DISSERTATION / DOCTORAL THESIS

Titel der Dissertation / Title of the Doctoral Thesis

„Essays in Labor Economics"

verfasst von / submitted by

Dipl.-Ing. Martin Kerndler, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Doctor of Philosophy (PhD)

Wien, 2018 / Vienna 2018

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 794 370 140
Dissertationsgebiet lt. Studienblatt / field of study as it appears on the student record sheet:	Volkswirtschaftslehre
Betreut von / Supervisor:	Dipl.-Vw. Dr. Michael Reiter Univ.-Prof. Dipl.-Vw. Monika Gehrig-Merz, PhD

Acknowledgments

This dissertation benefited from the comments and expertise of many people at the Institute of Advanced Studies and at the University of Vienna. First and foremost, I am indebted to Michael Reiter for his guidance and support throughout the doctorate. His critical thinking helped a lot to sharpen my arguments and to improve the quality of my research. Special thanks goes to Tamas Papp, who introduced me to the field of directed search and provided great assistance at the early stage of this dissertation project. I appreciate the support of Monika Gehrig-Merz and Iacopo Morchio, who readily provided feedback on my research ideas. I am also grateful to the active participants of the Macro research seminar for their comments and suggestions. Concerning the empirical analysis, I could benefit from repeated discussions with Rudolf Winter-Ebmer.

Moreover, I want to thank the board of the Vienna Graduate School of Economics for the financial support, which also allowed me to present my research at numerous workshops and conferences. Funding for the research stays at the IAB data center in Nuremberg is gratefully acknowledged. I am also grateful to my fellow students at the Vienna Graduate School of Economics, especially to Milica Nikolic and Phillip Hergovich, for their continuing mental support and the lively exchange of ideas.

In October 2017, I joined the research group in economics at TU Wien, filling the position of a postdoctoral researcher. I am grateful to Alexia Fürnkranz-Prskawetz, the head of the research group, for considering me for this position and for allowing me to start the next phase of my career already before completing the doctorate. I appreciate the trust and understanding of my colleagues, who allowed me to devote a significant part of my working time to completing the essays of this dissertation.

Last but by no means least, I am highly indebted to my family for giving me the opportunity to focus fully on my research and for their invaluable support concerning all non-research related issues. Thank you!

Contents

1 Introduction

What determines whether your employer sticks with you for another month or shows you the door? According to the canonical search and matching model of the labor market pioneered by Peter Diamond, Dale Mortensen, and Christopher Pissarides, it is only two factors that matter in this regard: the minimum wage that you are willing to work for (your *reservation wage*) and the maximum wage that your employer is willing to pay for your services. Abstracting from redundancy costs, the latter equals the additional revenue that you generate for the firm (your *match productivity*). The prediction of the Diamond-Mortensen-Pissarides (DMP) model is that employment continues as long as the match productivity exceeds the reservation wage. Otherwise it is better to part ways.[1]

Interestingly, it is only the minimum acceptable wage and the maximum affordable wage that should affect the separation decision. The actual wage does not play any role. This is because a very flexible wage setting process operates in the background of the DMP model. Whenever the match productivity exceeds the reservation wage, there is a nonempty set of wages (the *feasible set*) for which continuing the employment relation is mutually beneficial. If after a negative productivity shock the current wage no longer lies in this set, a bilateral agreement is reached to reduce the wage to a level within the feasible set. This wage cut, which is in the interest of both parties, avoids the looming layoff. In practice, however, there can be many factors that limit the wage-setting flexibility of firms and workers, especially when it comes to wage reductions. These limiting factors stem from legal regulations but also from fundamental market failures. Both sources are elaborated in the following.

The legal framework in many countries imposes minimum wages. These are either mandated by the government and cover all legal working relations in a country. Alternatively, minimum wages are set on the industry or occupational level by collective bargaining agreements that are negotiated between trade unions and employer representatives. In the likely case that the minimum wage exceeds the reservation wage of some workers, they limit the scope for bilateral wage adjustments and lead to excess layoffs. This is particularly relevant for workers with low average productivity, such as low skilled and inexperienced workers, who earn close to the

[1]Classical references for the DMP model include Pissarides (1990), Mortensen and Pissarides (1994), and Mortensen and Pissarides (1999).

minimum wage. By contrast, legal constraints on wage-setting seem less relevant for workers whose wage level is well above the mandated wage floor.[2]

Wage-setting constraints that arise from market failures potentially affect *all* workers. While many different forms of market failures are possible, let me focus on frictions that relate to asymmetric information. Asymmetric information refers to a situation in which one party possesses more accurate information than another party. For instance, it is likely that match productivity, which captures the worker's marginal contribution to the firm's revenue, is easier to quantify for the firm than for the individual worker. This creates an informational advantage for the employer in wage renegotiations, because the worker is uncertain about the match productivity and therefore about the size of rents that are to be split. In this case, downwards wage adjustments might be impossible altogether as the following argument demonstrates. Suppose that the employer observes a decrease in match productivity such that keeping up the match is no longer profitable for the firm at the prevailing wage. The firm informs the worker that she will be laid off unless the wage is reduced sufficiently. As long as the maximum wage that the firm is able to pay exceeds the worker's reservation wage, it is beneficial for both parties to agree on a lower wage and avoid the layoff. The problem is, however, that the worker cannot verify whether the firm's claim is truthful or cheap talk. If the worker accepts a proposed wage cut with positive probability, a profit-maximizing firm has the incentive to always pretend that a wage cut is necessary to avoid a layoff, even if this is not actually the case. Therefore, the firm's claims do not contain any information, and rational workers should not react to them.[3]

But not only the employer possesses relevant private information. The effort that a worker provides on the job can be very hard to monitor for the employer. Instead of truthful working, the employee could engage in opportunistic behavior, such as not showing up for work, pursuing personal interests during working time or even stealing and hiding corporate property. To avoid opportunistic behavior, firms typically pay a premium above the worker's reservation wage. The higher the premium, the higher is the welfare loss of the worker in case she is caught and laid off for disciplinary reasons. The scope for downwards wage adjustment is therefore limited by the induced rise in opportunistic behavior. Additionally, wage cuts might increase workers' effort to search for alternative jobs. These considerations are more important in jobs where work effort is hard to monitor and where hiring replacement workers is expensive, such as in typical white-collar occupations.[4]

The previous two paragraphs illustrate how wage rigidities that stem from market failures

[2] For instance, Diéz-Catalán and Villanueva (2015) argue that collectively agreed minimum wages have contributed to the soar of unemployment in Spain during the Great Recession. However, the effect is limited to workers who earned less than 20% above the minimum wage before the Recession, which applied to 24% of all job stayers covered by a collective agreement.

[3] One may argue that also the employer cannot perfectly quantify the marginal contribution of an individual worker. In this case, the employer's perception of match productivity determines the highest feasible wage. As long the employer's perception is unknown to the worker, the same argumentation goes through.

[4] The arguments outlined in this paragraph are at the core of the efficiency wage literature including Stiglitz (1974), Akerlof (1982), Shapiro and Stiglitz (1984), Akerlof and Yellen (1990), and many others. If both informational frictions were included in one model, the productivity of match in a given period may be written as $z = ye$, where the employer observes z while the worker knows e.

such as asymmetric information can increase layoff rates above the socially efficient level. In this thesis, I investigate the effects of such frictions on employment, output, and welfare in the economy and discuss optimal policy responses. The first two essays use models of directed search that are calibrated to the data, while the third essay conducts a microeconometric analysis.

The first essay "Contracting frictions and inefficient layoffs over the life-cycle" assumes that wages cannot respond to stochastic fluctuations in match productivity. As motivated above, such a restriction can be due to asymmetric information about match productivity. Embedded into a life-cycle model of the labor market, I find that workers of different age are affected by this friction very differently. While layoff probabilities increase above the efficient level at all ages, the friction particularly depresses the employment rate in the oldest age group. The intuition behind this finding is that all workers react to the friction by contracting lower wages, which increases vacancy creation of the firms. For prime-age workers, the higher job creation almost offsets the higher job destruction in the calibrated model, such that the net employment effect is small. This is not the case for elderly workers. Due to their shorter distance to retirement, they experience a relatively larger increase in the layoff probability and a smaller increase in the job-finding probability. The contracting friction is also found to lower the effectiveness of policy reforms. While reducing generosity of early retirement arrangements boosts employment among the elderly, these gains are lower in presence of the friction. Restricting access to early retirement should therefore be complemented by labor market policies that improve firms' willingness to keep elderly workers employed.

One policy option that proofs effective in this regard is severance pay. The second essay "Optimal severance pay under different contractual regimes" discusses the optimal design of a mandated severance pay scheme in a very similar model framework. It investigates analytically and numerically how contractual flexibility and worker's risk attitudes shape the socially optimal design of severance pay. If workers are either risk neutral or search frictions on the workers' side of the labor market are negligible, the optimal level of severance pay does not depend on the severity of bilateral contracting frictions. Otherwise, severance pay should be highest if contracting frictions are most severe. I also derive the optimal tenure-profile of severance pay and find that risk aversion and realistic contracting frictions alone are not sufficient to understand why severance pay is typically increasing in tenure. If the model is extended by a private effort decision of the worker, however, both the evolution of wages and severance pay over a career are qualitatively in line with the data.

Finally, the third essay "Size and persistence matters: wage and employment insurance at the micro level" conducts an empirical analysis at the microeconomic level. Using rich linked employer-employee data from Germany, I investigate how idiosyncratic shocks to firm-level productivity affect individual wages and layoff probabilities. Using a novel estimation strategy based on Kalman smoothing, I document that both the persistence and the size of productivity shocks matter in this context. While wages respond largely symmetrically to positive and negative permanent productivity shocks, transitory shocks trigger asymmetric wage responses. Negative shocks tend to reduce wages, while positive shocks are fully captured by the firm. Firms also adjust to shocks by dismissing workers, but only in response to negative permanent shocks. Differentiating between white- and blue-collar employment reveals important heterogeneity. Real wage cuts and employment loss after negative productivity shocks are in fact limited to blue-collar workers. Whereas white-collar workers appear to be fully insured against negative shocks, both in terms of wages and in terms of employment. The wage effects could be due to considerations about employee motivation and turnover: The effort of blue-collar workers is typically easier to monitor, which allows more downward wage flexibility without spurring opportunistic behavior. Due to lower hiring and training costs, blue-collar workers are also less expensive to replace if they shirk or decide to quit after a wage cut. The same argument could contribute to the finding that layoffs are more concentrated on blue-collar workers. Additionally, blue-collar employment may be easier to substitute by other production factors such as capital.

2 Contracting frictions and inefficient layoffs over the life-cycle

2.1 Introduction

For its *Employment Outlook 2013* (OECD, 2013), the OECD analyzed the incidence of job displacement and its economic consequences for different groups of workers. A "job displacement" was defined as an "involuntary job separation due to economic or technological reasons or as a result of structural change" (p.194). The report concludes on pages 225–226 that

> *"[S]ome workers are more prone to job displacement, and to negative consequences after displacement, than others. In particular, older workers and those with low education levels have a higher displacement risk, take longer to get back into work and suffer greater (and more persistent) earnings losses in most countries examined."*

Labor market conditions for older workers were found to be particularly tough in continental Europe, where old-age displacement rates are high, re-employment rates are low, and a large share of old individuals becomes inactive within one year of displacement. Since early exits from the labor force increase the financial pressure on the social welfare system, various measures have been proposed and were already implemented by national governments in order to facilitate re-integration of unemployed older workers into the labor market.[1] This indicates that policy-makers perceive hiring of elderly unemployed as inefficient and try to intervene. However, it is also not clear whether the job separations that rendered these elderly workers unemployed had been efficient in the first place.

Deviations from the socially optimal separation rate might arise from inadequately designed social welfare systems, but also from imperfections of private employment arrangements. Standard models of labor economics typically assume that job separations are at least *bilaterally efficient*. Bilateral efficiency means that apart from exogenous reasons, an employment spell ends if and only if the joint surplus of the firm–worker match becomes negative. At this point, parting ways is optimal for both the firm and the worker. This property arises from bilaterally efficient wage determination mechanisms such as generalized Nash bargaining or directed search

[1] Table 5.2 in OECD (2006) provides an overview of the measures taken. Konle-Seidl (2017) summarizes the estimated effects of programs implemented in Austria, Germany, France, the Netherlands, and Norway.

(Mortensen and Pissarides, 1999). It remains valid when these models are put into a life-cycle context (Chéron et al., 2011, 2013).

For older workers, bilateral efficiency of separations seems hard to align with empirical evidence. First, bilateral efficiency implies that observed job separations should to a large extent be considered optimal by both parties. If they were not, the wage should have adjusted to ensure ongoing employment. Survey evidence instead suggests that many displaced old workers would have preferred to continue work but were denied to.[2] Unfortunately, it remains unclear from these surveys whether the respondents would have accepted a wage cut in order to remain employed. More convincing evidence against bilateral efficiency is presented by Frimmel et al. (2018). If separations were bilaterally efficient, the timing of a separation should only depend on the age-productivity profile of the firm–worker match and the worker's outside option, but not directly on the wage profile. In fact, the only role for wages should be the determination of the present discounted value for firms, which influences job creation (Hornstein et al., 2005). Frimmel et al. (2018) instead document a direct causal effect of wages on separations of older workers even after controlling for productivity and outside options. Using Austrian social security data, the authors analyze the age at which workers aged 57 to 65 exit their last job before retirement. They find a large variation in job exit ages between similar firms and show that part of these differences can be explained by differences in the age profile of wages. According to the authors' estimates, a one standard deviation increase in the steepness of the wage-age profile relative to the industry average leads to a 5.5 (6.9) months earlier job exit of blue (white) collar workers on average.[3]

The above evidence suggests that bilateral efficiency may fail because wages are not renegotiated. Since firms within the same industry are subject to the same labor market regulations, this is likely due to a market failure in the form of incomplete private employment contracts. To assess the consequences of such a market failure, this essay proposes and analyzes an age-structured labor market model with a contracting friction. Wages can only depend on the worker's age, but not on the productivity of the firm–worker match, which is subject to stochastic shocks. This restriction leads to situations in which paying the contracted wage is not profitable for the firm after the productivity shock is observed. The resulting layoff is *ex post* bilaterally inefficient if the productivity of the match would have exceeded the reservation productivity. I assess the micro- and macroeconomic effects of this contracting friction on different age groups and investigate the interaction between the friction and public policy.

First, I find that although the contracting friction increases the layoff probability at all ages,

[2]Dorn and Sousa-Poza (2010) report that a substantial amount of transitions to early retirement happens "not by choice" of the worker. The share is particularly high in continental Europe (Germany 50%, France 41%, Sweden 37.5%, Spain 32.5%) but also reaches 28.9% in the United Kingdom. Marmot et al. (2003) reports a similar share for the UK using a different data set. According to the 2012 wave of the European Labour Force Survey, 28% of the economically inactive persons in age 50–69 who received a pension at the day of the interview would have wished to stay longer in employment. The share exceeds 70% if job loss and/or unsuccessful job search was their main reason to retire (Eurostat, 2012, Graph 6.2).

[3]The estimations include worker and industry fixed effects as well as worker-specific incentives to retire. The steepness of the wage-age profile is instrumented by the lagged unemployment rate of prime-age workers 10 years before job exit to rule out reverse causality and worker self-selection.

it particularly depresses employment rates of the elderly. All workers react to the friction by contracting lower wages, which increases vacancy posting of the firms. For prime-age workers, the higher job creation almost offsets the higher job destruction in the calibrated model, such that the net employment effect is small. This is not the case for elderly workers. Due to their shorter distance to retirement, they experience a relatively larger increase in the layoff probability and a smaller increase in the job-finding probability. Second, I demonstrate that the positive macroeconomic effects of reducing generosity of early retirement are lower in presence of the contracting friction. The model suggests that reforms to the early retirement system should be accompanied by labor market policies that increase firms' willingness to keep elderly workers in employment. Otherwise the reform is likely to generate inefficiently high unemployment among the elderly—a common fear of politicians and labor unions.

The essay is structured as follows. Section 2.2 briefly summarizes the literature on inefficient layoffs and motivates the particular friction considered in this essay. Section 2.3 introduces the model. Section 2.4 derives the equilibrium and comparative static effects. The analytical results are complemented by a numerical assessment in Section 2.6, which illustrates the role of the friction when an early retirement reform is enacted and investigates complementary labor market reforms. Section 2.7 concludes. Section 2.A contains an overview of all defined functions, variables, and parameters. All proofs and additional lemmas are delegated to Section 2.B.

2.2 Sources of inefficient layoffs

Labor market outcomes arise from the interaction of workers' labor supply and firms' labor demand. Both margins may be distorted by governmental policies and/or market-inherent frictions, thereby resulting in an inefficient allocation of labor. The relation between public policy and the labor market exit of older workers has been intensively studied in the literature during the last decade. Fisher and Keuschnigg (2008), Jaag et al. (2010), and Hairault et al. (2015) argue that the social welfare system distorts individual behavior by introducing implicit taxes into the labor participation and retirement decision, unless the pension formula is actuarially fair at the optimal retirement age. Because wages are determined by generalized Nash bargaining in these papers, job separations are nevertheless *bilaterally* efficient.

This property might break down if the ability of private agents to renegotiate wages is restricted. Dustmann and Schönberg (2009) report that the wage floors that unionized firms face in Germany lead to fewer wage cuts and more layoffs of young workers. Guimarães et al. (2017) find lower hiring and higher separations rates in Portuguese firms to which collectively bargained wages are extended. Diéz-Catalán and Villanueva (2015) argue that the wage floors set by collective bargaining agreements increased the incidence of job loss during the Great Recession in Spain. But even without legal restrictions on wage setting, efficient wage renegotiation might fail due to market-inherent contracting frictions. Mechanisms that have been considered in this regard include asymmetric information about the size of the match surplus (Hashimoto, 1981; Hall and Lazear, 1984), adverse selection (Weiss, 1980), and moral hazard

(Lazear, 1979; Ramey and Watson, 1997). The presence of these market failures endogenously constrains the set of wage contracts that can be implemented in equilibrium. Further, contracting frictions and governmental policies may interact and re-enforce each other. Winter-Ebmer (2003) investigates the extension of unemployment insurance (UI) benefit duration for workers above age 50 introduced in 1988. The resulting increase in separation rates was significantly larger for workers with more than 10 years tenure than for workers with shorter tenure. Since high-tenured workers are likely to be more productive on average, the additional separations triggered by the UI reform were mainly driven by wage cost considerations of the employer rather than by match productivity and were therefore bilaterally inefficient.

In this essay, I embed a market-inherent contracting friction into a directed search model of the labor market with life-cycle dynamics in the manner of Menzio et al. (2016). Because search is directed, the agents internalize the search externalities they impose on other market participants (Shimer, 1996; Moen, 1997). Yet, neither private agents nor the government can overcome the search or the contracting friction. The contracting friction is modeled as in Alvarez and Veracierto (2001) and Boeri et al. (2017):

(i) the productivity of a firm-worker match is stochastic in each period,

(ii) wage contracts are written before productivity realizes and may not be contingent on productivity,

(iii) wage renegotiation is not possible.

The inability to renegotiate wages is the most restrictive assumption in this set and can be rationalized by asymmetric information. Suppose that the realized productivity draw is private knowledge of the firm. An employer can increase her own profit by making the worker agree on a wage cut. This creates an innate incentive to cheat on the worker and pretend that a wage cut is required to a prevent a layoff, even if this is not the case. A rational worker anticipates the employer's motives and opposes wage reductions. Alternative microfoundations for the absence of renegotiation may include employer's considerations about employee motivation, fairness, and the use of wage contracts as a screening device for new hires.

The contracting friction introduces above implies that for some productivity realizations the pre-negotiated wage level is *ex post* inappropriate to sustain the match, because one of the parties would suffer a loss and instead walks away. As the worker's outside option is deterministic in the model, it will be the firm that in some cases finds the contracted wage too high to keep up employment. The worker is then laid off, which is bilaterally inefficient if the match productivity would have exceeded the reservation productivity. When a bilaterally inefficient layoff occurs, *ex post* it would have been superior for both parties if they had contracted a lower wage *ex ante*, although the agents had correctly anticipated the probability of a layoff in the wage-setting process.

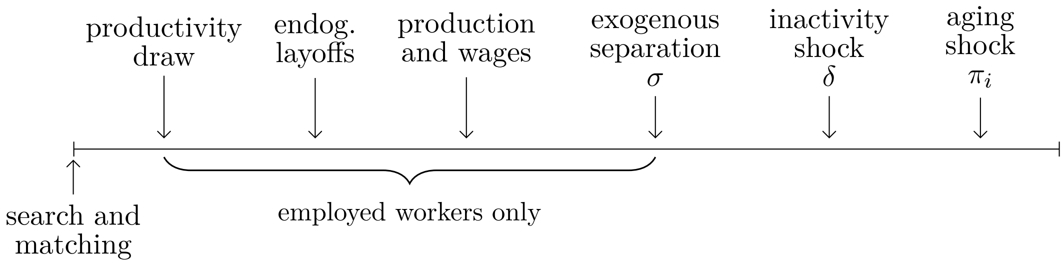


Figure 2.1. Timing within a period

2.3 Model setup

2.3.1 Individuals

Time is discrete with $t = 0, 1, 2, \ldots$. In each period, a unit mass of identical, risk averse individuals is born. Every individual lives through two stages of life: prime working age (m) and old working age (o). The aging process is stochastic. Each period, prime-age individuals proceed to old working age with probability $\pi_m > 0$, and individuals in old working age reach the normal retirement age with probability $\pi_o > 0$, at which they leave the model.[4] In any period, individuals can either be employed or unemployed. Unemployed individuals receive a period income b_m (b_o) while in the first (second) stage of their life. This income comprises the value of leisure or home production, z_i, and government transfers, g_i, such that $b_i = z_i + g_i$ for $i \in \{m, o\}$. Employed individuals who are in the first stage of their life are considered as *prime-age workers* (m). Employed individuals who are in the second stage of their life are either referred to as senior workers and as old workers. A *senior worker* (s) already started her current job during prime age. Whereas an *old worker* (o) started her current job when she was already in old working age. This distinction is necessary because the equilibrium wage will depend both on the worker's current age and the age at which she was hired.

The timing within a period is illustrated in Figure 2.1. At the beginning of a period, unemployed workers apply to vacancies that offer some wage contract ω_i. With probability $p(\theta_i)$ this application is successful, and a new firm–worker match is formed. Firm and worker then commit to the wage contract but not to actual employment. That is, either party can leave the match at any time.

The period output y_i that a matched worker can generate is stochastic and emerges from a distribution that may depend on the worker type $i \in \{m, s, o\}$. Productivity is drawn at the beginning of a match and renewed when the aging shock hits. In any other period, a new draw happens with probability $\phi \in [0, 1]$. The draws are independent across individuals, periods, and age groups. After the productivity of the current period is observed by the firm, it may terminate the match. Doing so is optimal if the firm surplus from the match turns out to be negative, that is, if the wage stream promised to the worker exceeds the sum of today's output

[4]I do not explicitly model youth and retirement beyond the normal retirement age. The model, however, takes into account early retirement.

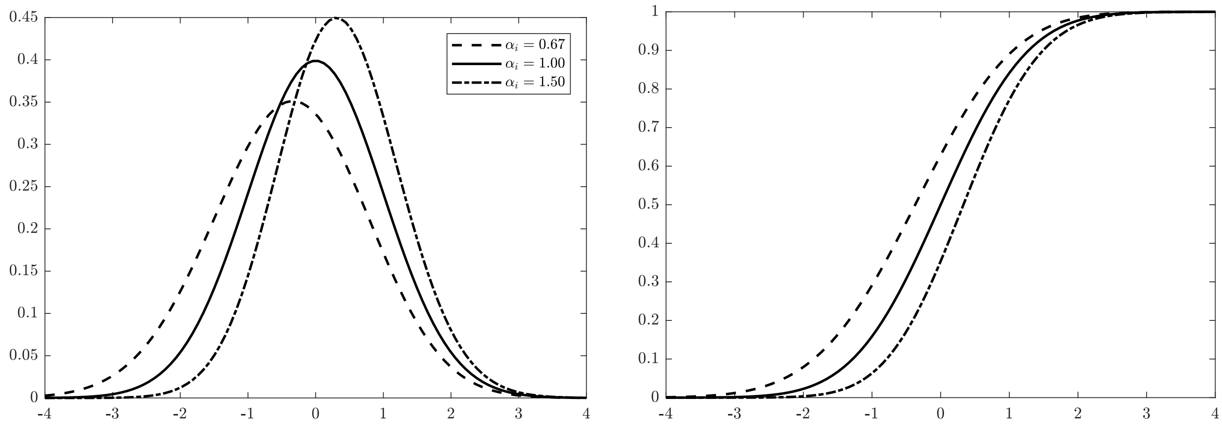


Figure 2.2. Density and distribution function of the normal distribution with $\mu_i = 0$, $s_i = 1$, and different levels of α_i.

and expected future output. If the match is profitable for the firm, production takes place and wages are paid according to the specified contract ω_i.

After the production stage, the match ends for exogenous reasons with probability $\sigma \geq 0$. Old individuals (regardless of their employment status) may additionally experience an inactivity shock with probability $\delta \geq 0$, after which they do not participate in the labor market any more. That is, they permanently stop all work and search activities. This could, for instance, capture a health shock that destroys the worker's production capacity, or a labor market exit for non-economic reasons. The aging shock hits at the very end of the period.

2.3.2 Productivity

The productivity of a match with a type i worker is a realization of the random variable Y_i for $i \in \{m, s, o\}$. These random variables satisfy some general properties.

Assumption 2.1. *Denote the distribution function of Y_i as F_i for $i \in \{m, s, o\}$. The distribution functions differ only in terms of a location parameter $\mu_i \in \mathbb{R}$, a scale parameter $s_i > 0$, and a shape parameter $\alpha_i > 0$. In particular, there exists a random variable Z with cdf F such that $F_i(y) = F(\frac{y-\mu_i}{s_i})^{\alpha_i}$ for $i \in \{m, s, o\}$ and the following properties hold:*

(i) the cdf F is twice continuously differentiable, the associated density f has support on the whole real line,

(ii) the random variable Z satisfies $0 \leq \mathbb{E}Z < \infty$,

(iii) the hazard rate $h := \frac{f}{1-F}$ is strictly increasing, while $\frac{h'}{h}$ is non-increasing,

(iv) the conditional expectation $\mathbb{E}[Z - a | Z \geq a]$ is convex in a.

According to the first part of the assumption, the three distribution functions are members of the same family of parametric distributions. For given shape parameter α_i, this is a location-scale family. The parameter μ_i governs the mean of the distribution, while s_i governs its dispersion. Prominent examples for such families are the normal distribution family and the logistic distribution family. To control the skewness of the distribution, I additionally introduce a shape parameter α_i. Figure 2.2 illustrates how the density function and cumulative distribution function are affected by changes in α_i, taking the standard normal distribution as reference ($F = \Phi$). For $\alpha_i = 1$, the distribution is symmetric around the mean. For $\alpha_i > 1$, the distribution becomes skewed to the right and the weight of the upper tail increases. For $\alpha_i < 1$, the weight of the lower tail increases.

Part (ii) of Assumption 2.1 is innocuous as the distribution family can always be reparameterized appropriately. The properties demanded in part (iii) and (iv) are satisfied by many frequently used distributions, including the normal and logistic family, see Section 2.B.1.

