2025a) and Table 81000-0129 (Persons in employment: Germany, years, industries, Statistisches Bundesamt (Destatis), Genesis Online, 2025b).

The VGRdL data is provided by a workgroup consisting of the 16 statistical offices of the German federal states and the German Statistical Office. Two distinct publications by the workgroup have been used for analysis: "Arbeitnehmerentgelt, Bruttolöhne und -gehälter in den Ländern der Bundesrepublik Deutschland" (Series 1, Volume 2; Arbeitskreis "Volkswirtschaftliche Gesamtrechnung der Länder", 2025a) and "Bruttoanlageinvestitionen in den Ländern der Bundesrepublik Deutschland" (Series 1, Volume 3; Arbeitskreis "Volkswirtschaftliche Gesamtrechnung der Länder", 2025b).

The main data source for the computation of scarcity statistics is the German Sample of Integrated Labour Market Biographies (SIAB) from the BA. Data access was provided via a Scientific Use File (SUF) supplied by the Research Data Centre (FDZ) of the BA at the Institute for Employment Research (IAB). The SUF is referred to as the SIAB Regional File (or SIAB-R), the version used here is Version 7519 v1 (Frodermann et al., 2021a). The SIAB is a 2% random sample of the labor market biographies of all individuals in Germany that are employed, unemployed or otherwise attached to the labor force. Two key features of the data are exploited: The structure of the dataset allows for observing an individuals' employment status over time, i.e. labor market stocks and flows can be observed; the sampling contains a 2% random draw from all sources, i.e. population averages for all status groups can be estimated. Earnings data in the SIAB is topcoded, earnings above the upper bound are imputed using the methodology of Gartner (2005). Occupations in the SIAB are coded according to the Classification of Occupations (KldB) of the German Statistical Office. Data is coded according to KldB issue 1988 (until 2011) and issue 2010 (from 2012 onward). Geographic information is categorized by German federal states and by labor market regions. Labor market regions are derived from the 40 labor market units as defined by the IAB, where some regions have been combined for compatibility with the SIAB-R.[11]

In addition, various datasets provided by the BA are used throughout the analysis. Information on employment is extracted from three series of publications: "Beschäftigte nach ausgewählten Merkmalen (Zeitreihe Quartalszahlen)"(Employed Persons by Selected Characteristics (Time Series Quarterly Data); Statistik der Bundesagentur für Arbeit, 2024d), "Beschäftigte nach Wirtschaftszweigen (WZ 2008) (Quartalszahlen)" (Employed Persons by Economic Sectors (WZ 2008) (Quarterly Data); Statistik der Bundesagentur für Arbeit, 2014–2024), and "Regionalreport über Beschäftigte (Quartalszahlen)" (Regional Report on Employed Persons (Quarterly Data); Statistik der Bundesagentur für Arbeit, 2013–2024). The first dataset contains information on total employment by age group; the second dataset (compiled from individual publications from the series between March 2014 and September 2024) contains information on total employment by age group and economic sector (classified according to WZ 2008); the third dataset (compiled from individual publications between December 2013 and December 2024) contains information on total employment by age and NUTS-1/ NUTS-3 region. Information on unemployment is taken from the publication "Arbeitslose – Kreise und Gemeinden (Monats- und Jahreszahlen)" (Unemployed Persons-Counties and Municipalities (Monthly and Yearly Data); Statistik der Bundesagentur für Arbeit, 2024a), which contains information on total unemployment by NUTS-4 region. Finally, information on vacancies is compiled from individual publications of the series "Gemeldete Arbeitsstellen (Monatszahlen)" (Reported Job Vacancies (Monthly Figures); Statistik der Bundesagentur für Arbeit, 2016–2023) between December 2016 and December 2023. The publications contain information on vacancies by NUTS-1/ NUTS-3 region and occupation group (KldB 2010 2-digit codes). All BA data captures the entirety of employed/ unemployed/ vacancies that have been registered with the BA.

For the robustness exercises, a bespoke dataset of labor market statistics by region (IAB labor market region) and occupation main group (KldB 2010 2-digit

11. See Appendix 2.A.3 for further details on classifications.

codes) has been provided by the BA (Statistik der Bundesagentur für Arbeit, 2024b). The data has been used to compute scarcity sub-indicators following the methodology outlined in Appendix 2.A.2. BA scarcity data by IAB labor market region has been aggregated to match SIAB-R regional classification using weighted averages: employment-weighted averages for *change in the share of foreign workers* and *wage growth* based on employment, vacancy-weighted averages for *vacancy duration* and *searcher-vacancy-ratio*, and unemployment-weighted averages for *exit rate from unemployment*. Aggregation weights for BA scarcity statistics are obtained from Statistik der Bundesagentur für Arbeit (2013–2024, 2016–2023, 2024a).

Finally, real interest rate data is provided by the BB via its online database. The data series used is "Real interest rates of German banks / New business / Households' deposits, overnight / SUR101" (Deutsche Bundesbank, 2025). Monthly real interest rates are aggregated (simple time-averaging) to obtain yearly values.

2.A.2 Methodology

The construction of scarcity indices builds on the methodology employed by the BA in their analysis of labor scarcity by occupation: First, five sub-indicators relating to different aspects of relative labor shortages are computed for each observation cell. The indicators are *vacancy duration*, *searcher-vacancy-ratio*, *change in the share of foreign workers*, *exit rate from unemployment*, and *wage growth*.[12] *Searcher-vacancy-ratio* is defined as the ratio of the number of searchers in a given observation unit to the number of vacancies in that same unit and measures labor market tightness. *Change in the share of foreign workers* is defined as the difference between the share of foreign workers in the current year and the share of foreign workers three years earlier and measures the relative inflow of human capital from outside the economy. *Exit rate from unemployment* is defined as the ratio of the number of transitions from unemployment in an observation unit to employment over the total number of unemployed workers within that unit and captures how quickly idle workers are matched to vacant positions. *Wage growth* is defined as the percentage change of the median wage in the current year vs. three years earlier and captures wage competition within a given observation unit. Further details on the definition, interpretation, and implication on overall labor scarcity of the statistics can be found in the BA methodology report (Statistik der Bundesagentur für Arbeit, 2020).

Individual scarcity statistics are transformed into indicators using thresholds. For each statistic, a value of zero is assigned, if the statistic is below the low threshold, a value of one is assigned, if the statistic is between the low and medium

12. The original methodology included a sixth sub-indicator (*occupation-specific unemployment rate*), which is not included here due to data limitations.

Table 2.A.1. Thresholds for scarcity indicators

	Threshold		
Statistic	Low	Medium	High
Exit rate from unemployment	5.0	10.0	14.1
Change in foreigner share	0.0	2.0	4.0
Unit-specific unemployment rate (in %)	5.0	4.0	3.0
Searcher-vacancy-ratio	4.4	3.4	2.4
Vacancy duration (in days)	40.0	60.0	80.0
Wage growth (in %)	3.3	6.8	10.2

Notes: Thresholds for the computation of scarcity sub-indicators from labor market statistics.
Source: Statistik der Bundesagentur für Arbeit (2024c), own computation.

thresholds, and a value of two is assigned, if the statistic is between the medium and high thresholds. A value of three is assigned if the statistic is above the high threshold. Note that, according to the methodology of the BA, thresholds for a subset of statistics (*vacancy time* and *searcher-vacancy-ratio*) are fixed, while the thresholds for the remaining statistics (*change in the share of foreign workers*, *exit rate from unemployment*, and *wage growth*) are variable. Variable thresholds are, in principle, recomputed every year based on the data and represent the mean (medium threshold) and the mean plus/ minus one standard deviation (for high and low threshold, respectively). Recomputing thresholds reduces comparability across years. As this exercise aims at identifying changes in investment behavior resulting from changes in labor scarcity over time, I apply the thresholds used in the latest BA report (Statistik der Bundesagentur für Arbeit, 2024c) for all years. Moreover, as the data used for the analysis does not feature skill levels, I aggregate thresholds over levels. The thresholds (rounded to one decimal place) are shown in Table 2.A.1. Finally, the five indicators are combined into a single scarcity index by simple averaging. The index is only computed for cells with information on at least three sub-indicators.

2.A.3 Classifications

For the geographic segmentation of the data, two distinct classifications are used: labor market regions as defined by the Institute for Employment Research (IAB) and the political structure of Germany, i.e. NUTS-1 (federal states) and NUTS-3 (counties and independent cities) regional codes. Labor market regions are defined such that cross-region exchange of workers is minimal (approx. 10% of workers commute to another labor market region). For compatibility with the SIAB-R, the 50 IAB regions are coarsened to 40 *labor market regions*, and the 400 NUTS-3 regions are aggregated to 328 units, which I refer to as *counties*. The coarsening keeps the hierarchical structure of the NUTS classification intact, i.e. the 328 *counties* can be aggregated into the 16 *federal states*. The coarsening of IAB labor market regions is summarized in Table 2.A.2. For a complete mapping of NUTS-3 regions to *counties*, see Table A9 in Frodermann, Ganzer, Schmucker, and Berge (2021b).

For occupations, the German Classification of Occupations (KldB) is employed. Data before 2011 is encoded using the 1988 issue of the KldB, data from 2012 onward is encoded using the 2010 issue. For compatibility with data from the SIAB-R, the 328 KldB 1988 3-digit codes are coarsened to 120 units, and the 144 KldB 2010 3-digit codes are coarsened to 126 units, which will be referred to as *occupation groups*.[13] To maintain the hierarchical structure of the classification, the coarsening of occupation groups requires the combination of some KldB 1988 level-2 and level-1 codes.[14] The resulting harmonized classification features for KldB 1988 5 level-1 units, 29 level-2 units, and 120 level-3 units, and for KldB 2010 10, 37, and 126 units, respectively. I refer to level-1 units as *occupation areas*, to level-2 units as *occupation main groups*, and to level-3 units as *occupation groups*.

For industries, the German Classification of Economic Activities, Edition 2008, (WZ 2008) is used. Industries are coarsened from WZ 2008 sections (letter codes) into three industry domains ('production', 'non-government services', and 'government services'). 'Production' entails WZ2 2008 sections A–F, 'non-government services' entails sections G–N, 'government services' entails sections O–U. Throughout the analysis, industry domain 'government service' is disregarded to ensure that decisions of market participants are primarily driven by market forces (as opposed to other forces, e.g. administrative considerations or political motives).

2.A.4 Robustness

Scarcity regression. To ensure robustness of the results to the choice of data used for the analysis, I repeat the analysis with an alternative data source for labor scarcity. Instead of computing scarcity statistics from the SIAB, I use a bespoke dataset from the BA (Statistik der Bundesagentur für Arbeit, 2024b). The dataset

13. Note that 3-digit-codes (*occupation groups*) refers to aggregation level 4 in KldB 1988 and level 3 in KldB 2010. For a complete mapping of KldB 3-digit-codes to *occupation groups* see Table A6 (for KldB 1988) and Table A7 (for KldB 2010) in Frodermann et al. (2021b).

14. For level 2, occupation sections IIa and IIIa, IIIk and IIIk, and IIIo and IIIp, respectively, are combined; for level 1, occupation areas II and III are combined.

Table 2.A.2. Geographic classification: labor market regions

IAB Labor market region		Labor market region (coarsened)	
02000	Hamburg	01	Hamburg
03101	Braunschweig/Wolfsburg	02	Braunschweig/Wolfsburg
05711	Bielefeld/Paderborn	03	Bielefeld/Paderborn/Göttingen/Erfurt/Leipzig
03159	Göttingen		
16051	Erfurt		
14713	Leipzig		
03241	Hannover	04	Hannover
03403	Oldenburg (O.)	05	Bremen/Oldenburg(O.)
04011	Bremen		
03404	Osnabrück	06	Osnabrück
05113	Düsseldorf-Ruhr	07	Düsseldorf-Ruhr
05315	Köln	08	Köln
05334	Aachen	09	Aachen
05515	Münster	10	Münster
05970	Siegen	11	Siegen
06412	Frankfurt a.M.	12	Frankfurt a.M./Mannheim/Saarbrücken
08222	Mannheim		
10041	Saarbrücken		
06611	Kassel	13	Kassel
07111	Koblenz	14	Koblenz
07211	Trier	15	Trier
08111	Stuttgart	16	Stuttgart
08212	Karlsruhe	17	Karlsruhe
08311	Freiburg i.Br.	18	Freiburg i.Br.
08317	Offenburg	19	Offenburg
08326	Villingen-Schwenningen	20	Villingen-Schwenningen
08335	Konstanz	21	Konstanz
08336	Lörrach	22	Lörrach
08421	Ulm	23	Ulm
09162	München	24	München/Ravensburg
08436	Ravensburg		
09262	Passau	25	Passau
09362	Regensburg	26	Regensburg
09464	Hof	27	Hof/Wunsiedel i.F./Weiden i.d.OPf.
09363	Weiden i.d.OPf.		
09479	Wunsiedel i.F.		
09462	Bayreuth	28	Bayreuth
16054	Suhl	29	Suhl/Coburg
09463	Coburg		
09564	Nürnberg	30	Nürnberg
09662	Schweinfurt	31	Schweinfurt
09663	Würzburg	32	Würzburg
11000	Berlin	33	Berlin
13003	Rostock	34	Rostock
13071	Mecklenburgische Seenplatte	35	Mecklenburgische Seenplatte
13075	Greifswald/Stralsund	36	Greifswald/Stralsund
14511	Chemnitz	37	Chemnitz
14612	Dresden	38	Dresden
15003	Magdeburg	39	Magdeburg
15085	Harz	40	Harz

Notes: Classification of labor market regions according to IAB definition (left column) and coarsened for compatibility with regional information in the SIAB dataset (right column).

Table 2.A.3. Effects of workforce aging on labor scarcity: alternative data source

	SIAB	BA
$lag_5(Share\ 5\ years\ to\ retirement)$	0.88***	1.63**
	(0.23)	(0.60)
R^2	0.46	0.40
Adj. R^2	0.45	0.39
Num. obs.	7,032	2,994

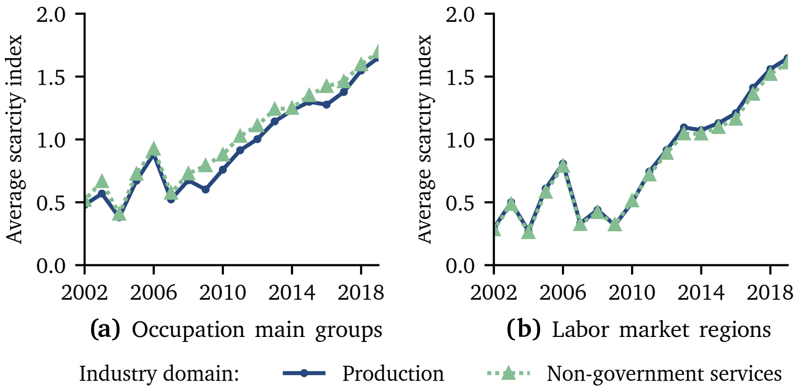
***$p < 0.001$; **$p < 0.01$; *$p < 0.05$

Notes: Coefficient estimates for β of (2.2). Observation units are labor market regions and occupation main groups. Labor market regions are coarsened from IAB labor market regions (for details see Table 2.A.2); Occupation main groups are KldB 2010 2-digit codes. Estimation based on SIAB dataset uses data from 2012–2019, estimation based on BA dataset uses data from 2015–2023.
Source: BA (Statistik der Bundesagentur für Arbeit, 2024b), SIAB (Frodermann et al., 2021a), own computation.

contains information on median income, share of foreign workers, vacancy durations, ratio of searchers to vacancies, and exit rates from unemployment by IAB labor market region and occupation main group (KldB 2010 2-digit codes). From these statistics, I construct the labor scarcity index as outlined above. The range of available data varies across statistics, allowing to compute the labor scarcity index for 2015–2023. For comparability, data by IAB labor market region is aggregated to the 40 IAB regions used in the analysis (see Table 2.A.2). Scarcity statistics by industry are then again computed using employment weights from the SIAB. The results are summarized in Table 2.A.3. The coefficient estimates are positive and statistically significant in both regressions. Further, the coefficient from the estimation based on the BA dataset (2015–2023) is higher than the coefficient estimate based on the SIAB. This indicates that the effect of aging might have increased over time.

Next, I assess the consequences of applying a different segmentation for the observation units. In the main text, observation units are combinations of labor market regions and occupation groups. Here, I compute scarcity statistics and employment weights by industry for occupation groups only and for German federal states as regions. Figure 2.A.1 depicts the resulting average scarcity statistics by industry domain. The key features of the data, i.e. stable and moderate levels until

Figure 2.A.1. Average labor scarcity by industry domain: alternative data segmentations



Notes: Average labor scarcity index by industry domain. Data by industry is computed using equation (2.1) from scarcity by occupation main group (Kldb 1988 level-2-codes/ KldB 2010-2-digit codes, left panel) and by state (right panel). For KldB 1988, some level-2-codes are combined for consistency with SIAB data (see Appendix 2.A.3). Industry domains are aggregated from WZ08 industry sections (1-letter codes); 'Production' entails sections A–F, 'non-government services' entails sections G–N.
Source: SIAB (Frodermann et al., 2021a), own computation.

2009, followed by a significant increase until 2019, and strong alignment across industries, remain unchanged. I also repeat the regression exercise described in Section 2.3 for these alternative segmentations. Table 2.A.4 summarizes the results.

The pattern discussed in the main text remains intact: the coefficient estimate for the demographic variable for the 2000–2011 period is insignificant, the estimate for the 2012–2019 period is positive and statistically significant (here even at the 0.1% confidence level). Moreover, the coefficient estimate for the segmentation based on regions only (estimated over the entire observation period) is positive and significant at the 1% confidence level. Thus, the results do not depend on the particular segmentation applied for the computation of scarcity statistics used in the aggregation.

Finally, I assess robustness with respect to the choice of the demographic variable. I repeat the estimation of equation (2.2) for three demographic indicators across two segmentations: 5-year lag of median age, 5-year lag of the share of employed workers within 5 years of statutory retirement, and the 10-year lag of the share of employed workers within 10 years of statutory retirement as indicators, and labor market regions and federal states as segmentations. Table 2.A.5 summarizes the results. The coefficient estimates for the demographic variables are positive and statistically significant for most combinations of variable and segmentation. Note that the estimation is conducted on the full dataset, i.e. including the period 2000–2011 for which coefficient estimates have been insignificant in other specifications.

Table 2.A.4. Effects of workforce aging on labor scarcity: alternative data segmentations

	Occupation group		Region
	KldB 1988 (2000-2011)	KldB 2010 (2012-2019)	Federal state (2000-2019)
$lag_5(Share\ 5\ years\ to\ retirement)$	−5.85	15.84***	15.43**
	(5.43)	(4.22)	(5.05)
R^2	0.39	0.60	0.11
Adj. R^2	0.34	0.54	0.06
Num. obs.	348	286	320

***$p < 0.001$; **$p < 0.01$; *$p < 0.05$

Notes: Coefficient estimates for β of (2.2) for different segmentations of observation units. Observation units are occupation main groups according to KldB 1988 (left), occupation main groups according to KldB 2010 (middle) and German federal states (right). For KldB 1988, some level-2-codes are combined for consistency with SIAB data (see Appendix 2.A.3). Estimation based on KldB 1988 uses data from 2000–2011, estimation based on KldB 2010 uses data from 2012–2019, estimation based on federal states uses data from 2000–2019.
Source: SIAB (Frodermann et al., 2021a), own computation.

Table 2.A.5. Effects of workforce aging on labor scarcity: alternative demographic indicators

	Labor market regions			Federal states		
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
$lag_5(Age)$	0.34^{***}			-0.00		
	(0.04)			(0.05)		
$lag_5(Share\ 5\ years\ to\ retirement)$		7.69			15.43^{**}	
		(4.39)			(5.05)	
$lag_{10}(Share\ 10\ years\ to\ retirement)$			7.20^{***}			6.59^{***}
			(1.99)			(1.84)
R^2	0.21	0.12	0.13	0.08	0.11	0.12
Adj. R^2	0.17	0.08	0.09	0.03	0.06	0.08
Num. obs.	800	800	788	320	320	315

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Notes: Coefficient estimates for β of (2.2) for different independent variables. *age* is the median age of full-time employed, *share 5 years to retirement* and *share 10 years to retirement* are the shares of full-time employed that are within 5 and 10 years to statutory retirement age, respectively. Observation units are labor market regions (models 1–3) and German federal states (models 4–6). Labor market regions are coarsened from IAB labor market regions (for details see Table 2.A.2). Estimation of all models uses data from 2000–2019.
Source: SIAB (Frodermann et al., 2021a), own computation.

Overall, the available evidence suggests that the increase in average labor scarcity observed in the data is in part driven by changes in the demographic composition of the work force. These results are robust across different data sources, different segmentations of observation units, and different measures of the demographic composition.

Investment regression. Next, I run several robustness checks on the results regarding firms' responses to labor scarcity. To ensure that the results on investment, wages and scarcity are driven by demographic factors, I employ a two-stage approach: I first compute predicted values for labor scarcity from regression (2.2) and average fitted labor scarcity indices by industry (using equation (2.3)). Then, I re-run the regression using average predicted labor scarcity. For the prediction of average scarcity, note the following: The regression setup for estimating the effects of aging on scarcity uses deviations of demographic indicators from the economy-wide average for identification. Using the estimated model (2.2) thus yields the scarcity predicted from the deviations. To obtain predictions of scarcity from levels of the demographic indicators, I need to add the scarcity predicted by the economy-wide averages of the demographic variables. Specifically, I compute the predicted average scarcity by occupation unit as[15]

$$\hat{\hat{z}}_{jkt} = \hat{\beta}(x_{jkt} - \bar{x}_t) + \hat{\gamma}_j + \hat{\sigma}_k + \hat{\beta}\bar{x}_t \tag{2.A.1}$$

and average predicted labor scarcity is computed as

$$\hat{\hat{z}}_{ijt} = \sum_k \bar{w}_{ijk}\hat{\hat{z}}_{jkt} \tag{2.A.2}$$

I then estimate the following adjusted regression equation:

$$y_{ijt} = \tilde{\beta}_1(\hat{\hat{z}}_{ijt} * Production_i) + \tilde{\beta}_2(\hat{\hat{z}}_{ijt} * Services_i) + \tilde{\gamma}_i + \tilde{\delta}_j + \tilde{\sigma}_t + \tilde{\varepsilon}_{ijt} \tag{2.A.3}$$

where y_{ijt} are again investment in machines and other, investment in structures, and wages per hour in industry i and state j in year t, and $\hat{\hat{z}}_{ijt}$ are average labor scarcity indices predicted by demographics. Table 2.A.7 summarizes the results. As can be seen, while the statistical significance of the results is lower, the qualitative results of Section 2.3 remain intact: production firms invest in machines and technology, whereas services firms increase wages.

Next, I address concerns of simultaneity bias. Investment could potentially affect local labor demand, thereby affecting measured labor scarcity. Note that while this may be of concern for investment, a similar channel seems unplausible for wages: increases in current wages should, if anything, decrease labor demand,

15. The underlying assumption is that the economy-wide demographic trend affects scarcity with the same coefficient estimated for the cross-sectional variation.

Table 2.A.6. Effects of labor scarcity on investment and wages: predicted regressors

	$\log\left(\frac{\text{Investment}}{\text{Worker}}\right)$		$\log\left(\frac{\text{Wage}}{\text{Hour}}\right)$
	Machines & other	Structures	
$\widehat{Scarcity}$ * Production	0.09*	−0.04	0.00
	(0.04)	(0.04)	(0.01)
$\widehat{Scarcity}$ * Non-govt. services	0.06	0.04	0.02*
	(0.04)	(0.04)	(0.01)
R^2	0.74	0.92	0.96
Adj. R^2	0.73	0.92	0.96
Num. obs.	640	640	640

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Notes: Coefficient estimates for $\tilde{\beta}_1$ and $\tilde{\beta}_2$ of (2.A.3). Fitted average labor scarcity by industry domain and state is computed using (2.A.1) and (2.A.2) and fitted values for labor scarcity from regression (2.2). Industry domains are aggregated from WZ08 industry sections (1-letter codes); 'Production' entails sections A–F, 'non-government services' entails sections G–N.
Source: SIAB (Frodermann et al., 2021a), VGRdL (Arbeitskreis "Volkswirtschaftliche Gesamtrechnung der Länder", 2025a,b), own computation.

Table 2.A.7. Effects of labor scarcity on investment and wages: lagged regressors

	$\log\left(\frac{\text{Investment}}{\text{Worker}}\right)$		$\log\left(\frac{\text{Wage}}{\text{Hour}}\right)$
	Machines & other	Structures	
$lag_1(Scarcity)$ * Production	0.08**	0.00	0.01
	(0.02)	(0.03)	(0.01)
$lag_1(Scarcity)$ * Non-govt. services	0.05*	0.06*	0.02**
	(0.02)	(0.03)	(0.01)
R^2	0.77	0.92	0.96
Adj. R^2	0.75	0.92	0.96
Num. obs.	608	608	608

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Notes: Coefficient estimates for β_1 and β_2 of (2.4) using 1-year-lagged average labor scarcity. Average labor scarcity by industry domain and state is computed using (2.3). Industry domains are aggregated from WZ08 industry sections (1-letter codes); 'Production' entails sections A–F, 'non-government services' entails sections G–N.
Source: SIAB (Frodermann et al., 2021a), VGRdL (Arbeitskreis "Volkswirtschaftliche Gesamtrechnung der Länder", 2025a,b), own computation.

thereby alleviating scarcity. To ensure that the results are robust against any of these concerns, I repeat the analysis with lagged scarcity measures (it is unlikely that investment in the current year affects labor scarcity in the previous year). Table 2.A.6 summarizes the results. Using lagged variables, there is weak evidence that scarcity also affects other investment categories and investment by services firms. The evidence is, however, insufficient to reject the null hypothesis of no effect at the 1% significance level for these cells. The coefficient estimates for investment in machines and other at production firms, and for hourly wages at services firms, remain significant at the 1% level. The magnitudes of coefficient estimates are consistent over all three specifications, indicating that the results are indeed robust to changes in the regression setup. In summary, the presented evidence suggests that firms respond differentially to demography-driven increases in labor scarcity depending on their industry domain: production firms invest in machines and technology and services firms offer higher wages.

Appendix 2.B Proofs and Derivations

This Appendix contains the proofs of all propositions, lemmas, and corollaries in the main body. It proceeds by the same order as Section 2.4, i.e. it starts out with the worker model, followed by the production model.

Workers. *Proof of Proposition 1.* The proof is by induction. Consider the last period of working age: By Assumption (W2), the absolute value of human capital upon retirement can be factored out of the retirement value:

$$\Omega_t(b_t) = \Omega_t(h_t^{S^w}\bar{b}_t) = \frac{(h_t^{S^w}\bar{b}_t)^{1-\gamma}}{1-\gamma}\sum_{m=1}^{S^r}\beta^m \Pi_{n=0}^{m-1}\zeta_{t+m,S^w+m} = (h_t^{S^w})^{1-\gamma}\Omega_t(\bar{b}_t)$$

Thus,

$$\begin{aligned} V_t\left(S^w, h_t, j, \varepsilon_t\right) &= \frac{\left[(1-\tau^p)\omega_t^j h_t \varepsilon_t^j\right]^{1-\gamma}}{1-\gamma} + \Omega_t(b_t) \\ &= h_t^{1-\gamma}\left(\frac{\left[(1-\tau^p)\omega_t^j \varepsilon_t^j\right]^{1-\gamma}}{1-\gamma} + \Omega_t(\bar{b}_t)\right) \\ &= h_t^{1-\gamma} v_{t,S^w}\left(j, \varepsilon_t\right) \end{aligned}$$

where the final equality defines the expression $v_{t,S^w}(j,\varepsilon_t)$.
For the induction step, consider some age $0 < s < S^w$ and assume that

$$V_{t+1,s+1}\left(h_{t+1}, l, \varepsilon_{t+1}\right) = \left(h_{t+1}\right)^{1-\gamma} v_{t+1,s+1}\left(l, \varepsilon_{t+1}\right)$$

Then,

$$\begin{aligned} V_{t,s}\left(h_t, j, \varepsilon_t\right) &= \frac{\left[(1-\tau^p)\omega_t^j h_t \varepsilon_t^j\right]^{1-\gamma}}{1-\gamma} \\ &\quad + \beta\zeta_{t,s}\max_l\left\{\chi_{t,s}^{j,l}\mathbb{E}_{\varepsilon_{t+1}}\left[V_{t+1,s+1}\left(h_t\tau_{t,s}^{j,l}\varepsilon_t^l, l, \varepsilon_{t+1}\right)\right]\right\} \\ &= h_t^{1-\gamma}\frac{\left[(1-\tau^p)\omega_t^j \varepsilon_t^j\right]^{1-\gamma}}{1-\gamma} \\ &\quad + \beta\zeta_{t,s}\max_l\left\{\chi_{t,s}^{j,l}\mathbb{E}_{\varepsilon_{t+1}}\left[\left(h_t\tau_{t,s}^{j,l}\varepsilon_t^l\right)^{1-\gamma} v_{t+1,s+1}\left(l, \varepsilon_{t+1}\right)\right]\right\} \\ &= h_t^{1-\gamma}\left(\frac{\left[(1-\tau^p)\omega_t^j \varepsilon_t^j\right]^{1-\gamma}}{1-\gamma}\right. \\ &\qquad \left. + \beta\zeta_{t,s}\max_l\left\{\chi_{t,s}^{j,l}\mathbb{E}_{\varepsilon_{t+1}}\left[(\tau_{t,s}^{j,l}\varepsilon_t^l)^{1-\gamma} v_{t+1,s+1}(l, \varepsilon_{t+1})\right]\right\}\right) \\ &= h_t^{1-\gamma} v_{t,s}(j, \varepsilon_t) \end{aligned}$$

Hence, if the condition holds for age $s+1$, it also holds for age s and, by induction, the property holds for all ages in all periods. ■

Proof of proposition 2. The proof closely follows the proof of Theorem 1 in Dvorkin and Monge-Naranjo (2019). First, I solve for the expressions for the expectations over the value function. Let $\Xi^{j,l}_{t,s} = \chi^{j,l}_{t,s}(\tau^{j,l}_{t,s})^{1-\gamma} v^l_{t+1,s+1}$. Taking expectations on both sides of equation (2.8) yields

$$\begin{aligned}
v^j_{t,s} &= \mathbb{E}_{\varepsilon_t}\left[v_{t,s}\left(j,\varepsilon^j_t\right)\right] \\
&= \mathbb{E}_{\varepsilon_t}\left[\frac{\left[(1-\tau^p)\,\omega^j_t\varepsilon^j_t\right]^{1-\gamma}}{1-\gamma} + \beta\zeta_{t,s}\max_l\left\{\chi^{j,l}_{t,s}\left(\tau^{j,l}_{t,s}\varepsilon^l_t\right)^{1-\gamma} v^l_{t+1,s+1}\right\}\right] \\
&= \mathbb{E}_{\varepsilon_t}\left[\left(\varepsilon^j_t\right)^{1-\gamma}\right]\frac{\left[(1-\tau^p)\,\omega^j_t\right]^{1-\gamma}}{1-\gamma} + \beta\zeta_{t,s}\mathbb{E}_{\varepsilon_t}\left[\max_l\left\{\chi^{j,l}_{t,s}\left(\tau^{j,l}_{t,s}\varepsilon^l_t\right)^{1-\gamma} v^l_{t+1,s+1}\right\}\right] \\
&= \Upsilon^j\frac{\left[(1-\tau^p)\,\omega^j_t\right]^{1-\gamma}}{1-\gamma} + \beta\zeta_{t,s}\mathbb{E}_{\varepsilon_t}\left[\max_l\left\{\chi^{j,l}_{t,s}\left(\tau^{j,l}_{t,s}\varepsilon^l_t\right)^{1-\gamma} v^l_{t+1,s+1}\right\}\right] \\
&= \Upsilon^j\frac{\left[(1-\tau^p)\,\omega^j_t\right]^{1-\gamma}}{1-\gamma} + \beta\zeta_{t,s}\mathbb{E}_{\varepsilon_t}\left[\max_l\left\{\Xi^{j,l}_{t,s}\left(\varepsilon^l_t\right)^{1-\gamma}\right\}\right]
\end{aligned} \tag{2.B.1}$$

