Title

Essays on Macroeconomics and Labor Market Frictions

Author

Maruyama, Toshitaka

Publication Date

2024

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
Los Angeles

Essays on Macroeconomics and Labor Market Frictions

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Economics

by

Toshitaka Maruyama

2024

ABSTRACT OF THE DISSERTATION

Essays on Macroeconomics and Labor Market Frictions

by

Toshitaka Maruyama
Doctor of Philosophy in Economics
University of California, Los Angeles, 2024
Professor Lee Edward Ohanian, Chair

This dissertation consists of three chapters. Chapter 1 presents new empirical evidence and a theoretical framework, which suggests that firing costs impede human capital accumulation within firms. To that end, I investigate the impact of a German reform in 2004 that removed firing costs exclusively for smaller establishments. Applying a difference-in-differences approach to administrative matched employee-employer data, I find that the targeted smaller establishments experienced an increase in the proportion of employees participating in training compared with others. To clarify the underlying mechanism and quantify the aggregate effect of the reform, I develop an on-the-job search model with human capital and firing costs. Reducing firing costs in this model stimulates job creation by firms, and thereby increases the probability of workers transitioning to firms in which their skills are more valuable. This encourages workers to accumulate human capital. The calibrated model suggests that the German reform could result in a 0.58% increase in aggregate productivity, with 0.49% attributed to human capital accumulation and the rest to more efficient resource allocation.

Chapter 2, which is co-authored with Yasutaka Koike-Mori and Koki Okumura, develops an endogenous growth model that incorporates on-the-job search and examines the alloca-

tion of inventors across firms, knowledge diffusion, and its impact on growth. In our model, inventors play dual roles: they engage in in-house R&D and transfer knowledge from previous employers to new ones when changing jobs. Using an administrative panel dataset on German inventors matched to their employing establishments and patents, we find that, relative to general workers, inventors are more likely to transition to less productive establishments and suffer a higher wage growth via the transition. We also find that the knowledge base of establishments measured by patents grows faster when a significant proportion of their inventors originate from establishments possessing a larger knowledge base. We then calibrate the model to reflect these empirical findings and examine the effects of innovation policy. While subsidies to frontier firms discourage knowledge diffusion from these firms to technologically lagging firms, these subsidies also encourage innovation within frontier firms. The former negative effect dominates in the short term, but the latter positive effect dominates in the long run.

Chapter 3, which is co-authored with Tomohide Mineyama, estimates the user cost of labor (UCL) - an allocative labor cost measure used in analyzing economies with ongoing employment relationships—while controlling for cyclical changes in job-match quality, which has been identified as a potential problem in previous studies. We use a novel dataset that exploits school (nonemployment)-to-employment flows to control for this measurement problem. The estimated UCL remains highly procyclical after making this correction, with an estimated wage semi-elasticity of 2.69 with respect to a one percentage point change in the unemployment rate. The semi-elasticity is around twice as large as that of the new-hire wage. We also find that the corrected UCL cyclicality is asymmetric, with the UCL rising in booms but remaining flat in downturns. To account for this asymmetry, we develop a directed search model with asymmetric information in which firms use wages as a screening tool to receive applications from targeted workers. The model economy generates asymmetric labor market dynamics that are consistent with the data.

The dissertation of Toshitaka Maruyama is approved.

Pierre-Olivier Weill

Hugo Andres Hopenhayn

Ariel Tomas Burstein

Andrew Granger Atkeson

Lee Edward Ohanian, Committee Chair

University of California, Los Angeles

2024

To my beloved family

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

Acknowledgement is a lengthy term with four syllables, but it falls short of adequately expressing gratitude. I am deeply indebted to the many people who provided unwavering support throughout the journey that is a Ph.D.

I couldn't have wished for a more exceptional main advisor than Lee E. Ohanian. Lee has consistently been extraordinarily friendly and encouraging since my first year. He has consistently maintained contact and has always welcomed me for discussions at any time. He also continuously believed in my potential and devoted substantial time to refining my skills and polishing this dissertation. I am genuinely grateful for the opportunity to learn and flourish under his guidance.

I would also like to extend my thanks to the other members of my dissertation committee, Andy Atkeson, Ariel Burstein, Hugo Hopenhayn, and Pierre-Olivier Weill. Each of them provided valuable advice and guidance, and I am honored to have had their support throughout this process. What I gained from their classes and study groups served as the fundamental groundwork for this dissertation.

The successful completion of this dissertation within four years is greatly indebted to the contributions of my co-authors, Tomohide Mineyama, Yasutaka Koike-Mori, and Koki Okumura. Their expertise and collaborative efforts have proven invaluable, and I have accumulated a wealth of knowledge from our shared endeavors over the years. I consider myself extremely fortunate to have had the opportunity to work alongside such talented economists.

My study at UCLA has been enriched by the support of the many faculty and staff members. I would like to express my sincere thanks to David Baqaee, Saki Bigio, Pablo Fajgelbaum, Joao Guerreiro, Daniel Haanwinckel, Gary D. Hansen, Oleg Itskhoki, Adriana Lleras-Muney, Ichiro Obara, and Jonathan Vogel. I also thank Kosuke Aoki, Adrien Auclert, Masao Fukui, Shinnosuke Kikuchi, Marianna Kudlyak, Toshihiko Mukoyama, Koki Oikawa, Masayuki Okada, Matt Rognlie, Liyan Shi, Ludwig Straub, Conor Walsh, and participants in various seminars for their helpful comments and conversations. I am also thankful to my financial sponsor, the Bank of Japan.

The beginning of my first year was overshadowed by the tragedy of the pandemic. I was unable to come to the U.S. and meet our colleagues and faculty members in person. Nevertheless, in the end, my Ph.D. journey turned out to be incredibly enjoyable, thanks to my friends, including Nobuhiro Abe, Ekaterina Gurkova, Bangyu He, Akira Ishide, Pepe Inguazo, Ali Ismal, Hiroyuki Kubota, Aristotle Magganas, Visarut Malsukhum, Antonio Martner, Patrick Molligo, Benjamin Pirie, Christopher Saw, and Angela Wu.

Lastly, but certainly not least, I extend my gratitude to my family members who have always been there for me, offering encouragement and cheering me on at every step. Without their support, completing this dissertation would have been an immensely challenging endeavor. I am profoundly grateful for their love and encouragement.

VITA

2016	**B.A. in Economics** The University of Tokyo
2022	**M.A. in Economics** University of California, Los Angeles

CHAPTER 1

Do Firing Costs Increase Human Capital Accumulation? Evidence from Germany

1.1 Introduction

A large literature studies the impact of firing costs on aggregate productivity. Some studies have found that these costs can impede job creation, which leads to less effective resource allocation and lower productivity (e.g., Hopenhayn and Rogerson 1993). In contrast, others argue that these costs can enhance human capital accumulation within firms and consequently increase productivity, since training in firms is worthwhile only for long-term employment relations (e.g., Nickell and Layard 1999). While the former perspective has received attention in empirical studies (e.g., Autor et al. 2007), the relationship between human capital investment and firing costs has not been explored.

This paper quantitatively evaluates the impact of firing costs on human capital accumulation within firms. To this end, I use the German employment protection reform in 2004 as a natural experiment. Under German legislation, terminating employment without a valid reason is prohibited, and employees have the right to contest their termination in court. Prior to 2004, only establishments with up to 5 employees were exempt from this protection. The reform in 2004 expanded this exemption to include establishments with up to 10 employees, which resulted in reduced firing costs for establishments with 6 to 10 employees and improved job-to-job transitions within these establishments (Bauernschuster 2013). This study conducts both empirical and theoretical examinations of the reform's impacts on human capital accumulation and aggregate productivity. To the best of my knowledge, this is the first study to provide empirical evidence on the causal impact of firing costs on human capital accumulation and quantify its effects on the aggregate economy.

I proceed in three steps. First, I empirically examine how this reform influenced human capital investment. To do this, I apply a difference-in-differences approach to matched employee-employer data that links the labor market biographies of workers with their employers. The data are sourced from German social security records and span the period from 1975 to 2019. I use data on employer-sponsored training – which encompasses in-company vocational training, traineeships, and external vocational training – as the measure for human capital investment.

My main empirical finding is that establishments with 6 to 10 employees experienced an approximately 1.0% increase in the proportion of employees participating in training compared with other establishments. Also, establishments with fewer than 6 employees experienced a more than 1.0% increase in the proportion. These results suggest that reducing firing costs enhanced human capital accumulation within these smaller establishments. This challenges arguments made in previous studies regarding the impact of firing costs on human capital accumulation. It also implies that workplace training contributes to the development of general human capital and thus supports the findings of previous studies (e.g., Loewenstein and Spletzer 1999; Pischke 2001).

The second part of the paper develops an on-the-job search model with endogenous skill accumulation and firing costs to elucidate the mechanism behind the empirical findings. Workers can accumulate general human capital on-the-job by allocating a fraction of their working time, following Ben-Porath (1967). Whereas investing in human capital incurs the opportunity cost of production and results in lower current wages, it increases human capital in subsequent periods. Workers also engage in on-the-job search in the spirit of Burdett and Mortensen (1998). When a worker transitions to a poaching firm, they can receive a portion of the surplus from the new match and experience a wage gain. In cases in which jobs are terminated, firms are obligated to pay firing costs.

The model describes how the reduction of firing costs can enhance human capital investment. A decrease in firing costs results in an increase in the expected profit of job creation, which prompts firms to generate more job vacancies. This, in turn, improves the probability that workers will transition to jobs in which their skills are more valuable. Since more skilled workers experience higher wage gains upon changing jobs, the increasing likelihood of moving up the job ladder motivates workers to further accumulate human capital.

I calibrate the model to key characteristics of the German labor market to quantify the magnitude of the German reform on aggregate productivity. According to the model, the German reform contributed to a 0.58% increase in aggregate productivity. Of this, 0.49% is attributed to the rise in human capital and the rest is explained by the reallocation to higher productivity firms. Despite the limited number of firms directly affected by the reform, the magnitude is notably significant. This is mainly attributable to the externality of human capital investment proposed by Acemoglu (1997). Because the human capital of workers increases in directly affected firms, other firms can poach higher-skilled workers from these affected firms. Hence, in addition to the rise in job creation by affected firms, there is an increase in job creation by other firms. This induces further reallocation and human capital accumulation across the economy.

Related literature. This paper builds on theoretical literature regarding on-the-job human capital investment dating to Becker (1962). Since a full survey of this literature is beyond the scope of this paper, I cite a few relevant studies. Acemoglu and Pischke (1998) show that a firm's informational advantage regarding their workers' abilities gives them ex post monopsony power, which incentivizes them to invest in general human capital. Nevertheless, their partial equilibrium model does not account for worker bargaining position and firm heterogeneity. Lentz and Roys (2024) develop a general equilibrium model in which the risk-averse workers conduct on-the-job search over heterogeneous firms. They find that search friction can reduce training due to misallocation toward less productive firms with less investment. In addition, since training in each firm increases with a worker's outside option, the search friction decreases the average outside option and training. While they employ the bargaining protocol of Postel–Vinay and Robin (2002), Engbom (2022) uses the protocol of Cahuc et al. (2006), in which workers can receive part of the surplus of a new match, and shows that wages can grow more over the life-cycle in more fluid labor markets. My model extends Engbom (2022) to incorporate firing costs.

Some research explores how employment protection affects the investment of human capital. Wasmer (2006) expands on the approach of Mortensen and Pissarides (1994) by introducing human capital investment. He shows that, in the absence of employment protection, high turnover rates can motivate workers to accumulate general skills instead of those specific to the firm. I provide empirical support using matched employee data and the natural experiment. An empirical study by Ueda and Claessens (2020) shows that employment protection in the U.S. between 1970 and 1990 has benefited the growth of knowledge-intensive industries. In my study, matched employee-employer German data enable us to observe changes in a direct measure for training. Furthermore, Doepke and Gaetani (2024) develop a Diamond-Mortensen-Pissarides search model with endogenous human capital accumulation to understand the difference in the college premium between the U.S. and Germany. The theory in this paper considers on-the-job search, which generates the chance that a worker can extract rent from a new match. I also provide empirical causal inference whereby a reduction in firing costs increases human capital investment.

This paper is inspired by Shi (2023), who develops an on-the-job search model with general human capital accumulation. Her model predicts that noncompete contracts between firms incentivize firms to invest in human capital. Since her model adopts the protocol of Postel–Vinay and Robin (2002), the incumbent firm-worker match cannot extract rent from a new match. Instead, the match can only extract rent from the compensation paid by the poaching firm to release the worker from the contract. Because the compensation increases with the skill of the poached worker, the noncompete contract generates higher human capital investment. In contrast, the theory in this paper adopts the protocol of Cahuc et al. (2006), which is a benchmark in the literature. Furthermore, while Shi (2023) demonstrates that firms with noncompete contracts for executives in the U.S. tend to spend more on intangible investment, I directly use the measure for training.

This paper also contributes to the broader discussion of firing costs and productivity. Autor et al. (2007) use the same natural experiment as Ueda and Claessens (2020) and demonstrates that the decrease in firing costs increased the firm's productivity. Hopenhayn and Rogerson (1993) develop a general equilibrium model of firm dynamics and show that as a consequence of the costs, firms fail to fully adjust their labor force in response to shocks, which decreases allocative efficiency and aggregate productivity. However, Atkenson et al. (1996) point out that the impact of firing costs in Hopenhayn and Rogerson (1993) varies largely depending on the curvature of the production function. Moreover, Hopenhayn (2014) attributes a relatively small quantitative impact of firing costs on aggregate productivity via misallocation, whereas extensions of the model in Hopenhayn and Rogerson (1993) by Boedo and Mukoyama (2012) and Da-Rocha et al. (2019) generate relatively larger effects. Other work examines the effect of firing costs on R&D investment, such as Mukoyama and Osotimehin (2019). Since employment protection also influences R&D, we should use data for training and exclude other intangible investments in order to test the mechanism of our interest. Through its empirical and theoretical analysis, this paper offers new insights into the effect of firing costs on productivity.

The paper is organized as follows. In Section 1.2, I establish empirical findings. In Section 1.3, I propose a simplified model to clarify the mechanism behind the findings. In Section 1.4, I estimate the quantitative model and assess the impact of the reform, and Section 1.5 concludes. The appendix contains additional empirical findings and a more detailed description of the quantitative model.

1.2 Empirical Findings

In this section, I empirically examine the effects of the German reform on human capital investment in firms, using matched employer-employee data. The results serve as motivation for the model presented in Section 1.3.

1.2.1 German Institutional Setting

According to the OECD (2004), German workers enjoy high employment protection compared with workers in other countries. The German Protection Against Dismissal Act deems termination of employment without a "valid reason" as illegal if the employee has worked for more than 6 months. Valid reasons are limited to situations related to (i) the employee's long-term sick leave, (ii) the behavior of the employee, including theft and fraud, or (iii) the business of the employer, such as restructuring. In the case of (iii), the employer must consider specific social criteria, such as tenancy, age, maintenance obligations, or disabilities, when determining which employees to let go. The burden of proof lies with the employer, and the employee can challenge the termination in court. If the appeal succeeds, the employer must either cancel the termination or pay monetary compensation. Jahn and Schnabel (2001) estimate that in 2001, 27% of terminations were challenged in court, with three out of four succeeding.

An exemption has been added to the act. Until 2003, establishments with up to 5 employees were not subject to the protection. In 2004, the German government initiated a reform that extended the exemption range to include establishments with up to 10 employees. The new threshold was announced in November 2003 and implemented in January 2004, and thus establishments only learned the new threshold shortly before the implementation.

The reform was part of a broader labor market reform package known as Agenda 2010 (or Hartz I-IV), which aimed to reduce the high structural unemployment rate. The reform encompassed various other reforms, such as the elimination of limitations on working hours, relaxed regulations for temporary workers, restructuring of public employment security offices, and reduction of unemployment benefit periods. Appendix 4.1.1 details the reform. It is worth noting that these reforms, except for the employment protection reform of our interest, do not impose a threshold dependent on the establishment's size. This implies that if we can observe a relative shift in behavior among establishments with 6 to 10 employees compared with other establishments in 2004, it can be attributed to the firing cost reform.

Bauernschuster (2013) regards this firing costs reform as a reduction in firing costs for establishments with 6 to 10 employees and shows that this reform increased the hiring rates of these establishments compared with other establishments. Leveraging the same natural experiment, I offer causal empirical evidence on how the reduction in firing costs influenced human capital investment within establishments. Appendix 4.1.2 shows that, in our data, we can also observe an increase in worker transitions within these establishments following the reform.

1.2.2 Data

I use matched employee-employer data in Germany, the Sample of Integrated Labor Market Biographies (Stichprobe der Integrierten Arbeitsmarktbiografien - SIAB). The SIAB data contain a 2% random sample of individual accounts drawn from the Integrated Employment Biographies (IEB) of the Institute for Employment Research. The IEB combines data from five sources, each of which may contain information from various administrative procedures. It comprises all individuals in Germany who hold at least one of the following employment statuses: employment subject to social security, marginal part-time employment, receipt of benefits according to the German Social Code III or II, official registration as a job seeker at the German Federal Employment Agency, or (planned) participation in programs of active labor market policies. For the more detailed structure of the data, I refer to Dauth and Eppelsheimer (2020).

The SIAB data provides information on individuals, such as their unique ID, age, gender, educational attainment, daily wage, and a dummy for participation in training each year. The training includes in-company vocational training, traineeships, and external vocational training. The SIAB also includes establishment-level variables, such as establishment ID, the total number of employees, and the mean gross daily wage of full-time workers. Each individual ID is linked with their employer's ID.

I generate additional establishment-level variables, including the mean age of employees, the proportion of women, and the proportion of college graduates. Furthermore, I calculate the proportion of employees participating in training within each establishment each year. Since the total spending for training in each establishment is unobservable, I employ the proportion as a proxy for human capital investment. I opt for the proportion rather than the number of participants because I can only observe information about a subset of workers in each establishment. For example, I may observe 5 workers in one establishment with 6 to 10 employees, but I can only observe 1 worker in another establishment with more than 10 employees.

Previous studies establish that workplace training contributes to the development of general human capital. For instance, Pischke (2001) examines data from the German Socioeconomic Panel and finds that 61% of workers who engaged in workplace training obtained a certificate that could be used for job transitions. Similarly, Loewenstein and Spletzer (1999) use the National Longitudinal Survey of Youth in the U.S. and find that 89% of workers who received company-sponsored training found the skills acquired there to be useful for a different employer. Moreover, 76% of employers stated that most skills learned in their sponsored training are applicable to other employers. These findings support the argument that most firm-specific human capital is actually a combination of general human capital, and firms use general human capital with different weights (Lazear 2009).

I focus on full-time workers aged 21 to 54 and exclude establishments with a mean wage of full-time workers below 15 euros. This criterion follows the analysis by Card et al. (2013) using SIAB data. Furthermore, to implement a difference-in-differences (DiD hereafter)

Table 1.1: Descriptive Statistics for Establishment-level Variables

Period:1991-2019	All continuous est.			Treatment group		
Name of variables	Mean	S.D.	N of obs.	Mean	S.D.	N of obs.
Mean daily wage, euro	87.8	42.6	3,637,026	69.5	32.7	142,967
Mean age of employees	40.4	11.6	3,637,026	37.3	11.8	142,967
Proportion of women, %	37.5	45.2	3,637,026	45.4	48.9	142,967
Proportion of college graduates, %	14.4	32.1	3,637,026	9.6	28.5	142,967
Proportion of employees participating in training, %	54.3	46.4	3,637,026	51.3	49.0	142,967

Notes: Variables are establishment-level. The sample is restricted to full-time workers aged 21 to 54. I also exclude establishments with a mean wage of full-time workers below 15 euros and limit the sample to establishments that were operational in at least 2003 and 2004. The treatment group for the analysis consists of establishments that had 6 to 10 employees in 2003.

approach for the 2004 reform, I restrict my analysis to establishments that were operational during the period covering both 2003 and 2004, which I refer to as continuous establishments. The treatment group for the analysis consists of continuous establishments that had 6 to 10 employees in 2003, and other continuous establishments form the control group. While the original SIAB data span from 1975 to 2019, I use data from 1991 to focus on the periods after German reunification in 1990. To capture the long-run effect, the baseline analysis uses data spanning from 1991 to 2019. As a robustness check, I also conduct a regression using a shorter data period.

Table 1.1 provides descriptive statistics for establishment-level variables from 1991 to 2019. On average, 54.3% of employees participate in training every year, with a large standard deviation of 46.4. This value is higher than those reported in other earlier surveys. Pischke (2001) finds that 35% of employees received workplace training in his 1991-92 data, and 27% in the 1985-86 data.

Several characteristics of the treatment group are worth mentioning. The mean wage in the treatment group is lower than in others, which indicates a higher likelihood of being lower-productivity establishments. These establishments tend to hire fewer college graduates and have younger employees, which suggests that their employees are likely to possess less human capital. Also, the proportion of employees participating in training is lower in the treatment group than in the control group.

1.2.3 Regression Framework

I employ the following DiD framework:

$$\mathrm{y}_{et} = \alpha + \beta_1 \cdot D_{\text{reform}} \cdot 1_{6 \leq N \leq 10} + \beta_2 \cdot X_{et} + \alpha_e + \alpha_t + \varepsilon_{et}$$

The variable y_{et} represents the proportion of employees participating in training in establishment e in year t.[1] D_{reform} is a dummy variable representing periods after the firing cost reform and takes the value of 1 after 2004 and 0 otherwise. The variable $1_{6 \leq N \leq 10}$ equals 1 for establishments in the treatment group and 0 otherwise. The coefficient of β_1 is of interest, because it captures the effect of the firing cost reform on training by establishments with 6 to 10 employees compared with other establishments.

The vector of control variables, X_{et}, includes the total number of employees, mean age of employees, proportion of college graduates, proportion of women, and mean wage. The terms α_e and α_t denote the establishment fixed effect and the time fixed effect, respectively, covering D_{reform} and $1_{6 \leq N \leq 10}$. I estimate standard errors using clustering by establishment and year, and account for the correlation of ε_{et} over time and within the cross-section.

[1] I also estimate an alternative model using worker-level variables with a training participation dummy as the dependent variable. The estimated coefficient of our interest is positive but not statistically significant. This lack of significance might be due to the limited sample size of workers in the treatment group, since many workers left their workplaces after 2004.

1.2.4 Estimation Result

Table 1.2 presents the results. F-values listed at the bottom of the table are the results of a Wald test that assesses whether the linear trend of the treatment group is parallel to that of the control group prior to 2004. The F-value is sufficiently low, which indicates that the hypothesis of the parallel trends assumption cannot be rejected.

The coefficient of the first column is statistically significantly positive. This implies that after the reform, establishments with 6 to 10 employees increased the proportion of employees participating in training by 0.87 percentage points compared with other establishments.

Table 1.2: Impact of Firing Cost Reform on Training

	Proportion of Employees Participating in Training (%)		
	(i)	(ii)	(iii)
$D_{\text{reform}} \cdot 1_{6 \leq N \leq 10}$	0.87**	1.06**	1.01*
	(0.38)	(0.41)	(0.60)
$D_{\text{reform}} \cdot 1_{N<6}$		1.54***	1.30**
		(0.45)	(0.60)
Control	√	√	√
Time & Est. FE	√	√	√
Period	1991-2019	1991-2019	1998-2008
N of obs.	3,069,399	3,069,399	1,362,808
F value of parallel-trends test	0.32	0.34	0.26
Adjusted R^2	0.68	0.68	0.86

Notes: Control variables are the total number of employees, proportion of college graduates, proportion of women, and mean wage. SEs clustered by establishments and year are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Smaller establishments with fewer than 6 employees might also change human capital investment, since changes in hiring by establishments with 6 to 10 employees can influence separation from these smaller establishments. To examine this, in the second column I introduce the interaction term $D_{\text{reform}} \cdot 1_{N<6}$, where $1_{N<6}$ equals 1 for continuous establishments whose total number of employees was fewer than 6 in 2003 and 0 otherwise. In other words, continuous establishments that had fewer than 6 employees in 2003 are excluded from the control group and included in the treatment group. Other specifications remain the same as in the first column.

The coefficient in the second row of the second column is statistically significantly positive. Establishments with fewer than 6 employees increased the proportion of employees participating in training by 1.54 percentage points compared with establishments with more than 10 employees. Since these small establishments with a positive effect are excluded from the control group, the coefficient in the first row becomes larger than in the first column.

It is important to note that the small coefficient does not necessarily imply that the reform's effect on the aggregate economy is small, since the coefficient from the DiD approach cannot capture the aggregate effects in the general equilibrium. Indeed, the quantitative model in Section 1.4 demonstrates that larger establishments, which are included in the control group, can also increase human capital investment due to the externality of human capital investment by smaller establishments.

As a robustness check, I examine a shorter sample from 1998 to 2008 – i.e., within 5 years from 2003. The third column of Table 1.2 shows that the estimated coefficients in this specification are close to those in the second column with a lower significance level.

In summary, findings in this section indicate that the reduction of firing costs in 2004 enhanced training in establishments with fewer than 10 employees. In subsequent sections, I develop a model to explain the underlying mechanism behind these findings and quantify the impact of the reform on the aggregate economy.

1.3 Model

Motivated by the empirical findings in the previous section, I develop a model for on-the-job search that includes endogenous skill accumulation and firing costs. This section presents a simplified model that yields an analytical result and clarifies the underlying mechanism. The quantitative model is detailed in Appendix 4.2, and its quantitative results are introduced in Section 1.4.

1.3.1 Environment

Time is divided into two periods, and there are no aggregate shocks. The population consists of a unit mass of hand-to-mouth workers, each with linear preferences for a single output good, and there is no discounting. Workers enter the market at $t = 1$ and exit at the end of $t = 2$. At the initial entry point in $t = 1$, workers possess initial general human capital denoted by $h_1 = 1$. I also refer to the human capital as the worker's skill. The economy features a single output good, which serves as the numeraire, and it is produced through one worker-one firm matches.[2]

Two types of firms exist in the economy: a high-productivity firm with productivity z_h and a low-productivity firm with productivity z_l, where z_l is less than z_h.

If a firm with productivity z and a worker with human capital h form a match, the output of the match is represented as $y = zh$. This implies that worker skill and firm productivity act as complements, following Acemoglu (1997). Consequently, the marginal value of human capital is higher in the high-productivity firm. The model abstracts from purely firm-specific human capital.

The worker can accumulate skills on-the-job by allocating a fraction of working time,

[2] I assume that a single worker is employed by a single firm to render the model tractable, and I assume that firm size represents firm productivity. If we were to consider multiple workers, the firm-workers match would need to trace the human capital levels of each worker within a single firm and the model is unsolvable.

denoted by i. The growth rate of human capital is expressed by the equation

$$\frac{h_2 - h_1}{h_1} = \frac{1}{\eta}\Big(izh_1\Big)^{\eta}, \tag{1.1}$$

where $\eta > 0$, h_2 is human capital at $t = 2$ and izh is the opportunity cost of training. The opportunity cost is forgone production. Hence, the cost of human capital accumulation for $t = 2$ is front-loaded to worker wage at $t = 1$, which depends on marginal net output. Given the initial human capital value of $h_1 = 1$, equation (1.1) can be simplified as $h_2 - 1 = (iz)^{\eta}/\eta$. Human capital does not depreciate.

Sequence of Events

The exact sequence of events in this economy is as follows. At the beginning of the first period, a worker enters the market with an initial skill level of $h_1 = 1$. I assume that the worker is always matched with a low-productivity firm. Upon being matched, the low-productivity firm generates an output of $(1-i)zh_1 = (1-i)z$. The firm offers human capital investment, denoted by i, along with a wage to the matched worker.

During production in the first period, workers engage in on-the-job search. The high-productivity firm generates v vacancies by paying the vacancy cost, c. When the firm creates v vacancies, the number of meetings is determined by $m = v^{\alpha}$, where $\alpha \in (0, 1)$ represents the elasticity of matches with respect to vacancies. The job-finding rate is hence $p = m/1 = v^{\alpha}$, and the worker-finding rate for the high-productivity firm is $q = m/v = v^{\alpha-1}$.

Given that $z_h > z_l$, workers move to the high-productivity firm if they receive an offer with a probability of p. In the event the worker transitions to the high-productivity firm, the low-productivity firm exits the market.

At the beginning of the second period, following the worker's decision to move, the job at the firms that employed the worker may face termination. The job is destroyed with an exogenous probability denoted by δ. This is equivalent to assuming that a negative shock, that is significant enough to inevitably lead to job loss, happens with a probability of δ.

We can incorporate endogenous job destruction by including fixed operational costs and introducing random productivity shocks, following Hopenhayn (1992). However, this does not qualitatively change the result.

When the job is destroyed, the affected firm (either the high-productivity or the low-productivity firm) is obligated to pay the firing cost. The amount of firing costs varies depending on the type of firm. The low-productivity firm incurs no firing cost, while the how-productivity firm bears a positive firing cost ε.

If the job is destroyed, the firm ceases production, and the worker exits the market. This assumption, wherein a firm must cover the firing cost before production, is by no means necessary for my results, but it simplifies the exposition. This assumption is relaxed in the continuous-time quantitative model in Section 1.4.

If the job is not terminated with a probability of $1 - \delta$, the firm employing the worker produces the output of zh_2. The productivity z is z_h if the worker transitions to the high-productivity firm and z_l if the worker remains at the low-productivity firm. Since this is the last period, the firm does not invest in human capital but pays wages to the worker. At the conclusion of the second period, all agents exit the market.

1.3.2 Worker

In terms of wage determination, I employ the bargaining protocol introduced by Cahuc et al. (2006). Wages are paid as a piece rate, r of net output, $w = r(1 - i)zh_t$. The worker and firm bargain over the piece rate, r, and human capital investment, i. Hence, the cost of human capital is loaded onto the current wage. Following the generalized Nash bargaining approach, the firm and worker jointly determine human capital investment to maximize their joint value (e.g., Engbom 2022; Shi 2023). The joint value is then divided through wages.

Let $W(r, i)$ represent the value of the worker with a piece rate r and human capital investment i, given by

$$W(r,i) = r(1-i)z_l + (1-p)(1-\delta)rz_lh_2 \\ + p(1-\delta)\Big[(1-\delta)z_lh_2 + \beta\{(1-\delta)(z_h-z_l)h_2 - \delta\varepsilon\}\Big] \tag{1.2}$$

The first term represents the wage in the first period, in which the worker receives the piece rate of output, $r(1-i)z_l$. The second term represents the expected value if the worker remains in the low-productivity firm with a probability of $1-p$. If the job is terminated with δ, the worker exits the market and receives nothing. Otherwise, the worker obtains the piece rate of output in the second period, rz_lh_2. There is no human capital investment in the second period, since there is no incentive for investments in the last period.

The third term represents the expected value if the worker moves to a high-productivity firm. If the job is terminated with a probability of δ, the worker receives nothing. I assume that both high- and low-productivity firms engage in Bertrand competition to attract workers. Hence, the worker's outside option becomes the expected full value of the existing match, $(1-\delta)z_lh_2$. The expected full value of the new match is $(1-\delta)z_hh_2 - \delta\varepsilon$. Therefore, if the job is not terminated after moving to the high-productivity firm with $p(1-\delta)$, the worker receives the expected full value by remaining at the low-productivity firm, $(1-\delta)z_lh_2$, and a portion of the additional expected value gained by moving to the high-productivity firm, $\beta\{(1-\delta)(z_h-z_l)h_2 - \delta\varepsilon\}$. Here, β is the bargaining power of the worker with respect to the poaching high-productivity firm, and I assume that $0 < \beta < 1$. I denote this wage term by $R(h_2) \equiv (1-\delta)z_lh_2 + \beta\{(1-\delta)(z_h-z_l)h_2 - \delta\varepsilon\}$ for the later expression.

1.3.3 Low-productivity Firm

Let $F_l(r,i)$ represent the value of the low-productivity firm with piece rate r and human capital investment i, given by

$$F_l(r,i) = (1-r)(1-i)z_l + (1-p)(1-\delta)(1-r)z_lh_2 \tag{1.3}$$

The first term represents the net profit after paying the wage. The second term accounts for the expected profit if the worker does not move to the high-productivity firm with a probability of $1-p$. With a probability of $1-\delta$, the job is not terminated and the firm can attain the net profit $(1-r)z_l h_2$.

Let $J(i)$ be the joint value of the worker and the low-productivity firm. From (1.2) and (1.3), this is given by

$$\begin{aligned} J(i) &= W(r,i) + F_l(r,i) \\ &= (1-i)z_l + (1-\delta)(z_l h_2 + p\Phi h_2) - p\Omega\varepsilon, \end{aligned} \tag{1.4}$$

where the job-finding rate is $p = v^\alpha$, $\Phi \equiv -\delta z_l + (1-\delta)\beta(z_h - z_l)$, and $\Omega \equiv p(1-\delta)\beta\delta > 0$. Φ is positive under reasonable parameter values.[3] Since the piece rate determines how to split the joint value, the joint value itself is independent of piece rate r.

It is worth emphasizing that the worker can obtain a portion of the additional value by transitioning to the high-productivity firm, denoted by $(1-\delta)\beta(z_h - z_l)h_2$. When viewed from the standpoint of the incumbent firm-worker match, this implies that the match is able to extract an equivalent amount of rent from the high-productivity firm with a probability of p. The term $p\Phi h_2$ in (1.4) includes this extracted rent.

The match jointly determines human capital investment i to solve the following maximization problem:

$$\operatorname*{Max}_{i} \quad (1-i)z_l + (1-\delta)\{z_l h_2 + v^\alpha \Phi h_2\} - v^\alpha \Omega\varepsilon \tag{1.5}$$

[3] In the calibration of Section 1.4, I obtain $\delta = 0.003$ and $\beta = 0.321$. In this case, the positive Φ requires $(z_h - z_l)/z_l > 0.009$.

1.3.4 High-productivity Firm

To close the model, let $F_h(v)$ denote the value of the high-productivity firm with a vacancy v, given by

$$F_h(v) = q\{-\delta\varepsilon + (1-\delta)(z_h h_2 - R(h_2))\} - c,$$

where the worker-finding rate is $q = v^{\alpha-1}$ and $R(h_2)$ represents the wage with human capital h_2, $(1-\delta)z_l h_2 + \beta\{(1-\delta)(z_h - z_l)h_2 - \delta\varepsilon\}$ in equation (1.2). The first term in the equation represents the expected profit when the worker is found. If the job is terminated with δ, the firm is obliged to pay the firing cost of ε. Otherwise, the job is not terminated, and the firm can achieve a net profit of $z_h h_2 - R(h_2)$ after paying a wage, $R(h_2)$, to the worker. The high-productivity firm incurs a fixed cost of c to create vacancies.

The high-productivity firm creates vacancies as long as $F_h(v) \geq 0$. Hence, the free entry condition for the high-productivity firm is given by

$$v^{\alpha-1}(1-\delta)(z_h h_2 - R(h_2)) - v^{\alpha-1}\delta\varepsilon - c = 0. \tag{1.6}$$

1.3.5 Equilibrium

I assume that the collected firing costs are disbursed to the worker as a lump sum payment of T. Then, given the parameters $(z_h, z_l, c, \beta, \eta, \delta, \varepsilon, \alpha, h_1, \varepsilon)$, a stationary search equilibrium consists of (J, i, v, h_2, T) such that:

(i) The level of human capital h_2 adheres to equation (1.1),

(ii) Human capital investment i solves (1.5), given v,

(iii) The number of vacancies v aligns with the free entry condition described in equation (1.6),

(iv) The market clears for each period.

1.3.6 Solution

The solution (i^*, v^*) can be obtained as a fixed point of the following two equations:

$$i(v) = \frac{1}{z_l}\{(1-\delta)(z_l + v^\alpha \Phi)\}^{\frac{1}{1-\eta}}, \tag{1.7}$$

$$v(i) = \left[\frac{1}{c}\Big\{(1-\delta)\Upsilon(1+\frac{1}{\eta}(iz_l)^\eta) - \delta^2\varepsilon\Big\}\right]^{\frac{1}{1-\alpha}}, \tag{1.8}$$

where $\Upsilon \equiv z_h - (1-\delta)z_l - \beta(1-\delta)(z_h - zl)$. A nonnegative value of $v(i)$ requires a positive value of Υ.

Equation (1.7) is the solution to the maximization problem in (1.5). It illustrates that $i(v)$ increases with v. If the high-productivity firm creates more vacancies, the incumbent firm-worker match is more likely to extract rent from the poaching firm. Consequently, the match increases human capital investment to raise the rent, which depends on the level of human capital in the next period.

Equation (1.8) is derived from the free entry condition in (1.6). It reveals that $v(i)$ grows with i. If a worker possesses higher human capital, the high-productivity firm creates more vacancies to poach her. Furthermore, $v(i)$ decreases with firing cost ε, since the cost reduces the expected profit from creating vacancies. These arguments result in the following proposition:

Proposition 1: Human capital investment by a low-productivity firm is a decreasing function of the firing cost for a high-productivity firm.

I can also derive the effect of firing costs on aggregate productivity. Aggregate (average) productivity is equivalent to aggregate output, given that there is a unit amount of workers. Let A_t represent aggregate productivity in period t. Then, the aggregate productivity for each period can be expressed as follows:

$$A_1 = z_l$$

$$\begin{aligned}A_2 &= (1-\delta)(1-p)z_l h_2 + (1-\delta)p z_h h_2 \\ &= (1-\delta)\{z_l + (v^*)^\alpha (z_h - z_l)\}\left\{1 + \frac{1}{\eta}(i^* z_l)^\eta\right\}\end{aligned} \tag{1.9}$$

In transitioning from the second equation to the third equation, I substitute $p = (v^*)^\alpha$ and $h_2 = 1 + (i^* z_l)^\eta/\eta$. The term in the second bracket of equation (1.9), $z_l + (v^*)^\alpha(z_h - z_l)$, represents the impact through misallocation. A smaller v signifies a lesser share of the high-productivity firm, resulting in an overall reduction in aggregate productivity. This effect is referred to as the "reallocation effect" in Section 1.4. The third bracket in equation (1.9), $1 + 1/\eta(i^* z_l)^\eta$, illustrates the influence of human capital investment. Diminished human capital investment in the initial period leads to a decrease in aggregate productivity. This effect is termed the "skill accumulation effect" in Section 1.4. Since firing cost reduce both human capital investment and vacancies, equation (1.9) gives rise to the following proposition:

Proposition 2: Aggregate productivity in the second period is a decreasing function of the firing cost for a high-productivity firm.

1.3.7 Understanding the Empirical Findings

Proposition 1 can explain the underlying mechanism behind the findings presented in Section 1.2. The empirical analysis reveals that, following the firing cost reform, establishments with 6 to 10 employees increased their training measures compared with establishments with more than 10 employees. Furthermore, establishments with fewer than 6 employees experienced a greater increase in the measure.

Suppose there are two types of firms, *high-productivity firms* and *low-productivity firms*, among establishments with 6 to 10 employees. In addition, considering the positive correlation between a firm's size and its productivity, we can categorize establishments with fewer than 6 employees as *low-productivity firms*.

Because the reform reduced firing costs for *high-productivity firms*, Proposition 1 suggests that this decrease leads to an uptick in human capital investment by *low-productivity firms*. The diminished firing costs motivate *high-productivity firms* to generate more job vacancies, which attracted workers from *low-productivity firms*. Therefore, the matches between *low-productivity firms* and workers were more likely to extract rent from *high-productivity firms* through worker transitions, and resulted in an increase in human capital investment by these matched pairs. Hence, overall human capital investment by establishments with fewer than 10 employees was increased.

1.4 Quantitative Analysis

I now turn to a quantitative assessment of the impact of the firing cost reform on the German economy. To that end, I develop a quantitative model and estimate it targeting moments in the German labor market. I focus on the long-run steady state. The model is computationally solved using the continuous-time tool developed by Achdou et al. (2021).

1.4.1 Quantitative Model

This section provides an overview of the quantitative model, with a particular emphasis on distinctions from the simplified model in Section 1.3. Appendix 4.2 offers a comprehensive description of the quantitative model.

The quantitative model operates in continuous and infinite time. The economy consists of a unit mass of infinitely lived workers and some mass of firms. All workers possess linear preferences over a single output good discounted at rate ρ. Workers draw different initial skills h_0 from a Pareto distribution Λ with tail indices $1/\sigma$. A firm draws idiosyncratic productivity z from an exogenous offer Pareto distribution Γ with tail indices $1/\xi$, and productivity is time-invariant. The production function is given by $y = zh$.

The growth rate of human capital is given by $\dot{h} = \mu/\eta\Big(izh\Big)^{\eta}$, where $\mu > 0$ and $\eta \in (0, 1)$. If firms create v vacancies, the number of meetings is $m = v^{\alpha}$. Hence, the job-finding rate is $p = v^{\alpha}$, and the worker-finding rate of firms is $q = v^{\alpha-1}$. Jobs are destroyed with a probability of δ. When a job is destroyed, the worker and firm exit the market.

I set a threshold for firing costs $\underline{z}$. If firm productivity z is lower than $\underline{z}$, no firing cost is incurred when the job is destroyed. However, if the productivity is higher than $\underline{z}$, the firing cost amounts to ε. I assume that the reform changes the threshold $\underline{z}$ from $\underline{z}_0$ to $\underline{z}_1$, and evaluate the impact of this change in $\underline{z}$ on the German economy.

The aggregate productivity is given by $A = \int_h \int_z (1 - i(z.h))zhg(z, h)dzdh$, where $g(z, h)$ is the endogenous distribution of matches over productivity and human capital, and solves the Kolmogorov forward equation in Appendix 4.2.

1.4.2 Calibration

I calibrate the model to match the characteristics of the labor market and human capital investment in Germany after the firing cost reform in 2004.

Externally Set

I assign standard values to three parameters, as outlined in Panel A of Table 1.3. The discount rate ρ implies an annual real interest rate of 4%. The curvature of the matching function, α, is set to be 0.5 following Petrongolo and Pissarides (2001). I use the values of the curvature of human capital accumulation and the worker bargaining power calibrated by Engbom (2022). The range of human capital h and firm productivity z is $h \in [1, 100]$ and $z \in [1, 100]$, respectively.

Table 1.3: Parameter Values

Parameter		Estimate
	Panel A. Externally Set	
ρ	Discount rate	0.03
α	Elasticity of matches	0.5
η	Curvature of human capital accumulation	0.497
β	Worker bargaining power	0.321
	Panel B. Direct Match to Data	
δ	Job destruction rate	0.003
$\underline{z}_0$	Threshold for firing costs before reform	3
$\underline{z}_1$	Threshold for firing costs after reform	6
	Panel C. Internally Estimated	
c	Vacancy cost (ratio to the maximum output)	2.79×10^{18}
μ	Drift of human capital accumulation	0.230
σ	Shape of initial human capital dispersion	0.340
ξ	Shape of firm productivity distribution	0.140
τ	Firing cost (ratio to the maximum output)	0.292

Notes: List of model parameters and calibrated values. All parameters in Panel C are calibrated jointly via the method of simulated moments. The maximum output is 1,000.