2.3.3 Firms, search, and matching

The economy is populated by a continuum of identical firms. Each firm consists of a single job and uses a linear production technology using only labor. Firms can freely enter the labor market, but posting a vacancy is involved with a period cost $c > 0$. The search and matching process follows the principles of competitive search (Shimer, 1996; Moen, 1997). Firms can age-direct their hiring process, such that prime-age and old-age job seekers search in different segments of the labor market. The labor market equilibrium is therefore independent of the age distribution in the economy.

In each labor market segment $i \in \{m, o\}$, firms post vacancies together with a wage contract ω_i, which yields a potentially infinite number of submarkets. Job seekers of type i costlessly observe these wage offers and apply to a submarket where an application yields the highest expected present discounted surplus for them. Within each submarket, JS_i applicants and V_i vacancies are randomly matched by a constant returns to scale matching technology $M(JS_i, V_i)$. As shown by Acemoglu and Shimer (1999), the labor market equilibrium can be characterized as the solution to a conceptually simple maximization problem (see below). Under standard assumptions, the equilibrium is unique and given by a pair (θ_i^*, ω_i^*). The variable θ_i is the labor market tightness, defined as the number of vacancies per applicant, $\theta_i = V_i/JS_i$. For future reference, the probability of filling a vacancy is defined as $q(\theta_i) = \frac{M(JS_i,V_i)}{V_i} = M(\frac{1}{\theta_i}, 1)$, and the probability that an application turns into a match is $p(\theta_i) = \frac{M(JS_i,V_i)}{JS_i} = \theta_i q(\theta_i)$.

The wage contracts ω_i posted by the firms are by assumption independent of productivity, but may depend on the worker's age. Therefore, prime-age job seekers look for wage contracts that specify a pair of wages $\omega_m = (w_m, w_s)$. The wage w_m applies as long as the worker is in prime working age, and the wage w_s applies thereafter. The contracts offered to old job seekers specify a single wage, $\omega_o = (w_o)$.

2.3.4 Government

The government plays a passive role in the model. The transfers g_i that non-employment individuals receive are financed by a lump sum tax τ levied on the whole population. In Section 2.6 I allow for additional government spending and/or revenue from labor market policies.

2.4 Equilibrium with the contracting friction

The model is solved assuming a demographic and economic steady state. The equilibrium consists of a set of wage contracts (ω_m^*, ω_o^*), labor market tightnesses (θ_m^*, θ_o^*), search values (V_m, V_o), and a lump sum tax τ^* that satisfy the following conditions:

(1) *labor market equilibrium of old job seekers*, i.e. taking τ^* and $(\theta_m^*, \omega_m^*, V_m)$ as given, the triple $(\theta_o^*, \omega_o^*, V_o)$ forms a directed search equilibrium:

- firms maximize profit under free entry, $q(\theta_o^*)\mathbb{E}J_o^+(\omega_o^*) = c$,
- job seekers apply optimally, $V_o = \max_{(\theta_o,\omega_o)} p(\theta_o)\mathbb{E}W_o^+(\omega_o) \geq p(\theta_o^*)\mathbb{E}W_o^+(\omega_o^*)$,

(2) *labor market equilibrium of prime-age job seekers*, i.e. taking τ^* and $(\theta_o^*, \omega_o^*, V_o)$ as given, the triple $(\theta_m^*, \omega_m^*, V_m)$ forms a directed search equilibrium:

- firms maximize profit under free entry, $q(\theta_m^*)\mathbb{E}J_m^+(\omega_m^*) = c$,
- job seekers apply optimally, $V_m = \max_{(\theta_m,\omega_m)} p(\theta_m)\mathbb{E}W_m^+(\omega_m) \geq p(\theta_m^*)\mathbb{E}W_m^+(\omega_m^*)$,

(3) *balanced budget*, i.e. taking $(\theta_o^*, \omega_o^*, V_o)$ and $(\theta_m^*, \omega_m^*, V_m)$ as given, τ^* balances the government budget.

Due to directed search, the labor market equilibrium on the labor market of old job seekers actually does not depend on $(\theta_m^*, \omega_m^*, V_m)$. The labor market equilibria can therefore be solved recursively. Section 2.4.1 considers the labor market equilibrium of old job seekers, before I turn to prime-age job seekers in Section 2.4.2. Section 2.4.3 defines aggregate economic measures and the equilibrium tax level. The analysis proceeds under the following functional restrictions:

Assumption 2.2. *Firms are risk neutral. Workers are risk averse with instantaneous utility function u defined on the interval (d, ∞) where $d \in \mathbb{R} \cup \{-\infty\}$ and $\lim_{x\to d} u(x) = -\infty$. It is three times differentiable with $u' > 0$, $u'' < 0$, $u''' \geq 0$, and $\lim_{x\to\infty} u'(x) = 0$. The matching function is Cobb-Douglas, which implies $q(\theta) = A\theta^{-\gamma}$ where $A > 0$ and $\gamma \in (0, 1)$.*

The assumptions on the utility function encompass, for example, the CARA and CRRA specifications. The specific form of the matching function makes the analysis of comparative static effects more tractable. The main results of this essay also hold for more general matching functions with varying matching elasticity $\varepsilon(\theta) = -\frac{q'(\theta)\theta}{q(\theta)}$. The main advantage of a constant elasticity $\varepsilon(\theta) = \gamma$ is that the optimal wage contract does not depend on the labor market tightness.

For the sake of tractability, the shape parameter of the distribution function is set to $\alpha_i = 1$ throughout this section.

Assumption 2.3. *Assume that $\alpha_i = 1$ for all $i \in \{m, s, o\}$.*

Under Assumption 2.3, the monotonicity properties of the hazard rate h demanded by Assumption 2.1 also apply to the hazard rates of the productivity distributions Y_i, that are given by $h_i := \frac{f_i}{1-F_i}$ for $i \in \{m, s, o\}$.

2.4.1 Labor market equilibrium of old job seekers

Following Acemoglu and Shimer (1999), the labor market equilibrium on the labor market of old job seekers is characterized as the solution to the constrained maximization problem

$$V_o := \max_{(\theta_o, w_o)} p(\theta_o)\mathbb{E}W_o^+(w_o) \quad \text{s.t.} \quad q(\theta_o)\mathbb{E}J_o^+(w_o) = c. \tag{2.1}$$

Intuitively, an old unemployed individual maximizes her expected surplus from applying to a vacancy with characteristics (θ_o, w_o), which is $p(\theta_o)\mathbb{E}W_o^+(w_o)$. With probability $p(\theta_o)$, the application is successful and generates an expected worker surplus of $\mathbb{E}W_o^+(w_o)$. Otherwise, the individual remains unemployed and her surplus over unemployment is zero by definition. Due to free entry, the value of vacant job is zero in equilibrium, such that the expected firm surplus of posting a vacancy just makes up for the posting cost c. This gives rise to the free entry condition $q(\theta_o)\mathbb{E}J_o^+(w_o) = c$, where $q(\theta_o)$ is the probability that the vacancy turns into a match, and $\mathbb{E}J_o^+(w_o)$ denotes the expected firm surplus of this match.

At the production stage, firm and worker surplus evolve over time according to

$$J_o(w_o; y) = y - w_o + \beta_o[\phi\mathbb{E}J_o^+(w_o) + (1-\phi)J_o(w_o; y)], \tag{2.2}$$

$$W_o(w_o) = u(w_o - \tau) - u(b_o - \tau) + \beta_o[\phi\mathbb{E}W_o^+(w_o) + (1-\phi)W_o(w_o) - V_o], \tag{2.3}$$

where $\beta_o := \beta(1-\pi_o)(1-\sigma)(1-\delta)$ is the *effective* time discount factor and $\beta \in [0, 1)$ is the pure time discount factor. Since the model is solved in a steady state, time indices are dropped altogether. The firm surplus $J_o(w_o; y)$ comprises the instantaneous profit $y - w_o$ and future profits discounted with the effective discount factor β_o. With probability ϕ a new productivity is drawn next period, which generates an expected surplus of $\mathbb{E}J_o^+(w_o)$. With probability $1-\phi$, the current draw prevails, and the surplus is the same as in the current period. The same logic applies to the surplus function of the worker. The instantaneous surplus over unemployment is captured by the difference in utility $u(w_o - \tau) - u(b_o - \tau)$ where τ is the lump sum tax. The continuation value of the match is diminished by the value of search V_o that unemployed workers pursue in the next period (employed workers do not search on the job).

At the layoff stage, the worker is dismissed if and only if firm surplus is negative, $J_o(w_o; y) < 0$. This can be rewritten in the form $y < \underline{y}_o(w_o) := w_o - \beta_o\phi\mathbb{E}J_o^+$, where $\underline{y}_o(w_o)$ is the *layoff threshold*. In case of a layoff, the firm is left with a vacant job, which generates a value of zero. Taking this into account, firm surplus at the search stage is $\mathbb{E}J_o^+(w_o) = \int_{\underline{y}_o(w_o)}^{\infty} J_o(w_o; y)\, dF_o(y)$. By equation (2.2), $J_o(w_o; y) = \frac{y - \underline{y}_o(w_o)}{1-\beta_o(1-\phi)}$, and therefore the layoff threshold solves

$$\underline{y}_o - w_o + \frac{\beta_o\phi}{1-\beta_o(1-\phi)} \int_{\underline{y}_o}^{\infty} y - \underline{y}_o \, dF_o(y) = 0. \tag{2.4}$$

The following proposition establishes that the layoff threshold is well-defined and summarizes how it reacts to changes in the model parameters.

Proposition 2.1. *For any $w_o \in \mathbb{R}$, equation* (2.4) *uniquely defines a layoff threshold $\underline{y}_o$. The layoff threshold is increasing in w_o and decreasing in β_o, ϕ, μ_o, and s_o.*

The proof of this proposition and all other propositions can be found in Section 2.B.3. *Ceteris paribus*, a higher wage decreases firm profit such that a higher productivity level is necessary for the firm to break even. The remaining parameters examined in Proposition 2.1 all increase future expected firm profit, and therefore the firm is willing to accept lower profits today. For future reference, define expected firm surplus conditional on retention as $J_o(\underline{y}_o(w_o)) = \mathbb{E}[J_o(w_o; Y_o)|Y_o \geq \underline{y}_o(w_o)] = \frac{\mathbb{E}[Y_o - \underline{y}_o(w_o)|Y_o \geq \underline{y}_o(w_o)]}{1-\beta_o(1-\phi)}$, which only depends on w_o via the layoff threshold $\underline{y}_o(w_o)$. To simplify notation, dependence of $\underline{y}_o$ on the wage is omitted in the following.

Expected worker surplus at the search stage is $\mathbb{E}W_o^+(w_o) = (1-F_o(\underline{y}_o))W_o(w_o)$. Substituting this back into (2.3) yields $W_o(w_o) = \frac{u(w_o-\tau)-u(b_o-\tau)-\beta_o V_o}{1-\beta_o(1-\phi F_o(\underline{y}_o))}$. In her optimal application decision, the worker takes the value V_o as given. Yet, in equilibrium $V_o = p(\theta_o^*)\mathbb{E}W_o^+(w_o^*)$ must hold.[5]

Equilibrium conditions

The first order optimality conditions of problem (2.1) can be summarized as

$$u'(w_o^* - \tau) = \frac{1-\gamma}{\gamma}\frac{W_o(w_o^*)}{J_o(\underline{y}_o^*)} + (1-\beta_o(1-\phi))h_o(\underline{y}_o^*)\frac{\partial \underline{y}_o^*}{\partial w_o}W_o(w_o^*), \tag{2.5}$$

$$q(\theta_o^*)\mathbb{E}J_o^+(w_o^*) = c, \tag{2.6}$$

where $\underline{y}_o^* = y_o(w_o^*)$ is defined in (2.4). The left-hand side of equation (2.5) captures the utility gain from a marginally higher wage, whereas the right-hand side combines the marginal costs of a higher wage. The first term on the right-hand side is standard in the literature and reflects the search friction. The higher the wage, the lower the worker's probability of finding a job. The second term on the right-hand side is novel and stems from the contracting friction. In case of a layoff, the worker loses the match surplus $W_o(w_o^*)$. The product $H_o(w_o) = h_o(\underline{y}_o)\frac{\partial \underline{y}_o}{\partial w_o}$ reflects the link between wage level and job security. It combines the marginal effect of w_o on the firm's layoff threshold y_o, measured by the partial derivative $\frac{\underline{y}_o^*}{\partial w_o} = \frac{1-\beta_o(1-\phi)}{1-\beta_o(1-\phi F_o(\underline{y}_o^*))} > 0$, and the hazard rate $h_o(\underline{y}_o^*)$. The latter determines how sensitive the retention probability responds to a change in the layoff threshold, since in general terms $h_o(x) = \frac{f_o(x)}{1-F_o(x)} = -\frac{\partial \ln(1-F_o(x))}{\partial x}$. The product $H_o(w_o)$ can therefore be interpreted as the marginal rate of substitution between the wage w_o and the log probability of retention $\ln(1-F_o(\underline{y}_o))$. If $H_o(w_o) = 0$, the retention probability is inelastic to the wage and the worker does not act against the risk. In this case, condition (2.5) implies that the worker earns a share γ of the joint surplus of employment $\frac{W_o(w_o^*)}{u'(w_o^*-\tau)} + J_o(\underline{y}_o^*)$. This is the usual finding when bargaining is bilaterally efficient as in Acemoglu and Shimer (1999). With $H_o(w_o) > 0$ it is no longer true. The higher $H_o(w_o)$, the more the worker is willing to decrease her wage in favor of a higher retention probability. This reduces the worker's share

[5] Since the worker's reservation wage is independent of match productivity, the possibility of voluntary quits can be safely ignored.

in match surplus below γ, and the firm earns an additional rent.[6]

The labor market equilibrium on the labor market of the old job seekers is characterized by the conditions (2.4)–(2.6), together with $V_o = p(\theta_o^*)\mathbb{E}W_o^+(w_o^*)$. For the special case that old age lasts for one period only ($\pi_o = 1$), existence and uniqueness of a labor market equilibrium can be established analytically. The threshold productivity then equals the wage, $\underline{y}_o(w_o) = w_o$, and the worker's reservation wage is her unemployment income b_o.

Proposition 2.2. *Let $\pi_o = 1$. For given tax level τ, a unique labor market equilibrium of old job seekers (θ_o^*, w_o^*, V_o) exists and satisfies $w_o^* > b_o$.*

Since the optimal wage w_o^* exceeds the worker's reservation wage b_o, part of the layoffs that occur in equilibrium are bilaterally inefficient. If the informational friction could be overcome, it would be optimal to maintain all matches with productivity $Y_o \geq b_o$, because in this case the value the individual generates in employment exceeds the value of non-employment. Due to the contracting friction, however, also matches with $Y_o \in (b_o, w_o^*)$ are dissolved because of negative firm profit. The probability for such a bilaterally inefficient layoff is $F_o(w_o^*) - F_o(b_o)$.

Comparative static effects

To obtain comparative static effects, I continue to assume that old age lasts for one period only, $\pi_o = 1$. Equation (2.5) then can be expressed as

$$\Phi(w_o^*) = u'(w_o^* - \tau) - \frac{1-\gamma}{\gamma}\frac{W_o(w_o^*)}{J_o(w_o^*)} - h_o(w_o^*)W_o(w_o^*) = 0, \tag{2.7}$$

where $W_o(w_o) = u(w_o - \tau) - u(b_o - \tau)$ and $J_o(w_o) = \mathbb{E}[Y_o - w_o | Y_o \geq w_o]$ since $\underline{y}_o(w_o) = w_o$. A marginal change in one of the model parameters in general spurs two effects to which the worker responds.

The first effect, which I refer to as *income effect (IE)*, captures the worker's reaction to changes in the surplus functions W_o and J_o, and the distribution function F_o. The income effect of an arbitrary parameter ξ on the equilibrium wage is

$$\left(\frac{\partial w_o^*}{\partial \xi}\right)^{IE} = -\Phi'(w_o^*)^{-1}\left\{\frac{1-\gamma}{\gamma}\frac{W_o(w_o^*)}{J_o(w_o^*)^2}\frac{\partial J_o(w_o^*)}{\partial \xi} - \left[\frac{1-\gamma}{\gamma}\frac{1}{J_o(w_o^*)} + h_o(w_o^*)\right]\frac{\partial W_o(w_o^*)}{\partial \xi}\right\}$$

where $\Phi'(w_o^*) < 0$. In absence of a contracting friction, only this income effect occurs. With a contracting friction, however, also the worker's valuation of risk may change. This corresponds to a change in the hazard function h_o on the right-hand side of (2.7) and triggers a *substitution effect (SE)*,

$$\left(\frac{\partial w_o^*}{\partial \xi}\right)^{SE} = \Phi'(w_o^*)^{-1}\frac{\partial h_o(w_o^*)}{\partial \xi}W_o(w_o^*).$$

[6]This is similar to the *informational rent* highlighted by Kennan (2010). Lemma 2.B.2(i) can be used to show that the optimal worker share in surplus always lies in the interval $(\frac{\gamma}{1+\gamma}, \gamma)$.

The marginal effect of an arbitrary parameter ξ on the equilibrium layoff probability is

$$\frac{dF_o(w_o^*)}{d\xi} = \frac{\partial F_o(w_o^*)}{\partial \xi} + f_o(w_o^*)\frac{\partial w_o^*}{\partial \xi} = \underbrace{\frac{\partial F_o(w_o^*)}{\partial \xi} + f_o(w_o^*)\left(\frac{\partial w_o^*}{\partial \xi}\right)^{IE}}_{IE} \underbrace{+ f_o(w_o^*)\left(\frac{\partial w_o^*}{\partial \xi}\right)^{SE}}_{SE}. \tag{2.8}$$

It combines the direct effect of ξ on the productivity distribution and the indirect effect through the equilibrium wage w_o^*. By the free entry condition (2.6), the equilibrium job-finding probability is determined by expected firm surplus $\mathbb{E}J_o^+(w_o^*)$. Higher expected surplus boosts vacancy-posting, which increases the labor market tightness θ_o^* and the job-finding probability $p(\theta_o^*)$. Expected firm surplus is also affected by parameter changes through a direct distributional effect and an indirect wage effect,

$$\begin{aligned} \frac{d\mathbb{E}J_o^+(w_o^*)}{d\xi} &= -\int_{w_o^*}^{\infty} \frac{\partial F_o(y)}{\partial \xi}\, dy - (1 - F_o(w_o^*))\frac{\partial w_o^*}{\partial \xi} \\ &= \underbrace{-\int_{w_o^*}^{\infty} \frac{\partial F_o(y)}{\partial \xi}\, dy - (1 - F_o(w_o^*))\left(\frac{\partial w_o^*}{\partial \xi}\right)^{IE}}_{IE} \underbrace{-(1 - F_o(w_o^*))\left(\frac{\partial w_o^*}{\partial \xi}\right)^{SE}}_{SE}. \end{aligned} \tag{2.9}$$

From the above expressions it is easy to see how a change in the worker's valuation of risk, h_o, affects the labor market equilibrium through the substitution effects. If the retention probability becomes locally more sensitive to the wage, $\frac{\partial h_o(w_o^*)}{\partial \xi} > 0$, the worker substitutes away from wage income in favor of a higher retention probability and a higher job-finding probability. The opposite happens if $\frac{\partial h_o(w_o^*)}{\partial \xi} < 0$. In the following, I illustrate the comparative static effects of the most relevant model parameters.

Unemployment income. An increase in b_o, for instance due to higher unemployment or early retirement benefits, lowers worker surplus W_o. Because the productivity distribution is unaffected, there is no change in J_o and h_o, and also no substitution effect. The income effect increases the equilibrium wage since the worker's outside option improves. This increases the layoff probability and lowers the job-finding probability.

Old-age productivity. The productivity parameters μ_o and s_o affect expected firm surplus and the hazard function, but not worker surplus. The sign of the partial derivatives of h_o and J_o are established in Lemma 2.B.1 and Lemma 2.B.2 in Section 2.B, respectively. An increase in the location parameter μ_o shifts the productivity distribution to the right, which raises firm surplus and lowers the hazard for given wage. Both the higher productivity (IE) and the lower valuation of risk (SE) increase the equilibrium wage. Furthermore, the distribution function decreases for given wage, $\frac{\partial F_o(w_o^*)}{\partial \mu_o} = -f_o(w_o^*) < 0$. Proposition 2.3 establishes that this negative direct effect dominates the positive wage effect in (2.8) and (2.9) because the wage increase is less than proportional, $\frac{\partial w_o^*}{\partial \mu_o} < 1$. As a result, the equilibrium layoff probability decreases and the job-finding probability increases when the productivity distribution shifts to the right.

Proposition 2.3. *A marginal increase in the location parameter μ_o increases the equilibrium wage w_o^*, lowers the layoff probability $F_o(w_o^*)$, and increases the job-finding probability $p(\theta_o^*)$.*

An increase in the scale parameter s_o has potentially ambiguous effects on the labor market equilibrium. Under additional assumptions, however, it is possible to derive analytical results.

Proposition 2.4. *A marginal increase in the scale parameter s_o exerts a positive income effect on w_o^*. The substitution effect is positive if and only if $\frac{w_o^*-\mu_o}{s_o} > \hat{z}$, where $\hat{z} < 0$ is the unique root of $h(z) + h'(z)z$.*

Assume that $w_o^* \leq \mu_o$. Then the layoff probability increases in s_o, and the job-finding probability increases if either $\frac{\partial w_o^*}{\partial s_o} \leq 0$ or $\gamma \leq \frac{J_o(w_o^*)+w_o^*-\mu_o}{J_o(w_o^*)+[1-J_o(w_o^*)h_o(w_o^*)](w_o^*-\mu_o)}$.

Wage. The firm benefits from a more dispersed productivity distribution because the mass of very productive workers is increasing, while the increasing mass of unproductive workers is laid off at no cost. As a result, the average productivity per retained worker increases, $\frac{\partial J_o(w_o^*)}{\partial s_o} > 0$, generating a positive income effect on w_o^*. The substitution effect can be positive or negative, depending on the reaction of the hazard function. For $\frac{w_o^*-\mu_o}{s_o} < \hat{z}$, the hazard function increases as the retention probability $1-F_o$ becomes locally more sensitive to the wage (cf. Lemma 2.B.1). In response, workers are willing to give up part of their wage in favor of higher job security. However, if wages are sufficiently high such that $\frac{w_o^*-\mu_o}{s_o} > \hat{z}$, increasing uncertainty actually decreases the willingness to substitute wages for job security because the retention rate becomes locally less responsive to the wage. This non-monotonic behavior occurs because an increase in s_o makes the distribution function steeper at the tails of the distribution, while it becomes flatter in the middle. The equilibrium wage therefore unambiguously increases if $\frac{w_o^*-\mu_o}{s_o} > \hat{z}$, while the wage response is analytically not clear otherwise.

Layoffs. A higher scale parameter s_o increases the distribution function for $w_o^* \leq \mu_o$ and decreases it for $w_o^* \geq \mu_o$. I consider the first case more relevant for real world applications, such that $\frac{\partial F_o(w_o^*)}{\partial s_o} = -\frac{w_o^*-\mu_o}{s_o^2} f_o(w_o^*) \geq 0$. It can be shown that under this condition, the positive income effect always offsets the potentially negative substitution effect in (2.8), such that the equilibrium layoff probability increases. Therefore, even if the worker responds to higher uncertainty by contracting a lower wage, layoffs become more likely.

Hiring. The direct effect of s_o on the job-finding probability is positive since $-\int_{w_o^*}^{\infty} \frac{\partial F_o(y)}{\partial s_o} dy = \frac{1-F_o(w_o^*)}{s_o}[J_o(w_o^*) + w_o^* - \mu_o] \geq 0$ (see proof of Proposition 2.4). Intuitively, the higher expected productivity per retained worker more than compensates the firm for the lower retention probability of the workers. If the equilibrium wage decreases in s_o, this further increases firm surplus, and the job-finding probability unambiguously increases as evident from (2.9). If $\frac{\partial w_o^*}{\partial s_o} > 0$, the upper boundary on γ established by Proposition 2.4 ensures that the wage increase does not offset the direct distributional effect. Intuitively, the lower γ, the more of the additional match surplus per retained worker is captured by the firm, and the less the equilibrium wage increases.

2.4.2 Labor market equilibrium of prime-age job seekers

I now turn to the search problem of prime-age job seekers, who look for a wage contract $\omega_m = (w_m, w_s)$. As above, the directed search equilibrium on the labor market of prime-age job seekers can be characterized as the solution to the optimization problem

$$V_m := \max_{(\theta_m,\omega_m)} p(\theta_m)\mathbb{E}W_m^+(\omega_m) \quad \text{s.t.} \quad q(\theta_m)\mathbb{E}J_m^+(\omega_m) = c.$$

At the production stage, firm and worker surplus evolve according to

$$J_m(\omega_m; y) = y - w_m + \beta_m[\phi\mathbb{E}J_m^+(\omega_m) + (1-\phi)J_m(\omega_m; y)] + \beta\pi_m(1-\sigma)\mathbb{E}J_s^+(w_s), \tag{2.10}$$

$$\begin{aligned} W_m(\omega_m) = u(w_m - \tau) - u(b_m - \tau) + \beta_m[\phi\mathbb{E}W_m^+(\omega_m) + (1-\phi)W_m(\omega_m) - V_m] \\ + \beta\pi_m(1-\sigma)[\mathbb{E}W_s^+(w_s) - V_o]. \end{aligned} \tag{2.11}$$

where $\beta_m := \beta(1-\pi_m)(1-\sigma)$ is the effective discount factor of a prime-age worker. If the worker receives the aging shock at the end of the period, she becomes a senior worker. Matches with senior workers generate an expected surplus of $\mathbb{E}J_s^+(w_s)$ and $\mathbb{E}W_s^+(w_s)$, which are defined in the same way as $\mathbb{E}J_o^+(w_o)$ and $\mathbb{E}W_o^+(w_o)$ above, except that the distribution function F_o has to be exchanged for F_s.

Likewise, the layoff threshold of a senior worker is defined as in (2.4). The layoff threshold of a prime-age worker is denoted by $\underline{y}_m(\omega_m)$ and characterized by the equation

$$\underline{y}_m - w_m + \frac{\beta_m\phi}{1-\beta_m(1-\phi)}\int_{\underline{y}_m}^{\infty} y - \underline{y}_m \, dF_m(y) + \beta\pi_m(1-\sigma)\mathbb{E}J_s^+(w_s) = 0. \tag{2.12}$$

Compared to equation (2.4), matches with prime-age workers bear an additional continuation value, $\beta\pi_m(1-\sigma)\mathbb{E}J_s^+(w_s)$, because of their larger distance from retirement age. This reflects the *horizon effect* highlighted by Chéron et al. (2013). Everything else equal, the layoff thresholds satisfy $\underline{y}_m < \underline{y}_s$, such that prime-age workers are less likely to be laid off compared to senior workers. The properties established in Proposition 2.1 apply also to $\underline{y}_m$ and $\underline{y}_s$. Expected firm surplus at the search stage is $\mathbb{E}J_m^+(\omega_m) = (1-F_m(\underline{y}_m))J_m(\underline{y}_m)$ where $J_m(\underline{y}_m) := \frac{\mathbb{E}[Y_m - \underline{y}_m | Y_m \geq \underline{y}_m]}{1-\beta_m(1-\phi)}$ is expected firm surplus conditional on employment. Expected worker surplus is $\mathbb{E}W_m^+(\omega_m) = (1-F_m(\underline{y}_m))W_m(\omega_m)$ where $W_m(\omega_m) = \frac{u(w_m-\tau) - u(b_m-\tau) - \beta_m V_m + \beta\pi_m(1-\sigma)[\mathbb{E}W_s^+(w_s) - V_o]}{1-\beta_m(1-\phi F_m(\underline{y}_m))}$.