Now, consider only the expected value term in the last line of (2.B.1):

$$\begin{aligned}
&\mathbb{E}_{\varepsilon_t}\left[\max_l\left\{\Xi^{j,l}_{t,s}\left(\varepsilon^l_t\right)^{1-\gamma}\right\}\right] \\
&\quad= \sum_{l=1}^{J}\int_0^\infty \Xi^{j,l}_{t,s}x^{1-\gamma}Pr\left(l \in \arg\max_k\left\{\Xi^{j,k}_{t,s}\left(\varepsilon^k_t\right)^{1-\gamma}\right\}|\varepsilon^l_t = x\right)f_{\varepsilon^l}(x)dx \\
&\quad= \sum_{l=1}^{J}\int_0^\infty \Xi^{j,l}_{t,s}x^{1-\gamma}Pr\left(\Xi^{j,l}_{t,s}x^{1-\gamma} > \Xi^{j,k}_{t,s}\left(\varepsilon^k_t\right)^{1-\gamma}\ \forall k \neq l\right)f_{\varepsilon^l}(x)dx \\
&\quad= \sum_{l=1}^{J}\int_0^\infty \Xi^{j,l}_{t,s}x^{1-\gamma}\prod_{k\neq l}Pr\left(\Xi^{j,l}_{t,s}x^{1-\gamma} > \Xi^{j,k}_{t,s}\left(\varepsilon^k_t\right)^{1-\gamma}\right)f_{\varepsilon^l}(x)dx \\
&\quad= \sum_{l=1}^{J}\int_0^\infty \Xi^{j,l}_{t,s}x^{1-\gamma}\prod_{k\neq l}Pr\left(\varepsilon^k_t < \left(\frac{\Xi^{j,l}_{t,s}}{\Xi^{j,k}_{t,s}}\right)^{\frac{1}{1-\gamma}}x\right)f_{\varepsilon^l}(x)dx \\
&\quad= \sum_{l=1}^{J}\int_0^\infty \Xi^{j,l}_{t,s}x^{1-\gamma}\prod_{k\neq l}F_{\varepsilon^k}\left(\left(\frac{\Xi^{j,l}_{t,s}}{\Xi^{j,k}_{t,s}}\right)^{\frac{1}{1-\gamma}}x\right)f_{\varepsilon^l}(x)dx
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{l=1}^{J} \int_0^{\infty} \Xi_{t,s}^{j,l} x^{1-\gamma} \prod_{k \neq l} e^{-\left[\left(\frac{\Xi_{t,s}^{j,l}}{\Xi_{t,s}^{j,k}}\right)^{\frac{1}{1-\gamma}} \frac{x}{\lambda^k}\right]^{-\kappa}} \frac{\kappa}{\lambda^l} \left(\frac{x}{\lambda^l}\right)^{-1-\kappa} e^{-\left(\frac{x}{\lambda^l}\right)^{-\kappa}} dx \\
&= \sum_{l=1}^{J} \int_0^{\infty} \Xi_{t,s}^{j,l} (\lambda^l)^{\kappa} e^{-\left[\sum_{k \neq l}\left(\frac{\Xi_{t,s}^{j,l}}{\Xi_{t,s}^{j,k}}\right)^{\frac{1}{1-\gamma}} \frac{x}{\lambda^k}\right]^{-\kappa}} \kappa x^{-\gamma-\kappa} e^{-(\lambda^l)^{\kappa} x^{-\kappa}} dx \\
&= \sum_{l=1}^{J} \int_0^{\infty} \Xi_{t,s}^{j,l} (\lambda^l)^{\kappa} e^{-\left[\sum_{k \neq l}\left(\frac{\Xi_{t,s}^{j,k}}{\Xi_{t,s}^{j,l}}\right)^{\frac{\kappa}{1-\gamma}} (\lambda^k)^{\kappa}\right] x^{-\kappa} - (\lambda^l)^{\kappa} x^{-\kappa}} \kappa x^{-\gamma-\kappa} dx \\
&= \sum_{l=1}^{J} \Xi_{t,s}^{j,l} (\lambda^l)^{\kappa} \int_0^{\infty} e^{-\left[\sum_{k=1}^{J}\left(\frac{\Xi_{t,s}^{j,k}}{\Xi_{t,s}^{j,l}}\right)^{\frac{\kappa}{1-\gamma}} (\lambda^k)^{\kappa}\right] x^{-\kappa}} \kappa x^{-\gamma-\kappa} dx
\end{aligned}
\tag{2.B.2}
$$

Note that, for the fourth equality, the transformation inside the probability term is valid for both $0 < \gamma < 1$ and $\gamma > 1$: in the former case, $\Xi_{t,s}^{j,k}$ is positive and the transformation $(\cdot)^{\frac{1}{1-\gamma}}$ is monotonically increasing, meaning the direction of the inequality is unchanged in both steps; in the latter case, $\Xi_{t,s}^{j,k}$ is negative and the transformation $(\cdot)^{\frac{1}{1-\gamma}}$ is monotonically decreasing, meaning the direction of the inequality is reversed twice. The integral in (2.B.2) is solved by change of variables with

$$z \equiv \left[\sum_{k=1}^{J}\left(\frac{\Xi_{t,s}^{j,k}}{\Xi_{t,s}^{j,l}}\right)^{\frac{\kappa}{1-\gamma}} (\lambda^k)^{\kappa}\right] x^{-\kappa} \tag{2.B.3}$$

$$x = z^{-\frac{1}{\kappa}} \left[\sum_{k=1}^{J}\left(\frac{\Xi_{t,s}^{j,k}}{\Xi_{t,s}^{j,l}}\right)^{\frac{\kappa}{1-\gamma}} (\lambda^k)^{\kappa}\right]^{\frac{1}{\kappa}} \tag{2.B.4}$$

$$dx = -\frac{1}{\kappa} z^{-\frac{1}{\kappa}-1} \left[\sum_{k=1}^{J}\left(\frac{\Xi_{t,s}^{j,k}}{\Xi_{t,s}^{j,l}}\right)^{\frac{\kappa}{1-\gamma}} (\lambda^k)^{\kappa}\right]^{\frac{1}{\kappa}} dz \tag{2.B.5}$$

Note that $\lim_{x\to 0} z(x) = \infty$ and $\lim_{x\to\infty} z(x) = 0$. Substituting equations (2.B.3), (2.B.4), and (2.B.5) and accounting for the limiting behavior of z, then implies

$$\int_0^\infty e^{-\left[\sum_{k=1}^J \left(\frac{\Xi_{t,s}^{j,k}}{\Xi_{t,s}^{j,l}}\right)^{\frac{\kappa}{1-\gamma}} (\lambda^k)^\kappa\right] x^{-\kappa}} \kappa x^{-\gamma-\kappa} dx$$

$$= \int_\infty^0 \left\{ e^{-z}\kappa \left[z^{-\frac{1}{\kappa}} \left[\sum_{k=1}^J \left(\frac{\Xi_{t,s}^{j,k}}{\Xi_{t,s}^{j,l}}\right)^{\frac{\kappa}{1-\gamma}} (\lambda^k)^\kappa \right]^{\frac{1}{\kappa}} \right]^{-\gamma-\kappa} \left(-\frac{1}{\kappa}\right) z^{-\frac{1}{\kappa}-1} \left[\sum_{k=1}^J \left(\frac{\Xi_{t,s}^{j,k}}{\Xi_{t,s}^{j,l}}\right)^{\frac{\kappa}{1-\gamma}} (\lambda^k)^\kappa \right]^{\frac{1}{\kappa}} \right\} dz$$

$$= \left[\sum_{k=1}^J \left(\frac{\Xi_{t,s}^{j,k}}{\Xi_{t,s}^{j,l}}\right)^{\frac{\kappa}{1-\gamma}} (\lambda^k)^\kappa \right]^{\frac{1-\gamma}{\kappa}-1} \left(-\int_\infty^0 e^{-z} z^{-\frac{1-\gamma}{\kappa}} dz \right)$$

$$= \left[\sum_{k=1}^J \left(\frac{\Xi_{t,s}^{j,k}}{\Xi_{t,s}^{j,l}}\right)^{\frac{\kappa}{1-\gamma}} (\lambda^k)^\kappa \right]^{\frac{1-\gamma}{\kappa}-1} \left(\int_0^\infty e^{-z} z^{\left(1-\frac{1-\gamma}{\kappa}\right)-1} dz \right)$$

$$= \left[\sum_{k=1}^J \left(\frac{\Xi_{t,s}^{j,k}}{\Xi_{t,s}^{j,l}}\right)^{\frac{\kappa}{1-\gamma}} (\lambda^k)^\kappa \right]^{\frac{1-\gamma}{\kappa}-1} \Gamma\left(1 - \frac{1-\gamma}{\kappa}\right)$$

where $\Gamma(\cdot)$ denotes the Gamma function. Substituting back into equation (2.B.2) yields

$$\mathbb{E}_{\varepsilon_t}\left[\max_l \left\{\Xi_{t,s}^{j,l}(\varepsilon_t^l)^{1-\gamma}\right\}\right] = \Gamma\left(1 - \frac{1-\gamma}{\kappa}\right) \sum_{l=1}^J \Xi_{t,s}^{j,l} (\lambda^l)^\kappa \left[\sum_{k=1}^J \left(\frac{\Xi_{t,s}^{j,k}}{\Xi_{t,s}^{j,l}}\right)^{\frac{\kappa}{1-\gamma}} (\lambda^k)^\kappa \right]^{\frac{1-\gamma}{\kappa}-1} \tag{2.B.6}$$

Now, for the case of $0 < \gamma < 1$, consumption utility and value functions are strictly positive. Consequently, $\Xi_{t,s}^{j,l} > 0$ for all ages and sectors. Thus, the common denominator in the summation can be factored out, which results in

$$
\begin{aligned}
&\mathbb{E}_{\varepsilon_t}\left[\max_l\left\{\Xi_{t,s}^{j,l}\left(\varepsilon_t^l\right)^{1-\gamma}\right\}\right] \\
&\quad = \Gamma\left(1-\frac{1-\gamma}{\kappa}\right)\sum_{l=1}^{J}\Xi_{t,s}^{j,l}\left(\lambda^l\right)^{\kappa}\left[\left(\Xi_{t,s}^{j,l}\right)^{-\frac{\kappa}{1-\gamma}}\sum_{k=1}^{J}\left(\Xi_{t,s}^{j,k}\right)^{\frac{\kappa}{1-\gamma}}\left(\lambda^k\right)^{\kappa}\right]^{\frac{1-\gamma}{\kappa}-1} \\
&\quad = \Gamma\left(1-\frac{1-\gamma}{\kappa}\right)\sum_{l=1}^{J}\left(\Xi_{t,s}^{j,l}\right)^{\frac{\kappa}{1-\gamma}}\left(\lambda^l\right)^{\kappa}\left[\sum_{k=1}^{J}\left(\Xi_{t,s}^{j,k}\right)^{\frac{\kappa}{1-\gamma}}\left(\lambda^k\right)^{\kappa}\right]^{\frac{1-\gamma}{\kappa}-1} \\
&\quad = \Gamma\left(1-\frac{1-\gamma}{\kappa}\right)\left[\sum_{k=1}^{J}\left(\Xi_{t,s}^{j,k}\right)^{\frac{\kappa}{1-\gamma}}\left(\lambda^k\right)^{\kappa}\right]^{\frac{1-\gamma}{\kappa}-1}\left[\sum_{l=1}^{J}\left(\Xi_{t,s}^{j,l}\right)^{\frac{\kappa}{1-\gamma}}\left(\lambda^l\right)^{\kappa}\right] \\
&\quad = \Gamma\left(1-\frac{1-\gamma}{\kappa}\right)\left[\sum_{l=1}^{J}\left(\Xi_{t,s}^{j,l}\right)^{\frac{\kappa}{1-\gamma}}\left(\lambda^l\right)^{\kappa}\right]^{\frac{1-\gamma}{\kappa}} \\
&\quad = \Gamma\left(1-\frac{1-\gamma}{\kappa}\right)\left[\sum_{l=1}^{J}\left(\chi_{t,s}^{j,l}v_{t+1,s+1}^{l}\right)^{\frac{\kappa}{1-\gamma}}\left(\tau_{t,s}^{j,l}\lambda^l\right)^{\kappa}\right]^{\frac{1-\gamma}{\kappa}}
\end{aligned}
$$

Substituting back into equation (2.B.1) yields the expression in the proposition.

For the case of $\gamma > 1$, consumption utility and value functions are strictly negative. Consequently, directly factoring $\Xi_{t,s}^{j,l}$ out of the quotient is problematic. Note that equation (2.B.6) can be rewritten as

$$
\begin{aligned}
\mathbb{E}_{\varepsilon_t}\left[\max_l\left\{\Xi_{t,s}^{j,l}(\varepsilon_t^l)^{1-\gamma}\right\}\right] = & -\Gamma\left(1-\frac{1-\gamma}{\kappa}\right) \\
& \sum_{l=1}^{J}\left\{\left|\Xi_{t,s}^{j,l}\right|\left(\lambda^l\right)^{\kappa}\left[\sum_{k=1}^{J}\left(\frac{\left|\Xi_{t,s}^{j,k}\right|}{\left|\Xi_{t,s}^{j,l}\right|}\right)^{\frac{\kappa}{1-\gamma}}(\lambda^k)^{\kappa}\right]^{\frac{1-\gamma}{\kappa}-1}\right\}
\end{aligned}
$$

Following the same steps as for above yields

$$
\begin{aligned}
\mathbb{E}_{\varepsilon_t}\left[\max_l\left\{\Xi_{t,s}^{j,l}(\varepsilon_t^l)^{1-\gamma}\right\}\right] &= -\Gamma\left(1-\frac{1-\gamma}{\kappa}\right)\left[\sum_{l=1}^{J}\left(\left|\Xi_{t,s}^{j,l}\right|\right)^{\frac{\kappa}{1-\gamma}}(\lambda^l)^{\kappa}\right]^{\frac{1-\gamma}{\kappa}} \\
&= -\Gamma\left(1-\frac{1-\gamma}{\kappa}\right)\left[\sum_{l=1}^{J}\left(\chi_{t,s}^{j,l}\left|v_{t+1,s+1}^{l}\right|\right)^{\frac{\kappa}{1-\gamma}}\left(\tau_{t,s}^{j,l}\lambda^l\right)^{\kappa}\right]^{\frac{1-\gamma}{\kappa}} \\
&= -\Gamma\left(1-\frac{1-\gamma}{\kappa}\right)\left[\sum_{l=1}^{J}\left(-\chi_{t,s}^{j,l}v_{t+1,s+1}^{l}\right)^{\frac{\kappa}{1-\gamma}}(\tau_{t,s}^{j,l}\lambda^l)^{\kappa}\right]^{\frac{1-\gamma}{\kappa}}
\end{aligned}
$$

Again, substituting back into equation (2.B.1) yields the expression in the proposition.

Next, I solve for sector choice probabilities. The probability of a worker of age s that is currently attached to sector j to choose sector l in period t is given by

$$\begin{aligned}
\mu_{t,s}^{j,l} &= Pr\left(l = \arg\max_k \left\{\Xi_{t,s}^{j,k}\left(\varepsilon_t^k\right)^{1-\gamma}\right\}\right) \\
&= \int_0^\infty Pr\left(\Xi_{t,s}^{j,k}\left(\varepsilon_t^k\right)^{1-\gamma} < \Xi_{t,s}^{j,l}x^{1-\gamma} \ \forall k \neq l\right) f_{\varepsilon^l}(x)dx \\
&= \int_0^\infty \prod_{k\neq l} Pr\left(\Xi_{t,s}^{j,k}\left(\varepsilon_t^k\right)^{1-\gamma} < \Xi_{t,s}^{j,l}x^{1-\gamma}\right) f_{\varepsilon^l}(x)dx \\
&= \int_0^\infty \prod_{k\neq l} Pr\left(\varepsilon_t^k < \left(\frac{\Xi_{t,s}^{j,l}}{\Xi_{t,s}^{j,k}}\right)^{\frac{1}{1-\gamma}} x\right) f_{\varepsilon^l}(x)dx \\
&= \int_0^\infty \prod_{k\neq l} F_{\varepsilon^k}\left(\left(\frac{\Xi_{t,s}^{j,l}}{\Xi_{t,s}^{j,k}}\right)^{\frac{1}{1-\gamma}} x\right) f_{\varepsilon^l}(x)dx \\
&= \int_0^\infty \prod_{k\neq l} e^{-\left[\left(\frac{\Xi_{t,s}^{j,l}}{\Xi_{t,s}^{j,k}}\right)^{\frac{1}{1-\gamma}} \frac{x}{\lambda^k}\right]^{-\kappa}} e^{-\left(\frac{x}{\lambda^l}\right)^{-\kappa}} \frac{\kappa}{\lambda^l}\left(\frac{x}{\lambda^l}\right)^{-1-\kappa} dx \\
&= \left(\lambda^l\right)^\kappa \int_0^\infty e^{-\left[\sum_{k=1}^J \left(\frac{\Xi_{t,s}^{j,k}}{\Xi_{t,s}^{j,l}}\right)^{\frac{\kappa}{1-\gamma}} \left(\lambda^k\right)^\kappa\right] x^{-\kappa}} \kappa x^{-1-\kappa} dx
\end{aligned} \tag{2.B.7}$$

The direction of the inequality is preserved in the forth equality by the same reasoning as above. The integral in equation (2.B.7) is again solved with a change of variables with $u \equiv x^{-\kappa}$, implying $x = u^{-\frac{1}{\kappa}}$, $dx = -\frac{1}{\kappa}u^{-\frac{1}{\kappa}-1}du$, $\lim_{x\to 0} u(x) = \infty$, and $\lim_{x\to\infty} u(x) = 0$. Substituting all of this yields

$$\begin{aligned}
\mu_{t,s}^{j,l} &= \left(\lambda^l\right)^\kappa \int_0^\infty e^{-\left[\sum_{k=1}^J \left(\frac{\Xi_{t,s}^{j,k}}{\Xi_{t,s}^{j,l}}\right)^{\frac{\kappa}{1-\gamma}} \left(\lambda^k\right)^\kappa\right] x^{-\kappa}} \kappa x^{-1-\kappa} dx \\
&= \left(\lambda^l\right)^\kappa \int_\infty^0 e^{-\left[\sum_{k=1}^J \left(\frac{\Xi_{t,s}^{j,k}}{\Xi_{t,s}^{j,l}}\right)^{\frac{\kappa}{1-\gamma}} \left(\lambda^k\right)^\kappa\right] u} \kappa\left(u^{-\frac{1}{\kappa}}\right)^{-1-\kappa}\left(-\frac{1}{\kappa}\right)u^{-\frac{1}{\kappa}-1}du \\
&= \left(\lambda^l\right)^\kappa \int_0^\infty e^{-\left[\sum_{k=1}^J \left(\frac{\Xi_{t,s}^{j,k}}{\Xi_{t,s}^{j,l}}\right)^{\frac{\kappa}{1-\gamma}} \left(\lambda^k\right)^\kappa\right] u} du \\
&= \left(\lambda^l\right)^\kappa \left[\sum_{k=1}^J \left(\frac{\Xi_{t,s}^{j,k}}{\Xi_{t,s}^{j,l}}\right)^{\frac{\kappa}{1-\gamma}} \left(\lambda^k\right)^\kappa\right]^{-1}
\end{aligned} \tag{2.B.8}$$

For $0 < \gamma < 1$, $\Xi^{j,l}_{t,s}$ is strictly positive and can directly be factored out of the sum:

$$\begin{aligned}
\mu^{j,l}_{t,s} &= (\lambda^l)^\kappa \left[\sum_{k=1}^{J}\left(\frac{\Xi^{j,k}_{t,s}}{\Xi^{j,l}_{t,s}}\right)^{\frac{\kappa}{1-\gamma}}(\lambda^k)^\kappa\right]^{-1} \\
&= (\lambda^l)^\kappa \left(\Xi^{j,l}_{t,s}\right)^{\frac{\kappa}{1-\gamma}}\left[\sum_{k=1}^{J}\left(\Xi^{j,k}_{t,s}\right)^{\frac{\kappa}{1-\gamma}}(\lambda^k)^\kappa\right]^{-1} \\
&= \frac{\left(\tau^{j,l}_{t,s}\lambda^l\right)^\kappa\left(\chi^{j,l}_{t,s}v^l_{t+1,s+1}\right)^{\frac{\kappa}{1-\gamma}}}{\sum_{k=1}^J\left(\tau^{j,k}_{t,s}\lambda^k\right)^\kappa\left(\chi^{j,k}_{t,s}v^k_{t+1,s+1}\right)^{\frac{\kappa}{1-\gamma}}}
\end{aligned}$$

For $\gamma > 1$, rewriting equation (2.B.8) with absolute values gives

$$\begin{aligned}
\mu^{j,l}_{t,s} &= \left(\lambda^l\right)^\kappa \left[\sum_{k=1}^{J}\left(\frac{-\left|\Xi^{j,k}_{t,s}\right|}{-\left|\Xi^{j,l}_{t,s}\right|}\right)^{\frac{\kappa}{1-\gamma}}\left(\lambda^k\right)^\kappa\right]^{-1} \\
&= \left(\lambda^l\right)^\kappa \left(\left|\Xi^{j,l}_{t,s}\right|\right)^{\frac{\kappa}{1-\gamma}}\left[\sum_{k=1}^{J}\left(\left|\Xi^{j,k}_{t,s}\right|\right)^{\frac{\kappa}{1-\gamma}}\left(\lambda^k\right)^\kappa\right]^{-1} \\
&= \frac{\left(\tau^{j,l}_{t,s}\lambda^l\right)^\kappa\left(\chi^{j,l}_{t,s}\left|v^l_{t+1,s+1}\right|\right)^{\frac{\kappa}{1-\gamma}}}{\sum_{k=1}^J\left[\left(\tau^{j,k}_{t,s}\lambda^k\right)^\kappa\left(\chi^{j,k}_{t,s}\left|v^k_{t+1,s+1}\right|\right)^{\frac{\kappa}{1-\gamma}}\right]} \\
&= \frac{\left(\tau^{j,l}_{t,s}\lambda^l\right)^\kappa\left(-\chi^{j,l}_{t,s}v^l_{t+1,s+1}\right)^{\frac{\kappa}{1-\gamma}}}{\sum_{k=1}^J\left[\left(\tau^{j,k}_{t,s}\lambda^k\right)^\kappa\left(-\chi^{j,k}_{t,s}v^k_{t+1,s+1}\right)^{\frac{\kappa}{1-\gamma}}\right]}
\end{aligned}$$

yielding the expressions from the proposition. Defining $\Xi^l_{t,0} = \chi^l_{t,0}(\tau^l_{t,0})^{1-\gamma}v^l_{t+1,1}$ and following the same steps as above yields the results for the initial sector choice. ■

Proof of Lemma 3. For any period t, the mass of workers of age 1 that are attached to sector j is given by

$$\phi^j_{t,1} = \phi_{t-1,0}\mu^j_{t-1,0} \quad \forall j$$

or, written as vector product $\phi_{t,1} = \phi_{t-1,0}\mathscr{M}_{t-1,0}$. Similarly, for any age $s > 1$, the vector $\phi_{s,t}$ is determined by the mass of workers of age $s-1$ by sector in the previous period $(t-1)$ and the corresponding sector choice probabilities $\mathscr{M}_{t-1,s+1}$:

$$\phi_{t,s} = \phi_{t-1,s-1}\zeta_{t-1,s-1}\mathscr{M}_{t-1,s-1} \quad \forall s > 1 \;\; \forall t \tag{2.B.9}$$

Iterating backwards over (2.B.9) then yields for the workers of age s

$$\phi_{t,s} = \phi_{t-s,0}\prod_{m=1}^{s-1}\zeta_{t-s+m,m}\mathcal{M}_{t-s,0}\prod_{m=1}^{s-1}\mathcal{M}_{t-s+m,m} = \phi_{t-s,0}\tilde{\zeta}_{t,s-1}\mathcal{M}_{t-s,0}\prod_{m=1}^{s-1}\mathcal{M}_{t-s+m,m}$$

Stacking the expressions for all ages yields the term in the Lemma. ■

Proof of Lemma 4. The proof closely follows the proof of Lemma 3 in Dvorkin and Monge-Naranjo (2019). By definition, the expected value of the labor market opportunity for sector l, conditional on l being the preferred sector next period, is given by

$$\begin{aligned}
&\mathbb{E}_{\varepsilon_t}\left[\varepsilon_t^l \,\middle|\, l \in \arg\max_k\left\{\Xi_{t,s}^{j,k}\left(\varepsilon_t^k\right)^{1-\gamma}\right\}\right] \\
&\quad= \frac{\int_0^\infty xPr\left(\Xi_{t,s}^{j,k}\left(\varepsilon_t^k\right)^{1-\gamma} < \Xi_{t,s}^{j,l}(x)^{1-\gamma} \;\; \forall k \neq l\right)f_{\varepsilon^l}(x)dx}{\int_0^\infty Pr\left(\Xi_{t,s}^{j,k}\left(\varepsilon_t^k\right)^{1-\gamma} < \Xi_{t,s}^{j,l}(x)^{1-\gamma} \;\; \forall k \neq l\right)f_{\varepsilon^l}(x)dx} \\
&\quad= \left(\mu_{t,s}^{j,l}\right)^{-1}\int_0^\infty x\prod_{k\neq l}F_{\varepsilon^k}\left(\left(\frac{\Xi_{t,s}^{j,l}}{\Xi_{t,s}^{j,k}}\right)^{\frac{1}{1-\gamma}}x\right)f_{\varepsilon^l}(x)dx \\
&\quad= \left(\mu_{t,s}^{j,l}\right)^{-1}\int_0^\infty x\prod_{k\neq l}e^{-\left[\left(\frac{\Xi_{t,s}^{j,l}}{\Xi_{t,s}^{j,k}}\right)^{\frac{\kappa}{1-\gamma}}\left(\lambda^k\right)^\kappa\right]x^{-\kappa}}e^{-(\lambda^l)^\kappa x^{-\kappa}}\left(\lambda^l\right)^\kappa\kappa x^{-1-\kappa}dx \\
&\quad= \left(\mu_{t,s}^{j,l}\right)^{-1}\left(\lambda^l\right)^\kappa\int_0^\infty e^{-\left[\sum_{k=1}^J\left(\frac{\Xi_{t,s}^{j,l}}{\Xi_{t,s}^{j,k}}\right)^{\frac{\kappa}{1-\gamma}}\left(\lambda^k\right)^\kappa\right]x^{-\kappa}}x^{-\kappa}\kappa dx
\end{aligned} \tag{2.B.10}$$

The integral is again solved by change of variables with z as defined above. Substituting (2.B.3), (2.B.4), and (2.B.5) into the integral in (2.B.10) and accounting for integration bounds yields

$$\int_0^{\infty} e^{-\left[\sum_{k=1}^{J}\left(\frac{\Xi_{t,s}^{j,l}}{\Xi_{t,s}^{j,k}}\right)^{\frac{\kappa}{1-\gamma}}(\lambda^k)^{\kappa}\right]x^{-\kappa}} x^{-\kappa}\kappa dx$$

$$= \int_{\infty}^{0} e^{-z} z \left[\sum_{k=1}^{J}\left(\frac{\Xi_{t,s}^{j,l}}{\Xi_{t,s}^{j,k}}\right)^{\frac{\kappa}{1-\gamma}}(\lambda^k)^{\kappa}\right]^{-1} \kappa\left(-\frac{1}{\kappa}\right) z^{\frac{1}{\kappa}-1}\left[\sum_{k=1}^{J}\left(\frac{\Xi_{t,s}^{j,l}}{\Xi_{t,s}^{j,k}}\right)^{\frac{\kappa}{1-\gamma}}(\lambda^k)^{\kappa}\right]^{\frac{1}{\kappa}} dz$$

$$= \int_0^{\infty} e^{-z} z^{(1-\frac{1}{\kappa})-1} dz \left[\sum_{k=1}^{J}\left(\frac{\Xi_{t,s}^{j,l}}{\Xi_{t,s}^{j,k}}\right)^{\frac{\kappa}{1-\gamma}}(\lambda^k)^{\kappa}\right]^{\frac{1-\kappa}{\kappa}}$$

$$= \Gamma\left(1-\frac{1}{\kappa}\right)\left[\sum_{k=1}^{J}\left(\frac{\Xi_{t,s}^{j,l}}{\Xi_{t,s}^{j,k}}\right)^{\frac{\kappa}{1-\gamma}}(\lambda^k)^{\kappa}\right]^{\frac{1-\kappa}{\kappa}} \tag{2.B.11}$$

Substituting (2.B.11) back into (2.B.10) yields

$$\mathbb{E}_{\varepsilon_t}\left[\varepsilon_t^l \middle| l \in \arg\max_k \left\{\Xi_{t,s}^{j,k}\left(\varepsilon_t^k\right)^{1-\gamma}\right\}\right]$$

$$= \left(\mu_{t,s}^{j,l}\right)^{-1}(\lambda^l)^{\kappa}\,\Gamma\left(1-\frac{1}{\kappa}\right)\left[\sum_{k=1}^{J}\left(\frac{\Xi_{t,s}^{j,l}}{\Xi_{t,s}^{j,k}}\right)^{\frac{\kappa}{1-\gamma}}(\lambda^k)^{\kappa}\right]^{\frac{1-\kappa}{\kappa}}$$

$$= \Gamma\left(1-\frac{1}{\kappa}\right)\lambda^l\left(\mu_{t,s}^{j,l}\right)^{-1}\left[\frac{(\lambda^l)^{\kappa}}{\sum_{k=1}^{J}\left(\frac{\Xi_{t,s}^{j,l}}{\Xi_{t,s}^{j,k}}\right)^{\frac{\kappa}{1-\gamma}}(\lambda^k)^{\kappa}}\right]^{1-\frac{1}{\kappa}}$$

$$= \Gamma\left(1-\frac{1}{\kappa}\right)\lambda^l\left(\mu_{t,s}^{j,l}\right)^{-\frac{1}{\kappa}}$$

which is the expression from the proposition. ■

Proof of Lemma 5. In any given period, the total human capital of workers of age s in a given sector j consists of the human capital that all surviving workers that chose j in the previous period have accumulated.

$$h_{t,s}^{j} = \zeta_{t-1,s-1}\sum_{l=1}^{J}\mu_{t-1,s-1}^{l,j}h_{t-1,s-1}^{l}\tau_{t-1,s-1}^{l,j}\bar{\varepsilon}_{t-1,s-1}^{l,j}$$

Substituting the expression for $\bar{\varepsilon}_{t-1,s-1}^{l,j}$ from Lemma 4 yields

$$h_{t,s}^{j} = \Gamma\left(1-\frac{1}{\kappa}\right)\lambda^j\zeta_{t-1,s-1}\sum_{l=1}^{J}h_{t-1,s-1}^{l}\tau_{t-1,s-1}^{l,j}\left(\mu_{t-1,s-1}^{l,j}\right)^{1-\frac{1}{\kappa}} \tag{2.B.12}$$

Rewriting equation 2.B.12 in vector notation gives

$$\begin{aligned} h_{t,s} &= \Gamma\left(1-\frac{1}{\kappa}\right)\lambda\zeta_{t-1,s-1}\circ\left[h_{t-1,s-1}\left[\mathcal{T}_{t-1,s-1}\circ\left(\mathcal{M}_{t-1,s-1}\right)^{1-\frac{1}{\kappa}}\right]\right] \\ &= \zeta_{t-1,s-1}\tilde{\Gamma}\circ\left[h_{t-1,s-1}\tilde{\mathcal{M}}_{t-1,s-1}\right] \end{aligned} \qquad (2.B.13)$$

For any period, the total human capital of newborn workers is given by the initial capital level in that period times the mass of newborn workers in that period, $h_{t,0}=h_t^0\phi_{t,0}$. Iterating backwards over equation 2.B.13, human capital for any age $s>1$ can be expressed as

$$h_{t,s} = h^0_{t-s}\phi_{t-s,0}\tilde{\zeta}_{t,s-1}\left(\tilde{\Gamma}\right)^s\circ\left[\tilde{\mathcal{M}}_{t,0}\prod_{m=1}^{s-1}\tilde{\mathcal{M}}_{t-s+m,m}\right] \qquad (2.B.14)$$

Stacking the expressions for all ages yields the term in the Lemma. ■

Proof of Proposition 6. The first step of the proof is again by induction. Note that with constant survival probabilities, the function that returns the discounted expected retirement value is also time-invariant:

$$\Omega_t(b_t) = u(b_t)\sum_{m=1}^{S^r}\beta^m\Pi_{n=0}^{m-1}\zeta_{t+m,S^w+m} = u(b_t)\sum_{m=1}^{S^r}\beta^m\Pi_{n=0}^{m-1}\zeta_{S^w+m} = \Omega(b_t)$$

Next, suppose that the expected value of human capital upon retirement is constant over time, $h_t^{S^w}=h^{S^w}\ \forall t$. This, together with Assumption (W7), implies that the ratio of human capital upon retirement to the annuity payment is also constant, $\bar{b}_t=\frac{h_t^{S^w}}{b_t}=\frac{h^{S^w}}{b}=\bar{b}$. With Proposition 2, using (W5) and (W7), the expected value of being attached to sector j in the last age of working life then is

$$v^j_{t,S^w} = \Upsilon^j_t\frac{\left[(1-\tau^p)\omega^j_t\right]^{1-\gamma}}{1-\gamma}+\Omega_t(\bar{b}_t) = \Upsilon^j\frac{\left[(1-\tau^p)\omega^j\right]^{1-\gamma}}{1-\gamma}+\Omega(\bar{b}) = v^j_{S^w}\quad\forall t$$

which represents the induction start. Next, suppose that for some age s and some period t, the expected value of being attached to sector j for the next age $s+1$ is time-invariant for all sectors (induction assumption). Then, from Proposition 2, for $0<\gamma<1$, using (W5)

$$\begin{aligned} v^j_{t,s} &= \Upsilon^j_t\frac{\left[(1-\tau^p)\omega^j_t\right]^{1-\gamma}}{1-\gamma}+\beta\zeta_{t,s}\Gamma\left(1-\frac{1-\gamma}{\kappa}\right)\left[\sum_{l=1}^{J}\left(\chi^{j,l}_{t,s}v^l_{t+1,s+1}\right)^{\frac{\kappa}{1-\gamma}}\left(\tau^{j,l}_{t,s}\lambda^j\right)^{\kappa}\right]^{\frac{1-\gamma}{\kappa}} \\ &= \Upsilon^j\frac{\left[(1-\tau^p)\,\omega^j\right]^{1-\gamma}}{1-\gamma}+\beta\zeta_s\Gamma\left(1-\frac{1-\gamma}{\kappa}\right)\left[\sum_{l=1}^{J}\left(\chi^{j,l}_{s}v^l_{s+1}\right)^{\frac{\kappa}{1-\gamma}}\left(\tau^{j,l}_{s}\lambda^j\right)^{\kappa}\right]^{\frac{1-\gamma}{\kappa}} \\ &= v^j_s\quad\forall t,s,j \end{aligned}$$

The case of $\gamma > 1$ follows analogously. This completes the induction step and, hence, expected values are time-invariant for all periods, ages, and sectors:

$$v^j_{t,s} = v^j_s \quad \forall t,s,j$$

Consequently, initial sector choice vectors and sector choice matrices by age are time-invariant as well, yielding the first statement of the proposition. Finally notice that, by (W6), initial human capital is constant over time. With time-invariant initial levels and accumulation governed by time-invariant transition matrices, expected human capital upon retirement is constant over time as well, confirming the initial supposition.