Direct Match to Data

Two parameters are set to directly match moments from German labor market data, as summarized in Panel B of Table 1.3. The job destruction rate is equal to $\delta = 0.003$, consistent with a 0.3% monthly employment-to-nonemployment transition probability in the microdata after 2004. This is close to the value reported by Jolivet et al. (2006), which is 11.2% within 3 years.

The threshold for firing costs $\underline{z}$ is determined based on the relative size of firms exempt from firing costs under the assumption that firm size represents firm productivity. As described in Section 1.2.1, establishments with fewer than 6 employees are exempt from firing costs before the reform, and those with fewer than 10 employees are exempt from the costs after the reform. In the microdata, the value of firm size at the 90th percentile after 2004 is 167, and the maximum value of firm productivity in the model is 100. Therefore, I consider the 90th percentile as the maximum size and set $\underline{z}_0 = 3$ ($\approx 6/167 \times 100$) and $\underline{z}_1 = 6$ ($\approx 10/167 \times 100$).

Internal Calibration Using a Simulated Method of Moments

I employ a simulated method of moments to estimate the five parameters listed in Panel C of Table 1.3. These parameters are represented by the vector $\Theta = \{c, \mu, \sigma, \zeta, \tau\}$ and estimated by minimizing the objective function

$$\mathcal{L}(\Theta) = (\hat{m} - m(\Theta))' W^{-1} (\hat{m} - m(\Theta)),$$

where $\hat{m}$ is a vector of empirical moments and $m(\Theta)$ are their model counterparts. The diagonal components of matrix W have the same weights, while all non-diagonal components are set to zero.

I base the estimation on five moments outlined in Table 1.4. Targeted moments are based on values after the firing cost reform, specifically during 2014 and 2019. To compute model counterparts, I employ the model with the firing cost threshold $\underline{z} = 6$.

Although the estimation is joint, providing a heuristic discussion of which moments particularly inform each parameter is useful. The drift of human capital accumulation μ is informed by the ratio of the amount of aggregate training investment to gross domestic output in Germany from EU KLEMS. The training investment in EU KLEMS is estimated based on data from the EU Continuing Vocational Training Survey, which includes firms with more than 10 employees. Appendix 4.1.3 provides more explanation. A larger ratio in

Table 1.4: Targeted Moments

Targeted moment	Model	Data	Data source
Training investment / output	0.004	0.006	EU KLEMS
Maximum wage / Mean wage	0.505	0.475	SIAB
Maximum firm size / Mean firm size	0.490	0.492	SIAB
Employment-to-employment transition probability	0.018	0.018	SIAB
Firing costs / output	0.111	0.130	Cahuc et al. (2016)

Notes: List of targeted moments in data and simulated moments from the model. In the first column, GDP is employed as the output for calculating the targeted moment. For targeted moments in the second and third columns, I use the value at the 90th percentile in the data as the maximum value in the calculations.

the data leads to a larger estimated value of μ, because a larger μ increases the marginal productivity of human capital investment, which results in a larger ratio of investment in the model.

The parameter in the Pareto distribution for initial human capital σ is informed by the ratio of the maximum wage to the mean wage in my dataset, with the maximum wage corresponding to the value at the 90th percentile. A higher σ leads to a flatter tail of initial human capital, which causes the ratio to shrink. Similarly, the parameter in the Pareto distribution for firm productivity ξ is informed by the ratio of the maximum firm size to the mean firm size in my dataset, with the maximum firm size corresponding to the value at the 90th percentile.

The vacancy cost c is informed by a monthly employment-to-employment transition probability. The large estimated parameter of c reflects the low employment-to-employment transition probability in my German data.[4] A higher c results in fewer vacancies created and a

[4]It is important to note that my model abstracts from matching efficiency due to the inability to identify vacancy cost and matching efficiency separately (see Free Entry Condition in Appendix 4.2). The quantitative results below remain unchanged even when incorporating matching efficiency and estimating the efficiency

smaller transition probability.

To inform the firing cost τ, I use the value of the average ratio of firing costs to output in EU countries calculated by Cahuc et al. (2016), in which higher τ corresponds to a higher ratio. Since Germany is one of the European countries with the highest firing costs (e.g., OECD 2004), the estimation using the average ratio in European countries can yield a conservative value of τ.

In summary, Table 1.4 demonstrates that the overall fit of the moments within the internal calibration is reasonably well despite the nonlinearity of the model. Therefore, the model is well suited for quantitative assessment of the firing cost reform.

1.4.3 Quantitative Assessment

This section assesses the impact of the firing cost reform on the German economy using the calibrated model. I assume that the reform expands the firing cost threshold from $\underline{z}_0$ to $\underline{z}_1$. Firms with productivity less than the threshold do not incur firing costs when a job is destroyed.

Table 1.5 summarizes the model prediction for the effect of the firing cost reform. The table delineates changes within all firms, firms with productivity less than the new threshold $\underline{z}_1$, and those with productivity exceeding $\underline{z}_1$ in the first, second, and third columns, respectively.

The initial row in the first column of the table shows a 0.58% increase in aggregate productivity resulting from the reform. This impact is notable when considering that real gross domestic output in Germany grew by 0.70% from 2003 to 2004.

The shift in aggregate productivity is divided into two effects, as discussed in Section 1.3.6. The first is the "skill accumulation effect," whereby an upswing in human capital investment alters the overall human capital level in the economy. This encompasses aug-

parameter by the normalized vacancy cost.

Table 1.5: Model Assessment of the Impact of the Firing Cost Reform

Change (%)	Total	Firms w/ productivity $\leq \underline{z}_1$	Firms w/ productivity $> \underline{z}_1$
Aggregate productivity	0.58	0.59	0.57
Skill accumulation effect	0.49	0.51	0.49
Reallocation effect	0.09	0.08	0.10
Human capital investment	0.01	0.01	0.01
Share of N (%)	100	6.0	94.0

Notes: The table illustrates the impact of the firing cost reform on aggregate variables based on the quantitative model. Preceding the reform, firms with productivity levels lower than the threshold $\underline{z}_0$ are exempt from firing costs. The reform extends the threshold from $\underline{z}_0$ to $\underline{z}_1$. Values in the second and third rows indicate the contributions of changes in human capital accumulation and reallocation to the overall change in aggregate productivity. The contribution of the reallocation effect is computed by multiplying the value of each productivity grid by the change in distribution across the grid, respectively. The remaining change in aggregate productivity is ascribed to the skill accumulation effect. Values in the last row represent the share of the number of firms before the reform.

mentation of the human capital of incumbent workers through an increase in investment by incumbent firm-worker match, as well as enhancement of the human capital of poached workers from another firm through an increase in another firm's investment. The second is the "reallocation effect," whereby the redistribution of workers from lower-productivity firms to higher-productivity firms enhances aggregate productivity. The contribution of the reallocation effect is computed by multiplying the value of each productivity grid by the change in distribution across the grid, respectively. The remaining change in aggregate productivity is ascribed to the skill accumulation effect.

The second and third rows in the first column of Table 1.5 illustrate the contributions of these two effects to the overall change in aggregate productivity. The table reveals that the primary factor that influences the change in aggregate productivity is the skill accumulation effect.

Nevertheless, the magnitude of the reallocation effect surpasses what previous studies have suggested. For instance, Hopenhayn and Rogerson (1993) calculate that aggregate productivity experiences a 0.8% increase through reallocation when firing costs for all firms transition from the value equivalent to 6 months' wages to zero. As shown in the third row of the first column in Table 1.5, the model calculates that aggregate productivity rises by 0.09% through reallocation despite the smaller share of firms directly affected by the reform (6.0% of firms with productivity less than $\underline{z}_1$ before the reform) and the lower firing costs (11% of output, as shown in Table 1.4).

The externality of human capital accumulation amplifies the reallocation effect in my model. This is elucidated using the second and third columns of Table 1.5. First, I begin by examining the change within firms with productivity less than $\underline{z}_1$ in the second column of the table. As the threshold shifts from $\underline{z}_0$ to $\underline{z}_1$, firms with productivity between $\underline{z}_0$ and $\underline{z}_1$ generate more vacancies (see Equation (1.8)). Due to the higher job-finding rate, there is an increased likelihood of workers from low-productivity firms moving to high-productivity firms within firms with productivity less than $\underline{z}_1$. This effect is incorporated into the reallocation effect in the second column of the table.

Furthermore, this process heightens the incentive for matches between workers and firms with productivity less than $\underline{z}_1$ to invest more in human capital (see Equation (1.7)). Therefore, the human capital level for all workers in those firms increases. This implies that firms with productivity exceeding $\underline{z}_1$ can also poach workers with more human capital from those firms than before the reform. These effects are included in the skill accumulation effect in the second and third columns of the table.

This positive externality of human capital investment enhances the benefits of creating vacancies for firms with productivity exceeding $\underline{z}_1$ (see equation (1.8)), which results in the creation of more vacancies. Since the production function assumes that human capital and firm productivity are complements ($y = zh$), the additional gain by poaching workers with more human capital is higher for higher-productivity firms. This boosts job creation by firms with productivity exceeding $\underline{z}_1$, especially higher-productivity firms. Hence, this triggers reallocation from firms with productivity less than $\underline{z}_1$ to those with productivity exceeding $\underline{z}_1$ *and* reallocation within firms with productivity exceeding $\underline{z}_1$. This effect is shown in the positive reallocation effect in the third column of the table. Given that reallocation takes place across all firms, the reallocation effect in the overall economy is relatively more substantial.

The fourth row illustrates the alteration in human capital investment. In addition, there is an increase in human capital investment even within firms with productivity exceeding $\underline{z}_1$. This can be attributed to the reallocation occurring across these firms, and leads to a higher probability of workers from low-productivity firms transitioning to high-productivity firms within those exceeding $\underline{z}_1$. Therefore, this dynamic intensifies the incentive for matches between workers and these firms to elevate the investment in human capital. The disparity in the skill accumulation effect between the second and third columns is more substantial than the difference in human capital investment, which is attributed to the concavity of human capital accumulation.

1.5 Conclusion

This study aims to investigate the impact of reform regarding firing costs in Germany on human capital investment. I conduct an empirical analysis using a difference-in-differences framework and leverage matched employee-employer data. The findings reveal that the reduction of firing costs for smaller establishments resulted in an increase in their training compared with larger establishments. Building on this observation, I develop an on-the-job search model that considers endogenous skill accumulation and firing costs. The model indicates that the reform in Germany results in a 0.58% increase in aggregate productivity, with 0.49% arising from human capital accumulation and the remaining 0.09% from reallocation.

There are limitations to my analyses. The model's tractability comes at the expense of ignoring firm size. Potential future research includes developing a tractable framework for tracking human capital for multiple workers within a single firm and incorporating human capital investment in the analysis of Hopenhayn and Rogerson (1993). In addition, I acknowledge that my paper does not account for some potential benefits associated with firing costs. It also does not discuss the effect on inequality, as explored in recent papers such as by Doepke and Gaetani (2024) and Daruich et al. (2023). Moreover, firing costs may mitigate the cost of job displacement during economic downturns (e.g., Schmieder et al. 2023). Although the paper provides new insights into the negative impacts of firing costs, a comprehensive evaluation must also take these benefits into account.

CHAPTER 2

Inventor Mobility, Knowledge Diffusion, and Growth

2.1 Introduction

Inventors play an essential role in both innovation within firms and knowledge diffusion between firms, which are important sources of economic growth. As Arrow (1962) stated "mobility of personnel among firms provides a way of spreading information," the mobility of inventors between firms has been considered an essential source of knowledge diffusion between firms.[1] Thus, policies related to labor markets for inventors are likely to have a significant impact on the firm productivity and economic growth.

This study provides an endogenous growth model to analyze the market for inventors and its impact on firm productivity and economic growth. In our model, inventors play dual roles: (i) they participate in in-house R&D efforts, enhancing the firm's technology, and (ii) they facilitate the transfer of knowledge from their former employers to their new ones when they change jobs. To quantify the model, we utilize data on inventors and patents linked to administrative labor market career information about inventors and their employing establishments in Germany. With these data, we document three novel empirical observations regarding inventors' job transitions, wage changes accompanying these transitions, and the influence of inventor inflows on the future innovation activity of recruiting firms. Finally, we discipline the model to align with these empirical findings and explore the consequences of

[1] For evidence from recent studies, see, Jaffe et al. (1993); Almeida and Kogut (1999); Song et al. (2003); Hoisl (2007); Rosenkopf and Almeida (2003); Breschi and Lissoni (2009); Singh and Agrawal (2011); Kaiser et al. (2015); Rahko (2017); Braunerhjelm et al. (2020).

inventor labor market policies.

In the theoretical part, we introduce an endogenous growth model that features the labor market where inventors and firms interact, and knowledge spills over across firms via inventor job transitions. Heterogeneous firms offer job openings, considering the knowledge diffusion from the inventors' prior employers, and inventors and firms match randomly in a frictional labor market. We focus on on-the-job search given our interest in how the inter-firm mobility of inventors influences knowledge spillover. This model is the first endogenous economic growth model that considers the endogenous job flows of inventors across firms and the knowledge diffusion through the job flows of inventors. The model's strength lies in its ability to endogenously generate both net and gross job flow of inventors and knowledge spillovers resulting from these flows, which are responsive to economic conditions and policy changes.

The empirical section documents novel findings for the job flows of inventors and their consequences. First, we examine the patterns of the mobility of inventors — defined as workers who have created patents — using inventor biography data from Germany, which links labor market biographies and their employing establishments recorded in the German social security data to patent register data. We find that a large proportion of inventors move to less productive establishments. This result is robust to the use of different productivity measures: establishment size, average wage, and number of patent citations. This job flow pattern of inventors is in contrast to general workers, who are more likely to move to more productive firms, as established in previous literature (e.g., Haltiwanger et al. (2018)). Moreover, we find that the wages of inventors grow more than those of general workers when they change jobs. This finding suggests that firms compensate for knowledge diffusion when they hire new inventors.

Then, we investigate how inventors' job flows influence the knowledge base of establishments, as measured by patents. We find that when a larger proportion of inventors comes from establishments with a more extensive knowledge base, the knowledge base of establishments grows faster over the next three to five years. Furthermore, we apply an instrumental variable to inventor flows and obtain significant results with the same sign as in the OLS specification. These results suggest the presence of knowledge diffusion through inventor flows.

In the quantitative section of this paper, we calibrate the model to match the key characteristics of the joint distribution of German inventors and firm dynamics observed in the microdata. We show that the calibrated model fits the target and non-target moments well, confirming that the model is well-suited to study counterfactual exercises.

We initially apply the calibrated model to conduct comparative statics analyses on matching efficiency. The model suggests that a decrease in matching efficiency, followed by a reduction in inventor mobility, leads to a decline in the economic growth rate. According to INV-BIO data, inventor mobility in Germany has been diminishing since the 1990s. Similarly, Akcigit and Goldschlag (2023a) report a decline in inventor mobility in the U.S. beginning in the early 2000s. Consequently, our model offers a framework for understanding the relationship between the observed decrease in inventor mobility and the deceleration of aggregate productivity growth in developed countries over recent decades.

Finally, we analyze the transition dynamics in our model to evaluate the effects of labor market policies on inventors. A key issue for those overseeing innovation policy is identifying which firms should be granted subsidies. In this context, we investigate the shift from an initial Balanced Growth Path (BGP) without subsidies to a new BGP with subsidies directed at technologically frontier firms. Frontier firms are characterized as those ranking in the upper half of the productivity distribution, with weighting based on the number of inventors. In the short term, subsidies to frontier firms reduce aggregate output by impeding the mobility of inventors from these leading firms to less advanced ones, thus hampering knowledge transfer. In contrast, over the long term, this policy boosts aggregate output by accelerating the growth rate at the technological frontier. Therefore, the impact of targeted subsidies on specific groups of firms hinges on whether policymakers focus on short-term or long-term economic effects.

Related Literature. Our paper is related to the literature on endogenous growth theory, particularly the diffusion of technology and knowledge, including Luttmer (2007), Lucas (2009), Lucas and Moll (2014), Perla and Tonetti (2014), Akcigit et al. (2018), Buera and Oberfield (2020), Shi and Hopenhayn (2020), Benhabib et al. (2021), and Prato (2022). Buera and Lucas (2018) surveys this topic. Perla and Tonetti (2014) and Lucas and Moll (2014) advanced the literature by modeling agents who choose to invest in technology diffusion. This approach enables the investigation of incentives, externalities, and welfare-improving policies. Our formulation of the knowledge diffusion function is based on the semi-endogenous growth model proposed by Buera and Oberfield (2020), which investigates international knowledge diffusion. Their model provides a micro-foundation for the knowledge diffusion function and expresses knowledge diffusion as a synergy of novel ideas and insights drawn from others. Similar to ours, Benhabib et al. (2021) and Shi and Hopenhayn (2020) address the interaction between R&D innovation and knowledge diffusion. In particular, our model applies the firms' innovation process formulated by Benhabib et al. (2021) to generate a realistic stationary productivity distribution. In the technology diffusion literature, our work is most closely related to Akcigit et al. (2018), who explicitly model inventors and analyze their role in knowledge diffusion among inventors. We depart from this literature by focusing on knowledge diffusion among firms due to inventor mobility. Moreover, we introduce a new perspective to this literature by incorporating labor market frictions, emphasizing the interaction between inventors and firms.

An expansive body of empirical research supports the concept of knowledge spillovers facilitated by job transitions of inventors. In one of the first such studies, Almeida and Kogut (1999) show that locations with greater intraregional labor mobility between firms tend to have more localized knowledge flows. Song et al. (2003) illustrate that mobile inventors build upon ideas from their previous firm more often than other inventors at the hiring firm. Rosenkopf and Almeida (2003) analyzes firm pairs, showing that those with higher labor mobility also have greater subsequent knowledge flow. These pioneering studies have inspired further research to facilitate our understanding of the connection between job transitions of inventors and knowledge spillovers (Hoisl, 2007; Breschi and Lissoni, 2009; Singh and Agrawal, 2011; Kaiser et al., 2015; Rahko, 2017; Braunerhjelm et al., 2020). Mawdsley and Somaya (2016) provide a review of these studies[2]. While most studies in this field struggle with limitations related to drawing causal inferences, some papers address the endogeneity problem. Singh and Agrawal (2011) employ a difference-in-differences approach to compare pre-move and post-move citation rates for poached inventors' previous and comparable control patents, concluding that acquiring firms intensify their use of inventions from the inventors' previous employers. Kaiser et al. (2015) use lagged mobility and industry mobility averages as instrumental variables for inventor mobility, uncovering a significantly positive impact of incoming inventors on their new employers' patent activity. Our paper is the first to integrate these insights into an endogenous growth model, emphasizing the interaction of inventor mobility and knowledge diffusion across firms. Furthermore, our study is novel in that it compares the patterns of job changes and the associated wage changes between inventors and general workers, providing evidence that suggests knowledge transmission and compensation for it.

Our paper also relates to the literature on frictional labor markets. In particular, our study benefits from recent developments in the modeling of multi-worker firms and on-the-job search, including Schaal (2017), Elsby and Gottfries (2021), and Bilal et al. (2023). The contemporary presence of on-the-job search and a non-constant return to scale revenue function in employment makes, in general, the firm problem intractable because we need to track the distribution of wages within each firm. To address the intractability, we assume

[2] A related area of study is the relationship between geography and knowledge diffusion. Early research by Jaffe et al. (1993) suggested a higher probability of cited patents originating from the same location as the citing ones. Breschi and Lissoni (2009) further improved this approach by introducing inventor mobility as a control, revealing that spatial proximity's effect on knowledge diffusion is cut by more than half. This suggests that the critical role of geography in knowledge transfer primarily results from inventors seldom relocating across regions.

that a firm posts a privately efficient number of vacancies, following Bilal et al. (2023). This assumption reduces the state variables to firm productivity and the number of inventors, thereby rendering the model tractable. Based on Bilal et al. (2022), Bilal et al. (2023) presented an endogenous growth model where the productivity distribution of incumbents determines the productivity of entrant firms. This model introduces an endogenous growth rate akin to Luttmer (2007). However, their model abstracts away the knowledge spillovers through worker mobility and its implication for economic growth.

Herkenhoff et al. (2018) and Shi (2023) explore models wherein knowledge diffuses among firms or colleagues via worker mobility. However, these studies consider models where firms employ only one or two workers at most. In contrast, our model allows firms to hire an arbitrary number of inventors unless it is profitable. Furthermore, while these papers examine more generalized workers, we restrict our focus to inventors and investigate the impact on economic growth.

The rest of the paper proceeds as follows. Section 2.2 describes the theory. Section 2.3 introduces the data and empirical results. Section 2.4 presents the calibration of the model and the quantitative policy counterfactual. Section 2.5 concludes.

2.2 Model

This section introduces an endogenous growth model featuring the role of the labor market, where inventors and firms match, and knowledge diffusion across firms due to inventors' job flows. Time is continuous, and the horizon is infinite. Inventors play two roles: (i) they engage in R&D activities in the firm to which they belong; (ii) when they switch jobs, they transfer knowledge from their previous employer to the new one, thereby enhancing productivity—this is referred to as knowledge diffusion. Inventors are homogeneous, except in terms of which firm they belong to[3]. We focus on on-the-job search since we are interested in the effect of the inter-firm mobility of inventors on knowledge diffusion. Inventors and firms are randomly matched in a frictional labor market. Firms make hiring decision by internalizing the marginal benefits of their contributions through internal R&D and knowledge diffusion. As we will discuss later, the state variable of a firm is summed up to the productivity of the firm, z, and the number of inventors employed, n. We construct a BGP equilibrium where aggregate variables grow at a constant rate g and inventor and firms' productivity distributions are stationary. Section 2.4 presents the transition dynamics.

Household

The representative household is composed of $\mathtt{n}$ individuals who supply inelastically one unit of time to the labor market for inventor. The size of the population is constant. Individuals work as inventors and receive wage payments from their firms. There is full insurance within the family, and thus the household problem can be split into a choice of aggregate consumption and a stage where the consumption is distributed across household members. The latter stage is irrelevant to labor market dynamics, so we focus on the former. The household discounts the future at the rate ρ. It derives utility from consumption, which we assume is logarithmic:

$$\int_0^\infty e^{-\rho t} \log \hat{C}(t) dt.$$

Variables with hats indicate that they are variables before detrending. We assume that the household trades shares in a mutual fund that owns all firms in the economy and trades a risk-free bond in zero net supply. As is standard, this implies that firms discount future payoffs at a constant risk-free rate $r(t) = \rho + g(t)$ in equilibrium on a BGP.

[3]Since this study considers constant wage contracts, wages can differ even among inventors who belong to the same firm. However, as we will discuss later, it is not necessary to track the distribution of wages within firms when characterizing the equilibrium

Production Technology

There is a unit mass of a continuum of firms. These firms produce a homogeneous product. Each firm has heterogeneous productivity $\hat{z}$. For simplicity, firm output equals firm productivity. As we will discuss later, in equilibrium in this model, firm productivity support is finite, and a maximum value of firm productivity exists. Let $\bar{z}(t)$ denote the maximum productivity of any firm, which we interpret as the technology frontier.

Matching Technology

Each firm employs a continuum of inventors n. Firms and inventors meet in a frictional labor market. Let $\hat{Z}(t)$ denote the aggregate productivity. A firm pays a cost $c(v)\hat{Z}(t)$ to post v vacancies. The cost function $c(v)$ is increasing and concave, and satisfies $c(0) = 0$ and $c'(0) = 0$. We focus on on-the-job search and assume that firms cannot lay off inventors, and inventors cannot voluntarily quit their jobs. Therefore, there are no unemployed inventors. Each vacancy randomly matches at a rate of A with an inventor who is working at other firms. For simplicity, we assume that the vacancy matching rate A is exogenous and does not depend on labor market tightness. An inventor meets a firm at rate $A\mathtt{v}$ where $\mathtt{v}$ is the total number of vacancies. An inventor incurs no cost of the search. As the vacancy cost is multiplied by $\hat{Z}(t)$, the vacancy cost grows as the economy grows. The rationale for this assumption is that as the economy grows, the price of resources for the vacancy (e.g., wages for human resources) also grows at the same rate.

Evolution of Firms' Productivity

We assume that firms' productivity changes due to the following three reasons: (i) innovation, (ii) knowledge diffusion, and (iii) leapfrog.

Innovation. The productivity of firms, with productivity $\hat{z}$ and inventor count n, increases at a rate of $\gamma(n)\hat{z}$, where $\gamma(\cdot)$ is an increasing and concave function. Consequently, the rate of productivity growth attributed to in-house R&D innovation tends to be higher for firms that employ a larger number of inventors.

Knowledge diffusion. When a firm with productivity $\hat{z}$ poaches an inventor from a firm with productivity $\hat{z}'$, the poaching firm's productivity increases by $\alpha(\hat{z}'/\hat{z})\hat{Z}(t)$ where $\alpha(\cdot)$ is an increasing and concave function. Therefore, a firm gains more knowledge when it poaches an inventor from a firm with higher relative productivity. While better insights lead to higher growth, the concavity of $\alpha(\cdot)$ implies that if the productivity difference between the poaching and poached firm is large, it becomes difficult for the poaching firms to utilize that knowledge.

Leapfrog. Finally, following Benhabib et al. (2021), we assume that firms can leapfrog to the frontier of the productivity distribution $\overline{z}(t)$ with an arrival rate $\eta > 0$. The possibility that firms can leap to the technology frontier represents an opportunity for the innovation process to yield significant insights rather than just steady incremental progress. This assumption establishes a stationary distribution with an upper bound on productivity for each period. The existence of this upper bound in the productivity distribution is crucial, as it ensures that the effect of knowledge diffusion does not become overly pronounced.

Contractual Environment

The contemporary presence of random search, on-the-job search, and a non-constant return-to-scale revenue function in employment generally makes the firm problem intractable. This is because computing optimal retention and vacancy policies requires keeping track of the entire wage distribution. Following Bilal et al. (2022), we make two assumptions regarding the contractual environment, ensuring that the state vector consists only of firm size and productivity.

Assumption 1. (**Bertrand Competition**) In a meeting with an employed inventor, the two firms Bertrand compete through a sequential auction. First, the poaching firm makes a take-it-or-leave-it wage offer. Second, the targeted firm makes a take-it-or-leave-it counteroffer to the worker. Finally, the inventor decides.

Assumption 2. (**Privately Efficient Vacancy Posting**) The firm posts a privately efficient number of vacancies, which is the one that maximizes the sum of the values of the firm and its workers.

We also assume that the information structure is such that everything relevant to payoffs is observable by both firms and inventors. Thus, we rule out private information by assumption.

While Assumption 1 is standard in the on-the-job search literature, Assumption 2 might be viewed as somewhat stringent. The latter assumption is necessary to simplify the model's analytical characterization and quantitative analysis. Under these assumptions, decisions made by both the employer and employees are privately efficient, as if they were maximizing their total joint value. As a result, the state variables of the joint value function are reduced to firm size and productivity. Hence, there is no need to track each firm's wage distribution to determine equilibrium allocations.

Distributions and Aggregate Variables

Let $\hat{F}(\hat{z}, n, t)$ be the cumulative distribution function of firms such that

$$1 = \int d\hat{F}(\hat{z}, n, t).$$

We assume that the total mass of firms in the economy is normalized to one. xThe distribution should also satisfy the inventor market clearing condition:

$$\mathtt{n} = \int n d\hat{F}(\hat{z}, n, t).$$

Let $\hat{v}(\hat{z}, n, t)$ be the amount of vacancy a firm $(\hat{z}, n)$ post at time t. The total mass of vacancy is given by

$$\mathtt{v(t)} = \int \hat{v}(\hat{z}, n, t) d\hat{F}(\hat{z}, n, t).$$

Because firms produce a homogeneous product, the aggregate output is given by

$$\hat{Z}(t) = \int \hat{z} d\hat{F}(\hat{z}, n, t).$$

Let $\hat{f}(\hat{z}, n, t)$ be a density of $\hat{F}(\hat{z}, n, t)$. Let define employment-weighted density

$$\hat{f}_n(\hat{z}, n, t) \equiv \frac{n\hat{f}(\hat{z}, n, t)}{\mathtt{n}}$$

and $\hat{F}_n(\hat{z}, n, t)$ the corresponding cumulative distribution. Also, let define vacancy-weighted distributions

$$\hat{f}_v(\hat{z}, n, t) = \frac{\hat{v}(\hat{z}, n, t)\hat{f}(\hat{z}, n, t)}{\mathtt{v}}$$

and $\hat{F}_v(\hat{z}, n, t)$ the corresponding cumulative distribution.

Condition for Successful Poaching

Define the poaching indicator function $\hat{\mathbb{1}}_P$ that takes 1 if the poaching successes and takes 0 otherwise. Let $\hat{\Omega}(\hat{z}, n, t)$ denote the joint value of an organization composed of a firm with productivity $\hat{z}$ and its n inventors at time t. Then, the poaching indicator function is expressed as

$$\hat{\mathbb{1}}_P(\hat{z}, n, \hat{z}', n', t) = \begin{cases} 1 & \text{if} \quad \hat{\Omega}_n(\hat{z}, n, t) + \alpha(\hat{z}'/\hat{z})\hat{Z}(t)\hat{\Omega}_z(\hat{z}, n, t) > \hat{\Omega}_n(\hat{z}', n', t) \\ 0 & \text{otherwise} \end{cases}$$

The first term $\hat{\Omega}_n(\hat{z}, n, t)$ is the derivative of the joint value with respect to n, which represents the change in the joint value resulting from an increase in the stock of inventors. This term captures the marginal contribution of the inventor to the firm's in-house R&D activity. The term $\alpha(\hat{z}'/\hat{z})Z\hat{\Omega}_z(\hat{z}, n, t)$ represents the change in the joint value resulting from an increase in firm productivity when the firm hires a new inventor. This term emerges because hiring a new inventor facilitates the transfer of ideas from the firm where the inventor previously worked. The poaching of the inventor is successful if the total marginal value of the inventor for the poaching firm $(\hat{z}, n)$ exceeds the value for the poached firm $(\hat{z}', n')$.

Hamilton-Jacobi-Bellman Equation

The following Hamilton-Jacobi-Bellman (HJB) equation determine the joint value $\hat{\Omega}(\hat{z}, n, t)$:

$$
\begin{aligned}
& r(t)\hat{\Omega}(\hat{z}, n, t) - \frac{\partial \hat{\Omega}(\hat{z}, n, t)}{\partial t} \\
&= \max_{\hat{v} \geq 0} \ \hat{z} - c(\hat{v})\hat{Z}(t) \\
&+ \underbrace{A\hat{v} \int \left[\hat{\Omega}_n(\hat{z}, n, t) + \alpha(\hat{z}'/\hat{z})\hat{Z}(t)\hat{\Omega}_z(\hat{z}, n, t) - \hat{\Omega}_n(\hat{z}', n', t)\right]^+ d\hat{F}_n(\hat{z}', n', t)}_{\text{Poaching Hire}} \\
&+ \underbrace{\gamma(n)\hat{z}\hat{\Omega}_z(\hat{z}, n, t)}_{\text{In-house R\&D}} \\
&+ \underbrace{\eta\left[\hat{\Omega}(\overline{z}, n, t) - \hat{\Omega}(\hat{z}, n, t)\right]}_{\text{Leapfrog}}
\end{aligned}
\tag{2.1}
$$

When a firm (z, n) hires a new inventor, the total value increases by $\hat{\Omega}_n(\hat{z}, n, t) + \alpha(\hat{z}'/\hat{z})\hat{Z}\hat{\Omega}_z(\hat{z}, n, t) - \hat{\Omega}_n(\hat{z}', n', t)$. The first and second term is the gain in value to the firm and incumbent inventors due to the new hire. The third term is the value the firm and incumbent inventors give the new inventor, which equals the highest value its former employer would pay to retain them. As mentioned earlier, the poaching is successful if this difference is positive.

Conversely, an incumbent inventor may quit and move to a higher marginal value firm. The firm and remaining inventors will lose $\hat{\Omega}_n(z, n, t)$ and are thus prepared to increase the inventor's value by $\hat{\Omega}_n(z, n, t)$ to retain them. Knowing this, the external firm hires the inventor by offering the inventor exactly $\hat{\Omega}_n(z, n, t)$. Therefore, the joint value of the firm, remaining inventors, and poached inventor are unchanged, and no "Poached Quit" term appears in (2.1).

Kolmogorov Forward Equation

The Kolmogorov forward equation (KFE) describes the evolution of the firm's distribution across productivity and the number of inventors. To characterize the KFE, we derive the drifts for changes in firm-level productivity and the number of inventors. The drift for the firm-level productivity change is given by

$$\hat{\mu}_z(\hat{z},n,t) \equiv \underbrace{\gamma(n)\hat{z}}_{\text{In-house R\&D}} + \underbrace{A\hat{v}(\hat{z},n,t)\hat{Z}(t)\int \hat{\mathbb{1}}_P(\hat{z},n,\hat{z}',n',t)\alpha(\hat{z}'/\hat{z})d\hat{F}_n(\hat{z}',n',t)}_{\text{Knowledge diffusion}}. \tag{2.2}$$

The first term on the right-hand side represents productivity growth due to in-house R&D. The second term accounts for the firms' productivity growth resulting from knowledge diffusion. When firms post $\hat{v}$ vacancies, these vacancies match with $A\hat{v}$ inventors. Owing to the randomness of the matchings, the original firms of these inventors are taken from the inventor-weighted firm distribution $\hat{F}_n$. If a poaching attempt is successful ($\hat{\mathbb{1}}_P(\hat{z},n,\hat{z}',n',t) = 1$), the productivity of the poaching firm increases by $\alpha(\hat{z}'/\hat{z})\hat{Z}(t)$. Note that our definition of $\hat{\mu}_z(\hat{z},n,t)$ does not include changes in productivity due to leapfrogging, and we need to include an additional term to incorporate leapfrogging in the KFE equation, which we will describe below.

The drift for the change in the number of inventors is determined by

$$\hat{\mu}_n(\hat{z},n,t) \equiv \underbrace{A\hat{v}(\hat{z},n,t)\int \hat{\mathbb{1}}_P(\hat{z},n,\hat{z}',n',t)d\hat{F}_n(z',n',t)}_{\text{Poaching hire}} - \underbrace{A\mathtt{v}\frac{n}{\mathtt{n}}\int \hat{\mathbb{1}}_P(\hat{z}',n',\hat{z},n,t)d\hat{F}_v(\hat{z}',n',t)}_{\text{Poached by other firms}} \tag{2.3}$$

The first term on the right-hand side illustrates the increase in the number of inventors owing to poaching hires from other firms, while the second term represents a decrease in the number of inventors as they are poached by other firms.

Given the above definition of $\hat{\mu}_z(\hat{z},n,t)$ and $\hat{\mu}_n(\hat{z},n,t)$, the KFE is presented as

$$\frac{\partial}{\partial t}\hat{f}(\hat{z},n,t) = -\underbrace{\frac{\partial}{\partial n}\left(\hat{\mu}_z(\hat{z},n,t)\hat{f}(\hat{z},n,t)\right)}_{N \text{ of inventor change}} - \underbrace{\frac{\partial}{\partial \hat{z}}\left(\hat{\mu}_n(\hat{z},n,t)\hat{f}(\hat{z},n,t)\right)}_{\text{Productivity change}}$$
$$\underbrace{-\eta\hat{f}(\hat{z},n,t) + \eta\int_0^{\overline{z}}\hat{f}(\hat{z},n,t)d\hat{z}\hat{\boldsymbol{\delta}}(\overline{z})}_{\text{Leapfrog}} \tag{2.4}$$

where $\hat{\boldsymbol{\delta}}(\overline{z})$ is the Dirac delta function, which is zero everywhere except $\hat{z} = \overline{z}$ where it is infinite and satisfies $\int \hat{\boldsymbol{\delta}}(\overline{z})dz = 1$.

Technology Frontier

Here, we argue that the technology frontier is finite, and we characterize its growth rate. If $\overline{z}(0) < \infty$, then $\overline{z}(t)$ will remain finite for all t. This is because it evolves from the firms' productivity growth in the interval infinitesimally close to $\overline{z}(t)$, and the firm's growth rate is finite. Furthermore, the growth rate of the technology frontier is determined by the productivity growth rate of firms that possess the highest growth rate among those at the technology frontier. This is because these firms will be at the technology frontier in the next instant. The following lemma formally characterizes the productivity growth rate of the technology frontier:

Lemma 1. (Growth Rate of the Technology Frontier) *If $\overline{z}(0) < \infty$, then $\overline{z}(t) < \infty \ \forall t < \infty$ and*

$$g(t) \equiv \frac{\overline{z}'(t)}{\overline{z}(t)} = \max_{n \in \left\{n \ | \ \hat{f}(\overline{z}(t),n,t)>0\right\}} \frac{\hat{\mu}_z(\overline{z}(t),n,t)}{\overline{z}(t)}$$

Normalization

In the following, we examine economies in a BGP equilibrium, where the distribution remains constant when appropriately scaled, and aggregate output experiences constant growth. It is convenient to transform this system into a set of stationary equations for computing BGP equilibria. While we could standardize using any variable that grows at the same rate as the aggregate economy, it is expedient to normalize variables relative to the technology frontier $\overline{z}(t)$. Define the normalized values and functions as follows:

$$z \equiv \hat{z}/\overline{z}(t)$$

$$Z(t) \equiv \hat{Z}(t)/\overline{z}(t)$$

$$\Omega(z,n,t) = \Omega(\hat{z}/\overline{z}(t),n,t) \equiv \hat{\Omega}(\hat{z},n,t)/\overline{z}(t) \tag{2.5}$$

$$F(z,n,t) = F(\hat{z}/\overline{z}(t),n,t) \equiv \hat{F}(\hat{z},n,t) \tag{2.6}$$

$$\mathbb{1}_P(z,n,z',n',t) = \mathbb{1}_P(\hat{z}/\overline{z}(t),n,\hat{z}'/\overline{z}(t),n',t) \equiv \hat{\mathbb{1}}_P(\hat{z},n,\hat{z}',n',t) \tag{2.7}$$

$$v(z,n,z',n',t) = v(\hat{z}/\overline{z}(t),n,\hat{z}/\overline{z}(t),n',t) \equiv \hat{v}(\hat{z},n,\hat{z}',n',t) \tag{2.8}$$

The technology frontier is normalized to $\overline{z}(t)/\overline{z}(t) = 1$. The above normalizations make the value functions, productivity distributions, and growth rates stationary.

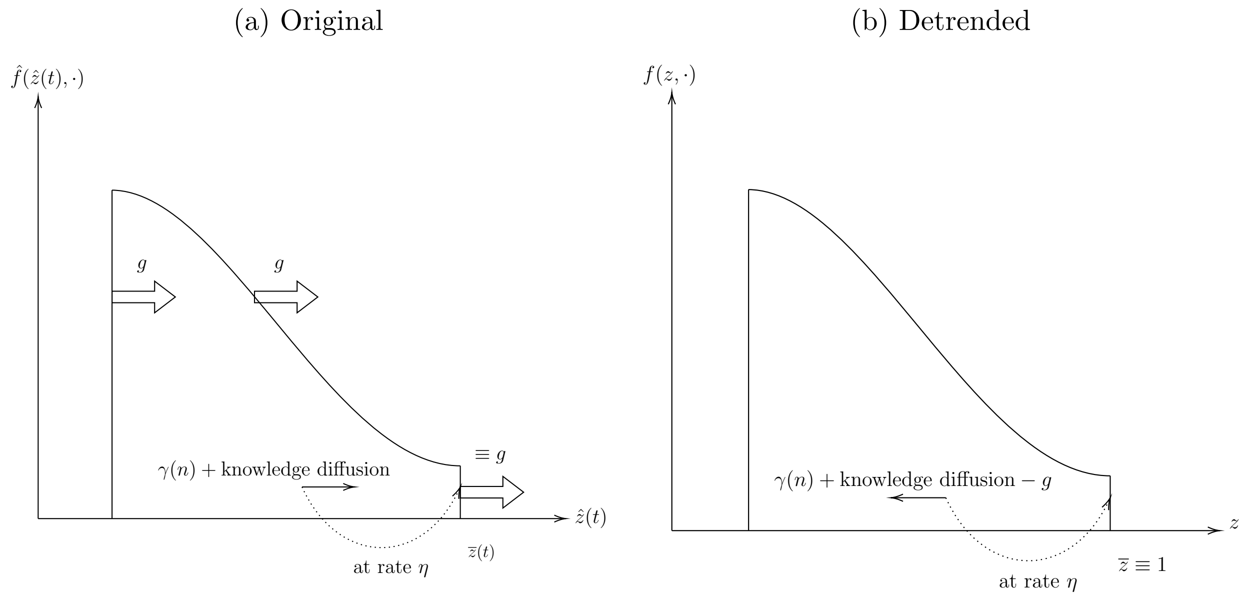
See the Figure 2.1 for an illustration of the original and detrended distributions. Nothing prevents the distribution from spreading without knowledge diffusion, driving the productivity variance to infinity. However, because of knowledge diffusion, as the distribution extends, productivity growth due to knowledge diffusion increases, and these forces compress the distribution.

Balanced Growth Path

Now, we describe a BGP equilibrium where aggregate productivity grows at a constant rate, and distributions are stationary. Define the growth rate of aggregate productivity to be $g_Z(t) \equiv \hat{Z}'(t)/\hat{Z}(t)$. That is, $g_Z(t) = g_Z$ and $F(z,n,t) = F(z,n)$ for all t. Aggregate output is given by

$$\hat{Z}(t) = \int \hat{z} d\hat{F}(\hat{z},n,t)$$

Figure 2.1: Marginal Density for Productivity of Firms



Notes: Illustration of original and detrended marginal distribution for firms' productivity on the BGP.

$$=\overline{z}(t)\int zdF(z,n,t)$$

On a BGP, the detrended productivity distribution is constant: $F(z,n,t) = F(z,n)$. Therefore, $g_Z = \hat{Z}'(t)/\hat{Z}(t) = \overline{z}'(t)/\overline{z}(t) = g$, and we obtain the following lemma:

Lemma 2. (Growth Rate of the Technology Frontier and Aggregate Productivity) *On a BGP, the aggregate productivity growth rate equals the technology frontier's growth rate. That is,* $g_Z = g$.

The following definition summarizes the characteristics of our BGP equilibrium.