Equilibrium conditions

The first order conditions for an optimal wage contract $\omega_m^* = (w_m^*, w_s^*)$ with $w_s^* > b_o$ are

$$u'(w_m^* - \tau) = \frac{1-\gamma}{\gamma}\frac{W_m(\omega_m^*)}{J_m(\underline{y}_m^*)} + (1-\beta_m(1-\phi))h_m(\underline{y}_m^*)\frac{\partial \underline{y}_m^*}{\partial w_m}W_m(\omega_m^*), \tag{2.13}$$

$$u'(w_s^* - \tau) = u'(w_m^* - \tau) + (1-\beta_o(1-\phi))h_s(\underline{y}_s^*)\frac{\partial \underline{y}_s^*}{\partial w_s}W_s(w_s^*), \tag{2.14}$$

$$q(\theta_m^*)\mathbb{E}J_m^+(\omega_m^*) = c, \tag{2.15}$$

where the layoff threshold $\underline{y}_m^* = \underline{y}_m(\omega_m^*)$ is defined in (2.12) and $\underline{y}_s^* = \underline{y}_s(w_s^*)$ is defined analogous to (2.4). Condition (2.13) resembles equation (2.5) and determines the optimal split of expected total job surplus from employment $\frac{W_m(\omega_m)}{u'(w_m-\tau)} + J_m(\underline{y}_m)$. Workers again face a trade-off between wages and job security, as an increase in either w_m or w_s increases the layoff threshold $\underline{y}_m$ and thereby the layoff probability. How strongly workers respond to the layoff risk depends on the product $H_m(\omega_m) = h_m(\underline{y}^*)\frac{\partial \underline{y}_m^*}{\partial w_m}$, which measures how sensitive the prime-age retention probability $1 - F_m(\underline{y}_m)$ reacts to changes in w_m.

While (2.13) determines the present value that the worker receives in optimum, condition (2.14) pins down the optimal intertemporal wage profile that implements this value. It reflects a trade-off between consumption smoothing (in the absence of savings this has to be accomplished by the wage contract) and old-age job security. In absence of uncertainty, $H_s(w_s) = h_s(\underline{y}_s^*)\frac{\partial \underline{y}_s^*}{\partial w_s} = 0$, the optimal contract features a flat wage profile, $w_m^* = w_s^*$. By condition (2.14), risk considerations let the worker contract a lower wage in the second period such that $w_m^* > w_s^*$. The reason is that a higher w_s increases the layoff risk in old age (through $\underline{y}_s$) but also during prime age (through the lower continuation value in $\underline{y}_m$). Whereas a higher w_m increases the layoff risk only during prime age. This generates an incentive to front-load wage income. According to (2.14), how much wages should fall in late working age depends on the marginal rate of substitution between wage income and job security, $H_s(w_s)$, and the utility loss in case of a layoff, $W_s(w_s)$.

To theoretically establish existence and uniqueness of an equilibrium, I assume that prime age and old age each last for only one period, which corresponds to $\pi_m = \pi_o = 1$. Figure 2.3 visualizes the two equations (2.13)–(2.14) in the (w_m, w_s)-space. Condition (2.13) defines a decreasing curve, which I refer to as the surplus sharing (SS) curve in Figure 2.3. It connects all wage combinations that implement the optimal surplus sharing rule. Condition (2.14) defines the upwards sloping consumption smoothing (CS) curve. The CS curve is flat for $w_m \leq b_o$ because the worker's participation constraint, $W_s(w_s) = w_s - b_o \geq 0$, binds in old age. The unique intersection of the two curves defines the optimal wage contract $\omega_m^* = (w_m^*, w_s^*)$.

Proposition 2.5. *Let $\pi_m = \pi_o = 1$ and $b_m \leq b_o$. For given tax level τ, a unique labor market equilibrium of prime-age job seekers $(\theta_m^*, \omega_m^*, V_m)$ exists. There exists a $\bar{b}_o > b_m$, such that for $b_o \in [b_m, \bar{b}_o)$ the wage contract is interior and the wage level is decreasing with age, $w_m^* > w_s^* > b_o$. For $b_o \geq \bar{b}_o$, the optimal contract satisfies $w_m^* \leq w_s^* = b_o$.*

Proposition 2.5 establishes that unless old workers enjoy very high outside options, the optimal contract pays above the reservation wage in old age, $w_s^* > b_o$. Because the CS curve lies below the 45 degrees line, the optimal wage contract is then decreasing in age due to the risk considerations highlighted above. If b_o is much higher than b_m, however, the worker's participation constraint $w_s^* = b_o$ may become binding in old age. The worker is then indifferent between work and unemployment. In Figure 2.3 this would correspond to an intersecting point that lies in the flat part of the CS curve. This case does not appear to be very relevant in practice. Although the baseline calibration of the model given in Table 2.2 grants a 30% higher

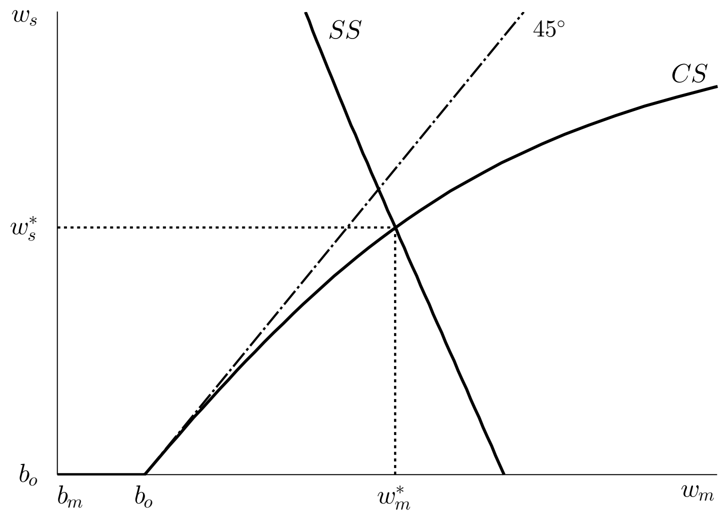


Figure 2.3. Wage determination of prime-age job seekers.

unemployment income to senior workers compared to prime-age workers, the optimal contract is still interior, as can be seen from Table 2.3.

Comparative static effects

How the labor market equlibrium of prime-age job seekers responds to parameter changes depends on how the SS and CS curve are affected. Throughout the section, I assume that ω_m^* is an interior solution as illustrated in Figure 2.3 and that each stage of the life-cycle deterministically lasts for one period ($\pi_m = \pi_o = 1$). This implies that the layoff threshold of a senior worker is $\underline{y}_s(w_s) = w_s$, while the layoff threshold of a prime-age worker is $\underline{y}_m(\omega_m) = w_m - \beta(1-\sigma)\mathbb{E}J_s^+(w_s)$.

Prime-age productivity. I first discuss how the parameters of the prime-age productivity distribution, μ_m and s_m, affect the equilibrium. The results are very similar to those of Section 2.4.1. From the first order conditions (2.13)–(2.14) it can be seen that these parameters only affect the SS curve. An increase in μ_m moves the SS curve to the right. As a result, the new intersecting point exhibits higher wages in both periods. Since the slope of the CS curve is less than 1, the prime-age wage increases more than the senior wage, such that the wage decline at the end of the career becomes more pronounced. Provided that the income effect dominates the substitution effect, the same wage effects are observed for an increase in s_m (compare Proposition 2.4).

The job-finding probability $p(\theta_m^*)$ and the layoff probability of prime-age workers $F_m(\underline{y}_m^*)$ are affected by changes in the productivity parameters both directly through the distribution function and indirectly through the response of equilibrium wages that affect the layoff threshold $\underline{y}_m^* = \underline{y}_m(\omega_m^*)$. By contrast, the layoff probability of senior workers, $F_s(w_s^*)$, depends on the prime-age productivity distribution only through the equilibrium wage. The two layoff probabilities may therefore react differently to parameter changes.

Proposition 2.6. *A marginal increase in the location parameter μ_m increases the equilibrium wages (w_m^*, w_s^*) in both periods, increases the job-finding probability $p(\theta_m^*)$, and decreases the layoff probability of prime-age workers $F_m(\underline{y}_m^*)$. Due to the higher wage, the layoff probability of senior workers $F_s(w_s^*)$ increases.*

Let $\underline{y}_m^* \leq \mu_m$. Then a marginal increase in the scale parameter s_m increases the layoff probability of prime-age workers. The job-finding probability increases if either $\frac{\partial \underline{y}_m^*}{\partial s_m} < 0$ or $\gamma \leq \frac{J_m(\underline{y}_m^*)+\underline{y}_m^*-\mu_m}{J_m(\underline{y}_m^*)+[1-J_m(\underline{y}_m^*)h_m(\underline{y}_m^*)](\underline{y}_m^*-\mu_m)}$.

The economic intuition underlying these results is tantamount to Proposition 2.3 and Proposition 2.4, and not repeated at this point.

Senior productivity. Changes in the parameters μ_s and s_s alter the productivity distribution of senior workers, which affects both the SS and the CS curve. This makes analytical predictions less clear-cut. I start the discussion with the CS curve. It is easy to see from (2.14) that the curve always goes through the point $(w_m, w_s) = (b_o, b_o)$ and has a slope less than 1 as indicated in Figure 2.3. The CS curve becomes steeper if h_s decreases, since a lower hazard increases the optimal degree of consumption smoothing. A change in the CS curve constitutes a pure *substitution effect* in the manner of Section 2.4.1 because it is caused by an altered hazard function h_s. The SS curve, by contrast, is affected by the productivity parameters of senior workers through the continuation values $\mathbb{E}J_s^+(w_s)$ and $\mathbb{E}W_s^+(w_s)$, which enter the terms $\underline{y}_m$ and $W_m(\omega_m)$. Any change in the SS curve therefore constitutes an *income effect.* In absence of the contracting friction, only the income effect would be present.

A higher μ_s increases retention probabilities and expected output per employed worker in old age. This translates into higher firm and worker surplus during prime age and lowers the layoff threshold $\underline{y}_m$. Since $W_m(\omega_m)$ and $J_m(\underline{y}_m)$ both increase, the effect on the surplus ratio in (2.13) is in general ambiguous. Under an additional assumption, however, the effect on firm surplus dominates.

Proposition 2.7. *Assume that in equilibrium $\gamma \leq \frac{W_m(\omega_m^*)}{u'(w_s^*-\tau)J_m(\underline{y}_m^*)}$.*[7] *Then a marginal increase in the location parameter μ_s raises w_s^*, while the effect on w_m^* is ambiguous. The IE acts to increase both w_s^* and w_m^*, the SE acts to increase w_s^* and reduce w_m^*.*

Under the assumption of Proposition 2.7, higher productivity at the senior stage raises prime-age firm surplus more than prime-age worker surplus. To restore optimal surplus sharing, the worker increases both w_m and w_s due to an income effect, and the SS curve shifts to the right as illustrated in Figure 2.4. Additionally, a higher μ_s makes the CS curve steeper. Since a higher μ_s lowers the hazard function h_s, workers are less inclined to give up wage income for job security. The new intersection point in Figure 2.4 features an unambiguously higher w_s^*, while w_m^* may increase or decrease. The higher expected surplus in old age lets w_m^* increase by an

[7]Note that $u'(w_m^* - \tau) \leq \frac{1}{\gamma}\frac{W_m(\omega_m^*)}{J_m(\underline{y}_m^*)}$ by (2.13) and Lemma 2.B.1(i). Therefore the assumption is satisfied if w_s^* is not substantially lower than w_m^*.

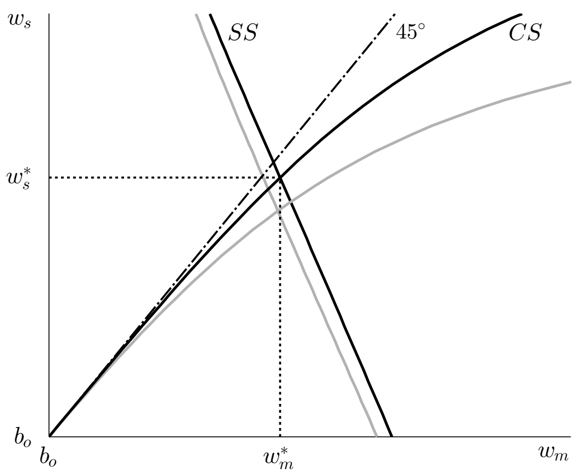


Figure 2.4. Wage response to an increase in μ_s.

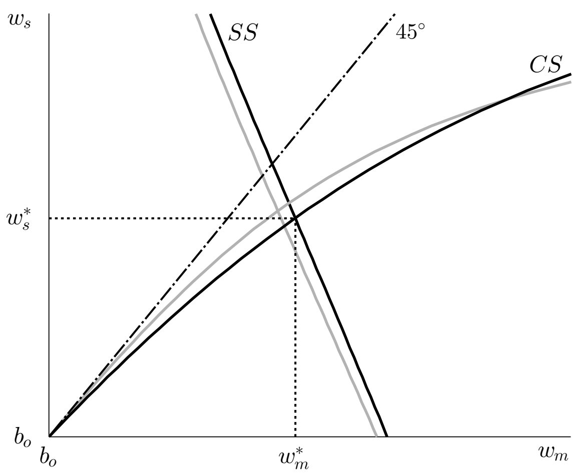


Figure 2.5. Wage response to an increase in s_s.

income effect, while the reduction in layoff risk in old age leads the worker to substitute away from w_m^*.

A larger dispersion s_s also increases expected firm surplus in old age, which translates into a higher firm surplus and a smaller layoff threshold during prime age. Old-age expected worker surplus, $\mathbb{E}W_s^+(w_s) = (1 - F_s(w_s))W_s(w_s)$, by contrast, declines in s_s through the lower retention probability, which then also lowers worker surplus during prime age. Therefore, a more dispersed productivity distribution shifts the SS curve unambiguously to the right in Figure 2.5. *Ceteris paribus*, the worker's share in match surplus falls, to which she responds by demanding higher wages in both periods. The effect of s_s on the CS curve is not monotone because the sign of $\frac{\partial h_s(w_s)}{\partial s_s}$ depends on whether $\frac{w_s - \mu_s}{s_s} \gtreqless \hat{z}$ (cf. Lemma 2.B.1). For w_s sufficiently low, an increase in s_s increases the worker's valuation of risk. This makes the CS curve flatter because the optimal degree of consumption smoothing decreases. The opposite happens for high w_s, as evident from Figure 2.5. In the figure, the curve becomes flatter around the old intersection point because $\frac{\partial h_s(w_s^*)}{\partial s_s} > 0$. The higher layoff hazard leads the worker to give up part of w_s^* in favor of w_m^* to increase the old-age retention rate $1 - F_s(w_s^*)$.

Unemployment income. Since the unemployment incomes b_m and b_o do not affect the hazard functions, the response of equilibrium wages is due to income effects that are driven by changes in match surplus. A higher b_m *ceteris paribus* decreases prime-age worker surplus due to better outside options. To restore optimal surplus sharing, the worker increases wages in both periods. This is captured by the outwards shift of the SS curve in Figure 2.6. Since b_m does not affect the CS curve, the new optimum exhibits a higher w_m^*, a higher w_s^*, and a lower ratio w_s^*/w_m^*. The higher wages translate into higher layoff probabilities in both periods and a lower job-finding probability.

Higher unemployment income for older workers, b_o, has the same effect on the SS curve as b_m. Additionally, the CS curve moves upwards in Figure 2.7 because a layoff at the senior stage becomes less costly for the worker. As a result, w_s^* increases at the expense of w_m^*. In total,

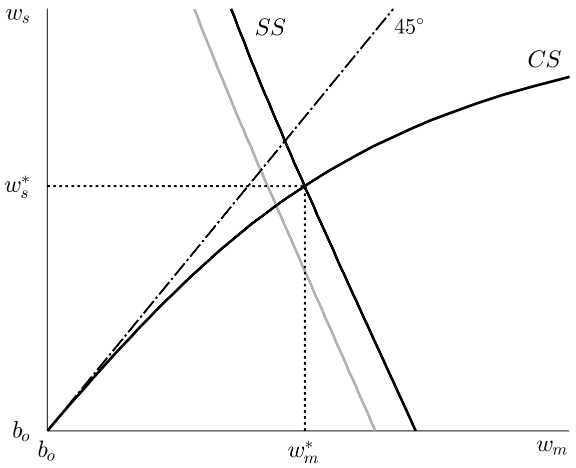


Figure 2.6. Wage response to an increase in b_m.

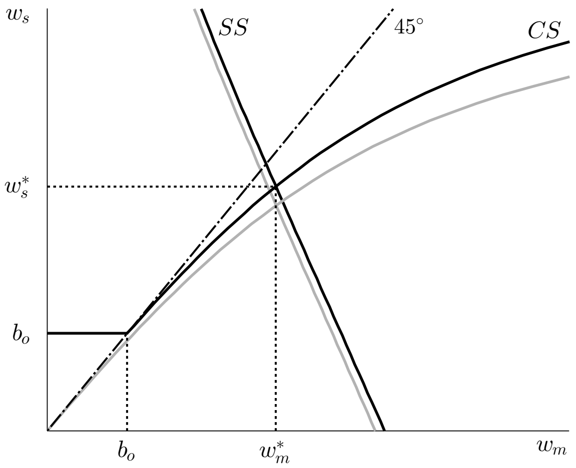


Figure 2.7. Wage response to an increase in b_o.

there are two upwards forces on w_s^*, which unambiguously increases, accompanied by a higher layoff probability in old age. The effect on the prime-age wage w_m^* is not clear. As long as w_m^* does not substantially decrease, however, higher b_o will also increase layoffs among prime-age workers (through a higher $\underline{y}_m^*$) and lower the job-finding probability.

Proposition 2.8. *An increase in* b_m *raises* w_m^* *and* w_s^**, and lowers* w_s^*/w_m^**. This increases layoff probabilities for prime-age and senior workers, and lowers the job-finding probability* $p(\theta_m^*)$*. An increase in* b_o *raises* w_s^* *and thereby the layoff rate* $F_s(w_s^*)$*, while the effect on* w_m^* *is ambiguous.*

These observations suggest that a change in outside options of a certain age group has stronger wage (and likely employment) effects on that age group, although workers are optimizing intertemporally.

2.4.3 Demography and economic aggregates

For simplicity, I assume a stationary demography. In each period, the inflow into an age group equal its outflow. Since the mass of newborns is normalized to 1, in steady state there is as mass $N_1 = \frac{1}{\pi_m}$ of prime-age individuals and a mass $N_2 = \frac{1}{\pi_o}$ of individuals in old working age. The total mass of the population is $N = N_1 + N_2$. By assumption, all prime-age individuals participate in the labor market, while older individuals become non-participants with a probability δ each period. Their participation rate equals $lf_2 = \frac{\pi_o}{1-(1-\pi_o)(1-\delta)}$ in steady state.

Employment. In steady state, the mass of type i workers remains constant over time,

$$\begin{aligned}
E_m &= p(\theta_m^*)(1 - F_m(\underline{y}_m^*))JS_m + (1 - \pi_m)(1 - \sigma)(1 - \phi F_m(\underline{y}_m^*))E_m, \\
E_o &= p(\theta_o^*)(1 - F_o(\underline{y}_o^*))JS_o + (1 - \pi_o)(1 - \sigma)(1 - \delta)(1 - \phi F_o(\underline{y}_o^*))E_o, \\
E_s &= \pi_m(1 - \sigma)E_m(1 - F_s(\underline{y}_s^*)) + (1 - \pi_o)(1 - \sigma)(1 - \delta)(1 - \phi F_s(\underline{y}_s^*))E_s,
\end{aligned}$$

where the stocks refer to the mass of employed workers at the production stage (cf. Figure 2.1). The prime-age employment rate is $e_1 = \frac{E_m}{N_1}$, while the old-age employment rate is $e_2 = \frac{E_s+E_o}{N_2}$. In each of the equations above, the second term of the sum captures the mass of workers that remain in the respective employment state, while the first term measures the inflow of new workers. The inflow of senior workers (s) equals the mass of aging prime-age workers who have been retained by their employer. The inflow of prime-age (m) and old workers (o) amounts to the new hires, where JS_m and JS_o are the mass of job seekers in the respective labor market, given by

$$
\begin{aligned}
JS_m &= 1 + (1-\pi_m)(N_1 - (1-\sigma)E_m),\\
JS_o &= \pi_m[N_1 - (1-\sigma)E_m] + (1-\pi_o)(1-\delta)[lf_2 N_2 - (1-\sigma)E_o].
\end{aligned}
$$

The mass of type i job seekers differs from the mass of unemployed individuals due to the timing convention of Figure 2.1. An individual who is employed at the production stage may be hit by an exogenous separation shock at the end of the period and become a job seeker. Prime-age job seekers comprise newborn individuals (normalized to 1) and individuals unemployed at the end of the period who remain in prime age. Old job seekers consist of unemployed prime-age individuals hit by the aging shock (first term) and unemployed old individuals who are still participating (second term).

When calibrating the model, I target two features of the cross-sectional distribution of tenure and unemployment. The first target measures the share of matches of prime-age workers that have tenure of less than one period. In each period, $E_m^0 = p(\theta_m^*)JS_m$ new matches with prime-age workers are created. Thereof, $E_m^1 = E_m^0(1 - F_m(\underline{y}_m^*))(1-\pi_m)(1-\sigma)$ workers complete at least a full period in their new job. For $s \geq 2$, the mass of matches with s periods of tenure evolves according to $E_m^s = E_m^{s-1}(1-\pi_m)(1-\sigma)(1-\phi F_m(\underline{y}_m^*))$. From these expressions, the cross-sectional share of matches that are in their first period can be computed as

$$
e_m^0 := \frac{E_m^0}{\sum_{s=0}^{\infty} E_m^s} = \frac{1-(1-\pi_m)(1-\sigma)(1-\phi F_m(\underline{y}_m^*))}{1-(1-\pi_m)(1-\sigma)(1-\phi)F_m(\underline{y}_m^*)}.
$$

The second target refers to the duration of prime-age unemployment and captures the cross-sectional share of unemployed individuals whose duration in unemployment is less than one period. Unemployment spells are interrupted whenever a new match is formed, even if this match is dissolved before the production stage. Since the period probability of staying in prime age and unemployed is $(1-p(\theta_m^*))(1-\pi_m)$, the mass of workers with s periods of uninterrupted unemployment satisfies $U_m^s = U_m^{s-1}(1-p(\theta_m^*))(1-\pi_m)$. The share of short-term unemployed in all unemployed is therefore

$$
u_m^0 := \frac{U_m^0}{\sum_{s=0}^{\infty} U_m^s} = 1 - (1-p(\theta_m^*))(1-\pi_m).
$$

Output. Output per age group is the value of produced goods net of vacancy posting costs,

$$Y_1 = \mathbb{E}[Y_m | Y_m \geq \underline{y}_m^*] E_m - c\theta_m^* J S_m,$$
$$Y_2 = \mathbb{E}[Y_s | Y_s \geq \underline{y}_s^*] E_s + \mathbb{E}[Y_o | Y_o \geq \underline{y}_o^*] E_o - c\theta_o^* J S_o.$$

Vacancy posting costs are subtracted from gross output as in Acemoglu and Shimer (1999), because only the remainder acts to increase welfare in the economy (see below).

Government budget. The government provides transfers g_m and g_o to unemployed prime-age and old individuals, respectively. Aggregate public expenditures per age group are therefore $G_1 = (N_1 - E_m)g_m$ and $G_2 = (N_2 - E_s - E_o)g_o$. The government collects a total tax revenue of τN. The equilibrium tax level that balances the budget is thus $\tau^* = \frac{G_1 + G_2}{N}$.

Welfare. To quantify the welfare cost of the contracting friction, I define welfare as the sum of utility within each age group,

$$\mathcal{W}_1 = E_m u(w_m^* - \tau) + (N_1 - E_m) u(b_m - \tau),$$
$$\mathcal{W}_2 = E_s u(w_s^* - \tau) + E_o u(w_o^* - \tau) + (N_2 - E_o - E_s) u(b_o - \tau),$$

and total welfare as $\mathcal{W} = \mathcal{W}_1 + \mathcal{W}_2$. Since firms earn zero expected profit, firm dividends can be neglected altogether. To convert utility levels into consumption equivalents, I compute the per capita income x that would generate the same level of welfare in the economy, i.e. $Nu(x) = \mathcal{W}$. This implies $x = u^{-1}(\mathcal{W}/N)$.

2.5 Equilibrium without the contracting friction

To quantify the welfare and employment loss that is caused by the contracting friction, I compare the equilibrium defined in Section 2.4 to the equilibrium of a counterfactual economy in which wages can be productivity-contingent. In this economy, wage contracts specify wages schedules $w_i : \mathbb{R} \to \mathbb{R}$ which can be arbitrary measurable functions of contemporaneous match productivity. I maintain the assumption that employment only occurs if both parties receive non-negative rents. Since wages can be productivity contingent, however, matches with positive joint surplus are never destroyed endogenously in equilibrium. Layoffs are therefore bilaterally efficient, and the layoff threshold of the firm becomes the reservation productivity y_i^r, implicitly defined by $W_i(w_i; y_i^r) = J_i(w_i; y_i^r) = 0$.

2.5.1 Labor market equilibrium of old job seekers

Firm and worker surplus at the production stage satisfy equations (2.2)–(2.3), except that w_o has to be replaced by $w_o(y)$. Expected firm and worker surplus at the search stage are

$$\mathbb{E}J_o^+(w_o) = \int_{y_o^r}^{\infty} J_o(w_o; y)\, dF_o(y) = \frac{\int_{y_o^r}^{\infty} y - w_o(y)\, dF_o(y)}{1 - \beta_o(1 - \phi F_o(y_o^r))},$$

$$\mathbb{E}W_o^+(w_o) = \int_{y_o^r}^{\infty} W_o(w_o; y)\, dF_o(y) = \frac{\int_{y_o^r}^{\infty} u(w_o(y) - \tau) - u(b_o - \tau) - \beta_o V_o\, dF_o(y)}{1 - \beta_o(1 - \phi F_o(y_o^r))}.$$

Since $J_o(w_o; y) \geq 0$ requires $w_o(y) \leq y + \beta_o \phi \mathbb{E}J_o^+(w_o)$, the reservation productivity y_o^r where both parties are indifferent between employment and non-employment satisfies

$$u(y_o^r + \beta_o \phi \mathbb{E}J_o^+(w_o) - \tau) - u(b_o - \tau) + \beta_o \phi \mathbb{E}W_o^+(w_o) - \beta_o V_o = 0. \tag{2.16}$$

The equilibrium on the labor market for old job seekers is characterized as in (2.1) but with the additional condition that $J_o(w_o; y) \geq 0$ for all $y \geq y_o^r$, which is the firm's layoff constraint. The first order optimality conditions can be summarized as[8]

$$w_o^\bullet(y) = \min\{\overline{w}_o^\bullet, y + \beta_o \phi \mathbb{E}J_o^+(w_o^\bullet)\} \text{ for } y \geq y_o^r, \tag{2.17}$$

$$u'(\overline{w}_o^\bullet - \tau) = \frac{1-\gamma}{\gamma} \frac{\mathbb{E}W_o^+(w_o^\bullet)}{\mathbb{E}J_o^+(w_o^\bullet)} + \frac{\beta_o \phi}{1 - \beta_o(1 - \phi F_o(y_o^r))} \Delta_o, \tag{2.18}$$

$$q(\theta_o^\bullet) \mathbb{E}J_o^+(w_o^\bullet) = c, \tag{2.19}$$

where $\Delta_o := \int_{y_o^r}^{\underline{y}_o^\bullet} u'(w_o^\bullet(y) - \tau) - u'(\overline{w}_o^\bullet - \tau)\, dF_o(y)$ and $\underline{y}_o^\bullet = \underline{y}_o(\overline{w}_o^\bullet)$ is given by (2.4). According to condition (2.17), the optimal wage schedule is piecewise linear. Provided that match productivity is sufficiently high, the worker earns a constant wage $\overline{w}_o^\bullet$ because of the preference for smooth consumption. For low enough productivity draws, however, the firm cannot afford this pay because $J_o(\overline{w}_o^\bullet, y) < 0$. In this case, the firm pays the maximum it can afford, which is the wage that grants the whole match surplus to the worker, $J_o(w_o^\bullet(y); y) = 0$. The profitability threshold, below which the firm earns no rent, is given by $\underline{y}_o^\bullet = \underline{y}_o(\overline{w}_o^\bullet)$ with $\underline{y}_o$ defined in equation (2.4). Hence with productivity-contingent wages, there are two productivity thresholds. If match productivity is below the reservation productivity, $y < y_o^r$, the match is dissolved. For $y \in [y_o^r, \underline{y}_o^\bullet]$, the match continues but the firm's layoff constraint is binding, $J_o(w_o^\bullet(y), y) = 0$. Only for productivity draws above the firm's profitability threshold, $y > \underline{y}_o^\bullet$, both firm and worker enjoy strictly positive rents. This is also visible from Figure 2.8 where the thick solid line corresponds to the wage schedule $w_o^\bullet(y)$.