Next, consider the vector of masses of workers of age s by sector in period t, $\{\phi^1_{t,s}, \ldots, \phi^J_{t,s}\}$:

$$\{\phi^1_{t,s}, \ldots, \phi^J_{t,s}\} = \phi_{t-s,0}\tilde{\zeta}_{t,s-1}\mathcal{M}_{t-s,0}\prod_{m=1}^{s-1}\mathcal{M}_{t-s+m,m} = \phi_{t,0}(1+n)^{-s}\tilde{\zeta}_{s-1}\mathcal{M}_0\prod_{m=1}^{s-1}\mathcal{M}_m$$

Dividing by the total population size in period t yields the population shares by sector of workers of age s

$$\begin{aligned}\{\tilde{\phi}^1_{t,s}, \ldots, \tilde{\phi}^J_{t,s}\} &= \frac{\{\phi^1_{t,s}, \ldots, \phi^J_{t,s}\}}{\phi_t} = \frac{\phi_{t,0}(1+n)^{-s}\tilde{\zeta}_{t,s-1}}{\phi_{t,0}\bar{\phi}}\mathcal{M}_0\prod_{m=1}^{s-1}\mathcal{M}_m \\ &= \tilde{\phi}_s\mathcal{M}_0\prod_{m=1}^{s-1}\mathcal{M}_m\end{aligned}$$

which is time-invariant, i.e. $\tilde{\phi}^j_{t,s} = \tilde{\phi}^j_s\ \forall t,s,j$. Stacking the expression for all ages yields the constant distribution of workers by age and sector $\tilde{\Phi}$.

Finally, consider the vector of total human capital of workers of age s by sector in period t. Again, applying assumptions and results from above yields

$$\begin{aligned}h_{t,s} &= h^0_{t-s}\phi_{t-s,0}\tilde{\zeta}_{t,s-1}\left(\tilde{\Gamma}\right)^s \circ \left[\tilde{\mathcal{M}}_{t-s,0}\prod_{m=1}^{s-1}\tilde{\mathcal{M}}_{t-s+m,m}\right] \\ &= h^0\phi_{t,0}(1+n)^{-s}\tilde{\zeta}_{s-1}\left(\tilde{\Gamma}\right)^s \circ \left[\tilde{\mathcal{M}}_0\prod_{m=1}^{s-1}\tilde{\mathcal{M}}_m\right]\end{aligned}$$

Dividing by the total population size in period t yields

$$\begin{aligned}\tilde{h}_{t,s} &= \frac{h_{t,s}}{\phi_t} = h^0\frac{\phi_{t,0}(1+n)^{-s}\tilde{\zeta}_{s-1}}{\phi_{t,0}\bar{\phi}}\left(\tilde{\Gamma}\right)^s \circ \left[\tilde{\mathcal{M}}_0\prod_{m=1}^{s-1}\tilde{\mathcal{M}}_m\right] \\ &= h^0\tilde{\phi}_s\left(\tilde{\Gamma}\right)^s \circ \left[\tilde{\mathcal{M}}_0\prod_{m=1}^{s-1}\tilde{\mathcal{M}}_m\right] \quad \forall t,s\end{aligned}$$

Again, the final statement of the proposition follows from stacking the expression for all ages. ■

Production. *Proof of Proposition 7.* The minimal unitary cost functions can be derived from the corresponding cost minimization problems to the profit maximization problems (2.15) and (2.17). For this, first note that the intermediate input production function (2.14) can be further broken down into a Cobb-Douglas production function with capital K_t^j and bundles of labor and machines B_t^j as inputs and ν as the input share of bundles. Bundles of labor and machines are then a CES-combination with machine-augmenting technology, labor weight η and elasticity of substitution σ_j. Denote by $c_t^{B^j}$ the minimal cost of producing one unit of bundles B_t^j and consider the case of $\sigma_j \neq 1$. The bundle production function is then

$$B_t^j = \left[(1-\eta)\left(\theta_t^j M_t^j\right)^{\frac{\sigma_j-1}{\sigma_j}} + \eta\left(H_t^j\right)^{\frac{\sigma_j-1}{\sigma_j}}\right]^{\frac{\sigma_j}{\sigma_j-1}} \tag{2.B.15}$$

The minimal cost is derived by solving the cost minimization problem

$$\begin{aligned} &\min_{\{H_t^j, M_t^j\}} w_t^j H_t^j + r_t^{M,j} M_t^j \\ &\text{s.t. } B_t^j = 1 \end{aligned} \tag{2.B.16}$$

Reformulating the optimization problem as Lagrangian and taking derivatives w.r.t. the variables H_t^j, M_t^j, and the Lagrange multiplier yields the first order conditions

$$w_t^j = \lambda \frac{\partial B_t^j\left(H_t^j, M_t^j\right)}{\partial H_t^j} = \lambda \frac{\eta\left(H_t^j\right)^{\frac{\sigma_j-1}{\sigma_j}}}{(1-\eta)\left(\theta_t^j M_t^j\right)^{\frac{\sigma_j-1}{\sigma_j}} + \eta\left(H_t^j\right)^{\frac{\sigma_j-1}{\sigma_j}}} \frac{B_t^j}{H_t^j} \tag{2.B.17}$$

$$r_t^{M,j} = \lambda \frac{\partial B_t^j\left(H_t^j, M_t^j\right)}{\partial M_t^j} = \lambda \frac{(1-\eta)\left(\theta_t^j M_t^j\right)^{\frac{\sigma_j-1}{\sigma_j}}}{(1-\eta)\left(\theta_t^j M_t^j\right)^{\frac{\sigma_j-1}{\sigma_j}} + \eta\left(H_t^j\right)^{\frac{\sigma_j-1}{\sigma_j}}} \frac{B_t^j}{M_t^j} \tag{2.B.18}$$

$$1 = \left[(1-\eta)\left(\theta_t^j M_t^j\right)^{\frac{\sigma_j-1}{\sigma_j}} + \eta\left(H_t^j\right)^{\frac{\sigma_j-1}{\sigma_j}}\right]^{\frac{\sigma_j}{\sigma_j-1}} \tag{2.B.19}$$

Dividing (2.B.17) by (2.B.18) and rearranging yields

$$\theta_t^j M_t^j = \left(\frac{1-\eta}{\eta} \frac{w_t^j \theta_t^j}{r_t^{M,j}}\right)^{\sigma_j} H_t^j \tag{2.B.20}$$

Substituting (2.B.20), the bundle production function becomes

$$\begin{aligned} B_t^j &= \left[(1-\eta)\left(\left(\frac{1-\eta}{\eta} \frac{w_t^j \theta_t^j}{r_t^{M,j}}\right)^{\sigma_j} H_t^j\right)^{\frac{\sigma_j-1}{\sigma_j}} + \eta\left(H_t^j\right)^{\frac{\sigma_j-1}{\sigma_j}}\right]^{\frac{\sigma_j}{\sigma_j-1}} \\ &= \left(\frac{w_t^j}{\eta}\right)^{\sigma_j} \left[(1-\eta)^{\sigma_j}\left(\frac{r_t^{M,j}}{\theta_t^j}\right)^{1-\sigma_j} + \eta^{1-\sigma_j}\left(w_t^j\right)^{\sigma_j}\right]^{\frac{\sigma_j}{\sigma_j-1}} H_t^j \end{aligned} \tag{2.B.21}$$

In combination with (2.B.19), this implies

$$H_t^j = \left(\frac{w_t^j}{\eta}\right)^{-\sigma_j} \left[(1-\eta)^{\sigma_j} \left(\frac{r_t^{M,j}}{\theta_t^j}\right)^{1-\sigma_j} + \eta^{1-\sigma_j} \left(w_t^j\right)^{\sigma_j}\right]^{\frac{\sigma_j}{1-\sigma_j}}$$

Finally, the cost of producing bundle quantity B_t^j (which is equal to one by the optimization constraint) can be derived as

$$\begin{aligned} c_t^{B^j} &= w_t^j H_t^j + r_t^{M,j} M_t^j = w_t^j H_t^j + r_t^{M,j} \left(\frac{1-\eta}{\eta} \frac{w_t^j \theta_t^j}{r_t^{M,j}}\right)^{\sigma_j} \frac{H_t^j}{\theta_t^j} \\ &= \left[w_t^j + \frac{r_t^{M,j}}{\theta_t^j} \left(\frac{1-\eta}{\eta} \frac{w_t^j \theta_t^j}{r_t^{M,j}}\right)^{\sigma_j}\right] \left(\frac{w_t^j}{\eta}\right)^{-\sigma_j} \\ &\qquad \left[(1-\eta)^{\sigma_j} \left(\frac{r_t^{M,j}}{\theta_t^j}\right)^{1-\sigma_j} + \eta^{1-\sigma_j} \left(w_t^j\right)^{\sigma_j}\right]^{\frac{\sigma_j}{1-\sigma_j}} \\ &= \left[(1-\eta)^{\sigma_j} \left(\frac{r_t^{M,j}}{\theta_t^j}\right)^{1-\sigma_j} + \eta^{1-\sigma_j} \left(w_t^j\right)^{\sigma_j}\right]^{\frac{1}{1-\sigma_j}} \end{aligned}$$

For $\sigma_j = 1$, the bundle production function (2.B.15) becomes

$$B_t^j = \left(M_t^j\right)^{1-\eta} \left(H_t^j\right)^{\eta}$$

and the corresponding first-order conditions of the cost minimization problem are

$$w_t^j = \lambda \eta \frac{B_t^j}{H_t^j} \tag{2.B.22}$$

$$r_t^{M,j} = \lambda (1-\eta) \frac{B_t^j}{M_t^j} \tag{2.B.23}$$

$$1 = \left(M_t^j\right)^{1-\eta} \left(H_t^j\right)^{\eta} \tag{2.B.24}$$

Following the same steps as before yields

$$\theta_t^j M_t^j = \frac{1-\eta}{\eta} \frac{w_t^j \theta_t^j}{r_t^{M,j}} H_t^j$$

and

$$B_t^j = \left(\frac{1-\eta}{\eta} \frac{w_t^j \theta_t^j}{r_t^{M,j}}\right)^{1-\eta} H_t^j$$

which, together with the final first-order condition yields

$$H_t^j = \left(\frac{1-\eta}{\eta}\frac{w_t^j \theta_t^j}{r_t^{M,j}}\right)^{\eta-1}$$

Finally, constructing the cost of producing B_t^j (which is again equal to one)

$$c_t^{B^j} = \left(\frac{w_t^j}{\eta}\right)^{\eta}\left(\frac{r_t^{M,j}}{(1-\eta)\theta_t^j}\right)^{1-\eta} \tag{2.B.25}$$

Combining the cases (i.e. equation (2.B.16) for $\sigma_j \neq 1$ and equation (2.B.25) for $\sigma_j = 1$) results in the expression from the proposition. The minimal unitary cost for producing one unit of intermediate input G^j, for producing one unit of intermediate output Q^j, and for producing one unit of the final good Y can then be derived from the cost of producing one unit of the bundle B_t^j by using the well-known unitary cost functions for standard Cobb-Douglas and CES production functions.

Next, unit demand functions, i.e. input demand functions to produce one unit of final output Y, are derived. For this, note that equation (2.B.21) holds irrespective of the quantity of bundles B_t^j that is to be produced. Inverting the equation and substituting (2.B.16) results in

$$H_t^j = \left(\eta\frac{c_t^{B^j}}{w_t^j}\right)^{\sigma_j} B_t^j \tag{2.B.26}$$

In combination with (2.B.20), (2.B.26) simplifies to

$$M_t^j = \left(\frac{1-\eta}{\eta}\frac{w_t^j \theta_t^j}{r_t^{M,j}}\right)^{\sigma_j}\left(\theta_t^j\right)^{-1} H_t^j = \left((1-\eta)\frac{\theta_t^j c_t^{B^j}}{r_t^{M,j}}\right)^{\sigma_j}\left(\theta_t^j\right)^{-1} B_t^j \tag{2.B.27}$$

The expressions from the proposition then follow again from applying the well-known solutions for unit demand functions of standard Cobb-Douglas and CES production functions. ■

Proof of Proposition 8. This proof consists of several parts: separation of optimizations, normalization of the Bellman equation, and properties of the normalized Bellman equation.

Separation of optimizations. First, notice that the price of technology set by the monopolists' only enters the Bellman equation via the profit function. The profit function in turn only depends on the current stock of technology, but not on next periods' stock of technology. Moreover, note that differentiability of $\Pi^j(\cdot)$ is sufficient to derive the first-order condition on the price of technology (irrespective of differentiability of the value functions). Taking derivatives of the objective function in problem (2.18) w.r.t. $p_t^{T,j}$ yields

$$\frac{1-\alpha}{\alpha}\left(p_t^{T,j} - \frac{x_t^j}{1-\alpha}\right)\left(p_t^{T,j}\right)^{-\frac{1+\alpha}{\alpha}} \left\{\left[(1-\alpha)\varphi^j p_t^{Q^j}\right]^{\frac{1}{\alpha}} G^j\left(H_t^j, M_t^j, K_t^j, \theta_t^j\right)\right\} = 0 \tag{2.B.28}$$

Strictly positive per capita stocks of input factors imply that the only solution to (2.B.28) is $p_t^{T,j} = (1-\alpha)^{-1}x_t^j$.[16] Normalizing by the price of the output good results in the solution stated in the proposition. As the solution to the price-setting problem is independent of the level of technology (in the current or any future period), the (static) price-setting problem can be separated from the (dynamic) technology creation problem and the solution to the price-setting problem can be substituted into (2.18), yielding

$$\Pi_t^{j*}\left(\theta_t^j\right) = (1-\alpha)\pi_t^{G^j} G^j\left(H_t^j, M_t^j, K_t^j, \theta_t^j\right)$$

Normalization of the Bellman equation. For the normalization, first note that real monopolists' per-capita profits are

$$\begin{aligned}
\tilde{\Pi}_t^{j*}\left(\theta_t^j\right) &= \frac{\Pi_t^{j*}\left(\theta^j\right)}{\phi_t} = \frac{(1-\alpha)\pi_t^{G^j} G^j\left(H_t^j, M_t^j, K_t^j, \theta_t^j\right)}{\phi_t} \\
&= (1-\alpha)\pi_t^{G^j} G^j\left(\frac{H_t^j}{\phi_t}, \frac{M_t^j}{\phi_t}, \frac{K_t^j}{\phi_t}, \theta_t^j\right) \\
&= (1-\alpha)\pi_t^{G^j} G^j\left(\tilde{H}_t^j, \tilde{M}_t^j, \tilde{K}_t^j, \theta_t^j\right)
\end{aligned}$$

where the second line follows from the constant returns to scale property of G^j and the last line employs the definition of per-capita quantities of human capital,

16. This is the standard solution of monopoly pricing settings with constant demand elasticity; the optimal price is a constant markup over marginal cost and the markup is given by the elasticity.

machines, and structures. Moreover, Assumption (P2) implies that

$$\begin{aligned}
I_t^{\theta,j} &= A(\theta^j)C_t\left(\theta_{t+1}^j-(1-\delta^\theta)\theta_t^j\right) = \phi_t A(\theta^j)\tilde{C}\left(\theta_{t+1}^j-(1-\delta^\theta)\theta_t^j\right)\\
\Leftrightarrow \tilde{I}_t^{\theta,j} &= \frac{I_t^{\theta,j}}{\phi_t} = A(\theta^j)\tilde{C}\left(\theta_{t+1}^j-(1-\delta^\theta)\theta_t^j\right)\\
\Leftrightarrow \left[A(\theta^j)\right]^{-1}\tilde{I}_t^{\theta,j} &= \tilde{C}\left(F_t\left(I_t^{\theta,j},\theta_t^j\right)\right)\\
\Leftrightarrow F_t\left(I_t^{\theta,j},\theta_t^j\right) &= \tilde{C}^{-1}\left(\left[A(\theta^j)\right]^{-1}\tilde{I}_t^{\theta,j}\right) \equiv \tilde{F}\left(\tilde{I}_t^{\theta,j},\theta_t^j\right)
\end{aligned}$$

Note that the technology creation function $\tilde{F}(\cdot)$ is time-invariant. Using the above, the law of motion for technology θ^j can be written in terms of per-capita investment as

$$\theta_{t+1}^j = (1-\delta^\theta)\theta_t^j+\tilde{F}\left(\tilde{I}_t^{\theta,j},\theta_t^j\right) \tag{2.B.29}$$

Under assumptions (P1)–(P2) and substituting the solution to the price setting problem, the optimization problem thus becomes

$$\begin{aligned}
V_t^j\left(\theta_t^j\right) &= \max_{\tilde{I}_t^{\theta,j}}\left\{\tilde{\Pi}_t^{j*}\left(\theta_t^j\right)-\tilde{I}_t^{\theta,j}+\frac{1}{1+\rho_t}V_{t+1}^j\left(\theta_{t+1}^j\right)\right\}\\
\text{s.t.}\quad \tilde{\Pi}_t^{j*}\left(\theta_t^j\right) &= (1-\alpha)\pi_t^{G^j}G^j\left(\tilde{H}_t^j,\tilde{M}_t^j,\tilde{K}_t^j,\theta_t^j\right)\\
\theta_{t+1}^j &= (1-\delta^\theta)\theta_t^j+\tilde{F}\left(\tilde{I}_t^{\theta,j},\theta_t^j\right)\\
\tilde{I}_t^{\theta,j} &\in \left[0,\tilde{\Pi}_t^{j*}(\theta_t^j)\right]
\end{aligned} \tag{2.B.30}$$

With constant per capita stocks of inputs and constant prices, the per-capita profit function is time-invariant as well. That means that none of the functions in the constraints change over time and the problem in period t is identical to the problem in $t+1$ (except for the starting value). Consequently, time-subscripts can be dropped and the Bellman equation becomes a proper recursion.

Properties of the Bellman equation. I now demonstrate that, under the assumptions made above, a solution to (2.B.30) exists. For this, let $\theta^{j\prime}$ denote next period technology level in sector j. Let Θ denote the domain of technology levels. Denote by $\Gamma^j:\Theta\to\Theta$ the constraint correspondence and by $w^j:X^j\to\mathbb{R}$ the reward function, where

$$X^j = \left\{\left(\theta^j,\theta^{j\prime}\right)\in\Theta\times\Theta\;:\;\theta^{j\prime}\in\Gamma^j\left(\theta^j\right)\right\}$$

denotes the set of feasible combinations of technology today and technology next period. The optimization problem can then be re-written as

$$V^j\left(\theta^j\right) = \max_{\theta^{j\prime}\in\Gamma^j\left(\theta^j\right)}\left\{w^j\left(\theta^j,\theta^{j\prime}\right)+\frac{1}{1+\rho}V^j\left(\theta^{j\prime}\right)\right\} \tag{2.B.31}$$

$$\text{s.t.}\quad w^j\left(\theta^j,\theta^{j\prime}\right) = \tilde{\Pi}^{j*}\left(\theta^j\right)-A\left(\theta^j\right)\tilde{C}\left(\theta^{j\prime}-\left(1-\delta^\theta\theta^j\right)\right) \tag{2.B.32}$$

$$\Gamma\left(\theta^j\right) = \left[\left(1-\delta^\theta\right)\theta^j,\left(1-\delta^\theta\right)\theta^j+\tilde{F}\left(\tilde{\Pi}^{j*}\left(\theta^j\right),\theta^j\right)\right] \tag{2.B.33}$$

By definition, technology levels are non-negative, i.e. $\Theta = \mathbb{R}_0^+$, which is a convex subset of $\mathbb{R}$. For given input prices and per capita stocks of inputs, the monopolists' per capita profit under optimal pricing $\tilde{\Pi}^{j*}\left(\theta^j\right)$ is non-negative and, for $0 < \sigma_j < 1$, bounded over Θ. To see this, note that the profit function is continuous for $\theta^j \in \Theta$, and boundedness at the limits of its domain is sufficient for boundedness over the domain. The limits of profits as θ^j approaches the bounds of Θ are, for $0 < \sigma_j < 1$

$$\begin{aligned}
\lim_{\theta^j \to 0} \tilde{\Pi}^{j*}\left(\theta^j\right) &= 0 \\
\lim_{\theta^j \to \infty} \tilde{\Pi}^{j*}\left(\theta^j\right) &= (1-\alpha)\pi^{G^j}\eta^{\frac{\sigma_j \nu}{\sigma_j - 1}}\left(\tilde{K}^j\right)^{1-\alpha}\left(\tilde{H}^j\right)^{\nu}
\end{aligned}$$

which are both finite. Moreover, profits are increasing and concave in θ^j. To see this, let

$$m^j\left(\theta^j\right) = \frac{(1-\eta)\left(\theta^j\tilde{M}^j\right)^{\frac{\sigma_j-1}{\sigma_j}}}{(1-\eta)\left(\theta^j\tilde{M}^j\right)^{\frac{\sigma_j-1}{\sigma_j}} + \eta\left(\tilde{H}^j\right)^{\frac{\sigma_j-1}{\sigma_j}}}$$

and note that $m^j\left(\theta^j\right) \in [0,1]$. Taking derivatives of $\tilde{\Pi}^{j*}$ twice yields

$$\tilde{\Pi}^{j*}\left(\theta^j\right) = (1-\alpha)\pi^{G^j}G^j\left(\tilde{H}^j,\tilde{M}^j,\tilde{K}^j,\theta^j\right) \geq 0 \tag{2.B.34}$$

$$\frac{d\tilde{\Pi}^{j*}\left(\theta^j\right)}{d\theta^j} = \nu\tilde{\Pi}^{j*}\left(\theta^j\right)\frac{m^j\left(\theta^j\right)}{\theta^j} > 0 \tag{2.B.35}$$

$$\frac{d^2\tilde{\Pi}^{j*}\left(\theta^j\right)}{d\left(\theta^j\right)^2} = -\nu\tilde{\Pi}^{j*}\left(\theta^j\right)\frac{m^j\left(\theta^j\right)}{\theta^j}\frac{(1-\nu)m^j\left(\theta^j\right) + \frac{1}{\sigma_j}\left[1-m^j\left(\theta^j\right)\right]}{\theta^j} < 0 \tag{2.B.36}$$

Under assumptions (P2)–(P3), the technology creation function $\tilde{F}(\cdot)$ is continuous and bounded over closed sets. Moreover, the reward function is continuously differentiable over X^j and boundedness of profits translates into boundedness of the reward function: as the monopolists' can always choose not to invest at all, the reward function is bounded above by

$$\begin{aligned}
\bar{w}^j\left(\theta^j\right) &= \tilde{\Pi}^{j*}\left(\theta^j\right) \\
&\geq \tilde{\Pi}^{j*}\left(\theta^j\right) - A\left(\theta^j\right)\tilde{C}\left(\theta^{j\prime} - \left(1-\delta^{\theta}\right)\theta^j\right) \\
&= w^j\left(\theta^j,\theta^{j\prime}\right) \quad \forall\left(\theta^j,\theta^{j\prime}\right) \in X^j
\end{aligned}$$

Under assumptions (P2)–(P4), the monotonicity and concavity of $\tilde{\Pi}^{j*}$ carry over into monotonicity (in θ^j) and concavity of the reward function w^j and monotonicity and concavity of the upper bound of the constraint correspondence Γ^j. To see this, note that the first and second partial derivative of w^j w.r.t. θ^j are

$$\begin{aligned}
\frac{\partial w^j\left(\theta^j,\theta^{j\prime}\right)}{\partial\theta^j} &= \frac{d\tilde{\Pi}^{j*}\left(\theta^j\right)}{d\theta^j} - \frac{dA\left(\theta^j\right)}{d\theta^j}\tilde{C}\left(\Delta\theta^j\right) \\
&\quad + \left(1-\delta^\theta\right)A\left(\theta^j\right)\frac{d\tilde{C}\left(\Delta\theta^j\right)}{d\Delta\theta^j} \\
&> 0 \\
\frac{\partial^2 w^j\left(\theta^j,\theta^{j\prime}\right)}{\partial\left(\theta^j\right)^2} &= \frac{d^2\tilde{\Pi}^{j*}\left(\theta^j\right)}{d\left(\theta^j\right)^2} - \frac{d^2A\left(\theta^j\right)}{d\left(\theta^j\right)^2}\tilde{C}\left(\Delta\theta^j\right) \\
&\quad + 2\left(1-\delta^\theta\right)\frac{dA\left(\theta^j\right)}{d\theta^j}\frac{d\tilde{C}\left(\Delta\theta^j\right)}{d\Delta\theta^j} \\
&\quad - \left(1-\delta^\theta\right)^2 A\left(\theta^j\right)\frac{d^2\tilde{C}\left(\Delta\theta^j\right)}{d\left(\Delta\theta^j\right)^2} \\
&< 0
\end{aligned}$$

where the inequalities follow directly from the properties of $\tilde{\Pi}^{j*}(\cdot)$, $A(\cdot)$, and $\tilde{C}(\cdot)$.

The determinant of the Hessian of w^j is

$$\begin{aligned}
\det\left(H_{w^j}\right) &= A\left(\theta^j\right)\frac{d^2A\left(\theta^j\right)}{d\left(\theta^j\right)^2}\tilde{C}\left(\Delta\theta^j\right)\frac{d^2\tilde{C}\left(\Delta\theta^j\right)}{d\left(\Delta\theta^j\right)^2} - \left[\frac{dA\left(\theta^j\right)}{d\theta^j}\frac{d\tilde{C}\left(\Delta\theta^j\right)}{d\Delta\theta^j}\right]^2 \\
&\quad - \frac{d^2\tilde{\Pi}^{j*}\left(\theta^j\right)}{d\left(\theta^j\right)^2}A\left(\theta^j\right)\frac{d^2\tilde{C}\left(\Delta\theta^j\right)}{d\left(\Delta\theta^j\right)^2} \\
&> A\left(\theta^j\right)\frac{d^2A\left(\theta^j\right)}{d\left(\theta^j\right)^2}\tilde{C}\left(\Delta\theta^j\right)\frac{d^2\tilde{C}\left(\Delta\theta^j\right)}{d\left(\Delta\theta^j\right)^2} - \left[\frac{dA\left(\theta^j\right)}{d\theta^j}\frac{d\tilde{C}\left(\Delta\theta^j\right)}{d\Delta\theta^j}\right]^2 \\
&\geq 0
\end{aligned}$$

where the first inequality follows from $\frac{d^2\tilde{C}\left(\Delta\theta^j\right)}{d\left(\Delta\theta^j\right)^2} > 0$ and $\frac{d^2\tilde{\Pi}^{j*}\left(\theta^j\right)}{d\left(\theta^j\right)^2} < 0$ and the last inequality follows from Assumption (P4).[17] The first principal minor being strictly negative and the determinant being strictly positive are sufficient conditions for the H_{w^j} to be negative definite and, therefore, w^j to be strictly concave.