Definition 1. (**Balanced Growth Path**) A *BGP equilibrium* consists of: (i) a joint value function $\Omega(z,n)$; (ii) a vacancy policy $v(z,n)$; (iii) a stationary distribution of firms $f(z,n)$; (iv) vacancy- and employment-weighted distributions $f_v(z,n)$ and $f_n(z,n)$; (v) poaching indicator function $\mathbb{1}_P(z,n,z',n')$; (vi) the aggregate productivity Z and the total vacancies v, and (vii) the economic growth rate g such that

1. The joint value $\Omega(z,n)$ satisfies the HJB equation

$$
\begin{aligned}
\rho\Omega(z,n) =& z - c(v(z,n))Z \\
&+ Av(z,n)\int\left[\Omega_n(z,n) + \alpha(z'/z)Z\Omega_z(z,n) - \Omega_n(z',n')\right]^+ dF_n(z',n') \\
&+ (\gamma(n)-g)\, z\Omega_z(z,n) \\
&+ \eta\left[\Omega(1,n) - \Omega(z,n)\right]
\end{aligned}
$$

2. The vacancy policy $v(z,n)$ satisfies the first order condition

$$
c_v(v(z,n))Z = A\int\left[\Omega_n(z,n) + \alpha(z'/z)Z\Omega_z(z,n) - \Omega_n(z',n')\right]^+ dF_n(z',n') \qquad (2.9)
$$

3. A density function $f(z,n)$ satisfies the KFE equation

$$
0 = -\frac{\partial}{\partial n}\left(\mu_n(z,n)f(z,n)\right) - \frac{\partial}{\partial z}\left(\mu_z(z,n)f(z,n)\right) - \eta f(z,n) + \eta\int_0^1 f(z',n)dz'\boldsymbol{\delta}(1)
$$

where the drift of the number of employed inventors $\mu_n(z,n)$ and productivity $\mu_z(z,n)$ are given by

$$
\begin{aligned}
\mu_n(z,n) \equiv& Av(z,n)\int \mathbb{1}_P(z,n,z',n')dF_n(z',n') - A\mathtt{v}\frac{n}{\mathtt{n}}\int \mathbb{1}_P(z',n',z,n)dF_v(z',n') \\
\mu_z(z,n) \equiv& (\gamma(n)-g)\, z + Av(z,n)Z\int \mathbb{1}_P(z,n,z',n')\alpha(z'/z)dF_n(z',n')
\end{aligned}
$$

4. Vacancy- and employment-weighted distributions are consistent:

$$
\begin{aligned}
f_v(z,n) &= \frac{v(z,n)f(z,n)}{\mathtt{v}} \\
f_n(z,n) &= \frac{nf(z,n)}{\mathtt{n}}
\end{aligned}
$$

5. Poaching indicator function $\mathbb{1}_P(z,n,z',n')$ is given by

$$
\mathbb{1}_P(z,n,z',n') = \begin{cases} 1 & \text{if} \quad \Omega_n(z,n) + \alpha(z'/z)Z\Omega_z(z,n) > \Omega_n(z',n') \\ 0 & \text{otherwise} \end{cases}
$$

6. The aggregate productivity Z and the total vacancies $\mathtt{v}$ rate are given by

$$Z = \int zdF(z,n)$$
$$\mathtt{v} = \int v(z,n)dF(z,n)$$

7. The inventor market clearing condition is satisfied:

$$\mathtt{n} = \int ndF(z,n)$$

Appendix 4.3 comprehensively derivates the normalized system. We also establish some properties of the joint value function in the Appendix. In it, we show that the following properties hold: (i) Ω is increasing in productivity: $\Omega_z > 0$; (ii) Ω is increasing in the number of inventors: $\Omega_n > 0$.

The equilibrium of the model is solved numerically in Section 2.4. Before that, we turn to the description of the empirical results.

2.3 Data and Empirical Findings

In this section, we investigate the job flows of inventors — workers who have created patents — between establishments using inventor biography data from Germany. The results provide motivation for our model, and we use these results to discipline the numerical model, as explained in Section2.4.

2.3.1 Data

Our analyses utilize two administrative data sets, ”Linked Inventor Biography Data 1980–2014” (INV-BIO) and ”Sample of Integrated Labor Market Biographies” (Stichprobe der Integrierten Arbeitsmarktbiografien — SIAB).[4]

[4]More detailed information is presented in Appendix 4.4.

The INV-BIO data combines labor market biographies recorded in the German social security data (Integrated Employment Biographies — IEB) with patent register data from the European Patent Office (EPO). This data set tracks information about 152,350 inventors who have registered their patents to the EPO from 1980 to 2014. The information includes their unique ID, age, gender, level of education, daily wage, and the number of citations received by the patents associated with each inventor in the EPO's records. The data also contains information about the establishments employing the inventors, such as the establishment ID, the total number of their employees, and the mean daily wage of their full-time employees. The advantage over patent-based datasets used in previous studies (e.g., EPO patent data by Akcigit et al. (2018)) is that we can use social security information to keep track of inventors' flows even when they are not creating patents.

The SIAB data is a 2% random sample from IEB. This data set contains the same information about individuals and their employing establishments as INV-BIO, except for patent-related information. In the absence of the patent data, we identify inventors in SIAB using a 3-digit occupation code, as described in Section 2.3.3. The data set covers 3,322,316 individuals from 1980 to 2019.

Since merging datasets is not allowed, we use the two datasets separately for each analysis: when comparing the movement patterns of inventors and other workers in Section 2.3.3, we use SIAB, which includes both, but otherwise we use INVBIO, a dataset focused exclusively on information about inventors.

2.3.2 Inventor Flows in INV-BIO

First, we adopt an approach similar to Haltiwanger et al. (2018) to characterize inventor flows using INV-BIO. We assign each establishment to a percentile rank according to patent information or productivity measure. We then compute the transition probabilities of inventor flows between these ranks.[5]

We utilize three different measures as proxies for the knowledge quality or productivity level.[6] The first measure is based on the forward citations for patents that establishments have created. Measuring patent quality through forward citations is widely employed in the literature about patent creation (e.g., Pakes (1986); Hall et al. (2001); Akcigit et al. (2018)). In particular, Akcigit et al. (2018) measures the idea quality of inventor teams based on the number of forward citations their patent receives. Similarly, our measure for an establishment e in year t, z_{et} is given by:

$$z_{et} = \frac{\sum_{j=-2}^{0} \text{citations}_{et+j}}{3}, \tag{2.10}$$
$$\text{where citations}_{et} = \sum_{i} \text{citations}_{it} \times \frac{n_{ie}}{n_i}.$$

citations$_{it}$ denotes the count of forward citations that occur five years after year t for patent i, which is created by a team including inventors employed at establishment e. Note that the team developing the patent can consist of inventors from different establishments. n_{ie} represents the number of inventors at establishment e in the team, while n_i represents the total number of inventors in the team, including those affiliated with different establishments. We multiply citations$_{it}$ by n_{ie}/n_i to adjust for the contribution made by inventors from establishments other than e.[7] Therefore, citations$_{et}$ is the count of five-year forward citations for patents that establishment e created, adjusting for the contributions of other establishments. Following Akcigit et al. (2018), we use the three-year backward average as the measure. The

[5] Establishments could be classified into different percentiles based on the measure each year. The ranks of the origin and destination establishments are determined based on the measure from the previous year, preceding the movement of inventors.

[6] We assume that the knowledge quality and productivity level are positively correlated. In fact, the three measures are positively correlated with each other as described in Appendix 4.4.

[7] In other words, we start by dividing the count of forward citations by the total number of inventors involved in the team for each inventor's patents. Afterward, we aggregate these values for the inventors who are employed at establishment e.

other measures are the number of employees (establishment size) and the mean wage of full-time employees, following standard practice in the literature, as summarized by Moscarini and Postel-Vinay (2018).[8]

Table 2.1: Transition Probabilities of Inventor Flows

(A) Rank by Citation/Inventor

Share of flows (%)		Destination establishment rank				
		≤ 50%	50-60	60-70	70-80	80-100
	≤ 50%	2.3	0.2	0.3	0.4	4.3
Origin	50-60	1.7	0.2	0.2	0.3	3.0
establishment	60-70	1.9	0.2	0.3	0.3	3.6
rank	70-80	2.2	0.2	0.2	0.4	4.2
	80-100	19.5	2.0	2.4	3.4	46.7

(B) Rank by Establishment Size

Share of flows (%)		Destination establishment rank				
		≤ 50%	50-60	60-70	70-80	80-100
	≤ 50%	2.6	0.9	0.7	0.8	6.3
Origin	50-60	0.4	0.5	0.6	0.4	2.3
establishment	60-70	0.5	0.2	0.7	0.9	3.0
rank	70-80	0.6	0.3	0.4	1.3	4.8
	80-100	5.8	2.5	3.4	4.7	55.7

Notes: Detailed description is presented below the panel (C) in the next page.

Table 2.1 shows the transition probabilities of inventor flows from origin to destination.[9] There is a substantial movement of inventors from higher ranks to lower ranks, denoted by

[8]On-the-job search models with heterogeneous productivity firms (e.g., Postel-Vinay and Robin (2002)) predict that more productive firms offer higher wages and attract more workers, leading to their growth in size.

[9]Appendix 4.4 shows the distribution of inventors according to each of the three measures. It reveals a notable concentration of inventors within specific establishments. Irrespective of the type of measure, more than half of the inventors are found in establishments ranked above the 80th percentile, and only approximately 10 percent belong to establishments below the 50th percentile. This aligns with Akcigit and Goldschlag (2023b)'s finding that inventors are concentrated in large incumbents in the U.S.

Table 2: Inventor Flows across Establishments (cont.)

(C) Rank by Mean Wage

Share of flows (%)		Destination establishment rank				
		≤50%	50-60	60-70	70-80	80-100
	≤ 50%	2.9	1.1	1.0	1.1	3.5
Origin	50-60	0.8	0.9	1.1	0.9	2.3
establishment	60-70	1.0	0.9	2.0	2.1	4.1
rank	70-80	1.4	1.0	1.7	3.6	7.3
	80-100	5.5	3.6	5.4	7.2	37.6

Notes: This table shows transition probabilities of inventor flows across percentiles of establishments. The inventors staying in the same establishment are excluded. The percentile rank in panel (A) is based on the three-year backward average of forward patent citation counts. Panel (B) is based on the number of employees, and panel (C) is based on the mean wage of full-time employees. Establishments could be classified into different percentiles based on these measures each year. The ranks of the origin and destination establishments are determined based on the measure from the previous year, preceding the movement of inventors. The sample encompasses data from 1980 to 2014. The values in the table represent the proportion of inventor flows in each cell in relation to the total flows in INV-BIO.

the red-colored cells. The sum of values in these red cells amounts to 33.7% in panel (A), 18.8% in panel (B), and 28.5% in panel (C). This suggests that a large portion of inventors move from higher-ranked establishments to lower-ranked ones.[10] Appendix 4.4 shows that this pattern is observable even when the sample is limited to job flows accompanied by wage increases.

This pattern is not found in previous literature on worker flows. For example, Haltiwanger et al. (2018) construct transition probabilities of worker flows based on the mean wage of firms, and they observe a higher probability of flows to higher ranks compared to lower ranks. This discrepancy suggests that the tendency for many flows to lower ranks is a distinctive

[10]Another observable pattern is that the values along the diagonal are considerably high, particularly in the bottom right of each panel: 49.9% in panel (A), 60.8% in panel (B), and 50.8% in panel (C). This indicates that many inventors tend to move within the same rank, especially within the highest rank. This can be observed in the literature on worker flows (e.g., Haltiwanger et al. (2018)).

characteristic specific to inventors.

2.3.3 Inventor Flows in Comparison with Worker Flows

Next, we compare inventor flows with worker flows. We utilize the SIAB for the comparison since INV-BIO lacks information on workers other than inventors.

To identify inventors within SIAB, we use a 3-digit occupation code. We find that the majority of inventors in INV-BIO are affiliated with specific occupations, each with their corresponding shares: research and development (20.2%), machine-building and operations (19.8%), mathematics, biology, and physics (19.1%), and mechatronics, energy, and electronics (18.8%). These four occupations account for nearly 80% of the inventors in INV-BIO. We thus consider workers in these four occupations within SIAB to likely be inventors.

Table 2.2: Identified Inventors in SIAB and Inventors in INV-BIO

Summary statistics		SIAB		INV-BIO
(1980 - 2014)		Workers	Identified inventors	Inventors
Daily wage, Euro	Mean	59.0	78.9	156.2
	S.D.	47.2	52.1	30.0
Age	Mean	38.7	38.4	42.4
	S.D.	12.9	12.4	9.0
Females, %		45.2	14.8	5.7
N of obs., thousand		21,344	2,871	420

Notes: This table compares the summary statistics between workers in SIAB and the inventors in INV-BIO. Identified inventors in SIAB are workers who work in the following four occupations: "research and development", "machine-building and operations", "mathematics, biology, and physics", and "mechatronics, energy, and electronics." The worker in the table includes the identified inventors. The summary statistics are calculated using a pooled sample with daily wage, age, and gender filled in.

Table 2.2 presents a comparison of summary statistics between the two data sets. The mean daily wage of workers in the four occupations (identified inventors) in SIAB falls between that of workers in SIAB and that of inventors in INV-BIO. Furthermore, the proportion of female workers among the identified inventors lies between the two groups. These findings suggest that our identified inventors include both actual inventors and a portion of non-inventor workers. Therefore, the result of the subsequent comparison between workers and identified inventors should be considered conservative due to the presence of attenuation bias.

We estimate the following Probit model:

$$P(D_{it} = 1) = \Phi(\beta_0 + \beta_1 I_{it} + \beta_2 X_{it}) \tag{2.11}$$

Individual i is the job changer without an unemployment spell. I_{it} serves as a dummy for the inventors, taking a value of one if individual i works in one of the four occupations in year t, and zero otherwise. D_{it} equals one if individual i moves from a more productive establishment to a less productive one in year t, and zero if the move is to a more productive establishment. Note that we investigate the moving between establishments rather than ranks here. For constructing D_{it}, we use the number of employees or mean wage as a proxy for productivity. The vector of control variables, X_{it}, includes age, a square of age, gender, and educational attainment. To avoid the incidental parameter problem, we estimate the model without incorporating fixed effects.[11] The function Φ is the cumulative distribution function of the standard normal distribution.

The coefficient of our interest is β_1. The positive β_1 implies that inventors are more likely to move to less productive establishments than other workers. Standard errors (SEs) are clustered by destination establishment and year, supposing the presence of persistent establishment and year specific shocks.

Table 2.3 presents results. The first column uses the establishment size as the productivity measure, and the second column uses the mean wage as the measure. The results show that the probability of an inventor transitioning to an establishment with fewer employees or a

[11] The estimated value of β_1 in the linear model with the fixed effects is also significantly positive as in Table 2.3. See Appendix 4.4.

Table 2.3: Estimation Result for Inventor Flows

	(1) $P(D_{it}=1)$				(2) $\Delta \log w_{it}$	
	Whole sample		Sample with wage ↑			
I_{it}	.077***	.036***	.052***	.012***	.017***	.021***
	(.004)	(.004)	(.004)	(.004)	(.005)	(.004)
D_{it}					-.078***	-.084***
					(.006)	(.005)
$D_{it} \times I_{it}$					.016***	-.002
					(.006)	(.006)
Control	√	√	√	√	√	√
Fixed Effects					√	√
Measure for D_{it}	Size	Mean wage	Size	Mean wage	Size	Mean wage
N	3,572,567	3,533,344	2,082,939	2,060,714	859,888	859,861
Adj. R^2	.019	.016	.005	.003	.13	.13

Notes: Control variables include age, a square of age, gender, and educational attainment. Fixed effects include year, year × industry, and destination establishment fixed effects. I_{it} equals to one if individual i works in one of the four occupations ("research and development", "machine-building and operations", "mathematics, biology, and physics", and "mechatronics, energy, and electronics") in year t, and zero otherwise. D_{it} equals one if individual i moves to a less productive establishment in year t, and zero otherwise. The productivity measure is based on the establishment size or the mean wage in year $t-1$. The sample spans from 1980 to 2019. SEs clustered by year and establishments are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

lower mean wage is higher than other workers, implying that inventors are more likely to move to a less productive establishment than other workers. Moving to the third and fourth columns, we narrow down the sample to job changers who experience wage increases, and the coefficients for I_{it} are still significantly positive. This suggests that many inventors move to less productive establishments and experience wage increases.

To further examine the association between the direction of flows and wages, we run the following regression:

$$\log w_{it} - \log w_{it-1} = \beta_0 + \beta_1 D_{it} + \beta_2 I_{it} + \beta_3 D_{it} I_{it} + \beta_4 X_{it} + \alpha + \varepsilon_{it} \tag{2.12}$$

The variable w_{it} represents the daily wage of individual i after a job change, while w_{it-1} represents the wage before the job change. The vector of fixed effects α includes year, year $\times$ industry, and destination establishment fixed effects. The definition of other variables remains the same as in the equation (2.11).

The last two columns of Table 2.3 show the estimation results. The coefficients for I_{et} are significantly positive, indicating that inventors experience greater wage increases by around 2% through job changes than other general workers.

The coefficients of D_{it} are negative, meaning that workers tend to experience fewer wage increases when moving to less productive establishments. However, the coefficients for $D_{it} \times I_{it}$ are significantly positive in the fifth column. The positive coefficient implies that inventors experience fewer wage decreases by moving to less productive establishments than general workers.

Knowledge transfer with inventor mobility has the potential to explain these results. That is, an inventor who worked in a high-productivity establishment can transfer that knowledge to a low-productivity establishment when changing jobs. Therefore, establishments are more willing to poach inventors from more productive establishments than other general workers and compensate inventors for the benefits.

2.3.4 Empirical Evidence of Knowledge Diffusion

The result in the previous section suggests the presence of knowledge diffusion via inventor flows. This section further investigates how the inventor flows influence the productivity growth of establishments.

Our specification is given by:

$$\log z_{et+j} - \log z_{et} = \beta_0 + \beta_1 \textit{H-Share}_{et} + \beta_2 X_{et} + \alpha_e + \alpha_t + \varepsilon_{et} \tag{2.13}$$

z_{et} is the knowledge quality of establishment e in year t, as defined in (2.10) of Section 2.3.2. We use three, four, or five-year forward citations for z_{et}. The variable $\textit{H-Share}_{et}$ represents the percentage share of inventor inflows from establishments with higher knowledge base measured by patent citations to total inventor inflows to establishment e. $\log n_{et}$ is the log of the number of inventors. The vector of control variables X_{et} includes the log of the establishment size, number of inventors, mean wage, and z_{et}. The vector of fixed effects α includes year, year $\times$ industry, and establishment fixed effects. The equation (2.13) is estimated using INV-BIO from 1980 to 2019. Standard errors (SEs) are clustered by destination establishment and year.

The results are reported in Table 2.4. The table shows that when more inventors come from productive establishments, the knowledge growth of the poaching establishments is higher over a period of three, four, or five years.

However, the coefficient of $\textit{H-Share}_{et}$ is susceptible to the endogeneity problem. The unobservable expectation for $\log z_{et+j} - \log z_{et}$ can be correlated with the realized $\log z_{et+j} - \log z_{et}$ and $\textit{H-Share}_{et}$. To address this issue of omitted variable bias[12], we employ an instrumental variable (IV) strategy. In this approach, we utilize the patent citation rank for establishments in their states from the previous year (referred to as $\textit{Regional Rank}_{et-1}$) as an instrument for $\textit{H-Share}_{et}$. In the first stage, the $\textit{Regional Rank}_{et-1}$ is expected to be highly correlated with $\textit{H-Share}_{et}$. A lower knowledge rank indicates a higher number of establishments with greater knowledge level in the state. Consequently, the share of inventors poached from these highly knowledgeable establishments ($\textit{H-Share}_{et}$) is more likely to be higher. The instrument can

[12]This omitted variable bias can have either an upward or a downward effect. If there is an expectation of high productivity growth, establishments may offer higher wages to attract more inventors from highly productive establishments, leading to an upward bias. On the other hand, if lower productivity growth is anticipated, establishments may attempt to offset the lower growth by poaching inventors from more productive establishments, resulting in a downward bias.

Table 2.4: Estimation Result for Knowledge Growth

	$\log z_{et+j} - \log z_{et}$				
	$j = 3$	$j = 4$	$j = 5$	$j = 3$	$j = 3$
H-Share$_{et}$ (%)	.0022***	.0028***	.0027***	.0024***	.0024***
	(.0004)	(.0004)	(.0004)	(.0003)	(.0003)
Control	√	√	√	√	√
Fixed Effects	√	√	√	√	√
Citation	3y fwd	3y fwd	3y fwd	4y fwd	5y fwd
N	24,625	22,270	19,982	26,791	26,451
Adj R^2	.21	.27	.36	.23	.25

Notes: Control variables are z_{et}, log of a number of employees and mean wage. Fixed effects include year, year × industry, and establishment fixed effects. The sample spans from 1980 to 2014. z_{et} is the forward citation measure (backward 3-year moving average). *H-Share*$_{et}$ is the share of the number of inventors who moved from establishments with a higher $z_{e't-1}$ at $t-1$. If there are no inflows, *H-Share*$_{et}$ is set to zero. SEs clustered by year and establishment are reported in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, .

be considered to satisfy the exclusion condition when taking into account the fixed effects and control variables. To address the issue of mean reversion in knowledge quality, we add z_{et-1} as one of the control variables.

Table 2.5 shows the result using the IV. In the first stage, we find a significant correlation between *Regional Rank*$_{et-1}$ and *H-Share*$_{et}$. Specifically, if the knowledge level is relatively lower within the state (indicated by a higher value of *Regional Rank*$_{et-1}$), *H-Share*$_{et}$ tends to be higher.

The impact of poaching from more knowledgeable establishments is highly significant and even larger than the results obtained from the OLS in Table 2.4. In the first column in Table 2.5 using the three-year forward citations, 1% increase in the share of inventors from more productive establishments increases citations by around 10% relative to the unconditional mean. In sum, our results suggest that establishments can enhance their knowledge growth

Table 2.5: Estimation Result for Knowledge Growth using IV

	$\log z_{et+j} - \log z_{et}$				
	$j=3$	$j=4$	$j=5$	$j=3$	$j=3$
H-Share$_{et}$ (%)	.092***	.101***	.103***	.084***	.088***
	(.007)	(.007)	(.009)	(.005)	(.006)
Control	√	√	√	√	√
Fixed Effects	√	√	√	√	√
Citation	3y fwd	3y fwd	3y fwd	4y fwd	5y fwd
First Stage IV					
Regional Rank$_{et-1}$	24.4***	22.7***	22.8***	30.0***	29.3***
	(1.7)	(1.5)	(1.8)	(1.7)	(1.7)
N	22,213	20,052	17,996	23,137	23,609
F statistic	204.8	232.6	155.2	302.7	286.2

Notes: Control variables are the log of the establishment size, number of inventors, mean wage, z_{et}, and z_{et-1}. The sample spans from 1980 to 2014. z_{et} is the the forward citation measure at t. *H-Share*$_{et}$ is the share of the number of inventors who moved from establishments with a higher z at $t-1$. *Regional Rank*$_{et-1}$ is the establishment's rank of z_{et-1} among all establishments in the state (16 states). A higher rank means lower z_{et-1}. The rank is normalized so that the maximum is equal to 1. SEs clustered by year and establishment are reported in parenthesis. *** $p < 0.01$, *** $p < 0.05$. The first stage F value is the Cragg-Donald Wald F statistic.

by recruiting inventors from high-productivity establishments.

2.4 Quantitative Analysis

This section quantifies the effects of inventor job flows and knowledge transfers on innovation and productivity and studies the effects of counterfactual policy. To do this, we calibrate the model from Section 2.2 to match the data described in Section 2.3. We subsequently demonstrate that the calibrated model closely fits the data for targeted moments, and we use it to examine the effects of policies related to the labor market for inventors.

2.4.1 Stochastic Process for In-House R&D Ability and Functional Forms

In this section, we introduce a stochastic process to characterize the unpredictable nature of in-house R&D ability, following the model presented in Benhabib et al. (2021). We set the functional form of the in-house R&D function to $\gamma(n, i) = \overline{\gamma}_i n^{\delta}$, where the index i represents the in-house R&D ability, which can either be high (h) or low (l). The innovation ability i follows a two-state Markov process. The R&D capacity is greater when in the h state than in the l state ($\overline{\gamma}_h > \overline{\gamma}_l$). The transition intensity (the rate at which the R&D ability changes) from the l to h state is denoted by λ_l, and the transition intensity from the h to l state is denoted by λ_h. We employ this two-state Markov process to formulate the stochastic innovation process because it allows the composition of firms at the technology frontier to change over time, which is essential for the existence of a stationary distribution on a BGP.

We will conduct our numerical exercises by calibrating the model with high transition rates. The characteristics of the stochastic process with conditional draws are similar to those with unconditional draws when the switching rates are high, as estimated in our calibrations.

For other functional forms, we assume that the vacancy cost function is $c(v) = \frac{\overline{c}}{\phi+1} v^{\phi+1}$. The knowledge diffusion rate function is $\alpha(z'/z) = \overline{\alpha}\,(z'/z)^{\beta}$ such that the knowledge diffusion rate is increasing in the productivity of the poached firm relative to the poaching firm.

2.4.2 Calibration

We calibrate the model along a BGP equilibrium to match features of the allocation of inventors across establishments and characteristics of inventor job flows between establishments.

Table 2.6: Parameter Values

Parameter	Description	Value
— Panel A. Externally Set or Normalized —		
ρ	Discount rate	0.0041
$\overline{z}$	Frontier productivity	1
$\mathtt{m}$	Measure of firms	1
$\overline{c}$	Vacancy cost coefficient	100
ϕ	Vacancy cost elasticity	3.45
$\overline{\gamma}_l$	l-type R&D coefficient	0
— Panel B. Direct Match to Data —		
$\mathtt{n}$	Measure of inventors	5
λ_h	Jump intensity: $h \to l$	0.02
λ_l	Jump intensity: $l \to h$	0.01
— Panel C. SMM Calibration —		
β	Diffusion curvature	0.33
$\overline{\alpha}$	Diffusion rate	0.0012
$\overline{\gamma}_h$	h-type R&D coefficient	0.0006
η	Leapfrog	0.0001
δ	R&D curvature	0.35
A	Matching efficiency	0.26

Notes: List of model parameters and calibrated values. In Panel C, all parameters are calibrated jointly for the SMM calibration.

Externally Set or Normalized

We normalize or set to standard values six parameters, as summarized in the Panel A of Table 2.6. The discount rate ρ implies an annual real interest rate of 5%. The first-order condition for vacancies implies that we cannot identify $\overline{c}$ and A separately, so we normalize $\overline{c}$. We normalize the productivity of technology frontier $\overline{z}$ and the measure of firms $\mathtt{m}$ to 1 without loss of generality. We also set l-type R&D coefficient $\overline{\gamma}_l$ to zero without loss of generality[13]. We use the value of the vacancy cost elasticity calibrated by Bilal et al. (2022).

Direct Match to Data

We set three parameters to directly match the moments from German inventor data, as summarized in Panel B of Table 2.6. The measure of inventor $\mathtt{n}$ is determined by the average number of inventors per establishment, given a unit measure of firms.

The transition rate of innovation ability λ_h and λ_l match the estimations from the two-state Markov transition matrix for the growth rate of the knowledge in establishments near the technology frontier in the spirit of Benhabib et al. (2021). Details are as follows. The knowledge in establishments is measured by patent citation, and we designate establishments within the top 10% of the 5-year forward citation measure, Z_{et} in Section 2.3.2, as the frontier each year. Among these, establishments exhibiting positive growth rates of the citations are categorized as being in the high state, whereas the remaining are seen as being in the low state. Based on this, we estimate the transition matrix.

Internal Calibration Using SMM

We estimate the six key parameters of the model listed in the Panel C of Table 2.6. These parameters are captured by the vector: $\Theta = \{\beta, \overline{\alpha}, \overline{\gamma}_h, \eta, \delta, A\}$ and estimated by minimizing the objective function

$$\mathcal{L}(\Theta) = (\hat{m} - m(\Theta))' W^{-1} (\hat{m} - m(\Theta))$$

[13] See Benhabib et al. (2021)

where $\hat{m}$ is a vector of empirical moments and $m(\Theta)$ are their model counterparts. The diagonal components of matrix W have the same weights. All non-diagonal components are zero. In the case of distributional information, the weights are adjusted so that the weights of the entire distribution add up to one. For example, the unconditional distribution of inventors is characterized by five quantile points and therefore weights 1/5. All non-diagonal components are zero. [14]

2.4.3 Results

Table 2.7: Targeted Moments

Moments	Data	Model
EE rate (%, monthly)	1.17	1.13
Growth rate (%, monthly)	0.16	0.13
Distribution of inventor by firm ranking	Figure 2.2a	
Distribution of inventor flow by poaching firm ranking	Figure 2.2b	

Notes: EE rate and growth rate are both monthly frequencies; EE rate is calculated by dividing the number of job changers by the number of inventors in the INV-BIO data.

Table 2.7 and Figure 2.2 summarize the target moments and parameters. Not only do the macro moments (EE rate and growth rate) in Table 2.7 provide a good fit, but Figure 2.2 also shows that the joint distribution of inventors and firms is well replicated. Overall, despite over-identification, the fit of the moments within the internal calibration is reasonably good.

Although the parameters are calibrated jointly, we will discuss the most relevant moments for each parameter. First, the EE rate primarily provides information about the matching

[14]Prior to estimation, we did the first 100 iterations for each of the six parameters and excluded regions that did not converge. Then the parameter space we explore is as follows: β ranges from 0.1 to 0.7, $\overline{\alpha}$ ranges from 0.001 to 0.003, $\overline{\gamma}_h$ ranges from 0.0001 to 0.001, η ranges from 0.00005 to 0.0003, δ ranges from 0.2 to 0.5, and A ranges from 0.05 to 0.3. For each of these parameter spaces, we take 20 grids and compute 2000 Halton grids.

efficiency A, which governs the size of the job flow. The growth rate has information mainly on γ_h. Both α and β are related to knowledge diffusion, where β controls for the sensitivity of knowledge diffusion to the difference in productivity between poaching firms and incumbents, and α adjusts the average size of knowledge diffusion. These parameters predominantly determine poaching firms' distribution over productivity in Figure 2.2b. Finally, Figure 2.2a, representing the relationship between productivity and inventors, primarily informs δ and η.

Next, we discuss the fit of the calibrated model for a important non-targeted regression result. In our empirical analysis, using equation (2.13), we find that the greater the share of inventor inflows from higher knowledge firms, the greater the productivity gains of the poaching firms. We examine whether the model can replicate this relationship. The corresponding equation of the model is given by

$$\hat{\mu}_z(\hat{z}, n, t) = \beta_0 + \beta_1 \textit{H-Share} + \beta_2 \log n + \varepsilon$$

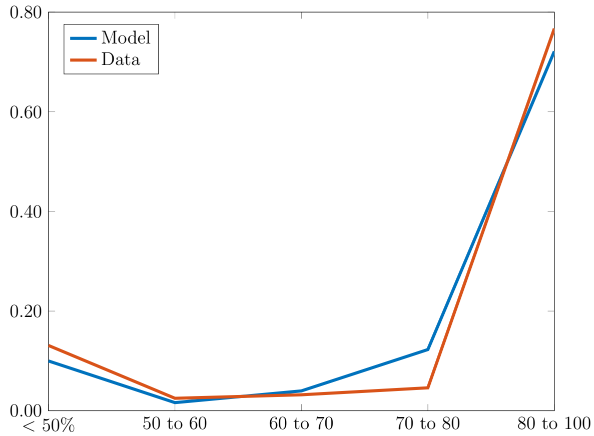
where $\hat{\mu}_z(\hat{z}, n, t)$ is the growth rate of productivity and *H-Share* is the fraction of $\hat{\mu}_n(\hat{z}, n, t)$ who transitioned from firms with higher productivity. We compute the coefficients using weighted least squares, with the density function of the firms $f(\hat{z}, n, t)$ in each grid as the sample. The model-implied coefficient β_1 is 0.003, well within the range of the OLS and IV estimates in Table 4, thereby successfully producing a reasonable quantitative magnitude.

2.4.4 Quantitative Exercises

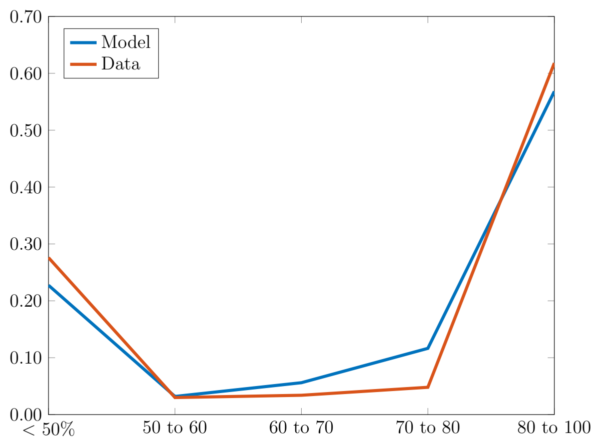
The previous section demonstrated that the calibrated model accurately aligns with the data for both targeted and non-targeted moments. Consequently, the model is well-suited for conducting counterfactual analyses. Initially, we will assess the impact of changes in matching efficiency in frictional labor markets for inventors. Subsequently, we will investigate the ramifications of a hypothetical policy intervention. In this policy analysis, we suggest a hypothetical policy that offers subsidies to frontier firms, aiming to foster innovation within

Figure 2.2: Inventor Distributions by Firm Productivity

(a) Density Weighted by the Mass of Inventors



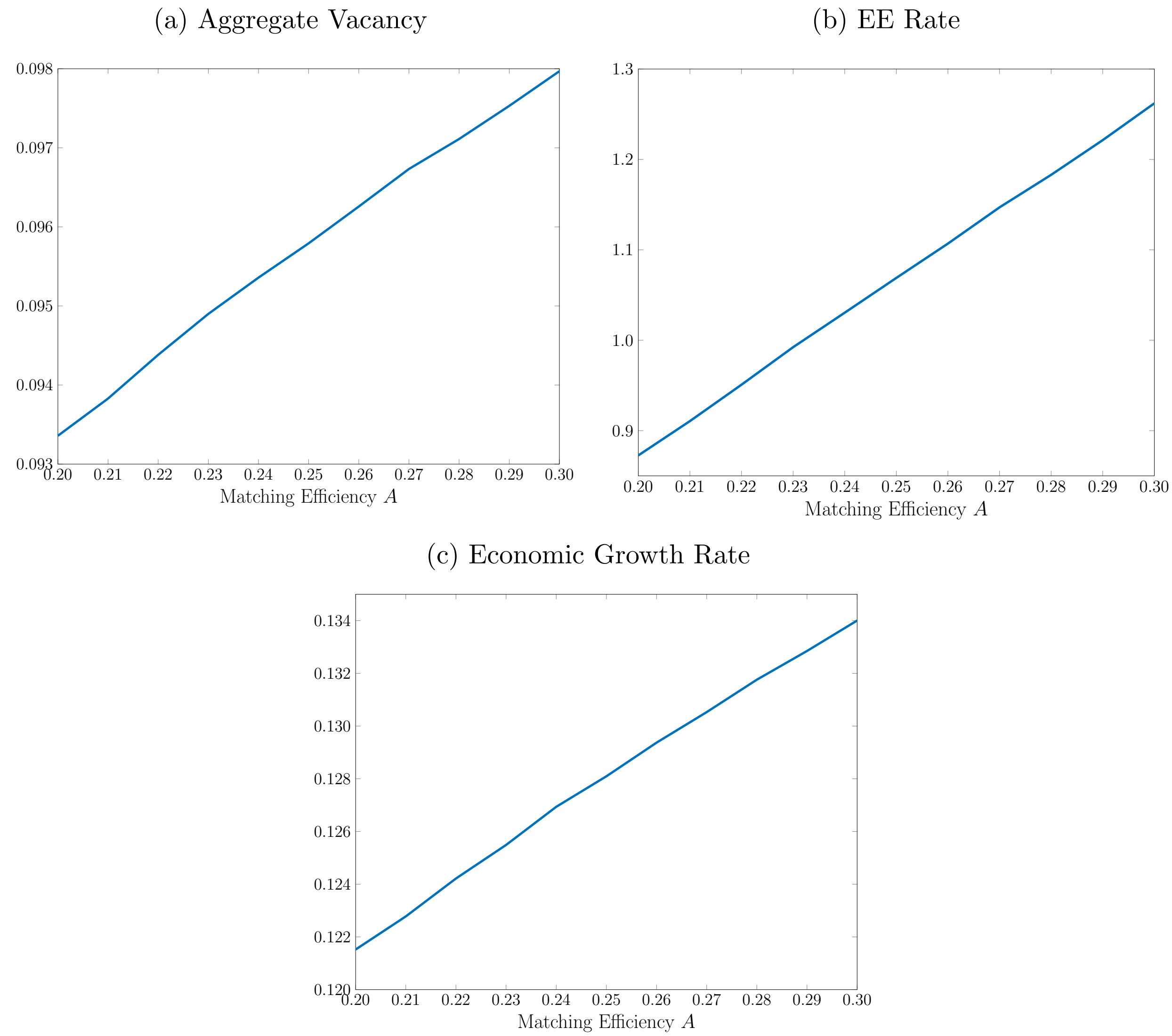
(b) Density Weighted by Inventor Inflows



Notes: The data in (a) and (b) are plots of the distribution in Table B.3. (A) and the marginal distribution of poaching firms in Table 2.1 (B). The corresponding model values are calculated using the productivity of the firms, the inventor and their joint density $f(\hat{z}, n, t)$ under the calibrated parameters.

these entities.

Figure 2.3: Comparative Statics: Aggregate Variables

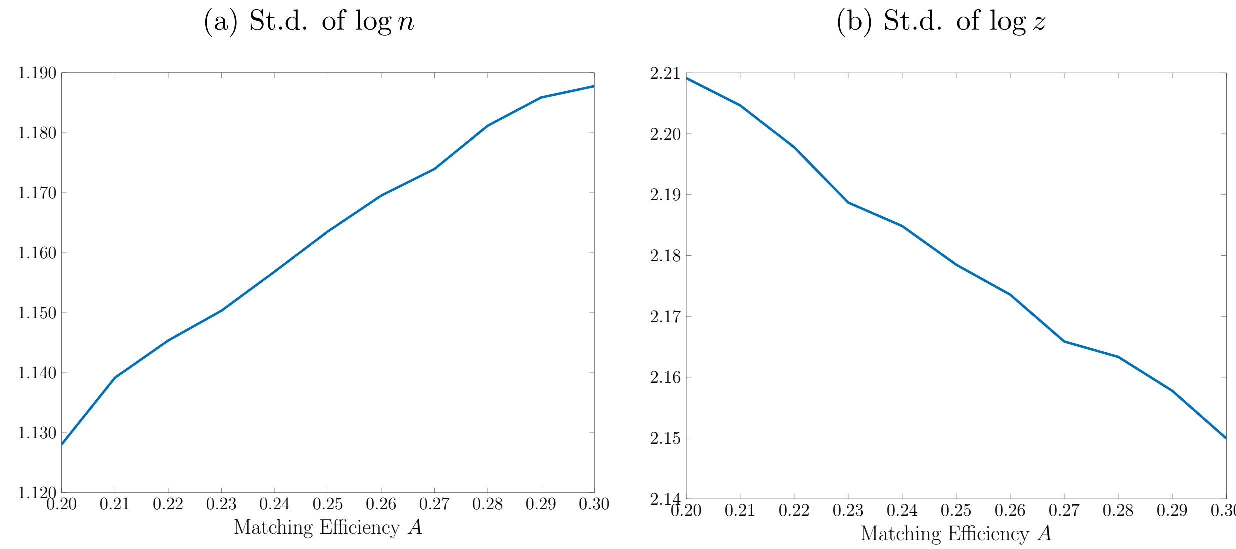


Notes: The figure display the comparative statics of varying matching efficiency A.

Quantifying the Impact of Inventor Market Frictions

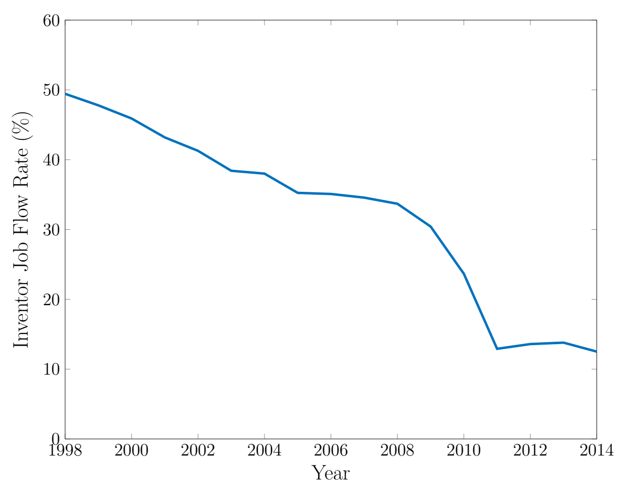
How do changes in frictions within the market for inventors impact the economy? Frictions in this market can emerge from a variety of sources, such as search and matching frictions, regulations related to labor, and agreements between employees and firms. To assess the impact of shifts in market frictions for inventors, we conduct a comparative static analysis

Figure 2.4: Comparative Statics: Distribution



Notes: The figure display the comparative statics of varying matching efficiency A.

Figure 2.5: Inventor Job Flow Rate in Germany



Notes: This figure shows the job flow rate for German inventors, calculated using INV-BIO data.

focusing on the matching efficiency parameter A.

As Figure 2.3 (a) shows, in our model, an increase in matching efficiency increases the vacancy posting of firms. Also, as Figure 2.3 (b) shows, higher matching efficiency and more vacancy postings lead to higher job-to-job transition rates for the inventor.

As matching efficiency increases, the economic growth rate increases (Figure 2.3 (c)) through the following mechanism. First, as the job-to-job transition rate increases, knowledge diffusion becomes more active. As a result, the dispersion of firm productivity decreases (Figure 2.4 (b)). Note that high-productivity firms grow mainly through in-house R&D, while low-productivity firms grow mainly through knowledge diffusion (See equation (2.2)). Since lower variance in firm productivity reduces growth through knowledge diffusion, it leads to a relative increase in inventor hiring by more productive firms. As a result, more inventors are attracted to firms in the technology frontier. As shown in Lemma 2, the economic growth rate in a BGP is determined by the productivity growth rate of frontier firms. Therefore, the economic growth rate also increases (Figure 2.3 (c)).

Our model links the observed decrease in inventor mobility in Germany and the US to the low economic growth in developed countries in recent years, as documented in the secular stagnation literature (e.g., Summers (2014); Eggertsson et al. (2019); Akcigit and Ates (2021)). In the INV-BIO data, we find a significant decline in the inventor job flow rate over time (Figure 2.5). Similarly, Akcigit and Goldschlag (2023a) document that inventor mobility has decreased since the early 2000s. Our model indicates that the decline in the matching efficiency, which subsequently leads to a reduction in job-to-job transitions, contributes to a lower economic growth rate. Thus, our model provides a framework linking the observed decrease in inventor mobility to the slowdown in aggregate productivity growth in developed countries over the past few decades.