Condition (2.18) determines the optimal level of the base wage $\overline{w}_o^\bullet$. The second term on the right-hand side captures that a higher base wage reduces the worker's ability to smooth consumption within a period as the firm's layoff constraint becomes binding in more states of the world (cf. Proposition 2.1). This effect, however, turns out to be quantitatively negligible, $\Delta_o \approx 0$, such that without the contracting friction the worker essentially earns a fraction γ of the joint match surplus. Remember that with the friction, the worker reduces her surplus share below γ in favor of a higher retention probability. The effect of the friction on equilibrium layoff and job-finding probabilities can also be discussed analytically.

[8] Equilibrium objects in the counterfactual economy are indicated by a dot $\bullet$ to distinguish them from the equilibrium objects with the friction that were indicated by an asterisk $*$.

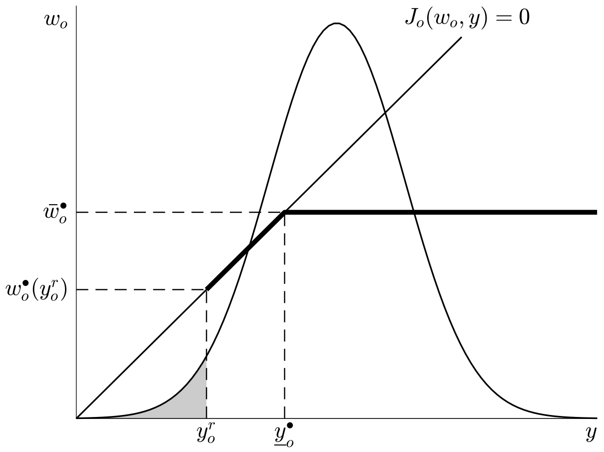


Figure 2.8. Labor market equilibrium of old job-seekers without the contracting friction.

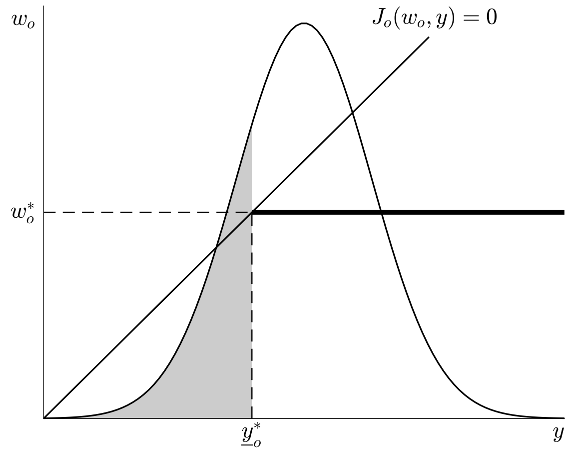


Figure 2.9. Labor market equilibrium of old job-seekers with the contracting friction.

Proposition 2.9. *Let $w_o^\bullet(y_r) < w_o^* < \overline{w}_o^\bullet$. Then the contracting friction increases both the equilibrium layoff probability, $F_o(\underline{y}_o^*) > F_o(y_o^r)$, and the equilibrium job-finding probability, $p(\theta_o^*) > p(\theta_o^\bullet)$.*

The first part of the assumption, $w_o^* < \overline{w}_o^\bullet$, holds in all conducted numerical experiments.[9] The second part, $w_o^\bullet(y_r) < w_o^*$, means that the equilibrium wage obtained under the friction lies above the reservation wage of the frictionless economy. This is a very weak assumption. If old age lasts for one period only, it is automatically satisfied since $w_o^\bullet(y_r) = b_o$ and $w_o^* > b_o$ by Proposition 2.2. In the general case, however, this condition seems necessary to ensure that the layoff probability is indeed higher with the friction.

Perhaps surprisingly, Proposition 2.9 also establishes that the contracting friction increases the equilibrium job-finding probability. In fact, if $w_o^* = \overline{w}_o^\bullet$, then the job-finding probability would be the same in both scenarios, $p(\theta_o^*) = p(\theta_o^\bullet)$. The reason is that in this case firm surplus, which fully determines hiring, is equal with both types of contracts. The argument is illustrated in Figure 2.8 and Figure 2.9. With the contracting friction, matches below the layoff threshold $\underline{y}_o^* = \underline{y}_o(w_o^*)$ are dissolved, which corresponds to the shaded area in Figure 2.9. Without the friction, layoffs only occur below the reservation productivity y_r as illustrated in Figure 2.8. Yet, the firm does not earn any surplus until the productivity exceeds $\underline{y}_o^\bullet = \underline{y}_o(\overline{w}_o^\bullet)$. Assuming $w_o^* = \overline{w}_o^\bullet$ we have that $\underline{y}_o^\bullet = \underline{y}_o^*$. Therefore, although more matches survive in absence of the friction, the firm earns zero profits on these additional matches, such that expected firm surplus is identical, $\mathbb{E}J_o^+(w_o^*) = \mathbb{E}J_o^+(w_o^\bullet)$. By the free entry conditions (2.6) and (2.19), this translates into identical labor market tightness and job-finding probability. In the likely case that the contracting friction gets workers to reduce their wage claims, $w_o^* < \overline{w}_o^\bullet$, the presence of the friction even increases expected firm profit and thus the job-finding probability as firms post

[9] This is not granted theoretically. *Ceteris paribus*, the friction decreases expected worker surplus while expected firm surplus remains unaffected. The reason is that any match that is destroyed by the friction was previously associated with zero firm surplus, $y \in [y_o^r, \underline{y}_o^\bullet)$. To restore optimal surplus sharing, the equilibrium wage increases. On the other hand, the friction implies a trade-off between wage and job security, which lowers the equilibrium wage. The latter effect seems to dominate in realistic calibrations.

more vacancies. Proposition 2.9 implies that the contracting friction increases labor turnover, while its effect on equilibrium employment is ambiguous.

2.5.2 Labor market equilibrium of prime-age job seekers

Firm and worker surplus at the production stage satisfy equations (2.10)–(2.11), except that w_i has to be replaced by $w_i(y)$ for $i \in \{m, s\}$. I only state the first order optimality conditions since the function definitions are very similar to the previous section. The optimal wage schedules $w_i^\bullet$ are again piecewise linear. For $y \geq \underline{y}_i^\bullet$ the worker receives a constant wage $\overline{w}_i^\bullet$, otherwise the worker earns the whole match surplus. The base wages $\overline{w}_m^\bullet$ and $\overline{w}_s^\bullet$ of the two wage schedules satisfy

$$u'(\overline{w}_m^\bullet - \tau) = \frac{1-\gamma}{\gamma} \frac{\mathbb{E}W_m^+(\omega_m^\bullet)}{\mathbb{E}J_m^+(\omega_m^\bullet)} + \frac{\beta_m \phi}{1 - \beta_m(1 - \phi F_m(y_m^r))} \Delta_m, \tag{2.20}$$

$$u'(\overline{w}_s^\bullet - \tau) = \mathbb{E}[u'(w_m^\bullet - \tau) | y \geq y_m^r] + \frac{\beta_o \phi}{1 - \beta_o(1 - \phi F_s(y_s^r))} \Delta_s, \tag{2.21}$$

where y_i^r is the reservation productivity of a type i worker. As in (2.18), the last term on the right-hand side of the first order equations are quantitatively negligible, such that the worker in expectation receives a share of joint surplus close to γ according to (2.20). The optimal age profile of wages is determined by condition (2.21). Since $w_m^\bullet(y_m^r) < \overline{w}_m^\bullet$ and utility is concave, $\mathbb{E}[u'(w_m^\bullet - \tau) | y \geq y_m^r] > u'(\overline{w}_m^\bullet - \tau)$. Condition (2.21) therefore implies $\overline{w}_s^\bullet < \overline{w}_m^\bullet$, such that the optimal wage profile is decreasing in age also in absence of the contracting friction. The underlying intuition is that a high senior wage $\overline{w}_s^\bullet$ reduces expected firm surplus at the senior stage, which decreases the firm's profitability threshold $\underline{y}_m(\overline{w}_m^\bullet)$ in prime age. *Ceteris paribus*, this reduces the states of the world in which a prime-age worker can enjoy smooth income. The intuition is therefore similar to that of (2.14), with the difference that now the marginal cost of a higher senior wage arises from less income smoothing within a period, without affecting the layoff probability. Whereas with the contracting friction, a higher senior wage leads to a higher layoff probability. The average wage decrease in old working age is therefore likely to be more pronounced in presence of the friction.

2.5.3 Economic aggregates

With productivity-contingent contracts, all demographic and aggregate economic variables are defined as in Section 2.4.3, replacing θ_i^* with $\theta_i^\bullet$ and $\underline{y}_i^*$ with y_i^r for $i \in \{m, s, o\}$. Aggregate welfare becomes

$$\begin{aligned} \mathcal{W}_1 &= E_m \overline{\mathcal{W}}_m + (N_1 - E_m) u(b_m - \tau), \\ \mathcal{W}_2 &= E_s \overline{\mathcal{W}}_s + E_o \overline{\mathcal{W}}_o + (N_2 - E_o - E_s) u(b_o - \tau), \end{aligned}$$

where $\overline{\mathcal{W}}_i = \frac{\int_{y_i^r}^\infty u(w_i^\bullet(y) - \tau)\, dF_i(y)}{1 - F_i(y_i^r)}$ is the average period utility of a type i worker.

parameter	value	parameter	value	parameter	value
μ_m, μ_s	1.0000	α_i	1.0000	π_m	0.3333
μ_o	0.9000	ϕ	0.1167	π_o	0.1000
s_m, s_s	0.2601	β	0.9709	γ	0.5000
s_o	0.2341	κ	3.0000		

Table 2.1. Parameters set directly

2.6 Numerical illustration and policy implications

To assess the quantitative importance of the contracting friction, I solve the model outlined in Section 2.4 numerically and compare it to the counterfactual economy without the friction described in Section 2.5. Additionally, I investigate how the presence of the friction affects the effectiveness of an early retirement reform. Finally, I compare several labor market policies and discuss their potential to reduce the aggregate costs caused by the contracting friction.

2.6.1 Calibration

A model period corresponds to a year. The future is discounted at an annual discount rate of 3%, which implies $\beta = 1/1.03 = 0.971$. Prime working age lasts from age 25 to 54, while old working age lasts from age 55 to 64. Therefore, the aging probabilities are set to $\pi_m = 1/30$ and $\pi_o = 1/10$. Productivity follows a normal distribution with mean μ_i and standard deviation s_i. In the baseline, $\alpha_i = 1$ for all worker types, such that the distributions are symmetric. The mean is normalized to $\mu_m = \mu_s = 1$ for prime-age and senior workers. For workers hired during old age, I assume a lower mean productivity of $\mu_o = 0.9$. This captures that learning and the adaption to new work requirements becomes more difficult with age, while workers can maintain high productivity in tasks that they are experienced in (Skirbekk, 2004, 2008). The standard deviations s_i are chosen such that for every worker type, productivity in the 90th percentile is twice as high as in the 10th percentile, which implies $s_m = s_s = 0.2601$ and $s_o = 0.2341$. As in Menzio et al. (2016) a productivity draw lasts for 8.5 years on average, such that $\phi = 0.1167$.[10]

Instantaneous utility exhibits constant absolute risk version, $u(w) = (1 - e^{-\kappa w})/\kappa$. This specification simplifies the analysis because it eliminates wealth effects. Additionally, it renders the labor market equilibria independent of the lump sum tax level. I set $\kappa = 3$, which in equilibrium implies rates of relative risk aversion between 2 and 3. The matching function is Cobb-Douglas $m(u, v) = Au^{\gamma}v^{1-\gamma}$ with elasticity $\gamma = 0.5$ (Petrongolo and Pissarides, 2001).

The remaining model parameters are calibrated to reflect important characteristics of the Austrian labor market in the year 2004, before a series of pension reforms became effective. I regard this as a good starting point to study the effect of a pension reform on the importance of the contracting friction. Austria runs a large scale publicly funded defined benefits pension

[10] Menzio et al. (2016) report a percentile ratio of three, but assume that information is perfect. Mas and Moretti (2009) report a ratio of 0.3 for supermarket cashiers, who perform a very standardized task. I choose an intermediate value that seems consistent with the data.

system, representative for continental Europe. In comparison with other countries, however, it is exceptionally generous with a net pension replacement rate well above 90% (OECD, 2006). Furthermore, until 2000, the age threshold for early retirement was 60 years for men, with a permanent reduction in pension benefits of only 2% for every year between the age of first benefit claiming and the normal retirement age of 65. Access to early retirement required 35 contribution years. To cope with the increasing demographic pressure, access to and discounts for early retirement were gradually reformed in 2000 and 2003 (see Section 2.6.3). Since there is a break in the Austrian labor market time series after 2003 and many 55 year olds could still retire according the old regulations in 2004, the targeted labor market characteristics refer to the year 2004, while the modeling of early retirement reflects the situation before 2000.

To proxy that a minimum number of contribution years was necessary to have access to early retirement benefits, I assume in the numerical model that workers who were employed at the time they entered old working age have access to a transfer g_o, while all other individuals can only collect unemployment benefits, $g_m < g_o$. The unemployment benefit g_m is calibrated to achieve a net replacement rate of 0.531. In Austria, unemployed individuals collect *Arbeitslosengeld* equal to 55% of their previous net wage during the initial months of unemployment. Thereafter, they can receive *Notstandshilfe* that grants up to 92% of the *Arbeitslosengeld* and therefore 50.6% of their last wage earnings. Weighting these figures with the stock of benefits recipients in both systems reported by Statistik Austria (2018) yields an average net replacement rate of 53.1% of the unemployment insurance (UI) system.

Workers eligible to early retirement benefits receive a transfer g_o. The net replacement rate of the Austrian pension system at normal retirement age is 93.2% (OECD, 2006). Assuming that the age of first benefit claiming is uniformly distributed in age 60–64, the average pension deduction is 6%. Up to age 60, only unemployment benefits can be collected, which replace 53.1% of the previous net wage, see above. The average unemployed worker in the age group 55 to 64 therefore faces a replacement rate of $\frac{0.531+0.932\cdot 0.94}{2} = 0.704$. This serves as calibration target for g_o.

The calibration targets that identify the parameters (A, σ, z_m, z_o, c) are taken from the OECD database (OECD, 2018) and refer to Austrian males in 2004 unless otherwise indicated. The matching technology A governs the job-finding probability and is identified by the cross-sectional share of prime-age unemployed with duration less than a year, $u_m^0 = 0.6383$. The parameters z_m, z_o, and σ all affect the layoff probability. The exogenous separation rate σ is pinned down by the cross-sectional share of matches with tenure less than a year, $e_m^0 = 0.1127$. This works because endogenous layoffs happen primarily at the beginning of a match (after the initial draw on average 8.5 years pass until the next productivity level realizes), while the probability for an exogenous layoff is independent of tenure. The valuations for leisure z_m and z_o affect layoff rates through the equilibrium wage and are used to target the empirical age profile of employment $(e_1, e_2) = (0.8807, 0.3662)$. The vacancy posting cost c targets an average labor market tightness of 0.714 in the economy. This figure relates the number of job vacancies reported by Eurostat (2018) to the number of unemployed.

parameter	value	calibration target
g_m	0.5180	UI replacement rate $g_m/w_m^* = 0.531$
g_o	0.6730	average of UI replacement rate and pension replacement rate with early retirement discounts $g_o/w_2^* = 0.704$
z_m	0.1788	employment rate 25 to 54 years $e_1 = 0.8807$
z_o	0.2553	employment rate 55 to 64 years $e_2 = 0.3662$
σ	0.0236	share of employed with tenure < 1 year, $e_m^0 = 0.093$
A	0.7406	share of unemployed with duration < 1 year, $u_m^0 = 0.6383$
c	0.9821	labor market tightness $\theta_m^* = 0.714$
δ	0.0535	potential labor force participation rate $lf_2 = 0.675$

Table 2.2. Calibrated parameter values and calibration targets

Finally, I construct a measure of potential labor force participation to pin down the inactivity shock δ. In the model, the labor force in old working age, $lf_2 N_2$, consists of all individuals that did not experience the δ shock. This shock stands in for health shocks or personal reasons to retire. The model labor force therefore encompasses all persons who are capable of working. Empirically reported measures of the labor force, by contrast, also subtract workers that are in principle able to work but do not participate in the labor market due to policy-related incentives. In a comparison of EU countries, with only 38.5% Austria had the lowest labor force participation rate in the age group 55 to 64 in 2004. By contrast, labor force participation was 92% in the age group 25 to 54, close to the EU average. While Ireland and the UK had similar labor market attachment during prime age, old-age labor force participation in these countries was much higher at 66.8% and 68.1%, respectively. I therefore assume that the maximum labor force participation rate that could have been attained in the Austrian economy by implementing adequate government policies was 67.5%. This corresponds to an exogenous retirement probability of $\delta = 0.0535$.[11]

The calibrated model parameters are given in Table 2.2. The ratio of unemployment income to mean productivity is $b_m = g_m + z_m = 0.7052$ for prime-age workers, which is close to the calibration of Costain and Reiter (2008) [0.745] for the US. By contrast, old unemployed with access to early retirement benefits can enjoy $b_o = g_o + z_o = 0.9204$, which is close to the small surplus calibration of Hagedorn and Manovskii (2008) [0.955].

2.6.2 Equilibrium

Panel (a) of Table 2.3 shows the equilibrium of the calibrated model. In line with Proposition 2.5, the optimal wage contract of prime-age job seekers is decreasing in age, $w_s^* < w_m^*$. However, the wage drop in old age is only 2.6%. Since senior workers have access to generous early retirement benefits, the utility loss from a layoff is small. The incentive to substitute between job security and wage income is therefore low, and the age-wage profile is almost flat. Part of

[11] Only the Scandinavian countries had even higher old-age participation rates in excess of 70%. This, however, is likely to be due to cultural norms.

individual variables	prime-age job seekers		old job seekers	
	m	s	n	o
wage w_i^*	0.975	0.950	0.888	1.000
layoff probability $F_i(\underline{y}_i^*)$	0.276	0.344	0.411	0.634
job-finding probability $p(\theta_i^*)$	0.626	———	0.256	0.123
per capita variables	prime age	old age	total	
job-finding rate	0.626	0.151	0.455	
endog. layoff rate	0.060	0.156	0.073	
employment rate	0.881	0.366	0.752	
gov. expenditures	0.062	0.415	0.150	
output	0.877	0.403	0.758	
welfare in cons. eq.	0.779	0.765	0.775	

individual variables	prime-age job seekers		old job seekers	
	m	s	n	o
base wage $\overline{w}_i^\bullet$	1.009	0.988	0.915	1.022
average wage $\mathbb{E}[w_i^\bullet \| y \geq y_i^r]$	0.991	0.983	0.897	1.004
layoff probability $F_i(y_i^r)$	0.161	0.313	0.261	0.504
job-finding probability $p(\theta_i^\bullet)$	0.498	———	0.217	0.105
per capita variables	prime age	old age	total	
job-finding rate	0.498	0.127	0.366	
endog. layoff rate	0.031	0.122	0.044	
employment rate	0.892	0.393	0.767	
gov. expenditures	0.056	0.398	0.141	
output	0.895	0.430	0.779	
welfare in cons. eq.	0.802	0.786	0.798	

Table 2.3. Equilibrium for the baseline economy

the old job seekers (type o in Table 2.3) also have access to early retirement benefits. These are only willing to accept very high-paying jobs, which results in a very low job-finding probability. By contrast, old job seekers who can only claim unemployment benefits (type n in Table 2.3) have a much lower wage demand, are fired less often and hired more frequently. Since most workers in the model population can enjoy very high outside options, the endogenous layoff rate is strongly increasing in age in Table 2.3(a), while the job-finding rate is decreasing. Government expenditures are 20% of output, the largest part thereof accrues to early retirement benefits.

To assess the quantitative effect of the contracting friction, I rerun the model allowing for state-contingent contracts, taking the parameterization of Table 2.2 as given. The corresponding equilibrium is given in panel (b) of Table 2.3. Comparing the aggregate employment rates, the friction depresses prime-age employment by 1.1 percentage points, while old-age employment is 2.7 percentage points lower. The reason for the smaller loss in prime-age employment is that although the layoff rate of prime-age workers is elevated by 2.9 percentage points under the friction, the job-finding rate is even 12.8 percentage points higher. The latter effect is due to lower equilibrium wages which stem from the worker's incentive to give up wage income for job security in presence of the friction (compare Proposition 2.9). Although the friction has the same qualitative effects on elderly individuals, they experience a much smaller increase in their job-finding rate under the friction (2.4pp) and a larger increase in their layoff rate (3.4pp). This is due to their shorter expected employment horizon. The calibrated model reveals that the cost of the contracting friction in terms of forgone output and welfare can be substantial. Comparing panels (a) and (b) of Table 2.3 reveals that the friction reduces aggregate welfare by 2.9% in consumption equivalents, while output is depressed by 2.7%.

In contrast to Table 2.3, Austrian earnings data shows little evidence for decreasing wages of high-tenured workers, compare Figure 2.C.3. One reason for this deviation may be that individuals do not act as farsightedly as in the model. To explore this possibility, Section 2.6.5 assumes that prime-age job seekers do not take into account that the shape of the wage contract affects their layoff probability in late working age. Compared to the baseline results, this increases the wage of senior workers and their layoff probability. The aggregate costs of the contracting friction are then even higher than predicted by Table 2.3. Human capital accumulation could also contribute to an increasing wage-tenure profile. In the model this can easily be captured by setting $\mu_s > \mu_m$. However, returns to tenure and experience after age 55 are likely to be small. A factor neglected by the model is the use of delayed compensation as motivational device. If work effort is unobservable, the parties can agree on a high wage in old age to reduce the incentive to shirk during prime age. Yet, if back-loading is beneficial, wages will already increase in tenure *during* prime age. Since the present model assumes a constant wage at each stage of the life-cycle, it is not suited to investigate this channel. This is left for future research.

2.6.3 The effect of an early retirement (ER) reform

In response to increasing longevity and the longer lifetime that individuals spend in retirement, most European countries have restricted access to early retirement and reduced benefit generosity to improve fiscal sustainability of the public pension system. For instance, the reforms implemented in Austria after 2000 increased the age threshold for early retirement to age 62, but this is conditional on more than 40 contribution years and a permanent pension deduction of 5.1 percent for every year of retirement before age 65 (OECD, 2005; Knell et al., 2006).

In the context of the model, I investigate the labor market effects of abolishing early retirement (ER) completely. I repeat the above analysis with the parameters of Table 2.2 but set $g_o = g_m = 0.518$, such that every old unemployed only receives the unemployment benefit. Since the UI replacement rate is much lower than the replacement rate of early retirement benefits, this is expected to boost employment of the elderly. The lower outside option makes layoffs more costly in old age, which leads to lower wages and higher retention probabilities. As evident from Table 2.4(a), the optimal wage contract of prime-age job seekers now features a 9.7% wage decrease in old age. Old job seekers after the ER reform only receive benefits from the UI system. They behave in the same way as the type n individuals in the pre-reform economy of

individual variables	prime-age job seekers		old job seekers
	m	s	n, o
wage w_i^*	0.978	0.883	0.888
layoff probability $F_i(\underline{y}_i^*)$	0.268	0.230	0.411
job-finding probability $p(\theta_i^*)$	0.641	———	0.256
per capita variables	**prime age**	**old age**	**total**
job-finding rate	0.641	0.256	0.535
endog. layoff rate	0.058	0.099	0.064
employment rate	0.888	0.484	0.787
gov. expenditures	0.058	0.267	0.110
output	0.881	0.507	0.787
welfare in cons. eq.	0.823	0.707	0.790

individual variables	prime-age job seekers		old job seekers
	m	s	n, o
base wage $\overline{w}_i^\bullet$	1.006	0.986	0.915
average wage $\mathbb{E}[w_i^\bullet \vert y \geq y_i^r]$	0.988	0.954	0.897
layoff probability $F_i(y_i^r)$	0.155	0.141	0.261
job-finding probability $p(\theta_i^\bullet)$	0.507	———	0.217
per capita variables	**prime age**	**old age**	**total**
job-finding rate	0.507	0.217	0.438
endog. layoff rate	0.030	0.053	0.034
employment rate	0.897	0.532	0.806
gov. expenditures	0.053	0.242	0.101
output	0.897	0.549	0.810
welfare in cons. eq.	0.843	0.745	0.815

Table 2.4. Equilibrium after the early retirement (ER) reform

Table 2.3(a), since this group of unemployed did not have access to ER benefits anyway.

Comparing Table 2.4(a) to Table 2.3(a) reveals that the reform boosts old-age employment by 11.8 percentage points. This is both due to fewer layoffs (–5.7pp) and more hiring (+10.5pp). The higher retention rate in old age also slightly increases prime-age employment by 0.7 percentage points. Government expenditures decrease by more than a quarter. This is due to fewer unemployed individuals and lower spending per unemployed. The early retirement reform increases aggregate output by 3.8% and aggregate welfare by 1.9%.

Despite the substantial positive economic effects of the reform, its effectiveness is reduced by the presence of the contracting friction. Comparing Table 2.4(b) to Table 2.3(b) reveals that without the friction, the reform would have increased the old-age employment rate by even 13.9 percentage points. Hence 2.1 percentage points and therefore 15% of the potential gain in old-age employment cannot unfold because of the market failure. The same applies to aggregate output and welfare, where 4% and 10% of the potential improvement is foregone, respectively.

worker type	before reform	after reform	difference
m	0.115	0.113	−0.002
s	0.031	0.089	+0.058
o	0.130	0.150	+0.020

Table 2.5. Difference in layoff probability, $F_i(\underline{y}_i^*) - F_i(y_i^r)$

As a result, the aggregate costs of the friction are higher after the early retirement reform than before.