Next, let $f^j\left(\theta^j\right)$ denote the upper bound of $\Gamma^j\left(\theta^j\right)$ and consider its first two derivatives:

17. If $A(\cdot)$ is linear, the determinant simplifies to $\det\left(H_{w^j}\right) = -\frac{d^2\tilde{\Pi}^{j*}\left(\theta^j\right)}{d\left(\theta^j\right)^2}A\left(\theta^j\right)\frac{d^2\tilde{C}\left(\Delta\theta^j\right)}{d\left(\Delta\theta^j\right)^2} > 0$.

$$f^j\left(\theta^j\right) = \left(1-\delta^\theta\right)\theta^j + \tilde{F}\left(\tilde{\Pi}^{j*}\left(\theta^j\right),\theta^j\right) \tag{2.B.37}$$

$$\frac{df^j\left(\theta^j\right)}{d\theta^j} = \left(1-\delta^\theta\right) + \frac{d\tilde{\Pi}^{j*}\left(\theta^j\right)}{d\theta^j}\frac{\partial\tilde{F}\left(\tilde{\Pi}^{j*}\left(\theta^j\right),\theta^j\right)}{\partial\tilde{I}^{\theta,j}} + \frac{\partial\tilde{F}\left(\tilde{\Pi}^{j*}\left(\theta^j\right),\theta^j\right)}{\partial\theta^j} \tag{2.B.38}$$

$$\begin{aligned}\frac{d^2f^j\left(\theta^j\right)}{d\left(\theta^j\right)^2} = {} & \frac{d^2\tilde{\Pi}^{j*}\left(\theta^j\right)}{d\left(\theta^j\right)^2}\frac{\partial\tilde{F}\left(\tilde{\Pi}^{j*}\left(\theta^j\right),\theta^j\right)}{\partial\tilde{I}^{\theta,j}} + \left(\frac{d\tilde{\Pi}^{j*}\left(\theta^j\right)}{d\theta^j}\right)^2\frac{\partial^2\tilde{F}\left(\tilde{\Pi}^{j*}\left(\theta^j\right),\theta^j\right)}{\partial\left(\tilde{I}^{\theta,j}\right)^2} \\ & + 2\frac{d\tilde{\Pi}^{j*}\left(\theta^j\right)}{d\theta^j}\frac{\partial^2\tilde{F}\left(\tilde{\Pi}^{j*}\left(\theta^j\right),\theta^j\right)}{\partial\tilde{I}^{\theta,j}\partial\theta^j} + \frac{\partial^2\tilde{F}\left(\tilde{\Pi}^{j*}\left(\theta^j\right),\theta^j\right)}{\partial\left(\theta^j\right)^2}\end{aligned} \tag{2.B.39}$$

The partial derivatives of $\tilde{F}(\cdot)$ and their properties can be derived from the definition of $\tilde{F}(\cdot)$ and the characteristics of $A(\cdot)$ and $\tilde{C}(\cdot)$:

$$\frac{\partial\tilde{F}\left(\tilde{I}^{\theta,j},\theta^j\right)}{\partial\tilde{I}^{\theta,j}} = \left[A\left(\theta^j\right)\frac{d\tilde{C}\left(\tilde{C}^{-1}\left(\left[A\left(\theta^j\right)\right]^{-1}\tilde{I}^{\theta,j}\right)\right)}{d\Delta\theta^j}\right]^{-1} > 0$$

$$\frac{\partial\tilde{F}\left(\tilde{I}^{\theta,j},\theta^j\right)}{\partial\theta^j} = \tilde{I}^{\theta,j}\left(-\frac{dA\left(\theta^j\right)}{d\theta^j}\right)\left[A\left(\theta^j\right)\frac{d\tilde{C}\left(\tilde{C}^{-1}\left(\left[A\left(\theta^j\right)\right]^{-1}\tilde{I}^{\theta,j}\right)\right)}{d\Delta\theta^j}\right]^{-1} > 0$$

i.e. $\tilde{F}(\cdot)$ is strictly increasing in both arguments.

Substituting the expressions for the partial derivatives into (2.B.38) and rearranging yields

$$\begin{aligned}\frac{df^j\left(\theta^j\right)}{d\theta^j} = {} & \left(1-\delta^\theta\right) + \frac{d\tilde{\Pi}^{j*}\left(\theta^j\right)}{d\theta^j}\left[A\left(\theta^j\right)\frac{d\tilde{C}\left(\tilde{C}^{-1}\left(\left[A\left(\theta^j\right)\right]^{-1}\tilde{I}^{\theta,j}\right)\right)}{d\Delta\theta^j}\right]^{-1} \\ & + \frac{d\tilde{\Pi}^{j*}\left(\theta^j\right)}{d\theta^j}\left[A\left(\theta^j\right)\right]^{-2}\frac{dA\left(\theta^j\right)}{d\theta^j}\left[\frac{d\tilde{C}\left(\tilde{C}^{-1}\left(\left[A\left(\theta^j\right)\right]^{-1}\tilde{I}^{\theta,j}\right)\right)}{d\Delta\theta^j}\right]^{-1}\end{aligned}$$

The upper bound is strictly increasing if and only if $\frac{df^j\left(\theta^j\right)}{d\theta^j} > 0$. Rearranging terms results in the condition

$$\left[A\left(\theta^j\right)\right]^{-1}\left(-\frac{dA\left(\theta^j\right)}{d\theta^j}\right) < 1 + (1-\delta^\theta)\frac{A\left(\theta^j\right)}{\tilde{\Pi}^{j*}\left(\theta^j\right)}\frac{d\tilde{C}\left(\tilde{C}^{-1}\left(\left[A\left(\theta^j\right)\right]^{-1}\tilde{I}^{\theta,j}\right)\right)}{d\Delta\theta^j}$$

which is ensured by Assumption (P3) (the second summand on the RHS is strictly positive, hence the weak inequality in (P3) is sufficient; note that the condition is stricter than necessary).

Taking the second partial derivatives of $\tilde{F}(\cdot)$ yields, for investment in new patents $\tilde{I}^{\theta,j}$,

$$\frac{\partial^2 \tilde{F}\left(\tilde{I}^{\theta,j},\theta^j\right)}{\partial\left(\tilde{I}^{\theta,j}\right)^2} = -\left[A\left(\theta^j\right)\right]^{-2}\left[\frac{d\tilde{C}\left(\tilde{C}^{-1}\left(\left[A\left(\theta^j\right)\right]^{-1}\tilde{I}^{\theta,j}\right)\right)}{d\Delta\theta^j}\right]^{-3} \frac{d^2\tilde{C}\left(\tilde{C}^{-1}\left(\left[A\left(\theta^j\right)\right]^{-1}\tilde{I}^{\theta,j}\right)\right)}{d\left(\Delta\theta^j\right)^2}$$

and, for current stock of patents θ^j,

$$\begin{aligned}\frac{\partial^2 \tilde{F}\left(\tilde{I}^{\theta,j},\theta^j\right)}{\partial\left(\theta^j\right)^2} &= -\tilde{I}^{\theta,j}\left[A\left(\theta^j\right)\right]^{-2}\frac{d^2A\left(\theta^j\right)}{d\left(\theta^j\right)^2}\left[\frac{d\tilde{C}\left(\tilde{C}^{-1}\left(\left[A\left(\theta^j\right)\right]^{-1}\tilde{I}^{\theta,j}\right)\right)}{d\Delta\theta^j}\right]^{-1} \\ &+ 2\tilde{I}^{\theta,j}\left[A\left(\theta^j\right)\right]^{-3}\left(\frac{dA\left(\theta^j\right)}{d\theta^j}\right)^2\left[\frac{d\tilde{C}\left(\tilde{C}^{-1}\left(\left[A\left(\theta^j\right)\right]^{-1}\tilde{I}^{\theta,j}\right)\right)}{d\Delta\theta^j}\right]^{-1} \\ &- \left(\tilde{I}^{\theta,j}\right)^2\left[A\left(\theta^j\right)\right]^{-4}\left(\frac{dA\left(\theta^j\right)}{d\theta^j}\right)^2\left[\frac{d\tilde{C}\left(\tilde{C}^{-1}\left(\left[A\left(\theta^j\right)\right]^{-1}\tilde{I}^{\theta,j}\right)\right)}{d\Delta\theta^j}\right]^{-3} \\ &\quad \frac{d^2\tilde{C}\left(\tilde{C}^{-1}\left(\left[A\left(\theta^j\right)\right]^{-1}\tilde{I}^{\theta,j}\right)\right)}{d\left(\Delta\theta^j\right)^2}\end{aligned}$$

and, for the cross partial derivative,

$$\begin{aligned}\frac{\partial^2 \tilde{F}\left(\tilde{I}^{\theta,j},\theta^j\right)}{\partial\tilde{I}^{\theta,j}\partial\theta^j} &= -\left[A\left(\theta^j\right)\right]^{-2}\frac{dA\left(\theta^j\right)}{d\theta^j}\left[\frac{d\tilde{C}\left(\tilde{C}^{-1}\left(\left[A\left(\theta^j\right)\right]^{-1}\tilde{I}^{\theta,j}\right)\right)}{d\Delta\theta^j}\right]^{-1} \\ &+ \left[A\left(\theta^j\right)\right]^{-3}\frac{dA\left(\theta^j\right)}{d\theta^j}\left[\frac{d\tilde{C}\left(\tilde{C}^{-1}\left(\left[A\left(\theta^j\right)\right]^{-1}\tilde{I}^{\theta,j}\right)\right)}{d\Delta\theta^j}\right]^{-3} \\ &\quad \frac{d^2\tilde{C}\left(\tilde{C}^{-1}\left(\left[A\left(\theta^j\right)\right]^{-1}\tilde{I}^{\theta,j}\right)\right)}{d\left(\Delta\theta^j\right)^2}\end{aligned}$$

Substituting into (2.B.39) yields

$$\frac{d^2 f^j\left(\theta^j\right)}{d\left(\theta^j\right)^2} = \left[A\left(\theta^j\right)\right]^{-1}\left[\frac{d\tilde{C}\left(\tilde{C}^{-1}\left(\left[A\left(\theta^j\right)\right]^{-1}\tilde{I}^{\theta,j}\right)\right)}{d\Delta\theta^j}\right]^{-1}$$
$$\left[A\left(\theta^j\right)\frac{d^2\tilde{\Pi}^{j*}\left(\theta^j\right)}{d\left(\theta^j\right)^2} - 2\frac{dA\left(\theta^j\right)}{d\theta^j}\frac{d\tilde{\Pi}^{j*}\left(\theta^j\right)}{d\theta^j}\right]$$
$$+ \tilde{\Pi}^{j*}\left(\theta^j\right)\left[A\left(\theta^j\right)\right]^{-3}\left[\frac{d\tilde{C}\left(\tilde{C}^{-1}\left(\left[A\left(\theta^j\right)\right]^{-1}\tilde{I}^{\theta,j}\right)\right)}{d\Delta\theta^j}\right]^{-1}$$
$$\left[2\left(\frac{dA\left(\theta^j\right)}{d\theta^j}\right)^2 - A\left(\theta^j\right)\frac{d^2A\left(\theta^j\right)}{d\left(\theta^j\right)^2}\right]$$
$$+ \left[A\left(\theta^j\right)\right]^{-4}\left[\frac{d\tilde{C}\left(\tilde{C}^{-1}\left(\left[A\left(\theta^j\right)\right]^{-1}\tilde{I}^{\theta,j}\right)\right)}{d\Delta\theta^j}\right]^{-3}\frac{d^2\tilde{C}\left(\tilde{C}^{-1}\left(\left[A\left(\theta^j\right)\right]^{-1}\tilde{I}^{\theta,j}\right)\right)}{d\left(\Delta\theta^j\right)^2}$$
$$\left[2\tilde{\Pi}^{j*}\left(\theta^j\right)\frac{d\tilde{\Pi}^{j*}\left(\theta^j\right)}{d\theta^j}A\left(\theta^j\right)\frac{dA\left(\theta^j\right)}{d\theta^j}\right.$$
$$\left. - \left(\tilde{\Pi}^{j*}\left(\theta^j\right)\frac{dA\left(\theta^j\right)}{d\theta^j}\right)^2 - \left(\frac{d\tilde{\Pi}^{j*}\left(\theta^j\right)}{d\theta^j}A\left(\theta^j\right)\right)^2\right]$$

The terms outside the brackets in all three summands are strictly positive and the term in brackets in the last summand is strictly negative. A sufficient (though stricter than necessary) condition for concavity of the upper bound function thus is for both bracketed terms in the first two summands to be non-positive. For the first summand, applying (2.B.36), non-positivity of the term in brackets is equivalent to

$$-\theta^j\left[A\left(\theta^j\right)\right]^{-1}\frac{dA\left(\theta^j\right)}{d\theta^j} \leq -\frac{1}{2}\theta^j\left(\frac{d\tilde{\Pi}^{j*}\left(\theta^j\right)}{d\theta^j}\right)^{-1}\frac{d^2\tilde{\Pi}^{j*}\left(\theta^j\right)}{d\left(\theta^j\right)^2}$$

With

$$-\theta^j\left(\frac{d\tilde{\Pi}^{j*}\left(\theta^j\right)}{d\theta^j}\right)^{-1}\frac{d^2\tilde{\Pi}^{j*}\left(\theta^j\right)}{d\left(\theta^j\right)^2} = (1-\nu)m\left(\theta^j\right) + \frac{1}{\sigma_j}\left[1 - m\left(\theta^j\right)\right]$$

and as $m^j\left(\theta^j\right) \in [0,1]$, Assumption (P3) ensures that this is satisfied for all $\theta^j \in \Theta$. For the second summand, nonnegativity of the term in brackets is equivalent to

$$\left[A\left(\theta^j\right)\right]^{-1}\left(\frac{dA\left(\theta^j\right)}{d\theta^j}\right)^2\left[\frac{d^2A\left(\theta^j\right)}{d\left(\theta^j\right)^2}\right]^{-1} \leq \frac{1}{2}$$

which again holds by Assumption (P3). Thus, the upper bound function $f^j\left(\theta^j\right)$ is concave. The lower bound of the constraint correspondence is concave as it is linear.

In summary, the reward function $w^j(\cdot)$ is continuous, bounded, continuously differentiable on the interior of X^j, increasing in θ^j and concave in $(\tilde{I}^{\theta,j}, \theta^j)$. The constraint correspondence $\Gamma^j\left(\theta^j\right)$ is non-empty, compact-valued (by boundedness of $\tilde{F}(\cdot)$ over closed sets), continuous, and convex (by concavity of the boundary functions). The discount factor is in $(0,1)$. Thus, by Theorems 4.8 and 4.11 in Stokey, Lucas, and Prescott (1989, pages 81, 85), a unique solution to (2.B.31) exists, is strictly concave, and is continuously differentiable at all interior points. Moreover, the optimal policy correspondence $\theta^{j*} : \Theta \to \Theta$ is a single-valued and continuous function.

To show that the value function is monotone, I consider a variant of the optimization problem. Suppose, for now, that the monopolists can choose to destroy technology at no cost. The cost of creating new technology remains unchanged, yet only occurs if there is a positive increment in technology. This variant can be expressed as

$$\begin{aligned} \tilde{V}^j\left(\theta^j\right) &= \max_{\theta^{j\prime}\in\tilde{\Gamma}^j\left(\theta^j\right)} \left\{\tilde{w}^j\left(\theta^j,\theta^{j\prime}\right) + \frac{1}{1+\rho}\tilde{V}^j\left(\theta^{j\prime}\right)\right\} \\ \text{s.t. } \tilde{w}^j\left(\theta^j,\theta^{j\prime}\right) &= \tilde{\Pi}^{j*}\left(\theta^j\right) - A\left(\theta^j\right)\tilde{C}\left(\max\left\{0, \theta^{j\prime} - \left(1-\delta^\theta\right)\theta^j\right\}\right) \\ \tilde{\Gamma}^j\left(\theta^j\right) &= \left[0, \left(1-\delta^\theta\right)\theta^j + \tilde{F}\left(\tilde{\Pi}^{j*}\left(\theta^j\right),\theta^j\right)\right] \end{aligned} \tag{2.B.40}$$

By the same reasoning as before, $\tilde{\Gamma}^j$ is non-empty, compact-valued, and continuous, $\tilde{w}(\cdot)$ is bounded above on $\tilde{X}^j = \left\{\left(\theta^j,\theta^{j\prime}\right)\in\Theta\times\Theta : \theta^{j\prime}\in\tilde{\Gamma}^j\left(\theta^j\right)\right\}$. Moreover, as $\tilde{\Pi}^{j*}(\cdot)$ is increasing in θ^j and $\tilde{F}(\cdot)$ is increasing in both arguments, $\theta_1^j \le \theta_2^j$ implies

$$0 \le \theta_1^{j\prime} \le \left(1-\delta^\theta\right)\theta_1^j + \tilde{F}\left(\tilde{\Pi}^{j*}\left(\theta_1^j\right),\theta_1^j\right) \le \left(1-\delta^\theta\right)\theta_2^j + \tilde{F}\left(\tilde{\Pi}^{j*}\left(\theta_2^j\right),\theta_2^j\right)$$

That is, $\theta_1^j \le \theta_2^j$ implies $\tilde{\Gamma}^j(\theta_1^j) \subseteq \tilde{\Gamma}^j(\theta_2^j)$, i.e. $\tilde{\Gamma}(\cdot)$ is monotone. Next, Taking derivatives of $\tilde{w}^j$ w.r.t. θ^j yields

$$\begin{aligned} \frac{\partial \tilde{w}^j\left(\theta^j,\theta^{j\prime}\right)}{\partial\theta^j} &= \frac{d\tilde{\Pi}^{j*}\left(\theta^j\right)}{d\theta^j} \\ &\quad + \mathbb{1}_{\{\theta^{j\prime}>(1-\delta^\theta)\theta^j\}}\left[\left(1-\delta^\theta\right)A\left(\theta^j\right)\frac{d\tilde{C}\left(\Delta\theta^j\right)}{d\Delta\theta^j} - \frac{dA\left(\theta^j\right)}{d\theta^j}\tilde{C}\left(\Delta\theta^j\right)\right] \\ &> 0 \end{aligned}$$

which follows from $\tilde{\Pi}^{j*}(\cdot)$ strictly increasing and the properties of $A(\cdot)$ and $\tilde{C}(\cdot)$ from assumptions (P2)–(P3). In other words, $\tilde{w}^j$ is strictly increasing in θ^j.[18] By

18. $\tilde{w}^j$ is not differentiable at $\theta^{j\prime} = \left(1-\delta^\theta\right)\theta^j$; however, as $\frac{d\tilde{\Pi}^{j*}\left(\theta^j\right)}{d\theta^j}$ is strictly increasing, $\tilde{w}^j$ is still strictly increasing in θ^j for all $\left(\theta^j,\theta^{j\prime}\right)\in\tilde{X}^j$.

Theorem 4.7 in Stokey et al. (1989, page 80), there exists a unique solution to (2.B.40) that is strictly increasing in θ^j.

Turning back to our original problem, note that while the monopolists can choose to destroy technology at no cost, there is also no benefit to it. More formally, for any given $\theta^j \in \Theta$: Any choice of $\theta^{j\prime}$ that is attainable under $\tilde{\Gamma}^j$, but not under Γ^j, is strictly smaller than $\left(1-\delta^{\theta}\right)\theta^j$ Note that any choice attainable under Γ^j is also attainable under $\tilde{\Gamma}^j$, i.e. $\Gamma^j\left(\theta^j\right) \subseteq \tilde{\Gamma}^j\left(\theta^j\right)$. For all choices that are attainable under both constraint correspondences, the reward function is identical:

$$\tilde{w}^j\left(\theta^j,\theta^{j\prime}\right) = w^j\left(\theta^j,\theta^{j\prime}\right) \quad \forall \theta^{j\prime} \in \tilde{\Gamma}^j\left(\theta^j\right) \cap \Gamma^j\left(\theta^j\right) = \Gamma^j\left(\theta^j\right)$$

Thus,

$$\begin{aligned}
\tilde{V}^j\left(\theta^j\right) &= \max_{\theta^{j\prime}\in\tilde{\Gamma}^j(\theta^j)} \left\{\tilde{w}^j\left(\theta^j,\theta^{j\prime}\right) + \frac{1}{1+\rho}\tilde{V}^j\left(\theta^{j\prime}\right)\right\} \\
&= \max\left\{ \max_{\theta^{j\prime}\in\Gamma^j(\theta^j)} \left\{\tilde{w}^j\left(\theta^j,\theta^{j\prime}\right) + \frac{1}{1+\rho}\tilde{V}^j\left(\theta^{j\prime}\right)\right\}, \right. \\
&\qquad \left. \max_{\theta^{j\prime}\in\tilde{\Gamma}^j(\theta^j)-\Gamma^j(\theta^j)} \left\{\tilde{w}^j\left(\theta^j,\theta^{j\prime}\right) + \frac{1}{1+\rho}\tilde{V}^j\left(\theta^{j\prime}\right)\right\} \right\} \\
&= \max_{\theta^{j\prime}\in\Gamma^j(\theta^j)} \left\{\tilde{w}^j\left(\theta^j,\theta^{j\prime}\right) + \frac{1}{1+\rho}\tilde{V}^j\left(\theta^{j\prime}\right)\right\} \\
&= \max_{\theta^{j\prime}\in\Gamma^j(\theta^j)} \left\{w^j\left(\theta^j,\theta^{j\prime}\right) + \frac{1}{1+\rho}\tilde{V}^j\left(\theta^{j\prime}\right)\right\}
\end{aligned}$$

The expression in the last line is identical to the definition of $V^j(\cdot)$, i.e. $V^j\left(\theta^j\right) = \tilde{V}^j\left(\theta^j\right)\ \forall \theta^j \in \Theta$. Thus, the solution to (2.B.31) is strictly increasing.

Finally, using the results derived above, the problem can be restated in its original terms as

$$V^j\left(\theta^j\right) = \max_{\tilde{I}^{\theta,j}\in\left[0,\tilde{\Pi}^{j*}(\theta^j)\right]} \left\{\tilde{\Pi}^{j*}\left(\theta^j\right) - \tilde{I}^{\theta,j} + \frac{1}{1+\rho}V^j\left(\left(1-\delta^{\theta}\right)\theta^j + \tilde{F}\left(\tilde{I}^{\theta,j},\theta^j\right)\right)\right\} \tag{2.B.41}$$

Differentiability of the value function allows for defining the investment policy function by the first-order condition. Taking derivatives of (2.B.41) w.r.t. $\tilde{I}^{\theta,j}$ yields

$$\frac{1}{1+\rho}\frac{\partial \tilde{F}\left(\tilde{I}^{\theta,j*}\left(\theta^j\right),\theta^j\right)}{\tilde{I}^{\theta,j}}\left(\left.\frac{dV^j(z)}{dz}\right|_{z=\left(1-\delta^{\theta}\right)\theta^j+\tilde{F}\left(\tilde{I}^{\theta,j*}(\theta^j),\theta^j\right)}\right) = 1 \tag{2.B.42}$$

Finally, with

$$\tilde{I}^{\theta,j*}\left(\theta^j\right) = A\left(\theta^j\right)\tilde{C}\left(\theta^{j*}\left(\theta^j\right) - \left(1-\delta^{\theta}\right)\theta^j\right)$$

where $\theta^{j*}(\theta^j)$ is the technology policy function, continuity of the investment policy function follows from continuity of the technology policy function. ∎

Power cost functions. Suppose for now that the both cost functions are power functions, i.e.

$$A\left(\theta^j\right) = \left(\theta^j + 1\right)^{-\varsigma_e} \quad \text{and} \quad \tilde{C}\left(\Delta\theta^j\right) = \iota\left(\Delta\theta^j\right)^d$$

This implies that

$$\begin{aligned}
-\left[A\left(\theta^j\right)\right]^{-1}\frac{dA\left(\theta^j\right)}{d\theta^j} &= \frac{\varsigma_e}{\theta^j + 1}\\
-\theta^j\left[A\left(\theta^j\right)\right]^{-1}\frac{dA\left(\theta^j\right)}{d\theta^j} &= \frac{\varsigma_e\theta^j}{\theta^j + 1}\\
\left[A\left(\theta^j\right)\right]^{-1}\left[\frac{dA\left(\theta^j\right)}{d\theta^j}\right]^2\left[\frac{d^2A\left(\theta^j\right)}{d\left(\theta^j\right)^2}\right]^{-1} &= \frac{\varsigma_e}{\varsigma_e + 1}\\
\tilde{C}\left(\Delta\theta^j\right)\left[\frac{d\tilde{C}\left(\Delta\theta^j\right)}{d\Delta\theta^j}\right]^{-2}\frac{d^2\tilde{C}\left(\Delta\theta^j\right)}{d\left(\Delta\theta^j\right)^2} &= \frac{\varsigma_d}{\varsigma_d - 1}
\end{aligned}$$

The conditions of assumptions (P2)–(P4) then are

$$\begin{aligned}
\frac{\varsigma_e}{\theta^j + 1} \leq 1 \ \forall\theta^j \in \Theta \quad &\Leftrightarrow \quad \varsigma_e \leq 1\\
\frac{\varsigma_e\theta^j}{\theta^j + 1} \leq \min\left\{\frac{1-\nu}{2}; \frac{1}{2\sigma_j}\right\} \ \forall\theta^j \in \Theta \quad &\Leftrightarrow \quad \varsigma_e \leq \frac{1-\nu}{2}\\
\frac{\varsigma_e}{\varsigma_e + 1} \leq \frac{1}{2} \quad &\Leftrightarrow \quad \varsigma_e \leq 1\\
\frac{\varsigma_e}{\varsigma_e + 1}\frac{\varsigma_d - 1}{\varsigma_d} \leq 1 \quad &\Leftrightarrow \quad \varsigma_d \geq \varsigma_e + 1
\end{aligned}$$

Note that $\nu \in [0, 1]$ implies $\frac{1-\nu}{2} \in \left[0, \frac{1}{2}\right]$. Hence, $\varsigma_e \leq \frac{1-\nu}{2}$ and $\varsigma_d \geq \varsigma_e + 1$ are sufficient conditions for (P2)–(P4) to hold.

Proof of Corollary 9. Substituting the investment policy function into (2.39) and taking derivatives w.r.t. θ^j yields

$$\frac{dV^j\left(\theta^j\right)}{d\theta^j} = \frac{d\tilde{\Pi}^{j*}\left(\theta^j\right)}{d\theta^j} + \frac{1}{1+n}\left[\left(1-\delta^\theta\right) + \frac{\partial\tilde{F}\left(\tilde{I}^{\theta,j*}\left(\theta^j\right), \theta^j\right)}{\partial\theta^j}\right]\left(\left.\frac{dV^j(z)}{dz}\right|_{z=\theta^{j*}(\theta^j)}\right) \tag{2.B.43}$$

Replacing the investment policy function with the technology policy function, the first-order condition (2.B.42) becomes

$$\frac{1}{1+\rho}\frac{\partial \tilde{F}\left(\tilde{I}^{\theta,j*}\left(\theta^j\right),\theta^j\right)}{\tilde{I}^{\theta,j}}\left(\left.\frac{dV^j(z)}{dz}\right|_{z=\theta^{j*}\left(\theta^j\right)}\right) = 1$$

Using the fact that the above condition equals one, equation (2.B.43) can be multiplied by this expression and rearranged to obtain

$$\frac{\frac{dV^j\left(\theta^j\right)}{d\theta^j}}{\left(\left.\frac{dV^j(z)}{dz}\right|_{z=\theta^{j*}\left(\theta^j\right)}\right)} = \frac{1}{1+\rho}\left[\frac{d\tilde{\Pi}^{j*}\left(\theta^j\right)}{d\theta^j}\frac{\partial \tilde{F}\left(\tilde{I}^{\theta,j*}\left(\theta^j\right),\theta^j\right)}{\partial \tilde{I}^{\theta,j}} + \left(1-\delta^\theta\right) + \frac{\partial \tilde{F}\left(\tilde{I}^{\theta,j*}\left(\theta^j\right),\theta^j\right)}{\partial \theta^j}\right] \tag{2.B.44}$$

To assess the behavior of this expression as θ^j approaches zero, I consider the limits of the marginal products of $\tilde{F}(\cdot)$ and of the monopolist's marginal profit. First, recall that investment is bounded by net profits. If the technology level equals zero, then net profits are zero, which means investment must also be zero, i.e. $\tilde{I}^{\theta,j*}(0) \leq \tilde{\Pi}^{j*}(0) = 0$. The definition of $\tilde{F}(\cdot)$, together with the properties of $A(\cdot)$ and $\tilde{C}(\cdot)$, then imply

$$\begin{aligned}
\lim_{\theta^j\to 0}\frac{\partial \tilde{F}\left(\tilde{I}^{\theta,j*}\left(\theta^j\right),\theta^j\right)}{\tilde{I}^{\theta,j}} &= \left[A(0)\frac{dC\left([A(0)]^{-1}\tilde{I}^{\theta,j*}(0)\right)}{d\Delta\theta^j}\right]^{-1} \\
&= \left[A(0)\frac{dC(0)}{d\Delta\theta^j}\right]^{-1} \\
&> 0 \\
\lim_{\theta^j\to 0}\frac{\partial \tilde{F}\left(\tilde{I}^{\theta,j*}\left(\theta^j\right),\theta^j\right)}{\theta^j} &= -\tilde{I}^{\theta,j*}(0)\,[A(0)]^{-2}\frac{dA(0)}{d\theta^j}\left[\frac{dC\left([A(0)]^{-1}\tilde{I}^{\theta,j*}(0)\right)}{d\Delta\theta^j}\right]^{-1} \\
&= -0\,[A(0)]^{-2}\frac{dA(0)}{d\theta^j}\left[\frac{dC(0)}{d\Delta\theta^j}\right]^{-1} \\
&= 0
\end{aligned}$$

The limit of the marginal profit of the monopolist as θ^j approaches zero then is

$$\begin{aligned}
\lim_{\theta^j\to 0}\frac{d\tilde{\Pi}^{j*}\left(\theta^j\right)}{d\theta^j} &= \lim_{\theta^j\to 0}\nu\tilde{\Pi}^{j*}\left(\theta^j\right)m^j\left(\theta^j\right)\left(\theta^j\right)^{-1} \\
&= \lim_{\theta^j\to 0}\nu(1-\alpha)\pi^{G^j}G^j\left(\tilde{H}^j,\tilde{M}^j,\tilde{K}^j,\theta^j\right)m^j\left(\theta^j\right)\left(\theta^j\right)^{-1} \\
&= \nu(1-\alpha)\pi^{G^j}\left(\tilde{K}^j\right)^{1-\nu} \\
&\quad \lim_{\theta^j\to 0}\left\{\left[(1-\eta)(\tilde{M}^j\theta^j)^{\frac{\sigma_j-1}{\sigma_j}}+\eta(\tilde{H}^j)^{\frac{\sigma_j-1}{\sigma_j}}\right]^{\frac{\sigma_j\nu}{\sigma_j-1}}\right. \\
&\quad \left.\frac{(1-\eta)(\theta^j\tilde{M}^j)^{\frac{\sigma_j-1}{\sigma_j}}}{(1-\eta)(\theta^j\tilde{M}^j)^{\frac{\sigma_j-1}{\sigma_j}}+\eta(\tilde{H}^j)^{\frac{\sigma_j-1}{\sigma_j}}}\left(\theta^j\right)^{-1}\right\} \\
&= \nu(1-\alpha)\pi^{G^j}(\tilde{K}^j)^{1-\nu}(1-\eta)(\tilde{M}^j)^{\frac{\sigma_j-1}{\sigma_j}} \\
&\quad \lim_{\theta^j\to 0}\left\{\left[(1-\eta)(\tilde{M}^j)^{\frac{\sigma_j-1}{\sigma_j}}+\eta(\tilde{H}^j)^{\frac{\sigma_j-1}{\sigma_j}}\left(\theta^j\right)^{\frac{1-\sigma_j}{\sigma_j}}\right]^{\frac{\sigma_j\nu}{\sigma_j-1}-1}\right. \\
&\quad \left.\left(\theta^j\right)^{\nu-1}\right\} \\
&= \nu(1-\alpha)\pi^{G^j}\left(\tilde{K}^j\right)^{1-\nu}(1-\eta)^{\frac{\sigma_j\nu}{\sigma_j-1}}\left(\tilde{M}^j\right)^{\nu}\lim_{\theta^j\to 0}\left(\theta^j\right)^{\nu-1}
\end{aligned}$$