In our model, changes in the distribution of firms and inventors play an essential role in the mechanism linking inventor mobility and economic growth rates, and indeed, the changes in the distributional characteristics of the model are consistent with the data. The first row of Table 2.8 shows the coefficient of variation of the number of inventors working in each establishment. Here, we compare the values for 1998, the first year of Figure 2.5, and 2014, the last year. Similarly, the second row shows the coefficient of variation of the

Table 2.8: The Change in the Second-Order Moments of the Distribution in the Data

	1998	2014
Coefficient of variation of N of inventors	0.93	1.12
Coefficient of variation of productivity	4.34	3.14

Notes: The first column shows the change in the coefficient of variation of the number of inventors working in each establishment in the INV-BIO data. We compare the values for 1998, the first year of Figure 2.5, and 2014, the last year. Similarly, the second column shows the change in the coefficient of variation of the establishments' innovativeness measured by patent citation.

establishments' innovativeness measured by patent citation. The direction of these changes in the data is consistent with the direction of change in the variance of z and the variance of n in our model (Figure 2.4) when the matching efficiency A decreases.

Policy Exercises: Subsidy for Firms Near the Technology Frontier

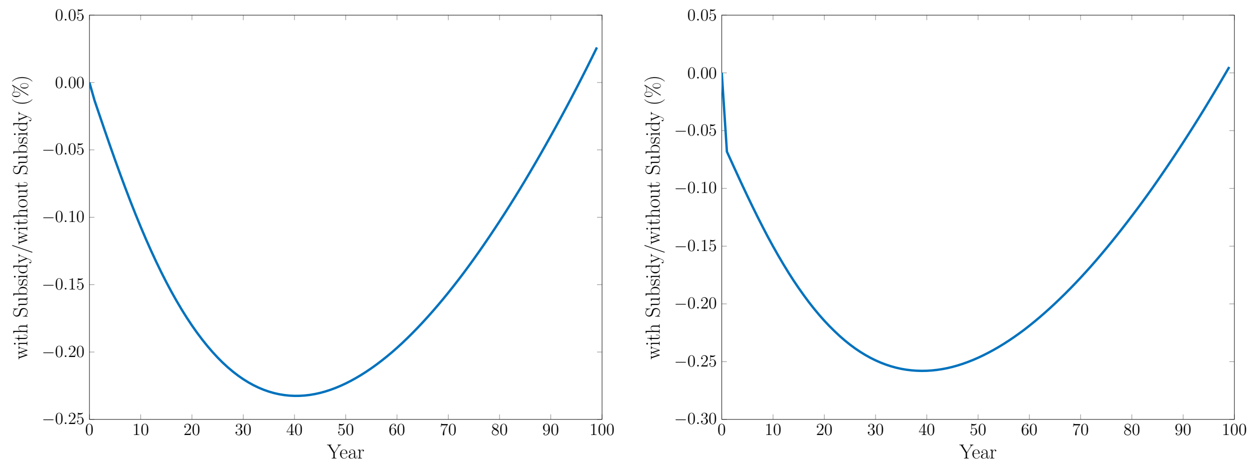
This section analyzes the consequences of subsidies for firms near the technology frontier. Policy-makers may consider encouraging R&D activity by offering subsidies to technologically progressive firms. Alternatively, they may wish to subsidize technologically lagging firms to promote knowledge diffusion through the movement of inventors. To find the more beneficial policy, we examine the transition from an initial BGP without subsidy to a new BGP with a subsidy aimed at technologically frontier firms.

We investigate the transition from an initial BGP with no subsidy to a new BGP with a subsidy rate 10% for frontier firms. We define frontier firms as those with productivity z in the top half of the distribution, weighted by the number of inventors[15]. We assume subsidies are financed by a constant rate tax on the bottom half of the distribution. This exercise

[15] In the distribution weighted by the number of inventors, the firms with productivity z in the top half of the distribution correspond to approximately the top 10% of firms with productivity in the unweighted distribution.

Figure 2.6: Transition Dynamics

(a) Output (b) Consumption



Notes: The figures display transitional dynamics on implementing a counterfactual subsidy for technologically frontier firms. Panel (a) shows the path for aggregate output relative to the old aggregate output. Panel (b) shows the path for aggregate consumption.

imposes a 17% tax on the production of these remaining firms. Subsidies and taxation changes are permanent. The agents do not anticipate the policy change until $t = 0$, and they are perfect foresight after $t = 1$.

The left panel of Figure 2.6 displays the output path relative to the baseline balanced growth path (BGP). Aggregate output decreases during the initial 40 years, but shows an increase in the long run. In the short run, this policy hinders job flows from technology-leading firms to laggard firms, thereby impeding knowledge diffusion to the latter. However, over the long term, it enhances the growth rate of the technology frontier, which ultimately fosters positive impacts on knowledge diffusion to laggard firms. Consequently, the rate of economic growth increases in the long term.

Lastly, we calculate the policy change's welfare effects by analyzing the economy's transitional dynamics, applying a discount rate ρ to future periods following policy implementation. Our results indicate a decline in welfare of 0.14%, measured in terms of consumption equivalent. This outcome is primarily driven by a short-term decline in aggregate productivity. However, this effect is largely offset by a long-term increase in productivity. Our analysis suggests that the effectiveness of a policy is dependent on the policymaker's time horizon.

2.5 Conclusion

This paper explores labor markets where inventors and firms interact, focusing on the implications for the distribution of inventors across firms, knowledge diffusion, and productivity growth. To examine these dynamics, we construct an endogenous growth model that incorporates frictions in labor markets for inventors. In our model, inventors (i) contribute to in-house R&D efforts, enhancing the firm's technology, and (ii) facilitate knowledge transfer from their previous employers to their new ones when they change jobs. Heterogeneous firms create job openings, considering the knowledge transfer from the inventors' prior employers. To quantify this framework, we utilize data on inventors and patents connected to administrative labor market career information about individuals and their employing establishments in Germany. We find three empirical findings: (i) inventors are more likely to transition to less productive establishments compared to general workers, (ii) inventors face significant increases in wages when changing jobs compared to general workers, and (iii) the number of patent citations increases more rapidly when a higher proportion of their inventors originates from higher productivity establishments. We calibrate our model to align with these empirical findings and demonstrate that the calibrated model closely fits both target and non-target moments, confirming its suitability for conducting counterfactual exercises. We then examine the transition from an initial Balanced Growth Path (BGP) without a subsidy to a new BGP with a subsidy targeting technologically frontier firms. In the short term, this subsidy decreases aggregate output by discouraging inventor mobility from frontier firms to laggard firms, thereby hindering knowledge diffusion. However, in the long term, the impact of the subsidy on output is reversed, as it enhances the growth rate of the technological frontier.

CHAPTER 3

Cyclicality and Asymmetry of the User Cost of Labor

3.1 Introduction

Wage cyclicality is a central question in macroeconomics as it is crucial in accounting for unemployment and other aggregate dynamics (e.g., Hall 2005; Christiano et al. 2005). The critical nature hinges on the premise that wage plays an allocative role in the labor market. However, observing this “allocative” wage in the data is challenging. Workers and firms often form a long-term employment contract, under which wages can be arbitrarily distributed across contract periods without changing the total value of the contract. In this situation, the cyclicality of period wages does not have meaningful implications (e.g., Barro 1977; Beaudry and DiNardo 1991).

To address this issue, Kudlyak (2014) proposes a concept of the user cost of labor (UCL). The UCL is the differential of the present discounted values of total labor costs at two points in time, analogous to the use cost of capital. For a firm that chooses whether it employs a worker in the current period or waits until the next, the UCL corresponds to the additional cost for using one unit of labor service in the current period—i.e., the marginal price of labor.

However, UCL’s cyclicality could be contaminated by cyclical upgrading of job-match quality (e.g., Gertler et al. 2020). Workers move from one job to another when a new offer is preferable to the existing match. This match upgrading is more likely to occur in booms (e.g., Barlevy 2002; Hagedorn and Manovskii 2013), which generates procyclicality in the average quality of new matches. Since the UCL is estimated from the wages of new matches over their tenure, the presence of procyclical match quality overstates the cyclicality of the UCL.

In this paper, we propose a novel methodology to control for the cyclical upgrading. Our central idea is to resort to the wage dynamics of new graduates. Since procyclical upgrading occurs because job changers can compare the quality of existing and new matches, new graduates who have not had jobs are plausibly assumed to be less affected by it. We implement this idea by exploiting unique Japanese wage data—the Basic Survey on Wage Structure (BSWS)—a nationally representative annual wage survey. As the BSWS records wages at each length of service, we can directly observe a sequence of wages from hiring until separation. We focus on the sample of new graduates to estimate the UCL that controls for the cyclical upgrading.

We use the estimated UCL series to investigate its cyclical properties. Our empirical findings are as follows. First, the wage cyclicality for new graduates is lower than for job changers. The semi-elasticity of new graduates' wages with respect to the unemployment rate is around half that of all newly hired workers, including job changers. This supports our assumption that new graduates are less affected by procyclical upgrading than job changers.

Second, the UCL remains highly procyclical after controlling for the cyclical upgrading. The semi-elasticity of the UCL with respect to the unemployment rate is around twice as large as that of the new-hire wage under the baseline specification. The result indicates that the UCL's high cyclicality arises from incumbent workers' wage rigidity. To see this, suppose that wages are persistently high (low) for workers hired in expansions (contractions). Their total labor cost reflects not only high (low) wages at the time of hiring, but also lasting wage differences across hiring cohorts over the course of their tenure. The UCL captures these dynamics of total labor cost.

Finally, the UCL's cyclicality is asymmetric. It rises in booms but it remains flat in recessions. Whereas asymmetry in wage adjustments—in particular, downward wage rigidity—has been widely reported for incumbent workers in the literature, it does not suffice for the asymmetric adjustments of the UCL if the new-hire wage is downwardly flexible. Indeed, we find that neither new-hire wage nor incumbent-worker wage declines in recessions in our dataset, causing the downward rigidity of the UCL. [1]

Theoretically, the UCL's higher cyclicality than the new-hire wage is in line with an implicit contract between risk-neutral firms and risk-averse workers (e.g., Harris and Holmstrom 1982; Beaudry and DiNardo 1991). In contrast, less is known about the asymmetry of its cyclicality.[2] Therefore, we propose a model to reconcile our empirical findings. Our benchmark is Rudanko (2009), who introduces an implicit contract to a directed search framework. Our extension involves productivity heterogeneity on both firm and worker sides. High-productivity firms seek to match with skilled workers, but they have limited ability to distinguish worker types. The setting gives rise to adverse selection, whereby unskilled workers may apply to high-productivity firms. In this setting, firms use wages as a screening tool to receive job applications only from the targeted type of workers. That is, firms keep the value of a posted contract high enough, even in recessions, that their labor market becomes too competitive for unskilled workers to apply to.

The model accounts for both UCL's overall high cyclicality and dampened responses in recessions. Consequently, it replicates asymmetric labor market dynamics consistent with the data, thereby generating relatively larger unemployment volatility. This result is in a stark contrast to existing studies on the UCL, which implies that high cyclicality of the UCL makes the unemployment volatility puzzle of Shimer (2005) more challenging to resolve (e.g., Kudlyak 2014).

[1]This finding is compatible with previous studies that have highlighted the downward flexibility of job changers' wages, because our new-hire wage is measured from firms' perspectives—i.e., a series of wages for new hires each year—whereas previous studies typically report downward flexibility at worker level—i.e., the wage changes of each worker before and after a job change. We argue that wages from firms' perspectives are relevant for measuring labor costs.

[2]A frequent explanation is a fairness constraint whereby the new-hire wage should be equal to the incumbent-worker wage (e.g., Gertler and Trigari 2009; Snell and Thomas 2010; Rudanko 2023). However, our empirical evidence suggests the presence of cohort wage differences.

Related Literature. This paper joins the wealth of literature on wage cyclicality. Its primary contribution is that it elaborates on a wage measure in the presence of a long-term contract. This issue was originally raised by Beaudry and DiNardo (1991) and recently tackled by Kudlyak (2014), who introduced the concept of the UCL; the UCL has been estimated by Kudlyak (2014), Basu and House (2016), Doniger (2021), and Bils et al. (2023).

Our analysis adds to this line of the literature by proposing a new method to address cyclical upgrading through job changes. In estimating the UCL, Basu and House (2016) and Doniger (2021) use a proxy for job-match quality proposed by Hagedorn and Manovskii (2013) on the premise that the labor market tightness in each worker's employment cycle represents the gain from cyclical upgrading during the period. We take a more direct approach by limiting the sample to new graduates assumed to be free from the upgrading.

Recent work by Bils et al. (2023) uses the long-run wage in a match as a measure of its quality, assuming that the impact of business cycles at the time of hiring on subsequent wages disappears at a sufficiently long horizon (8 years in their baseline analysis). Their UCL is based on the wages of workers who remain employed in the same firm for 8 years, who may be subject to selection bias to the extent that better matches last longer (e.g., Altonji and Shakotko 1987). Our approach does not impose an assumption on workers' tenure. We also verify that our sample of new graduates does not exhibit cyclical fluctuations in match quality when applying their methodology.

Gertler et al. (2020) use the sample of unemployment-to-employment flows in the Survey of Income and Program Participation (SIPP) to control for cyclical upgrading.[3] Our use of school-to-employment flows further addresses a concern regarding the potential selection bias of unemployed workers, including the possibility that workers entering unemployment in recessions may have left a poor match (e.g., Caballero and Hammour 1994; Bils et al. 2023). Moreover, the SIPP does not cover each worker's wages for a long enough period to

[3] To be clear, Gertler et al. (2020) do not estimate the UCL, but control for cyclical upgrading to examine the wage cyclicality of new hires.

estimate the UCL.

In addition, we introduce a new dataset, the BSWS, that tests for the robustness of previous studies' findings.[4] The UCL's high cyclicality after controlling for match-quality changes in our methodology is aligned with earlier findings by Kudlyak (2014), Basu and House (2016), and Bils et al. (2023) for the aggregate UCL, and Doniger (2021) for the UCL of workers with a bachelor's degree. We also provide evidence on the asymmetry of the UCL's cyclicality. Furthermore, whereas previous studies report contrasting results regarding the cyclicality of new-hire wages when controlling for job-match quality (acyclical: Basu and House 2016; cyclical: Bils et al. 2023), our study reveals significant cyclicality in new-hire wages.

This paper is related to a broader literature that studies various sources of composition biases over business cycles (e.g., Bils 1985; Solon et al. 1994; McLaughlin and Bils 2001). These include reallocation in recessions (e.g., Caballero and Hammour 1994) and booms (e.g., Barlevy 2002); firm-side characteristics (e.g., Carneiro et al. 2012); and skill mismatches (e.g., Figueiredo 2022). Our approach—focusing on new graduates—echoes Solon et al. (1997) and Martins et al. (2012), who limit their attention to entry-level jobs to lessen the variation in job-match quality.

At the same time, this paper revisits the cyclicality of new-hire and incumbent-worker wages. Numerous studies report greater wage cyclicality of new hires (e.g., a survey by Pissarides 2009). However, recent studies question this conclusion by controlling for worker productivity (Grigsby et al. 2021); job-match quality (Gertler and Trigari 2009; Gertler et al. 2020); and occupation (Black and Figueiredo 2022), and by focusing on job-level wages (Hazell and Taska 2020). Our empirical finding of no significant difference in the cyclicality of new-hire and incumbent-worker wages is in line with these studies. This paper

[4] All existing studies that estimated the UCL use the NLSY. Doniger (2021) complements the analysis by using the CPS, supplemented by the CPS Job Tenure and Occupational Mobility Supplement for job tenure, and confirms a result similar to that of the NLSY.

is also related to a large body of literature that studies the long-term effects of a recession on labor market outcomes. A number of studies report that new graduates entering the labor market in recessions experience persistently lower earnings than those who enter the labor market in booms (e.g., Kahn 2010; Oreopoulos et al. 2012; Genda et al. 2010). The high cyclicality of the UCL is consistent with these findings.

This paper also contributes to the model selection of wage setting. A seminal work by Beaudry and DiNardo (1991) finds that the current wage depends on the history of labor market conditions at the time of and after hiring. Thomas and Worrall (1998) formally show that this pattern is consistent with an implicit wage contract with limited commitment. Hagedorn and Manovskii (2013) challenge this view by arguing that on-the-job search in the spot market leads to cyclical upgrading through job changes, which replicates the observed history dependence. More recently, Bellou and Kaymak (2021) construct a measure of job-match quality that accounts for the dissolution of poor matches through quits in expansions and layoffs in contractions, and find evidence in favor of an implicit contract. With our direct approach of focusing on new graduates, we still find evidence for history dependence in line with an implicit contract.

On the theoretical front, our model proposes a novel explanation of the asymmetry of the UCL's cyclicality. The literature has offered various explanations for incumbent-workers' downward wage rigidity, including psychological factors (e.g., Bewley 1999); shirking (e.g., Shapiro and Stiglitz 1984); and collective bargaining (e.g., Holden 1994). For new hires, Menzio and Moen (2010) consider a lack of commitment to secure employment, while Fukui (2020) explores strategic complementarities among firms on the job ladder as a source of rigidity. Our proposed mechanism is closely related to the efficiency wage (e.g., Solow 1979; Yellen 1984) and adverse selection (e.g., Stiglitz 1976; Guerrieri et al. 2010). A critical difference from these studies is that in our model, a single wage policy can attain a separation equilibrium by resorting to the different trade-off between wage and job-finding rate each type of worker faces. Moen and Rosén (2011) examine private information on match quality and effort during employment, while we focus on asymmetric information before hiring.

Layout. The remainder of the paper is organized as follows. Section 3.2 defines the UCL and describes its measurement. Section 3.3 explains the BSWS data and discusses our assumptions. Section 3.4 presents our main empirical results, and Section 3.5 is devoted to robustness checks. Section 3.6 develops a model to account for the empirical results, and Section 3.7 concludes.

3.2 Methodology

Definition. Let PDV_t be the expected present discounted value (PDV) of future allocative wages, or the user cost of labor, $UCL_{t+\tau}$:

$$PDV_t := \mathbb{E}_t \left[\sum_{\tau=0}^{\infty} \beta^\tau (1-s)^\tau UCL_{t+\tau} \right], \tag{3.1}$$

where $\beta \in (0,1)$ is the discount factor and $s \in (0,1)$ is the separation rate. Rearranging this equation, the UCL can be written as the differential of the present values of current and future labor:

$$UCL_t = PDV_t - \mathbb{E}_t \left[\beta(1-s) PDV_{t+1} \right]. \tag{3.2}$$

Notice that PDV_t should be equal to the PDV of future remitted wages, $w_{t,t+\tau}$:

$$PDV_t = \mathbb{E}_t \left[\sum_{\tau=0}^{\infty} \beta^\tau (1-s)^\tau w_{t,t+\tau} \right], \tag{3.3}$$

where $w_{t,t+\tau}$ denotes the wage paid at $t+\tau$ to the worker hired at t. Equations (3.2) and (3.3) map the remitted wages to the UCL.

What Does the UCL Capture? Before turning to an empirical counterpart, it would be helpful to rewrite the UCL following the exposition used by Doniger (2021):

$$UCL_t = w_{t,t} + \mathbb{E}_t\left[\sum_{\tau=1}^{\infty}\beta^{\tau}(1-s)^{\tau}(w_{t,t+\tau} - w_{t+1,t+\tau})\right]. \tag{3.4}$$

Equation (3.4) states that the UCL consists of the wage at hiring, or the new-hire wage, $w_{t,t}$, and the sequence of wage differences between hiring cohorts, $w_{t,t+\tau} - w_{t+1,t+\tau}$ for $\tau = 1, 2, ...$

Two special cases yield intuition on its cyclicality. In the first case, suppose that all workers are treated equally regardless of the timing of hiring. In the absence of cohort differences, the UCL coincides with the new-hire wage, i.e., $UCL_t = w_{t,t}$ if $w_{t,t+\tau} = w_{t+1,t+\tau}$ for all $\tau \geq 1$. In this case, wage differences in the current period represent the marginal cost of labor. In the other extreme case, each cohort is treated differently throughout their tenure. Suppose, for simplicity, that wages are fixed permanently after hiring—i.e., $w_{t,t+\tau} = w_{t,t}$. It is immediate to show that $UCL_t = 1/(1-\beta(1-s))w_{t,t} - \beta(1-s)/(1-\beta(1-s))\mathbb{E}_t[w_{t+1,t+1}]$, which implies that the impact of $w_{t,t}$ is scaled up by $1/(1-\beta(1-s)) > 1$. These examples show that the cyclicality of the UCL is closely linked to that of the new-hire wage, but also depends on the extent to which cohort differences persist after hiring.

Estimation Strategy. Our objective is to construct a measure of the UCL from the data. As described in Section 3.3, we construct the real hourly wage, $w^c_{t,t+\tau}$, and the number of workers, $N^c_{t,t+\tau}$, from the BSWS, where t, τ, and c represent the time of hiring, length of service, and employer-employee characteristics, respectively. These observations allow us to compute the UCL accordingly:

$$UCL^c_t := w^c_{t,t} + \sum_{\tau=1}^{T-1}\beta^{\tau}(1-s^c)^{\tau}(w^c_{t,t+\tau} - w^c_{t+1,t+\tau}), \tag{3.5}$$

where the discount factor β is set to 0.97, following Basu and House (2016). The separation rate is obtained by taking the time-series mean of the average survival rate over tenure:

$$s^c = 1 - \frac{1}{t_{end}}\sum_{t=1}^{t_{end}}\left(\frac{N^c_{t,t+T-1}}{N^c_{t,t}}\right)^{\frac{1}{T-1}}, \tag{3.6}$$

with t_{end} being the end period of the sample. T is the length of the period in which the calculation of PDV is truncated, as described below. Note that the separation rate differs across employer-employee characteristics, c, in this specification, which accommodates the potential heterogeneity pointed out in recent studies (e.g., Hall and Kudlyak 2022; Gregory et al. 2021). In the literature on the UCL, Doniger (2021) allows for heterogeneity in the separation rate across educational attainments but not across other characteristics.

Assumptions and Discussion. Assumptions for mapping the definition of the UCL in equations (3.2)–(3.3) into its empirical counterpart in equations (3.5)–(3.6) are summarized below.

Assumption 1. Perfect foresight. The expectation operator in equations (3.2)–(3.3) is replaced by the realized values, since the expected wages are not observed in the data. In this sense, our UCL is an ex post measure. The assumption follows previous studies (Kudlyak 2014; Basu and House 2016). Sensitivity to this assumption is assessed in Section 3.5.1.

Assumption 2. Years of truncation. The calculation of PDV is truncated at some year T, which should be sufficiently long so as not to affect the cyclical properties of the UCL. We set $T = 10$ years to reflect the lower separation rates in the Japanese labor market, while Kudlyak (2014) and Basu and House (2016) use $T = 7$ years for U.S. data. In Section 3.5.1, we verify that a further increase in T has little impact on our results.

Assumption 3. Time-invariant separation rate. The separation rate, s, is assumed to be constant over time in each category, but it differs across categories. We find that the separation rate is only insignificantly responsive to business-cycle fluctuations in our sample, as shown in Section 3.5.2.[5]

[5] This is not inconsistent with an implicit contract. Even under the limited commitment of workers' participation, firms may adjust the incumbent-worker wage enough to avoid their quits. We obtain empirical evidence consistent with this hypothesis, presented in Section 3.5.2.

Alternative Wage Measures. For comparison purposes, we construct the average wage, $w_t^{ave,c}$, as follows:

$$w_t^{ave,c} := \frac{\sum_{\tau=0}^{\tau_{max}} N_{t-\tau,t}^c w_{t-\tau,t}^c}{\sum_{\tau=0}^{\tau_{max}} N_{t-\tau,t}^c}, \tag{3.7}$$

where τ_{max} is the maximum tenure in the data. The new-hire wage, $w_t^{new,c}$, is denoted by:

$$w_t^{new,c} := w_{t,t}^c. \tag{3.8}$$

3.3 Data

3.3.1 BSWS

Our primary dataset is the Basic Survey on Wage Structure (BSWS) of Japan, which is a nationally representative annual survey conducted by the Ministry of Health, Labour and Welfare (MHLW). Since the survey is classified as a fundamental statistic under the Statistics Act of Japan, survey subjects are obliged to report their data and incur a penalty if they fail to do so.

Subjects of the survey are firms that employ more than five full-time employees. Sampling proceeds in two steps. First, firms are randomly selected to represent the whole economy in terms of geography, industry, and firm size. In the second step, employees are chosen in each selected firm to represent various categories, including gender and age. The number of firms sampled is around 80,000 as of 2016 from a population of about 1,400,000. The number of workers covered in the survey is more than 1 million each year. The survey is conducted annually in July and collects data on monthly earnings and hours worked in June of that year and bonuses in the entire previous year.[6]

[6] Bonuses are not available for new hires due to this survey method. We impute them using the bonuses that workers with tenure of 1 year received in the previous year.

Our sample spans from 1981 to 2019, with the aim of examining wage-setting behavior under low and stable inflation in recent decades. Though the survey separately compiles data for regular and non-regular workers (e.g., part-timers), we focus on regular workers.[7] This is in line with the paper's objective, since regular workers are assumed to be subject to long-term employment. With years of truncation T set to 10 years, the UCL series are constructed annually from 1981 through 2010.

Employer-Employee Characteristics. We have access to semi-aggregate wage series for each employer-employee characteristic in each age-tenure pair in the BSWS, rather than individual workers' wages. Thus, our UCL dataset is a category (employer-employee characteristic)-by-year balanced panel. Employer-employee characteristics in the BSWS include educational attainment (junior high school; high school; junior college or technical school; or college);[8] gender (male or female); and firm size (large firms with 1,000 or more full-time employees; medium firms with 100 to 999; or small firms with 99 or fewer), as well as worker's age and tenure in the current employment.[9] We omit junior high school graduates due to their small sample size. The number of categories available is 18 ($= 3 \times 2 \times 3$). We discuss the implications of the semi-aggregate series in Section 3.3.2.

Panel (A) of Table 3.1 displays the number of employees in each category. Workers with a high school degree and male workers account for the largest fraction in the sample, whereas workers are distributed somewhat evenly across firm size.

---

[7]These worker types are labeled "full-time" and "part-time" workers, respectively, in the BSWS. We prefer to call them "regular" and "non-regular" workers, because the MHLW asks surveyed firms to differentiate the two types of workers according to their scope of responsibility and the type of wage and employment contract.

[8]Years of schooling to complete each educational attainment are 9 for junior high school, 12 for high school, 14 for junior college or technical school, and 16 for college.

[9]Firm size is assessed in each survey year, and thus could change over time. However, firm dynamics in Japan are more stable over time than those in U.S. data (e.g., Mukoyama 2009).

Table 3.1: Descriptive Statistics

(A) Number of Workers by Category

	High school	Junior college or technical school	College	Total
Education	9,518	4,097	7,960	21,575
	(44.1)	(19.0)	(36.9)	(100)
	Male	Female		Total
Gender	13,982	7,593		21,575
	(64.8)	(35.2)		(100)
	Large	Medium	Small	Total
Firm size	8,091	7,936	5,548	21,575
	(37.5)	(36.8)	(25.7)	(100)

Notes: Thousands of workers. The portion of total employees is in parentheses. As of 2019.

Remitted Wage Measure. Our measure of remitted wages, $w_{t,t+j}$, is real hourly earnings. Hourly earnings are computed by dividing total cash earnings by total hours worked. Total cash earnings are the sum of base pay, overtime pay, bonuses, and other cash payments such as commuting and family allowances. These are before taxes and the employee's contributions to social security. The inclusion of non-base components allows us to accommodate wage adjustments using various components of earnings.[10]. The firms surveyed report employees' earnings and hours worked according to their payroll records; Japanese law requires that firms record this information. Therefore, BSWS data do not suffer measurement errors, unlike workers' self-reported data. Panel (B) of Table 3.2 shows the average nominal hourly earnings in each category. Workers with a college degree, male workers, and those in large firms receive higher earnings on average.

[10]Though non-cash benefits are not covered by the BSWS, there is limited scope for arbitrary adjustments of non-cash benefits under the Japanese legal framework. More discussion is provided in Appendix 4.6.1

Table 3.2: Descriptive Statistics (cont.)

(B) Average Nominal Hourly Wage by Category

	High school	Junior college or technical school	College	Total
Education	2,024	2,239	3,068	2,450
	(82.6)	(91.4)	(125.2)	(100)
	Male	Female		Total
Gender	2,703	1,984		2,450
	(110.3)	(81.0)		(100)
	Large	Medium	Small	Total
Firm size	2,936	2,322	1,925	2,450
	(119.8)	(94.8)	(78.6)	(100)

Notes: JPY per hour. Wage levels relative to the average wage are in parentheses. As of 2019, given the average JPY/USD exchange rate of 109.0, the overall average hourly wage (2,450 JPY) is equivalent to 22.5 USD.

(C) Average Separation Rate by Category

	High school	Junior college or technical school	College	Total
Education	9.1%	9.1%	6.3%	8.0%
	Male	Female		Total
Gender	6.7%	11.2%		8.0%
	Large	Medium	Small	Total
Firm size	5.9%	7.9%	10.3%	8.0%

Notes: Annual rates within 10 years from hiring. Average from 1981 to 2019.

Wage series are deflated by the GDP implicit price deflator. This choice follows Basu and House (2016), on the premise that the UCL measures labor cost from the firm's perspective and thus reflects prices on the production side of the economy.

New Graduates. To control for the cyclical upgrading of job-match quality through job changes, we restrict the sample to new graduates, who do not have previous jobs by definition and thus are not subject to cyclical upgrading through job changes.[11]. New graduates are identified by their tenure, age, and educational attainment, as recorded in the BSWS. The standard age of new graduates is 18 years for high school graduates, 20 for junior college or technical school graduates, and 22 for college graduates. The BSWS records workers' age in the following groups: 17 years or below, 18–19 years, 20–21 years, 22–24 years, and 25 years and above in 5-year increments. We assume that new hires (workers with 0 years of tenure) aged 18–19 years with a high school degree, 20–21 years with a junior college or technical school degree, and 22–24 years with a college degree are new graduates.

Strictly speaking, it is not impossible that each age category includes non-new graduates or excludes new graduates. However, these risks are considered to be minimal for the following reasons.[12] First, it is not possible to graduate from high school before the age of 18, since in the Japanese education system schooling years are strictly linked to students' age up to high school. In addition, it is extremely rare to graduate with a higher degree earlier than at the standard age. Only those with outstanding qualities in certain fields, such as sports, can be admitted to university under the age of 18. Historically, there have been fewer than 100 such cases out of hundreds of millions of students. Second, separation rates are low and stable in the Japanese labor market. In our sample, annual separation rates are lower than 10% even in the first few years of employment. Thus, there are few job changes within an age group (e.g., a high school graduate who starts working at 18 years

[11]Since the BSWS does not distinguish workers hired from unemployment or another employment, we cannot follow exactly the same approach as Gertler et al. (2020). Instead, we focus on new graduates, as described above. New graduates constitute the majority of workers not hired from previous jobs in the Japanese labor market. We provide further discussion of the sample of new graduates in Appendix 4.6.2

[12]According to the Japan Household Panel Survey, which offers additional panel data for Japan (see to Section 3.5.3 for more details), only 0.6% of workers who have obtained a high school degree (3 out of 479 workers) change jobs before reaching the age of 19. Similarly, only 0.5% of workers with a junior college or technical school degree (2 out of 422 workers) change jobs before turning 21, and only 0.2% of workers with a college degree (1 out of 433) change jobs before they reach 24 years of age. This finding suggests that almost all workers classified as "new graduates" within the BSWS would remain in their initial employment.

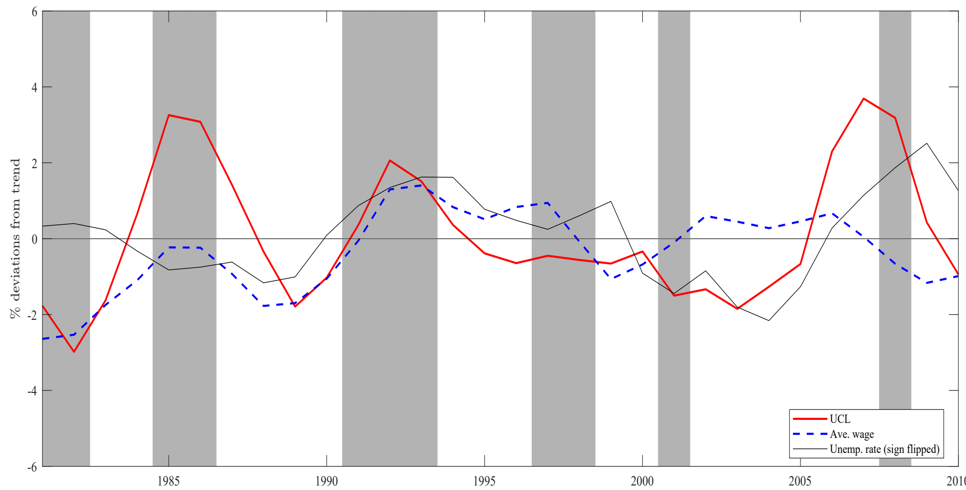
old, then quits and moves to another job before 19 years old is categorized in the age group 18–19 years old). Moreover, workers within the first 3 years of work experience are called "second new graduates." They usually undergo the same application and selection process and are employed under the same conditions as new graduates in the Japanese labor market practices.

New graduates are tracked after hiring according to their tenure and age. For example, 32-year-old college graduates with a tenure of 10 years in their current jobs are assumed to have started the jobs as new graduates. The length of service is reported in the following groups: 0 years (new hires), 1–2 years, 3–4 years, and 5 years and above in 5-year increments. We interpolate wages for each year of service by using the fourth-order polynomial.

Separation Rate. Panel (C) of Table 3.2 shows the average annual separation rates within 10 years after hiring. The rates tend to be higher for lower-wage workers, such as those with a non-college degree, female workers, and small firms. These features are similar to those reported by Farber (1999) for the U.S. economy. The overall average separation rate is 8.0% per year. The rate is substantially lower than that in the NLSY sample, in which Kudlyak (2014) reports that the average separation rate is around 30% per year. The difference can reflect labor market regulations and practices in the two economies, as well as the fact that our sample focuses on regular workers whose tenure tends to be long.

While the low separation rate could be understood in the context of a lifetime employment system in the Japanese labor market (e.g., Hashimoto and Raisian 1985), one might expect that job mobility has been increasing in recent years. However, a secular trend appears to be marginal in our dataset, with the average separation rate increasing from 7.7% in the first half of the sample (1981–1999) to 8.2% in the second half (2000–2019). Kambayashi (2017) documents that changes in Japanese labor market practices, including the increase in job mobility, have been caused by the spread of non-regular workers—who replace a substantial portion of self-employed workers. In contrast, labor market practices for regular

Figure 3.1: Estimated UCL and Average Wage for Male High School Graduates in Large Firms



Notes: HP-filtered values. The UCL and average wage are the 3-year centered average. The sign of the unemployment rate is flipped. Shaded areas indicate recessions, according to business-cycle dating by the Cabinet Office of Japan.

workers—the focus of our empirical analysis—have been largely protected.

Estimated Series. Before turning to our empirical analysis, Figure 3.1 shows the estimated UCL series for male high school graduates in large firms—the largest category in terms of the number of workers. Though disaggregated series tend to be noisy and affected by category-specific factors, the estimated UCL appears to be more synchronized with business cycles—for example, with the unemployment rate—than the average wage.

3.3.2 Assumption and Discussion

Since we use semi-aggregate wage series, wage dynamics are subject to composition changes within a category. Although the sample of new graduates is supposed to be free from cyclical upgrading through job changes, our implicit assumption is that other dimensions of unobserved heterogeneity—namely, worker and firm characteristics within each category—are, on average, independent of business-cycle fluctuations.[13] We discuss below the validity of this assumption. At the same time, we provide suggestive evidence in Section 3.5.3 on the acyclicality of job-match quality in our sample of new graduates using longitudinal panel data. Note also that in the regression analysis in Sections 3.4 and 3.5, the UCL and new-hire wage are constructed from the identical sample of new hires in each year (either new graduates or all new hires), and thus different cyclicalities between them are not attributable to composition bias.

First, various observed employer-employment characteristics are controlled for. In addition to the aforementioned 18 categories (education×gender×firm size), the sample of new graduates tracks workers with specific ages and years of service. For instance, an observation in our dataset is the wage of male high-school graduates aged 18–19 years old who work in a large firm with 0 years of tenure. Total data points are around 2,500 each year. Unobserved heterogeneity, if any, would cause bias only in each data point.

Second, regarding worker-side heterogeneity, entrance to a higher-level school and graduation timing can be endogenous. However, such endogenous choices would reduce the excess supply or shortage of new graduates over business cycles (e.g., fewer workers would choose to go into a job market in recessions), and thereby reduce the wage cyclicality. Moreover, since job posting is procyclical, a looser labor market in recessions causes more severe competition among workers, which leads to countercyclicality in the average productivity of new hires for a given job through sorting. This composition bias would understate the cyclicality of the UCL, rather than overstating it.

Third, on the firm side, if the job creation of higher-wage jobs is more cyclical, the average new-hire wage becomes procyclical. A long line of the literature has reported that

[13] In the literature of the UCL, while the original estimation of Kudlyak (2014) controls for individual fixed effects in the NLSY sample, Bils et al. (2023) do not include them while controlling for match quality by using the long-term wage in a match. Doniger (2021) also does not include them when using the CPS, due to the short length of the panel structure.

composition bias has the opposite effect. That is, the employment of lower-wage workers is more cyclical, which dampens cyclicality at the aggregate level (e.g., Bils 1985; Solon et al. 1994). However, a potential dimension that can generate procyclical bias is industry if higher-wage industries are more cyclical (e.g., McLaughlin and Bils 2001). Although industry-level data are available in the BSWS, we chose not to use them in the baseline analysis because the industry breakdown leads to a small sample in each category, which in turn cause noisy observations and frequent missing values. The issue is magnified for new graduates, whose sample sizes are relatively small. Alternatively, in Section 3.5.1, we examine potential composition bias due to the omission of industry variation by using the sample that includes job changers. The analysis indicates that controlling for industry heterogeneity has little impact on our results. While we do not preclude another dimension of the firm-side unobserved heterogeneity,[14] it should be noted that other studies that use worker-level wage data are also subject to this bias, if any.

Fourth, composition changes through separation are less concerning due to the low and stable separation rates in the Japanese labor market. We do not find significant changes in the separation rate over business cycles, as shown in Section 3.5.2.

3.4 Main Results

3.4.1 Cyclicality of the UCL

To examine the cyclicality of the UCL, we run the following regression:

$$y_t^c = \gamma_1 x_t + \alpha^c + \epsilon_t^c, \tag{3.9}$$

[14]For instance, Oreopoulos et al. (2012) report that using a Canadian employer-employee-matched dataset for college degree workers, persistent earnings declines for workers who enter the labor market in recessions are due to more workers taking jobs at poor quality firms, measured by firm size or average earnings. In this regard, we calculate the UCL for different categories, including firm size, and thereby control for the composition change in this dimension to some extent.

where α^c is the fixed effect for category (educational attainment$\times$gender$\times$firm size), and y_t^c and x_t are wage and cyclicality measures. Our wage measures, y_t^c, are the UCL, new-hire wage, and average wage. For a cyclicality measure, x_t, we follow the convention in the literature and use the unemployment rate in the baseline case. Since most new graduates start employment in April after an academic year that ends in March in the Japanese education system, working conditions, including wages, are presumably determined according to the economic conditions before the start of their jobs. Thus, we use the average unemployment rate from April of the previous year to March of the current year. We use an HP filter to extract the cyclical components for both y_t^c and x_t with a standard smoothing parameter of $\lambda = 100$ for annual data (Backus and Kehoe 1992). The robustness of these specifications is assessed in Section 3.5.1.

The coefficient of interest is γ_1. A positive value of γ_1 indicates the procyclicality of y_t^c. The weighted least square is used with the weight equal to the number of new graduates in each category, $N^{new,c} = 1/t_{end} \sum_{t=0}^{t_{end}} N_{t,t}^c$. Standard errors (SEs) are clustered by category, assuming the presence of persistent category-specific shocks. To ensure robustness, we also report SEs clustered by year to accommodate year-specific shocks that are not captured by the unemployment rate.

Table 3.3 reports the estimated coefficient of the unemployment rate, γ_1, in the cyclicality regression (3.9). Since the sign of the unemployment rate is reversed and the series is standardized, γ_1 represents the percentage change in wage measures to a one-standard-deviation decrease in the unemployment rate. Column (1) is our baseline specification of the UCL for new graduates, whereas column (2) shows the UCL for all workers including job changers, which corresponds to the series without correcting match quality changes for comparison purposes.

The lower cyclicality of the UCL for new graduates than for all workers implies the procyclical upgrading of job-match quality. However, even after correcting for procyclical upgrading—i.e., by limiting the sample to new graduates—the UCL remains highly procyclical: The estimated coefficient is close to twice as large as that for the average wage, $w_t^{ave,c}$, shown in column (5).

Table 3.3: Baseline Results: Cyclicality of the UCL, Average Wage, and New-Hire Wage

	(1)	(2)	(3)	(4)	(5)
	UCL		New-hire wage		Ave. wage
	New grad.	All workers	New grad.	All workers	All workers
x_t: Unemp. rate	0.866	1.032	0.471	0.788	0.525
(SE clust. by category)	(0.229)***	(0.192)***	(0.105)***	(0.143)***	(0.135)***
(SE clust. by year)	(0.307)***	(0.320)***	(0.276)*	(0.323)*	(0.238)**
R-squared	0.04	0.10	0.04	0.10	0.07
N of categories	18	18	18	18	18
N of observations	540	540	540	540	540

SEs clustered by category/year are reported in parentheses in the upper/lower line.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table shows the estimated values of γ_1 in (3.9). The unemployment rate is standardized and its sign is reversed. Estimated coefficients represent the percentage change of wage measures to a one-standard-deviation decrease in the unemployment rate.

Regarding the new-hire wage, $w_t^{new,c}$, when we focus on new graduates in column (3), its cyclicality is around half that for all newly hired workers in column (4). Although previous literature shows that the job quality of new graduates is influenced by economic conditions (e.g., Kahn 2010), our result suggests that new graduates are less affected by cyclical upgrading—at least less so than job changers. We further investigate this point in Section 3.5.3. Also, the cyclicality in column (3) becomes similar to that of the average wage in column (5). Put together, our result is in line with studies that control for workers' productivity and job-match quality (e.g., Gertler and Trigari 2009; Gertler et al. 2020; Grigsby et al. 2021), while other studies report that new-hire wages are highly cyclical without such correction.

It is notable that the UCL is more cyclical than the new-hire wage. This implies that the high cyclicality of the UCL arises not only from that of the new-hire wage, but also from the wage dynamics of incumbent workers. As discussed in Section 3.2, incumbent-workers' wage rigidity increases the cyclicality of the UCL. In that sense, our result is consistent with studies that document incumbent-workers' wage rigidity (e.g., Yamamoto 2007 for the Japanese labor market). The difference is clearer in the sample of new graduates (columns 1 and 3) than all workers (columns 2 and 4). This is partly because new graduates exhibit relatively lower separation rates, and thus the persistence of the incumbent-worker wage is more amplified when calculating their UCL.