The reason for this increasing gap is that layoff rates respond very differently to a reduction in outside options in the two contractual frameworks studied. With productivity-contingent wages, the layoff probability of older workers is determined by the reservation productivity defined in (2.16). The numerical analysis reveals that a reduction in unemployment income b_o triggers almost a one-for-one decrease in the reservation productivity, $\frac{\Delta y_o^r}{\Delta g_o} = 0.98$. With flat wages, on the other hand, layoffs are governed by the layoff threshold defined by equation (2.4). Since worker's unemployment income b_o does not show up explicitly in this equation, the only link between the equilibrium layoff probability $F_o(\underline{y}_o^*)$ and b_o comes through the equilibrium wage w_o^*, compare Section 2.4.1. Since the wage response to a change in unemployment income is less than proportional, $\frac{\Delta w_o^*}{\Delta g_o} = 0.72$, the layoff threshold does not decrease as much as the reservation productivity. As a result, the reform increases the gap in layoff probabilities, $F_o(\underline{y}_o^*) - F_o(y_o^r)$, by 2 percentage points from 0.13 to 0.15 in the last row of Table 2.5. Since with productivity-contingent wages layoffs are bilaterally efficient, these additional layoffs are bilaterally inefficient.

The gap in layoff probabilities increases even more senior workers. The second row of Table 2.5 reveals that without the contracting friction, their layoff probability would have decreased by 5.8 percentage points more in response to the ER reform. The reason is that intertemporal consumption smoothing implies a wage elasticity of only $\frac{\Delta w_s^*}{\Delta g_o} = 0.43$. While before the reform only one in ten layoffs of senior workers was bilaterally inefficient, this figure increases to four in ten after the reform. By contrast, the efficiency of layoffs of prime-age workers is hardly affected by lower outside options in old age.

2.6.4 Complementary labor market reforms

According to the above analysis, the early retirement reform increases employment, output, and welfare in the economy, but at the same time the detrimental effects of the friction gain in importance. The employment rate of the elderly remains 2.1 percentage points under its potential. At the same time, the welfare loss caused by the friction has increased to 3.1% and the loss in output to 2.8%. Labor market policies that reduce excessive layoffs may be beneficial. In this section I assess the potential of different labor market policies implemented after the ER reform to achieve the same labor market allocation (E_m, E_s, E_o) as in the frictionless economy without policy intervention (panel (b) of Table 2.4).[12] The goal of this exercise is *not* to design

[12]Here and in the following *frictionless* refers to the absence of the contracting friction. The search frictions are always present.

an optimal policy, but to assess the effort necessary to undo the employment distortions that are caused by the friction. I consider training programs, wage cost subsidies, layoff taxes, as well as severance pay. To compare the potentials and caveats of each of these labor market programs, the analysis takes the post-reform economy of Table 2.4 as a reference and discusses the effect of one additional labor market related policy measure. Since the equilibrium employment allocation is $(E_m, E_s, E_o) = (26.63, 3.50, 1.44)$ under the friction and $(E_m, E_s, E_o) = (26.90, 4.13, 1.19)$ without the friction, the labor market measures particularly aim at increasing retention rates of senior workers.[13]

Training

Consider first a reform that increases match productivity. While I focus on a training program, especially for elderly workers similar productivity-enhancing effects could be achieved by establishing a more age-friendly work environment, employee health programs, or organizing work in teams (OECD, 2006; Göbel and Zwick, 2013; Börsch-Supan and Weiss, 2016). The employment and welfare gains of such programs hinge on the size of the associated productivity gains as well as on setup and participation costs. To discipline the model, I use the cost-benefit link that has been estimated for the German WeGebAU program. This program provides government-sponsored training to low-skilled workers and to employed workers who are over 45 years old. Dauth and Toomet (2016) estimate causal effects and find that for workers above age 55, participation in the program increases the probability of remaining in paid employment by 5 percentage points in the two-year period following treatment. Whereas the probability only increased by 1.5 percentage points in the age group 45 to 55. Furthermore, the authors report that the average cost per participant was 1,720 euros annually, which amounts to 5.9% of annual average wage income in Germany.

To design a training program that implements the frictionless employment allocation, I alter the means of the productivity distributions (μ_m, μ_s, μ_o) and assume that the costs C_i necessary to reduce the layoff probability of one participant by one percentage point is in line with Dauth and Toomet (2016). Since the annual average wage after the pension reform is 0.964 in the model, I assume that $\frac{\Delta C_m}{\Delta F_m(\underline{y}_m^*)} = \frac{0.059 \cdot 0.964}{0.015} = 3.79$ and $\frac{\Delta C_s}{\Delta F_s(\underline{y}_s^*)} = \frac{\Delta C_o}{\Delta F_o(\underline{y}_o^*)} = \frac{0.059 \cdot 0.964}{0.05} = 1.14$. The productivity increase is considered as immediate, transferable across jobs, and valid until the worker leaves the age group in which training was provided. Hence training costs accrue twice for every worker, once in prime age and once in old age.[14]

Table 2.6(a) shows the equilibrium after implementation of the training program. To attain the frictionless employment allocation, the program should increase the means of the productivity distributions by $(\Delta\mu_m^*, \Delta\mu_s^*, \Delta\mu_o^*) = (0.007, 0.086, 0.021)$. Hence training should mainly focus on long-tenured old workers, such that their average productivity increases by 8.6%. Less

[13]In the economy without the contracting friction there are still several imperfections that a utilitarian social planner would address. Designing an optimal policy is therefore beyond the scope of this essay. An effort in this regard is taken by the second essay contained in this thesis.

[14]The transferability of skills reflects the nature of the WeGebAU program, which provides external courses to improve general human capital, see Dauth and Toomet (2016) for details.

individual variables	prime-age job seekers		old job seekers
	m	s	n, o
wage w_i^*	0.988	0.905	0.896
layoff probability $F_i(\underline{y}_i^*)$	0.256	0.141	0.384
job-finding probability $p(\theta_i^*)$	0.663	———	0.276
per capita variables	prime age	old age	total
job-finding rate	0.663	0.276	0.571
endog. layoff rate	0.054	0.067	0.056
employment rate	0.897	0.532	0.806
gov. expenditures	0.055	0.256	0.105
output	0.892	0.586	0.815
welfare in cons. eq.	0.840	0.726	0.807

individual variables	prime-age job seekers		old job seekers
	m	s	n, o
wage w_i^*	0.986	0.899	0.893
layoff probability $F_i(\underline{y}_i^*)$	0.254	0.141	0.379
job-finding probability $p(\theta_i^*)$	0.660	———	0.273
per capita variables	prime age	old age	total
job-finding rate	0.660	0.273	0.567
endog. layoff rate	0.054	0.066	0.056
employment rate	0.897	0.532	0.806
gov. expenditures	0.055	0.256	0.106
output	0.892	0.574	0.813
welfare in cons. eq.	0.838	0.724	0.806

Table 2.6. Equilibrium after the ER reform and implementation of a training program

effort is required for newly hired old workers and prime-age workers. In steady state, every year 6% of the workforce are enrolled in the training program. With the cost-benefit link estimated by Dauth and Toomet (2016), the annual training costs amount to 0.4% of aggregate output. In total, the program reduces government spending, since the program costs are more than compensated by lower expenditures on unemployment benefits. Moreover, the welfare cost of the contracting friction decreases from 3.1% to 1%, while the aggregate output even exceeds the level of the counterfactual frictionless economy where no policy is implemented.

While this experiment assumed that the productivity of every worker increases uniformly, it is likely that training has a larger effect on the productivity of low productive workers and a smaller effect on workers in the upper tail of the distribution. As evident from Figure 2.2, asymmetric returns to training can be captured by an increase in α_i, which at the same time increases the mean and lowers the variance of the distribution. I therefore repeat the above exercise, but keep μ_i at their baseline levels and instead alter α_i. The frictionless employment allocation is attained for $(\Delta\alpha_m^*, \Delta\alpha_s^*, \Delta\alpha_o^*) = (0.038, 0.337, 0.105)$. Table 2.6(b) shows that while

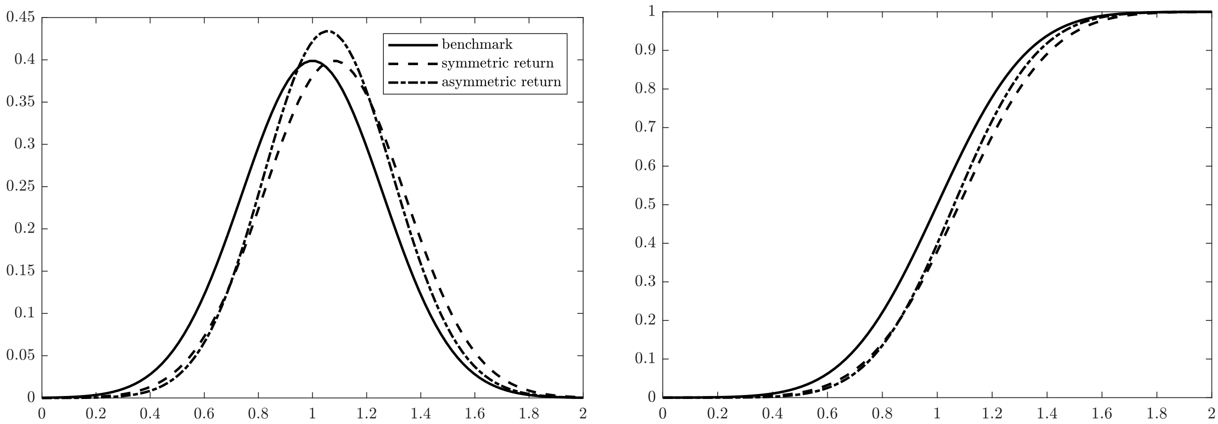


Figure 2.10. Density and distribution function of Y_s in the baseline ($\mu_s = \alpha_s = 1$), after the training program with symmetric returns ($\mu_s = 1.086, \alpha_s = 1$), and the program with asymmetric returns ($\mu_s = 1, \alpha_s = 1.337$).

wages are lower with asymmetric returns, the macroeconomic effects of the two scenarios are almost identical.

Figure 2.10 illustrates how the two training scenarios affect the productivity distribution of senior workers. With asymmetric returns, the productivity increase at the lower tail of the distribution hardly differs from the scenario with symmetric returns, while the upper tail of the distribution is close to the baseline calibration. Since it is primarily the lower tail of the distribution that determines employment levels, the effect of training on high productive workers hardly affects economic aggregates. What is key for the success of the program is that it boosts the productivity of low productive elderly workers. To increase cost-efficiency, government sponsored training programs should therefore target elderly workers with low productivity. This is corroborated by the observation of Staubli and Zweimüller (2013) that workers with low lifetime earnings (and therefore low average productivity) and poor health are particularly prone to end up unemployed if early retirement pathways are closed.

Wage cost subsidies

Layoffs can also be reduced by providing wage cost subsidies to firms. I assume that firms receive a transfer S_i from the government for every employed type i worker. The worker continues to earn w_i but only costs the employer $w_i - S_i$. The lower labor costs decrease the layoff threshold of the firm which is likely to increase equilibrium employment. The effect of the subsidy on layoff thresholds, firm surplus, and equilibrium conditions can be seen from Section 2.C.[15] The frictionless allocation of employment is achieved for the subsidy bundle $(S_m^*, S_s^*, S_o^*) = (0.007, 0.086, 0.021)$. Restricting access to early retirement should therefore be accompanied by wage cost subsidies for firms that employ senior and older workers. The government should reduce wage costs of long-tenured workers by about 10% and wages of newly hired old workers

[15]Because of surplus sharing, it is irrelevant whether the subsidy is paid to the firm (to decrease labor cost) or to the worker (to increase labor income). If w_i^* is the optimal wage in the first scenario, then $w_i^* - S_i$ is the optimal wage in the second scenario. Except for equilibrium wages, the equilibria are identical.

individual variables	prime-age job seekers		old job seekers
	m	s	n, o
wage w_i^*	0.988	0.905	0.896
layoff probability $F_i(\underline{y}_i^*)$	0.256	0.141	0.384
job-finding probability $p(\theta_i^*)$	0.663	——	0.276
per capita variables	**prime age**	**old age**	**total**
job-finding rate	0.663	0.276	0.571
endog. layoff rate	0.054	0.067	0.056
employment rate	0.897	0.532	0.806
gov. expenditures	0.060	0.280	0.115
output	0.885	0.847	0.801
welfare in cons. eq.	0.830	0.717	0.798

Table 2.7. Equilibrium after the ER reform and introduction of wage cost subsidies

by about 2.3%. The resulting increase in old-age firm surplus makes a wage cost subsidy for prime-age workers almost unnecessary.

Comparing Table 2.7 to Table 2.6(a) reveals that the labor market equilibria are identical, and only government expenditures, output, and welfare differ. Conceptually, reducing the cost of labor by x units has the same effect on firm profit as making a worker x units more productive. Therefore, the policy effects on firm surplus, layoff probabilities, and employment coincide, and $S_i^* = \Delta\mu_i^*$. Nevertheless, the macroeconomic effects of the two policies differ substantially. With training, the output loss caused by the friction is more than undone, while the subsidy is only able to close half of the gap. The wage subsidy also leads to smaller welfare gains because the equilibrium tax level is higher. This is because the subsidy program is much more expensive than the comparable training program. While the costs of the latter equal 0.4% of total output, the subsidy program costs 1.8% of output. To keep the budget balanced, a 14% higher tax level τ^* is necessary.

The low cost-effectiveness of wage subsidy programs is widely considered to be a large caveat (Boockmann, 2015). However, the calibrated model shows that wage subsidies are much cheaper than the high early retirement benefits that were in place initially. Comparing Table 2.7 to Table 2.3(a) shows that government expenditures are almost 24% lower. While the replacement rate for individuals with access to the early retirement scheme is 70% in the baseline calibration, the subsidy for old and senior workers only replaces 8% of wage income. At the same time, the number of benefit recipients is similar. While 49% of the old population were living on early retirement benefits initially, the wage subsidy in Table 2.7 is paid to 53% of the older population.[16]

[16] The model generates a cost-benefit link of the wage subsidy that is empirically plausible. Albanese and Cockx (2018) evaluate a wage subsidy program in Belgium that covers all workers above age 58 and amounts to a reduction of 4% of median wage cost. For employees who are at high risk of leaving to early retirement, they find a causal effect of a 2.2 percentage points higher short-run employment rate. In the model, the subsidy on average amounts to 8% of wage income and leads to a 4.8 percentage points higher old-age employment rate in the long-run.

individual variables	prime-age job seekers		old job seekers
	m	s	n, o
wage w_i^*	0.970	0.892	0.866
layoff probability $F_i(\underline{y}_i^*)$	0.218	0.141	0.232
job-finding probability $p(\theta_i^*)$	0.600	———	0.205
per capita variables	**prime age**	**old age**	**total**
job-finding rate	0.600	0.205	0.506
endog. layoff rate	0.045	0.050	0.045
employment rate	0.897	0.532	0.806
gov. expenditures	0.044	0.232	0.091
output	0.893	0.549	0.807
welfare in cons. eq.	0.840	0.733	0.810

Table 2.8. Equilibrium after the ER reform and introduction of layoff taxes

Layoff taxes

With a layoff tax, the firm has to pay a fine T_i to the government for displacing a type i worker. I assume that the penalty only accrues to endogenous separations and that firm owners have deep pockets that allow them to pay the penalty even if the match does not become productive at all. The employment allocation of the frictionless economy can be implemented with a tax bundle $(T_m^*, T_s^*, T_o^*) = (0.230, 0.378, 0.367)$. The equilibrium is summarized in Table 2.8. The layoff tax is increasing in age since the employment loss caused by the friction is highest for elderly workers. The tax applicable to layoffs of senior workers corresponds to 5 monthly wages.

The reported value of T_o^* should be interpreted with caution. Although taxing layoffs of workers who were hired during old age decreases their layoff probability, firms at the same time post fewer vacancies, anticipating higher separation costs. This prediction is in line with Behaghel et al. (2008), who report that hiring rates of over 50 year olds were oppressed substantially by a layoff tax in France. The calibrated model reveals that whether a layoff tax levied on workers hired during old age can have a positive net effect on employment crucially depends on the response of the equilibrium wage w_o^*. If the wage does not sufficiently decrease when the layoff tax is introduced, the tax destroys employment of type o workers instead of promoting it. Therefore, it might be recommendable to exempt newly hired old workers from layoff taxes and instead use a wage subsidy or a training program to promote their employment. In fact, combining layoff taxes $(T_m, T_s) = (0.249, 0.401)$ with a training program $\Delta\mu_o = 0.021$ also implements the frictionless employment allocation and is slightly superior in terms of welfare. Compared to the post-reform economy of Table 2.4, this policy bundle reduces the welfare cost of the friction from 3.1% to 0.5%, while foregone output reduces from 2.8% to 0.4%.

In general, using layoff taxes to correct the employment distortions gives rise to much higher aggregate welfare than wage subsidies and to slightly higher welfare than training programs. The reason is that layoff taxes do not require additional government spending but instead generate revenue. This lowers the equilibrium tax rate and uniformly increases utility in the economy.

individual variables	prime-age job seekers		old job seekers
	m	s	n, o
wage w_i^*	0.966	0.926	0.865
layoff probability $F_i(\underline{y}_i^*)$	0.145	0.141	0.213
job-finding probability $p(\theta_i^*)$	0.493	——	0.198
per capita variables	prime age	old age	total
job-finding rate	0.493	0.198	0.422
endog. layoff rate	0.027	0.048	0.031
employment rate	0.897	0.532	0.806
gov. expenditures	0.053	0.242	0.101
output	0.897	0.549	0.810
welfare in cons. eq.	0.842	0.743	0.814

Table 2.9. Equilibrium after the ER reform and introduction of severance pay

Severance pay

With severance pay, the fine (now denoted by P_i) is not paid to the government but directly to the displaced worker. The severance pay schedule that removes the employment distortions in the post-reform economy is $(P_m^*, P_s^*, P_o^*) = (0.723, 0.553, 0.418)$. As evident from Section 2.C, severance pay affects firm surplus and layoff thresholds in the same way as layoff taxes. For the worker, by contrast, severance pay acts like an increase in the outside option as layoffs become less painful. As a result, wage levels are higher with severance pay than with a layoff tax of the same size. A larger intervention is therefore necessary to reduce the layoff probability by a given amount, which implies $P_i^* > T_i^*$.

Interestingly, the wage increase for prime-age workers is so large that introducing severance pay may even reduce prime-age employment, compare the upper-left panel of Figure 2.C.1. For sufficiently low levels, the *insurance role* of severance pay seems to dominate its *penalty role* (Alvarez and Veracierto, 2001). Despite this non-monotonicity in employment, per-capita welfare is monotonically increasing. This is because displaced workers enjoy an income of b_m+P_m in their first period of unemployment instead of b_m. In the cross-section, this implies a more balanced consumption allocation compared to layoff taxes, which explains the higher welfare level in Table 2.9 compared to all previously considered labor market policies.[17]

Boeri et al. (2017) study the same contracting friction and demonstrate that severance pay can at the same time remove the distortions in the job-finding probability and in the layoff probability. This neat property does not hold in the present model because utility is not perfectly transferable between workers and firms due to risk aversion. Comparing Table 2.9 to Table 2.4(b) reveals that while severance pay can restore the equilibrium employment levels, the labor market

[17] The non-monotonicity in employment disappears if firms are granted a probation period during which a worker can be displaced at no cost. Although this dampens the negative effects of severance pay on hiring, it also reduces the effect on layoffs. Figure 2.C.2 reveals that with a probation period higher levels of severance pay are required to attain the desired employment levels. Furthermore, aggregate welfare is lower due to higher lump sum taxes.

is more rigid compared to the frictionless economy due to fewer firing and fewer hiring. Another implication of risk aversion is that workers always strictly prefer work over a layoff with severance pay. This is in contrast to Boeri et al. (2017) were workers are risk neutral and the optimal level of severance pay is such that apart from the first period of an employment spell, workers are always indifferent between work and being laid off with severance pay.

Similar to the layoff tax, the net employment effects of severance pay on old job seekers crucially depend on the equilibrium response of wages. A combination of severance pay and training might be a more robust policy and also proofs superior in terms of welfare. The bundle $(P_m, P_s, \Delta\mu_o) = (0.726, 0.558, 0.021)$ attains the highest welfare level of all labor market policies considered. The welfare loss relative to the counterfactual economy is only 0.1%.

It should be noted that in practice also other considerations may lead countries to implement a certain level of severance pay. The important message of the model is that in response to an early retirement reform, particularly the level of severance pay for long-tenured old workers should be increased. Before the reform, a bundle $(P_m, P_s, \Delta\mu_o) = (0.725, 0.155, 0.020)$ can remove the employment distortions of the friction. The early retirement reform therefore particularly increases the intervention that is necessary to prevent excess layoffs of senior workers.

2.6.5 Bounded rationality

In Tables 2.3 and 2.4 the labor market equilibrium is compared to the equilibrium of the counterfactual economy without the contracting friction. *Ceteris paribus*, the presence of the friction leads to suboptimal employment rates, but this is dampened by lower equilibrium wages. If workers recognize that lower wages can increase their retention probability, they are willing to substitute between the two margins. Panel (a) of Table 2.4 shows that especially senior workers are willing to reduce their wage after the pension reform, such that wages of long-tenured workers reduce by 10% in the last ten years before retirement. Results from the Structure of Earnings Survey indicate that the wage-tenure profile of males in Austria have indeed flattened after 2002. In 2002, the average hourly wage at 20–29 years tenure was 12.3% higher than at 10-19 years tenure. By 2014, this differential has declined to 7.3%, see Figure 2.C.3. Figure 2.C.4 shows that also the cross-sectional age-wage distribution became flatter after age 45. It will be interesting to see whether these trends continue in future waves of the study.

Nevertheless, it is questionable whether prime-age job seekers in reality behave as farsighted as assumed in the model. While they might anticipate that contracting a high wage today has adverse effects on their retention probability in the near future, it seems much more difficult to understand how the specificities of the wage contract will affect their chances to be retained once they turn 55. Additionally, the subjective odds of remaining in the firm until age 55 might not be very high *ex ante*, such that these considerations are neglected. To demonstrate how strong awareness of the trade-off between wage and old-age job security affects optimal wage contracts written during prime age, I perform the following counterfactual experiment. I assume that prime-age job seekers act as if their old-age layoff probability was beyond their control. This corresponds to setting $h_s = 0$ in the first order condition (2.14). An alternative interpretation

individual variables	prime-age job seekers		old job seekers
	m	s	n, o
wage w_i^*	0.972	0.972	0.888
layoff probability $F_i(\underline{y}_i^*)$	0.272	0.384	0.379
job-finding probability $p(\theta_i^*)$	0.633	———	0.256
per capita variables	prime age	old age	total
job-finding rate	0.633	0.256	0.514
endog. layoff rate	0.059	0.142	0.072
employment rate	0.884	0.434	0.771
gov. expenditures	0.060	0.293	0.118
output	0.879	0.459	0.774
welfare in cons. eq.	0.809	0.710	0.781

Table 2.10. Equilibrium after the ER reform with boundedly rational prime-age job seekers.

of this experiment is to change s_s to zero and interpret the resulting substitution effect.

As evident from condition (2.14), such a boundedly rational prime-age job seeker choses a flat contract, $w^* := w_m^* = w_s^*$. For the baseline parameterization, the optimal wage is $w^* = 0.974$ and close to the $w_s^* = 0.950$ chosen by a perfectly rational agent (Table 2.3(a)). This is because the utility loss in case of a layoff is small, such that workers have little incentive to act against the layoff risk. Economic aggregates with boundedly rational agents hardly differ from Table 2.3(a). Table 2.10 shows the equilibrium with boundedly rational agents after the pension reform has been implemented. Relative to before the reform, the optimal long-run wage reduces only marginally to $w^* = 0.972$ because the lower g_o hardly affects worker surplus at prime age due to discounting. Whereas under perfect rationality the optimal senior wage decreases to $w_s^* = 0.883$ as evident from Table 2.4(a). As a result, bounded rationality implies a much higher layoff probability of senior workers and a much lower employment rate in old age. The gap in old-age employment relative to the frictionless economy increases to 9.8pp, relative to 4.8pp under perfect rationality. Likewise, the cost of the friction in terms of welfare increases from 3.1% to 4.2%, while the loss in output increases from 2.8% to 4.4%. Therefore, if prime-age job seekers do not fully take into account the link between the age profile of wages and their old-age layoff probability, complementing early retirement reforms with appropriate labor market policies becomes even more pressing.

2.7 Conclusion

In this essay, I have analyzed an age-structured model of the labor market, where wage contracts are subject to a friction. Contracted wages can depend on age but not on productivity. If realized productivity is too low, honoring the *ex ante* optimal wage contract is not profitable for the firm and a layoff occurs. Since equilibrium wages in general exceed reservation wages, part of these layoffs are bilaterally inefficient.

The first key insight of the model is that the friction lowers equilibrium wages and thereby generates an additional rent for the employer. This leads to more vacancy posting, which partly counteracts the higher layoff rates. In the calibrated model, the two forces almost offset each other for prime-age workers, such that the contracting friction only slightly decreases prime-age employment. This is not the case for elderly workers. Elderly workers in long lasting matches unequivocally suffer from the higher job destruction rate, while for old job seekers, the increase in job creation is too small to compensate them for the higher job destruction. Therefore, the contracting friction particularly depresses employment rates in old working age.

The second key insight of the model is that the contracting friction dampens the positive economic effects of reforms to the early retirement system. In the numerical analysis, about 15% of the potential gain in old-age employment cannot be realized because of the friction. The reason is that with the friction, the layoff probability reacts less sensitively to changes in the worker's outside option. As a result, pushing back a government failure (granting excessive outside options to the elderly) increases the detrimental effects of the market failure. Reforms that make early retirement less attractive should therefore be accompanied by labor market policies that increase firms' willingness to keep elderly workers employed. The quantitative results suggest that increasing employment protection for long-tenured old workers is most effective in this regard. The urgency of labor market reforms increases if prime-age job seekers do not take into account that the age profile of wages affects their job security in late working age.