Thus, the marginal profit of the monopolist tends to positive infinity as θ^j approaches zero. Consequently,

$$\begin{aligned}
\lim_{\theta^j\to 0}\left(\frac{\frac{dV^j\left(\theta^j\right)}{d\theta^j}}{\left(\left.\frac{dV^j(z)}{dz}\right|_{z=\theta^{j*}\left(\theta^j\right)}\right)}\right) &= \frac{1-\delta^\theta}{1+\rho} \\
&\quad +\left[(1+\rho)A(0)\left(\left.\frac{d\tilde{C}(z)}{dz}\right|_{z=0}\right)\right]^{-1}\lim_{\theta^j\to 0}\frac{d\tilde{\Pi}^{j*}\left(\theta^j\right)}{d\theta^j}
\end{aligned}$$

i.e. the ratio of the derivative of the value function today and the derivative of the value function next period under the optimal policy tends to infinity as θ^j approaches zero. By strict concavity of the value function, $\frac{dV^j\left(\theta^j\right)}{d\theta^j}$ is strictly decreasing in θ^j. This implies that there exist values of θ^j that satisfy

$$\frac{\frac{dV^j\left(\theta^j\right)}{d\theta^j}}{\left(\left.\frac{dV^j(z)}{dz}\right|_{z=\theta^{j*}\left(\theta^j\right)}\right)} > 1 \quad \Leftrightarrow \quad \theta^{j*}\left(\theta^j\right) > \theta^j$$

Next, let $\bar{\tilde{\Pi}}^{j*}$ denote the upper bound of net per-capita profits over all possible technology levels. Bounded profits imply that investment, and therefore the creation of new technology, is bounded as well:

$$\tilde{F}\left(\tilde{I}^{\theta,j*}\left(\theta^j\right),\theta^j\right) \;\leq\; \tilde{F}\left(\tilde{\Pi}^{j*}\left(\theta^j\right),\theta^j\right) \;\leq\; \tilde{F}\left(\bar{\tilde{\Pi}},\theta^j\right) \quad \forall\theta^j \;\in\; \Theta$$

Now, consider again the law of motion for technology and rearrange to obtain the relative difference in technology levels between the current and the next period:

$$\frac{\theta^{j*}\left(\theta^j\right)-\theta^j}{\theta^j} \;=\; -\delta^\theta + \frac{\tilde{F}\left(\tilde{I}^{\theta,j*}\left(\theta^j\right),\theta^j\right)}{\theta^j} \;\leq\; -\delta^\theta + \frac{\tilde{F}\left(\bar{\tilde{\Pi}}^{j*},\theta^j\right)}{\theta^j}$$

Thus, if there exist technology levels θ^j such that the creation rate of new technology when investing the supremum of net profits over all technology levels falls short of the depreciation rate, the change in technology levels is negative, i.e. $\theta^{j*}\left(\theta^j\right) > \theta^j$. Now, consider the limit of the technology creation rate as θ^j approaches infinity:

$$\begin{aligned}
\lim_{\theta^j\to\infty}\frac{\tilde{F}\left(\bar{\tilde{\Pi}}^{j*},\theta^j\right)}{\theta^j} &= \lim_{\theta^j\to\infty}\frac{\partial\tilde{F}\left(\bar{\tilde{\Pi}}^{j*},\theta^j\right)}{\partial\theta^j}\\
&= -\bar{\tilde{\Pi}}^{j*}\lim_{\theta^j\to\infty}\left[A\left(\theta^j\right)\right]^{-2}\frac{dA\left(\theta^j\right)}{d\theta^j}\left[\left(\left.\frac{\tilde{C}(z)}{dz}\right|_{z=\left[A\left(\theta^j\right)\right]^{-1}\bar{\tilde{\Pi}}^{j*}}\right)\right]^{-1}\\
&= 0
\end{aligned}$$

where the first equality follows from l'Hospital's rule, the second from the definition of $\tilde{F}(\cdot)$ and the third from Assumption (P4). In other words, the convexity of $\tilde{C}(\cdot)$ is sufficient to outweigh the convexity of $A(\cdot)$ and the marginal product of $\tilde{F}(\cdot)$ w.r.t. technology is diminishing. Now, as the limit is zero, there exists some $\bar{\theta}^j$ such that the creation rate of new technology when investing the entire net profits falls short of the depreciation rate for all technology levels larger than $\bar{\theta}^j$:

$$\exists\bar{\theta}^j \text{ s. th. } \frac{\tilde{F}\left(\bar{\tilde{\Pi}}^{j*},\theta^j\right)}{\theta^j} \;<\; \delta^\theta \quad\Leftrightarrow\quad \theta^{j*}\left(\theta^j\right) \;<\; \theta^j \quad \forall\theta^j \;>\; \bar{\theta}^j$$

Thus, there exist technology levels such that $\theta^{j*}\left(\theta^j\right) > \theta^j$ and technology levels such that $\theta^{j*}\left(\theta^j\right) < \theta^j$. By continuity of the policy function, there must exist at least one technology level θ^{j*} such that $\theta^{j*}\left(\theta^j\right) = \theta^j$. This technology level is characterized by

$$\begin{aligned}
1 &= \frac{\frac{dV^j\left(\theta^{j*}\right)}{d\theta^j}}{\left(\left.\frac{dV^j(z)}{dz}\right|_{z=\theta^{j*}\left(\theta^{j*}\right)}\right)}\\
&= \frac{1}{1+\rho}\left[\frac{d\tilde{\Pi}^{j*}\left(\theta^{j*}\right)}{d\theta^j}\frac{\partial\tilde{F}\left(\tilde{I}^{\theta,j*}(\theta^{j*}),\theta^{j*}\right)}{\partial\tilde{I}^{\theta,j}} + (1-\delta^\theta) + \frac{\partial\tilde{F}\left(\tilde{I}^{\theta,j*}(\theta^{j*}),\theta^{j*}\right)}{\partial\theta^j}\right]
\end{aligned}$$

Rearranging produces

$$\frac{d\tilde{\Pi}^{j*}\left(\theta^{j*}\right)}{d\theta^j} = \left[\frac{\partial \tilde{F}\left(\tilde{I}^{\theta,j*}\left(\theta^{j*}\right),\theta^{j*}\right)}{\partial \tilde{I}^{\theta,j}}\right]^{-1}\left[\left(\rho+\delta^\theta\right) - \frac{\partial \tilde{F}\left(\tilde{I}^{\theta,j*}\left(\theta^{j*}\right),\theta^{j*}\right)}{\partial \theta^j}\right]$$

Note that, by definition

$$\tilde{I}^{\theta,j*}\left(\theta^{j*}\right) = A\left(\theta^{j*}\right)\tilde{C}\left(\theta^{j*}\left(\theta^{j*}\right) - \left(1-\delta^\theta\right)\theta^{j*}\right) = A\left(\theta^{j*}\right)\tilde{C}\left(\delta^\theta\theta^{j*}\right)$$

which implies

$$\begin{aligned}\tilde{C}^{-1}\left(\left[A\left(\theta^{j*}\right)\right]^{-1}\tilde{I}^{\theta,j*}\left(\theta^{j*}\right)\right) &= \tilde{C}^{-1}\left(\left[A\left(\theta^{j*}\right)\right]^{-1}A\left(\theta^{j*}\right)\tilde{C}\left(\delta^\theta\theta^{j*}\right)\right)\\ &= \tilde{C}^{-1}\left(\tilde{C}(\delta^\theta\theta^{j*})\right)\\ &= \delta^\theta\theta^{j*}\end{aligned}$$

Finally, substituting the expressions for the partial derivatives of $\tilde{F}(\cdot)$ and rearranging results in

$$\frac{d\tilde{\Pi}^{j*}\left(\theta^{j*}\right)}{d\theta^j} = \left(\rho+\delta^\theta\right)A\left(\theta^{j*}\right)\frac{d\tilde{C}(\delta^\theta\theta^{j*})}{d\Delta\theta^j} + \frac{dA\left(\theta^{j*}\right)}{d\theta^j}\tilde{C}\left(\delta^\theta\theta^{j*}\right)$$

which is the expression from the Corollary. ■

Uniqueness of stationary technology levels under power cost functions. Suppose again that direct and indirect cost of producing new patents are power cost functions. The condition for the stationary technology level then is

$$\frac{d\tilde{\Pi}^{j*}\left(\theta^{j*}\right)}{d\theta^j} = \left(\rho+\delta^\theta\right)\iota\varsigma_d\left(\delta^\theta\right)^{\varsigma_d-1}\frac{\left(\theta^{j*}\right)^{\varsigma_d-1}}{\left(\theta^{j*}+1\right)^{\varsigma_e}} - \varsigma_e\iota\left(\delta^\theta\right)^{\varsigma_d}\frac{\left(\theta^{j*}\right)^{\varsigma_d}}{\left(\theta^{j*}+1\right)^{\varsigma_e+1}} \tag{2.B.45}$$

and the stationary technology level is unique under the restrictions on the exponents derived above. To see this, consider again the derivative of the RHS of (2.B.45):

$$\begin{aligned}\frac{dRHS\left(\theta^j\right)}{d\theta^j} &= \iota\left(\delta^\theta\right)^{\varsigma_d-1}\frac{\left(\theta^j\right)^{\varsigma_d-2}}{\left(\theta^j+1\right)^{\varsigma_e+2}}\Big\{\left(\rho+\delta^\theta\right)\varsigma_d\left[\left(\varsigma_d-\varsigma_e-1\right)\theta^j+\left(\varsigma_d-1\right)\right]\\ &\qquad - \varsigma_e\delta^\theta\frac{\theta^j}{\theta^j+1}\left[\left(\varsigma_d-\varsigma_e-1\right)\theta^j+\varsigma_d\right]\Big\}\\ &= \iota\left(\delta^\theta\right)^{\varsigma_d-1}\frac{\left(\theta^j\right)^{\varsigma_d-2}}{\left(\theta^j+1\right)^{\varsigma_e+2}}\Big\{\left(\varsigma_d-\varsigma_e-1\right)\left[\rho+\delta^\theta\left(\varsigma_d-\frac{\theta^j}{\theta^j+1}\varsigma_e\right)\right]\theta^j\\ &\qquad + \varsigma_d\left[\rho\left(\varsigma_d-1\right)+\delta^\theta\left(\varsigma_d-\frac{\theta^j}{\theta^j+1}\varsigma_e-1\right)\right]\Big\}\end{aligned}$$

Recall that $\iota, \rho, \delta^\theta \geq 0$ and $\theta^j \in \Theta = \mathbb{R}_0^+$, implying that $\frac{\theta^j}{\theta^j+1} \in [0, 1)$. Together with the restriction $\varsigma_d \geq \varsigma_e + 1$, this ensures that all terms in the derivative are non-negative, i.e. the RHS is non-decreasing.

Chapter 3

Labor Market Mismatch

3.1 Introduction

Neither vacant jobs nor human capital are perfectly mobile. Most jobs are executed with installed physical capital (machines, offices, etc.) and cannot easily be transferred to another location. Moving a worker also takes time and resources. A given task in the production process usually cannot be replaced by another task one-for-one and requires a specific set of skills. While skills can often be complemented with adjacent ones, this typically comes at the cost of reduced efficiency. These frictions give rise to the possibility that human capital, or labor supply, and vacant jobs, or labor demand, are not perfectly aligned in terms of location and in terms of the content of work. This misalignment of vacant positions and idle workers will be referred to in this essay as labor market mismatch. Mismatch is economically meaningful in the sense that it represents a market inefficiency, making it a legitimate target for policymakers. A prerequisite for designing effective policies to reduce labor market mismatch is to understand its extent and implications.

It is not clear up front whether to think of this as vacancies being imperfectly allocated to searchers or searchers being imperfectly allocated to vacancies. Both distributions are highly endogenous objects, subject to factors such as moving cost, local preferences, but also path dependencies and interdependencies in location choice. In this essay, no stance is taken on what determines these distributions. Instead, I take the distribution of vacancies and the distribution of searching workers as given and attempt to quantify the magnitude of labor market mismatch in Germany. Since the underlying drivers of the distributions are abstracted from, this essay cannot offer a welfare analysis. What it can offer is to quantify potential effects on output and employment from removing barriers for job searchers to move between sectoral labor markets.

For this, it is necessary to account for differences across sectors in some key aspects. The theory on frictional labor markets posits that searching workers and open positions are complements in the creation of new jobs. That is, the ratio of vacancies to searching workers, or labor market tightness, is relevant to determine how many jobs are created. Moreover, sectoral labor markets might differ in their matching efficiency, i.e. in how efficiently they transform vacancies and searchers into new jobs. Once the jobs are created, there might be differences in how long they exist before being destroyed, and in how productive they are in the meantime. In other words, the job loss rate (and, hence, the expected duration of a newly created match) and the labor productivity matter in determining how valuable a newly created job is from the perspective of output and employment. Accounting for all of these factors requires rich data and a quantitative model.

Building on historical reports published by the German Federal Employment Agency (BA) and the German Sample of Integrated Labour Market Biographies (SIAB), I construct a dataset of vacant positions and searching workers spanning almost four decades. The long observation period contains the German reunification and several episodes of economic expansion and recession, offering unique opportunities for analysis. The data contains information on the geographic location, as well as the occupations of searchers and vacancies. This allows for analysing potential mismatch across two distinct dimensions of labor market segmentation. This novel dataset is then combined with a model of frictional labor markets. The model builds on the work of Şahin, Song, Topa, and Violante (2014). I extend their framework by deriving an explicit solution of the planner allocation of searching workers to sectoral labor markets. This allows me to not only measure the extent of mismatch present in the German labor market, but also to quantify the distributional consequences of reducing it – specifically, how many workers would need to be reallocated across regions and occupations. This, in turn, allows for a quantification not only of the benefits, but also of the costs involved in affecting the extent of labor market mismatch in Germany.

The essay proceeds as follows: Section 3.2 provides a brief overview of the related literature and the contribution of this essay to the literature. Section 3.3 presents the data and empirical results. The model framework, the derivation of the planner allocation and the measures used to quantify labor market mismatch are described in Section 3.4. Section 3.5 discusses the results. Finally, Section 3.6 concludes.

3.2 Related Literature and Contribution

This essay adds to the literature on spatial unemployment differences. Both the causes and consequences of persistent spatial labor market differences are a field of active research. Recently, Bilal (2021) and Kuhn, Manovskii, and Qiu (2021) have found that job separation rates are the key drivers of spatial labor market differences. Bilal (2021) has shown that these differences lead to agglomeration and congestion externalities and that place-based policies incentivizing moves to high-separation regions improve welfare. Heise and Porzio (2022) study how frictions impeding labor mobility across space affect the joint allocation of labor across firms and regions. They find that spatial labor market frictions reduce aggregate productivity by shielding low productivity firms from competition from other regions. Şahin et al. (2014) take no stance on the causes of spatial labor market differences and assess the aggregate unemployment effects of mismatch between unemployed workers and vacant jobs across U.S. regions. They build a multi-sector model of frictional labor markets and find that mismatch contributed to the increase in unemployment after the Great Recession

The essay is also connected to the broader literature on worker mobility. In this literature, occupational mobility is often interpreted as mobility between jobs over the worker life cycle (job ladder mechanisms). Examples are Moscarini and Postel-Vinay (2012) and Haltiwanger, Hyatt, Kahn, and McEntarfer (2018), who both document the interaction of these mechanisms with the business cycle. The study by Dvorkin and Monge-Naranjo (2019) provides insights into the aggregate consequences of worker mobility across (occupational) sectors of the economy. Recently, Amior (2024) provides an analysis of regional worker mobility and finds that aggregate mobility has declined substantially across all demographic groups over the past three decades. The analysis presented here adds to the literature on the implications of aggregate worker mobility across occupations and across regions.

The quantitative analysis in this essay is based on the theory of frictional labor markets in the tradition of Diamond (1982), Mortensen (1982), and Pissarides (1985). While this theoretical framework has traditionally focused on aggregate outcomes, recent empirical work by Chetty, Hendren, Kline, and Saez (2014) using administrative data has revealed substantial heterogeneity in labor market dynamics across different worker types and geographic areas. This highlights the importance of distributional analysis in understanding labor market frictions. Methodologically, this work is most closely related to the work on mismatch unemployment by Şahin et al. (2014) on which it builds extensively.

My contribution to the literature is twofold: In terms of methodology, this study extends the analytical framework developed by Şahin et al. (2014) by incorporating distributional consequences into the core analysis. While the original framework focused primarily on measuring the extent of aggregate mismatch and its impact on unemployment, the refinement developed here explicitly accounts for the potential distributional implications of alleviating mismatch at the sectoral level. In other words, while the approach of Şahin et al. (2014) is informative on the forces driving aggregate unemployment (and thus the potential benefits of policies affecting these forces), this study focuses on quantifying the distributional consequences of any such policies (and thus the costs and trade-offs associated with them) as well. This methodological advancement addresses a critical gap in the existing literature, where distributional impacts have typically been treated as secondary considerations despite their importance for policy design. By explicitly modeling how mismatch reduction affects different sectors, regions, and occupations differently, this work helps complete the picture of labor market mismatch that Şahin et al. (2014) began to paint.

Empirically, the framework is applied to a comprehensive and novel dataset comprising long time series of German employment and vacancy data, disaggregated by both region and occupation. This rich dataset spans almost four decades (1982–2019) and provides unprecedented granularity for analyzing labor market dynamics in one of Europe's largest economies. The data allows for an examination of heterogeneity in causes and consequences of labor market mismatch, both in terms of regional differences and differences across occupations. Moreover, the long observation period – spanning multiple business cycles – offers novel insights into the cyclicality of heterogeneity across regions and occupations, revealing how geographic and occupational mismatch vary over time and in response to different aggregate economic conditions.

3.3 Empirical Evidence

Economic theory suggests that resources should be allocated where they are utilized most efficiently. Interpreting vacant positions and searching workers as resources for the creation of newly filled positions, the question I try to address here is where these resources are currently utilized and how efficiently. This question goes beyond maximizing the number of filled positions; efficient utilization implies the creation of jobs where it is easiest to do so, where jobs are most productive, and where jobs are most stable. I therefore examine four aspects of potential differences across sectoral labor markets: (1) the sectoral ratio of vacancies to searching workers (labor market tightness), measuring how vacancies and searching workers are currently allocated across sectors; (2) the efficiency of the sectoral matching process, which is informative on how vacancies and searching workers translate into filled positions; (3) the sectoral labor productivity, measuring how productive filled positions are; and (4) the sectoral job loss rate, which is informative on how long a job is expected to exist. I now briefly introduce data and methods used in the analysis before presenting the empirical results examining geographic and occupational differences in the German labor market.

3.3.1 Data Sources

Throughout the analysis, I employ two distinct segmentations of labor markets: by geography and by occupation. The geographic segmentation builds on the political structure of Germany (NUTS regions) and the spatial structure of the labor market (IAB labor market regions). The occupational segmentation employs the Classification of Occupations (KldB) of the German Statistical Office. Using long time series of German data introduces the structural break of the German reunification. This is obvious for geographically segmented data, as new observation units are added. But it also holds true for occupational data: structural differences in the economies of the former German Democratic Republic (GDR) and the Federal Republic of Germany (FRG) may result in different outcomes when data is pooled across geographies. To account for this, data on the former GDR (including Berlin) is included from 1994 on only.[1]

For the analysis, I combine a variety of data sources: Information on vacancies is obtained from several sources of the statistical office of the BA. The vacancy data contains all positions registered with the Federal Employment Agency. Importantly, open positions that are not registered with the BA are not included. Moreover, it is important to note that there has been a major change in the BA vacancy reporting by occupation during the observation period: Data prior to 2011 is coded according to KldB issue 1998, data from 2012 onward is coded according to KldB issue 2010. As observation units have changed, comparisons across issues are typically not feasible.[2]

For information on employment, unemployment, new jobs, and job losses, I use the SIAB Regional File (or SIAB-R), Version 7519 v1 (Frodermann, Ganzer, Schmucker, and Berge, 2021a). The dataset contains information on age, employment status, job location, occupation, and earnings. I use observations from the SIAB between January 1982 and December 2019.

Information on regional labor productivity is obtained from the "Volkswirtschaftliche Gesamtrechnung der Länder" (National Accounts of the Federal States of Germany, VGRdL). The VGRdL dataset contains yearly information on GDP, employment, and hours worked by state from 1982 to 2019.[3] Data by county is available from 2000 to 2019.

To ensure compatibility across data sources, the same harmonization as in Section 2.3 is applied.[4] Geographic data is reported for the 16 *federal states*, 40

---

1. It is impossible to define a precise moment in time at which the labor markets of the former GDR and the FRG have been fully integrated. Instead, integration of labor markets has been a process that is potentially still ongoing. In addition, data quality and consistency are unclear in the years immediately following the political reunification. In light of all of this, 1994 has been chosen as the first year of "sufficient" integration of labor markets and data infrastructure.

2. In fact, there has been another change in reporting: Before 2000, reported vacancies include regular vacancies (the so-called primary labor market) as well as subsidized vacancies (the so-called secondary labor market); after 2000, only vacancies in the primary labor market are included in the reporting. The change in the scope of the reporting (primary and secondary labor market vs. primary only) has relatively minor implications, however, as the main motivation for the change has been that the secondary labor market had become much smaller by 2000 than it had been historically.

3. Total hours worked only available from 2000 onward.

labor market regions, and 328 *counties*. The harmonized occupational segmentation features for KldB 1988 5 *occupation areas*, 29 *occupation main groups*, and 120 *occupation groups*[5], and for KldB 2010 10, 37, and 126 units, respectively. Combining the different data sources, I construct several datasets. Table 3.1 provides a complete overview of segmentations and observation periods.

Table 3.1. Data availability

Segmentation	Observation period
Geography	
Federal states	1982–2019**
Labor market regions	2016–2019
Counties	2016–2019
Occupation	
KldB 1988*	1982–2011***
KldB 2010*	2012–2019

Notes: Data availability by segmentation.
* Aggregation levels 1 (occupation areas), 2 (occupation main groups), and 3 (occupation groups).
** Data from former GDR included from 1994 onward; for 1998–1999 only yearly frequency.
*** Data from former GDR included from 1994 onward; for 1982–1999 only yearly frequency.

3.3.2 Methodology

The analysis aims at identifying differences across sectoral labor markets. In principle, aggregate data is not required if sectoral data can be observed directly. For clarity of presentation and in preparation of the quantitative exercise in Section 3.5, it is however convenient to normalize sector-specific data with the respective observed aggregate measures.

First, consider aggregate and relative sector-specific labor market tightness. Aggregate tightness is denoted by Θ_t and defined as the ratio of aggregate vacancies v_t to aggregate searchers s_t. Analogously, the observed sector-specific tightness is defined as by $\theta_{it} = \frac{v_{it}}{s_{it}}$. The relative sector-specific labor market tightness is then defined as

$$\tilde{\theta}_{it} \equiv \frac{\theta_{it}}{\Theta_t}$$

4. See appendices 2.A and 3.A for additional details on the definition of observation units and the harmonization across data sources.

5. *Occupation groups* are comparable in terms of granularity to ISCO minor occupation groups/ 3-digit codes.

As monthly data on vacancies and searching workers is generally available, I compute aggregate and relative sector-specific labor market tightness at monthly frequency.

Next, I define the sector-specific relative productivity $\tilde{z}_{it}$ as

$$\tilde{z}_{it} \equiv \frac{z_{it}}{Z_t}$$

where z_{it} is the observed productivity sector i and period t, and Z_t is the observed economy-wide productivity. When defined this way, the sector-specific labor productivity is unit free and the measure from which it is computed can be interchanged without any rescaling. Where available, I use relative GDP per hour to measure labor productivity. Where hours worked are not available, I use relative GDP per worker. If no sectoral GDP measures are available, I use relative median wages, i.e. median wages within the sectors normalized with the economy-wide median wage.[6] Using relative wages to approximate relative labor productivity implicitly assumes that that differences in relative wages are driven primarily by differences in labor productivity. This holds true under perfect competition and in the absence of any frictions in the labor market, but is not necessarily the case if other forces (e.g. different degrees of unionization, differences in outside options in the wage bargaining, etc.) are present. For data segmented by geography, multiple data sources are available and all measures are highly correlated, supporting the assumption.[7]

Next, I examine relative sectoral-specific job loss rates. Job losses are observed as changes in employment status from one month to the next. More formally, I define the observed job loss rate δ_{it} as the share of workers that are employed in sector i in period t and unemployed in period $t+1$ of all workers employed in sector i and period t. The sector-specific relative job loss measure δ_{it} is then again defined as the observed sectoral job loss rates normalized with the observed economy-wide loss rates Δ_t:

$$\tilde{\delta}_{it} \equiv \frac{\delta_{it}}{\Delta_t}$$

Although data is available at monthly frequency, I compute yearly average job loss rates.[8]

6. Yearly relative GDP measures can be directly computed from the VGRdL data. Information on relative wages are obtained from the SIAB by computing monthly median wages of full time employed workers by sector, averaging by year, and normalizing with the same measure computed for the economy as a whole.

7. The correlation between relative GDP per hour and relative GDP per worker with relative wages at the state level is 0.9 and 0.95, respectively. For further details, see Appendix 3.A

8. Labor markets are seasonal and the SIAB features year-end-bunching of changes in employment status (resulting from the administrative nature of the data sources used in compiling

Finally, I assess sectoral matching efficiency. This requires additional assumptions on the job creation process. I employ the methodology of Şahin et al. (2014). The matching process is described by a Cobb-Douglas matching technology both on the aggregate level and for each sector individually. Formally, let h_t denote the total count of new jobs created in period t and assume that

$$h_t \;=\; \Phi_t v_t^{\alpha} s_t^{1-\alpha} \tag{3.1}$$

where Φ_t is the aggregate matching efficiency, v_t is the aggregate number of vacancies, s_t is the aggregate number of searching workers, and $\alpha \in [0,1]$ is the input share of vacancies. Sectoral matches h_{it} are assumed to follow

$$h_{it} \;=\; \Phi_t \tilde{\phi}_i v_{it}^{\alpha} s_{it}^{1-\alpha} \tag{3.2}$$

Note that sector-specific matching efficiency is again decomposed into aggregate efficiency Φ_t and relative sector-specific efficiency $\tilde{\phi}_i$, but the sector-specific component is assumed to be constant over time. Sector-specific matching efficiency is then estimated in a two-step procedure: First, aggregate efficiency is estimated from aggregate vacancies and searchers; Second, sector-specific relative efficiency measures are estimated, accounting for aggregate efficiency.

Note that, by construction,

$$\theta_{it} = \Theta_t \tilde{\theta}_{it} \qquad z_{it} = Z_t \tilde{z}_{it} \qquad \delta_{it} = \Delta_t \tilde{\delta}_{it} \qquad \phi_{it} = \Phi_t \tilde{\phi}_i$$

i.e. the observed sector-specific quantities are decomposed into an aggregate component and a relative sector-specific component.[9] For additional details on the computation of sector-specific labor market statistics, as well as the aggregate time series used for the normalization, see Appendix 3.A.

3.3.3 Geographic Segmentation

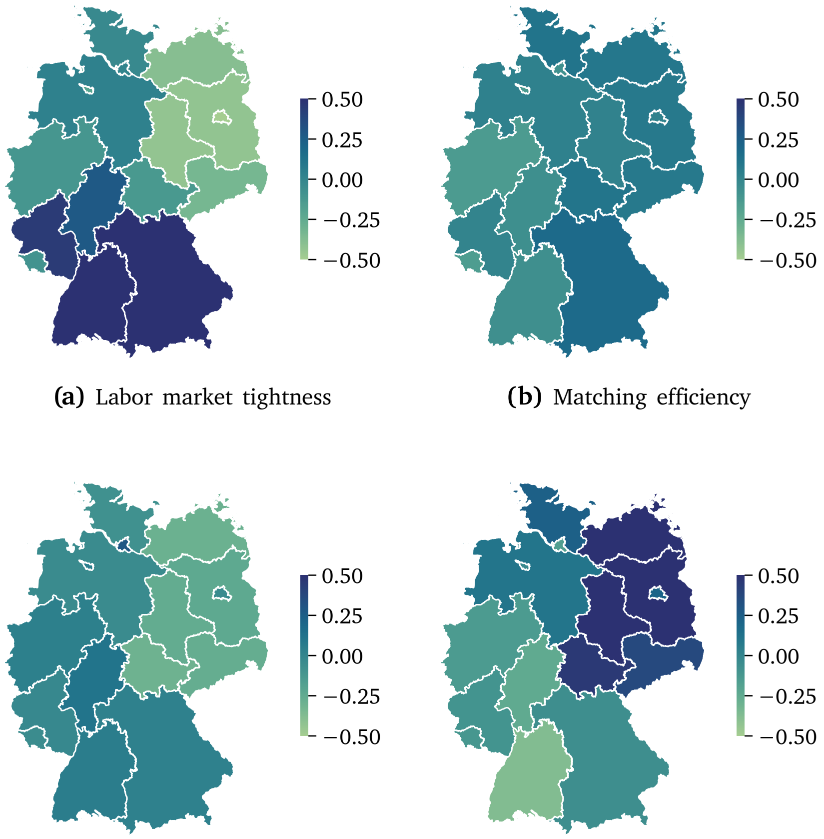
In the main analysis, observation units are federal states. To avoid any issues due to the change in observation units over time, the observation period for time-averaged statistics is January 1994–December 2019.

Figure 3.1 depicts average sector-specific labor market characteristics by federal state. For demonstrative purposes, data is presented as relative deviation from

the SIAB). Labor market mismatch in the sense of this essay is a structural phenomenon, thus seasonality and bunching complicate the measurement. Averaging job loss rates by calendar year mitigates these challenges.

9. Further note that relative deviations from aggregate levels can be obtained by subtracting one from relative sector-specific variables.

Figure 3.1. Geographic segmentation: sectoral labor market characteristics



(a) Labor market tightness

(b) Matching efficiency

(c) Productivity

(d) Job loss rate

Notes: Average relative deviation of labor market characteristics from economy-wide values (observation period 1994–2019). Relative labor market characteristics computed as outlined in Section 3.3.2.
Source: BA (Statistik der Bundesagentur für Arbeit, 2007–2019, 2023), SIAB (Frodermann et al., 2021a), VGRdL (Arbeitskreis "Volkswirtschaftliche Gesamtrechnung der Länder", 2024), own computation.

the economy-wide values. As can be seen, there is substantial heterogeneity in observed labor market tightness. Recall that, in the absence of any sectoral heterogeneity, economic intuition would suggest to equate tightness across sectors. With sectoral heterogeneity, however, regional differences in tightness may be efficient.

Heterogeneity in relative matching efficiency is relatively mild (cf. Figure 3.1b). Labor markets in the states with higher population density tend to be less efficient than labor markets in the states with lower population density. Relative differences in labor productivity are more pronounced and follow a strong divide between the states of the former GDR and the remaining states. Job loss rates show the largest relative differences across sectors and follow a southwest vs. northeast divide. The key observation is that these deviations are correlated. States in southern Germany exhibit a higher than average labor market tightness and labor productivity and a lower than average job loss rate, while the pattern is reversed for states in the former GDR.

This apparent correlation can be quantified by computing the correlation of labor market tightness with market characteristics. For this, I again use time-averages of sectoral labor market tightness and characteristics. I compute within-variable variation (measuring how dispersed a given variable is across sectors) and across-variable correlation (measuring how different characteristics tend to be associated with each other) on the sectoral average values. In conduct this exercise for all available segmentations of the data. Table 3.2 summarizes the results.

Table 3.2. Geographic segmentation: Variation of sectoral labor market characteristics

	Federal states	Labor market regions	Counties
Std.Dev. $\left(\bar{\tilde{\theta}}_i\right)$	0.417	0.603	0.719
Std.Dev. $\left(\tilde{\phi}_i\right)$	0.101	0.168	0.226
Std.Dev. $\left(\bar{\tilde{z}}_i\right)$	0.167	0.133	0.168
Std.Dev. $\left(\bar{\tilde{\delta}}_i\right)$	0.322	0.200	0.199
Corr. $\left(\bar{\tilde{\theta}}_i, \tilde{\phi}_i\right)$	-0.059	0.744	0.712
Corr. $\left(\bar{\tilde{\theta}}_i, \bar{\tilde{z}}_i\right)$	0.511	0.211	0.063
Corr. $\left(\bar{\tilde{\theta}}_i, \bar{\tilde{\delta}}_i\right)$	-0.728	-0.623	-0.579

Notes: Standard deviation and correlation structure of average labor market characteristics (observation period federal states: 1994–2019; labor market regions and counties: 2016–2019). Relative labor market characteristics computed as outlined in Section 3.3.2.
Source: BA (Statistik der Bundesagentur für Arbeit, 2007–2019; Bundesanstalt für Arbeit, 1982–2003a,b), SIAB (Frodermann et al., 2021a), own computation.