SEs tend to be larger when clustered by year. While the significance level is not affected for the UCL (columns 1-2), clustering by year leads to a slightly lower significance level for average and new-hire wages (columns 3–5). These results cause the difference in the cyclicality of the UCL and other wage measures to be even more apparent.

Comparison with Studies using U.S. Data. Table 3.4 provides a summary of studies by Basu and House (2016) and Bils et al. (2023), who estimate the semi-elasticity of the UCL and new-hire wages with respect to the unemployment rate change, while controlling for job-match quality with different methodologies.[15,16] The final row of the table presents our estimated values for semi-elasticity, obtained by dividing the values from columns (1) and (3) in Table 3.3 by the standard deviation of the unemployment rate in our sample (0.322).

Column (2) shows contrasting results in prior studies regarding the cyclicality of new-hire

[15] Doniger (2021) also estimates the UCL after controlling for job-match quality using the NLSY, while she estimates the UCL for workers with different degrees. Her estimated semi-elasticity of the UCL for workers beyond a college degree, with a high-school degree, and without a high-school degree is 15.49, 4.90, and -1.36, respectively.

[16] Basu and House (2016) and Doniger (2021) use a proxy for job-match quality proposed by Hagedorn and Manovskii (2013). Bils et al. (2023) use the long-run wage in a match as a measure of its quality.

Table 3.4: Cyclicality after Controlling for Job-Match Quality in Previous Studies

Semi-elasticity w.r.t Unemp. rate	(1) UCL	(2) New-hire wage	Data
Basu and House (2016)	4.77	0.69	NLSY
	(2.05)	(1.85)	
Bils et al. (2023)	4.81	2.35	NLSY
	(1.83)	(0.67)	
This paper	2.69	1.46	BSWS

Notes: The table reports the semi-elasticity of each measure with respect to a 1 p.p. change in the unemployment rate after controlling for job-match quality. We report the values of Table 6 in Basu and House (2016) and Table 7 in Bils et al. (2023). SEs are reported in parentheses.

wages. Specifically, Basu and House (2016) observe acyclicality, whereas Bils et al. (2023) find significant cyclicality. Our estimate for the new-hire wage falls within the intermediate range of these values with statistical significance.

In column (1), our estimated value for the UCL is approximately half the values reported in previous studies, while the SEs in the above two studies are large due to the small sample size of the NLSY (27 in Basu and House 2016 and 32 in Bils et al. 2023).

The smaller cyclicality of our UCL can be attributed to lower wage rigidity for incumbent workers in Japan.[17] As discussed in Section 3.2, the lower wage rigidity among incumbent workers implies that wage changes over time are less smooth, which leads to smaller cyclicality of the UCL. Previous studies that compare wage rigidity between Japan and the U.S. have consistently highlighted lower wage rigidity in Japan. For instance, Mineyama et al. (2022) observe that compensation per hour in Japan falls more than in the U.S. during

[17]Another possible explanation for this disparity in cyclicality is the methodology we employ to control for job-match quality. However, due to the limited availability of cohorts in the NLSY dataset, we are unable to estimate UCL in the U.S. using our approach.

past recessions. Kuroda and Yamamoto (2007) report a higher frequency of wage changes for Japanese workers than for the U.S., and attribute the lower wage rigidity to specific characteristics of the Japanese labor market, such as the significance of firm-specific human capital. The accumulation of firm-specific human capital under strong employment protection results in low job turnover, since workers cannot carry over the accumulated human capital when transitioning to a different firm. Consequently, an employment contract in the Japanese labor market often embeds wage adjustment mechanisms to render the long-term employment system sustainable; these include overtime payments and bonuses, and workers tend to accept frequent wage changes rather than quit (e.g., Kuroda and Yamamoto 2007).

Put together, the lower rigidity of incumbent workers' wages in Japan can cause our UCL to be less cyclical than those in the U.S., while the existence of incumbent workers' wage rigidity causes the UCL to be more cyclical than new-hire wage in our estimation.

3.4.2 History Dependence of the Incumbent-Worker Wage

To delve into the factors behind the high cyclicality of the UCL, we examine the history dependence of the incumbent-worker wage. Although our approach follows the literature initiated by Beaudry and DiNardo (1991), an innovation is that we focus on the sample of new graduates to remove cyclical upgrading through job changes. We regress wages at each tenure j, $w^c_{t-j,t}$, on the unemployment rate in the year of hiring, x^{lag}_t, and the lowest rate in the course of tenure, x^{min}_t, as well as controlling for the contemporaneous rate, x_t:

$$w^c_{t-j,t} = \gamma_2 x^{lag}_t + \delta x_t + \alpha^c + \epsilon^c_t, \tag{3.10}$$

$$\text{and} \quad w^c_{t-j,t} = \gamma_2 x^{lag}_t + \gamma_3 x^{min}_t + \delta x_t + \alpha^c + \epsilon^c_t. \tag{3.11}$$

A positive coefficient of γ_2 indicates that the labor market condition as of hiring persistently affects wages, which implies incumbent-workers' wage rigidity. Beaudry and DiNardo (1991) argue that this pattern is consistent with an implicit contract under which a risk-neutral firm offers insurance against business-cycle fluctuations to risk-averse workers. In contrast, Hagedorn and Manovskii (2013) show that the history dependence can be replicated in a spot labor market in the presence of cyclical upgrading through job changes. Since our sample of new graduates excludes job changes, it is well suited to test this hypothesis. In Table Table 3.5, the estimated positive γ_2 supports the presence of an implicit contract. The effects gradually decay as the length of service increases, which is also consistent with sequential wage updating under limited commitment, as discussed below.

Table 3.5: Effect of Labor Market Condition as of Hiring

	(1)	(2)	(3)	(4)
	Wages with the length of service			
	1–2y	3–4y	5–9y	10–14y
x_t^{lag}: Lagged unemp. rate	0.442	0.352	0.278	0.277
	(0.077)***	(0.108)***	(0.142)*	(0.097)**
	(0.224)*	(0.162)**	(0.240)	(0.326)
x_t: Current unemp. rate	0.324	0.812	1.163	0.652
	(0.101)***	(0.110)***	(0.134)***	(0.109)***
	(0.263)	(0.215)***	(0.195)***	(0.210)***
R-squared	0.14	0.24	0.40	0.11
N of categories	18	18	18	18
N of observations	522	522	522	522

Notes: The sample is restricted to new graduates. Dependent variables are wages of new graduates with each length of service. x_t^{lag} is the unemployment rate in the year of hiring. Wages and lagged unemployment rates are taken as the 2-year average to smooth volatile observations.

The lowest unemployment rate in the course of tenure, x_t^{min}, is a proxy for the most favorable labor market condition for workers. If workers' commitment to the continuation of a wage contract is limited, firms may raise wages in expansions to prevent workers from quitting for another job as long as match surplus remains. The positive estimate of γ_3 in Table 3.6 is consistent with limited commitment. Different columns in the panel report

Table 3.6: Effect of Labor Market Condition in the Course of Tenure

	(1)	(2)	(3)	(4)
	Wages with the length of service			
	1–2y	3–4y	5–9y	10–14y
x_t^{lag}: Lagged unemp. rate	0.442	0.424	0.190	0.175
	(0.077)***	(0.144)***	(0.147)	(0.092)*
	(0.224)*	(0.226)*	(0.258)	(0.234)
x_t^{min}: Lowest unemp. rate	0.324	–0.201	0.362	1.113
	(0.101)***	(0.174)	(0.082)***	(0.266)***
	(0.263)	(0.441)	(0.280)	(0.471)**
x_t: Current unemp. rate	–	0.947	0.966	0.514
	–	(0.185)***	(0.146)***	(0.122)***
	–	(0.397)**	(0.223)***	(0.178)***
R-squared	0.14	0.25	0.41	0.19
N of categories	18	18	18	18
N of observations	522	522	522	522

SEs clustered by category/year are reported in parentheses in the upper/lower line.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: The sample is restricted to new graduates. Dependent variables are wages of new graduates with each length of service. x_t^{min} is the lowest rate after hiring—i.e., $x_t^{min} = \min\{x_{t-j}, x_{t-j+1}, ..., x_t\}$. Wages and lagged unemployment rates are taken as the 2-year average to smooth volatile observations.

smaller γ_2 and larger γ_3 at a longer tenure.[18] The longer their tenure, the more likely workers receive a large enough shock to trigger a wage change. At the same time, their wages become less dependent on the labor market condition at the time of their hiring.

An implication of this history dependence of the incumbent-worker wage is higher cyclicality of the UCL compared with the new-hire wage. That is, the cost of labor is not just

[18]The estimated negative γ_3 in column (2), though it is not significantly different from zero, appears counterintuitive. This may arise from the high collinearity among x_t^{lag}, x_t^{min}, and x_t for workers with shorter tenure, which could make it difficult to identify the precise effect of each variable.

the remitted wage at each point in time; rather, it entails persistent wage differences across hiring cohorts over their entire tenure, as discussed in Section 3.2.

3.4.3 Asymmetry

Another area of potential interest is the asymmetry in cyclicality, given the substantial debate in the literature regarding the downward rigidity of wage adjustments (e.g., Tobin 1972 and many others) and asymmetry in labor market dynamics (e.g., Neftçi 1984 and many others). Table 3.7 shows the asymmetry of each wage measure. In columns (1)-(3), we split the HP-filtered unemployment rate into positive and negative values and estimate different coefficients for them. The table indicates that the cyclicality is considerably larger in booms across different wage measures.[19] In columns (4)-(6), we test for an alternative specification by splitting the sample by the direction of changes, rather than the sign of levels, and estimate different coefficients for increases and decreases in the unemployment rate. Wage measures in dependent variables are also taken as differences. The estimated coefficients are highly asymmetric, which confirms downward rigidity.

The result is surprising in light of the conventional wisdom regarding incumbent-worker wage's downward rigidity, since the UCL would be adjusted downward if the new-hire wage were to be downwardly flexible. The combination of the downward rigidity of new-hire and incumbent-worker wages can explain that of the UCL. In the literature, various studies find that wages frequently change when workers move to other jobs (e.g., Barattieri et al. 2014). However, at the job level, recent work by Hazell and Taska (2020) shows that wages posted by firms are downwardly rigid in U.S. data. Fukui et al. (2020) report a similar pattern in the Japanese labor market based on data during the COVID-19 pandemic. Our result is in line with these studies.

[19] The estimated negative coefficients for contractions in several specifications are not inconsistent with previous studies that report the countercyclicality of real wages in recessions (e.g., Gu et al. 2020). Fixed components of earnings, such as base wage, can be more expensive in recessions when converted to an hourly rate.

Table 3.7: Asymmetry in Cyclicality

	(1) UCL	(2) New-hire wage	(3) Average wage	(4) ΔUCL	(5) Δ New-hire wage	(6) ΔAverage wage
$x_t \times 1_{x_t \geqq 0}$	1.570	1.123	1.589			
	(0.362)***	(0.185)***	(0.189)***			
	(0.895)*	(0.540)***	(0.687)**			
$x_t \times 1_{x_t < 0}$	0.072	–0.264	–0.675			
	(0.414)	(0.128)*	(0.174)**			
	(0.582)	(0.393)	(0.546)			
$\Delta x_t \times 1_{\Delta x_t \geqq 0}$				2.397	1.535	0.366
				(0.738)***	(0.111)***	(0.175)*
				(1.323)*	(0.908)*	(0.701)
$\Delta x_t \times 1_{\Delta x_t < 0}$				-1.463	-0.878	0.090
				(0.158)	(0.366)**	(0.252)
				(1.311)	(1.060)	(0.564)
R-squared	0.05	0.07	0.15	0.02	0.03	0.01
Prob >F	0.03	0.00	0.00	0.00	0.03	0.89
N of categories	18	18	18	18	18	18
N of observations	540	540	540	522	522	522

SEs clustered by category/year are reported in parentheses in the upper/lower lines.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: The same notes as for Table 3.3 apply. The UCL and new-hire wage are those for new graduates, whereas the average wage is that for all workers. In columns (1)-(3), the independent variable (HP-filtered unemployment rate) is split into positive and negative values. In columns (4)-(6), both dependent and independent variables are differences, and the latter is split by increases and decreases. The values in "Prob > F" represent the p-value associated with the F test for the null hypothesis that the coefficient of positive values equals that of negative values.

3.5 Robustness Check

3.5.1 Specification

Table 3.8 assesses the robustness of our baseline result. Panel (A) examines different detrending methods. In columns (1)–(3), we employ the Hamilton (2018) filter with which filtered values are obtained as residuals by regressing the current value on the 2-to-5-period lagged values. We also use the year-on-year growth rate in columns (4)–(6). These alternative detrending methods yield a similar pattern of relative cyclicality. The difference in the estimated coefficients of the UCL and average wage is in the range of two to four times, which is somewhat higher than the baseline result with the HP filter.

Table 3.8: Robustness Check

(A) Alternative Methods of Detrending

	(1)	(2)	(3)	(4)	(5)	(6)
	Hamilton filter			Year-on-year changes		
	UCL	New-hire wage	Ave. wage	UCL	New-hire wage	Ave. wage
x_t: Unemp. rate	1.891***	1.309***	0.510**	0.503**	0.316**	0.290***
	(0.394)	(0.432)	(0.183)	(0.211)	(0.113)	(0.098)
R-squared	0.08	0.12	0.03	0.01	0.01	0.03
N of categories	18	18	18	18	18	18
N of observations	450	450	450	522	522	522

Notes: The same notes as for Table 3.3 apply. The UCL and new-hire wage are those for new graduates, whereas the average wage is that for all workers.

Table 3.9 explores different cyclicality measures. Since the UCL of new graduates is calculated from young workers' wages, Panel (B-I) examines youth unemployment rates as an independent variable. We use the unemployment rates for 10-year age groups (15–24 years, 25–34 years, ...) published by the Ministry of Internal Affairs and Communications.

Table 3.9: Robustness Check (cont.)

(B-I) Youth Unemployment Rate

	(1)	(2)	(3)	(4)	(5)	(6)
	Unemp. rate for 15–24 years old			Unemp. rate for 25–34 years old		
	UCL	New-hire wage	Ave. wage	UCL	New-hire wage	Ave. wage
x_t: Unemp. rate for age 15-24	0.931*** (0.243)	0.500*** (0.109)	0.496*** (0.132)			
x_t: Unemp. rate for age 24-35				0.923*** (0.252)	0.529*** (0.112)	0.610*** (0.145)
R-squared	0.04	0.05	0.10	0.04	0.05	0.08
N of categories	18	18	14	18	18	18
N of observations	540	540	540	540	540	540

(B-II) Alternative Cyclicality Measures

	(1)	(2)	(3)
	Job-opening-to-applicant ratio		
	UCL	New-hire wage	Ave. wage
x_t: Job opening ratio	1.390*** (0.321)	1.070*** (0.143)	0.898*** (0.204)
R-squared	0.06	0.13	0.11
N of categories	18	18	14
N of observations	540	540	540

SEs clustered by category are reported in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: The same notes as for Table 3.3 apply. The UCL and new-hire wage are those for new graduates, whereas the average wage is that for all workers.

The correlation among these unemployment rates across age groups is so high (above 0.9) that the estimated coefficients remain close to the baseline case, which uses the overall unemployment rate.

Also, Panel (B-II) uses the job-openings to job-applicants ratio in the Job/Employment Placement Services Statistics (JEPSS) as an alternative cyclicality measure. The JEPSS compiles vacancies posted and job-seekers registered in public employment security offices, which cover around 20% of all new hiring. The job-openings to job-applicants ratio is an indicator of labor market tightness, which is often defined in a search and matching model. The panel confirms the high cyclicality of the UCL. The relative magnitude of cyclicality remains unchanged from the baseline case when the job-opening to job-applicants ratio is used as a cyclicality measure.

Panel (C) examines industry-level data. As discussed in Section 3.3, an industry breakdown of the new graduates' sample renders the series noisy and frequently generates missing values. As a second-best approach, we focus on the sample of all workers, whose sample sizes are larger than those of new graduates, and examine potential industry composition bias. We use eight industries that are available throughout the sample period; details are presented in Appendix 4.6.4. The panel suggests that the industry breakdown does not materially alter the estimates, but the difference between the cyclicality of the UCL and average wage is a little larger than the baseline case. Consistently, Appendix Figure 4.2 shows that there is no clear relationship between the cyclicality of the number of new hires and their wage level across industries, which implies little bias due to industry composition over business cycles.

Panel (D) explores alternative assumptions in constructing the UCL. We first examine different truncation periods, T. Columns (2)–(4) confirm that the estimates are little affected by T. Next, we consider an alternative expectation assumption. Specifically, we assume that agents forecast future wages at each tenure with a random walk (RW), i.e., $\tilde{\mathbb{E}}[w_{t,t+j}] = w_{t-j,t}$ and $\tilde{\mathbb{E}}[w_{t+1,t+j}] = w_{t-j,t+1}$. This assumption implies that the best forecast for the j-years-ahead wage of a newly hired worker is the current wage of an incumbent worker with j years of tenure. Under the RW assumption, the estimated coefficient in column (5) is substantially larger than the baseline case of perfect foresight. We conjecture that the reality lies between

Table 3.10: Robustness Check (cont.)

(C) Industry-Level Data

	(1)	(2)	(3)	(4)	(5)	(6)
	UCL		New-hire wage		Ave. wage	
	Baseline	Industry	Baseline	Industry	Baseline	Industry
x_t: Unemp. rate	1.032***	1.143***	0.788***	0.744***	0.525***	0.323***
	(0.192)	(0.162)	(0.143)	(0.120)	(0.135)	(0.093)
R-squared	0.19	0.02	0.15	0.05	0.10	0.01
N of categories	18	144	18	144	18	144
N of observations	504	4,032	504	4,032	504	4,032

(D) Calculation of the UCL

	(1)	(2)	(3)	(4)	(5)	(6)
	Baseline	Different T			RW	Overall UCL
		15y	20y	25y	assumption	
x_t: Unemp. rate	0.866***	0.847***	0.851***	0.841***	1.289***	0.845***
	(0.229)	(0.220)	(0.217)	(0.213)	(0.291)	(0.114)
R-squared	0.04	0.04	0.05	0.05	0.01	0.09
N of categories	18	18	18	18	18	1
N of observations	540	540	540	540	540	28

In column (6), the HAC robust SE with a 3-period lag is reported.

Notes: Panel (C) uses the sample of all workers, while Panel (D) uses that of new graduates. Baseline results, reported in columns (1), (3), and (5) of Panel (C) and column (1) of Panel (D), are identical to those reported in Table 3.3.

the two cases; agents have imperfect knowledge about the future, and partly benchmark incumbent workers' wages when forming expectations. Finally, in column (6), we calculate the overall UCL across categories in the spirit of Kudlyak (2014). Details are provided in Appendix 4.6.3. The estimated coefficient is quite close to our baseline case, which supports the conceptual equivalence of the two measures.

Recent work by Doniger (2021) estimates the UCL for workers with different levels of educational attainment, and reports that the high cyclicality is concentrated for highly educated workers in U.S. data. Although our UCL series are category specific and accommodate heterogeneity in wage dynamics and separation rates, we do not find a significant difference in cyclicality across educational attainments in our sample of Japanese workers, as presented in Appendix 4.7. This may stem from the institutional settings in U.S. and Japanese labor markets. Among others, a dual labor market structure is often reported between regular and non-regular workers in Japan (e.g., OECD 2009), but our empirical analysis focuses on regular workers, who may tend to be subject to less heterogeneous employment conditions.

3.5.2 Separation Rate

Nekarda and Ramey (2020) dispute the assumption of exogenous separation rates in calculating the UCL. They argue that on-the-job search, along with incumbent-workers' wage rigidity, leads to a lower separation rate of workers hired in expansions, since these workers can enjoy the high wages of the current job during the course of their tenure. Those hired in recessions, who are stuck in a low-wage contract, are more willing to move to another job.

While we admit that the allocative role of the UCL can be altered with endogenous separation, our view is that its relevance is an empirical question, since there can be mitigating and offsetting forces for the channel, as highlighted by Nekarda and Ramey (2020). For example, firms may compensate workers to prevent their separation (e.g., Beaudry and DiNardo 1991). In addition, employment-to-unemployment flow is countercyclical in the data and reflects job losses in recessions (e.g., Fujita and Ramey 2009).

Table 3.11 shows the sensitivity of separation rates to labor market conditions. We calculate the separation rates of each hiring cohort during the 10 years after hiring and regress them on the unemployment rate in the year of hiring, x_t, in column (1). A negative coefficient is expected for x_t (the sign of the unemployment rate is flipped) according to the hypothesis proposed by Nekarda and Ramey (2020). However, the estimated coefficient is essentially zero. This result is in line with the story whereby firms compensate workers so

Table 3.11: History Dependence of Separation Rates

	(1)	(2)	(3)
x_t	0.001	0.001	0.001
	(0.002)	(0.003)	(0.003)
$x_t^{fore,ave}$		–0.009*	
		(0.004)	
$x_t^{fore,min}$			–0.006
			(0.004)
R-squared	0.74	0.74	0.74
N of categories	18	18	18
N of observations	498	480	498

SEs clustered by category are reported in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Dependent variables are the average annual separation rates within 10 years after hiring. For independent variables, x_t is the current unemployment rate, which represents the labor market condition as of hiring. $x_t^{fore,ave}$ and $x_t^{fore,min}$ are the average and lowest unemployment rates within 10 years after hiring, respectively, i.e., $x_t^{fore,ave} = 1/10(x_{t+1} + ... + x_{t+10})$ and $x_t^{fore,min} = \min\{x_{t+1}, ..., x_{t+10}\}$. Fixed effects for categories are included in the regression.

that they will remain in the current job in expansions. It is also consistent with the relatively high cyclicality of the average wage reported in Table 3.3. In columns (2) and (3), the average and lowest unemployment rates in the course of tenure, $x_t^{fore,ave}$ and $x_t^{fore,min}$, are expected to have positive coefficients if booms after hiring facilitate job changes. The regression does not support the hypothesis, while the weakly significant negative coefficient of $x_t^{fore,ave}$ is rather consistent with job losses in recessions. In a nutshell, the evidence suggests that our sample is well suited for assuming exogenous separation rates.

Note that our measure of the separation rate is the share of workers who leave their current jobs and includes both employment-to-employment (EE) and employment-to-unemployment (EU) flows. In contrast, previous studies often define the separation rate using the EU flow (e.g., Elsby et al. 2013). Since the EE flow is procyclical and the EU flow is countercyclical in the data, it is not surprising that our definition of separation rate is acyclical.[20]

3.5.3 Cyclicality of Job-Match Quality in Alternative Data

Our study's central assumption is that new graduates are less subject to cyclical changes in job-match quality than job changers. However, a large literature indicates that the wages for new graduates are also determined by economic conditions at the time they enter the labor market (e.g., Kahn 2010). While we showed the lower cyclicality of new graduates' wages compared with job changers in Table 3.3 of Section 3.4, this section further assesses the validity of our assumptions in the Japanese labor market. To this end, we use individual workers' panel data in the Japan Household Panel Survey (JHPS/ KHPS) instead of the BSWS.

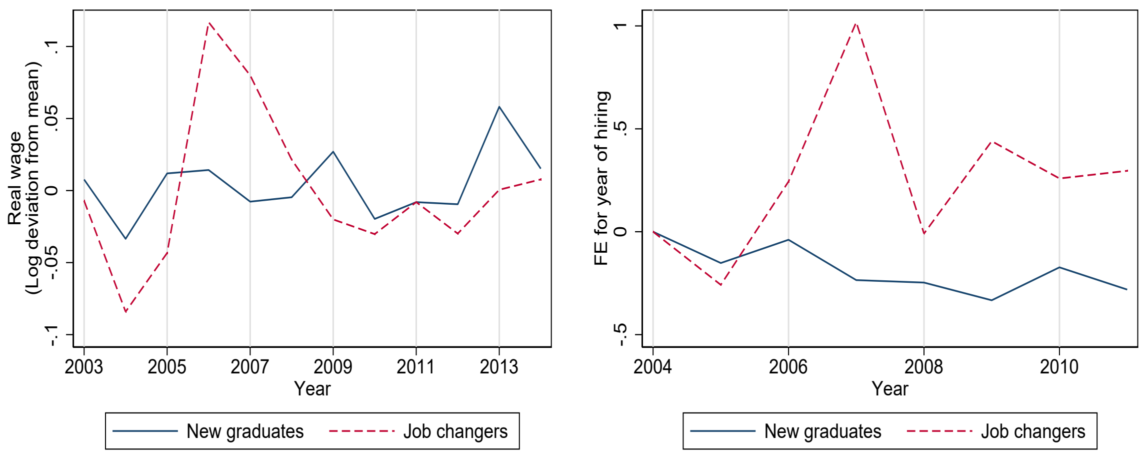
JHPS/KHPS. The JHPS/KHPS has been conducted annually since 2003 by the Panel Data Research Center at Keio University. Similar to the Panel Study of Income Dynamics (PSID) in the U.S. and the European Community Household Panel (ECHP) in Europe, tracks the same individuals' consumption and employment activities—including employment status, annual income, and hours worked—over time. Survey subjects are selected to represent the country's population aged 20 and older at the time of sampling. Subjects' spouses also receive the questionnaire and are asked to respond separately about their own employment conditions. As of 2014, 5,322 individuals answered the survey, with 1,703 individuals working as full-time employees.

We restrict the sample to new graduates and job changers. Full-time workers under 26

[20] While the decomposition of EE and EU flows is not available in the BSWS, we calculate these flows according to the Employment Trend Survey. The correlation with the unemployment rate is 0.61 for the EU flow ratio, -0.41 for the EE flow ratio, and 0.15 for our definition of separation rate. Each ratio is calculated as a share of the total regular workers and detrended by the HP filter. The sample period is from 1991 to 2019.

Figure 3.2: Time-Series of Fixed Effects for Individual Matches (Proxy for Job-Match Quality)

(A) Fixed Effect of a Match (B) Fixed Effect of Wages 8 Years after Hiring



Notes: In Panel (A), values are the average fixed effects of each match on real hourly wage after removing a quadratic time trend. The fixed effect is estimated as the mean of residual wages from hiring until separation for each job, after controlling for age, age square, educational attainment, and firm size. The number of identified new graduates and job changers is 83 and 502, respectively. In Panel (B), values are the coefficients of year fixed effects for wages for workers with 8 years of tenure with the same controls as Panel (A). The numbers of identified new graduates and job changers are 43 and 51, respectively, in this case.

years old when they started working are classified as new graduates. Job changers are workers below 55 years old who had changed jobs since the previous year; one question on the survey distinguishes between new graduates and job changers. We focus on male workers, as is often the case in previous studies. We identify 83 new graduates and 502 job changers. Our wage measure is real hourly earnings, in line with our empirical analysis in earlier sections. Hourly earnings are obtained by dividing annual total cash earnings—the sum of monthly earnings multiplied by 12 and bonuses—by total hours worked and are deflated by the GDP implicit price deflator.

Measure of Job-Match Quality. Under the assumption that the job-match quality lasts until separation, we capture job-match quality as a fixed effect of individual wages in a match. Specifically, we first regress the logarithm of wages on age, age square, educational attainment, and firm size. A fixed effect is estimated as the mean of residual wages from hiring until separation for each match. Note that this fixed effect also includes the unobserved time-invariant worker and firm productivity of a match. This approach reflects the spirit of Bils et al. (2023), who use the long-run wage in a match to control for its quality in estimating the UCL. As they explain, the approach differs from the use of workers' fixed effects because match-specific quality can fluctuate within a worker.

We also use Bils et al. (2023)'s proposed measure, which is the wage 8 years after hiring as an alternative measure of job-match quality. We run a regression of the logarithm of wages for workers with 8 years of tenure on age, age squared, educational attainment, firm size, and a fixed effect for the year of hiring.

Figure 3.2 shows the time-series of the estimated fixed effects as a proxy for job-match quality. Though the series are somewhat jagged, presumably due to small sample size, the fixed effects of job changers track the boom and bust around the global financial crisis in both panels. In contrast, fluctuations in the fixed effects of new graduates are muted and do not have a clear cyclical pattern. They suggest the presence of cyclical upgrading through job changes and acyclicality of the average match quality of new graduates.

We take the evidence as suggestive only, given the small sample size and short period. Furthermore, we cannot deny that the job quality of new graduates is affected by economic conditions in Japan, as discussed by Genda et al. (2010). However, the evidence in this section supports our assumption that new graduates are less affected by cyclical changes in job-match quality than job changers.

3.6 Model

This section develops a model to reconcile our empirical findings. Since we have obtained ample empirical support for the implicit wage contract of Beaudry and DiNardo (1991), our starting point is Rudanko (2009), who introduces an implicit wage contract to a search and matching framework. The resulting incumbent-workers' wage rigidity generates overall high cyclicality of the UCL. Our model is also intended to capture the asymmetric cyclicality of the UCL. To this end, the extension involves productivity heterogeneity on both sides of firms and workers. High-productivity firms seek to match skilled workers by using wages as a screening tool under imperfect information regarding the type of workers. We demonstrate that firms maintain the high value of a posted contract in recessions to keep their labor market competitive enough to exclude applications from unskilled workers. Consequently, the extended model replicates our empirical findings—namely, the acyclicality of the UCL in recessions despite its overall high cyclicality.

3.6.1 Environment

Matching Framework. The model environment follows that of Rudanko (2009), except for the heterogeneity of firms and workers. A firm-worker match is attained by directed search in the spirit of Moen (1997): Firms post a wage contract in a submarket and workers choose which submarket to apply to. Specifically, in the beginning of each period, firms post vacancies in a submarket i with a wage contract σ_i, which specifies state-contingent period wages:

$$\sigma_i(z_t) = \left\{ w_{it+\tau}(z^{t:t+\tau}) \in [\underline{w}, \overline{w}] \text{ for all } z^{t:t+\tau} = (z_t, z_{t+1}, ..., z_{t+\tau}) \right\}_{\tau=0}^{\infty}, \tag{3.12}$$

where z_t denotes aggregate productivity. z_t takes one of the values in a set, $Z = \{z_1, z_2, ...z_K\}$, with $z_k < z_{k+1}$ for all $k = 1, 2, ..., K$, and follows a stationary first-order Markov process with transition probabilities $\pi(z_{t+1}|z_t)$. Notice that the UCL can be defined as in Section 3.2.

Within each submarket, a worker-firm match is formed according to a matching function $m_i(u_{it}, v_{it})$, where u_t is the measure of unemployed workers searching for a measure v_t of vacancies. Following standard assumptions, $m_i(u_{it}, v_{it})$ is a constant return-to-scale Cobb-Douglas function:

$$m_i(u_{it}, v_{it}) = \kappa_i u_{it}^{\alpha_i} v_{it}^{1-\alpha_i}, \tag{3.13}$$

where the function is submarket-specific with parameters κ_i and α_i. Workers' job-finding rate is given by $m_i(u_{it}, v_{it})/u_{it} = m_i(1, \theta_{it}) = \mu_i(\theta_{it})$, with $\theta_{it} = v_{it}/u_{it}$ being a measure of labor market tightness. Similarly, the arrival rate of workers for a vacancy is given by $m_i(u_{it}, v_{it})/v_t = q_i(\theta_{it})(= \mu_i(\theta_{it})/\theta_{it})$.

Workers. There is a continuum of workers on a unit interval. Workers provide one unit of labor service and receive labor income if employed and benefits if unemployed.[21] They consume their income each period under preferences $\mathbb{E}_t[\sum_{\tau=0}^{\infty} u(c_{t+\tau})]$, with $u(\cdot)$ being a CRRA utility function. A wage contract terminates with an exogenous separation rate. Once it terminates, workers become unemployed and search for another job. Job search is costless for workers.

There are two types of workers, skilled (S) and unskilled (N). Their skill difference materializes when matched with a specific type of firm, which will be described shortly. Their wages differ depending on the type of firms they match with.

Let U_{ji} and E_{ji}^{σ} denote type-j worker's value of unemployment and employment, respectively, with a wage contract σ in submarket i. They can be written as

$$U_{ji}(z_t) = u(b) + \beta \mathbb{E}_t \left[\mu_i(\theta_i(z_{t+1})) E_{ji}^{\sigma}(z_{t+1}) + (1 - \mu_i(\theta_i(z_{t+1}))) U_{ji}(z_{t+1}) \right], \tag{3.14}$$

$$E_{ji}^{\sigma}(z_t) = u(w_i^{\sigma}(z_t)) + \mathbb{E}_t \left[\sum_{\tau=1}^{\infty} \beta^{\tau} (1 - s_i)^{\tau-1} \left\{ (1 - s_i) u(w_i^{\sigma}(z^{t:t+\tau})) + s_i U_{ji}(z_{t+\tau}) \right\} \right], \tag{3.15}$$

[21] We abstract variations in hours worked for simplicity. Thomas and Worrall (2007) show that the main implications of an implicit contract model are preserved under variable hours.

for $j = S, N$, where β is the subjective discount factor, b is the unemployment benefit, and s_i is the separation rate in each submarket i. Workers choose the submarket that maximizes the expected surplus of employment V^σ_{ji}:

$$V^\sigma_{ji}(z_t) = \mu_i(\theta_i(z_t))(E^\sigma_{ji}(z_t) - U_{ji}(z_t)). \tag{3.16}$$

Firms. There are two types of firms with different productivity, high (H) and low (L). Type-H firms offer non-routine jobs, and thereby achieve higher productivity than type-L firms, but production only occurs when matched with skilled workers. In contrast, type-L firms offer routine jobs, which are doable by both types of workers. The productivity of skilled and unskilled workers is identical when matched with type-L firms.

Type-H firms offer a state-contingent wage contract under limited commitment on the worker side, which is the focus of our analysis. Type-L firms are assumed to offer a wage contract under full commitment for simplicity.

Let F^σ_i denote type-i firm's value of a matched wage contract σ. It can be written as

$$F^\sigma_i(z_t) = a_i z_t - w^\sigma_i(z_t) + \mathbb{E}_t\left[\sum_{\tau=1}^{\infty} \beta^\tau (1 - s_i)^\tau (a_i z_{t+\tau} - w^\sigma_i(z^{t:t+\tau}))\right], \tag{3.17}$$

for $i = H, L$. Firms produce output $a_i z_t$ using one unit of labor and pay period wages determined under a contract σ. Once a contract terminates, the firm's expected value is zero under the free-entry assumption, as explained below. Firm-specific productivity a_i is given by

$$a_H = \begin{cases} \bar{a} > 1 & \text{if matched with a type-}S\text{ worker} \\ 0 & \text{if matched with a type-}N\text{ worker} \end{cases}, \tag{3.18}$$

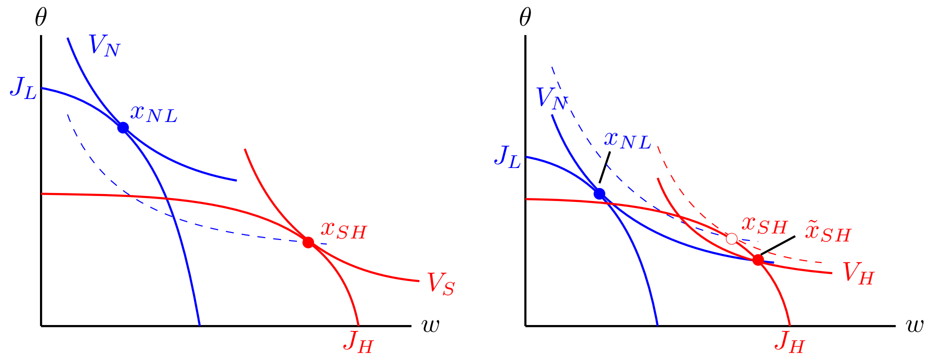
$$a_L = 1. \tag{3.19}$$

Firms have to pay vacancy posting cost k_i in each period. Free entry means that the expected profit of job posting J_i becomes zero:

$$J_i(z_t) = -k_i + q_i(\theta_i(z_t))F^\sigma_i(z_t) = 0. \tag{3.20}$$

Figure 3.3: Illustration of Equilibrium

(A) Non-Binding Exclusion Constraint (B) Binding Exclusion Constraint



3.6.2 Equilibrium

Separation of Submarkets. Type-H and L firms offer different wage contracts in each submarket. As Moen (1997) explains, two types of workers are separated into different submarkets if firms announce skill requirements and screen applications accordingly. In our case, type-H firms seek to match with skilled workers and type-L firms are indifferent in terms of worker type. Thus, a natural separation is the type-H firm–type-S worker pair and type-L firm–type-N worker pair, as depicted in Panel (A) of Figure 3.3.[22]

We consider an alternative circumstance in which the screening by type-H firms is imperfect. Specifically, if unskilled (type-N) workers apply to type-H firms, type-H firms can detect the worker type and decline the application with probability $p \in (0, 1)$. There remains a positive probability $(1 - p)$ with which type-H firms may hire the unintended type of worker (type-N). In such a case, production cannot occur due to the lack of required skill of the type-N worker, but the type-H firm has to pay wages because a contract is already signed.

To avoid this undesirable consequence, type-H firms use their posted wage contract as a

[22]As we will formalize shortly, a directed search equilibrium is located on a Pareto frontier on the (w, θ) plane as a result of workers' choice of submarket and firms' profit maximization.

screening tool. Specifically, a high value of a wage contract can render the submarket "too competitive" for type-N workers, while maintaining applications from type-S workers. This differentiation is possible with a single posted contract, because workers' objective function—i.e., the expected surplus—takes into account the job-finding rate, and therefore a change in the value of the wage contract can have different impacts on each type of worker. Since type-N workers face a lower job-finding rate than type-S workers in type-H firms' submarket due to screening probability p, the marginal benefit of a higher value contract is smaller. As a result, type-N workers may not tolerate the low labor market tightness (low job-finding rate) associated with a high value of the wage contract posted in the submarket.

This situation is illustrated in Panel (B) of Figure 3.3. The contact point of type-H firm's profit frontier and type-S worker's indifference curve (x_{SH}) renders type-N workers better off when they apply to type-H firm's submarket. On the other hand, a lower job-finding rate of type-L workers implies a flatter indifference curve on the (w, θ) plane, which leaves them in type-L firm's submarket when type-H firms post a wage contract with a high enough value ($\tilde{x}_{SH}$). This can be described as an exclusion constraint of type-N workers in type-H firms' problem:[23]

$$V^{\sigma}_{NL}(z_t) \geq V^{\sigma}_{NH}(z_t). \tag{3.21}$$

We assume that type-L firms take the type-H firms' screening as given and do not change their own behavior.

Equilibrium. Our definition of a directed search equilibrium is in line with that of Moen (1997) and Rudanko (2009) and is presented in Appendix 4.8. We confirm that risk-neutral firms post a fixed wage contract to risk-averse workers while offering wage increases in expansions to ensure workers' participation under their limited commitment, as shown in previous

[23]Type-S workers' participation in type-H firm's submarket, $V_{SH}(z_t) \geq V_{SL}(z_t)$, is also necessary to ensure the separation of workers in each submarket. However, this condition is satisfied under $\bar{a} > 1$ and reasonable values of other parameters for an intuitive reason.

studies. These account for incumbent-workers' wage rigidity.

The novelty of our model is the addition of the exclusion constraint (3.21). Although we leave the investigation of its impact for quantitative analysis in the following subsection, our intuition suggests that the constraint becomes more binding for lower z. Lower productivity implies a smaller surplus of a match for workers, since the value of employment declines, while that of unemployment is bounded by the unemployment benefit. Hence, the marginal benefit of ensuring higher labor market tightness diminishes. Workers lean toward a high wage contract with low market tightness offered by type-H firms, rather than a low wage contract of type-L firms, which enables type-H firms to post a high-value contract to satisfy the exclusion constraint.[24,25]

3.6.3 Numerical Analysis

Calibration. Since the numerical analysis aims to compare the qualitative implications of the model with empirical observations, calibration is conducted in a parsimonious manner to capture key features of the Japanese labor market. In particular, the Japanese labor market is known for the duality between regular and non-regular workers (e.g., OECD 2009). These different segments of the labor market exhibit quite different wage levels and employment practices, and the transition to the other segment is contained when workers start from either regular or non-regular employment in their initial jobs. We view these regular and

[24] Although our model is stylized, we do not preclude other channels in reality through which unskilled workers are more willing to apply to high-productivity firms in recessions. For example, Mukoyama et al. (2018) show that workers intensify their search efforts in recessions, while Engbom (2021) argues that unemployed workers apply for more jobs that they are less likely to be a good fit for.

[25] Downward rigidity arises for skilled workers who apply to high-productivity firms, whereas unskilled workers are not subject to it. As we explain shortly, regular and non-regular workers in the BSWS are proxies for skilled and unskilled workers in the model. Thus, one way to empirically test this proposition is to compare the cyclicality of the UCL between these two worker groups. However, the BSWS lacks a sufficiently long time-series for non-regular workers to examine the asymmetry in their UCL's cyclicality as it starts only in 2001. Instead, Appendix 4.7 shows no significant asymmetry in the average and new-hire wages' cyclicality for non-regular workers. This is in contrast to the findings for regular workers presented in Section 3.4.3, and supports the model's proposition that there is no downward rigidity on unskilled workers' UCL.

non-regular labor markets as a counterpart of employment created by type-H and L firms in the model. As our empirical analysis is based on regular workers' wages, our focus here is on the wage dynamics of type-H firms.

Calibrated parameters are listed in Panel (A) of Table 3.12. The time frequency of the model is quarterly. Parameters in the matching function are set according to the JEPSS. The number of vacancies and job-seekers recorded in the JEPSS enables us to construct empirical counterparts of the vacancy-filling rate, q, job-finding rate, μ, and labor market tightness, θ. We use these series of regular workers to calibrate the parameters of type-H firms and those of non-regular workers for type-L firms. Specifically, we regress the HP-filtered values of $\ln q_{it}$ on those of $\ln \theta_{it}$ in the sample for 1972–2019 to obtain $\alpha_H = 0.57$ and $\alpha_L = 0.78$. These estimates are broadly in line with those in U.S. and European countries' data surveyed by Petrongolo and Pissarides (2001). κ_i is set consistent with the time average, i.e., $\kappa_i = \bar{q}_i \bar{\theta}_i^{\alpha_i}$ with $\bar{x}$ denoting the average value in the sample, which yields $\kappa_H = 0.26$ and $\kappa_L = 0.38$. The separation rate for type-H firms, s_H, is the one in the BSWS data reported in Table 3.2, whereas that of type-L firms s_L is obtained from the Survey on Employment Trends (SET) conducted by the MHLW. Though somewhat smaller than the BSWS, the SET surveys around 15,000 firms following a sampling procedure similar to the BSWS. Job-posting costs k_i are calibrated to match μ_i in the model to the data using the simulated method of moments. The productivity of type-H firms $\bar{a}$ is targeted to the wage gap between type-H and L firms. These parameter values capture a sharp contrast between regular (type-H) and non-regular (type-L) labor markets. Regular workers enjoy high productivity (high $\bar{a}$) under a long-term contract (low s) while requiring a costly recruiting process on both the worker (low μ) and firm side (high k). The unemployment benefit b is set to the average replacement rate 0.55 reported in the SET. Regarding workers' preference, the subjective discount factor β is set to $0.97^{1/4}$. The period utility function is assumed to take the logarithm. The evolution of aggregate productivity is parameterized as a discretized AR(1) process with the persistence parameter $\rho_z = 0.90$ and the standard deviation of shocks $\sigma_z = 0.015$. The screening probability of type-H firms is set to $p = 0.83$ so that downward wage rigidity binds at the lower quartile of productivity.