2.A Notation

symbol	explanation
ω	wage contract, either $\omega_m = (w_m, w_s)$, $\omega_s = (w_s)$, or $\omega_o = (w_o)$
w, $w(y)$	period wage (Section 2.4), period wage schedule (Section 2.5)
$\overline{w}$	base wage of the wage schedule $w(y)$ (Section 2.5)
$u(.)$	utility function, defined on (d, ∞)
$\underline{y}(\omega)$	layoff threshold (Section 2.4), profitability threshold (Section 2.5)
y^r	reservation productivity
$F(.)$	cumulative distribution function of productivity distribution
$f(.)$	probability density function of productivity distribution
$h(.)$	hazard function of productivity distribution, $h = \frac{f}{1-F}$
$\hat{z}$	productivity level for which $h(z) + h'(z)z = 0$
θ	labor market tightness
$p(\theta)$	job-finding probability
$q(\theta)$	vacancy-filling probability
$J(\omega; y)$	firm surplus at the production stage
$J(\underline{y}(\omega))$	expected firm surplus at the search stage conditional on retention, $J(\underline{y}(\omega)) = \mathbb{E}[Y - \underline{y}(\omega) \mid Y \geq \underline{y}(\omega)]$
$\mathbb{E}J^+(\omega)$	expected firm surplus at the search stage, $\mathbb{E}J^+(\omega) = (1 - F(\underline{y}(\omega)))J(\underline{y}(\omega))$
$W(\omega; y)$	worker surplus at the production stage
$W(\omega)$	expected worker surplus conditional on retention
$\mathbb{E}W^+(\omega)$	expected worker surplus at the search stage $\mathbb{E}W^+(\omega) = (1 - F(\underline{y}(\omega)))W(\omega)$
V	maximized search value, $V = p(\theta^*)\mathbb{E}W^+(\omega^*)$
N	mass of population
E	mass of employed individuals
e	employment rate, E/N
lf	labor force participation rate
JS	mass of job seekers
Y	aggregate output
G	government expenditures
τ	lump sum tax
$\mathcal{W}$	aggregate welfare
S	wage subsidy
T	layoff tax
P	severance pay
*	indicates optimal level under the contracting friction
$^\bullet$	indicates optimal level without the contracting friction

Table 2.A.1. Overview of defined functions and variables

symbol	explanation
μ_i	location parameter of the productivity distribution
s_i	scale parameter (dispersion) of the productivity distribution
α_i	shape parameter of the productivity distribution
ϕ	probability of drawing a new match productivity
κ	coefficient of absolute risk aversion
b_i	unemployment income, $b_i = g_i + z_i$
g_i	government transfer to unemployed individuals
z_i	value of leisure, home production
π_m	transition probability from prime working age to old working age
π_o	transition probability from old working age to retirement age
β	time discount factor
β_i	effective discount factor, $\beta_m = \beta(1-\pi_m)(1-\sigma)$, $\beta_o = \beta(1-\pi_o)(1-\sigma)(1-\delta)$
σ	probability of an exogenous separation
δ	probability of an inactivity shock
A	level of matching technology
γ	elasticity of the matching function
c	period cost of posting a vacancy

Table 2.A.2. Overview of model parameters

2.B Mathematical appendix

2.B.1 Properties of the normal and logistic distribution

This section verifies that the hazard functions of the standard normal and the standard logistic distribution satisfy properties (iii) and (iv) of Assumption 2.1.

Normal distribution. The pdf of the standard normal distribution is $f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$ and the cdf is $F(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-t^2/2}\, dt$. The hazard function can be expressed as $h(z) = e^{-z^2/2}\big[\int_z^\infty e^{-t^2/2}\, dt\big]^{-1}$. The growth rate of the hazard is $\gamma_h(z) := \frac{h'(z)}{h(z)} = h(z) - z$, which implies $\gamma_h'(z) = h'(z) - 1 = h(z)[h(z) - z] - 1$. According to Sampford (1953), the hazard rate satisfies $0 < h(z)[h(z) - z] < 1$, which implies $\gamma_h(z) > 0$ and $\gamma_h'(z) < 0$ for $z \in \mathbb{R}$. Furthermore, convexity of the conditional expectation follows from $\mathbb{E}[Z - a|Z \geq a] = h(a) - a$ and the fact that the hazard rate of the normal distribution is strictly convex (Sampford, 1953).

Logistic distribution. The pdf of the standard logistic distribution is $f(z) = \frac{e^{-z}}{(1+e^{-z})^2}$, and the cdf is $F(z) = \frac{1}{1+e^{-z}}$. The hazard function is $h(z) = \frac{1}{1+e^{-z}} = F(z)$. Therefore, $\gamma_h(z) = h'(z)/h(z) = f(z)/F(z) = 1 - F(z) > 0$, and $\gamma_h'(z) = -f(z) < 0$. The conditional expectation $\mathbb{E}[Z - a|Z \geq a] = \frac{\ln(1+e^{-a})}{1+e^{a}}$ is strictly convex in a.

The same properties can be established for the Gumbel distribution and the Weibull distribution with shape parameter $k > 1$. The proofs are available by request from the author.

2.B.2 Additional lemmas

The hazard rate is a central object in the analysis. The following Lemma 2.B.1 summarizes important properties:

Lemma 2.B.1. *Consider the hazard rate* $h_i(y) = \frac{f_i(y)}{1-F_i(y)}$ *and define the elasticity* $\varepsilon_h(z) = \frac{h'(z)z}{h(z)}$. *Under Assumption 2.1 and Assumption 2.3, the partial derivatives satisfy the following properties:*

(i) $h_i'(y) > 0$,

(ii) $\frac{\partial h_i(y)}{\partial \mu_i} = -h_i'(y)$,

(iii) $\frac{\partial h_i(y)}{\partial s_i} \lesseqgtr 0$ *for* $\frac{y-\mu_i}{s_i} \gtreqless \hat{z}$, *where* $\hat{z} < 0$ *is characterized by* $\varepsilon_h(\hat{z}) = -1$.

Proof. The imposed assumptions imply that the density of Y_i is $f_i(y) = \frac{1}{s_i} f(z)$ where $z = \frac{y-\mu_i}{s_i}$. The hazard rate is therefore $h_i(y) = \frac{1}{s_i} h(z)$. Properties (i) and (ii) directly follow from monotonicity of h. Differentiation of h_i with respect to s_i gives $\frac{\partial h_i(y)}{\partial s_i} = -\frac{1}{s_i^2}[h(z) + h'(z)z]$. The sign of $\frac{\partial h_i(y)}{\partial s_i}$ is therefore the opposite of $k(z) := 1 + \varepsilon_h(z)$. Since $k(0) = 1$ and $h' > 0$, any root of k must lie in the negative domain. For $z < 0$, Assumption 2.1(iii) implies that $k'(z) = \frac{d}{dz}\left[\frac{h'(z)}{h(z)}\right] z + \frac{h'(z)}{h(z)} > 0$. Hence there exists a unique $\hat{z} < 0$ with $k(\hat{z}) = 0$. □

Another object that repeatedly occurs in the analysis are conditional expectations of the form $\mathbb{E}[Y_i - a | Y_i \geq a]$.

Lemma 2.B.2. *Consider the conditional expectation* $J_i(a) := \mathbb{E}[Y_i - a | Y_i \geq a] = \frac{\int_a^\infty y - a \, dF_i(y)}{1-F_i(a)}$. *Under Assumptions 2.1 and 2.3, the following properties hold:*

(i) $\max\{0, \mathbb{E}Y_i - a\} < J_i(a) < h_i(a)^{-1}$,

(ii) $\lim_{a \to -\infty} [J_i(a) + a] = \mathbb{E}Y_i$,

(iii) $\lim_{a \to \infty} [J_i(a) - h_i(a)^{-1}] = 0$,

(iv) $J_i'(a) < 0$, $\frac{\partial J_i(a)}{\partial \mu_i} = -J_i'(a)$, $\frac{\partial J_i(a)}{\partial s_i} > 0$

Proof. Since the integrand in $J_i(a)$ is non-negative, $J_i(a) > 0$ follows from the definition. The upper bound can be found using integration by parts and exploiting the monotonicity of the hazard function,

$$J_i(a) = \frac{\int_a^\infty 1 - F_i(y)\, dy}{1 - F_i(a)} = \frac{\int_a^\infty f_i(y)/h_i(y)\, dy}{1 - F_i(a)} < \frac{\int_a^\infty f_i(y)\, dy}{1 - F_i(a)} \frac{1}{h_i(a)} = \frac{1}{h_i(a)}.$$

This inequality also implies that $J_i(a)$ is monotonically decreasing, $J_i'(a) = -1 + J_i(a)h_i(a) < 0$. The existence of a second lower bound in (i) and the limit in (ii) can be shown together. Define the auxiliary function $l(a) := J_i(a) + a = \frac{\int_a^\infty y dF_i(y)}{1-F_i(a)}$. Substituting the above expression for the derivative yields $l'(a) = J_i'(a) + 1 = J_i(a)h_i(a) > 0$. Furthermore, $l(a)$ converges to $\mathbb{E}Y_i$ if a tends to $-\infty$. Therefore, $l(a) > \mathbb{E}Y_i$ for all $a \in \mathbb{R}$, and the bound is approached in the limit. Property (iii) follows from L'Hospital's rule, $\lim_{a\to\infty} J_i(a) = \lim_{a\to\infty} \frac{1-F_i(a)}{f_i(a)} = \lim_{a\to\infty} h_i(a)^{-1}$. Concerning the derivatives with respect to the parameters of the distribution, observe that for any parameter ξ it holds that

$$\frac{\partial J_i(a)}{\partial \xi} = \frac{\int_a^\infty \frac{\partial 1-F_i(y)}{\partial \xi} dy}{1-F_i(a)} - J_i(a)\frac{\frac{\partial 1-F_i(a)}{\partial \xi}}{1-F_i(a)}. \tag{2.B.1}$$

Substituting $F_i(a) = F(\frac{a-\mu_i}{s_i})$ reveals $\frac{\partial 1-F_i(y)}{\partial \mu_i} = f_i(y)$. Plugging this back into (2.B.1) reveals $\frac{\partial J_i(a)}{\partial \mu_i} = 1 - J_i(a)h_i(a) = -J_i'(a)$. The derivative with respect to s_i is $\frac{\partial 1-F_i(y)}{\partial s_i} = \frac{y-\mu_i}{s_i} f_i(y)$. Substituting this into (2.B.1) and collecting terms yields

$$\frac{\partial J_i(a)}{\partial s_i} = \frac{J_i(a)}{s_i} + \frac{a-\mu_i}{s_i}[1 - J_i(a)h_i(a)] \tag{2.B.2}$$

By property (i), the term in square brackets is positive, such that $\frac{\partial J_i(a)}{\partial s_i} > 0$ for $a \geq \mu_i$. To show that $\frac{\partial J_i(a)}{\partial s_i} > 0$ also for $a \leq \mu_i$, it is sufficient to verify that $l(a) = J_i(a) + a \geq \mu_i$. This holds because it has been shown above that $l(a) > \mathbb{E}Y_i$, and $\mathbb{E}Y_i \geq \mu_i$ follows from Assumption 2.1(ii) since $\mathbb{E}Y_i = \mu_i + s_i\mathbb{E}Z$ for $\alpha_i = 1$. □

2.B.3 Proofs omitted in the text

Proof of Proposition 2.1. Define the function on the left-hand side of (2.4) as $\Upsilon(a) = a - w_o + \lambda \int_a^\infty y - a\, dF_o(y)$ where $\lambda := \beta_o\phi/(1-\beta_o(1-\phi)) \in [0,1)$. Let $w_o \in \mathbb{R}$. It is easy to see that $\Upsilon(w_o) > 0$. Differentiation yields $\Upsilon'(a) = 1 - \lambda(1 - F_o(a)) > 0$ and hence Υ is strictly monotonically increasing on $\mathbb{R}$. By continuity, a unique root exists if $\lim_{a\to-\infty} \Upsilon(a) < 0$. Rewrite $\Upsilon(a) = \lambda \int_a^\infty y dF_o(y) - w_o + [1 - \lambda(1 - F_o(a))]a$. Taking the limit $a \to -\infty$, the first term converges to $\mathbb{E}Y_o$. Since the term in square brackets converges to $(1-\lambda) > 0$, the expression as a whole becomes unbounded, $\lim_{a\to-\infty} \Upsilon(a) = -\infty$, whereby a unique root exists. By the implicit function theorem, $\frac{\partial \underline{y}_o}{\partial \xi} = -\Upsilon'(\underline{y}_o)^{-1} \frac{\partial \Upsilon(\underline{y}_o)}{\partial \xi}$ for an arbitrary parameter ξ. Hence the marginal effect of ξ on $\underline{y}_o$ has the opposite sign of $\frac{\partial \Upsilon(\underline{y}_o)}{\partial \xi}$. Clearly, this partial derivative is negative for w_o, such that $\underline{y}_o$ increases. The partial derivative is positive for λ, which in turn is increasing in β_o and ϕ. To obtain the marginal effect with respect to the parameters of the productivity distribution, note that $\int_a^\infty y - a\, dF_o(y) = \int_a^\infty 1 - F_o(y)\, dy = \int_a^\infty 1 - F(\frac{y-\mu_o}{s_o})\, dy$. The survival function is increasing in μ_o since $\frac{\partial 1-F_o(y)}{\partial \mu_o} = f_o(y) > 0$. Concerning s_o, observe that $\frac{\partial}{\partial s_o} \int_a^\infty 1 - F_o(y)\, dy = \int_a^\infty \frac{y-\mu_o}{s_o} dF_o(y) = (1 - F_o(a))\frac{J_o(a)+a-\mu_o}{s_o} > (1 - F_o(a))\frac{\mathbb{E}Y_o-\mu_o}{s_o} = (1 - F_o(a))\mathbb{E}Z \geq 0$ by Lemma 2.B.2(i) and Assumption 2.1(ii). As a result, $\underline{y}_o$ is decreasing in both parameters. □

Proof of Proposition 2.2. Under $\pi_o = 1$, the equilibrium wage must satisfy $\Phi(w_o^*) = 0$, where

Φ is given in (2.7). Worker surplus $W_o(w) = u(w-\tau) - u(b_o-\tau)$ is increasing in w. Since $h_o'(w) > 0$ and $J_o'(w) < 0$ by Lemma 2.B.1, we have $\Phi'(w) < 0$ for all $w \in \mathbb{R}$. Furthermore, it is easy to see that $\Phi(b_o) = u'(b_o) > 0$, and that Lemma 2.B.2(iii) implies

$$\lim_{w\to\infty} \Phi(w) = \lim_{w\to\infty} \Big[u'(w-\tau) - \frac{h_o(w)W_o(w)}{\gamma}\Big]. \tag{2.B.3}$$

Since $u'(w)$ vanishes asymptotically by Assumption 2.2, the limit is strictly negative. By continuity, Φ has a unique root $w_o^* > b_o$. For given τ, the unique labor market equilibrium is therefore given by the triple (θ_o^*, w_o^*, V_o) where $\theta_o^* = \big[A\mathbb{E}J_o^+(w_o^*)/c\big]^{1/\gamma}$ and $V_o = p(\theta_o^*)\mathbb{E}W_o^+(w_o^*)$. □

Proof of Proposition 2.3. The increase in w_o^* follows from Lemma 2.B.1 and Lemma 2.B.2. Since $\frac{\partial F_o(w_o^*)}{\partial \mu_o} = -f_o(w_o^*)$ and $-\int_{w_o^*}^{\infty} \frac{\partial F_o(y)}{\partial \mu_o}\,dy = 1 - F_o(w_o^*)$, the two additional statements hold if and only if $\frac{\partial w_o^*}{\partial \mu} < 1$. By the implicit function theorem, this is equivalent to $-\frac{\partial \Phi(w_o^*)}{\partial \mu_o} < \Phi'(w_o^*)$. This inequality can be verified by substituting the respective expressions, taking into account that all terms in $\Phi'(w_o^*)$ are positive, that $\frac{\partial J_o(w_o)}{\partial \mu_o} = -J_o'(w_o)$, and that $\frac{\partial h_o(w_o)}{\partial \mu_o} = -h_o'(w_o)$. □

Proof of Proposition 2.4. The wage effect follows from Lemma 2.B.1 and Lemma 2.B.2. Since $\frac{\partial F_o(w_o^*)}{\partial s_o} = -\frac{w_o^*-\mu_o}{s_o} f_o(w_o^*)$, the layoff probability certainly increases if $-\frac{w_o^*-\mu_o}{s_o} + (\frac{\partial w_o^*}{\partial s_o})^{SE} \geq 0$ since the income effect on the wage is positive. The substitution effect can be written $(\frac{\partial w_o^*}{\partial s_o})^{SE} = \Phi'(w_o^*)^{-1}\frac{\partial h_o(w_o^*)}{\partial s_o} W_o(w_o^*)$ where $\frac{\partial h_o(w)}{\partial s_o} = -\frac{h_o(w)}{s_o} - h_o'(w)\frac{w-\mu_o}{s_o} < h_o'(w)\frac{\mu_o - w}{s_o}$. Assuming $w_o^* \leq \mu_o$ and noting $\Phi'(w_o^*) < -h_o'(w_o^*)W_o(w_o^*) < 0$ yields $(\frac{\partial w_o^*}{\partial s_o})^{SE} > \Phi'(w_o^*)^{-1}h_o'(w_o^*)\frac{\mu_o - w_o^*}{s_o}W_o(w_o^*) > -\frac{\mu_o-w_o^*}{s_o}$. Therefore, the above inequality holds, and the layoff probability is strictly increasing in s_o provided that $w_o^* \leq \mu_o$. To show that also the job-finding rate is increasing under certain circumstances, I first demonstrate that the wage response is bounded by $\frac{\partial w_o^*}{\partial s_o} < \gamma\frac{\partial J_o}{\partial s_o}$. Since the right hand-side is positive, this is trivial for $\frac{\partial w_o^*}{\partial s_o} \leq 0$. Otherwise the implicit function theorem gives $\frac{\partial w_o^*}{\partial s_o} = -\Phi'(w_o^*)^{-1}\frac{\partial \Phi(w_o^*)}{\partial s_o}$ where $\frac{\partial \Phi(w_o^*)}{\partial s_o}$ is strictly positive. Convexity of J_o implies $h_o'(w) \geq \frac{1-J_o(w_o)h_o(w_o)}{J_o(w_o)}$, which can be used to show $\Phi'(w_o^*) < -\frac{u'(w_o^*-\tau)}{\gamma J_o(w_o^*)}$ as well as $\frac{\partial \Phi(w_o^*)}{\partial s_o} \leq \frac{u'(w_o^*-\tau)}{s_o}\big[1 + \frac{1-J_o(w_o^*)h_o(w_o^*)}{J_o(w_o^*)}(w_o^* - \mu_o)\big]$. The latter bound is only valid if $w_o^* \leq \mu_o$. Combining the two inequalities yields $\frac{\partial w_o^*}{\partial s_o} < \gamma\{\frac{J_o(w_o^*)}{s_o} + [1 - J_o(w_o^*)h_o(w_o^*)]\frac{w_o^*-\mu_o}{s_o}\} = \gamma\frac{\partial J_o(w_o^*)}{\partial s_o}$. The direct effect in (2.9) is $-\int_{w_o^*}^{\infty}\frac{\partial F_o(y)}{\partial s_o}\,dy = \frac{\int_{w_o^*}^{\infty} y-\mu_o\,dF_o(y)}{s_o} = (1-F_o(w_o^*))\frac{J_o(w_o^*)+w_o^*-\mu_o}{s_o}$. The sign of the total effect therefore equals the sign of $J_o(w_o^*) + w_o^* - \mu_o - s_o\frac{\partial w_o^*}{\partial s_o}$. The first term is positive since $J_o(w_o^*) + w_o^* - \mu_o > \mathbb{E}Y_o - \mu_o = s_o\mathbb{E}Z \geq 0$ by Lemma 2.B.2(i) and Assumption 2.1(ii). Hence the job-finding probability unambiguously decreases if $\frac{\partial w_o^*}{\partial s_o} < 0$. Otherwise the bound on the wage change established above reveals $J_o(w_o^*) + w_o^* - \mu_o - s_o\frac{\partial w_o^*}{\partial s_o} > J_o(w_o^*) + w_o^* - \mu_o - \gamma\{J_o(w_o^*) + [1 - J_o(w_o^*)h_o(w_o^*)](w_o^* - \mu_o)\}$. The right-hand side is non-negative if and only if $\gamma \leq \frac{J_o(w_o^*)+w_o^*-\mu_o}{J_o(w_o^*)+[1-J_o(w_o^*)h_o(w_o^*)](w_o^*-\mu_o)}$. □

Proof of Proposition 2.5. An optimal wage contract with $w_o > b_o$ must satisfy the two first order equations $\Phi(w_m, w_s) = 0$ and $\Psi(w_m, w_s) = 0$, where

$$\Phi(w_m, w_s) = u'(w_m - \tau) - \frac{1-\gamma}{\gamma} \frac{W_m(\omega_m)}{J_m(\underline{y}_m)} - h_m(\underline{y}_m) W_m(\omega_m),$$
$$\Psi(w_m, w_s) = u'(w_s - \tau) - u'(w_m - \tau) - h_s(w_s) W_s(w_s).$$

and $\underline{y}_m = w_m - \beta(1-\sigma)\mathbb{E}J_s^+(w_s)$. Otherwise the optimal contract has the form (w_m, b_o), where w_m solves $\Phi(w_m, b_o) = 0$.

The CS curve $CS(w_m)$ is defined piecewise. Consider $w_m \geq b_o$. In this case $CS(w_m)$ is implicitly defined by $\Psi(w_m, w_s) = 0$. For given w_m, a unique root exists since $\Psi(w_m, b_o) \geq \Psi(b_o, b_o) = 0$, $\Psi(w_m, w_m) \leq 0$, and Ψ is strictly decreasing in w_s. These properties imply $CS(b_o) = b_o$ and $CS(w_m) \in (b_o, w_m)$ for $w_m > b$. Moreover, the curve is upwards sloping with a slope less than 1, $CS'(w_m) = -\frac{\partial \Psi}{\partial w_m} / \frac{\partial \Psi}{\partial w_s}\big|_{\Psi=0} = \frac{-u''(w_m-\tau)}{-u''(w_s-\tau)+h_s'(w_s)W_s(w_s)+h_s(w_s)u'(w_s-\tau)}\big|_{\Psi=0} < 1$ for $w_m > b_o b$. Since $\lim_{w_m\to\infty} u'(w_m) = 0$, the CS curve converges to a wage level $\overline{w}_s$ defined by $u'(\overline{w}_s - \tau) = h_s(\overline{w}_s) W_s(\overline{w}_s)$. Now consider the second possibility, $w_m < b_o$. In this case the level of w_s that satisfies $\Psi(w_m, w_s) = 0$ lies below b_o, which would violate the worker's participation constraint, $W_s(w_s) \geq 0$. Therefore, the optimal contract is a constrained one, $CS(w_m) = b_o$, and the curve is flat in this region.

The SS curve $SS(w_m)$ is monotonically decreasing since $\frac{\partial \Phi}{\partial w_m} < 0$ and $\frac{\partial \Phi}{\partial w_s} < 0$. Before proofing existence of an intersection, I verify that the SS curve is well-defined in the relevant range of wages. In particular, I show that for every $w_s \in [b_o, \overline{w}_s)$ there exists a w_m such that $\Phi(w_m, w_s) = 0$. First, $b_m \leq b_o$ ensures that $W_m(w_o^*, w_s) > 0$, while $\lim_{w_m\to d} W_m(w_m, w_s) = -\infty$ by Assumption 2.2. This ensures a $\hat{w}_m$ such that $W_m(\hat{w}_m, w_s) = 0$, which implies $\Phi(\hat{w}_m, w_s) = u'(\hat{w}_m - \tau) > 0$. On the other hand, Lemma 2.B.2 and Assumption 2.2 ensure that $\lim_{w_m\to\infty} \Phi(w_m, w_s) = -\lim_{w_m\to\infty} h(\underline{y}_m) W_m(\omega_m)/\gamma < 0$. Since Φ is continuous and strictly decreasing in w_m, for any fixed w_s there exists a unique w_m such that $\Phi(w_m, w_s) = 0$, and the SS curve is well-defined for $w_s \in [b_o, \overline{w}_s)$.

It remains to proof that the two curves intersect. Since $\Phi(b_m, b_o) > 0$, the SS curve lies above the CS curve at $w_m = b_m$. Furthermore, the SS curve strictly decreases and defines a unique w_m for every $w_s \in [b_o, \overline{w}_s)$. Since the CS curve is increasing and tends to $\overline{w}_s$ as $w_m \to \infty$, there exists a unique intersection. For given τ, the unique labor market equilibrium is therefore unique and given by the triple $(\theta_m^*, \omega_m^*, V_m)$ where $\omega_m^* = (w_m^*, w_s^*)$, $\theta_m^* = \left[A\mathbb{E}J_m^+(\omega_m^*)/c\right]^{1/\gamma}$, and $V_m = p(\theta_m^*)\mathbb{E}W_m^+(\omega_m^*)$. The equilibrium contract satisfies $w_s^* > b_o$ if and only if $\Phi(b_o, b_o) > 0$. Since $b \to \Phi(b, b)$ is strictly decreasing with $\Phi(b_m, b_m) > 0$ and $\lim_{b\to\infty} \Phi(b, b) < 0$, there exists a threshold $\bar{b}_o$ as postulated by the proposition. □

Proof of Proposition 2.6. The response in equilibrium wages can be expressed using the implicit function theorem as

$$\begin{pmatrix} \frac{\partial w_m^*}{\partial \xi} \\ \frac{\partial w_s^*}{\partial \xi} \end{pmatrix} = - \begin{pmatrix} \frac{\partial \Phi}{\partial w_m} & \frac{\partial \Phi}{\partial w_s} \\ \frac{\partial \Psi}{\partial w_m} & \frac{\partial \Psi}{\partial w_s} \end{pmatrix}^{-1} \begin{pmatrix} \frac{\partial \Phi}{\partial \xi} \\ \frac{\partial \Psi}{\partial \xi} \end{pmatrix}$$

where all partial derivatives are evaluated in the optimum ω_m^*. For $\xi \in \{\mu_m, s_m\}$, the derivative $\frac{\partial \Psi}{\partial \xi}$ is zero and we can rewrite $(\frac{\partial w_m^*}{\partial \xi}, \frac{\partial w_s^*}{\partial \xi})' = (-\frac{\partial \Psi}{\partial w_s}, \frac{\partial \Psi}{\partial w_m})' \frac{\partial \Phi}{\partial \xi} D^{-1}$ where $D = \frac{\partial \Phi}{\partial w_m}\frac{\partial \Psi}{\partial w_s} - \frac{\partial \Phi}{\partial w_s}\frac{\partial \Psi}{\partial w_m} > 0$ is the determinant of the Jacobian. Since the entries of the vector on the right-hand side are all positive, the two wage levels move in the same direction, and the sign of $\frac{\partial w_i^*}{\partial \xi}$ equals the sign of $\frac{\partial \Phi}{\partial \xi}$. Lemma 2.B.1 and Lemma 2.B.2 imply that $\frac{\partial \Phi}{\partial \mu_m} > 0$ such that the equilibrium wages increase in μ_m, while the wage effect of s_m is ambiguous.