The results confirm the patterns described: there is substantial variation across sectors in labor market tightness, as well as in market characteristics; variation in characteristics is highest for loss rates and lowest for matching efficiency;[10] tightness is positively correlated with productivity, yet only slightly so on the county level; tightness is strongly negatively correlated with job loss rates across all segmentations.

In summary, there is substantial geographic heterogeneity across key determinants of potential mismatch between vacant jobs and searching workers. In general, heterogeneity in loss rates is higher than in labor productivity and matching efficiency. Labor markets with particularly high tightness tend to have above average labor productivity and job stability.

3.3.4 Occupational Segmentation

Next, I conduct the same analysis with sectors defined by occupations. As discussed before, the classification of occupations changed from KldB 1988 to KldB 2010 in 2012. I therefore present results for both issues separately. If not stated otherwise, the reporting period for KldB 1988 is from January 1994 to December 2011 and for KldB 2010 from January 2012 to December 2019.

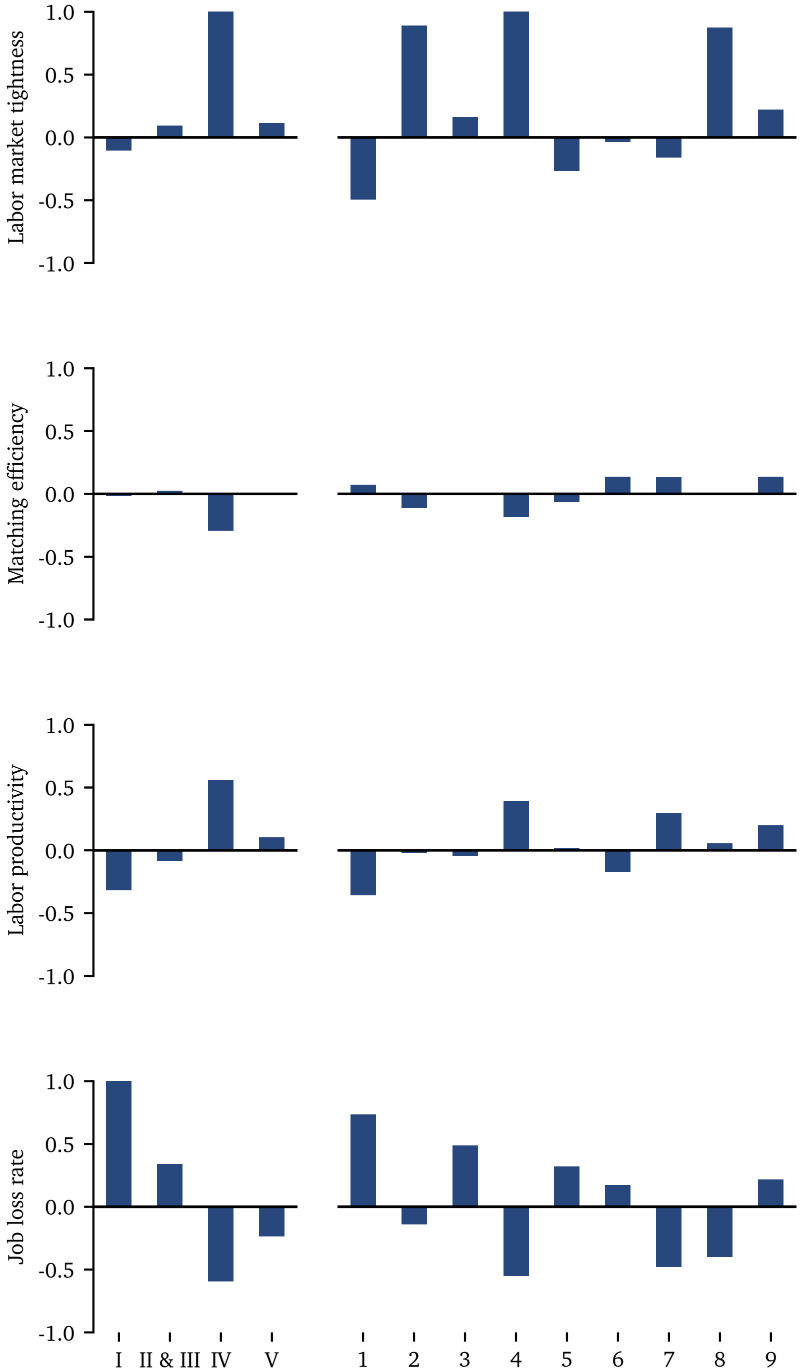
Figure 3.2 depicts relative deviation of labor market statistics from economy-wide averages for occupation areas (aggregation level 1).[11] Again, there are considerable differences in labor market tightness and labor market characteristics across occupation areas. Tightness is highly dispersed and tends to be lower than average in agriculture, and higher than average in technical occupations. Deviations in matching efficiency are minor. Matching efficiency in technical occupations is below the economy-wide average efficiency. Labor is most productive in technical and services occupations and job loss rates are highest in agriculture. As was the case for the geographic segmentation, labor market tightness appears to be positively correlated with labor productivity and negatively correlated with matching efficiency and job loss rates.

Next, I assess within-sector variation and across-sector correlation of labor market characteristics for the occupational segmentation. As mentioned before, the hierarchical structure of the KldB allows for conducting analyses on different aggregation levels. Table 3.3 summarizes the results. Again, the patterns observed above persist, irrespective of classification issue and aggregation level. This indicates that there are indeed structural relationships between variables. For the occupational segmentation, there is again substantial variation across sectors in labor market tightness, as well as in market characteristics. As was the case for the geographic segmentation, variation is highest for loss rates and lowest for

10. Note that the difference in the correlation coefficient of tightness with matching efficiency between federal states and the other segmentations might be due to the different observation periods.

11. For an overview of category labels, see Table 2.A.2 in the appendix.

Figure 3.2. Occupational segmentation: sectoral labor market characteristics



(a) KldB 1988 **(b)** KldB 2010

Notes: Average relative deviation of labor market characteristics from economy-wide values (observation period KldB 1988 1994–2011; KldB 2010 2012–2019). Relative labor market characteristics computed as outlined in Section 3.3.2. Category labels for occupation areas are summarized in Table 3.A.1.
Source: BA (Statistik der Bundesagentur für Arbeit, 2007–2019; Bundesanstalt für Arbeit, 1982–2003a,b), SIAB (Frodermann et al., 2021a), VGRdL (Arbeitskreis “Volkswirtschaftliche Gesamtrechnung der Länder”, 2024), own computation.

matching efficiency. The correlation structure described above is confirmed by the analysis and persists across classification issues and aggregation levels. Across all combinations, the correlations are decreasing in magnitude with increasing granularity of the data.

Table 3.3. Occupational segmentation: Variation of sectoral labor market characteristics

	KldB1988			KldB2010		
	1	2	3	1	2	3
$Std.Dev.\left(\bar{\tilde{\theta}}_i\right)$	0.594	1.035	1.250	0.654	1.075	2.542
$Std.Dev.\left(\tilde{\phi}_i\right)$	0.149	0.167	0.195	0.125	0.210	0.667
$Std.Dev.\left(\bar{\tilde{z}}_i\right)$	0.365	0.260	0.306	0.217	0.256	0.312
$Std.Dev.\left(\bar{\tilde{\delta}}_i\right)$	0.908	0.768	0.788	0.433	0.544	0.788
$Corr.\left(\bar{\tilde{\theta}}_i, \tilde{\phi}_i\right)$	-0.970	-0.320	-0.029	-0.747	-0.355	-0.172
$Corr.\left(\bar{\tilde{\theta}}_i, \bar{\tilde{z}}_i\right)$	0.939	0.581	0.399	0.507	0.320	0.046
$Corr.\left(\bar{\tilde{\theta}}_i, \bar{\tilde{\delta}}_i\right)$	-0.723	-0.444	-0.328	-0.690	-0.508	-0.228

Notes: Standard deviation and correlation structure of average labor market characteristics (observation period KldB 1988: 1994–2011; KldB 2010: 2012–2019). Relative labor market characteristics computed as outlined in Section 3.3.2.
Source: BA (Statistik der Bundesagentur für Arbeit, 2007–2019; Bundesanstalt für Arbeit, 1982–2003a,b), SIAB (Frodermann et al., 2021a), own computation.

In summary, sectoral labor markets exhibit significant differences in labor market tightness and the analyzed characteristics. As mentioned before, differences in tightness, i.e. in how searchers and vacancies are distributed across markets, can potentially be rationalized with differences in characteristics. The correlation structure indicates that sectors with relatively more productive and longer lasting jobs exhibit relatively fewer searchers per vacancy, whereas sectors in which the matching process is relatively more efficient exhibit relatively more searchers per vacancy. The latter observation is in line with economic intuition, whereas the former is at odds with it. Thus, assessing whether the observed variation in tightness across sectors can be accounted for by the observed variation in characteristics, or whether there is mismatch in the German labor market, requires a quantitative analysis. For this, I now introduce a model of labor market mismatch.

3.4 Model

I use a variant of the model developed by Şahin et al. (2014). I extend their framework in several ways: First, I allow for random in- and outflows between model agents inside the labor force and outside the labor force, allowing the model to be aligned with the data more closely. Second, I explicitly solve for the planner allocation of searchers to sectors, allowing me to assess the distributional consequences of implementing the planner allocation. Finally, I introduce additional measures of mismatch that not only capture immediate employment effects, but also account for long-term effects on employment, as well as immediate and long-term effects on output.

3.4.1 Environment

Time is discrete and indexed by t. The economy is populated by a continuum of homogeneous agents with mass one. Economic activity takes place in a large number of sectors indexed by $i \in I$. At the beginning of each period, an agent can be attached to a job in sector i (captured by the state e_{it}), be searching for employment (captured by the state s), or be detached from the labor force.

Existing jobs are destroyed with probability $\Delta_t \tilde{\delta}_{it}$, where the aggregate component Δ_t is common to all sectors. New jobs are created through search in frictional sectoral labor markets. New jobs in sector i, h_{it}, are produced from vacancies v_{it} and searching workers s_{it} using a Cobb-Douglas matching technology. Note that this implies that the matching efficiency is independent of the ratio of vacancies to searchers in the sector. The efficiency parameter of the matching function consists of an aggregate component, Φ_t, which is common to all sectors, and an idiosyncratic component, $\tilde{\phi}_{it}$. The distribution of vacancies across sectors is exogenous.[12]

All sectors directly produce the final output good.[13] There is no physical capital in the economy, i.e. output is produced using labor only. This implies that sectoral output only depends on the total number of workers and labor productivity within a given sector (which in turn is exogenous, i.e. independent of the size of the sectoral work force). The total number of workers in sector i is given by the sum of surviving pre-existing jobs and newly created jobs. Labor productivity consists of an aggregate component Z_t and a sector-specific component $\tilde{z}_{it}$.

In each period, exogenous flows into and out of the labor force occur, either into employment in a given sector, Δe_{it}, or into the searching state, Δs_t.

Aggregate productivity, aggregate matching efficiency and the aggregate job loss rate are assumed to be jointly distributed random variables that follow the conditional distribution function

$$\Gamma_{Z,\Phi,\Delta} = \Gamma(Z_t, \Phi_t, \Delta_t | Z_{t-1}, \Phi_{t-1}, \Delta_{t-1}) \tag{3.3}$$

12. Şahin et al. (2014) solve an extension of the model with endogenous vacancy creation by the planner. They find that, with a minimum set of standard assumptions and parameter values that are standard in the literature, their results are robust to endogenizing vacancy creation.

13. This is equivalent to assuming perfect substitution between sectoral inputs in the production of final goods from sectoral intermediates.

The vector of vacancies by sector $\boldsymbol{v} = \{v_{it}\}$ follows conditional distribution function

$$\Gamma_v(\boldsymbol{v}_t | \boldsymbol{v}_{t-1}, Z_{t-1}, \Phi_{t-1}, \Delta_{t-1}) \tag{3.4}$$

Sector-specific parameters $\tilde{\boldsymbol{z}}_t = \{\tilde{z}_{it}\}$, $\tilde{\boldsymbol{\phi}}_t = \{\tilde{\phi}_{it}\}$, and $\tilde{\boldsymbol{\delta}}_t = \{\tilde{\delta}_{it}\}$ are independent random variables following conditional distribution functions $\Gamma_{\tilde{\boldsymbol{z}}}\left(\tilde{\boldsymbol{z}}_t | \tilde{\boldsymbol{z}}_{t-1}\right)$, $\Gamma_{\tilde{\boldsymbol{\phi}}}\left(\tilde{\boldsymbol{\phi}}_t | \tilde{\boldsymbol{\phi}}_{t-1}\right)$, and $\Gamma_{\tilde{\boldsymbol{\delta}}}\left(\tilde{\boldsymbol{\delta}}_t | \tilde{\boldsymbol{\delta}}_{t-1}\right)$. Finally, the exogenous labor market flows follow the conditional distribution functions $\Gamma_{\Delta e_i}$ and $\Gamma_{\Delta s}$.

All agents receive utility from consuming the final good. Utility from non-participation is normalized to zero. Searchers incur disutility ξ (measured in units of output) from searching. Utility from working is assumed to be higher than utility from non-working, implying that all workers available for search will search in some sector i (as disutility from search is incurred in any case).

The model features a central planner that is subject to some frictions. I assume that the planner may allocate searchers freely across sectors, but the creation of matches is still governed by the matching function. In other words, the planner cannot directly match searchers with vacant positions, but can only assign searchers to frictional sectoral labor markets. The exercise thus consists not of assessing the extent of search frictions present in the market, but instead attempts to assess the impact of improving the allocation of searchers to markets while keeping search frictions in place. The planner takes two simultaneous decisions: the allocation of searchers to sectors $\{s_{it}\}_{i \in I}$, and the change in the labor force from the current period to the next Δl_t. The first choice represents the core mechanism in the model. The latter choice simplifies the analysis while leaving the core mechanism intact. As agents are homogeneous, maximizing the total utility of all workers is equivalent to maximizing total output net of disutility from search, which will be the objective of the planner.[14]

The timing of the model is as follows: At the beginning of each period, the vector of existing jobs by sector, $\boldsymbol{e}_t = \{e_{it}\}_{i \in I}$, and the total number of job searchers, s_t, are given. First, the vector of vacancies $\boldsymbol{v}_t = \{v_{it}\}_{i \in I}$, aggregate shocks $\{Z_t, \Phi_t, \Delta_t\}$, and the vectors of sector specific shocks $\{\tilde{\boldsymbol{z}}_t, \tilde{\boldsymbol{\phi}}_t, \tilde{\boldsymbol{\delta}}_t\}$ are drawn from their respective distribution functions. Workers in the searching state incur disutility from search ξ. The planner then chooses $\{s_{it}\}_{i \in I}$ and Δl_t. Separations occur and the matching process takes place and matches are produced according to matching function

$$h_{it} \;=\; \Phi_t \tilde{\phi}_{it} v_{it}^{\alpha} s_{it}^{1-\alpha} \tag{3.5}$$

Production occurs according to the production function

$$Y_t \;=\; \sum_{i \in I} Z_t \tilde{z}_{it} \left[\left(1 - \Delta_t \tilde{\delta}_{it}\right) e_{it} + h_{it}\right] \tag{3.6}$$

14. The underlying assumption is that total output will be distributed evenly across all agents.

The change in labor force Δl_t is imposed by moving non-working agents out of the labor force or vice versa. Finally, the exogenous in-/outflows into employed and searching agents materialize and the economy moves to the next period.

3.4.2 Planner Problem

The assumptions on the stochastic processes ensure that the planner problem can be expressed in recursive form. For this, let x' denote the next-period value of variable x and denote by Ξ the vector of state variables in period t

$$\Xi = \{\boldsymbol{v}, Z, \Phi, \Delta, \tilde{\boldsymbol{z}}, \tilde{\boldsymbol{\phi}}, \tilde{\boldsymbol{\delta}}\}$$

The planner problem then is

$$\begin{aligned}
V(\boldsymbol{e}, s|\Xi) &= \max_{\{\{s_i\}_{i\in I}, \Delta l\}} \left\{ -s\xi + \sum_{i\in I} \left[\left(1 - \Delta\tilde{\delta}_i\right) e_i + h_i \right] Z\tilde{z}_i + \beta \mathbb{E}\left[V\left(\boldsymbol{e}', s'|\Xi'\right) \right] \right\} \\
\text{s.t.} \sum_{i\in I} s_i &\leq s && (3.7) \\
h_i &= \Phi\tilde{\phi}_i v_i^{\alpha} s_i^{1-\alpha} \quad \forall i && (3.8) \\
e_i' &= \left(1 - \Delta\tilde{\delta}_i\right) e_i + h_i + \Delta e_i \quad \forall i && (3.9) \\
s' &= s + \sum_{i\in I} \left(\Delta\tilde{\delta}_i e_i - h_i \right) + \Delta l + \Delta s && (3.10) \\
s_i &\in [0, s] \quad \forall i && (3.11) \\
\Delta l &\in [-l, 1-l] && (3.12) \\
&\Gamma_{\Delta s}, \Gamma_{\Delta e_i} \text{ given} && (3.13) \\
&\Gamma_{Z,\Phi,\Delta}, \Gamma_{\tilde{\boldsymbol{z}}}, \Gamma_{\tilde{\boldsymbol{\phi}}}, \Gamma_{\tilde{\boldsymbol{\delta}}}, \Gamma_{\boldsymbol{v}} \text{ given} && (3.14)
\end{aligned}$$

To obtain closed form solutions to the planner problem, imposing additional structure on the processes for aggregate and idiosyncratic shocks and for labor force flows is necessary. For simplicity, assume that $\{Z, \Delta, \tilde{\boldsymbol{z}}, \tilde{\boldsymbol{\delta}}\}$ are all positive martingales. Furthermore, assume that the expected value for the exogenous labor force flows is zero, both for all employment states and the searching state. Additionally, assume that the disutility from search ξ and labor force flows are such that the planner solution is always interior. In other words, it is assumed that there will always be sufficiently many agents in the employment states, the searching state, and outside labor force from which to draw upon for the labor force adjustment Δl or the exogenous labor force flows $\Delta \boldsymbol{e}$ and Δs.

The first order condition for the planner allocation of searchers to sectors is then[15]

$$(1-\alpha)\,\Phi\tilde{\phi}_i \left(\frac{v_i}{s_i^*}\right)^{\alpha} \frac{Z\tilde{z}_i}{1 - \beta\left(1 - \Delta\tilde{\delta}_i\right)} = \mu \quad \forall i \qquad (3.15)$$

15. see Appendix 3.B.1 for detailed derivations.

The left-hand side of expression (3.15) captures the probability to create an additional job by allocating another searcher to sector i times the expected total discounted output produced by that job. The right-hand side represents the shadow value of an additional worker that is available for search. As the latter value is independent of the sector, the planner allocation may be obtained by equalizing the left-hand side across sectors.

Note that, in the absence of any sectoral heterogeneity, the planner simply equalizes labor market tightness across sectors. Thus, there may be mismatch even without any sectoral heterogeneity. When differences across sectors are accounted for, labor market tightness is weighted with relative matching efficiency (governing the relative probability of producing new matches from a given number of searchers), relative labor productivity (governing the output produced by a job if created), and the relative job loss rate (governing the expected duration of a created job) before equalizing across sectors.

Going forward, it is convenient to define the weighted relative efficiency of a sector by $\tilde{\chi}_i = \frac{\tilde{\phi}_i \tilde{z}_i}{1-\beta(1-\Delta\tilde{\delta}_i)}$. Rearranging equation (3.15) yields

$$\frac{s_i^*}{s_j^*} = \frac{v_i}{v_j}\left(\frac{\tilde{\chi}_i}{\tilde{\chi}_j}\right)^{\frac{1}{\alpha}} \quad \forall i,j \tag{3.16}$$

Setting an arbitrary base category and combining with the constraint on the total number of workers (3.7) then yields the planner allocation of searchers to sectors

$$s_i^* = \frac{\frac{v_i}{v_1}\left(\frac{\tilde{\chi}_i}{\tilde{\chi}_1}\right)^{\frac{1}{\alpha}}}{\sum_{j\in I}\frac{v_j}{v_1}\left(\frac{\tilde{\chi}_j}{\tilde{\chi}_1}\right)^{\frac{1}{\alpha}}} s \tag{3.17}$$

Equation (3.17), together with the total number of searchers s, determines the planner allocation of searchers to sectors. The planner allocation can then be compared with the decentralized (or observed) allocation, shedding light on the distributional aspects of labor market mismatch.

3.4.3 Measuring Mismatch

As noted above, this analysis sets out to measure the extent of mismatch between searching workers and vacant positions across sectors of the economy. The allocation of searchers to sectoral labor markets determines both the quantity and the quality of newly created jobs. Assigning more searchers to a sector with higher labor productivity will create new jobs that, on average, produce more output. Assigning more searchers to a sector with a lower job loss rate will result in newly created jobs that, on average, last longer. Consequently, there are several ways to measure mismatch in sectoral labor markets and I propose multiple measures, highlighting different aspects of mismatch.

The most straightforward way is to compare the number of matches produced under the planner allocation to the number of new jobs created under the decentralized allocation. More formally, let h^*_{it} denote the number of new jobs in sector i created in period t (using (3.5) and s^*_{it}). Now, define the short-term employment-based mismatch measure for period t as

$$\mathbb{H}^{st}_t = \frac{\sum_{i\in I} h^*_{it}}{\sum_{i\in I} h_{it}} - 1 \tag{3.18}$$

In words, $\mathbb{H}^{st}_t$ measures the relative change in the total number of newly created jobs in the current period from moving from the decentralized allocation $\{s_i\}_{i\in I}$ to the planner allocation $\{s^*_i\}_{i\in I}$.

If job loss rates vary across sectors, jobs created in different sectors differ in their expected duration. Depending on the application, it may therefore be of interest to not only account for immediate employment effects from imposing the planner allocation, but instead to look at long-term effects as well. One way of doing so is by weighting new jobs with their expected duration, or, put differently, with how many periods a job created today is expected to be productive before it gets destroyed. This yields the long-term employment-based measure

$$\mathbb{H}^{lt}_t = \frac{\sum_{i\in I} \left(\tilde{\delta}_{it}\right)^{-1} h^*_{it}}{\sum_{i\in I} \left(\tilde{\delta}_{it}\right)^{-1} h_{it}} - 1 \tag{3.19}$$

Usually, not all sectors of the economy are equally productive. If this is the case, i.e. if labor productivity differs across sectors, looking at employment changes only is not informative about output changes. I therefore propose two additional measures, based on the output produced by the newly created jobs. Analogously to the short-term employment measure, the short-term output measure captures the relative change in output produced in the immediate period from moving from the decentralized allocation to the planner allocation. Formally,

$$\mathbb{Y}^{st}_t = \frac{\sum_{i\in I} \tilde{z}_{it} h^*_{it}}{\sum_{i\in I} \tilde{z}_{it} h_{it}} - 1 \tag{3.20}$$

Again, some of the jobs created in a given period survive into the following periods and continue to generate output. To account for this, the long-term output measure compares the expected total discounted output created by jobs created in period t between planner allocation and decentralized allocation:

$$\mathbb{Y}^{lt}_t = \frac{\sum_{i\in I} \tilde{z}_{it} \left[1 - \beta\left(1 - \Delta_t \tilde{\delta}_{it}\right)\right]^{-1} h^*_{it}}{\sum_{i\in I} \tilde{z}_{it} \left[1 - \beta\left(1 - \Delta_t \tilde{\delta}_{it}\right)\right]^{-1} h_{it}} - 1 \tag{3.21}$$

The long-term measures $\mathbb{H}^{lt}_t$ and $\mathbb{Y}^{lt}_t$ implicitly assume that relative labor productivity, and aggregate and relative job loss rates remain constant for the entire duration of the new job.[16] For further details on the derivation of the measures, see Appendix 3.B.

Note that, in the absence of sectoral differences in productivity and job loss rates, all four measures coincide. In this case, the measures are non-negative by construction.[17] When sectoral heterogeneity in market characteristics is accounted for, the different measures will respond in different ways. Given that the objective of the planner and the quantity measured are not always aligned, accounting for more sectoral heterogeneity does not necessarily increase the measure. In fact, only the long-term output-based measure is fully aligned with the planner objective. Thus, this is the only measure that is non-negative by construction. All other measures can, in principle, take on both positive and negative values, depending on the relative impact of the differing sources of heterogeneity.

The difference between measures is thus informative about the importance of sectoral variation in labor productivity (comparing employment-based with output-based measures) and job loss rates (comparing short-term with long-term measures). Finally, note that, there is no production factor besides labor, no substitution across sectors, and matching is Cobb-Douglas in the model. Thus, implementing a different allocation of searchers, by assumption, does not affect matching efficiency, labor productivity or job loss rates. Further, as vacancies are exogenous, the distribution of vacancies is also, by assumption, unaffected. As the share of newly created jobs of total employment is below 0.5%, the assumption on labor productivity and job destruction appears reasonable. Describing the matching process by a Cobb-Douglas function (i.e. a process where efficiency is independent of labor market tightness) is standard in the literature. The job posting decision, however, is tightly linked to the probability of filling a vacancy and, hence, to the tightness in the respective labor market. Maintaining the Cobb-Douglas assumption on the matching process, the probability of filling a vacancy is decreasing in labor market tightness (or increasing in the number of searchers per vacancy). Accordingly, allocating fewer (more) searchers to a sector will, in reality, likely decrease (increase) the number of vacancies posted in that sector. It is therefore likely that implementing a different allocation of searchers to sectors will affect the distribution of vacancies. Modelling this channel requires a much richer model and additional structure on how firms decide to post vacancies and how firms and workers bargain for wages, which is beyond the scope of this analysis. Considering these limitations, the results should be seen as providing an upper bound on the extent and implications of labor market mismatch.

16. As the measures are derived from expected duration and expected total output, this corresponds to the martingale assumption on the distribution functions $\Gamma_{Z,\Delta,\Phi}$, Γ_z and Γ_δ.

17. As mentioned before, the planner solutions then consists in equalizing tightness across sectors, which cannot decrease the total number of new matches relative to the observed allocation.

3.5 Results

In this section, I apply the model developed in Section 3.4 to the German data presented in Section 3.3. It is important to note that the levels of all measures described in Section 3.4.3 depend on the underlying segmentation of the labor market. Comparisons of measure levels are not valid when the definition of sectors is not identical, be it across segmentations (e.g. geographic vs. occupational), across classifications (e.g. KldB 1988 vs. KldB 2010), across aggregation levels (e.g. KldB 3-digit codes vs KldB 1-digit codes) or if the number of sectors changes (e.g. by including the states of the former GDR after 1994). Therefore, the level of an individual measure is not very informative on its own, yet the comparison across measures and the development of measures over time are.

3.5.1 Parametrization

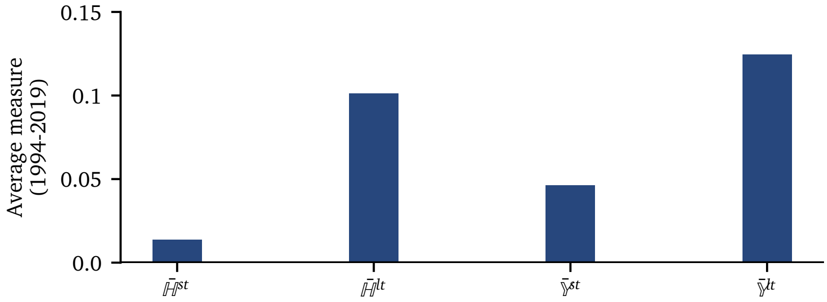
To align the model with the data, some additional assumptions are necessary. To match the worker stocks and flows between model and data, I assume the following: The distribution of workers by sector at the beginning of the period, $\{e_{it}\}_{i \in I}$, is given by the total employment in sector i in period t in the data. The decentralized allocation of searchers to sectors in period t $\{s_{it}\}_{i \in I}$, is given by the total unemployment in sector i and period t in the data. The total number of searchers in period t is then the equal to total unemployment in the data by construction. Next, I assume that the random in-/outflows into or out of the employment states Δe_{it} occur exclusively from those workers that have been already employed at the beginning of the period. This allows me to observe the total number of matches within a given period t and sector i as the total number of observations that are unemployed in sector i in period t and employed in sector i in period $t+1$. Similarly, the (target) labor force adjustment Δl_t and random flows into or out of the searching state Δs_{it} take place exclusively from workers that have been in the searching state at the beginning of the period. The total number of destroyed jobs is then given by the total number of observations with employment in sector i in period t and unemployment in period $t+1$ The above assumptions allow me to observe $\{\{e_{it}\}, s_t, \{\Delta\delta_{it}e_{it}\}, \{h_{it}\}\}$ directly from the data and to back out $\{\{\Delta e_{it}\}, \Delta s_t\}$ as residuals.

Regarding the matching process, I assume that the relative sector-specific matching efficiency is constant over time and given by the values estimated in Section 3.3. Relative sector-specific labor productivity and job loss rates, as well as aggregate loss rates are also given by the values computed in Section 3.3. With all required variables in place, I now compute geographic and occupational mismatch in the German labor market.

3.5.2 Geographic Segmentation

The sectoral labor market characteristics presented in Section 3.3 document substantial heterogeneity across federal states in Germany. These variations across sectors indeed result in labor market mismatch as measured here. Figure 3.3 shows the average mismatch by measure for the regional segmentation of the labor market.

Figure 3.3. Geographic segmentation: comparison of average mismatch measures

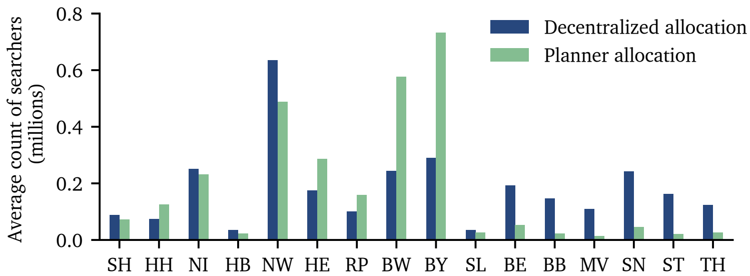


Notes: Average of mismatch measures over time (January 1994–December 2018). Measures as defined in Section 3.4.
Source: BA (Statistik der Bundesagentur für Arbeit, 2007–2019, 2023; Bundesanstalt für Arbeit, 1982–2003b), SIAB (Frodermann et al., 2021a), VGRdL (Arbeitskreis "Volkswirtschaftliche Gesamtrechnung der Länder", 2024), own computation.

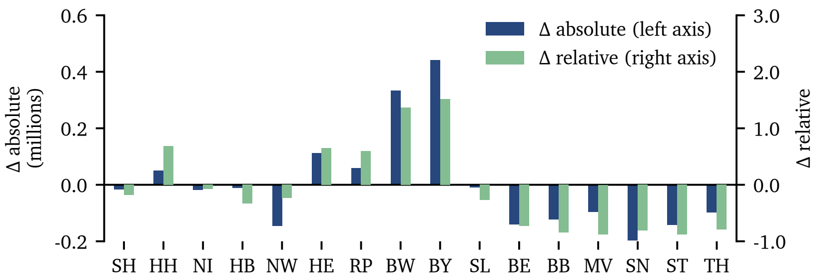
The values are comparable to the results of Şahin et al. (2014) for the US labor market, indicating that the extent of mismatch is similar in both economies. Recall that the measures show the effect of moving from the observed allocation to the planner allocation defined by equation (3.17). Notably, implementing the planner allocation would have negligible immediate employment effects: the increase in new matches is only roughly 1.5% and given that new matches constitute a fraction of total employment, the total employment effect is almost zero. The change in duration-weighted matches is much larger at roughly 10%, indicating that differences in loss rates indeed play an important role. Finally, short and long-term output from new matches would increase by ca. 5% and 13%, respectively. Again, note that this is the change in output from new matches only, the effect on total output is orders of magnitude smaller. The difference between the short-term measures is smaller than the differences to the respective long-term measures, indicating that while productivity differences play a role, differences in job loss rates dominate.[18]

18. For a systematic decomposition of the contributions of the input factors to overall mismatch, see Appendix 3.C.

Figure 3.4. Geographic segmentation: decentralized allocation vs. planner allocation



(a) Allocation of searchers to sectors



(b) Change in allocated searchers by sector

Notes: Average allocation of searchers under decentralized (observed) allocation vs. planner allocation (top panel) and average difference between allocations (bottom panel). Averages over time (January 1994–December 2018). Planner allocation obtained using (3.17).
Source: BA (Statistik der Bundesagentur für Arbeit, 2007–2019, 2023; Bundesanstalt für Arbeit, 1982–2003b), SIAB (Frodermann et al., 2021a), VGRdL (Arbeitskreis “Volkswirtschaftliche Gesamtrechnung der Länder”, 2024), own computation.