Table 3.12: Calibration

(A) Calibrated Parameters

Description	Symbol	Value	Source
Elasticity of q to θ in type-H firms	α_H	0.57	JEPSS
Elasticity of q to θ in type-L firms	α_L	0.78	JEPSS
Constant in matching function for type-H firms	κ_H	0.26	JEPSS
Constant in matching function for type-L firms	κ_L	0.38	JEPSS
Separation rate for type-H firms	s_H	0.022	BSWS
Separation rate for type-L firms	s_L	0.064	SET
Unemployment benefit	b	0.55	SET
Job-posting cost for type-H firms	k_H	1.151	Internally calibrated
Job-posting cost for type-L firms	k_L	0.161	Internally calibrated
Productivity of type-H firms	$\bar{a}$	1.57	Internally calibrated
Subjective discount factor	β	$0.97^{1/4}$	Externally fixed
Persistence of aggregate productivity	ρ_z	0.90	Externally fixed
Size of aggregate productivity shock	σ_z	0.015	Externally fixed
Probability of screening in type-H firms	p	0.85	Externally fixed

Notes: Abbreviations are as follows: JEPSS: Job/Employment Placement Services Statistics; SET: Survey on Employment Trends.

Simulation Results. Figure 3.4 shows the dynamics of the UCL, new-hire wage, and average wage of type-H firms (the counterpart of our empirical wage measures), as well as the underlying productivity process, in simulated model data. It is immediate to see

Table 3.13: Calibration (cont.)

(B) Targeted Moments

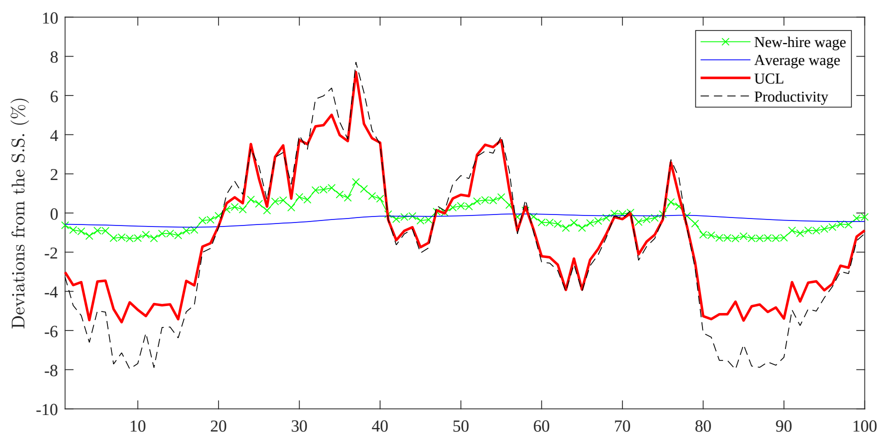
Description	Symbol	Model	Data	Source
Vacancy to job-seeker ratio of type-H firms	$\bar{\theta}_H$	0.80	0.80	JEPSS
Vacancy to job-seeker ratio of type-L firms	$\bar{\theta}_L$	1.51	1.51	JEPSS
Arrival rate of workers of type-H firms	$\bar{q}_H$	0.29	0.29	JEPSS
Arrival rate of workers of type-L firms	$\bar{q}_H$	0.28	0.28	JEPSS
Job-finding rate of type-H firms	$\bar{\mu}_H$	0.21	0.21	JEPSS
Job-finding rate of type-L firms	$\bar{\mu}_L$	0.38	0.38	JEPSS
Wage gap between type-H and L firms	$\bar{w}_H/\bar{w}_L$	1.51	1.51	BSWS

that the UCL is much more procyclical than the new-hire and average wages, in line with the empirical observation. Note that the average wage in the model is extremely rigid, presumably because it abstracts the intensive margin and wage adjustments of incumbent workers through overtime premiums and lump-sum bonuses.

It is also notable that the UCL does not fully track productivity in a downward phase. To interpret, the value of a posted wage contract remains high enough, even in a recession, to exclude type-N workers' participation in type-H firms' market.

The asymmetry of UCL's cyclicality is clearer in Figure 3.5, which compares the baseline case under imperfect information and an alternative case of perfect information. Under perfect information, type-H firms are able to detect worker type and decline unskilled workers' applications, if any, as a consequence of which skilled and unskilled workers are separated into the two submarkets. Thus, the flexibility of the UCL is present both in expansions and contractions. In the presence of imperfect information, in contrast, the room for downward adjustments of the UCL is limited. Since the UCL serves as the price of labor in the model, the limited adjustments in contractions lead to enlarged fluctuations in labor market tightness, θ_t, and the unemployment rate, u_t.

Figure 3.4: Model Simulation



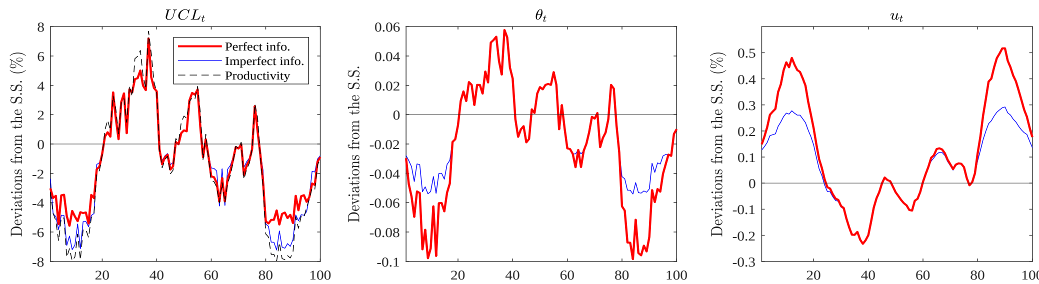
Notes: The policy function of each variable is obtained through the value function iteration on discretized grids. In the model simulation, aggregate productivity is generated as a continuous AR(1) series, and endogenous variables are simulated using the linear interpolation of the policy function.

In relation to the literature, the asymmetry of labor market dynamics has often been documented in previous studies (e.g., Neftçi 1984) and various models—in particular, models with downward wage rigidity—have been proposed (e.g., Fahr and Smets 2010; Dupraz et al. 2021). The novelty of this paper's model is that the asymmetry emerges endogenously through wage posting under heterogeneity of firms and workers.

Moreover, the resulting larger unemployment fluctuations contribute to addressing Shimer (2005)'s puzzle, which is also observed in the Japanese labor market (Esteban-Pretel et al. 2011). It is all the more challenging to address in the context of the UCL, because the larger cyclicality of the UCL than the average wage implies less unemployment fluctuation (Kudlyak 2014). Although our model is intended to derive qualitative implications and thus is not equipped to fully account for business cycle dynamics, the standard deviation of the unemployment rate is considerably higher in the presence of the imperfect information in our model, which partly accounts for the puzzle.[26]

[26]To be precise, the standard deviation of the unemployment rate is 37% higher under imperfect informa-

Figure 3.5: Model Simulation with and without Imperfect Information



3.7 Conclusion

In this paper, we examine the cyclical properties of the user cost of labor (UCL). Whereas the literature sees an advantage of the UCL as a measure of allocative wage, the estimation involves empirical challenges. These include collecting a sequence of wages from hiring to separation and controlling for cyclical changes in the average quality of new job matches. We address these challenges by tracking the wages of new graduates in a large-scale Japanese wage survey. Our approach confirms nonnegligible cyclical fluctuations in the average job-match quality, but the UCL remains highly procyclical after correcting for them. The estimated UCL is more procyclical than the new-hire wage, because the rigidity of the incumbent-workers' wage amplifies the wage differences among hiring cohorts.

The high cyclicality of the UCL aligns with a series of studies on the topic starting from Kudlyak (2014), who questions the conventional wisdom regarding a rigid real wage. That is, the high cyclicality of the UCL poses a difficulty for a standard macroeconomic model with respect to accounting for unemployment and other business-cycle fluctuations. Our analysis reconciles the challenge by demonstrating asymmetry of the UCL's cyclicality. The UCL rises during booms but remains relatively flat during downturns. As numerous studies have pointed out, business cycles are characterized by a sharp contraction followed

tion under our calibration—0.13 under imperfect information and 0.10 under perfect information. Note that the corresponding moment in the Japanese data is 0.22 from 1972 to 2019.

by a gradual recovery. These asymmetric business-cycle fluctuations are consistent with the asymmetry of UCL's cyclicality. Indeed, we propose a directed search model in which firms use wages to screen a certain type of applicants—which results in asymmetric fluctuations in the UCL—and show that the model replicates the asymmetric labor market dynamics.

We acknowledge caveats of our data, in that the use of semi-aggregate series in the BSWS leaves room for potential composition bias if it is systematically linked to business cycles. As we explain in earlier sections, available quantitative and qualitative evidence shows that the impact would be relatively contained or locate our empirical results at the conservative end of the spectrum. While noting each dataset's merits and limitations, this paper's approach would complement recent advances in the literature that strive to identify wage cyclicality cleanly.

We conclude with potential applications of our elaborated UCL measure. First, the UCL would be useful for exploring the cyclicality of price markup. A key issue in estimating price markup is measurement of the marginal factor price (e.g., Bils et al. 2018). The UCL can elaborate on the conventional labor share, based on the average wage, to measure the (inverse of) markup. Second, a medium- and long-run trend of the UCL would be worth investigating. Potential implications include those for the cross-sectional and intergenerational wage inequality that may not be fully captured by the remitted wage. The consequences for a secular trend in the labor share could be another area of exploration.

CHAPTER 4

Appendix

4.1 Chapter 1 Empirical Appendix

4.1.1 Agenda 2010

Agenda 2010, introduced by Chancellor Gerhard Schröder's government from 2003 to 2005, was a comprehensive set of policy measures also known as the Hartz Reforms, named for committee's head, Peter Hartz. The reforms comprised several stages (Hartz I-IV).

The first and second stages (Hartz I and II) took effect in January 2003. They encompassed the establishment of new Personnel Service Agencies, support for further vocational education from the German Federal Labor Agency, and deregulation of the temporary work sector. The employment protection reform we are interested in is part of Hartz III, which also involved restructuring the Federal Labor Office and became effective in January 2004. The final reform, known as Hartz IV, became effective in January 2005 and incorporated a reduction in the duration of unemployment benefits and a new definition of acceptable jobs, accompanied by sanctions for refusing such positions.

Within the realm of these reforms, the literature has particularly focused on the consequences of shortening the duration of unemployment benefits in Hartz IV. For instance, Krausea and Uhlig (2012) calibrate Diamond-Mortensen-Pissarides search models to the German economy, and demonstrate that the reduction in unemployment benefits led to a decrease in the unemployment rate from 10.8% to 8.0%.

4.1.2 Impact of the Reform on Job Changes

The reduction in firing costs for establishments with 6 to 10 employees can encourage the establishments to create more jobs and thus increase the movement between them. To investigate these job changes in our data, I estimate the following equation:

$$1(6 \leq N \leq 10 \rightarrow 6 \leq N \leq 10)_{it} = \alpha + \beta_3 \cdot D_{\text{reform}} + \beta_4 \cdot X_{et}^{'} + \alpha_e + \varepsilon_{it}$$

The sample includes all workers in establishments with 6 to 10 employees from 1991 to 2019. It is not limited to workers who were employed by continuous establishments that operated during the periods covering both 2003 and 2004. The dummy variable $1(6 \leq N \leq 10 \rightarrow 6 \leq N \leq 10)_{it}$ equals 1 if the individual i moves within establishments with 6 to 10 employees in year t and otherwise 0. D_{reform} equals 1 after 2004 when the reform was implemented, as in the previous section. The coefficient of β_3 is of interest, because it captures the impact of the reform on the transition probability within establishments with 6 to 10 employees.

The vector of control variables, $X_{et}^{'}$, includes the linear trend, quadratic trend, age, a square of age, educational attainment, gender, and the mean wage of the employer. The term α_e represents the establishment fixed effect for establishments, and in the case of a worker changing jobs, it denotes the fixed effect for establishments before the job change. Consequently, workers in establishments in which all employees remain over the period are excluded from the sample. I omit the individual fixed effects to retain the sample of workers who have never changed jobs. I estimate standard errors using clustering by establishment and year. Instead of a nonlinear model such as the Probit model, I employ a linear model to avoid the incidental parameter problem that can arise from the fixed effect.

Table 4.1 shows the results. The coefficient of D_{reform} is statistically significantly positive in the first column. After the reform, workers are more likely to move within establishments with 6 to 10 employees.

In the second column, the sample includes all workers in establishments with fewer than 6 employees. The dependent variable is $1(N < 6 \rightarrow 6 \leq N \leq 10)_{it}$, which equals 1 if individual

Table 4.1: Impact of the Firing Cost Reform on Job Changes

	$1(6 \leq N \leq 10 \rightarrow 6 \leq N \leq 10)_{it}$	$1(N < 6 \rightarrow 6 \leq N \leq 10)_{it}$
D_{reform}	0.006***	0.010***
	(0.002)	(0.003)
Control	✓	✓
Est. FE	✓	✓
Period	1991-2019	1991-2019
N of obs.	529,527	548,123
Adjusted R^2	0.16	0.48

Notes: Control variables include the linear trend, quadratic trend, age, a square of age, educational attainment, gender, and the mean wage of the employer. SEs clustered by establishments and year are reported in parentheses. *$p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

i moves from one establishment with fewer than 6 employees to another establishment with 6 to 10 employees in year t and otherwise 0. The other specification is the same as in the first column.

The second column shows that the coefficient of D_{reform} is statistically significantly positive. After the reform, workers are more likely to move from establishments with fewer than 6 employees to establishments with 6 to 10 employees. These findings suggest that establishments with 6 to 10 employees create more jobs following the reform.

4.1.3 Data for Training Investment in EU KLEMS

The training investment in EU KLEMS is derived from data obtained through the EU Continuing Vocational Training Survey (CVTS). This survey gathers information on companies' expenditures on vocational training for their employees. The CVTS provides details on training costs, which encompasses both the opportunity costs for employees participating in training and the labor costs associated with trainers. To ensure consistency with national accounts, the reported costs are adjusted by multiplying them by the ratio of the compensation of employees in the national account to the total labor costs in the CVTS. Given that CVTS data are only available for the years 2005, 2010, 2015, and 2020, a time series for training is created through linear interpolation.

4.2 Chapter 1 Theoretical Appendix

In this appendix, I explain the model used for the quantitative analysis.

4.2.1 Hamilton–Jacobi–Bellman Equation

Let $W(h, z, r, i)$ be the value of the worker when employed in a firm with productivity z when paid piece rate r and under training policy i, $F(h, z, r, i)$ the corresponding value to the firm; and $J(h, z)$ the maximized joint value of the match. I show below that the maximized joint value does not depend on how it is split.

Consider a meeting between a newborn worker and a firm. The worker and firm bargain over a training policy and a piece rate, which leads them to adopt the training policy that maximizes their joint value and piece rate $w^n(h, z, i)$ determined by

$$W(h, z, w^n, i) = \beta J(h, z) \tag{4.1}$$

Consider the case in which the worker contracts a poaching firm. The new and the old firm Bertrand compete for the worker, such that the worker's outside option becomes the full value of the least productive match, $J(j, z)$. Let $w^e(h, z', z, i)$ denote the pay of a worker who was employed in firm z but gets poached by a firm $z' > z$ under training policy i. It is given by

$$W(h, z', w^e(h, z', z, i), i) = J(h, z) + \beta(J(h, z') - J(h, z)) \tag{4.2}$$

If the poaching firm is less productive, $z' < z$, the worker remains with her current firm, but possibly with an updated piece rate given by the maximum of $w^e(h, z, z', i)$ and her previous wage, r. Let $r(h, z, r, i)$ be the lowest outside offer that leads to a renegotiation of the current contract, given by $J(h, r(h, z, r, i)) + \beta(J(h, z) - J(h, r(h, z, r, i))) = W(h, z, r, i)$.

Given an optimal training policy that maximizes the joint value as well as the splitting rules (4.1)-(4.2), the value of a worker, $W(h, z, r, i)$, evolves according to

$$\begin{aligned}(\rho+\delta)W(h,z,r,i) = {} & r(1-i(h,z))zh + \frac{\mu}{\eta}\Big(i(h,z)zh\Big)^{\eta}\frac{\partial W(h,z,r,i)}{\partial h} \\ & + p\int_{r(h,z,r,i)}^{z}\Big(J(h,z') + \beta(J(h,z) - J(h,z')) - W(h,z,r,i)\Big)d\Gamma(z') \\ & + p\int_{z}^{\infty}\Big(J(h,z) + \beta(J(h,z') - J(h,z)) - W(h,z,r,i)\Big)d\Gamma(z')\end{aligned} \tag{4.3}$$

subject to $0 \leq W(h, z, r, u) \leq J(h, z)$. The effective discount rate includes the discount rate ρ and the exogenous job destruction rate δ. The worker receives a share r of net output today and grows her human capital. At rate p, she receives outside offers from the distribution Γ. If the new match is worse than her current one, she will remain with her current firm but possibly with uploaded pay. Otherwise, she switches firms.

Let $F(h, z, r, i)$ be the value to a firm of productivity z when employing a worker with human capital h under piece rate r and the training policy that maximizes worker and joint surplus,

$$\begin{aligned}(\rho+\delta)F(h,z,r,i) = {} & (1-r)(1-i(h,z))zh + \frac{\mu}{\eta}\Big(i(h,z)zh\Big)^{\eta}\frac{\partial F(h,z,r,i)}{\partial h} \\ & + p\int_{r(h,z,r,i)}^{z}\Big((1-\beta)(J(h,z) - J(h,z')) - F(h,z,r,i)\Big)d\Gamma(z') \\ & - p(1-\Gamma(z))F(h,z,r,i) - \delta 1_{\underline{z}\leq z}\varepsilon\end{aligned} \tag{4.4}$$

subject to $0 \leq F(h, z, r, i) \leq J(h, z)$. The parameter ε is the firing cost, and $\underline{z}$ is the threshold for firing costs. If the productivity is lower than $\underline{z}$, no firing cost is incurred.

Combining the value of a worker (4.3) and the value of a firm (4.4), imposing the fact that $J(h, z) = W(h, z, r, i) + F(h, z, r, i)$ and collecting terms, I get the following Hamilton–Jacobi–Bellman equation;

$$(\rho+\delta)J(h,z) = (1-i(h,z))zh + \frac{\mu}{\eta}\Big(i(h,z)zh\Big)^{\eta}\frac{\partial J(h,z)}{\partial h} \tag{4.5}$$
$$+ p\beta\int_{z}^{\infty}\Big(J(h,z')-J(h,z)\Big)d\Gamma(z') - \delta 1_{\underline{z}\leq z}\varepsilon$$

subject to $J(h,z)\geq 0$. Since the piece rate r does not enter this expression, this verifies the conjecture that the joint value does not depend on how it is split. The optimal training policy is that which maximizes (4.5). Therefore, it is given by the first-order condition

$$\Big(i(h,z)zh\Big)^{1-\eta} = \mu\frac{\partial J(h,z)}{\partial h} \tag{4.6}$$

4.2.2 Kolmogorov Forward Equation

Let $g(h,z)$ be the distribution of workers over productivity and human capital. $g(z,h)$ solves the Kolmogorov forward equation

$$g(z,h) = -\Big(\delta + p(1-\Gamma(z))\Big)g(z,h) - \frac{\mu}{\eta}(i(z,h)zh)^{\eta}\frac{\partial g(z,h)}{\partial h} \tag{4.7}$$
$$+\gamma(z)p\int_{0}^{z} g(z',h)dz'.$$

Workers flow out due to exogenous separation at δ, up the job ladder at rate $p(1-\Gamma(z))$. With probability $\gamma(z)$, workers receive an offer from a firm with productivity z, which they accept if they are employed lower down the job ladder.

4.2.3 Free Entry Condition

The free entry condition is

$$c = (1-\beta)q\int_{0}^{\infty}\left(\int_{h}\int_{0}^{z}(J(z,h)-J(z',h))g(z',h)dz'\right)d\Gamma(z). \tag{4.8}$$

In return for the flow cost of a vacancy c, the firm contracts a potential hire at a rate q. The term in the bracket is the payoff from contracting an employed potential hire. The new potential match draws productivity from $\Gamma(z)$, and the firm gets a slice $1-\beta$ of any match formed.

4.2.4 Equilibrium

The value of a match (4.5), optimal training (4.6), the law of motion for the evolution of the aggregate state (4.7), and the free entry condition (4.8) fully characterize the allocation of the decentralized equilibrium. Given an allocation, wages are determined by the value of a worker (4.3) and the wage policies (4.1)-(4.2) under the optimal training policy. The market-clearing condition is satisfied, and collected firing costs are disbursed as lump sum payments for each worker equally.

4.3 Chapter 2 Theoretical Appendix

4.3.1 Normalizing the Distribution

Differentiate (2.6) with respect to $\hat{z}$ and n and use the definition $z \equiv \hat{z}/\overline{z}(t)$,

$$\begin{aligned} \frac{1}{\overline{z}(t)} \frac{\partial}{\partial z} \frac{\partial}{\partial n} F(\hat{z}/\overline{z}(t), n, t) &= \frac{\partial}{\partial \hat{z}} \frac{\partial}{\partial n} \hat{F}(\hat{z}, n, t) \\ f(z, n, t) &= \overline{z}(t) \hat{f}(\hat{z}, n, t) \end{aligned} \tag{4.9}$$

Differentiate (4.9) with respect to t and use the chain rule to obtain the transformation of the time derivative:

$$\frac{\partial}{\partial t} f(\hat{z}/\overline{z}(t), n, t) - \frac{\hat{z}}{\overline{z}(t)} \frac{\overline{z}'(t)}{\overline{z}(t)} \frac{\partial}{\partial z} f(\hat{z}/\overline{z}(t), n, t) = \overline{z}'(t) \hat{f}(\hat{z}, n, t) + \overline{z}(t) \frac{\partial}{\partial t} \hat{f}(\hat{z}, n, t).$$

Define the growth rates of the technology frontier as $g(t) \equiv \overline{z}'(t)/\overline{z}(t)$. Use the definition of $g(t)$ and the definition $z \equiv \hat{z}/\overline{z}$(t),

$$\frac{\partial}{\partial t} f(z, n, t) - z g(t) \frac{\partial}{\partial z} f(z, n, t) = g(t) \overline{z}(t) \hat{f}(\hat{z}, n, t) + \overline{z}(t) \frac{\partial}{\partial t} \hat{f}(\hat{z}, n, t)$$

Use (4.9),

$$\frac{\partial}{\partial t} f(z, n, t) - z g(t) \frac{\partial}{\partial z} f(z, n, t) - g(t) f(z, n, t) = \overline{z}(t) \frac{\partial}{\partial t} \hat{f}(\hat{z}, n, t). \tag{4.10}$$

Next, we derive the law of motion for firm-level detrended productivity. Note that

$$\begin{aligned}
\frac{d}{dt}z &= \frac{d}{dt}\left(\frac{\hat{z}}{\overline{z}(t)}\right) \\
&= \frac{\frac{d}{dt}\hat{z}}{\overline{z}(t)} - \frac{\hat{z}}{\overline{z}(t)}\frac{\overline{z}'(t)}{\overline{z}(t)} \\
&= \frac{\frac{d}{dt}\hat{z}}{\overline{z}(t)} - g(t)z
\end{aligned}$$

Therefore, the drift of detrended productivity is

$$\mu_z(z,n,t) = \frac{\hat{\mu}_z(z,n,t)}{\overline{z}(t)} - g(t)z \tag{4.11}$$

Use (2.2), and then, use (2.6), (2.7) , and (2.8)[1],

$$\mu_z(z,n,t) = (\gamma(n) - g(t))\, z + Av(z,n,t)Z(t)\int \mathbb{1}_P(z,n,z',n',t)\alpha(z'/z)dF_n(z',n',t).$$

Similarly, we derive the law of motion for firm-level employment growth, which is a function of detrended variables. Substitute (2.6), (2.7), and (2.8) into (2.3),

$$\begin{aligned}
\mu_n(z,n,t) \equiv& \hat{\mu}_n(\hat{z},n,t) \\
=& Av(z,n,t)\int \mathbb{1}_P(z,n,z',n',t)dF_n(z',n',t) - A\mathtt{v}\frac{n}{\mathtt{n}}\int \mathbb{1}_P(x,n,x',n',t)dF_v(z',n',t).
\end{aligned} \tag{4.12}$$

Multiply (2.4) by $\overline{z}(t)$,

$$\begin{aligned}
\overline{z}(t)\frac{\partial}{\partial t}\hat{f}(\hat{z},n,t) =& -\frac{\partial}{\partial n}\Big(\hat{\mu}_n(\hat{z},n,t)\overline{z}(t)\hat{f}(\hat{z},n,t)\Big) - \frac{\partial}{\partial \hat{z}}\Big(\hat{\mu}_z(\hat{z},n,t)\overline{z}(t)\hat{f}(\hat{z},n,t)\Big) \\
& - \eta\overline{z}(t)\hat{f}(\hat{z},n,t) + \eta\int_0^{\overline{z}} \overline{z}(t)\hat{f}(\hat{z},n,t)d\hat{z}\hat{\boldsymbol{\delta}}(\overline{z})
\end{aligned}$$

[1]For any set of functions h and $\hat{h}$ such that $h(z,n) = \hat{h}(\hat{z},n)$, using change of variables, $\int\int \hat{h}(\hat{z},n)\hat{f}(\hat{z},n)d\hat{z}dn = \int\int h(z,n)zf(z,n)(1/z)dzdn = \int\int h(z,n)f(z,n)dzdn$. Therefore, $\int \hat{h}(\hat{z},n)d\hat{F}(\hat{z},n) = \int h(z,n)dF(z,n)$.

Use (4.9), (4.11), and (4.12),[2]

$$
\begin{aligned}
\frac{\partial}{\partial t} f(z,n,t) =& zg(t)\frac{\partial}{\partial z} f(z,n,t) + g(t) f(z,n,t) \\
& - \frac{\partial}{\partial n}\left(\mu_n(z,n,t) f(z,n,t)\right) - \frac{\partial}{\partial z}\left(\left(\mu_z(z,n,t) - g(t)z\right) f(z,n,t)\right) \\
& - \eta f(z,n,t) + \eta \int_0^1 f(z,n,t) dz \boldsymbol{\delta}(1)
\end{aligned}
$$

Because $\frac{\partial}{\partial z}\left(zg(t) f(z,n,t)\right) = zg(t)\frac{\partial}{\partial z} f(z,n,t) + g(t) f(z,n,t)$, we get

$$
\begin{aligned}
\frac{\partial}{\partial t} f(z,n,t) =& - \frac{\partial}{\partial n}\left(\mu_n(z,n,t) f(z,n,t)\right) - \frac{\partial}{\partial z}\left(\mu_z(z,n,t) f(z,n,t)\right) - \eta f(z,n,t) \\
& + \eta \int_0^1 f(z,n,t) dz \boldsymbol{\delta}(1)
\end{aligned}
$$

4.3.2 Normalizing the Value Function

Differentiate (2.5) with respect to $\hat{z}$ and rearrange,

$$
\frac{\partial}{\partial z}\Omega(z,n,t) = \frac{\partial}{\partial z}\hat{\Omega}(\hat{z},n,t). \tag{4.13}
$$

Differentiate (2.5) with respect to n and rearrange,

$$
\frac{\partial}{\partial n}\Omega(z,n,t) = \frac{\partial}{\partial n}\hat{\Omega}(\hat{z},n,t)/\overline{z}(t). \tag{4.14}
$$

Rearrange and differentiate (2.5) with respect to t,

$$
\frac{\partial}{\partial t}\hat{\Omega}(\hat{z},n,t) = \overline{z}'(t)\Omega(\hat{z}/\overline{z}(t),n,t) - \hat{z}\frac{\overline{z}'(t)}{\overline{z}(t)}\frac{\partial}{\partial z}\Omega(\hat{z}/\overline{z}(t),n,t) + \overline{z}(t)\frac{\partial}{\partial t}\Omega(\hat{z}/\overline{z}(t),n,t).
$$

Divided by $\overline{z}(t)$ and use the definition of $g(t) \equiv \overline{z}'(t)/\overline{z}(t)$ and the definition $z \equiv \hat{z}/\overline{z}(\mathrm{t})$,

$$
\frac{1}{\overline{z}(t)}\frac{\partial}{\partial t}\hat{\Omega}(\hat{z},n,t) = g(t)\Omega(z,n,t) - g(t)z\frac{\partial}{\partial z}\Omega(z,n,t) + \frac{\partial}{\partial t}\Omega(z,n,t). \tag{4.15}
$$

[2]For any set of functions h and $\hat{h}$ such that $h(z,n) = \hat{h}(\hat{z},n)$, $\frac{\partial}{\partial \hat{z}}\hat{h}(\hat{z},n) = \frac{\partial}{\partial z}h(z,n)\frac{dz}{d\hat{z}} = \frac{\partial}{\partial z}h(z,n)\cdot\frac{1}{\overline{z}(t)}$.

Divide (2.1) by $\overline{z}(t)$ and then substitute (4.13), (4.14), and (4.15), and use (2.6),

$$
\begin{aligned}
&(r(t) - g(t))\,\Omega(z,n,t) + \frac{\partial}{\partial t}\Omega(z,n,t) \\
&= \max_{v\geq 0}\; z - c(v)Z(t) \\
&+ Av \int \left[\Omega_n(z,n,t) + \alpha(z'/z)Z(t)\Omega_z(z,n,t) - \Omega_n(z',n',t)\right]^+ dF_n(z',n',t) \\
&+ (\gamma(n) - g(t))\, z\Omega_z(z,n,t) \\
&+ \eta\left[\Omega(1,n,t) - \Omega(z,n,t)\right]
\end{aligned}
$$

Use $r(t) = \rho + g(t)$,

$$
\begin{aligned}
&\rho\Omega(z,n,t) + \frac{\partial}{\partial t}\Omega(z,n,t) \\
&= \max_{v\geq 0}\; z - c(v)Z(t) \\
&+ Av \int \left[\Omega_n(z,n,t) + \alpha(z'/z)Z(t)\Omega_z(z,n,t) - \Omega_n(z',n',t)\right]^+ dF_n(z',n',t) \\
&+ (\gamma(n) - g(t))\, z\Omega_z(z,n,t) \\
&+ \eta\left[\Omega(1,n,t) - \Omega(z,n,t)\right]
\end{aligned}
$$

4.3.3 $\Omega_z(z,n) > 0$

Rewrite the problem in terms of $x = \log z$. Denote with an abuse of notation, $\Omega(x,n) = \Omega(e^z,n)$ and $v(x,n) = v(e^z,n)$. Also, denote $\hat{\alpha}(x,x') = \alpha(\exp(x'-x))/\exp(x)$. Then, the HJB equation becomes

$$
\begin{aligned}
\rho\Omega(x,n) = {} & \exp(x) - c(v(x,n))Z \\
& + Av(x,n) \int \left[\Omega_n(x,n) + \hat{\alpha}(x,x')Z\Omega_x(x,n) - \Omega_n(x',n')\right]^+ dF_n(x',n') \\
& + (\gamma(n) - g)\,\Omega_x(x,n) \\
& + \eta\left[\Omega(0,n) - \Omega(x,n)\right].
\end{aligned}
\tag{4.16}
$$

Denote $\zeta(x,n) = \Omega_x(x,n)$. Differentiate the Bellman equation (4.16) w.r.t. x and use the envelope theorem,

$$\begin{aligned}\rho\zeta(x,n) =& \exp(x)\\ & Av(x,n)\frac{\partial}{\partial x}\int\left[\Omega_n(x,n)+\hat{\alpha}(x,x')Z\Omega_x(x,n)-\Omega_n(x',n')\right]^+ dF_n(x',n')\\ & + (\gamma(n)-g)\,\zeta_x(x,n)\\ & - \eta\zeta(x,n).\end{aligned} \tag{4.17}$$

Let define the poaching indicator function as

$$\mathbb{1}_P(x,n,x',n') = \begin{cases} 1 & \text{if} \quad \Omega_n(x,n)+\hat{\alpha}(x,x')Z\Omega_x(x,n)-\Omega_n(x',n') > 0 \\ 0 & \text{otherwise} \end{cases}$$

Then,

$$\begin{aligned}&\frac{\partial}{\partial x}\int\left[\Omega_n(x,n)+\hat{\alpha}(x,x')Z\Omega_x(x,n)-\Omega_n(x',n')\right]^+ dF_n(x',n')\\ &= \int\frac{\partial}{\partial x}\left[\Omega_n(x,n)+\hat{\alpha}(x,x')Z\Omega_x(x,n)-\Omega_n(x',n')\right]^+ dF_n(x',n')\\ &= \int \mathbb{1}_P(x,n,x',n')\frac{\partial}{\partial x}\left[\Omega_n(x,n)+\hat{\alpha}(x,x')Z\Omega_x(x,n)-\Omega_n(x',n')\right] dF_n(x',n')\\ &= \int \mathbb{1}_P(x,n,x',n')\left[\zeta_n(x,n)+\hat{\alpha}_x(x,x')Z\zeta_x(x,n)+\hat{\alpha}(x,x')Z\zeta_x(x,n)\right] dF_n(x',n')\end{aligned}$$

The second equality follows because for any differentiable function $f(x)$,

$$\frac{\partial}{\partial x}\left[f(x)\right]^+ = \begin{cases} f'(x) & \text{if} \quad f(x) > 0 \\ \text{not differentialble} & \text{if} \quad f(x) = 0 \\ 0 & \text{if} \quad f(x) < 0 \end{cases}$$

and at $f(x) = 0$, the derivative $\frac{\partial}{\partial x}\left[f(x)\right]^+$ is bounded by $\frac{\partial}{\partial x}\left[f(x)\right]^+ \in [\min\{0, f'(x)\}$ $, \max\{0, f'(x)\}]$, and the measure of (x',n') that satisfies $\Omega_n(x,n)+\hat{\alpha}(x,x')Z\Omega_x(x,n)-\Omega_n(x',n') = 0$ is zero for any (x,n). Therefore, (4.17) is rewritten as

$$\rho\zeta(x,n) = \exp(x)$$

$$
\begin{aligned}
&+ Av(x,n)\int \mathbb{1}_P(x,n,x',n')dF_n(x',n')\zeta_n(x,n) \\
&+ Av(x,n)Z\int \mathbb{1}_P(x,n,x',n')\hat{\alpha}_x(x,x')dF_n(x',n')\zeta(x,n) \\
&+ Av(x,n)Z\int \mathbb{1}_P(x,n,x',n'))\hat{\alpha}(x,x')dF_n(x',n')\zeta_x(x,n) \\
&+ (\gamma(n)-g)\,\zeta_x(x,n) \\
&- \eta\zeta(x,n)
\end{aligned}
$$

$$
\begin{aligned}
&\left(\rho+\eta - Av(x,n)Z\int \mathbb{1}_P(x,n,x',n')\hat{\alpha}_x(x,x')dF_n(x',n')\right)\zeta(x,n) \\
&\quad = \exp(x) \\
&\qquad + Av(x,n)\int \mathbb{1}_P(x,n,x',n')dF_n(x',n')\zeta_n(x,n) \\
&\qquad + \left(\gamma(n) + Av(x,n)Z\int \mathbb{1}_P(x,n,x',n')\hat{\alpha}(x,x')dF_n(x',n') - g\right)\zeta_x(x,n)
\end{aligned}
$$

Now, define the "effective discount rate"

$$
R(x,n) = \rho+\eta - Av(x,n)Z\int \mathbb{1}_P(x,n,x',n')\hat{\alpha}_x(x,x')dF_n(x',n')
$$

Define the stochastic process

$$
\begin{aligned}
dx_t &= \left\{\gamma(n_t) + Av(x_t,n_t)Z\int \mathbb{1}_P(x,n,x',n')\hat{\alpha}\left(x_t,x'\right)dF_n\left(x',n'\right) - g\right\}dt \\
dn_t &= \left\{Av(x_t,n_t)\int \mathbb{1}_P(x,n,x',n')dF_n\left(x',n'\right)\right\}dt
\end{aligned}
\tag{4.18}
$$

We can now use the Feynman–Kac formula (Pham (2009)) to go back to the sequential formulation:

$$
\zeta(z,n) = \mathbb{E}\left[\int_0^\infty e^{-\int_0^t R(x_\tau,n_\tau)d\tau}\exp(x)dt \mid x_0 = z, n_0 = n, \{x_t,n_t\} \text{ follows (4.18)}\right]
$$

Because $\exp(x)$ is positive, $\zeta(z,n)$ is positive. This concludes the proof.

4.3.4 $\Omega_n(z,n) > 0$

Denote $\zeta(x,n) = \Omega_x(x,n)$ and $\hat{\alpha}(x,x') = \alpha(\exp(x'-x))/\exp(x)$ Differentiate the Bellman equation (4.16) w.r.t. n,

$$\begin{aligned}\rho\zeta(x,n) = &+ Av(x,n)\int \mathbb{1}_P(x,n,x',n')dF_n(x',n')\zeta_n(x,n)\\ &+ Av(x,n)Z\int \mathbb{1}_P(x,n,x',n')\hat{\alpha}(x,x')dF_n(x',n')\zeta_x(x,n)\\ &+ (\gamma(n)-g)\,\zeta_x(x,n)\\ &+ \gamma'(n)\Omega_x(x,n)\\ &+ \eta\,[\zeta(0,n)-\zeta(x,n)]\end{aligned}$$

$$\begin{aligned}\rho\zeta(x,n) = &\gamma'(n)\Omega_x(x,n)\\ &+ Av(x,n)\int \mathbb{1}_P(x,n,x',n')dF_n(x',n')\zeta_n(x,n)\\ &+ \left(\gamma(n) + Av(x,n)Z\int \mathbb{1}_P(x,n,x',n')\hat{\alpha}(x,x')dF_n(x',n') - g\right)\zeta_x(x,n)\\ &+ \eta\,[\zeta(0,n)-\zeta(x,n)]\end{aligned}$$

Define the stochastic process

$$\begin{aligned}dx_t &= \left\{\gamma(n_t) + Av(x_t,n_t)Z\int \mathbb{1}_P(x,n,x',n')\hat{\alpha}\left(x_t,x'\right)dF_n\left(x',n'\right) - g\right\}dt + (0-x_t)\,dH_t\\ dn_t &= \left\{Av(x_t,n)\int \mathbb{1}_P(x,n,x',n')dF_n\left(x',n'\right)\right\}dt\end{aligned} \tag{4.19}$$

where H_t is a compensated Poisson process of intensity η. Again, we can use the Feynman–Kac formula to go back to the sequential formulation:

$$\zeta(z,n) = \mathbb{E}\left[\int_0^\infty e^{-\rho}\gamma'(n)\Omega_x(x,n)dt \mid x_0 = z, n_0 = n, \{x_t,n_t\} \text{ follows (4.19)}\right]$$

Because $\gamma'(n) > 0$ by assumption and $\Omega_x(x,n)$ is positive from the previous proof, $\zeta(z,n)$ is positive. This completes the proof.

4.4 Chapter 2 Empirical Appendix

4.4.1 Data

Our analyses utilize two administrative data sets, "Linked Inventor Biography Data 1980-2014" (INV-BIO) and "Sample of Integrated Labor Market Biographies" (Stichprobe der Integrierten Arbeitsmarktbiografien - SIAB). Both data sets are constructed by the Institute for Employment Research (IAB).

The SIAB data is a 2% random sample from (Integrated Employment Biographies - IEB). The IEB combines data from five different sources, each of which may contain information from various administrative procedures. It comprises all individuals in Germany who hold at least one of the following employment statuses: employment subject to social security, marginal part-time employment, receipt of benefits according to the German Social Code III or II, official registration as a job seeker at the German Federal Employment Agency, and (planned) participation in programs of active labor market policies (Dauth and Eppelsheimer 2020 for more detail).

The patent information contained in the INV-BIO dataset is sourced from register data recorded in PATSTAT, which includes bibliographical, procedural, and legal status information on patent applications handled by the European Patent Office. Additionally, data from DPMAregister, the online patent register of the German Patent and Trademark Office, is incorporated to enhance the PATSTAT data extract. The DPMAregister provides exclusive records of national patent applications that are not transferred to the European Patent Office or filed under the PCT (Patent Cooperation Treaty) route. As a result, the INV-BIO dataset comprises inventors who are listed on patent filings at the European Patent Office (EPO) between 1999 and 2011 and have been successfully linked with IEB (Dorner et al. 2018 for more detail).

Table B.1 shows the summary statistics for INV-BIO and SIAB, respectively. Table B.2 shows the correlation between the three measures used as the proxy for knowledge quality or productivity level in Section 2.3.2 and 2.3.3. We can observe positive correlations between them.

Table 4.2: Summary Statistics

(A) INV-BIO

Establishment level variables	Mean	S.D.	N of est. (thus.)
N of inventors (n_{et})	4.9	18.5	119
N of employees	688.9	2150.6	119
Mean daily wage, Euro	121.6	55.5	119
N of three-year forward citations for patents (three-year backward average, z_{et})	11.3	69.2	119
Share of inventors moving from higher productivity est. (H-Share$_{et}$), %	61.2	49.5	119
Total inventor inflows	1.67	5.30	119

(B) SIAB

Worker level variables	Mean	S.D.	N of workers (thus.)
Dummy for moving to less productive est. (D_{it})			
based on est. size	0.50	-	4,669
based on mean wage	0.52	-	4,583
Dummy for the identified inventors (I_{it})	0.10	-	5,691
Daily Wage, Euro	44.4	42.1	5,691
Age	33.7	12.9	5,691
Share of Women, %	47.3	-	5,691

4.4.2 Robustness Check of Empirical Analyses

The table from B.4 to B.6 shows the result of the robustness check for our empirical result. Table B.4 shows the transition matrix of inventor flows with wage increases, suggesting many flows from more productive establishments to less productive ones, even in this sample.