The effects of an arbitrary parameter ξ on layoffs and hiring are similar to (2.8)–(2.9)

$$\frac{dF_m(\underline{y}_m^*)}{d\xi} = \frac{\partial F_m(\underline{y}_m^*)}{\partial \xi} + f_m(\underline{y}_m^*)\frac{\partial \underline{y}_m^*}{\partial \xi}, \quad \frac{dF_s(w_s^*)}{d\xi} = \frac{\partial F_s(w_s^*)}{\partial \xi} + f_s(w_s^*)\frac{\partial w_s^*}{\partial \xi},$$

$$\frac{d\mathbb{E}J_m^+(\omega_m^*)}{d\xi} = -\int_{\underline{y}_m^*}^{\infty} \frac{\partial F_m(y)}{\partial \xi}\, dy - (1 - F_m(\underline{y}_m^*))\frac{\partial \underline{y}_m^*}{\partial \xi}.$$

For $\xi \in \{\mu_m, s_m\}$, the change in the layoff probability of senior workers is proportional to their wage response, $\frac{dF_s(w_s^*)}{d\xi} = f_s(w_s^*)\frac{\partial w_s^*}{\partial \xi}$, whereby $\frac{dF_s(w_s^*)}{d\mu_m} > 0$. By the definition of $\underline{y}_m^*$, observe

$$\frac{\partial \underline{y}_m^*}{\partial \xi} = \frac{\partial w_m^*}{\partial \xi} + \beta(1 - F_s(w_s^*))\frac{\partial w_s^*}{\partial \xi} = \frac{-\frac{\partial \Psi}{\partial w_s} + \beta(1-\sigma)(1 - F_s(w_s^*))\frac{\partial \Psi}{\partial w_m}}{D}\frac{\partial \Phi}{\partial \xi}. \tag{2.B.4}$$

Straightforward differentiation reveals that in optimum $\frac{\partial \Phi}{\partial w_s} = \beta(1-\sigma)(1 - F_s(w_s^*))[\frac{\partial \Phi}{\partial w_m} - u''(w_m^* - \tau)]$. The determinant can therefore be rewritten $D = -\frac{\partial \Phi}{\partial w_m}[-\frac{\partial \Psi}{\partial w_s} + \beta(1-\sigma)(1 - F_s(w_s^*))\frac{\partial \Psi}{\partial w_m}] + \beta(1-\sigma)(1 - F_s(w_s^*))u''(w_m^* - \tau)\frac{\partial \Psi}{\partial w_m}$. Substituting this into (2.B.4) and noting $u'' < 0$ reveals that $\frac{\partial \underline{y}_m^*}{\partial \xi} = \lambda(-\frac{\partial \Phi}{\partial \xi})/\frac{\partial \Phi}{\partial w_m} = \lambda \frac{\partial w_m^*}{\partial \xi}\big|_{w_s = w_s^*}$ for a $\lambda \in (0,1)$. The proofs of Proposition 2.3 and Proposition 2.4 can be replicated to show that $\frac{\partial w_m^*}{\partial \mu_m}\big|_{w_s = w_s^*} \in (0,1)$ and $\frac{\partial w_m^*}{\partial s_m}\big|_{w_s = w_s^*} \leq \gamma \frac{\partial J_m(\underline{y}_m^*)}{\partial s_m}$. Since $\lambda \in (0,1)$, the same bounds hold for $\frac{\partial \underline{y}_m^*}{\partial \mu_m}$ and $\frac{\partial \underline{y}_m^*}{\partial s_m}$. The remainder of the proof is then analogous to that of Proposition 2.3 and Proposition 2.4. □

Proof of Proposition 2.7. I demonstrate that the assumption on γ is sufficient for the SS curve to shift upwards if μ_s increases. The SS curve shifts upwards at the optimum if and only if $\frac{\partial \Phi(\omega_m^*)}{\partial \mu_s} = \frac{\partial \Phi(\omega_m^*)}{\partial W_m}\frac{\partial W_m(\omega_m^*)}{\partial \mu_s} + \frac{\partial \Phi(\omega_m^*)}{\partial \underline{y}_m}\frac{\partial \underline{y}_m^*}{\partial \mu_s} > 0$. It is easy to verify that $\frac{\partial \Phi(\omega_m^*)}{\partial W_m} = -\frac{u'(w_m^* - \tau)}{W_m(\omega_m^*)}$, and that the convexity of the conditional expectation J_m implies $\frac{\partial \Phi(\omega_m^*)}{\partial \underline{y}_m} < \frac{J_m'(\underline{y}_m^*)}{J_m(\underline{y}_m^*)}u'(w_m^* - \tau)$. Furthermore, $\frac{\partial W_m(\omega_m^*)}{\partial \mu_s} = \beta(1-\sigma)f_s(w_s^*)W_s(w_s^*)$ and $\frac{\partial \underline{y}_m^*}{\partial \mu_s} = -\beta(1-\sigma)(1-F_s(w_s^*))$. Combining all of the above yields $\frac{\partial \Phi(\omega_m^*)}{\partial \mu_s} > -\beta(1-\sigma)(1-F_s(w_s^*))u'(w_m^* - \tau)\Big[\frac{h_s(w_s^*)W_s(w_s^*)}{W_m(\omega_m^*)} + \frac{J_m'(\underline{y}_m^*)}{J_m(\underline{y}_m^*)}\Big]$. The term in square brackets has the same sign as $J_m'(\underline{y}_m^*)\frac{W_m(\omega_m^*)}{J_m(\underline{y}_m^*)} + h_s(w_s^*)W_s(w_s^*) = h_m(\underline{y}_m^*)W_m(\omega_m^*) - \frac{W_m(\omega_m^*)}{J_m(\underline{y}_m^*)} + h_s(w_s^*)W_s(w_s^*) = u'(w_s^* - \tau) - \frac{1}{\gamma}\frac{W_m(\omega_m^*)}{J_m(\underline{y}_m^*)}$, where the last identity exploits the two optimality conditions. The assumption on γ postulated by Proposition 2.7 therefore ensures $\frac{\partial \Phi(\omega_m^*)}{\partial \mu_s} > 0$. □

Proof of Proposition 2.9. Since $J_o(w_o^\bullet(y); y) > 0$ only for $y > \underline{y}_o^\bullet$, expected firm surplus can be rewritten as $\mathbb{E}J_o^+(w_o^\bullet) = \int_{y_o^r}^{\infty} J_o(w_o^\bullet; y)\, dF_o(y) = \frac{\mathbb{E}[Y_o - \underline{y}_o^\bullet | Y_o \geq \underline{y}_o^\bullet]}{1-\beta_o(1-\phi)}$. Since $w_o^\bullet(y_o^r) = y_o^r + \beta\phi\mathbb{E}J_o^+(w_o^\bullet)$, equation (2.5) reveals $\underline{y}_o(w_o^\bullet(y_r)) = y_o^r$. Monotonicity then implies $\underline{y}_o^* = \underline{y}_o(w_o^*) > y_o^r$. By the free entry conditions (2.6) and (2.19), the job-finding probability is only a function of expected firm surplus $\mathbb{E}J_o^+$. Define $I(a) = \frac{\int_a^\infty 1-F_o(y)\,dy}{1-\beta_o(1-\phi)}$, which is strictly decreasing in a. Under the friction, $\mathbb{E}J_o^+(w_o^*) = I(\underline{y}_o(w_o^*))$, while without the friction, $\mathbb{E}J_o^+(w_o^\bullet) = I(\underline{y}_o(\overline{w}_o^\bullet))$. Since $\underline{y}_o$ is strictly increasing and $w_o^* < \overline{w}_o^\bullet$ by assumption, we have $\mathbb{E}J_o^+(w_o^*) > \mathbb{E}J_o^+(w_o^\bullet)$. □

2.C Equilibrium with labor market policies

2.C.1 Surplus functions and optimality conditions

To study different labor market policies in Section 2.6.4, the model of Section 2.4 is extended by the following elements,

- a training program that changes the productivity distribution F_i and costs the public C_i per participant,
- a firm that employs a type i worker receives a wage cost subsidy S_i,
- a firm that (endogenously) lays off a type i worker pays a layoff tax T_i to the government and severance pay P_i to the displaced worker.

I only discuss the changes regarding old workers at this place. The same modifications apply to prime-age and senior workers. Due to the wage subsidy, an old worker earns w_o but costs the firm only $w_o - S_o$. This changes firm surplus at the production stage to

$$J_o(w_o; y) = \frac{y - (w_o - S_o) + \beta_o\phi\mathbb{E}J_o^+(w_o)}{1 - \beta_o(1-\phi)}.$$

Due to the layoff tax and the severance pay, the worker is laid off whenever $J_o(w_o; y) + T_o + P_o < 0$ which changes the layoff threshold to $\underline{y}_o(w_o) = w_o - S_o - \beta_o\phi\mathbb{E}J_o^+(w_o) - (1-\beta_o(1-\phi))(T_o + P_o)$. This allows to express firm surplus as $J_o(w_o; y) = \frac{y - \underline{y}_o(w_o)}{1-\beta_o(1-\phi)} - (T_o + P_o)$. Expected firm surplus has to take into account that for $y < \underline{y}_o(w_o)$ the firm incurs layoff costs,

$$\mathbb{E}J_o^+(w_o) = \frac{\int_{\underline{y}_o}^\infty y - \underline{y}_o\, dF_o(y)}{1 - \beta_o(1-\phi)} - (T_o + P_o).$$

This yields the implicit equation for the layoff threshold

$$\underline{y}_o - (w_o - S_o) + \frac{\beta_o\phi}{1 - \beta_o(1-\phi)}\int_{\underline{y}_o}^\infty y - \underline{y}_o\, dF_o(y) + (1-\beta_o)(T_o + P_o) = 0.$$

Worker surplus at the production stage is $W_o(w_o) = \frac{u(w_o-\tau) - u(b_o-\tau) + \beta_o(\mathbb{E}W_o^+(w_o) - V_o)}{1-\beta_o(1-\phi)}$, while expected surplus is $\mathbb{E}W_o^+(w_o) = (1 - F_o(\underline{y}_o))W_o(w_o) + F_o(\underline{y}_o)(u(b_o + P_o - \tau) - u(b_o - \tau))$. Substituting

$W_o(w_o)$ yields

$$\mathbb{E}W_o^+(w_o) = (1 - F_o(\underline{y}_o)\frac{u(w_o-\tau) - u(b_o-\tau) - \beta_o V_o}{1-\beta_o(1-\phi F_o(\underline{y}_o))} + F_o(\underline{y}_o)(1-\beta_o(1-\phi))\frac{u(b_o+P_o-\tau) - u(b_o-\tau)}{1-\beta_o(1-\phi F_o(\underline{y}_o))}.$$

The first order condition (2.5) becomes

$$u'(w_o^*-\tau) = \frac{1-\gamma}{\gamma}\frac{\mathbb{E}W_o^+(w_o^*)}{\mathbb{E}J_o^+(w_o^*)} + (1-\beta_o(1-\phi))h_o(\underline{y}_o^*)\frac{\partial \underline{y}_o^*}{\partial w_o}[W_o(w_o^*) + u(b_o-\tau) - u(b_o+P_o-\tau)]$$

where $\frac{y_o^*}{\partial w_o} = \frac{1-\beta_o(1-\phi)}{1-\beta_o(1-\phi F_o(\underline{y}_o^*))}$. Similar changes apply to the surplus functions of prime-age and senior workers and the first order conditions for ω_m. In the aggregate, wage subsidies, training, and layoff taxes change the composition of government expenditures,

$$G_1 = (N_1 - E_m)g_m + E_m S_m - L_m T_m - C_m p(\theta_m^*)Q_m,$$
$$G_2 = (N_2 - E_s - E_o)g_o + E_s S_s + E_o S_o - L_s T_s - L_o T_o - C_s \pi_m(1-\sigma)E_m - p(\theta_o^*)Q_o,$$

where L_i amounts to the mass of layoff events involving type i workers,

$$L_m = [JS_m p(\theta_m^*) + (1-\pi_m)(1-\sigma)\phi E_m]F_m(\underline{y}_m^*),$$
$$L_s = [\pi_m(1-\sigma)E_m + (1-\pi_o)(1-\sigma)(1-\delta)\phi E_s]F_s(\underline{y}_s^*),$$
$$L_o = [JS_o p(\theta_o^*) + (1-\pi_o)(1-\sigma)(1-\delta)\phi E_o]F_o(\underline{y}_o^*),$$

and Q_i denotes the mass of type i individuals who have not been employed in their age class before, which satisfy

$$Q_m = \pi_m N_1 + (1-\pi_m)(1-p(\theta_m^*))Q_m,$$
$$Q_o = \pi_m(1-p(\theta_o^*))[N_1 - (1-\sigma)E_m] + (1-\pi_o)(1-\delta)(1-p(\theta_o^*))Q_o.$$

Severance pay directly affects welfare, which is updated to

$$\mathcal{W}_1 = E_m u(w_m^* - \tau) + (N_1 - E_m - L_m)u(b_m - \tau) + L_m u(b_m + P_m - \tau),$$
$$\mathcal{W}_2 = E_s u(w_s^* - \tau) + E_o u(w_o^* - \tau) + (N_2 - E_o - E_s - L_o - L_s)u(b_o - \tau) + L_s u(b_p + P_s - \tau) + L_o u(b_o + P_o - \tau).$$

2.C.2 Quantitative effects

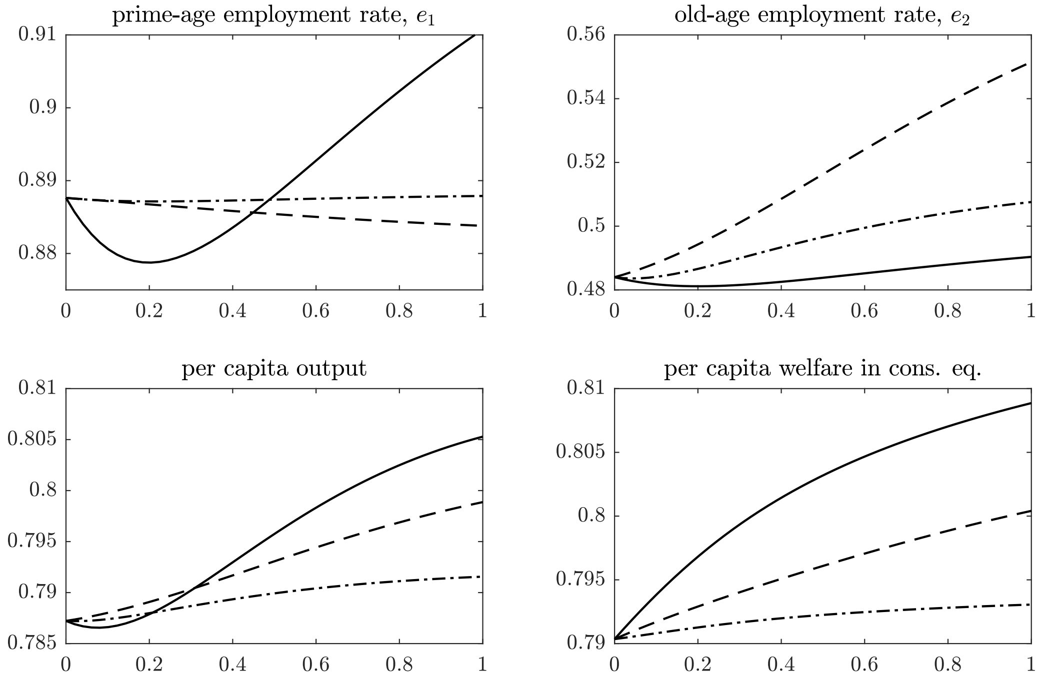


Figure 2.C.1. Effect of severance pay on employment, output and welfare, relative to Table 2.4(a); only one variable is altered at a time; solid line: P_m, dashed line: P_s, dash-dotted line: P_o

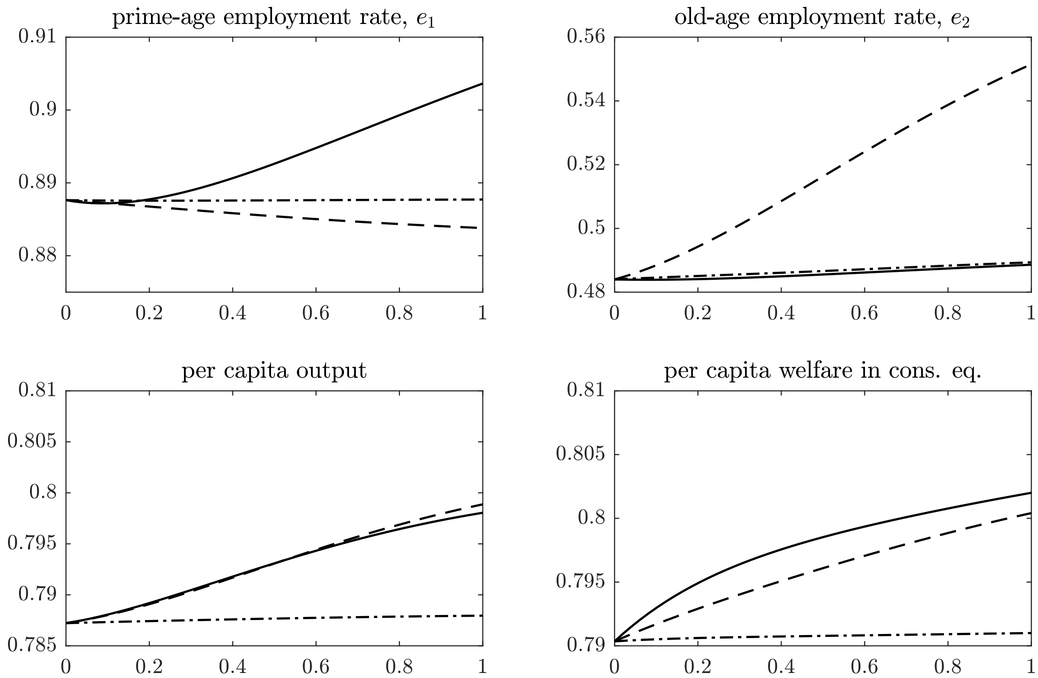


Figure 2.C.2. Effect of severance pay with a probation period on employment, output and welfare, relative to Table 2.4(a); only one variable is altered at a time; solid line: P_m, dashed line: P_s, dash-dotted line: P_o

2.C.3 Wage profiles in Austria

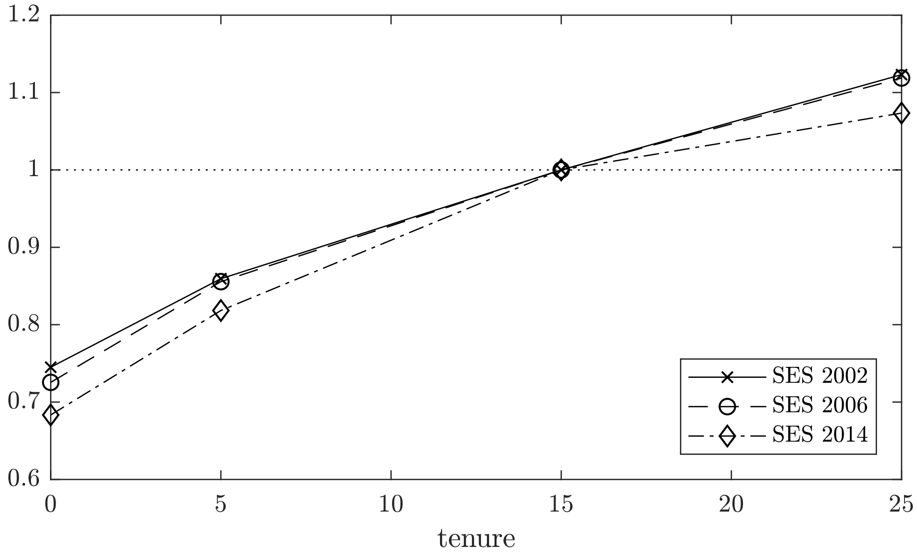


Figure 2.C.3. Hourly wage by tenure relative to tenure group 10–19, dependent employed males in the private sector in Austria, source: SES waves 2002, 2006, 2014 (Statistik Austria, 2006, 2009, 2017)

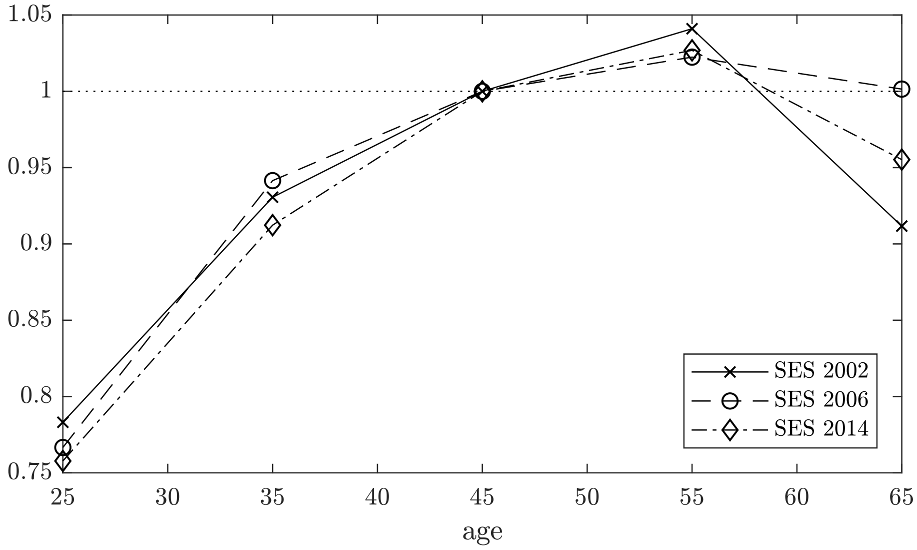


Figure 2.C.4. Hourly wage by age relative to age group 40–50, dependent employed males in the private sector in Austria, source: SES waves 2002, 2006, 2014 (Statistik Austria, 2006, 2009, 2017)

3 Optimal severance pay under different contractual regimes

3.1 Introduction

Severance pay, a financial compensation that the worker receives from her previous employer after a dismissal, is a widely observed policy around the world. If severance pay is not mandated by the federal government, it is typically regulated in industry-wide collective agreements, or negotiated privately between worker and firm.[1] Holzmann et al. (2012) identify three economic rationales for severance pay. Firstly, severance pay provides partial income insurance against a dismissal and attenuates the welfare loss of the worker. Second, severance pay levels are typically increasing in tenure, which encourages human capital investments of the worker and reduces moral hazard motives. Thirdly, employers perceive severance pay as a tax that makes dismissals more expensive, which reduces layoff rates. For this reason, severance pay is a key element of employment protection legislation. According to Boeri et al. (2017), almost half of the cross-country variation in the OECD index of the strictness of employment protection legislation is explained by severance pay.

While the potential of severance pay to increase social welfare is undisputed, the policy also has unintended consequences. Most importantly, severance pay inhibits job and worker mobility. In low productivity states, firms increase labor hoarding relative to the *laissez-faire* to avoid the costs of a layoff. This keeps low productive firms alive, such that average productivity in the economy is likely to decrease (Hall and Lazear, 1984). Additionally, since severance pay usually accrues after employer-induced separations but not after voluntary quits, workers may be held back from moving to better performing firms (Kettemann et al., 2017). Anticipating separation costs, employers also reduce job creation. While the net effect on employment is ambiguous, the expected duration of unemployment increases. For this reason, high employment protection in Europe is often seen as the source for the sclerotic European labor markets (Bentolila and Bertola, 1990; Lazear, 1990). Therefore, while severance pay may in general be welfare improving, excessive levels of severance can reduce social welfare. Apart from the size of severance pay, the welfare effects also depend on the group of eligibile persons. If only a certain subpopulation

[1] A comprehensive overview and classification of the existing systems in 183 countries is given in Annex B of Holzmann et al. (2012).

of workers is eligible to severance pay, firms may substitute away from these workers. Less stringent employment protection is in fact one reason for the surge of temporary and fixed term contracts in southern European countries (Blanchard and Landier, 2002; Boeri and Garibaldi, 2007; OECD, 2014). Similar substitution effects occur if layoff costs are differentiated along other dimensions, such as age (Behaghel et al., 2008).

Having said this, it is not trivial that severance pay has a welfare effect at all. Lazear (1990) shows that if contracts are complete and all agents are risk neutral, firm and worker can privately undo any mandated severance pay by adjusting wages appropriately. Severance pay then neither affects labor market dynamics nor social welfare. To establish a welfare case, the literature has departed from these idealized conditions in two directions. One strand of literature maintains risk neutrality and imposes contractual frictions that may be rationalized by asymmetric information or moral hazard. Recent studies in this vein include Fella (2012) and Boeri et al. (2017). In both papers, wages cannot adjust to productivity fluctuations. A second strand of literature has focused on the insurance role of severance pay. Workers are assumed to be risk averse while contracts are complete, see Cozzi and Fella (2016) and Lalé (2018) for instance. Alvarez and Veracierto (2001), Blanchard and Tirole (2008), and Fella and Tyson (2013) combine both ideas and consider risk averse workers and contractual frictions in the same model. The purpose of this essay is to provide a systematic treatment of how risk attitudes of the worker and the contractual flexibility of workers and firms jointly affect the optimal design of severance pay.

I demonstrate this in a frictional search and matching model with idiosyncratic productivity shocks. Contracts are written at the beginning of the match and determined by directed search along the lines of Acemoglu and Shimer (1999).[2] Contracting is over a wage schedule and a productivity threshold below which the match is dissolved. Three contractual regimes are compared. In the *full commitment* regime, the parties commit to wages and the separation rule. In the *limited commitment* regime there is only commitment to the wage while employment is at will. In the *incomplete information* regime wages cannot depend on productivity and employment is at will.

The analysis is first carried out in a stationary setting where private contracts as well as mandated severance pay do not depend on tenure. Under risk neutrality, the model confirms previous findings of Lazear (1990) and Boeri et al. (2017) on the efficacy of severance pay. With risk averse workers, I proof two novel results. The optimal size of severance pay is the same for all three contractual regimes if search frictions on the worker's side of the labor market are negligible. Otherwise, the optimal size of severance pay should increase in the severity of the contracting friction. These findings continue to hold with dynamic contracts and tenure-dependent severance pay.

I also demonstrate that risk aversion and incomplete information are not sufficient to rationalize severance pay schedules that are increasing in tenure. However, if the model is extended by a private effort decision of the worker, the optimal tenure-profile of wages and severance

[2] Acemoglu and Shimer (1999) extend the ideas of Shimer (1996) and Moen (1997) to risk averse individuals.

pay are qualitatively in line with the data. This suggests that among the three rationales put forward by Holzmann et al. (2012) and sketched in the first paragraph, ensuring human capital investment is key to understand the dynamic elements of existing severance pay systems.

The essay is structured as follows. Section 3.2 investigates optimal severance pay in a stationary setting. In Section 3.3, the analysis is extended to dynamic contracts. Section 3.4 concludes. Proofs that are short and insightful are included in the text. The remainder is delegated to Section 3.A.

3.2 Stationary contracts

The economy is populated by a mass of individuals that live forever. In each period, an individual is either employed or unemployed. An unemployed individual has a home production of b and searches on the labor market at no cost. Employed individuals engage in market production and earn wage income. There is no on-the-job search, and individuals consume their earnings in every period. They have a strictly concave utility function u that belongs to a parametric family which approaches linear utility in the limit $\eta \to 0$.[3]

Assumption 3.1. *The utility function u belongs to a parametric family of utility functions $\{u_\eta : \eta \in (0, \infty)\}$ where $u' > 0$ and $u'' < 0$, as well as $\lim_{\eta \to 0} u_\eta = \mathrm{id}$.*

Firms are risk neutral and consist of a single job that is either vacant or occupied. Firms can freely enter or leave the market, which implies that the value of a vacant job is zero in equilibrium. The labor market equilibrium is determined by directed search along the lines of Acemoglu and Shimer (1999). Each firm can post a vacancy together with a contract $(w, \underline{y})$. Firms offering the same contract form a submarket. Unemployed individuals costlessly observe all contracts and apply to a submarket where an application yields the highest expected present discounted surplus for them. In any submarket, A applicants and V vacancies are randomly matched by a constant returns to scale matching technology $M(A, V)$. The probability of filling a vacancy is $q(\theta) = \frac{M(A,V)}{V} = M(1/\theta, 1)$, and the probability that an application turns into a match is $p(\theta) = \frac{M(A,V)}{A} = \theta q(\theta)$, where $\theta = V/A$ is the labor market tightness of the submarket. Denote with $\varepsilon(\theta) := -\frac{q'(\theta)\theta}{q(\theta)}$ the elasticity of the vacancy-filling probability.

The productivity of a firm–worker pair is subject to idiosyncratic productivity shocks that are independent over time and arrive with probability ϕ. If no shock hits in a given period, the productivity of the previous period prevails. To simplify the exposition, the productivity distribution is assumed to be continuous with support on the whole real line. The cdf is denoted by F.