The measures discussed above are informative about the potential aggregate gains from reducing mismatch in the labor market. The cost of doing so depends to a large degree on how many searchers would need to be moved to a different sector. The novelty of the approach applied this essay lies in explicitly solving for the planner allocation. Thus, the distributional consequences can be quantified.

Figure 3.4 depicts the observed allocation of searchers and the planner allocation, as well as the difference between allocations. Two observations can be made: First, the implications for the allocation of searchers across sectors are substantial. Under the planner allocation, hundreds of thousands of workers are allocated to a different state. Second, the difference between allocations follows clear cross-regional patterns that coincide with the patterns observed in the empirical analysis: While states in the Northwest remain relatively unaffected, there is a substantial shift from the East to the South. The impact on the East is particularly noteworthy: except for Berlin, all states of the former GDR have 80–90% fewer searchers allocated under the planner allocation.

Table 3.4. Geographic segmentation: average count and share of searchers allocated to different sector

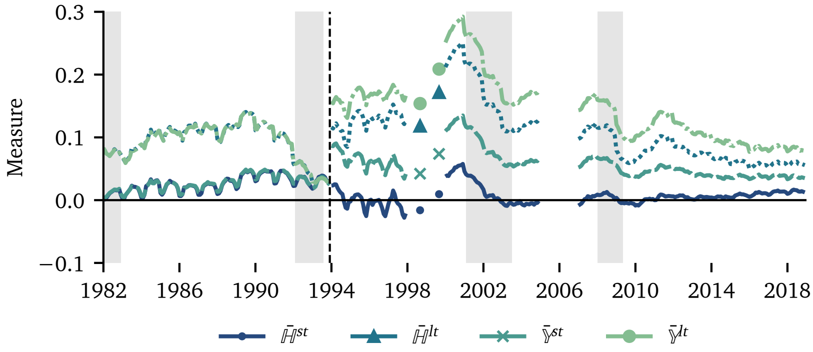
		Searchers allocated to different sector	
Segmentation	# of sectors	Average count (thousands)	Average share
Federal states	16	1,015	34.0 %
Labor market regions	40	663	36.3 %
Counties	328	655	39.8 %

Notes: Average count and share of workers that are allocated to a different sector under the planner allocation compared to the decentralized (observed) allocation by labor market segmentation. Averages over time (federal states: 1994–2011; labor market regions and counties: 2016–2019). Planner allocation obtained using (3.17).
Source: (Statistik der Bundesagentur für Arbeit, 2007–2019, 2023; Bundesanstalt für Arbeit, 1982–2003a), SIAB (Frodermann et al., 2021a), VGRdL (Arbeitskreis "Volkswirtschaftliche Gesamtrechnung der Länder", 2024), own computation.

The available data allows for replicating the analysis for different geographic segmentations. Table 3.4 summarizes the average count and share of workers assigned to a different sector under the planner allocation vs. the observed allocation over the respective observation period. On average, about one third of all searchers are allocated to a different sector. For the 1994–2019 observation period, this amounts to ca. 1 Million workers being assigned to another state.

Finally, I assess how geographic mismatch in the labor market has evolved over time. The unusually long observation period allows for unique insights on the German labor market over many decades, including on the impact of German

Figure 3.5. Geographic segmentation: mismatch measures over time



Notes: Observation units are federal states. Measures for January 1982–December 1993 based on data excl. former territories of the GDR (incl. West Berlin), for January 1994–December 2019 based on data for all of Germany. Measures as defined in Section 3.4. Shaded areas are recession dates as defined by the German Council of Economic Experts.
Source: BA (Statistik der Bundesagentur für Arbeit, 2007–2019, 2023; Bundesanstalt für Arbeit, 1982–2003b), SIAB (Frodermann et al., 2021a), VGRdL (Arbeitskreis "Volkswirtschaftliche Gesamtrechnung der Länder", 2024), SVR (Sachverständigenrat, 2017), own computation.

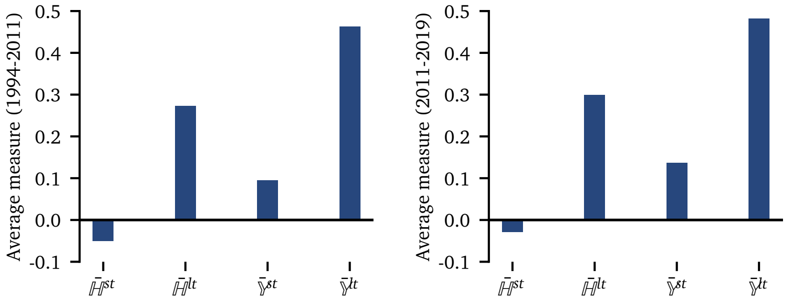
reunification. Figure 3.5 depicts the four mismatch measures, computed month-by-month, between 1982 and 2019. As mentioned before, directly comparing measures from before and after 1994 is not meaningful. It is, however, informative to compare trajectories and relative differences between the measures.

In the period before 1994, mismatch features yearly periodicity, but is overall relatively flat. The similarity between the two short-term measures and the two long-term measures, respectively, indicates that sectoral differences in labor productivity were small prior to 1994. During the reunification recession (1990–1993), mismatch according to the long-term measures is decreasing, while the short-term measures remain unaffected. After including data from the states of the former GDR in the analysis, the difference between measures has increased markedly, indicating that the dataset including all 16 states exhibits more heterogeneity in sectoral characteristics. Moreover, the measures initially diverge, which suggests a further polarization of sectors. Mismatch in the labor market reaches a peak in the early 2000s, after which it has generally been decreasing and the measures are converging again, indicating that sectoral heterogeneity w.r.t. labor productivity and job loss rates has decreased over this period. Overall, it is apparent that regional labor market mismatch is decreasing during times of economic downturn. In all three recessions since the 1990s, mismatch has decreased according to all measures used in this analysis.

3.5.3 Occupational Segmentation

I now repeat the analysis using occupations as sectors. Figure 3.6 depicts averages of the mismatch measures for the occupational segmentation. Though not directly comparable, there are remarkable similarities between the classifications. Note that the short-term employment-based measure $\bar{\mathbb{H}}^{st}$ is negative in both cases. This means that the planner would choose to create fewer new jobs than are created under the decentralized allocation, implying that differences in productivity and durability of jobs are dominating differences in matching efficiency. Moreover, differences between short and long-term measures are relatively large. Combined with the evidence on cross-sector variation presented above, this indicates that differences in job stability is the main driver of occupational mismatch (complementing the results of Kuhn et al., 2021, on the drivers of geographic differences in unemployment).

Figure 3.6. Occupational segmentation: comparison of average mismatch measures



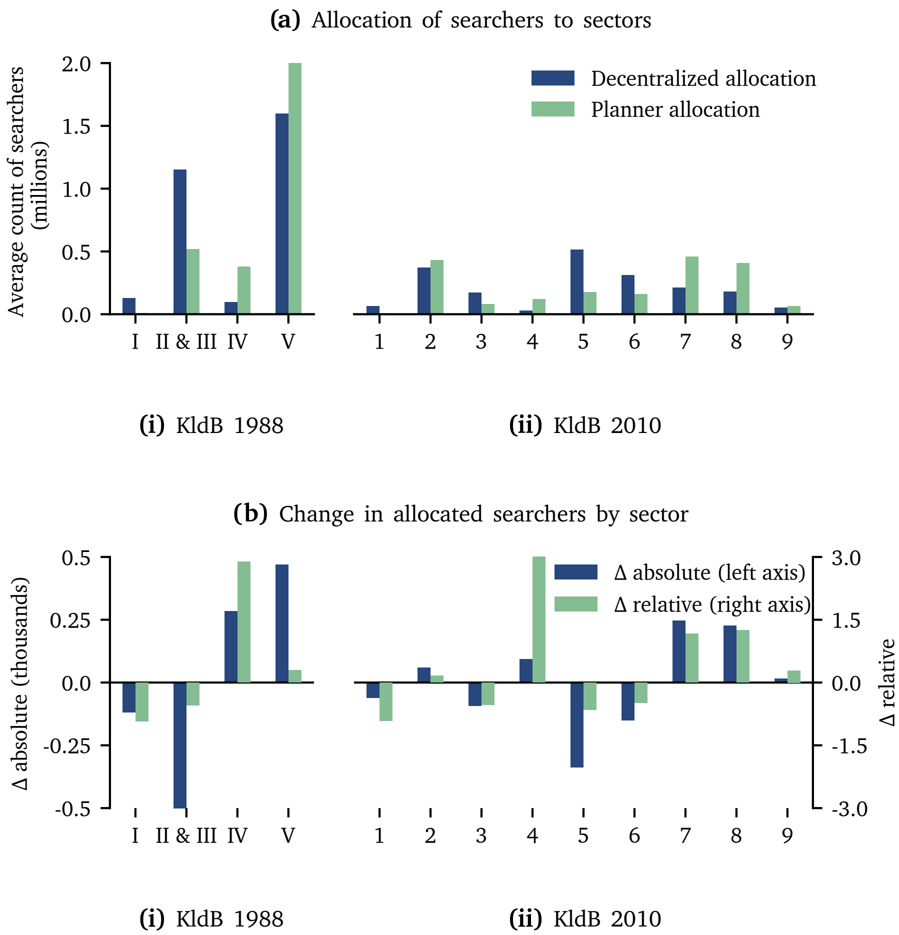
(a) KldB 1988 **(b)** Kldb 2010

Notes: Average of mismatch measures over time (KldB 1988: January 1994–December 2011, KldB 2010: January 2012–December 2018). Measures as defined in Section 3.4.
Source: BA (Statistik der Bundesagentur für Arbeit, 2007–2019, 2023; Bundesanstalt für Arbeit, 1982–2003a), SIAB (Frodermann et al., 2021a), VGRdL (Arbeitskreis “Volkswirtschaftliche Gesamtrechnung der Länder”, 2024), own computation.

The distributional implications of implementing the planner allocation are depicted in Figure 3.7. They are again substantial. As expected from the correlation structure of the market characteristics, the planner would allocate fewer searchers to agriculture and traffic and security and more searchers to technical occupations and selected services occupations.

Total net differences are of the same order of magnitude as in the geographic segmentation. On average, the roughly 750k searchers (ca. 25%) are allocated to a different KldB 1988 occupation area under the planner allocation. For KldB 2010, about one third of all searchers, or approximately 640k searchers, are allocated to a different sector. Table 3.5 summarizes average count and share of searchers allocated to a different sector under the planner allocation vs. the decentralized

Figure 3.7. Occupational segmentation: decentralized allocation vs. planner allocation



Notes: Average allocation of searchers under decentralized (observed) allocation vs. planner allocation (top panel) and average difference between allocations (bottom panel). Averages over time (Kldb 1988: January 1994–December 2011; KldB 2010: January 2012–December 2019). Planner allocation obtained using (3.17).
Source: (Statistik der Bundesagentur für Arbeit, 2007–2019, 2023; Bundesanstalt für Arbeit, 1982–2003a), SIAB (Frodermann et al., 2021a), VGRdL (Arbeitskreis "Volkswirtschaftliche Gesamtrechnung der Länder", 2024), own computation.

Table 3.5. Occupational segmentation: average count and share of searchers allocated to different sector

			Searchers allocated to different sector	
Classification	Level	# of sectors	Average count (thousands)	Average share
KldB1988	3	63	1,525	51.2 %
	2	29	1,468	49.3 %
	1	4	752	24.3 %
KldB2010	3	126	1,034	53.7 %
	2	36	876	45.7 %
	1	10	641	33.3 %

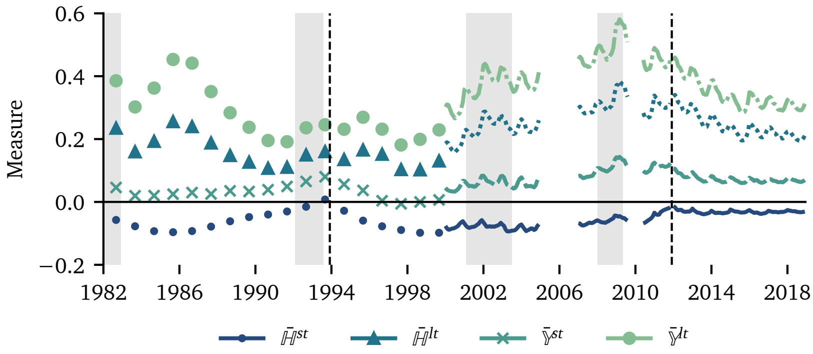
Notes: Average count and share of workers that are allocated to a different sector under the planner allocation compared to the decentralized (observed) allocation by labor market segmentation. Averages over time (KldB 1988: 1982–2011; KldB 2010: 2012–2019). Planner allocation obtained using (3.17).
Source: (Statistik der Bundesagentur für Arbeit, 2007–2019, 2023; Bundesanstalt für Arbeit, 1982–2003a), SIAB (Frodermann et al., 2021a), VGRdL (Arbeitskreis "Volkswirtschaftliche Gesamtrechnung der Länder", 2024), own computation.

allocation. As the definition of a sector becomes more granular, these numbers are increasing. For both classifications, more than 45% of all searching workers are allocated to a different occupation main and more than half are allocated to a different occupation group under the planner allocation.

Finally, Figure 3.8 depicts the development of occupational mismatch over time. First, note that, opposite to the geographic segmentation, there appears to be a positive correlation with the business cycle. This is evidence that regional and occupational mismatch are, at least in part, caused by different forces. Occupational mismatch has generally been decreasing over the 1980s. Until the mid 1990s, it has then increased, which may be attributed to the inclusion of East German workers and firms in the sectoral labor markets. From the mid 1990s to the end of 2009, mismatch has steadily increased and then again decreased until 2019.

In summary, there is evidence of both regional and occupational labor market mismatch in Germany. The magnitude, as measured here, is mild though, and the distributional implications of reducing it are drastic. The effort required to implement the planner allocation would be enormous. For a very rough back-of-the-envelope quantification of the monetary cost of moving alone, consider the official tax deduction amounts for moving ("Umzugskostenpauschale") of the German fiscal authorities. As of March 2024, the deductible is €964 (ca. €750 in 2015

Figure 3.8. Occupational segmentation: mismatch measures over time



Notes: Observation units are occupation main groups (see Appendix 2.A). Measures for January 1982–December 1993 based on data excl. former territories of the GDR (incl. West Berlin), for January 1994–December 2019 based on data for all of Germany. Data until 2011 is coded according to KldB issue 1988, data from 2012 on is coded according to KldB issue 2010. Measures as defined in Section 3.4. Shaded areas are recession dates as defined by the German Council of Economic Experts.
Source: (Statistik der Bundesagentur für Arbeit, 2007–2019, 2023; Bundesanstalt für Arbeit, 1982–2003a), SIAB (Frodermann et al., 2021a), VGRdL (Arbeitskreis "Volkswirtschaftliche Gesamtrechnung der Länder", 2024), SVR (Sachverständigenrat, 2017), own computation.

prices) per person (see Bundesministerium der Finanzen, 2024). Given that actual costs are likely higher than the tax-deductible, actual monetary costs of moving 1 million workers once likely exceed €750 million. Similarly, consider the monetary cost to retrain workers to enable them to work in a different occupation. Even if average retraining cost are as low as €2,000[19] and only half of relocated searchers required retraining, the total monetary cost of implementing the planner allocation once would be upward of €1 billion, which is in the same order of magnitude as the moving cost derived above. Note that these cost are not one-offs: implementing the planner allocation in every period is much more involved than moving or retraining the average count of affected searchers once. A more robust quantification, however, would require a simulation of the allocation of searchers over time, which in turn requires a richer model.

19. In 2024, the German "Aufstiegs-BAföG" (federal programm supporting career advancement training across more than 700 eligible programs, including master craftsperson, technician, specialist, and business administrator qualifications) funded 189,700 participants with a total budget of €1.083 billion, of which €911 million have been grants (Statistisches Bundesamt, 2025). This yields an average grant of ca. €4,800 per participant (ca. €3,800 in 2015 prices).

3.6 Conclusion

This essay combines the rich and well-researched SIAB dataset with long time series for vacancies from historical reports from the ANBA to document sectoral differences in key labor market characteristics in Germany. This unique dataset is employed to assesses the magnitude of mismatch between labor demand as measured by open positions and labor supply as measured by searching workers. The labor market is segmented by region and by occupation. The evidence suggests that heterogeneity in sectoral job loss rates, labor productivity, matching efficiency and labor market tightness is significant, both in terms of location and occupation. Moreover, labor market characteristics are correlated: lower labor market tightness and lower matching efficiency tend to come together with higher job loss rates and lower labor productivity. This indicates that employment and output could potentially be increased by relocating searching workers from sectors with below average tightness, matching efficiency, labor productivity and above average job loss rates to sectors with the opposite characteristics. In terms of geography, this would entail relocating workers from the East to the West and the South of Germany. In terms of occupations, this means moving workers from agriculture, traffic and logistics, and commercial and services occupations to manufacturing, business administration, and the health care and education sector.

Applying a quantitative model à la Şahin et al. (2014), in which the allocation of searchers to sectors is chosen by a planner that can freely allocate searchers to sectors, yet not avoid search and matching frictions in the sectoral labor markets, confirms this intuition. The resulting mismatch, as measured here, is however moderate. I propose the use of four measures of mismatch: short- and long-term measures of employment effects and output effects. Depending on the measure and the segmentation used, imposing the planner allocation changes the average number of new jobs per month between -5% and +25% compared to the observed allocation. Output generated from new jobs would increase between 5% and 40%, however, as new matches account for less than 0.5% of total employment, total output remains almost unchanged. The distributional consequences, on the other hand, are substantial. The planner allocation corresponds to an average relocation of about one third of all searching workers, or approximately 1 million workers, across states or occupation groups in total and up to 90% of the searching workers in some states, in particular in the East. It is noteworthy that the observed correlation of labor market characteristics, yields considerable moves from weaker economic regions to stronger economic regions, thus exacerbating existing regional differences. A rough quantification of moving and retraining costs indicates that implementing the planner allocation once would require the investment of upward of €1 billion.

The exercise indicates that, even absent adjustment cost to location or occupation, the upper bound for potential employment and output gains from resolving labor market mismatch is relatively insignificant. At the same time, implementing the planner allocation has drastic distributional implications and amplifies existing regional differences. Although the analyses conducted in this essay do not allow for welfare assessments, careful consideration of the tradeoff between output and inequality appears to be of key importance when assessing policies aimed at reducing labor market mismatch.

Finally, I find that regional labor market mismatch in Germany is countercyclical, while occupational mismatch is uncorrelated or even positively correlated with the business cycle. This indicates that the forces driving the mismatch might differ across segmentations. What exactly constitutes these forces and why the distribution of searchers across occupations and across locations are affected differently are questions that are left for future research.

References

Amior, Michael. 2024. "Education and Geographical Mobility: The Role of the Job Surplus." *American Economic Journal: Economic Policy* 16 (4): 341–81. [153]

Arbeitskreis "Volkswirtschaftliche Gesamtrechnung der Länder". 2024. *Bruttoinlandsprodukt, Bruttowertschöpfung in den Ländern der Bundesrepublik Deutschland 1991 bis 2023 - Reihe 1, Ländererergebnisse Band 1*. German. Volkswirtschaftliche Gesamtrechnung der Länder. Stuttgart. URL: https://www.vgrdl.de/ (visited on 06/16/2025). [159, 162, 171–178, 182, 186, 187, 194]

Bilal, Adrien. 2021. "The Geography of Unemployment." Working Paper 29269. National Bureau of Economic Research. [152]

Bundesanstalt für Arbeit. 1982–2003a. "Sondernummer der Amtlichen Nachrichten der Bundesanstalt für Arbeit (ANBA). Strukturanalyse – Bestände sowie Zu- und Abgänge an Arbeitslosen und offenen Stellen." Special issue. Nürnberg. [160, 162, 163, 173, 175–178, 182, 194]

Bundesanstalt für Arbeit. 1982–2003b. "Sondernummer der Amtlichen Nachrichten der Bundesanstalt für Arbeit (ANBA). Arbeitsstatistik – Jahreszahlen." Special issue. Nürnberg. [160, 162, 163, 171, 172, 174, 182]

Bundesministerium der Finanzen. 2024. *Umzugskosten. Anhang 29 IV. 2. – Steuerliche Anerkennung von Umzugskosten nach R 9.9 Absatz 2 Lohnsteuer-Richtlinien*. German. Bundesministerium der Finanzen. URL: https://esth.bundesfinanzministerium.de/lsth/2024/B-Anhaenge/Anhang-29-Druck/IV/IV-2/inhalt.html (visited on 09/10/2025). [178]

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. 2014. "Where is the land of opportunity? The geography of intergenerational mobility in the United States." *Quarterly Journal of Economics* 129 (4): 1553–623. [153]

Diamond, Peter A. 1982. "Wage Determination and Efficiency in Search Equilibrium." *Review of Economic Studies* 49 (2): 217–27. [153]

Dvorkin, Maximiliano, and Alexander Monge-Naranjo. 2019. "Occupation Mobility, Human Capital and the Aggregate Consequences of Task-Biased Innovations." Working Papers 2019-064. Human Capital, and Economic Opportunity Working Group. [153]

Frodermann, Corinna, Andreas Ganzer, Alexandra Schmucker, and Philipp Vom Berge. 2021a. *Factually anonymous Version of the Sample of Integrated Labour Market Biographies (SIAB-Regionalfile) – Version 7519 v1*. [155, 159, 160, 162, 163, 171–178, 182, 186, 187, 194]

Frodermann, Corinna, Andreas Ganzer, Alexandra Schmucker, and Philipp Vom Berge. 2021b. "Sample of Integrated Labour Market Biographies Regional File (SIAB-R) 1975-2019." FDZ-Datenreport 05/2021 (en). Nürnberg. [182]

Haltiwanger, John C., Henry R. Hyatt, Lisa B. Kahn, and Erika McEntarfer. 2018. "Cyclical Job Ladders by Firm Size and Firm Wage." *American Economic Journal: Macroeconomics* 10 (2): 52–85. [153]

Heise, Sebastian, and Tommaso Porzio. 2022. "Labor Misallocation Across Firms and Regions." Working Paper 30298. National Bureau of Economic Research. [153]

Kuhn, Moritz, Iourii Manovskii, and Xincheng Qiu. 2021. "The Geography of Job Creation and Job Destruction." Working Paper 29399. National Bureau of Economic Research. [152, 175, 192]

Mortensen, Dale T. 1982. "The Matching Process as a Noncooperative Bargaining Game." In *The Economics of Information and Uncertainty*. NBER Chapters. National Bureau of Economic Research, Inc, 233–58. [153]

Moscarini, Giuseppe, and Fabien Postel-Vinay. 2012. "The Contribution of Large and Small Employers to Job Creation in Times of High and Low Unemployment." *American Economic Review* 102 (6): 2509–39. [153]

Pissarides, Christopher A. 1985. "Short-Run Equilibrium Dynamics of Unemployment, Vacancies, and Real Wages." *American Economic Review* 75 (4): 676–90. [153]

Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung. 2017. "Jahresgutachten 2017/18: Für eine zukunftsorientierte Wirtschaftspolitik." Jahresgutachten. Wiesbaden: Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung. [174, 178, 183]

Şahin, Ayşegül, Joseph Song, Giorgio Topa, and Giovanni L. Violante. 2014. "Mismatch Unemployment." *American Economic Review* 104 (11): 3529–64. [152–154, 158, 163, 164, 171, 179, 183]

Statistik der Bundesagentur für Arbeit. 2007–2019. *Gemeldete Arbeitsstellen (Monatszahlen)*. Nürnberg. URL: https://statistik.arbeitsagentur.de/SiteGlobals/Forms/Suche/Einzelheftsuche_Formular.html?topic_f=gemeldete-arbeitsstellen (visited on 03/27/2025). [159, 160, 162, 163, 171–178, 182, 194]

Statistik der Bundesagentur für Arbeit. 2023. *Bestand an gemeldeten Arbeitsstellen nach Berufsgruppen der KldB 1988*. Custom dataset. Nürnberg. Order no. 335724. [159, 171–178, 182, 194]

Statistik der Bundesagentur für Arbeit. 2025. *Arbeitsstellen im Zeitverlauf (Monatszahlen)*. Nürnberg. URL: https://statistik.arbeitsagentur.de/SiteGlobals/Forms/Suche/Einzelheftsuche_Formular.html?nn=1459928&topic_f=zr-stea-1950 (visited on 03/27/2025). [182, 187]

Statistisches Bundesamt. 2025. *Aufstiegs-BAföG: Zahl der Geförderten im Jahr 2024 nahezu unverändert*. URL: https://www.destatis.de/DE/Presse/Pressemitteilungen/2025/06/PD25_229_214.html (visited on 09/10/2025). Press release. [178]

Appendix 3.A Empirical Analysis

3.A.1 Data Sources

The "Amtliche Nachrichten der Bundesagentur für Arbeit" (Bulletin of the BA, ANBA) publishes several annual special issues with detailed statistics on the German labor market. I use the 1982–2003 issues of the "Arbeitsstatistik: Jahreszahlen" (Labor Statistics: Yearly Data; Bundesanstalt für Arbeit, 1982–2003b) for data on vacancies by federal state in monthly frequency[20] and the 1982–2003 issues of the "Strukturanalyse" (Structural Analysis; Bundesanstalt für Arbeit, 1982–2003a) for data on vacancies by occupation (KldB 1988–3-digit codes) as of end of September for the respective years. Both sources report vacancies according to the old reporting standard, i.e. pooling primary and secondary labor market figures. Since December 2007, the BA publishes the monthly report "Gemeldete Arbeitsstellen (Monatszahlen)" (Registered vacancies (monthly figures); Statistik der Bundesagentur für Arbeit, 2007–2019) by geographic unit.[21] The report includes vacancies by occupation (KldB 2010–3-digit codes). Aggregate data on vacancies is taken from the BA dataset "Gemeldete Stellen (Zeitreihe Monatszahlen)" (Registered vacancies (monthly time series); Statistik der Bundesagentur für Arbeit, 2025). Finally, the BA has provided a custom dataset containing monthly vacancy data by state and KldB88 2-digit codes from January 2000 to November 2011 (Statistik der Bundesagentur für Arbeit, 2023).

Data on wages and employment transitions is taken from the German Sample of Integrated Labour Market Biographies (SIAB). Data access was again provided via a Scientific Use File (SIAB Regional File, referred to as SIAB-R; the version used here is Version 7519 v1, see Frodermann et al., 2021a). The SIAB-R has been supplied by the Research Data Centre (FDZ) of the BA at the Institute for Employment Research (IAB).[22]

From the VGRdL, I use the dataset "Bruttoinlandsprodukt, Bruttowertschöpfung in den Ländern der Bundesrepublik Deutschland 1991 bis 2023" (Series 1, Volume 1; see Arbeitskreis "Volkswirtschaftliche Gesamtrechnung der Länder", 2024). The dataset contains data on GDP, employment, and hours worked by NUTS-1 and NUTS-3 regions. Data on GDP and employment by federal state (NUTS-1) is available for 1982–2020, data on hours worked and data by counties and independent cities (NUTS-3) is available for 2000–2020.

---

20. From 1998 onwards, only September figures are available.

21. Individual reports by federal states since December 2007 and, in addition, by county since December 2016.

22. For more details on the SIAB-R, see the description in Section 2.A or the data report (Frodermann, Ganzer, Schmucker, and Berge, 2021b).

Business cycle information, in particular recession dates, are obtained from the Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung (German Council of Economic Experts, SVR). The recession dates for the German economy presented here are taken from Sachverständigenrat (2017).

3.A.2 Classification

Compatibility across data sources again requires the coarsening of the geographic and occupational segmentations.[23] The 400 NUTS-3 regions are coarsened into 328 *counties*. The 328 *counties* can be aggregated into 16 *federal states*, or into 40 *labor market regions* (derived from IAB labor market regions). The harmonized occupational segmentation features 5 *occupation areas*, 29 *occupation main groups*, and 120 *occupation groups* for KldB 1988, and 10, 37, and 126 units, respectively, for KldB 2010. The *occupation areas* are summarized in Table 3.A.1.

3.A.3 Methodology

In the first step, the aggregate component of the matching efficiency Φ_t is estimated using information on aggregate vacancies, searching workers and new jobs. Taking logarithms of equation (3.1) and rearranging yields the linear equation

$$\log\left(\Phi_t\right) \;=\; \log\left(\frac{h_t}{s_t}\right) - \alpha\log\left(\frac{v_t}{s_t}\right) \tag{3.A.1}$$

This equation can be estimated using OLS. Note however that, as Şahin et al. (2014) point out, the vacancy posting process is subject to endogeneity. To account for this, they suggest the use of a polynomial time trend in combination with structural breaks. I add two structural breaks to the regression equation: when data from the states of the former GDR is first included (1994) and when the reporting of vacancies in the BA data is switched from the old reporting system to the new system (2000). In the main specification, I do not estimate the Cobb-Douglas parameter, but take it from the literature and fix it to $\alpha = 0.5$. The aggregate matching efficiency can then be estimated using the regression function

$$Y_t \;=\; \beta_{pre}\mathbb{1}_{\{t\leq 1994\}} + \beta_{mid}\mathbb{1}_{\{1994<t\leq 2000\}} + \beta_{post}\mathbb{1}_{\{t>2000\}} + \sum_{j=1}^{l}\beta_j t^j + \varepsilon_t \tag{3.A.2}$$

where $Y_t = \log\left(\frac{h_t}{s_t}\right) - \alpha\log\left(\frac{v_t}{s_t}\right)$. The aggregate matching efficiency is then predicted using the coefficient estimates from (3.A.2)

$$\hat{\Phi}_t \;=\; \exp\left(\hat{\beta}_{pre}\mathbb{1}_{\{t\leq 1994\}} + \hat{\beta}_{mid}\mathbb{1}_{\{1994<t\leq 2000\}} + \hat{\beta}_{post}\mathbb{1}_{\{t>2000\}} + \sum_{j=1}^{l}\hat{\beta}_j t^j\right)$$

23. The same adjustments have been made for the analysis in Chapter 2; for details, see Appendix 2.A.

Table 3.A.1. Geographic segmentation: occupation areas

Code	Label
I	Agricultural producers, animal breeders, fishery
II & III	Miners, mineral extractors, manufacturing occupations
IV	Engineering and technical occupations
V	Service occupations
VI	Other occupations

(a) KldB 1988

Code	Label
1	Agriculture, forestry, animal breeding, and horticulture
2	Mining, production, and manufacturing
3	Construction, architecture, and building engineering
4	Natural sciences, geography, and computer science
5	Transportation, logistics, safety, and security
6	Commercial services, trade, hospitality and tourism
7	Professional, administrative, and legal services
8	Medical, social, teaching and educational services
9	Literature, social and economic sciences, media, arts
0	Military

(b) KldB 2010

Notes: Classification of *occupation areas* (aggregation level 1) according to KldB issue 1988 (top) and issue 2010 (bottom). For Kldb 1988, categories II and III are merged for compatibility with the SIAB-R.