Table 4.3: Correlation between Three Measures

Correlation	z_{et}	Est. size	Mean Wage
z_{et}	1.00		
Est. size	0.44	1.00	
Mean wage	0.11	0.08	1.00

Table 4.4: Distribution of Inventors

(A) Rank by Citation

Establishment percentile rank	≤50%	50-60	60-70	70-80	80-100
Share of inventors (%)	13.1	2.5	3.2	4.6	76.7

(B) Rank by Establishment Size

Establishment percentile rank	≤50%	50-60	60-70	70-80	80-100
Share of inventors (%)	9.3	3.9	5.2	7.5	74.0

(C) Rank by Mean Wage

Establishment percentile rank	≤50%	50-60	60-70	70-80	80-100
Share of inventors (%)	10.2	6.6	10.5	14.8	57.8

Notes: This table shows the distribution of inventors across percentiles of establishments. The percentile in panel (A) is based on the three-year backward average of forward patent citation counts. Panel (B) is based on the establishment size, and panel (C) is based on the mean wage of full-time employees. Establishments could be classified into different percentiles based on these measures each year. The sample encompasses data from 1980 to 2014. The values in the table represent the proportion of inventors within each percentile in relation to the total number of inventors in INV-BIO.

Instead of Probit model in Section 2.3.3, we estimate the following equation to control for fixed effects,

$$D_{it} = \beta_0 + \beta_1 I_{it} + \beta_2 X_{it} + \alpha_e + \alpha_t + \varepsilon_{it} \tag{4.20}$$

Definitions of variables are the same as in Section 2.3.3. Table B.5 shows that inventors are more likely to move to less productive establishments conditional on fixed effects.

4.5 Chapter 2 Quantitative Appendix

4.5.1 Numerical Solution to Joint Value HJB Equation

Change of variables

We want to solve

$$\begin{aligned}
\rho\Omega(z,n,i) = &\max_{v\geq 0} z - c(v)Z \\
&+ Av\int \left[\Omega_n(z,n,i) + \alpha(z'/z)Z\Omega_z(z,n,i) - \Omega_n(z',n',i')\right]^+ dF_n(z',n',i) \\
&+ (\gamma(n,i) - g)\, z\Omega_z(z,n,i) \\
&+ \eta\left[\Omega(1,n,i) - \Omega(z,n,i)\right] \\
&+ \lambda_i\left[\Omega(z,n,-i) - \Omega(z,n,i)\right]
\end{aligned}$$

As in our quantitative exercise, let $c(v) = \frac{\bar{c}}{\phi+1}v^{\phi+1}$. The first order condition for vacancies is

$$\begin{aligned}
\bar{c}v(z,n,i)^{\phi}Z &= A\int \left[\Omega_n(z,n,i) + \alpha(z'/z)Z\Omega_z(z,n,i) - \Omega_n(z',n',i')\right]^+ dF_n(z',n',i') \\
v(z,n,i) &= \left\{\frac{A}{\bar{c}Z}\int \left[\Omega_n(z,n,i) + \alpha(z'/z)Z\Omega_z(z,n,i) - \Omega_n(z',n',i')\right]^+ dF_n(z',n',i')\right\}^{\frac{1}{\phi}}
\end{aligned}$$

Consider a change of variables. Let $\tilde{z} = \log z, \tilde{n} = \log n$. Now, define $\tilde{\Omega}(\tilde{z},\tilde{n},i) = \Omega\left(e^{\tilde{z}}, e^{\tilde{n}}, i\right) = \Omega(z,n,i)$. Applying the chain rule to $\Omega(z,n,i) = \widetilde{\Omega}(\log z, \log n, i)$, and re-arranging:

$$\Omega_n(z,n,i)n = \widetilde{\Omega}_{\tilde{n}}(\widetilde{z},\tilde{n},i)$$

Table 4.5: Transition Probabilities of Inventor Flows with wage increase

(A) Rank by Citation/Inventor

Share of flows (%)		Destination establishment rank				
		$\leq$ 50%	50-60	60-70	70-80	80-100
	$\leq$ 50%	2.7	0.2	0.3	0.4	4.1
Origin	50-60	2.1	0.2	0.2	0.3	3.1
establishment	60-70	2.3	0.2	0.3	0.4	3.6
rank	70-80	2.7	0.2	0.3	0.4	3.6
	80-100	19.0	1.8	2.1	3.1	45.9

(B) Rank by Establishment Size

Share of flows (%)		Destination establishment rank				
		$\leq$ 50%	50-60	60-70	70-80	80-100
	$\leq$ 50%	3.8	1.2	0.9	0.9	6.6
Origin	50-60	0.4	0.8	0.9	0.5	2.1
establishment	60-70	0.4	0.2	1.1	1.2	2.8
rank	70-80	0.5	0.3	0.4	1.9	4.9
	80-100	3.5	1.5	2.1	3.1	58.2

(C) Rank by Mean Wage

Share of flows (%)		Destination establishment rank				
		$\leq$ 50%	50-60	60-70	70-80	80-100
	$\leq$ 50%	4.4	1.5	1.3	1.4	4.4
Origin	50-60	0.9	1.3	1.6	1.1	2.5
establishment	60-70	0.8	0.9	2.6	2.9	4.1
rank	70-80	0.7	0.6	1.6	4.8	7.4
	80-100	2.0	2.6	2.8	4.9	42.4

$$\Omega_z(z,n,i)z = \widetilde{\Omega}_{\tilde{z}}(\tilde{z},\tilde{n},i)$$

or equivalently,

$$\Omega_n(z,n,i) = \frac{\widetilde{\Omega}_{\tilde{n}}(\tilde{z},\tilde{n},i)}{e^{\tilde{n}}}$$

Table 4.6: Estimation Result for Inventor Flows (Linear Model)

	D_{it}			
	Whole sample		Sample with wage ↑	
I_{it}	.008***	.010**	.009***	.014***
	(.003)	(.004)	(.003)	(.004)
Control	√	√	√	√
Fixed Effects	√	√	√	√
Measure for D_{it}	Size	Mean wage	Size	Mean wage
N	2,938,537	2,959,368	1,609,460	1,617,613
Adj. R^2	.25	.22	.21	.20

$$\Omega_z(z, n, i) = \frac{\widetilde{\Omega}_{\tilde{z}}(\widetilde{z}, \tilde{n}, i)}{e^{\tilde{z}}}$$

As in our quantitative exercise, let $\alpha(z'/z) = \overline{\alpha}\,(z'/z)^{\beta}$. Then, we can rewrite it as

$$\begin{aligned}\alpha(z'/z) &= \overline{\alpha}\left(e^{\tilde{z}'}/e^{\tilde{z}}\right)^{\beta} \\ &= \overline{\alpha}e^{\beta(\tilde{z}'-\tilde{z})}\end{aligned}$$

As in our quantitative exercise, let $\gamma(n, i) = \overline{\gamma}_i n^{\delta}$. Then, we can rewrite it as

$$\gamma(n, i) = \overline{\gamma}_i e^{\delta\tilde{n}}$$

Define $\tilde{F}(\tilde{z}, \tilde{n}, i) = F(e^{\tilde{z}}, e^{\tilde{n}}, i) = F(z, n, i)$. The relationship between density of $F(z, n, i)$ and $\tilde{F}(\tilde{z}, \tilde{n}, i)$ is given by

$$\begin{aligned}f(z, n, i) &= \frac{\partial}{\partial z}\frac{\partial}{\partial n}F(z, n, i) \\ &= \frac{\partial}{\partial z}\frac{\partial}{\partial n}\tilde{F}(\log z, \log n, i) \\ &= \frac{1}{zn}\tilde{f}(\tilde{z}, \tilde{n}, i)\end{aligned}$$

When we change the variables from (z, n) to $(\tilde{z}, \tilde{n})$, the Jacobian is zn. Therefore, for any set of functions h and $\tilde{h}$ such that $h(z, n, i) = \tilde{h}(\tilde{z}, \tilde{n}, i)$,

$$\int \tilde{h}(\tilde{z}, \tilde{n}, i) d\tilde{F}(\tilde{z}, \tilde{n}, i) = \int h(z, n) dF(z, n, i)$$

The total output can be expressed as

$$\mathtt{Z} = \int e^{\tilde{z}} d\tilde{F}(\tilde{z}, \tilde{n}, i)$$

The total mass of the inventor can be expressed as

$$\mathtt{n} = \int e^{\tilde{n}} d\tilde{F}(\tilde{z}, \tilde{n}, i)$$

Let define

$$\tilde{f}_{\tilde{n}}(\tilde{z}, \tilde{n}, i) = \frac{e^{\tilde{n}} \tilde{f}(\tilde{z}, \tilde{n}, i)}{\mathtt{n}}$$

and $\tilde{F}_{\tilde{n}}(\tilde{z}, \tilde{n}, i)$ the corresponding cumulative distribution. Then, the inventor-weighted distribution $f_{\tilde{n}}(z, n, i)$ can be rewritten as

$$\begin{aligned} f_n(z, n, i) &= \frac{n f(z, n, i)}{\mathtt{n}} \\ &= \frac{1}{z} \cdot \frac{1}{n} \frac{e^{\tilde{n}} \tilde{f}(\tilde{z}, \tilde{n}, i)}{\mathtt{n}} \\ &= \frac{1}{z} \cdot \frac{1}{n} \tilde{f}_{\tilde{n}}(\tilde{z}, \tilde{n}, i) \end{aligned}$$

Therefore, for any set of functions h and $\tilde{h}$ such that $h(z, n, i) = \tilde{h}(\tilde{z}, \tilde{n}, i)$,

$$\int \tilde{h}(\tilde{z}, \tilde{n}, i) d\tilde{F}_{\tilde{n}}(\tilde{z}, \tilde{n}, i) = \int h(z, n) dF_n(z, n, i)$$

Define $\tilde{v}(\tilde{z}, \tilde{n}, i) = v(e^{\tilde{z}}, e^{\tilde{n}}, i) = v(z, n, i)$. Then,

$$\begin{aligned} \tilde{v}(\tilde{z}, \tilde{n}, i) = &\Big\{ \frac{A}{\bar{c}Z} \int [\widetilde{\Omega}_{\tilde{n}}(\widetilde{z}, \tilde{n}, i)/e^{\tilde{n}} + \overline{\alpha} e^{\beta(\tilde{z}' - \tilde{z})} Z \widetilde{\Omega}_{\tilde{z}}(\widetilde{z}, \tilde{n}, i)/e^{\tilde{z}} \\ &- \widetilde{\Omega}_{\tilde{n}}(\widetilde{z}', \tilde{n}', i')/e^{\tilde{n}}]^{+} d\tilde{F}_{\tilde{n}}(\tilde{z}', \tilde{n}', i') \Big\}^{\frac{1}{\phi}} \end{aligned}$$

The Bellman equation can be rewritten as

$$\rho \tilde{\Omega}(\tilde{z}, \tilde{n}, i) = e^{\tilde{z}} - \frac{\overline{c}}{\phi + 1} \tilde{v}(\tilde{z}, \tilde{n}, i)^{\phi+1} Z$$

$$
\begin{aligned}
&+ A\tilde{v}(\tilde{z},\tilde{n},i)\int[\widetilde{\Omega}_{\tilde{n}}(\widetilde{z},\tilde{n},i)/e^{\tilde{n}} + \overline{\alpha}e^{\beta(\tilde{z}'-\tilde{z})}Z\widetilde{\Omega}_{\tilde{z}}(\widetilde{z},\tilde{n},i)/e^{\tilde{z}} \\
&- \widetilde{\Omega}_{\tilde{n}}(\widetilde{z}',\tilde{n}',i')/e^{\tilde{n}'}]^{+}d\tilde{F}_{\tilde{n}}(\tilde{z}',\tilde{n}',i') \\
&+ \left(\overline{\gamma}_i e^{\delta\tilde{n}} - g\right)\widetilde{\Omega}_{\tilde{z}}(\widetilde{z},\tilde{n},i) \\
&+ \eta\left[\tilde{\Omega}(0,\tilde{n},i) - \tilde{\Omega}(\tilde{z},\tilde{n},i)\right] \\
&+ \lambda_i\left[\tilde{\Omega}(\tilde{z},\tilde{n},-i) - \tilde{\Omega}(\tilde{z},\tilde{n},i)\right]
\end{aligned}
$$

Let define poaching indicator function with transformed variable $\tilde{\mathbb{1}}_P(\tilde{z},\tilde{n},i,\tilde{z}',\tilde{n}',i')$ as

$$
\tilde{\mathbb{1}}_P(\tilde{z},\tilde{n},i,\tilde{z}',\tilde{n}',i') = \begin{cases} 1 & \text{if} \quad \widetilde{\Omega}_{\tilde{n}}(\widetilde{z},\tilde{n},i)/e^{\tilde{n}} + \overline{\alpha}e^{\beta(\tilde{z}'-\tilde{z})}Z\widetilde{\Omega}_{\tilde{z}}(\widetilde{z},\tilde{n},i)/e^{\tilde{z}} > \widetilde{\Omega}_{\tilde{n}}(\widetilde{z}',\tilde{n}',i')/e^{\tilde{n}} \\ 0 & \text{otherwise} \end{cases}
$$

The Bellman equation can be rewritten as

$$
\begin{aligned}
\rho\tilde{\Omega}(\tilde{z},\tilde{n},i) = {} & e^{\tilde{z}} - \frac{\overline{c}}{\phi+1}\tilde{v}(\tilde{z},\tilde{n},i)^{\phi+1}Z \\
&+ A\tilde{v}(\tilde{z},\tilde{n},i)/e^{\tilde{n}}\int\left[\tilde{\mathbb{1}}_P(\tilde{z},\tilde{n},i,\tilde{z}',\tilde{n}',i')\right]d\tilde{F}_{\tilde{n}}(\tilde{z}',\tilde{n}',i')\widetilde{\Omega}_{\tilde{n}}(\widetilde{z},\tilde{n},i) \\
&+ A\tilde{v}(\tilde{z},\tilde{n},i)\overline{\alpha}Z/e^{\tilde{z}}\int\left[\tilde{\mathbb{1}}_P(\tilde{z},\tilde{n},i,\tilde{z}',\tilde{n}',i')e^{\beta(\tilde{z}'-\tilde{z})}\right]d\tilde{F}_{\tilde{n}}(\tilde{z}',\tilde{n}',i')\widetilde{\Omega}_{\tilde{z}}(\widetilde{z},\tilde{n},i) \\
&- A\tilde{v}(\tilde{z},\tilde{n},i)\int\left[\tilde{\mathbb{1}}_P(\tilde{z},\tilde{n},i,\tilde{z}',\tilde{n}',i')\widetilde{\Omega}_{\tilde{n}}(\widetilde{z}',\tilde{n}',i')/e^{\tilde{n}'}\right]d\tilde{F}_{\tilde{n}}(\tilde{z}',\tilde{n}',i') \\
&+ \left(\overline{\gamma}_i e^{\delta\tilde{n}} - g\right)\widetilde{\Omega}_{\tilde{z}}(\widetilde{z},\tilde{n},i) \\
&+ \eta\left[\tilde{\Omega}(0,\tilde{n},i) - \tilde{\Omega}(\tilde{z},\tilde{n},i)\right] \\
&+ \lambda_i\left[\tilde{\Omega}(\tilde{z},\tilde{n},-i) - \tilde{\Omega}(\tilde{z},\tilde{n},i)\right]
\end{aligned}
$$

$$
\begin{aligned}
\rho\tilde{\Omega}(\tilde{z},\tilde{n},i) = {} & e^{\tilde{z}} - \frac{\overline{c}}{\phi+1}\tilde{v}(\tilde{z},\tilde{n},i)^{\phi+1}Z \\
&- A\tilde{v}(\tilde{z},\tilde{n},i)\int\left[\tilde{\mathbb{1}}_P(\tilde{z},\tilde{n},i,\tilde{z}',\tilde{n}',i')\widetilde{\Omega}_{\tilde{n}}(\widetilde{z}',\tilde{n}',i')/e^{\tilde{n}'}\right]d\tilde{F}_{\tilde{n}}(\tilde{z}',\tilde{n}',i') \\
&+ A\tilde{v}(\tilde{z},\tilde{n},i)/e^{\tilde{n}}\int\left[\tilde{\mathbb{1}}_P(\tilde{z},\tilde{n},i,\tilde{z}',\tilde{n}',i')\right]d\tilde{F}_{\tilde{n}}(\tilde{z}',\tilde{n}',i')\widetilde{\Omega}_{\tilde{n}}(\widetilde{z},\tilde{n},i) \\
&+ \overline{\gamma}_i e^{\delta\tilde{n}}\widetilde{\Omega}_{\tilde{z}}(\widetilde{z},\tilde{n},i)
\end{aligned}
$$

$$+ A\tilde{v}(\tilde{z},\tilde{n},i)\overline{\alpha}Z/e^{\tilde{z}} \int \left[\tilde{\mathbb{1}}_P(\tilde{z},\tilde{n},i,\tilde{z}',\tilde{n}',i')e^{\beta(\tilde{z}'-\tilde{z})}\right] d\tilde{F}_{\tilde{n}}(\tilde{z}',\tilde{n}',i')\widetilde{\Omega}_{\tilde{z}}(\tilde{z},\tilde{n},i)$$

$$- g\widetilde{\Omega}_{\tilde{z}}(\tilde{z},\tilde{n},i)$$

$$+ \eta\left[\tilde{\Omega}(0,\tilde{n},i) - \tilde{\Omega}(\tilde{z},\tilde{n},i)\right]$$

$$+ \lambda_i\left[\tilde{\Omega}(\tilde{z},\tilde{n},-i) - \tilde{\Omega}(\tilde{z},\tilde{n},i)\right]$$

Implicit method

We solve the Bellman equation using an implicit method. Let Δ denote step-size and τ the iteration of the algorithm. Then given $\tilde{\Omega}^{\tau-1}(\tilde{z},\tilde{n},i)$, the implicit method gives an update

$$\frac{1}{\Delta}\left[\tilde{\Omega}^{\tau}(\tilde{z},\tilde{n},i) - \tilde{\Omega}^{\tau-1}(\tilde{z},\tilde{n},i)\right] + \rho\tilde{\Omega}(\tilde{z},\tilde{n},i) =$$

$$e^{\tilde{z}} - \frac{\overline{c}}{\phi+1}\tilde{v}(\tilde{z},\tilde{n},i)^{\phi+1}Z - A\tilde{v}(\tilde{z},\tilde{n},i)\int \left[\tilde{\mathbb{1}}_P(\tilde{z},\tilde{n},i,\tilde{z}',\tilde{n}',i')\widetilde{\Omega}^{\tau-1}_{\tilde{n}}(\tilde{z}',\tilde{n}',i')/e^{\tilde{n}'}\right] d\tilde{F}_{\tilde{n}}(\tilde{z}',\tilde{n}',i')$$

$$+ A\tilde{v}(\tilde{z},\tilde{n},i)/e^{\tilde{n}} \int \left[\tilde{\mathbb{1}}_P(\tilde{z},\tilde{n},i,\tilde{z}',\tilde{n}',i')\right] d\tilde{F}_{\tilde{n}}(\tilde{z}',\tilde{n}',i')\widetilde{\Omega}^{\tau}_{\tilde{n}}(\tilde{z},\tilde{n},i)$$

$$+ \left(\overline{\gamma}_i e^{\delta\tilde{n}} + A\tilde{v}(\tilde{z},\tilde{n},i)\overline{\alpha}Z/e^{\tilde{z}} \int \left[\tilde{\mathbb{1}}_P(\tilde{z},\tilde{n},i,\tilde{z}',\tilde{n}',i')e^{\beta(\tilde{z}'-\tilde{z})}\right] d\tilde{F}_{\tilde{n}}(\tilde{z}',\tilde{n}',i') - g\right)\widetilde{\Omega}^{\tau}_{\tilde{z}}(\tilde{z},\tilde{n},i)$$

$$+ \eta\left[\tilde{\Omega}^{\tau}(0,\tilde{n},i) - \tilde{\Omega}^{\tau}(\tilde{z},\tilde{n},i)\right]$$

$$+ \lambda_i\left[\tilde{\Omega}^{\tau}(\tilde{z},\tilde{n},-i) - \tilde{\Omega}^{\tau}(\tilde{z},\tilde{n},i)\right]$$

Rearranging this expression:

$$\left(\frac{1}{\Delta} + \rho + \eta\right)\tilde{\Omega}^{\tau}(\tilde{z},\tilde{n},i)$$

$$- A\tilde{v}(\tilde{z},\tilde{n},i)/e^{\tilde{n}} \int \left[\tilde{\mathbb{1}}_P(\tilde{z},\tilde{n},i,\tilde{z}',\tilde{n}',i')\right] d\tilde{F}_{\tilde{n}}(\tilde{z}',\tilde{n}',i')\widetilde{\Omega}^{\tau}_{\tilde{n}}(\tilde{z},\tilde{n},i)$$

$$- \left(\overline{\gamma}_i e^{\delta\tilde{n}} + A\tilde{v}(\tilde{z},\tilde{n},i)\overline{\alpha}Z/e^{\tilde{z}} \int \left[\tilde{\mathbb{1}}_P(\tilde{z},\tilde{n},i,\tilde{z}',\tilde{n}',i')e^{\beta(\tilde{z}'-\tilde{z})}\right] d\tilde{F}_{\tilde{n}}(\tilde{z}',\tilde{n}',i') - g\right)\widetilde{\Omega}^{\tau}_{\tilde{z}}(\tilde{z},\tilde{n},i)$$

$$- \lambda_i\left[\tilde{\Omega}^{\tau}(\tilde{z},\tilde{n},-i) - \tilde{\Omega}^{\tau}(\tilde{z},\tilde{n},i)\right]$$

$$= e^{\tilde{z}} + \eta\tilde{\Omega}^{\tau}(0,\tilde{n},i) - \frac{\overline{c}}{\phi+1}\tilde{v}(\tilde{z},\tilde{n},i)^{\phi+1}Z$$

$$- A\tilde{v}(\tilde{z},\tilde{n},i)\int \left[\tilde{\mathbb{1}}_P(\tilde{z},\tilde{n},i,\tilde{z}',\tilde{n}',i')\widetilde{\Omega}^{\tau-1}_{\tilde{n}}(\tilde{z}',\tilde{n}',i')/e^{\tilde{n}'}\right] d\tilde{F}_{\tilde{n}}(\tilde{z}',\tilde{n}',i')$$

$$+\frac{1}{\Delta}\tilde{\Omega}^{\tau-1}(\tilde{z},\tilde{n},i)$$

We now discretize $\tilde{n}$ on an evenly spaced $N_{\tilde{n}} \times 1$ grid and $\tilde{z}$ on an evenly spaced $N_{\tilde{z}} \times 1$. Stack these according to:

$$\begin{pmatrix} \widetilde{z}_1, \tilde{n}_1, h \\ \widetilde{z}_2, \tilde{n}_1, h \\ \vdots \\ \widetilde{z}_{N_{\tilde{z}}}, \tilde{n}_1, h \\ \vdots \\ \widetilde{z}_1, \tilde{n}_{N_{\tilde{n}}}, h \\ \vdots \\ \widetilde{z}_{N_{\tilde{z}}}, \tilde{n}_{N_{\tilde{n}}}, h \\ \widetilde{z}_1, \tilde{n}_1, l \\ \vdots \\ \widetilde{z}_{N_{\tilde{z}}}, \tilde{n}_{N_{\tilde{n}}}, l \end{pmatrix}$$

The above equation can be rewritten in vector form as:

$$\left(\frac{1}{\Delta}+\rho+\eta\right)\tilde{\Omega}^{\tau}-\mu_n\widetilde{\Omega}^{\tau}_{\tilde{n}}-\mu_z\widetilde{\Omega}^{\tau}_{\tilde{z}}-\Lambda\tilde{\Omega}^{\tau}=\pi+\frac{1}{\Delta}\tilde{\Omega}^{\tau-1} \tag{4.21}$$

where

- the element of $N_{\tilde{z}} \times N_{\tilde{n}} \times 2$ vector $\tilde{\Omega}^{\tau}$ consists of $\tilde{\Omega}^{\tau}(\tilde{z},\tilde{n},i)$,
- the element of $N_{\tilde{z}} \times N_{\tilde{n}} \times 2$ vector $\tilde{\Omega}^{\tau}_{\tilde{n}}$ consists of $\tilde{\Omega}^{\tau}_{\tilde{n}}(\tilde{z},\tilde{n},i)$,
- the element of $N_{\tilde{z}} \times N_{\tilde{n}} \times 2$ vector $\tilde{\Omega}^{\tau}_{\tilde{z}}$ consists of $\tilde{\Omega}^{\tau}_{\tilde{z}}(\tilde{z},\tilde{n},i)$,
- the element of $N_{\tilde{z}} \times N_{\tilde{n}} \times 2$ vector μ_n consists of

$$A\tilde{v}(\tilde{z},\tilde{n},i)/e^{\tilde{n}}\int\left[\tilde{\mathbb{1}}_P(\tilde{z},\tilde{n},i,\tilde{z}',\tilde{n}',i')\right]d\tilde{F}_{\tilde{n}}(\tilde{z}',\tilde{n}',i'),$$

- the element of $N_{\tilde{z}} \times N_{\tilde{n}} \times 2$ vector μ_z consists of

$$\overline{\gamma}_i e^{\delta \tilde{n}} + A\tilde{v}(\tilde{z}, \tilde{n}, i)\overline{\alpha} Z / e^{\tilde{z}} \int \left[\tilde{\mathbb{1}}_P(\tilde{z}, \tilde{n}, i, \tilde{z}', \tilde{n}', i') e^{\beta(\tilde{z}' - \tilde{z})} \right] d\tilde{F}_{\tilde{n}}(\tilde{z}', \tilde{n}', i') - g,$$

  and

- the element of $N_{\tilde{z}} \times N_{\tilde{n}} \times 2$ vector π consists of

$$\begin{aligned} & e^{\tilde{z}} + \eta \tilde{\Omega}^{\tau}(0, \tilde{n}, i) - \frac{\overline{c}}{\phi + 1} \tilde{v}(\tilde{z}, \tilde{n}, i)^{\phi+1} Z \\ & - A\tilde{v}(\tilde{z}, \tilde{n}, i) \int \left[\tilde{\mathbb{1}}_P(\tilde{z}, \tilde{n}, i, \tilde{z}', \tilde{n}', i') \widetilde{\Omega}_{\tilde{n}}^{\tau-1}(\tilde{z}', \tilde{n}', i') / e^{\tilde{n}'} \right] d\tilde{F}_{\tilde{n}}(\tilde{z}', \tilde{n}', i') \end{aligned}$$

$(N_{\tilde{z}} \times N_{\tilde{n}} \times 2) \times (N_{\tilde{z}} \times N_{\tilde{n}} \times 2)$ matrix Λ is

$$\Lambda = \begin{pmatrix} -\lambda_h I_{N_{\tilde{z}} \times N_{\tilde{n}}} & \lambda_h I_{N_{\tilde{z}} \times N_{\tilde{n}}} \\ \lambda_l I_{N_{\tilde{z}} \times N_{\tilde{n}}} & -\lambda_l I_{N_{\tilde{z}} \times N_{\tilde{n}}} \end{pmatrix}$$

where $I_{N_{\tilde{z}} \times N_{\tilde{n}}}$ is $N_{\tilde{z}} \times N_{\tilde{n}}$ identity matrix. Let $D_{\tilde{n}}$ be the $(N_{\tilde{z}} \times N_{\tilde{n}} \times 2) \times (N_{\tilde{z}} \times N_{\tilde{n}} \times 2)$ matrix that, when pre-multiplying $\tilde{\Omega}^{\tau}$, gives an approximation of $\tilde{\Omega}^{\tau}_{\tilde{n}}$. Analogously, define $D_{\tilde{n}}$:

$$\begin{aligned} \widetilde{\Omega}^{\tau}_{\tilde{n}} &= D_{\tilde{n}} \widetilde{\Omega}^{\tau} \\ \widetilde{\Omega}^{\tau}_{\tilde{z}} &= D_{\tilde{z}} \widetilde{\Omega}^{\tau} \end{aligned}$$

To compute the derivative matrices $D_{\tilde{n}}$ and $D_{\tilde{z}}$, we follow an upwind scheme. That is, we use a forward approximation when the drift of the state variable is positive and a backward approximation when the drift of the state is negative. Using these, we can write (4.21) as

$$\left[\left(\frac{1}{\Delta} + \rho + \eta \right) - \mu_n D_{\tilde{n}} - \mu_z D_{\tilde{z}} - \Lambda \right] \tilde{\Omega}^{\tau} = \pi + \frac{1}{\Delta} \tilde{\Omega}^{\tau-1}.$$

The implicit method works by updating $\widetilde{\Omega}^{\tau}$ through the above equation.

4.5.2 Numerical Solution to Kolmogorov Forward Equation

The total mass of the inventor can be expressed as

$$\mathtt{v} = \int \tilde{v}(\tilde{z}, \tilde{n}, i) d\tilde{F}(\tilde{z}, \tilde{n}, i)$$

Let define

$$\tilde{f}_{\tilde{v}}(\tilde{z}, \tilde{n}, i) = \frac{\tilde{v}(\tilde{z}, \tilde{n}, i)\tilde{f}(\tilde{z}, \tilde{n}, i)}{\mathtt{v}}$$

and $\tilde{F}_{\tilde{v}}(\tilde{z}, \tilde{n}, i)$ the corresponding cumulative distribution. We construct the Kolmogorov forward equation in terms of the transformed variables.

$$\begin{aligned} 0 = &- \frac{\partial}{\partial \tilde{n}} \left(\tilde{\mu}_{\tilde{n}}(\tilde{z}, \tilde{n}, i)\tilde{f}(\tilde{z}, \tilde{n}, i) \right) - \frac{\partial}{\partial \tilde{z}} \left(\tilde{\mu}_{\tilde{z}}(\tilde{z}, \tilde{n}, i)\tilde{f}(\tilde{z}, \tilde{n}, i) \right) \\ &- \eta \tilde{f}(\tilde{z}, \tilde{n}, i) + \eta \int_0^1 \tilde{f}(\tilde{z}', \tilde{n}, i) dz' \boldsymbol{\delta}(0) \\ &- \lambda_i \tilde{f}(\tilde{z}, \tilde{n}, i) + \lambda_{-i} \tilde{f}(\tilde{z}, \tilde{n}, -i) \end{aligned}$$

where [3]

$$\begin{aligned} \tilde{\mu}_{\tilde{n}}(\tilde{z}, \tilde{n}, i) = &A \frac{\tilde{v}(\tilde{z}, \tilde{n}, i)}{e^{\tilde{n}}} \int \left[\tilde{\mathbb{1}}_P(\tilde{z}, \tilde{n}, i, \tilde{z}', \tilde{n}', i') \right] d\tilde{F}_{\tilde{n}}(\tilde{z}', \tilde{n}', i') \\ &- A \frac{\mathtt{v}}{\mathtt{n}} \int \tilde{\mathbb{1}}_P(\tilde{z}', \tilde{n}', i', \tilde{z}, \tilde{n}, i) d\tilde{F}_{\tilde{v}}(\tilde{z}', \tilde{n}', i') \\ \tilde{\mu}_{\tilde{z}}(\tilde{z}, \tilde{n}, i) = &\overline{\gamma}_i e^{\delta \tilde{n}} + A \tilde{v}(\tilde{z}, \tilde{n}, i) \overline{\alpha} Z / e^{\tilde{z}} \int \left[\tilde{\mathbb{1}}_P(\tilde{z}, \tilde{n}, i, \tilde{z}', \tilde{n}', i') e^{\beta(\tilde{z}' - \tilde{z})} \right] d\tilde{F}_{\tilde{n}}(\tilde{z}', \tilde{n}', i') - g \end{aligned}$$

We can vectorize this in the same way as above, and obtain

$$0 = -D_{\tilde{n}} \tilde{\mu}_{\tilde{n}} \tilde{f} - D_{\tilde{z}} \tilde{\mu}_{\tilde{z}} \tilde{f} - \eta \tilde{f} + \eta \tilde{f}_0 + \Lambda' \tilde{f}$$

where

- the element of $N_{\tilde{z}} \times N_{\tilde{n}} \times 2$ vector $\tilde{f}$ is the stacked as the value function,
- the element of $N_{\tilde{z}} \times N_{\tilde{n}} \times 2$ vector $\tilde{\mu}_{\tilde{n}}$ consists of $\tilde{\mu}_{\tilde{n}}(\tilde{z}, \tilde{n}, i)$,

[3] Note that $\frac{dn/n}{dt} = \frac{d \log n}{dt} = \frac{d\tilde{n}}{dt}$ and $\frac{dz/z}{dt} = \frac{d \log z}{dt} = \frac{d\tilde{z}}{dt}$.

- the element of $N_{\tilde{z}} \times N_{\tilde{n}} \times 2$ vector $\tilde{\mu}_{\tilde{z}}$ consists of $\tilde{\mu}_{\tilde{z}}(\tilde{z}, \tilde{n}, i)$, and

- the element of $N_{\tilde{z}} \times N_{\tilde{n}} \times 2$ vector $\tilde{f}_0$ consists of $\int_0^1 \tilde{f}(\tilde{z}', \tilde{n}, i) dz' \boldsymbol{\delta}(0)$[4].

To construct the derivative matrices, we use a backward approximation when the drift of the state variable is positive, and a forward approximation when the drift of the state is negative. This expression can be rearranged to yield

$$\left(-D_{\tilde{n}}\tilde{\mu}_{\tilde{n}} - D_{\tilde{z}}\tilde{\mu}_{\tilde{z}} - \eta + \Lambda'\right)\tilde{f} = -\eta\tilde{f}_0$$

and the distribution of individuals is updated according to the above equation.

4.5.3 Solving the Transition Path

We illustrate how to solve the transition path. We solve the transition path in terms of transformed variables and later recover the non-transformed values over the transition. Finally, we explain how to calculate the consumption-equivalent welfare gain from the policy change.

Perfect Foresight Equilibrium

First, we define the perfect foresight equilibrium, which differs from the balanced growth path equilibrium in that it has a time notation and, the HJB equation and KFE have time derivative terms. Note that the perfect foresight equilibrium is detrended by the productivity of the technology frontier $\overline{z}(t)$ at each period.

Definition 2. (Perfect Foresight Equilibrium) A *perfect foresight equilibrium* consists of: (i) a joint value function $\Omega(z, n, t)$; (ii) a vacancy policy $v(z, n, t)$; (iii) a stationary distribution of firms $f(z, n, t)$; (iv) vacancy- and employment-weighted distributions $f_v(z, n, t)$ and $f_n(z, n, t)$; (v) poaching indicator function $\mathbb{1}_P(z, n, z', n', t)$; (vi) the aggregate productivity

[4] Due to the Dirac delta function $\boldsymbol{\delta}(0)$, the elements of $\tilde{f}_0$ can take positive value only if $\tilde{z} = 0$.

$Z(t)$, the aggregate consumption $C(t)$, and the total vacancies $\mathtt{v}(t)$, and (vii) the economic growth rate $g(t)$ such that

1. The joint value $\Omega(z,n,t)$ satisfies the HJB equation

$$\begin{aligned}
&\rho\Omega(z,n,t) + \frac{\partial}{\partial t}\Omega(z,n,t) = \\
&\quad z - c(v(z,n,t))Z(t) \\
&\quad + Av(z,n,t)\int\left[\Omega_n(z,n,t) + \alpha(z'/z)Z(t)\Omega_z(z,n,t) - \Omega_n(z',n',t)\right]^+ dF_n(z',n',t) \\
&\quad + (\gamma(n) - g(t))\, z\Omega_z(z,n,t) \\
&\quad + \eta\left[\Omega(1,n,t) - \Omega(z,n,t)\right]
\end{aligned}$$

2. The vacancy policy $v(z,n,t)$ satisfies the first order condition

$$c_v(v(z,n,t))Z = A\int\left[\Omega_n(z,n,t) + \alpha(z'/z)Z\Omega_z(z,n,t) - \Omega_n(z',n',t)\right]^+ dF_n(z',n',t)$$

3. A density function $f(z,n,t)$ satisfies the KFE equation

$$\begin{aligned}
\frac{\partial}{\partial t}f(z,n,t) = &-\frac{\partial}{\partial n}\left(\mu_n(z,n,t)f(z,n,t)\right) - \frac{\partial}{\partial z}\left(\mu_z(z,n,t)f(z,n,t)\right) - \eta f(z,n,t) \\
&+ \eta\int_0^1 f(z,n,t)dz\boldsymbol{\delta}(1)
\end{aligned}$$

where

$$\begin{aligned}
\mu_z(z,n,t) \equiv& \left(\gamma(n) - g\right)z + Av(z,n,t)Z\int \mathbb{1}_P(z,n,z',n',t)\alpha(z'/z)dF_n(z',n',t) \\
\mu_n(z,n,t) \equiv& Av(z,n,t)\int \mathbb{1}_P(z,n,z',n',t)dF_n(z',n',t) \\
&- A\mathtt{v}\frac{n}{\mathtt{n}}\int \mathbb{1}_P(z',n',z,n,t)dF_v(z',n',t)
\end{aligned}$$

4. Vacancy- and employment-weighted distributions are consistent:

$$\begin{aligned}
f_v(z,n,t) &= \frac{v(z,n,t)f(z,n,t)}{\mathtt{v}(t)} \\
f_n(z,n,t) &= \frac{nf(z,n,t)}{\mathtt{n}(t)}
\end{aligned}$$

5. Poaching indicator function $\mathbb{1}_P(z, n, z', n', t)$ is

$$\mathbb{1}_P(z, n, z', n', t) = \begin{cases} 1 & \text{if} \quad \Omega_n(z, n, t) + \alpha(z'/z)Z\Omega_z(z, n, t) > \Omega_n(z', n', t) \\ 0 & \text{otherwise} \end{cases}$$

6. The aggregate productivity $Z(t)$, aggregate consumption $C(t)$, and the total vacancies $\mathtt{v}(t)$ satisfy rate satisfy

$$\begin{aligned} Z(t) &= \int z dF(z, n, t) \\ C(t) &= \left\{ 1 - \int c(v(z, n, t)) dF(z, n, t) \right\} Z(t) \\ \mathtt{v}(t) &= \int v(z, n, t) dF(z, n, t) \end{aligned}$$

7. The inventor market clearing condition is satisfied:

$$\mathtt{n}(t) = \int n dF(z, n, t)$$

Now, we recover the non-transformed values over the transition from the transformed values. Without loss of generality, let normalize the productivity of the technology frontier at time $t = 0$ to 1: $\overline{z}(0) = 1$. Then,

$$\begin{aligned} \overline{z}(t) &= \exp\left(\int_0^t g(\tau) d\tau \right) \\ \hat{Z}(t) &= \overline{z}(t) Z(t) \\ \hat{C}(t) &= \overline{z}(t) C(t) \end{aligned}$$

Solution Algorithm

To solve the transition path, we guess a path for optimal behavior of firms over a discretized grid for productivity, number of inventors and time. Subsequently, we iterate on

1. Given a path for the distribution of firms and inventors, update optimal behavior of firms and inventors backwards in time;

2. Given a path for behavior, update the evolution of the distribution of firms and inventor forward in time

3. If the updated path for the distributions and behavior are close enough to the original path, stop. Otherwise return to 1.

Recover Non-Transformed Values

Now, we recover the non-transformed values over the transition from the transformed values. Without loss of generality, let normalize the productivity of the technology frontier at time $t = 0$ to 1: $\overline{z}(0) = 1$. Then,

$$\begin{aligned} \overline{z}(t) &= \exp\left(\int_0^t g(\tau) d\tau\right) \\ \hat{Z}(t) &= \overline{z}(t) Z(t) \\ \hat{C}(t) &= \overline{z}(t) C(t) \end{aligned}$$

Consumption-Equivalent Welfare Gains

Definition 3. (Consumption-Equivalent Welfare Gains from Policy Change) Consider an economy without policy change and the associated consumption path $\left\{\hat{C}(t)\right\}_{t\geq 0}$. *The consumption equivalent welfare gains from policy change* is the scalar $\mathcal{L}$ such that the consumer is indifferent between the consumption path $\left\{\mathcal{L} \times \hat{C}(t)\right\}_{t\geq 0}$ and the consumption path generated by the policy change.

Let $V\left(\left\{\hat{C}(t)\right\}_{t\geq 0}\right)$ define the welfare:

$$V\left(\left\{\hat{C}(t)\right\}_{t\geq 0}\right) \equiv \int_0^\infty e^{-\rho t} \log \hat{C}(t) dt.$$

Then,

$$V\left(\mathcal{L} \times \left\{\hat{C}(t)\right\}_{t\geq 0}\right) = \int_0^\infty e^{-\rho t} \log\left(\mathcal{L} \times \hat{C}(t)\right) dt$$

$$
\begin{aligned}
&= \int_0^\infty e^{-\rho t} \log \mathcal{L} dt + \int_0^\infty e^{-\rho t} \log \hat{C}(t) dt \\
&= \frac{\log \mathcal{L}}{\rho} + \int_0^\infty e^{-\rho t} \log \hat{C}(t) dt \\
&= \frac{\log \mathcal{L}}{\rho} + V\left(\left\{\hat{C}(t)\right\}_{t\geq 0}\right)
\end{aligned}
$$

This implies

$$
\frac{\log \mathcal{L}}{\rho} = V\left(\mathcal{L} \times \left\{\hat{C}(t)\right\}_{t\geq 0}\right) - V\left(\left\{\hat{C}(t)\right\}_{t\geq 0}\right)
$$

or equivalently,

$$
\mathcal{L} = \exp\left[\rho\left\{V\left(\mathcal{L} \times \left\{\hat{C}(t)\right\}_{t\geq 0}\right) - V\left(\left\{\hat{C}(t)\right\}_{t\geq 0}\right)\right\}\right]
$$

Let $\left\{\hat{C}'(t)\right\}_{t\geq 0}$ denote the consumption path generated by policy change. Then, the consumption equivalent welfare gains from policy change are calculated as

$$
\mathcal{L} = \exp\left[\rho\left\{V\left(\left\{\hat{C}'(t)\right\}_{t\geq 0}\right) - V\left(\left\{\hat{C}(t)\right\}_{t\geq 0}\right)\right\}\right]
$$

4.6 Chapter 3 Data Appendix

4.6.1 Non-Cash Benefits

The BSWS compiles total cash earnings, and thus does not cover non-cash benefits. To provide an overview of non-cash benefits in the Japanese labor market, Table 4.7 shows the average share using the General Survey on Working Conditions (GSWC). The survey has been conducted periodically since 2001 by the MHLW to supplement the BSWS, and covers a broader range of compensation types and other working conditions. The table indicates that the average share of non-cash benefits is stable over time, at a little below 20% of total compensation.[5] Legal welfare expenses—i.e., employers' contributions to social insurance—

[5] The share is lower than in U.S. data, mainly due to the lower public health insurance premium in Japan. Gu et al. (2020) report that for the U.S. economy, the average share of benefit expenditures is 27% in 1982 (25% if excluding non-production bonus, premium pay, and shift differential, which are likely paid in cash) and 32% in 2018 (29%) in a sample of employer cost surveys conducted by the Bureau of Labor Statistic (BLS). They find that health insurance premiums account for 9.5% in 2018, with a rising trend.

show a slight upward trend, but a decline in non-legal welfare expenses and retirement expenses offset it.

Table 4.7: Average Share of Non-Cash Benefits

	2001	2005	2010	2015
Total compensation	**100.0**	**100.0**	**100.0**	**100.0**
Cash earnings	**81.7**	**81.0**	**81.5**	**80.9**
Scheduled hourly earnings	59.6	60.2	62.1	60.3
Non-scheduled hourly earnings	5.4	5.7	5.6	5.8
Bonus and other earnings	16.7	15.0	13.8	14.8
Non-cash benefits	**18.3**	**19.0**	**18.5**	**19.1**
Legal welfare expenses	9.3	10.0	10.8	11.4
Health and long-term care insurance	3.0	3.4	3.6	4.0
Welfare pension insurance	5.1	5.2	5.8	6.2
Labor insurance	1.2	1.4	1.3	1.0
Other legal welfare expenses	0.1	0.1	0.1	0.2
Non-legal welfare expenses	2.3	2.1	2.0	1.6
House expenses	1.1	1.0	1.0	0.8
Food expenses	0.2	0.2	0.2	0.1
Other welfare expenses	0.9	0.8	0.8	0.7
Retirement expenses	5.8	6.0	5.0	4.5
In-kind benefits	0.3	0.2	0.1	0.1
Training	0.3	0.3	0.1	0.2
Others	0.4	0.4	0.4	1.2

Notes: Figures are average shares for all full-time workers. Legal welfare expenses are employers' contributions to each social security item. Non-legal items are non-cash benefits, such as the expenses associated with company housing and canteens. The GSWC is available since 2001.

Source: BSWS, GSWC.

It is worth noting that the Japanese legal framework leaves little room for discretionary changes in legal welfare expenses, which account for more than half of total non-cash benefits. Social insurance, including health and long-term care insurance, welfare pension insurance, and labor (employment and accident) insurance, are mandatory for all full-time workers. Contributions are determined by law, based on employees' earnings and family conditions, and must be split evenly between employers and employees. Consequently, these components are considered to be loosely proportional to cash earnings.

4.6.2 Sample of New Graduates

Our central approach is to employ the sample of new graduates. While the approach is aimed at eliminating the cyclical upgrading of job-match quality through job changes, one implication is that our empirical results hinge on the specificities of new graduates, if any, in terms of wage cyclicality. In this section, we discuss the representativeness of new graduates' wages in the Japanese labor market and potentially desirable features for our analysis.

First, new graduates compose a sizable part of the marginal labor market in Japan; they account for 61.5% of total workers not hired from previous jobs.[6] Moreover, 40.3% of regular workers aged 20 to 54 do not change jobs after graduation, and the average length of tenure is 12.4 years in our BSWS sample. These reflect salient features of the Japanese labor market: the simultaneous recruiting of new graduates and lifetime employment. In other words, new graduates' wages comprise a central part of firms' wage-setting behavior and their total labor cost.

Second, the recruitment process for new graduates is standardized to a considerable extent. For college graduates, the recruitment schedule, including the start dates for job advertisements and the selection process, is guided by the Japan Business Federation, an economic organization formed by major corporations. Consequently, recruitment for most firms proceeds simultaneously with job matches sequentially formed. For high school grad-

[6]The ratio is the average for regular workers from 1998 to 2020, the available data period in the Employment Trend Survey conducted by the MHLW.

uates, high schools and public employment offices play crucial roles in matching firms and students (see Genda et al. 2010 for more details), which contributes to clearing the market. These institutional factors may facilitate wage determination and reflect relevant labor supply and demand.

Third, a uniform base wage tends to be offered to all new graduates within a firm, which leaves little room for discretionary wage setting according to job-match quality. Base wages for new graduates are often published on corporate websites. While it is possible to differentiate new graduates' earnings by using other compensation components, such as bonuses, the fraction of these variable components is, on average, lower for workers with shorter tenure.

4.6.3 Overall UCL

We calculate the overall UCL in the following steps, which are analogous to those taken by Kudlyak (2014).

1. We run the following regression:
   $\ln w^c_{t,t+\tau} = const. + \alpha^c + \theta Z^c_{t,t+\tau} + \sum_{d_0=1}^{T} \sum_{d=d_0}^{T} \chi_{d_0,d} D^c_{d_0,d} + \varepsilon^c_{t,t+\tau},$
   where $w^c_{t,t+\tau}$ is the real wage of worker hired in year t with tenure τ in category c, α^c is the category fixed effect, and $\varepsilon^c_{t,t+\tau}$ is the error term such that $\varepsilon^c_{t,t+\tau} \sim N(0, \sigma^c_\varepsilon)$. $Z^c_{t,t+\tau}$ is a vector of controls, in which we include tenure and tenure squared. $D^c_{d_0,d}$ is a dummy variable that takes the value 1 if $d_0 = t$ and $d = t + \tau$ and 0 otherwise.

2. Using the estimates of the coefficients, we obtain the projected wages:
   $\widehat{\ln w_{t,t+\tau}} = \widehat{const.} + \hat{\theta} Z_{t,t+\tau} + \hat{\chi}_{t,t+\tau}$.

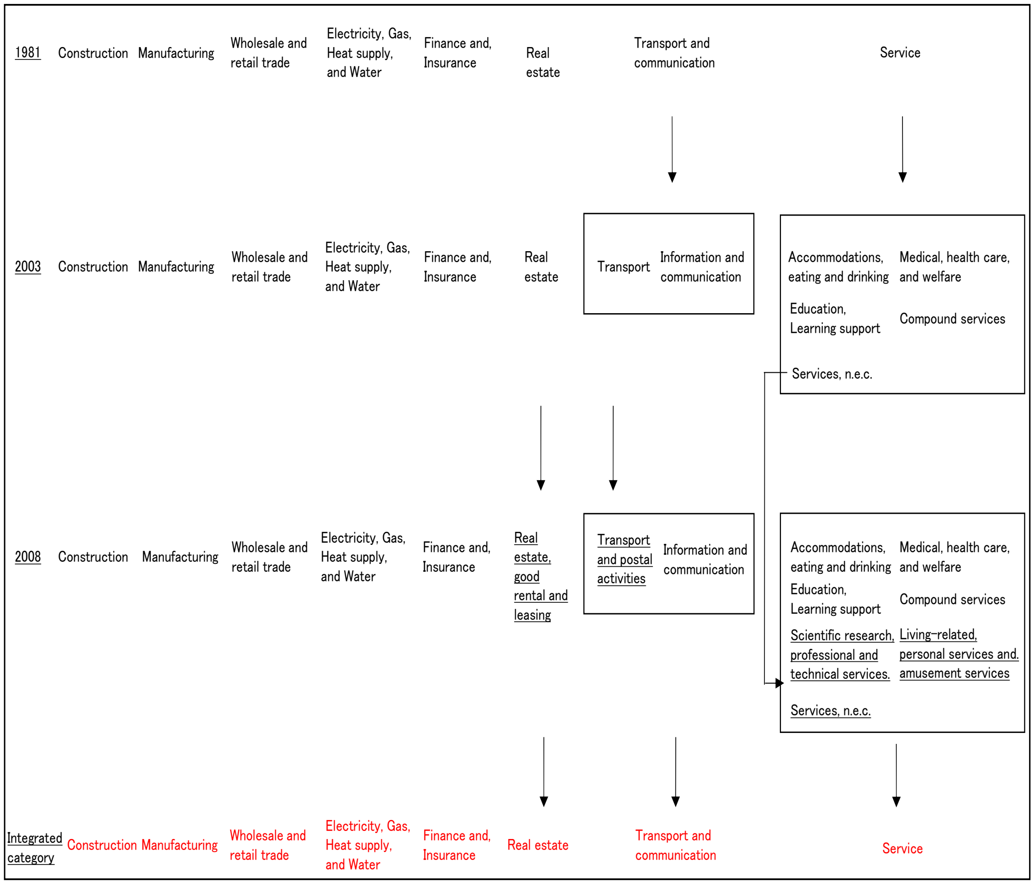
3. Using the projected wages, we compute the overall UCL as follows:
   $UCL_t = \widehat{w_{t,t}} + \sum_{\tau=1}^{T-1} \beta(1-\bar{s})^\tau (\widehat{w_{t,t+\tau}} - \widehat{w_{t+1,t+\tau}})$,
   where $\widehat{w} = \exp\left(\widehat{\ln w}\right)$ and $\bar{s}$ is the average separation rate across categories.

4.6.4 Industry-Level Data

Industry category. Industry categories in the BSWS changed during the period covered by this study, in line with revisions of the Japan Standard Industrial Category (JSIC). For our analysis, we consolidate industry categories into those available throughout our sample period. Specifically, we use the following eight categories: (1) Construction, (2) Manufacturing, (3) Wholesale and Retail trade, (4) Finance and insurance, (5) Real estate, (6) Electricity, gas, heat supply, and water, (7) Transport and communications, and (8) Services. The consolidation at each time of the JSIC revision is shown in Figure 4.1.

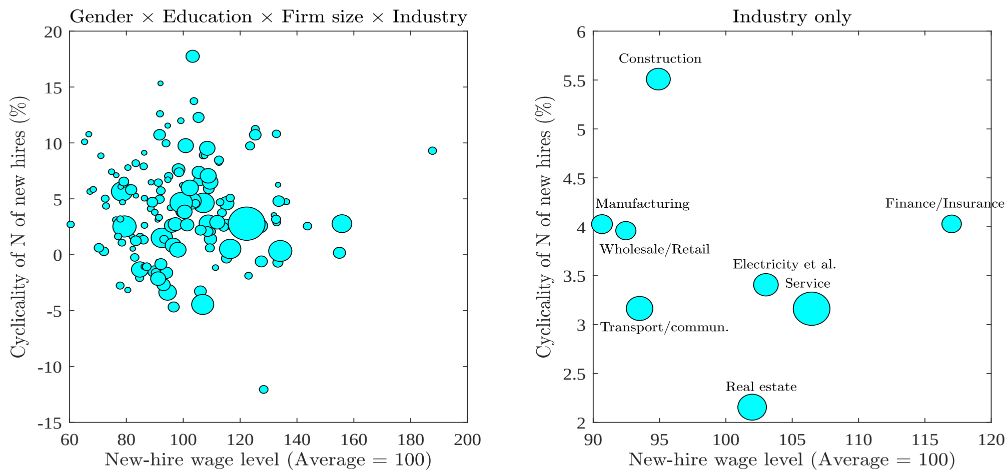
Industry composition bias. If higher-paid industries adjust the number of new hires more cyclically over business cycles, the average wage of new hires becomes procyclical at the aggregate level (e.g., McLaughlin and Bils 2001). As we discuss in the main text, we chose not to use the industry breakdown in the baseline analysis because it leads to small samples in some categories. Nevertheless, here we explore the potential direction and magnitude of the composition bias by focusing on the sample of all workers, which offers relatively larger sample sizes than that of new graduates. Figure 4.2 plots the relationship between new hires' wage levels (horizontal axis) and the cyclicality of the number of new hires (vertical axis) in each category. A positive association of these two measures gives rise to the composition bias described above. Interestingly, the correlation is close to zero or negative (–0.02 in the left panel and –0.21 in the right panel). For example, in the right panel, in which industry aggregates are displayed, industries with a high wage level but moderate cyclicality (finance and insurance) and with a low wage level and high cyclicality (construction) mitigate the correlation. It is also notable that the industry wage differences are relatively small compared with U.S. data. A simple calculation following McLaughlin and Bils (2001) suggests that composition bias due to the omission of an industry breakdown on the average wage is –0.01%, or essentially zero.

Figure 4.1: Transition of Industry Categories



4.7 Chapter 3 Empirical Appendix

Figure 4.2: Industry Composition Bias



Notes: The left panel plots 4,032 categories for each gender, education, firm size, and industry characteristic. The right panel aggregates characteristics other than industry, and thus the number of plots corresponds to that of industries (8).

4.8 Chapter 3 Theoretical Appendix

Posted wage contract. Firms post a wage contract to maximize its value while ensuring the worker's value at the Pareto frontier. It is convenient to start from a type-L firm's problem, which can be represented by dynamic programming:

$$f_L^\sigma(z, E_L^\sigma, U_L(z)) = \max_{w_L, \{E_L^\sigma(z')\}_{z' \in Z}} z - w_L + \beta \mathbb{E}_z \left[(1 - s_L) f_L^\sigma(z', E_L^\sigma(z'), U_L(z')) \right], \tag{4.22}$$

$$\text{s.t.} \quad E_L^\sigma = u(w_L) + \beta \mathbb{E}_z \left[(1 - s_L) E_L^\sigma(z') + s_L U_L(z') \right]. \tag{4.23}$$

The notation for worker type is dropped, since workers' values do not depend on worker type in type-L firms. The first-order conditions (FOCs), along with the envelope condition, lead to $1/u'(w) = 1/u'(w(z'))$ for all $z' \in Z$, which implies a fixed-wage contract. This is a well-known result, whereby a risk-neutral firm offers perfect insurance against aggregate fluctuations to risk-averse workers.

A type-H firm's problem is boiled down to dynamic programming with additional con-

Table 4.8: Heterogeneity in Cyclicality

	(1) UCL	(2) UCL	(3) UCL	(4) Average wage for non-regular workers	(5) New hire wage for non-regular workers
x_t: Unemp. rate	0.957***	1.249***	0.834***		
	(0.225)	(0.309)	(0.232)		
$x_t \times 1_{\text{high school}}$	-0.220				
	(0.278)				
$x_t \times 1_{\text{male}}$		-0.733**			
		(0.343)			
$x_t \times 1_{\text{small firm}}$			0.189		
			(0.371)		
$x_t \times 1_{x_t \geqq 0}$				0.393***	0.466***
				(0.107)	(0.112)
$x_t \times 1_{x_t < 0}$				0.316**	0.301***
				(0.128)	(0.059)
R-squared	0.04	0.05	0.04	0.21	0.33
N of categories	18	18	18	18	18
N of observations	540	540	540	144	144

Notes: The same notes as for Table 3.3 apply. The UCL and new hire wage are those for new graduates, whereas the average wage is that for all workers. In column (1), $1_{\text{high school}}$ takes a value of one if workers are high school graduates and zero otherwise. In column (2), 1_{male} takes a value of one if workers are male and zero otherwise. In column (3), $1_{\text{small firm}}$ takes a value of one if workers are employed by small firms and zero otherwise. In columns (4) and (5), the independent variable (HP-filtered unemployment rate) is split into positive and negative values.

straints:

$$f_H^\sigma(z, E_{SH}^\sigma, U_{SH}(z)) = \max_{w_H, \{E_{SH}^\sigma(z')\}_{z' \in Z}} \bar{a}z - w_H + \beta \mathbb{E}_z \left[(1 - s_H) f_H^\sigma(z', E_{SH}^\sigma(z'), U_{SH}(z')) \right], \tag{4.24}$$

$$\text{s.t.} \quad E_{SH}^\sigma = u(w_H) + \beta \mathbb{E}_z \left[(1 - s_H) E_{SH}^\sigma(z') + s_H U_{SH}(z')) \right], \tag{4.25}$$

$$E_{SH}^\sigma(z') \geq U_{SH}(z') \quad \text{for all } z' \in Z, \tag{4.26}$$

$$V_{NL}^\sigma(z) \geq V_{NH}^\sigma(z). \tag{4.27}$$

Equation (4.26) denotes the worker's participation constraint in the life of the contract due to the limited participation assumption. Equation (4.27) is type-N workers' exclusion constraint, which only appears in the initial period of the contract. $V_{NL}^\sigma(z)$ is obtained by solving a type-L firm's dynamic programming as above, whereas $V_{NH}^\sigma(z)$ is given by

$$V_{NH}^\sigma(z) = p\mu_H(\theta_H(z))(E_{NH}^\sigma(z) - U_{NH}(z)), \tag{4.28}$$

$$E_{NH}^\sigma(z) = u(w_H) + \beta \mathbb{E}_z \left[(1 - s_H) E_{NH}^\sigma(z') + s_H U_{NH}(z') \right], \tag{4.29}$$

$$U_{NH}(z) = u(b) + \beta \mathbb{E}_z \left[p\mu_H(\theta_H(z')) E_{NH}^\sigma(z') + (1 - p\mu_H(\theta_H(z'))) U_{NH}(z') \right]. \tag{4.30}$$

Notice that the job-finding probability is discounted by the screening probability p.

As shown by Thomas and Worrall (1998) and verified in a search and matching framework by Rudanko (2009), period wages take a max function under a worker's limited commitment: $w' = \max\{w, w^*(z')\}$, where $w^*(z')$ is the wage that lets the participation constraint (4.26) hold with equality, i.e., $E_{SH}^\sigma(z') = U_{SH}(z')$. This can be seen in the optimality condition $1/u'(w) = 1/u'(w(z')) - \eta(z')$, where $\beta(1 - s_H)\pi(z'|z)\eta(z')$ is the Lagrangian of (4.26). A fixed wage continues if (4.26) holds with inequality ($\eta(z') = 0$ due to the complementary slackness condition), while a firm raises the wage to ensure the worker's participation if the outside option would exceed the value of the current match under a fixed wage ($\eta(z') > 0$).

Definition of equilibrium. A directed search equilibrium is defined in line with Moen (1997) and Rudanko (2009). An equilibrium consists of a worker's employment value, $E_{ji}^\sigma(z_t)$;

a worker's unemployment value, $U_{ji}(z_t)$; a firm's value function; $f_i^{\sigma}(z_t)$; a market tightness, $\theta_i(z_t)$; and a wage contract, $\sigma_i(z_t)$, for $j = S, N$, $i = H, L$, and all $z_t \in Z$ such that

1. (Free entry) Firms enter a labor market and post vacancies with the associated contract $\sigma_i(z_t)$ until the value of posting a vacancy becomes zero—i.e., equation (20) in the main text is satisfied.

2. (Firm's optimization) Given U, the firm's value function $f_i^{\sigma}(z_t)$ solves the dynamic programming (4.22) and (4.24).

3. (Worker's unemployment value) The worker's unemployment value, $U_{ji}(z_t)$, evolves consistent with other elements of the equilibrium—i.e., equation (14) in the main text is satisfied.

4. (Pareto efficiency) There is no alternative pair $(\hat{\theta}(z_t), \hat{\sigma}(z_t))$ with which the net surpluses of workers and firms are at least as much as those with $(\theta(z_t), \sigma(z_t))$ and it is strictly more for one party—i.e.,

$$\mu_i(\hat{\theta}_i(z_t))(E_{ji}^{\hat{\sigma}}(z_t) - U_{ji}(z_t)) > \mu_i(\theta_i(z_t))(E_{ji}^{\sigma}(z_t) - U_{ji}(z_t)), \tag{4.31}$$

and

$$-k_i + q_i(\hat{\theta}_i(z_t)) f_i^{\hat{\sigma}}(z_t) > 0. \tag{4.32}$$

Acemoglu, D., "Training and Innovation in an Imperfect Labour Market," *Economic Review of Studies* 64 (1997), 445–464.

Acemoglu, D. and J.-S. Pischke, "Why Do Firms Train? Theory and Evidence," *Quarterly Journal of Economics* 113 (1998), 79–119.

Achdou, Y., J. Han, J.-M. Lasry, P.-L. Lions and B. Moll, "Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach," *Review of Economic Studies* 89 (2021), 45–89.

Akcigit, U. and S. T. Ates, "Ten Facts on Declining Business Dynamism and Lessons from Endogenous Growth Theory," *American Economic Journal: Macroeconomics* 13 (2021), 257–298.

Akcigit, U., S. Caicedo, E. Miguelez, S. Stantcheva and V. Sterzi, "Dancing with the Stars: Innovation Through Interactions," National Bureau of Economic Research, 2018.

Akcigit, U. and N. Goldschlag, "Measuring the Characteristics and Employment Dynamics of U.S. Inventors," *SSRN Electronic Journal* (2023a).

———, "Where Have All the "Creative Talents" Gone? Employment Dynamics of US Inventors," NBER Working Paper No. 31085, 2023b.

Almeida, P. and B. Kogut, "Localization of Knowledge and the Mobility of Engineers in Regional Networks," *Management Science* 45 (1999), 905–917.

Altonji, J. G. and R. A. Shakotko, "Do Wages Rise with Job Seniority?," *Review of Economic Studies* 54 (1987), 437–459.

Arrow, K. J., “Economic Welfare and the Allocation of Resources for Invention,” in C. K. Rowley, ed., *Readings in Industrial Economics: Volume Two: Private Enterprise and State Intervention* (London: Macmillan Education UK, 1962), 219–236.

Atkenson, A., A. Khan and L. Ohanian, “Are Data on Industry Evolution and Gross Job Turnover Relevant for Macroeconomics?,” *Carnegie-Rochester Conference Series on Public Policy* 44 (1996), 215–239.

Autor, D., W. Kerr and A. Kugler, “Does Employment Protection Reduce Productivity? Evidence from US States,” *Economic Journal* 117 (2007), 189–217.

Backus, D. and P. Kehoe, “International Evidence on the Historical Properties of Business Cycles,” *American Economic Review* 82 (1992), 864–888.

Barattieri, A., S. Basu and P. Gottschalk, “Some Evidence on the Importance of Sticky Wages,” *American Economic Journal: Macroeconomics* 6 (2014), 70–101.

Barlevy, G., “The Sullying Effect of Recessions,” *Review of Economic Studies* 69 (2002), 65–96.

Barro, R. J., “Long-term Contracting, Sticky Prices, and Monetary Policy,” *Journal of Monetary Economics* 3 (1977), 305–316.

Basu, S. and C. L. House, “Allocative and Remitted Wages: New Facts and Challenges for Keynesian Models,” *Handbook of Macroeconomics* 2 (2016), 297–354.

Bauernschuster, S., “Dismissal Protection and Small Firms’ Hirings: Evidence from a Policy Reform,” *Small Business Economics* 40 (2013), 293–307.

Beaudry, P. and J. DiNardo, “The Effect of Implicit Contracts on the Movement of Wages Over the Business Cycle: Evidence from Micro Data,” *Journal of Political Economy* 99 (1991), 665–688.

Becker, G. S., "Investment in Human Capital: A Theoretical Analysis," *Journal of Political Economy* 70 (1962), 9–41.

Bellou, A. and B. Kaymak, "The Cyclical Behavior of Job Quality and Real Wage Growth," *American Economic Review: Insights* 3 (2021), 83–96.

Ben-Porath, Y., "The Production of Human Capital and the Life Cycle of Earnings," *Journal of Political Economy* 75 (1967), 352–365.

Benhabib, J., J. Perla and C. Tonetti, "Reconciling Models of Diffusion and Innovation: A Theory of the Productivity Distribution and Technology Frontier," *Econometrica* 89 (2021), 2261–2301.

Bewley, T. F., *Why Wages Don't Fall During a Recession* (Harvard University Press, 1999).

Bilal, A., N. Engbom, S. Mongey and G. Violante, "Labor Market Dynamics When Ideas are Harder to Find," *The Economics of Creative Destruction* (2023).

Bilal, A., N. Engbom, S. Mongey and G. L. Violante, "Firm and Worker Dynamics in a Frictional Labor Market," *Econometrica* 90 (2022), 1425–1462.

Bils, M., "Real Wages Over the Business Cycle: Evidence from Panel Data," *Journal of Political Economy* 93 (1985), 666–689.

Bils, M., P. J. Klenow and B. A. Malin, "Resurrecting the Role of the Product Market Wedge in Recessions," *American Economic Review* 108 (2018), 1118–1146.

Bils, M., M. Kudlyak and P. Lins, "The Quality-Adjusted Cyclical Price of Labor," Federal Reserve Bank of San Francisco Working Paper 2023-10, 2023.

Black, I. and A. Figueiredo, "The False Illusion of Wage Cyclicality," mimeo, 2022.

Boedo, H. M. and T. Mukoyama, "Evaluating the Effects of Entry Regulations and Firing Costs on International Income Differences," *Journal of Economic Growth* 17 (2012), 143–170.

Braunerhjelm, P., D. Ding and P. Thulin, "Labour Market Mobility, Knowledge Diffusion and Innovation," *European Economic Review* 123 (2020), 103386.

Breschi, S. and F. Lissoni, "Mobility of skilled workers and co-invention networks: an anatomy of localized knowledge flows," *Journal of Economic Geography* 9 (2009), 439–468.

Buera, F. J. and R. E. Lucas, "Idea Flows and Economic Growth," *Annual Review of Economics* 10 (2018), 315–345.

Buera, F. J. and E. Oberfield, "The Global Diffusion of Ideas," *Econometrica* 88 (2020), 83–114.

Burdett, K. and D. Mortensen, "Wage Differentials, Employer Size, and Unemployment," *International Economic Review* 39 (1998), 257–273.

Caballero, R. J. and M. L. Hammour, "The Cleansing Effect of Recessions," *American Economic Review* 84 (1994), 1350–68.

Cahuc, P., O. Charlot and F. Malherbet, "Explaining the Spread of Temporary Jobs and Its Impact on Labor Turnover," *International Economic Review* 57 (2016), 533–571.

Cahuc, P., F. Postel-Vinay and J.-M. Robin, "Wage Bargaining with On-the-Job Search: Theory and Evidence," *Econometrica* 74 (2006), 326–364.

Card, D., J. Heining and P. Kline, "Workplace Heterogeneity and the Rise of West German Wage Inequality," *Quarterly Journal of Economics* 128 (2013), 967–1015.

Carneiro, A., P. Guimarães and P. Portugal, "Real Wages and the Business Cycle: Accounting for Worker, Firm, and Job Title Heterogeneity," *American Economic Journal: Macroeconomics* 4 (2012), 133–152.

Christiano, L. J., M. Eichenbaum and C. L. Evans, “Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy,” *Journal of Political Economy* 113 (2005), 1–45.

Da-Rocha, J.-M., D. Restuccia and M. M. Tavares, “Firing Costs, Misallocation, and Aggregate Productivity,” *Journal of Economic Dynamics and Control* 98 (2019), 60–81.

Daruich, D., S. D. Addario and R. Saggio, “The Effects of Partial Employment Protection Reforms: Evidence from Italy,” *Review of Economic Studies* 90 (2023), 2880–2942.

Dauth, W. and J. Eppelsheimer, “Preparing the Sample of Integrated Labour Market Biographies (SIAB) for Scientific Analysis: A Guide,” *Journal for Labour Market Research* 54 (2020).

Doepke, M. and R. Gaetani, “Why Didn’t the College Premium Rise Everywhere? Employment Protection and On-the-Job Investment in Skills,” *American Economic Journal: Macroeconomics* forthcoming (2024).

Doniger, C. L., “The Ways the Cookie Crumbles: Education and the Margins of Cyclical Adjustment in the Labor Market,” Finance and Economics Discussion Series 2021-019. Washington: Board of Governors of the Federal Reserve System, 2021.

Dorner, M., D. Harhoff, F. Gaessler, K. Hoisl and F. Poege, “Linked Inventor Biography Data 1980-2014,” FDZ-Datenreport 03/2018 en, 2018.

Dupraz, S., E. Nakamura and J. Steinsson, “A Plucking Model of Business Cycles,” mimeo, 2021.

Eggertsson, A., N. R. Mehrotra and J. A. Robbins, “A Model of Secure Stagnation:

Theory and Quantitative Evaluation," *American Economic Journal Macroeconomics* 11 (2019), 1–48.

Elsby, M. W., B. Hobijn and A. Sahin, "Unemployment Dynamics in the OECD," *Review of Economics and Statistics* 95 (2013), 530–548.

Elsby, M. W. L. and A. Gottfries, "Firm Dynamics, On-the-Job Search, and Labor Market Fluctuations," *Review of Economic Studies* 89 (2021), 1–50.

Engbom, N., "Contagious Unemployment," mimeo, 2021.

———, "Labor Market and Human Capital Accumulation," NBER Working Paper No. 29698, 2022.

Esteban-Pretel, J., R. Nakajima and R. Tanaka, "Japan's Labor Market Cyclicality and the Volatility Puzzle," RIETI Discussion Paper Series 11-E-040, 2011.

Fahr, S. and F. Smets, "Downward Wage Rigidities and Optimal Monetary Policy in a Monetary Union," *Scandinavian Journal of Economics* 112 (2010), 812–840.

Farber, H. S., "Mobility and Stability: The Dynamics of Job Change in Labor Markets," in O. C. Ashenfelter and D. Card, eds., *Handbook of Labor Economics*volume 3B (North Holland, 1999), 2439–2483.

Figueiredo, A., "Wage Cyclicality and Labor Market Sorting," *AER: Insights* 4 (2022), 425–442.

Fujita, S. and G. Ramey, "The Cyclicality of Separation and Job Finding Rates," *International Economic Review* 50 (2009), 415–430.

Fukui, M., "A Theory of Wage Rigidity and Unemployment Fluctuations with On-the-Job Search," mimeo, 2020.

Fukui, M., S. Kikuchi and L. Goalist Co., “Job Creation during the COVID-19 Pandemic in Japan,” CREPEDP-73, 2020.

Genda, Y., A. Kondo and S. Ohta, “Long-Term Effects of a Recession at Labor Market Entry in Japan and the United States,” *Journal of Human Resources* 45 (2010), 157–196.

Gertler, M., C. Huckfeldt and A. Trigari, “Unemployment Fluctuations, Match Quality, and the Wage Cyclicality of New Hires,” *Review of Economic Studies* 87 (2020), 1876–1914.

Gertler, M. and A. Trigari, “Unemployment Fluctuations with Staggered Nash Wage Bargaining,” *Journal of Political Economy* 117 (2009), 38–86.

Gregory, V., G. Menzio and D. G. Wiczer, “The Alpha Beta Gamma of the Labor Market,” NBER Working Paper 28663, 2021.

Grigsby, J., E. Hurst and A. Yildirmaz, “Aggregate Nominal Wage Adjustments: New Evidence from Administrative Payroll Data,” *American Economic Review* 111 (2021), 428–471.

Gu, G. W., E. Prasad and T. Moehrle, “New Evidence on Cyclical Variation in Average Labor Costs in the United Staes.,” *Review of Economics and Statistics* 102 (2020), 966–979.

Guerrieri, V., R. Shimer and R. Wright, “Adverse Selection in Competitive Search Equilibrium,” *Econometrica* 78 (2010), 1823–1862.

Hagedorn, M. and I. Manovskii, “Job Selection and Wages over the Business Cycle,” *American Economic Review* 103 (2013), 771–803.

Hall, B. H., A. B. Jaffe and M. Trajtenberg, “The NBER patent citation data file: Lessons, insights and methodological tools.,” NBER Working Paper No. 8498, 2001.

Hall, R. E., "Employment Fluctuations with Equilibrium Wage Stickiness," *American Economic Review* 95 (2005), 50–65.

Hall, R. E. and M. Kudlyak, "Job-Finding and Job-Losing: A Comprehensive Model of Heterogeneous Individual Labor-Market Dynamics," NBER Working Paper 25625, 2022.

Haltiwanger, J. C., H. R. Hyatt and L. B. Kahn, "Cyclical Job Ladders by Firm Size and Firm Wage," *American Economic Journal: Macroeconomics* 10 (2018), 52–85.

Hamilton, J. D., "Why You Should Never Use the Hodrick-Prescott Filter," *Review of Economics and Statistics* 100 (2018), 831–843.

Harris, M. and B. Holmstrom, "A Theory of Wage Dynamics," *Review of Economic Studies* 49 (1982), 315–333.

Hashimoto, M. and J. Raisian, "Employment Tenure and Earnings Profiles in Japan and the United States," *American Economic Review* 75 (1985), 721–735.

Hazell, J. and B. Taska, "Downward Rigidity in the Wage for New Hires," mimeo, 2020.

Herkenhoff, K., J. Lise, G. Menzio and G. Phillips, "Production and Learning in Teams," Technical Report, National Bureau of Economic Research, Cambridge, MA, 2018.

Hoisl, K., "Tracing Mobile Inventors—The Causality between Inventor Mobility and Inventor Productivity," *Research Policy* 36 (2007), 619–636.

Holden, S., "Wage Bargaining and Nominal Rigidities," *European Economic Review* 38 (1994), 1021–1039.

Hopenhayn, H. and R. Rogerson, "Job Turnover and Policy Evaluation: A General Equilibrium Analysis," *Journal of Political Economy* 101 (1993), 915–938.

HOPENHAYN, H. A., "Entry, Exit, and firm Dynamics in Long Run Equilibrium," *Econometrica* 60 (1992), 1127–1150.

———, "Firms, Misallocation, and Aggregate Productivity: A Review," *Annual Review of Economics* 6 (2014), 735–770.

JAFFE, A. B., M. TRAJTENBERG AND R. HENDERSON, "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations," *The Quarterly Journal of Economics* 108 (1993), 577–598.

JAHN, E. AND C. SCHNABEL, "Bestandsschutz durch Abfindungen: Höhere Rechtssicherheit und Eff," *Wirtschaftsdie* 83 (2001), 219–223.

JOLIVET, G., F. POSTEL-VINAY AND J.-M. ROBIN, "The Empirical Content of the Job Search Model: Labor Mobility and Wage Distributions in Europe and the US," *European Economic Review* 50 (2006), 877–907.

KAHN, L. B., "The Long-Term Labor Market Consequences of Graduating from College in a Bad Economy," *Labour Economics* 17 (2010), 303–316.

KAISER, U., H. C. KONGSTED AND T. RØNDE, "Does the mobility of R&D labor increase innovation?," *Journal of Economic Behavior & Organization* 110 (2015), 91–105.

KAMBAYASHI, R., *The World of Regular and Irregular Workers–Basic Issues in Contemporary Japanese Labor Economics [Translated from Japanese.]* (Keio University Press, 2017), japanese original title: Seiki no Secaki, Hiseiki no Sekai—Gendai Nihon Roudou Keizaigaku no Kihon Mondai.

KRAUSEA, M. U. AND H. UHLIG, "Transitions in the German Labor Market: Structure and Crisis," *Journal of Monetary Economics* 59 (2012), 64–79.

KUDLYAK, M., "The Cyclicality of the User Cost of Labor," *Journal of Monetary Economics* 68 (2014), 53–67.

Kuroda, S. and I. Yamamoto, "Why Are Nominal Wages Downwardly Rigid, but Less So in Japan? An Explanation Based on Behavioral Economics and Labor Market/Macroeconomic Differences," *Monetary and Economic Studies* 25 (2007), 45–88.

Lazear, E. P., "Firm-Specific Human Capital: A Skill-Weights Approach," *Journal of Political Economy* 117 (2009), 914–940.

Lentz, R. and N. Roys, "Training and Search On the Job," *Review of Economic Dynamics* forthcoming (2024).

Loewenstein, M. A. and J. R. Spletzer, "General and Specific Training: Evidence and Implications," *Journal of Human Resources* 34 (1999), 710–733.

Lucas, R. E., "Ideas and Growth," *Economica* 76 (2009), 1–19.

Lucas, R. E. and B. Moll, "Knowledge Growth and the Allocation of Time," *The Journal of Political Economy* 122 (2014), 1–51.

Luttmer, E. G. J., "Selection, Growth, and the Size Distribution of Firms," *The Quarterly Journal of Economics* 122 (2007), 1103–1144.

Martins, P. S., G. Solon and J. P. Thomas, "Measuring What Employers Do about Entry Wages over the Business Cycle: A New Approach," *American Economic Journal: Macroeconomics* 4 (2012), 36–55.

Mawdsley, J. K. and D. Somaya, "Employee Mobility and Organizational Outcomes: An Integrative Conceptual Framework and Research Agenda," *Journal of Management* 42 (2016), 85–113.

McLaughlin, K. J. and M. Bils, "Interindustry Mobility and the Cyclical Upgrading of Labor," *Journal of Labor Economics* 19 (2001), 94–135.

Menzio, G. and E. R. Moen, "Worker Replacement," *Journal of Monetary Economics* 57 (2010), 623–636.

Mineyama, T., W. Hirata and K. Nishizaki, “Optimal Inflation Rates in a Non-linear New Keynesian Model: The Case of Japan and the U.S,” *International Journal of Central Banking* 18 (2022), 1–46.

Moen, E. R., “Competitive Search Equilibrium,” *Journal of Political Economy* 105 (1997), 385–411.

Moen, E. R. and A. Rosén, “Incentives in Competitive Search Equilibrium,” *Review of Economic Studies* 78 (2011), 733–761.

Mortensen, D. T. and C. A. Pissarides, “Job Creation and Job Destruction in the Theory of Unemployment,” *Review of Economic Studies* 61 (1994), 881–894.

Moscarini, G. and F. Postel-Vinay, “The Cyclical Job Ladder,” *American Review of Economics* 10 (2018), 165–188.

Mukoyama, T., “On the Establishment Dynamics in the United States and Japan,” *Monetary and Economic Studies* 27 (2009), 53–74.

Mukoyama, T. and S. Osotimehin, “Barriers to Reallocation and Economic Growth: The Effects of Firing Costs,” *American Economic Journal: Macroeconomics* 11 (2019), 235–270.

Mukoyama, T., C. Patterson and A. Şahin, “Job Search Behavior over the Business Cycle,” *American Economic Journal: Macroeconomics* 10 (2018), 190–215.

Neftçi, S. N., “Are Economic Time Series Asymmetric over the Business Cycle?,” *Journal of Political Economy* 92 (1984), 307–328.

Nekarda, C. J. and V. A. Ramey, “The Cyclical Behavior of the Price-Cost Markup,” *Journal of Money, Credit and Banking* 2 (2020), 319–353.

Nickell, S. and R. Layard, “Labor Market Institutions and Economic Performance,” *Handbook of Labor Economics* 3 (1999), 3029–3084.

OECD, *Employment Outlook* (OECD, 2004).

———, "Overcoming the Global Crisis: The Need for a new Growth Model," Economic Survey of Japan 2009, 2009.

Oreopoulos, P., T. von Wachter and A. Heisz, "The Short- and Long-Term Career Effects of Graduating in a Recession," *American Economic Journal: Applied Economics* 4 (2012), 1–29.

Pakes, A. S., "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks," *Econometrica* 54 (1986), 755–784.

Perla, J. and C. Tonetti, "Equilibrium Imitation and Growth," *The Journal of Political Economy* 122 (2014), 52–76.

Petrongolo, B. and C. Pissarides, "Looking into the Black Box: A Survey of the Matching Function," *Journal of Economic Literature* 39 (2001), 390—-431.

Pham, H., *Continuous-time Stochastic Control and Optimization with Financial Applications* (Springer Science & Business Media, 2009).

Pischke, J.-S., "Continuous training in Germany," *Journal of Population Economics* 14 (2001), 523–548.

Pissarides, C. A., "The Unemployment Volatility Puzzle: Is Wage Stickiness the Answer?," *Econometrica* 77 (2009), 1339–1369.

Postel-Vinay, F. and J.-M. Robin, "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity," *Econometrica* 70 (2002), 2295–2350.

Postel–Vinay, F. and J. Robin, "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity," *Econometrica* 70 (2002), 9–41.

PRATO, M., “The Global Race for Talent: Brain Drain, Knowledge Transfer, and Growth,” , 2022.

RAHKO, J., “Knowledge spillovers through inventor mobility: the effect on firm-level patenting,” *Journal of Technology Transfer* 42 (2017), 585–614.

ROSENKOPF, L. AND P. ALMEIDA, “Overcoming Local Search through Alliances and Mobility,” *Management Science* 49 (2003), 751–766.

RUDANKO, L., “Labor Market Dynamics under Long-term Wage Contracting,” *Journal of Monetary Economics* 56 (2009), 170–183.

———, “Firm Wages in a Frictional Labor Market,” *American Economic Journal: Macroeconomics* 15 (2023), 517–550.

SCHAAL, E., “Uncertainty and unemployment,” *Econometrica* 85 (2017), 1675–1721.

SCHMIEDER, J., T. V. WATCHER AND J. HEINING, “The Costs of Job Displacement over the Business Cycle and Its Sources: Evidence from Germany,” *American Economic Review* 113 (2023), 1208–1254.

SHAPIRO, C. AND J. E. STIGLITZ, “Equilibrium Unemployment as a Worker Discipline Device,” *American Economic Review* 74 (1984), 434–444.

SHI, L., “Optimal Regulation of Noncompete Contracts,” *Econometrica* 91 (2023), 425–463.

SHI, L. AND H. A. HOPENHAYN, “Knowledge Creation and Diffusion with Limited Appropriation,” , 2020.

SHIMER, R., “The Cyclical Behavior of Equilibrium Unemployment and Vacancies,” *American Economic Review* 95 (2005), 25–49.

SINGH, J. AND A. AGRAWAL, “Recruiting for Ideas: How Firms Exploit the Prior Inventions of New Hires,” *Management Science* 57 (2011), 129–150.

Snell, A. and J. P. Thomas, “Labor Contracts, Equal Treatment, and Wage-Unemployment Dynamics,” *American Economic Journal: Macroeconomics* 2 (2010), 98–127.

Solon, G., R. Barsky and J. A. Parker, “Measuring the Cyclicality of Real Wages: How Important is Composition Bias?,” *Quarterly Journal of Economics* 109 (1994), 1–25.

Solon, G., W. Whatley and A. H. Stevens, “Intrafirm Job Mobility over the Business Cycle: Two Case Studies,” *Industrial and Labor Relations Review* 50 (1997), 402–415.

Solow, R. M., “Another Possible Source of Wage Stickiness,” *Journal of Macroeconomics* 1 (1979), 79–82.

Song, J., P. Almeida and G. Wu, “Learning–by–Hiring: When Is Mobility More Likely to Facilitate Interfirm Knowledge Transfer?,” *Management Science* 49 (2003), 351–365.

Stiglitz, J. E., “Prices and Queues as Screening Devices in Competitive Markets,” in P. Dasgupta, D. Gale, O. Hart and E. Maskin, eds., *Economic Analysis of Markets and Games: Essays in Honor of Frank Hahn* (MIT Press, 1976), 128–166.

Summers, L. H., “U.S. Economic Prospects: Secular Stagnation, Hysteresis, and the Zero Lower Bound,” *Business economics* 49 (2014), 65–73.

Thomas, J. and T. Worrall, “Self-Enforcing Wage Contracts,” *Review of Economic Studies* 55 (1998), 541–553.

Thomas, J. P. and T. Worrall, “Limited Commitment Models of the Labour Market,” *Scottish Journal of Political Economy* 54 (2007), 750–773.

Tobin, J., “Inflation and Unemployment,” *American Economic Review* 62 (1972), 1–18.

Ueda, K. and S. Claessens, “Basic Employment Protection, Bargaining Power, and Economic Outcomes,” *Journal of Law, Finance, and Accounting* 5 (2020), 179–229.

WASMER, E., “General versus Specific Skills in Labor Markets with Search Frictions and Firing Costs,” *American Economic Review* 96 (2006), 811–831.

YAMAMOTO, I., “Nominal Wage Flexibility after the Lost Decade in Japan,” *Mita Business Review* 50 (2007), 1–14, (in Japanese).

YELLEN, J. L., “Efficiency Wage Models of Unemployment,” *American Economic Review* 74 (1984), 200–2005, Papers and Proceedings of the Ninety-Sixth Annual Meeting of the American Economic Association (May 1984).