In the second step, sector-specific relative matching efficiency parameters are estimated. For this, the fitted values for the time series of aggregate matching efficiency $\hat{\Phi}_t$ are substituted into equation (3.2). Again taking logarithms and rearranging yields another linear function

$$\log\left(\tilde{\phi}_i\right) \;=\; \log\left(\frac{h_{it}}{s_{it}}\right) - \alpha \log\left(\frac{v_{it}}{s_{it}}\right) - \log\left(\hat{\Phi}_t\right) \tag{3.A.3}$$

This equation can again be estimated using linear methods, i.e. I run an OLS panel regression on

$$y_{it} \;=\; \gamma_i + \varepsilon_{it} \tag{3.A.4}$$

where $y_{it} = \log\left(\frac{h_{it}}{s_{it}}\right) - \alpha \log\left(\frac{v_{it}}{s_{it}}\right) - \log\left(\hat{\Phi}_t\right)$. Sector-specific relative matching efficiency is then predicted as $\hat{\phi}_i = \exp\left(\hat{\gamma}_i\right)$. The endogeneity problem also exists for sector-specific data. Assuming that it affects vacancy posting in all sectors in the same way, it is accounted for by normalizing sectoral data with the aggregate time series.

Figure 3.A.1 depicts the aggregate data series used for deriving sector-specific variables. Figure 3.A.1a shows the predicted aggregated matching efficiency from estimating (3.A.2) with α fixed at 0.5 (main specification) and with α estimated as additional regression parameter. As can be seen, both aggregate time series feature very similar trends. As the aggregate series are used for normalizing all sectors, level differences are not relevant for the purpose of this analysis, supporting the modelling choice of fixing α outside the model. Moreover, the structural break in 2000, which represents the change in vacancy reporting from both primary and secondary labor market to primary only, has minor impact on the aggregate series, supporting the decision to pool data across reporting regimes.

Figure 3.A.1b depicts the aggregate time series of GDP per hour, GDP per worker, and median daily wage, all indexed to 2000. As can be seen, all productivity measures exhibit very similar time trends. Again, as aggregate series are used to normalize sector-specific data, level differences are irrelevant. The fact that the measures are highly correlated supports the assumption that they can be used as productivity measures more or less interchangeably, at least in the aggregate.

Finally, Figure 3.A.1c depicts the time series of the aggregate job loss rate. The gap in the time series results from data missing in the SIAB-R.

3.A.4 Robustness

Productivity measures. As outlined before, approximating relative labor productivity with relative wages implicitly assumes that all other forces that determine wages besides productivity are similar across observation units. The validity of this approximation can be tested for where multiple productivity measures are available, which is the case for regional data. To assess the assumption, I compute relative GDP per hour, relative GDP per worker, and relative median wages by county, by labor market region, and by state and examine two measures: within-measure variation and across-measure correlation. Table 3.A.2 summarizes the results. Relative GDP per hour and relative GDP per worker exhibit indeed very

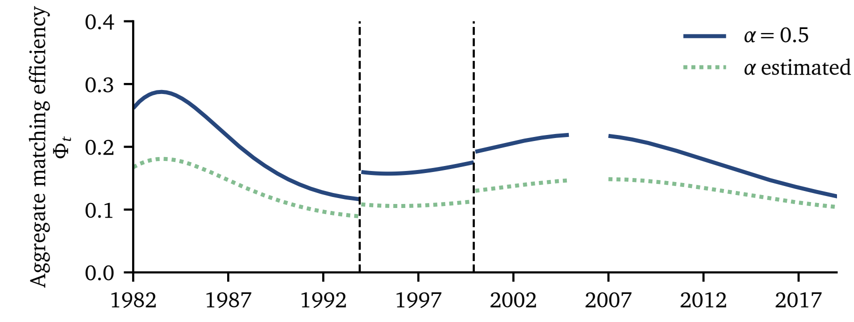
Table 3.A.2. Variation and correlation of productivity measures

	Federal states	Labor market regions	Counties
Std.Dev.(rel. GDP per hour)	0.165	0.138	0.177
Std.Dev.(rel. GDP per worker)	0.158	0.123	0.174
Std.Dev.(rel. median wage)	0.129	0.108	0.125
Corr.(rel. GDP per hour, rel. GDP per worker)	0.988	0.986	0.986
Corr.(rel. GDP per hour, rel. median wage)	0.947	0.922	0.845
Corr.(rel. GDP per worker, rel. median wage)	0.904	0.871	0.791

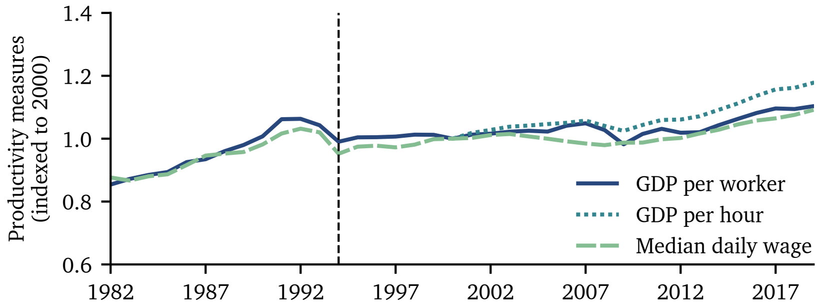
Notes: Variation of productivity measures within segmentations and correlation between productivity measures. Standard deviations are computed over available data, correlations between measures for overlapping time periods.
Data from former GDR from 1994 onwards.
Source: SIAB (Frodermann et al., 2021a), VGRdL (Arbeitskreis "Volkswirtschaftliche Gesamtrechnung der Länder", 2024), own computation.

similar within-measure variation and are very highly correlated. The correlation between relative GDP per hour and relative GDP per worker with relative wages at the state level is 0.9 and 0.95, respectively. Even at the region level, the correlations are 0.8 and 0.85, respectively. Relative wages exhibit a slightly lower variance compared to the GDP measures (standard deviation 0.13 vs. 0.16). If these results transfer to other segmentations, differences in relative wages are a reasonable, though potentially slightly too low, approximation of relative sectoral labor productivity differences.

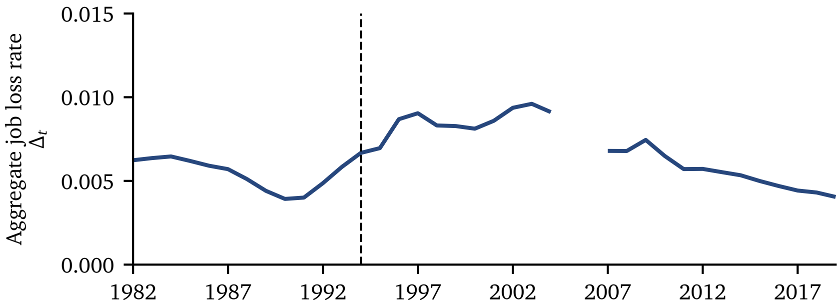
Figure 3.A.1. Aggregate data series



(a) Matching efficiency



(b) Labor productivity



(c) Job loss rates

Notes: Aggregate time series for matching efficiency (top panel), labor productivity measures (indexed to 2000, middle panel), and job loss rate (bottom panel). Matching efficiency obtained using eqs. (3.A.1) to (3.A.3). Job loss rate computed as ratio over employed-to-unemployed transitions over total employment. Dashed vertical lines represent the introduction of East German states (1994) and the change in the vacancy reporting system (2000, only relevant for matching efficiency). The gaps in matching efficiency and loss rate time series are due to missing data in SIAB.
Source: BA (Statistik der Bundesagentur für Arbeit, 2025), SIAB (Frodermann et al., 2021a), VGRdL (Arbeitskreis "Volkswirtschaftliche Gesamtrechnung der Länder", 2024), own computation.

Appendix 3.B Proofs and Derivations

3.B.1 Solution of the Planner Problem

First, note that the matching function satisfies the lower Inada condition w.r.t. s_i, thus ruling out that the planner assigns zero searchers to any sectors. Moreover, considering only interior solutions, condition (3.12) is satisfied by assumption. With

$$h_i\left(s_i\right) = \Phi\tilde{\phi}_i v_i^{\alpha} s_i^{1-\alpha} \tag{3.B.1}$$

$$e_i'\left(e_i, s_i\right) = \left(1 - \Delta\tilde{\delta}_i\right)e_i + h_i\left(s_i\right) + \Delta e_i \tag{3.B.2}$$

$$s'\left(\{e_i, s_i\}_{i\in I}, s, \Delta l\right) = s + \sum_{i\in I}\left[\Delta\tilde{\delta}_i e_i - h_i\left(s_i\right)\right] + \Delta l + \Delta s \tag{3.B.3}$$

the planner problem is equivalent to

$$V\left(\{e_i\}_{i\in I}, s|\Xi\right) = \max_{\{\{s_i\}_{i\in I}, \Delta l\}}\left\{ -\xi s + \sum_{i\in I} Z\tilde{z}_i\left[\left(1-\Delta\tilde{\delta}_i\right)e_i + h_i(s_i)\right] + \beta\mathbb{E}_{\Xi}\left[V\left(\{e_i'(\cdot)\}_{i\in I}, s'(\cdot)|\Xi'\right)\right] + \mu\left(s - \sum_{i\in I} s_i\right)\right\}$$
$$\text{s.t. } \Gamma_{Z,\Delta,\Phi}, \Gamma_{\{v_i\}}, \Gamma_{\{z_i\}}, \Gamma_{\{\delta_i\}}, \Gamma_{\{\phi_i\}}, \Gamma_{\{\Delta e_i\}}, \Gamma_{\Delta s} \tag{3.B.4}$$

The partial derivatives of (3.B.1) w.r.t. the control variables are

$$\frac{\partial h_i}{\partial s_i} = (1-\alpha)\Phi\tilde{\phi}_i\left(\frac{v_i}{s_i}\right)^{\alpha} \quad ; \quad \frac{\partial e_i'}{\partial s_i} = \frac{\partial h_i}{\partial s_i} \quad ; \quad \frac{\partial s'}{\partial s_i} = -\frac{\partial h_i}{\partial s_i} \tag{3.B.5}$$

$$\frac{\partial e_i'}{\partial \Delta l} = 0 \quad ; \quad \frac{\partial s'}{\partial \Delta l} = 1 \tag{3.B.6}$$

For convenience, let denote V_{e_i} and V_s denote the partial derivatives of the value function w.r.t. to its first and second argument, respectively.

First-order conditions are obtained by taking derivatives of (3.B.4) w.r.t. the control variables. For s_i, the FOCs are

$$\frac{\partial h_i}{\partial s_i}\left(\tilde{z}_i + \beta\mathbb{E}_{\Xi}\left[V_{e_i}' - V_s'\right]\right) = \mu \quad \forall i \tag{3.B.7}$$

For Δl, the FOC is

$$\mathbb{E}_{\Xi}\left[V_s'\right] = 0 \tag{3.B.8}$$

The envelope conditions are derived by taking derivatives of (3.B.4) w.r.t. the state variables. Using

$$\frac{\partial e_i'}{\partial e_i} = \left(1-\Delta\tilde{\delta}_i\right) \quad ; \quad \frac{\partial s'}{\partial e_i} = \Delta\tilde{\delta}_i$$
$$\frac{\partial e_i'}{\partial s} = 0 \quad ; \quad \frac{\partial s'}{\partial s} = 1$$

the envelope conditions for $\{e_i\}$ is given by

$$V_{e_i} = Z\tilde{z}_i\left(1-\Delta\tilde{\delta}_i\right) + \beta\left(1-\Delta\tilde{\delta}_i\right)\mathbb{E}_{\Xi}\left[V_{e_i}'\right] + \beta\Delta\tilde{\delta}_i\mathbb{E}_{\Xi}\left[V_s'\right] \tag{3.B.9}$$

and for s by

$$V_s = \mu - \xi$$

Substituting (3.B.8) in (3.B.9) and iterating forward (making use on the martingale assumption on $\{Z, \Delta, \{\tilde{z}_i\}, \{\tilde{\delta}_i\}\}$) yields

$$V_{e_i} = \frac{1-\Delta\tilde{\delta}_i}{1-\beta\left(1-\Delta\tilde{\delta}_i\right)} Z\tilde{z}_i \tag{3.B.10}$$

Substituting eqs. (3.B.5), (3.B.8) and (3.B.10) back into (3.B.7) and applying the martingale property once more, results in

$$(1-\alpha)\,\Phi\tilde{\phi}_i\left(\frac{v_i}{s_i^*}\right)^{\alpha}\frac{Z\tilde{z}_i}{1-\beta\left(1-\Delta\tilde{\delta}_i\right)} = \mu \quad \forall i$$

which is the expression from Section 3.4.2.

3.B.2 Derivation of Mismatch Measures

The approach to measure the long-term impact of mismatch on employment is based on the expected number a new job participates in production. By assumption, all new jobs created in a given period reach the production stage of that period, as well as the employment state at the beginning of the next period. Within the following period, a share of $\Delta_{t+1}\tilde{\delta}_{i,t+1}$ jobs are destroyed and the remaining $(1-\Delta_{t+1}\tilde{\delta}_{i,t+1})$ jobs reach the production stage. Of these jobs, a share $\Delta_{t+2}\tilde{\delta}_{i,t+2}$ are destroyed in the period thereafter, and so on. Formally, let $\tilde{n}_{it}$ denote the expected number of periods that a job created in sector i at time t will participate in production before being destroyed. Then,

$$\begin{aligned}
\tilde{n}_{it} &= 1 + \mathbb{E}_{\Xi}\left[\left(1 - \Delta_{t+1}\tilde{\delta}_{i,t+1}\right)\right] + \mathbb{E}_{\Xi}\left[\left(1 - \Delta_{t+1}\tilde{\delta}_{i,t+1}\right)\left(1 - \Delta_{t+2}\tilde{\delta}_{i,t+2}\right)\right] + \dots \\
&= 1 + \mathbb{E}_{\Xi}\left[\prod_{l=1}^{1}\left(1 - \Delta_{t+l}\tilde{\delta}_{i,t+l}\right)\right] + \mathbb{E}_{\Xi}\left[\prod_{l=1}^{2}\left(1 - \Delta_{t+l}\tilde{\delta}_{i,t+l}\right)\right] + \dots \\
&= 1 + \sum_{j=1}^{\infty}\mathbb{E}_{\Xi}\left[\prod_{l=1}^{j}(1 - \Delta_{t+l}\tilde{\delta}_{i,t+l})\right] \\
&= 1 + \sum_{j=1}^{\infty}\left(1 - \Delta_t\tilde{\delta}_{i,t}\right)^j \\
&= \sum_{j=0}^{\infty}\left(1 - \Delta_t\tilde{\delta}_{i,t}\right)^j \\
&= \left[1 - \left(1 - \Delta_t\tilde{\delta}_{it}\right)\right]^{-1} \\
&= \left(\Delta_t\tilde{\delta}_{it}\right)^{-1}
\end{aligned}$$

The relative change in expected total participations in the production process from all matches created in period t thus is

$$\mathbb{H}_t^{lt} = \frac{\sum_{i\in I}\tilde{n}_{it}h_{it}^*}{\sum_{i\in I}\tilde{n}_{it}h_{it}} - 1 = \frac{\sum_{i\in I}\left(\Delta_t\tilde{\delta}_{it}\right)^{-1}h_{it}^*}{\sum_{i\in I}\left(\Delta_t\tilde{\delta}_{it}\right)^{-1}h_{it}} - 1 = \frac{\sum_{i\in I}\left(\tilde{\delta}_{it}\right)^{-1}h_{it}^*}{\sum_{i\in I}\left(\tilde{\delta}_{it}\right)^{-1}h_{it}} - 1$$

Similarly, the expected total output produced by a job created in period t can be computed. For the output-based measure, however, future output is discounted with discount rate R. The expected total discounted output per job created in period t in sector i is denoted by $\tilde{y}_{it}$ and is given by

$$\begin{aligned}
\tilde{y}_t &= Z_t\tilde{z}_{it} + \sum_{j=1}^{\infty}R^{-j}\mathbb{E}_{\Xi}\left[\prod_{l=1}^{j}\left(1 - \Delta_{t+l}\tilde{\delta}_{i,t+l}\right)Z_{t+l}\tilde{z}_{i,t+l}\right] \\
&= Z_t\tilde{z}_{it} + \sum_{j=1}^{\infty}\left[R^{-1}\left(1 - \Delta_t\tilde{\delta}_{i,t}\right)\right]^j Z_t\tilde{z}_{it} \\
&= Z_t\tilde{z}_{it}\sum_{j=0}^{\infty}\left[R^{-1}\left(1 - \Delta_t\tilde{\delta}_{i,t}\right)\right]^j \\
&= Z_t\tilde{z}_{it}\left[1 - R^{-1}\left(1 - \Delta_t\tilde{\delta}_{it}\right)\right]^{-1}
\end{aligned}$$

With $R^{-1} = \beta$, the relative change in expected total discounted output from moving from the decentralized allocation to the planner allocation then is

$$\mathbb{Y}_t^{lt} = \frac{\sum_{i\in I}\tilde{y}_t h_{it}^*}{\sum_{i\in I}\tilde{y}_t h_{it}} - 1 = \frac{\sum_{i\in I}\tilde{z}_{it}\left[1 - \beta\left(1 - \Delta_t\tilde{\delta}_{it}\right)\right]^{-1}h_{it}^*}{\sum_{i\in I}\tilde{z}_{it}\left[1 - \beta\left(1 - \Delta_t\tilde{\delta}_{it}\right)\right]^{-1}h_{it}} - 1$$

Appendix 3.C Quantitative Analysis

Decomposition of Mismatch Measures. By construction, setting relative sector-specific variables to one is equivalent to assuming that all sectors operate at the aggregate value of the given variable. Sectoral differences in characteristics can thus be accounted for or disregarded by either applying the relative sector-specific values derived in Section 3.3 or by setting all relative sector-specific values to one for that variable. As I assess three sector-specific characteristics, there is a total of six different configurations accounting for or disregarding sectoral differences. Note that each configuration results in a different planner allocation. The values presented in Section 3.5 are the results for the configuration of accounting for sectoral differences in all three variables. The configuration that accounts for no sector-specific differences in characteristics may still be subject to mismatch. As mentioned before, in the absence of any sectoral heterogeneity, the planner would equalize labor market tightness across sectors. This, in some sense, is the simplest measure of labor market mismatch and, in fact, all measures presented here coincide by construction if sectoral heterogeneity in characteristics is disregarded. Moreover, again by construction, this measure is guaranteed to be non-negative.

The decomposition exercise now aims at identifying the contributions of the different sources of heterogeneity to total mismatch. For this, denote by $\bar{\mathbb{X}}^k_{m,l,p}$ the (time-)averaged mismatch measure $\mathbb{X}$ with time-horizon k and configuration of sectoral heterogeneity $\{m, l, p\}$, where $m, l, p \in \{0, 1\}$ denote whether sector-specific matching efficiency (m), loss rates(l), and productivity (p) have been accounted for or not.[24] Moreover, denote the marginal effect of sectoral heterogeneity in characteristic j, as measured by measure $\bar{\mathbb{X}}^k$, by $\bar{\mathbb{X}}^k_j$. The effect of accounting for a given source of heterogeneity for a given configuration can then be computed as the difference in the average measure, with and without heterogeneity in that variable. This effect may differ depending on the configuration (i.e. on which other sources of heterogeneity are being accounted for). Consequently, the marginal effect of a given source of sectoral heterogeneity is measured as the average change in the mismatch measure from including the respective source across orders of inclusion. For instance, the marginal effect of sectoral differences in matching efficiency on the short-term match-based measure is

$$\bar{\mathbb{H}}^{st}_m = \frac{1}{4} \sum_{l \in \{0,1\}} \sum_{p \in \{0,1\}} \left(\bar{\mathbb{H}}^{st}_{1,l,p} - \bar{\mathbb{H}}^{st}_{0,l,p} \right)$$

The marginal effects for all other sources of heterogeneity and all other mismatch measures are defined analogously.

24. E.g. $\bar{\mathbb{H}}^{st}_{1,0,0}$ denotes the average short-term match-based measure when accounting for sectoral heterogeneity in matching efficiency only.

Figure 3.C.1 depicts the marginal effects of accounting for sectoral differences in labor market characteristics for the geographic segmentation. Recall that all measures are defined as the relative difference of some quantity under the planner allocation vs. the observed allocation.[25] Several qualitative and quantitative observations can be made from the decomposition. As outlined before, the sign of the marginal effects depends on the source of heterogeneity and on the measure. Accounting for sectoral heterogeneity in matching efficiency positively contributes across all measures. This is to be expected, as it essentially constitutes that the planner allocates the resources for the matching process more efficiently, disregarding all other factors. Accounting for sectoral heterogeneity in labor productivity positively contributes to the output-based measures and negatively contributes to the match-based measures. The reason is that the planner decides to create fewer jobs that are more productive vis-à-vis the configurations that do not account for differences in productivity. Finally, accounting for sectoral heterogeneity in loss rates contributes negatively to short-term measures and positively to long-term measures. This is again due to the planner's objective: fewer, but more stable jobs are chosen over more, but less stable jobs in the configurations not accounting for differences in job loss rates. Thus, when accounting for sectoral heterogeneity in characteristics, it depends on both the applied measure and the relative importance of the sources of heterogeneity whether mismatch increases or decreases compared to the simple measure.[26]

Quantitatively, there are two key observations: First, in line with the evidence on within-sector variation presented in Section 3.3, the marginal effect of sectoral differences in matching efficiency is relatively small across all measures. This indicates that geographic differences in matching efficiency are of minor importance for labor market mismatch. The largest marginal effects are the contributions of sectoral heterogeneity in loss rates for the long-term measures. As outlined before, this is, in part, due to how the measures are constructed. However, the marginal effects of differences in job loss rates for the short-term measures are sizeable as well. This indicates that heterogeneity in job loss rates (i.e. in job stability) is central for labor market mismatch, replicating the findings of Kuhn et al. (2021).

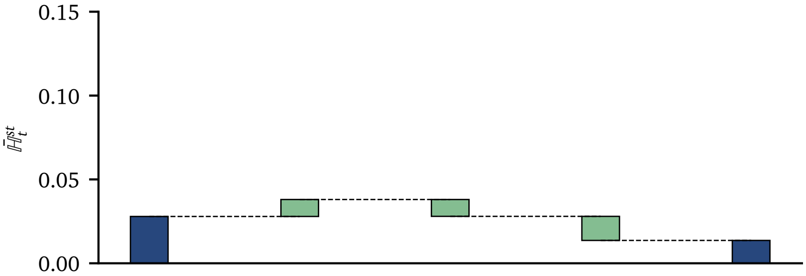
Second, the differences between the simple measure (focusing on heterogeneity in tightness alone) and the measures accounting for all sources of sectoral heterogeneity are substantial, irrespective of measure definition. As already men-

25. For the match-based measure, this quantity is total matches created; for the output-based measure, this is total output produced by the new matches. The short- and long-term measures account for immediate-period effects and effects over the total expected duration of the new matches, respectively.

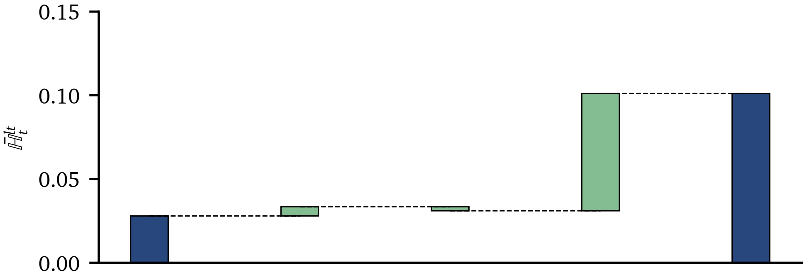
26. Or whether the measure is even positive; In fact, the only measure that accounts for sectoral heterogeneity in labor market characteristics and is guaranteed to be nonnegative is the long-term output-based measure.

tioned in Section 3.5, the variation across measures is also substantial. The decomposition thus reiterates the importance of aligning the measure with the objective: When e.g. designing a policy that aims at reducing labor market mismatch, the choice of mismatch measure has potentially severe consequences on the assessment of the policy.

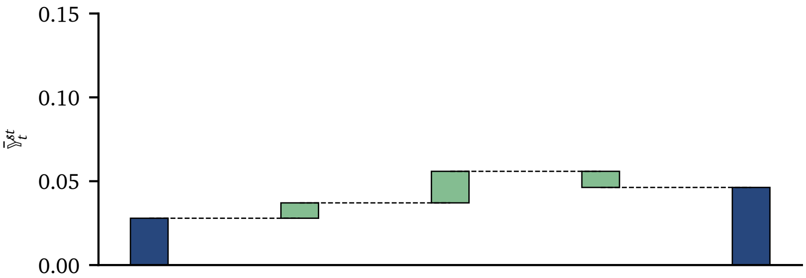
Figure 3.C.1. Geographic segmentation: decomposition of mismatch measures



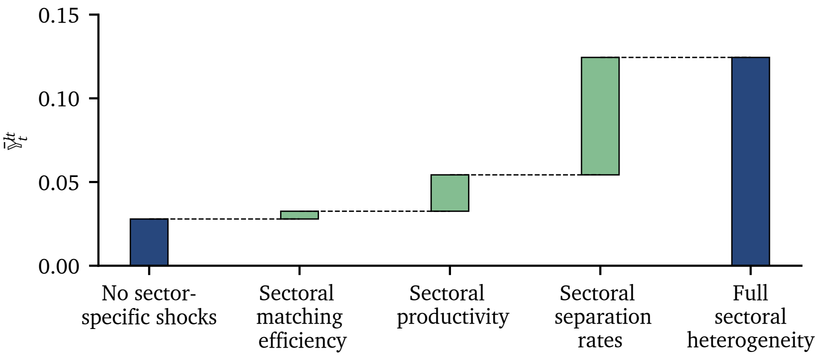
(a) Short-term employment-based



(b) Long-term employment-based



(c) Short-term output-based



(d) Long-term output-based

Notes: Marginal contributions of sources of heterogeneity computed by averaging effects over order of inclusion.
Source: BA (Statistik der Bundesagentur für Arbeit, 2007–2019, 2023; Bundesanstalt für Arbeit, 1982–2003a), SIAB (Frodermann et al., 2021a), VGRdL (Arbeitskreis "Volkswirtschaftliche Gesamtrechnung der Länder", 2024), own computation.

Conclusion

Human capital shapes the economy through distinct but interconnected mechanisms. These mechanisms operate at different dimensions of economic activities. Analysing them requires taking on different perspectives. This constitutes the first insight of this thesis: there is not a single, universally applicable concept of human capital that is useful for all perspectives of economic analysis. From the microeconomic perspective, accumulating individual human capital is a key motive behind worker decisions and a key determinant of labor market outcomes. Understanding these motives is central to designing effective policies. In Chapter 1, I have demonstrated that aligning policy aims with the self-interest of workers allows the policymaker to target UI benefits to those workers that have the highest benefit from public insurance, without requiring complex or conditional policies.

From the perspective of macroeconomic aggregates, human capital represents a key input to production and not only determines how much output is produced, but also how it is produced. Chapter 2 highlights the importance of input availability in determining industry structure and technological setup of an economy as a whole. The demographic transition that we currently observe in virtually all developed economies represents a massive disruption in the relative availability of inputs to production. Capital abundance and labor shortage set the stage for substitution by innovation. Importantly, this is not uniform across sectors of the economy. The substitutability of machines and human capital, determines whether a sector ends up being high-tech or low-tech. This substitutability ultimately traces back to the very foundations of human capital, i.e. it relates to the question of what a human is capable of vs. what a machine is capable of. The analysis demonstrates that secular decline and technological polarization are not merely the result of advances in technology, but that, at least in part, advances in technology, alongside the transition of labor to services and increasing levels of automation in production, are the result of a fundamental change in the availability of human capital.

Focussing on distributional aspects of macroeconomic availability of human capital, I have shown in Chapter 3 that searching workers and vacant positions are not sufficient to create new jobs. Relative location and occupational content matter, and a measure of labor market efficiency that only accounts for how the aggregate number of vacancies and the aggregate number of idle workers are transformed into the aggregate number of new jobs fails to address important dimensions of market efficiency. In truth, the efficiency measured in the analysis still only captures a part of the picture. Taking the distribution of vacancies and searchers, or even more broadly of all individuals participating in economic activities, as exogenous abstracts from important processes that are governed by human capital.

This latter point addresses the second key observation of the thesis: The phenomena described here, while initially appearing to be unrelated, are intricately connected. Differences in wages across sectors, relating to differences in productivity, are one determinant of workers' sectoral choice. But so are differences in the opportunity to accumulate human capital in order to increase future income. Consequently, if these motives are not accounted for, we end up making inaccurate predictions about the aggregate share of workers that choose to work in a given sector of the economy, which in turn is related to labor productivity in the respective sector. In the same spirit, the current geographic location and occupational training a worker possesses are the result of all the choices the workers have made over their previous life. Thus, how well the distribution of workers and vacant positions match ultimately boils down to individual decisions that again cannot be appropriately modeled if the human capital channel is not sufficiently understood. Similarly, labor market mismatch might well be caused by workers that are "stuck" in occupations that are no longer asked for, because the economy is reconfiguring as a response to demographic pressures. Any assessment of whether idle workers should be re-trained thus requires a thorough understanding of the determinants of sectoral demand for human capital.

Finally, a common shortcoming of the essays in this thesis also relates to the very concept of human capital. It is, by definition, unobservable. This is the third insight: Any measure of human capital is only appropriate for the specific perspective for which it has been designed. In that sense, there can be no such thing as a comprehensive study of human capital. We can define approximate measures and try to examine it from many different angles, but ultimately all of these approaches will be limited by the perspective they take. So are the analyses presented here. There are vast components of human capital that are not addressed here: personal and professional networks, transferability of human capital across generations, etc. The difficulty in measuring human capital complicates any quantification, and any attempt at conducting welfare analyses is challenging at best.

Difficult does not mean impossible. So, going forward, there are many more aspects and mechanisms that warrant examination. This thesis can serve as a stepping stone in this journey, furthering the methodology and expanding the data available for the analysis of human capital, thus opening paths for future research.

References

Gerhardt, Holger. No date. *TeXTemplates*. URL: https://github.com/HolgerGerhardt/TeXTemplates. [198]

Heiler, Simon. 2025a. *Replication code for the research project "Demographic Change and Technological Choice"*. Version v1.0.0. URL: https://doi.org/10.5281/zenodo.17102400. Initial Release. [198]

Heiler, Simon. 2025b. *Replication Code for the research project "Labor Market Mismatch in Germany"*. Version v1.0.0. URL: https://doi.org/10.5281/zenodo.17102519. Initial Release. [198]

Heiler, Simon. 2025c. *Replication code for the research project "Worker Heterogeneity and Optimal Unemployment Insurance"*. Version v1.0.0. URL: https://doi.org/10.5281/zenodo.17102414. Initial Release. [198]

von Gaudecker, Hans-Martin. 2019. *Templates for Reproducible Research Projects in Economics*. URL: https://doi.org/10.5281/zenodo.2533241. [198]

Appendix Resources

Replication. Code to replicate all results presented in this thesis is available (for Chapter 1, see Heiler, 2025c; for Chapter 2, see Heiler, 2025a; for Chapter 3, see Heiler, 2025b). Most data sources used in the analysis are publicly available (see sections 1.3, 2.3, and 3.3, and appendices 1.A, 2.A, and 3.A for details on the data sources). Data files containing the datasets constructed for the analyses are available from the author upon request.

Templates and tools. The project code for all chapters is embedded in a template for reproducible projects in computational economics by von Gaudecker (2019). This dissertation has been created using a template by Gerhardt (no date).

Software. Analyses were conducted using Python 3.11, R 4.4.3, and Stata. Complete computational environments are specified in the `environment.yml` files included in the replication code repositories.