[3] Also severance pay is consumed within a single period. Depending on the size of severance pay, this can give rise to a very unbalanced consumption profile during periods of unemployment. Since individuals are risk averse, this has adverse welfare consequences. In Section 3.2.4 I generalize the model by allowing individuals to smooth the consumption of severance pay over several periods. While the qualitative predictions continue to hold, the extended model is no longer analytically tractable. The main analysis is therefore carried out under the more restrictive assumption that individuals always live hand-to-mouth.

The employment contract $(w, \underline{y})$ that worker and firm sign at the beginning of the match contains a potentially productivity-contingent wage schedule $w : \mathbb{R} \to \mathbb{R}$ and a separation threshold $\underline{y}$. The match ends endogenously if productivity falls below $\underline{y}$. With probability $\delta \in [0, 1)$, the match ends for exogenous reasons. Firm and worker discount the future with the common discount factor $\beta \in (0, 1)$.

Consider period $t \geq 0$ of the match. If period t productivity is above the separation threshold, $y \geq \underline{y}$, the firm's value of employment is

$$J_t(w, y) = y - w(y) + \tilde{\beta}[\phi \mathbb{E} J_{t+1}(w, \underline{y}) + (1 - \phi) J_{t+1}(w, y)],$$

where $\tilde{\beta} = \beta(1 - \delta)$ is the effective discount rate. It equals the instantaneous profit plus the continuation value, which takes into account that a new productivity is drawn with probability ϕ. Since the value of a vacancy is zero, $J_t(w, y)$ is also the firm surplus over non-employment. Likewise, the value of employment for the worker is

$$N_t(w, y) = u(w(y)) + \tilde{\beta}[\phi \mathbb{E} N_{t+1}(w, \underline{y}) + (1 - \phi) N_{t+1}(w, y)].$$

The value of unemployment is $U_t = u(b) + \tilde{\beta}(U_{t+1} + v)$ where v is the maximal value attainable by sending a job application. This is endogenously determined in equilibrium but taken as given by the worker. Worker surplus of employment over unemployment is $W_t := N_t - U_t$ and satisfies

$$W_t(w, y, v) = u(w(y)) - u(b) + \tilde{\beta}[\phi \mathbb{E} W_{t+1}(w, \underline{y}, v) + (1 - \phi) W_{t+1}(w, y, v) - v].$$

For $t \geq 0$, the expressions $\mathbb{E} J_t(w, \underline{y})$ and $\mathbb{E} W_t(w, \underline{y}, v)$ refer to expected firm and worker surplus if a new productivity is drawn in period t. Assuming that the government mandates severance pay P for any endogenous separation, expected surplus is $\mathbb{E} J_t(w, \underline{y}) = \int_{\underline{y}}^{\infty} J_t(w, y)\, dF(y) - F(\underline{y})P$ for the firm and $\mathbb{E} W_t(w, \underline{y}, v) = \int_{\underline{y}}^{\infty} W_t(w, y, v)\, dF(y) + F(\underline{y})\Delta$ for the worker, where $\Delta := u(b+P) - u(b)$. If the new draw is above the threshold $\underline{y}$ the match is productive and the agents earn the surplus defined above. Otherwise a separation takes place and the firm has to pay P to the worker. The value Δ is the worker surplus of a separation over unemployment. To see this, note that the worker's value in case of a separation in period t is $L_t = u(b + P) + \tilde{\beta}[U_{t+1} + v]$. Since severance pay only accrues in the period of the separation and worker live hand-to-mouth, the continuation value is the same as for a worker who has been unemployed throughout the whole period. This implies $\Delta = L_t - U_t = u(b + P) - u(b)$.[4]

[4]Section 3.2.4 provides a generalization of the model where severance pay need not be consumed all at once. The quantitative predictions, however, remain the same. Furthermore, in real labor markets, the eligibility to severance pay typically depends on whether the separation is initiated by the employer (layoff) or initiated by the worker (quit). For the present analysis it is safe to ignore this differentiation since contracting $w(y) = \infty$ for $y < \underline{y}$ always makes the firm lay off the worker below the contracted separation threshold without affecting surplus functions. If wages cannot depend on productivity, it turns out that all endogenous separations that happen in equilibrium are actually layoffs.

3.2.1 Exogenous severance pay

This section characterizes the labor market equilibrium for given severance pay P. The results serve as the basis for optimal policy considered in Section 3.2.2. In both sections, I focus on stationary equilibria where not only the decision variables but also the value functions are constant over time. Expected surplus is then

$$\mathbb{E}J(w,\underline{y}) = \frac{\int_{\underline{y}}^{\infty} y - w(y)\, dF(y) - (1-\tilde{\beta}(1-\phi))F(\underline{y})P}{1-\tilde{\beta}(1-\phi F(\underline{y}))}, \tag{3.1}$$

$$\mathbb{E}W(w,\underline{y},v) = \frac{\int_{\underline{y}}^{\infty} u(w(y)) - u(b) - \tilde{\beta}v\, dF(y) + (1-\tilde{\beta}(1-\phi))F(\underline{y})\Delta}{1-\tilde{\beta}(1-\phi F(\underline{y}))}. \tag{3.2}$$

The formal equilibrium definition follows Acemoglu and Shimer (1999).

Definition 3.1 (Stationary Equilibrium)**.** *A stationary equilibrium with severance pay P consists of a function $\Theta^*(w,\underline{y}) \geq 0$, a separation threshold $\underline{y}^* \in \mathbb{R}$, a (Lebesgue measurable) wage schedule $w^* : [\underline{y}^*, \infty) \to \mathbb{R}$, and a value $v^* \geq 0$ such that*

- *firms maximize profit under free entry, $q(\Theta^*(w,\underline{y}))\mathbb{E}J(w,\underline{y}) \leq c$ for all $(w,\underline{y})$, with equality for $(w^*,\underline{y}^*)$,*
- *job seekers apply optimally, $v^* \geq p(\Theta^*(w,\underline{y}))\mathbb{E}W(w,\underline{y},v^*)$ for all $(w,\underline{y})$ and $\Theta^*(w,\underline{y}) \geq 0$ with complementary slackness, where $v^* = p(\Theta^*(w^*,\underline{y}^*))\mathbb{E}W(w^*,\underline{y}^*,v^*)$.*

Along the lines of directed search, the equilibrium specifies the market tightness Θ^* for all feasible pairs $(w,\underline{y})$. This implies that no agent has an incentive to deviate from the equilibrium $(w^*,\underline{y}^*)$. In the following, two directed search equilibria are regarded as identical if they only differ in their off-equilibrium prescriptions.

Definition 3.2. *Two equilibria are identical if they give rise to the same equilibrium objects $(\theta^*, w^*, \underline{y}^*, v^*)$ where $\theta^* = \Theta^*(w^*,\underline{y}^*)$.*

As demonstrated by Acemoglu and Shimer (1999), the equilibrium defined above can be characterized by the constrained optimization problem

$$V(v) = \max_{(\theta,w,\underline{y})} p(\theta)\mathbb{E}W(w,\underline{y},v) \quad \text{s.t.} \quad q(\theta)\mathbb{E}J(w,\underline{y}) = c, \tag{3.3}$$

together with the equilibrium condition

$$V(v) = v. \tag{3.4}$$

To explore the effect of contractual frictions on the equilibrium, I add additional constraints to the contracting problem (3.3) that reflect realistic limitations of the contract space:

- Solving (3.3)–(3.4) without additional constraints gives rise to the *full commitment* (FC) solution. Firm and worker commit to the wage contract w and to the separation rule that the match breaks up endogenously if and only if $y < \underline{y}$.
- Under *limited commitment* (LC), the parties commit to the wage contract w, but the separation decision is taken privately. For the negotiated separation rule to be self-enforcing, the additional constraint $J(w, y) + P \geq 0$ for $y \geq \underline{y}$ is imposed in (3.3).
- As a third scenario, I assume that employment contracts cannot depend on y at all. This corresponds to the additional constraints $w(y) = w$ for $y \geq \underline{y}$ and $J(w, \underline{y}) + P = 0$ in (3.3). This type of friction has already been analyzed in connection with severance pay by Alvarez and Veracierto (2001) and others. Since it can arise from asymmetric information about match productivity, I refer to this scenario as *incomplete information* (II).[5]

Full commitment. The Lagrangian for the optimization problem (3.3) under full commitment is $\mathcal{L}^{FC} = p(\theta)\mathbb{E}W(w, \underline{y}, v) + \lambda[q(\theta)\mathbb{E}J(w, \underline{y}) - c]$. Differentiation with respect to $w(y)$ reveals that the optimal contract provides full insurance, $w^*(y) = w^*$ for $y \geq \underline{y}^*$. The optimal constant wage w^*, the separation threshold $\underline{y}^*$, and the tightness θ^* satisfy

$$\frac{1 - \varepsilon(\theta^*)}{\varepsilon(\theta^*)} \frac{\mathbb{E}W(w^*, \underline{y}^*, v)}{\mathbb{E}J(w^*, \underline{y}^*)} = u'(w^*), \tag{3.5}$$

$$\frac{W(w^*, \underline{y}^*, v) - \Delta}{u'(w^*)} + J(w^*, \underline{y}^*) + P = 0, \tag{3.6}$$

together with the free entry condition $q(\theta^*)\mathbb{E}J(w^*, \underline{y}^*) = c$. Equation (3.5) implies that the worker in equilibrium receives a share $\varepsilon(\theta^*)$ of expected joint surplus $\mathbb{E}W^*/u'(w^*) + \mathbb{E}J^*$. This is a generalization of the familiar Nash sharing rule to the case of risk averse workers. Equation (3.6) shows that the match should be terminated if and only if the joint surplus of employment, $W^*/u'(w^*) + J^*$ falls short of the joint surplus of a separation, $\Delta/u'(w^*) - P$.

Limited commitment. Under limited commitment, the Lagrangian of the optimization problem is $\mathcal{L}^{LC} = p(\theta)\mathbb{E}W(w, \underline{y}, v) + \lambda[q(\theta)\mathbb{E}J(w, \underline{y}) - c] + \int_{\underline{y}}^{\infty} \mu(y)[J(w, y) + P]\, dF(y)$. The optimal wage schedule turns out to be piecewise linear, $w^*(y) = \min\{w^*, y + \tilde{\beta}\phi\mathbb{E}J(w^*, \underline{y}^*) + (1 - \tilde{\beta}(1 - \phi))P\}$ for $y \geq \underline{y}^*$. The remaining first order conditions can be summarized as

$$\frac{1 - \varepsilon(\theta^*)}{\varepsilon(\theta^*)} \frac{\mathbb{E}W(w^*, \underline{y}^*, v)}{\mathbb{E}J(w^*, \underline{y}^*)} = u'(w^*) - \frac{\tilde{\beta}\phi}{1 - \tilde{\beta}(1 - \phi F(\underline{y}^*))} \int_{\underline{y}^*}^{\infty} u'(w^*(y)) - u'(w^*)\, dF(y), \tag{3.7}$$

$$\frac{W(w^*, \underline{y}^*, v) - \Delta}{u'(w^*(\underline{y}^*))} + J(w^*, \underline{y}^*) + P = 0, \tag{3.8}$$

[5] This can be justified as follows. Suppose that the realized productivity draw is private knowledge of the firm. An employer can increase her own profit by making the worker agree on a wage cut. This creates an innate incentive to cheat on the worker and pretend that a wage cut is required to a prevent a layoff, even if this is not the case. A rational worker anticipates the employer's motives and does not believe claims about productivity.

together with the free entry condition. If firms cannot commit to a separation rule, the optimal wage is no longer a constant but a piecewise linear function of y. For good productivity realizations, the worker earns the constant wage w^*. If productivity is sufficiently bad, a wage decrease might be necessary to meet the firm's layoff constraint $J(w,y)+P=0$. The productivity threshold y^+ below which this constraint binds is defined by $y^+ = w^* - \tilde{\beta}\phi\mathbb{E}J(w^*, \underline{y}^*) - (1-\tilde{\beta}(1-\phi))P$. The optimal solution under limited commitment is therefore characterized by two thresholds: the separation threshold $\underline{y}^*$ and the profitability threshold y^+. For $y < \underline{y}^*$ the match ends, for $y \in [\underline{y}^*, y^+]$ the match continues and worker earns the whole match surplus. For $y > y^+$, the worker earns the constant wage w^* and the firm earns a positive rent. Although the worker earns the whole match surplus in certain states of the world, condition (3.7) shows that the worker's share in *expected* match surplus is lower than under full commitment. The reason is that marginally increasing w^* raises y^+, which reduces the set of productivity states where consumption smoothing is possible. This additional marginal cost is captured by the second term on the left-hand side of (3.7) and lowers the worker's surplus share below $\varepsilon(\theta^*)$ if $\underline{y}^* < y^+$. If $\underline{y}^* \geq y^+$, by contrast, the layoff constraint does not bind at $y = \underline{y}^*$ and the first order conditions reveal that the optimal contracts under full and limited commitment coincide.[6]

Incomplete information. The Lagrangian of the optimization problem under incomplete information is $\mathcal{L}^{II} = p(\theta)\mathbb{E}W(w,\underline{y},v) + \lambda[q(\theta)\mathbb{E}J(w,\underline{y}) - c] + \mu[J(w,\underline{y}) + P]$. The necessary first order conditions are

$$\frac{1-\varepsilon(\theta^*)}{\varepsilon(\theta^*)}\frac{\mathbb{E}W(w^*,\underline{y}^*,v)}{\mathbb{E}J(w^*,\underline{y}^*)} = u'(w^*) - \frac{f(\underline{y}^*)}{1-F(\underline{y}^*)}\frac{(1-\tilde{\beta}(1-\phi))^2}{1-\tilde{\beta}(1-\phi F(\underline{y}^*))}[W(w^*,\underline{y}^*,v) - \Delta], \quad (3.9)$$

$$J(w^*,\underline{y}^*) + P = 0, \quad (3.10)$$

together with the free entry condition. Equation (3.9) reveals that the worker receives a share in expected joint surplus below $\varepsilon(\theta^*)$ if the she values employment more than a separation, $W(w^*,\underline{y}^*,v) > \Delta$. The reason is that a higher wage w^* increases the separation threshold $\underline{y}^*$ determined by (3.10), which lowers the worker's retention probability. The worker hence faces a trade-off between wage and retention probability and is willing to give up some of the former for the latter.

3.2.2 Optimal severance pay

Assume that a social planner seeks to maximize lifetime utility of the individuals by setting the level of severance pay P, taking into account the optimal decision rules of firms and workers described above. Denote with $V(v,P)$ the value function (3.3) given search value v and severance pay level P. Assume that $V(v,P)$ is continuously differentiable in both arguments. This implies that for given P, the equilibrium value of search v^* is unique. The reason is that $V(v,P)$ is then continuous and strictly decreasing in v, such that equation (3.4) has at most one solution. This

[6]Whether y^+ is above or below $\underline{y}^*$ depends on the parameterization of the model.

solution is henceforth denoted by $v^*(P)$. For given P, lifetime utility of an individual is

$$\mathcal{W}(P) = \mathbb{E}_0\Big[\sum_{t=0}^{\infty}\beta^t u(c_t)\Big] = \frac{u(b) + v^*(P)}{1-\beta},$$

assuming that workers start off unemployed. Clearly, the planner also has other options to increase welfare in the economy, such as unemployment benefits that act as an increase in b. As argued by Blanchard and Tirole (2008), however, unemployment benefits alone cannot counter the welfare loss of the contracting frictions considered here, while severance payments can achieve this under certain conditions. The analysis therefore focuses on optimal severance pay.[7]

Optimality conditions

Clearly, maximizing $\mathcal{W}(P)$ with respect to P is equivalent to finding the highest search value $v^*(P)$. If v^* is continuously differentiable in P, the necessary condition for optimal policy is $\frac{dv^*(P)}{dP} = 0$. Since the labor market equilibrium is determined by directed search, this derivative is easy to evaluate. By the fixed point condition (3.4) any equilibrium must satisfy $V(v^*(P), P) = v^*(P)$. Implicit differentiation with respect to P yields

$$\frac{dv^*(P)}{dP} = \frac{\frac{\partial V(v^*(P),P)}{\partial P}}{1 - \frac{\partial V(v^*(P),P)}{\partial v}}. \tag{3.11}$$

To evaluate the partial derivatives on the right-hand side, the envelope theorem can be applied to the Lagrangians $\mathcal{L}$ given in the previous section. This important observation is summarized in the following lemma which is proven in Section 3.A.

Lemma 3.1. *For any P the sign of $\frac{dv^*(P)}{dP}$ equals the sign of $\frac{\partial \mathcal{L}}{\partial P}$ with the Lagrangian $\mathcal{L}$ given in Section 3.2.1 and evaluated in equilibrium. The first order condition for socially optimal severance pay is $\frac{\partial \mathcal{L}}{\partial P} = 0$.*

Fella and Tyson (2013) consider severance pay as part of the compensation package that firm and worker negotiate. In the framework studied here, it is irrelevant whether severance pay is part of private employment contracts or optimally set by a government. This is evident from Lemma 3.1 since $\frac{\partial \mathcal{L}}{\partial P} = 0$ is just the additional first order condition that arises if severance pay were part of the contracting problem. Due to directed search, the agents internalize the search and matching externalities that they exert on the other agents in the economy. As a consequence, the privately optimal level of severance pay (satisfying $\frac{\partial \mathcal{L}}{\partial P} = 0$) equals the socially optimal level of severance pay (satisfying $\frac{dv^*(P)}{dP} = 0$).

[7]The interplay between unemployment benefits and severance pay has partly been analyzed in the first essay of this thesis. The lower the unemployment benefit (which took the form an early retirement benefit), the larger the efficiency loss caused by the contracting friction, and the larger the necessary intervention by severance pay. Blanchard and Tirole (2008) elaborate on the optimal provision of both severance pay and unemployment benefits in a one-period model.

Applying Lemma 3.1, the welfare maximizing level P^* under full commitment satisfies

$$\frac{1-\varepsilon(\theta^*)}{\varepsilon(\theta^*)}\frac{\mathbb{E}W(w^*,\underline{y}^*,v^*)}{\mathbb{E}J(w^*,\underline{y}^*)} = u'(b+P^*). \tag{3.12}$$

Combining this with (3.5) reveals that $w^* = b + P^*$ if individuals are risk averse. Optimal severance pay is such that the worker does not experience any change in income in the period of the separation. Nevertheless, Proposition 3.5 below shows that the worker is worse off in present discounted value terms, $W(w^*,\underline{y}^*,v^*) > \Delta^*$.

The first order condition under limited commitment is

$$\frac{1-\varepsilon(\theta^*)}{\varepsilon(\theta^*)}\frac{\mathbb{E}W(w^*,\underline{y}^*,v^*)}{\mathbb{E}J(w^*,\underline{y}^*)} = u'(b+P^*)+\frac{1-\tilde{\beta}}{1-\tilde{\beta}(1-\phi F(\underline{y}^*))}\frac{\int_{\underline{y}^*}^{\infty} u'(w^*(y))-u'(w^*)\,dF(y)}{F(\underline{y}^*)}. \tag{3.13}$$

The last term on the right-hand side reflects that *ceteris paribus* severance pay decreases the profitability threshold of the firm and thereby increases the range of productivity realizations for which consumption smoothing is possible.

Assuming incomplete information yields the condition for optimal policy

$$\frac{1-\varepsilon(\theta^*)}{\varepsilon(\theta^*)}\frac{\mathbb{E}W(w^*,\underline{y}^*,v^*)}{\mathbb{E}J(w^*,\underline{y}^*)} = u'(b+P^*)+\frac{f(\underline{y}^*)}{F(\underline{y}^*)}\frac{(1-\tilde{\beta})(1-\tilde{\beta}(1-\phi))}{1-\tilde{\beta}(1-\phi F(\underline{y}^*))}[W(w^*,\underline{y}^*,v^*)-\Delta^*]. \tag{3.14}$$

The last term on the right-hand side reflects that *ceteris paribus* higher severance pay increases the continuation value of the firm which reduces the worker's layoff probability by (3.10).

The remainder of this section compares and contrasts the socially optimal level of severance pay between the three contractual regimes. I start by summarizing results for risk neutral individuals, which have partly been derived in the literature already. I then turn to risk averse individuals.

Optimal severance pay with risk neutral individuals

The following results apply to the limiting case $\eta \to 0$ in which workers become risk neutral, $u(x) = x$.[8] The first proposition verifies the neutrality of severance pay that has been highlighted by Lazear (1990) if agents are risk neutral and contracts are complete. The neutrality result remains valid if firms can only commit to wages but not to a separation rule.

Proposition 3.1. *Let individuals be risk neutral ($\eta \to 0$). Apart from the wage, the directed search equilibrium is independent of P under full and limited commitment.*

The proof is given in Section 3.A and relies on the fact that for any equilibrium $(\theta^*, w_P^*, \underline{y}^*, v^*)$ with severance pay P, there exists a wage schedule w_0^* such that $(\theta^*, w_0^*, \underline{y}^*, v^*)$ is an equilibrium for $P = 0$. Hence as long as wages can be productivity-contingent, the effect of any legally imposed level of severance pay can be undone by private wage arrangements. In equilibrium,

[8] The limit argument is necessary since strict concavity of the utility function is required to derive $w^*(y) = w^*$ and similar results. Only the expected wage is pinned down when linear utility is assumed from the outset.

neither expected firm profits nor worker welfare depend on P. If wages cannot depend on productivity and there is no commitment to a separation rule, however, severance pay does have a welfare effect.

Proposition 3.2. *Let individuals be risk neutral ($\eta \to 0$). With incomplete information, optimal severance pay is strictly positive, $P^* > 0$, and satisfies $W(w^*, \underline{y}^*, v^*) = \Delta^*$. The resulting labor market equilibrium coincides with the equilibrium under full and limited commitment.*

Proof. With linear utility, the optimality conditions (3.9) and (3.14) imply $W(w^*, \underline{y}^*, v^*) = \Delta^*$. The first order conditions (3.9)–(3.10) then coincide with (3.5)–(3.6). Since (3.9) implies $\mathbb{E}W^* > 0$ and $\mathbb{E}W^* = \Delta^* = u(b + P^*) - u(b)$, optimal severance pay is strictly positive. □

Since $W(w^*, \underline{y}^*, v^*) = \Delta^*$, the optimal level of severance pay makes the worker indifferent between work and separation. She therefore enjoys perfect insurance against job loss. As a result, the trade-off between wage income and the retention rate vanishes in (3.9). The optimal policy then restores the efficient hiring and firing probabilities at the same time. This neat property of severance pay has previously been shown by Boeri et al. (2017). However, it is demonstrated below that it no longer applies if workers are risk averse.

Optimal severance pay with risk averse individuals

If individuals are risk averse ($\eta > 0$), severance pay is not neutral even if contracts are complete. This is because a positive level of severance pay allows more consumption smoothing across states of the world. The first important observation is that optimal severance pay is always strictly positive, $P^* > 0$. The proof can be found in Section 3.A.

Proposition 3.3. *Let individuals be risk averse. Then optimal severance pay is strictly positive in all three contractual regimes.*

How does the contractual regime affect the optimal level of severance pay? With risk neutral agents, Proposition 3.2 establishes that optimal severance pay is independent of the contractual regime. Implementing a level that sets workers indifferent between work and separation, $W(w^*, \underline{y}^*, v^*) = \Delta^*$, achieves the same welfare in all three cases studied. The same independence result applies to risk averse agents, provided that workers do not face search frictions in equilibrium, $p(\theta^*) = 1$. The reason behind this observation is that irrespective of the contractual framework, $W(w^*, \underline{y}^*, v^*) - \Delta^*$ is proportional to $u(w^*) - u(b + P^*) + \tilde{\beta}(1 - p(\theta^*))\Delta^*$. The worker's valuation of continued employment over a job loss combines the immediate change in utility and the expected utility change in future periods, where $1 - p(\theta^*)$ reflects the delay the worker faces in finding a new job. For $p(\theta^*) = 1$, the worker finds a new job immediately, such that optimal severance pay in all three regimes satisfies $W(w^*, \underline{y}^*, v^*) = \Delta^*$ and $w^* = b + P^*$.

Proposition 3.4. *Let individuals be risk averse, and assume that $p(\theta^*) = 1$ in an equilibrium with optimal severance pay. Then optimal severance pay P^* is identical in all three contractual regimes and satisfies $W(w^*, \underline{y}^*, v^*) = \Delta^*$. The resulting labor market equilibria are identical.*

Proof. Under the condition imposed, $W(w^*, \underline{y}^*, v^*) - \Delta^* = \frac{u(w^*) - u(b+P^*)}{1-\tilde{\beta}(1-\phi)F(\underline{y}^*)}$. It is then easy to verify that the triple $(w^*, \underline{y}^*, P^*)$ that satisfies $w^* = b + P^*$, $J(w^*, \underline{y}^*) + P^* = 0$, and $\frac{1-\varepsilon(\theta^*)}{\varepsilon(\theta^*)} \frac{\Delta^*}{\mathbb{E}J(w^*, \underline{y}^*)} = u'(w^*)$ solves the first order equations of all three contractual regimes and gives rise to the same θ^* and v^*. □

Proposition 3.4 presents a very strong result. It implies that a welfare-maximizing policy maker can be agnostic about the imperfections of private employment contracts. The optimal policy as well as the resulting labor market equilibrium do not depend on the severity of commitment problems and informational frictions at the micro level. While $p(\theta^*) = 1$ is unlikely to hold in standard calibrations, Proposition 3.4 suggests that the severity of search frictions on the worker's side determines how strongly severance pay should take into account contractual imperfections. This is confirmed by the numerical experiments of Section 3.2.3.

Let me now turn to the general case where workers face search frictions in equilibrium, $p(\theta^*) < 1$. Optimal severance pay is then such that the worker strictly prefers to stay employed.

Proposition 3.5. *Let individuals be risk averse, and let $p(\theta^*) < 1$. Then optimal severance pay is such that $W^*(w^*, \underline{y}^*, v^*) > \Delta^*$ under full commitment and incomplete information.*

Proof. In equilibrium $W(w^*, \underline{y}^*, v^*) - \Delta^* = \frac{u(w^*) - u(b+P^*) + \tilde{\beta}(1-p(\theta^*))\Delta^*}{1-\tilde{\beta}[1-\phi F(\underline{y}^*) - p(\theta^*)(1-F(\underline{y}^*))]}$. With full commitment, conditions (3.5) and (3.12) imply $w^* = b + P^*$. Since $\Delta^* > 0$ by Proposition 3.3, we have $W(w^*, \underline{y}^*, v^*) > \Delta^*$. With incomplete information, the statement is shown by contradiction. Assume that $W(w^*, \underline{y}^*, v^*) - \Delta^* = 0$. Then conditions (3.9) and (3.14) imply $w^* = b + P^*$. As above, this implies $W(w^*, \underline{y}^*, v^*) > \Delta^*$ and therefore leads to a contradiction. □

In contrast to risk neutrality, optimal severance pay with risk averse individuals is—apart from the special case considered in Proposition 3.4—such that workers enjoy a positive surplus of employment over a job loss with severance pay. Insofar as there is moral hazard on the side of the worker (a possibility that is not reflected by the baseline version of the model), a higher gap in the valuation of employment relative to separation reduces the incentive to shirk. If $W(w^*, \underline{y}^*, v^*) = \Delta^*$, workers could be tempted to reduce work effort and force a layoff unless courts are able to distinguish between economic dismissals and disciplinary dismissals for which typically no severance pay accrues. Section 3.3.3 explicitly considers the possibility of shirking in an environment with dynamic wage contracts.[9]

Under certain conditions, the optimal levels of severance pay can be compared between the full commitment and the incomplete information scenario. In the following, superscript II denotes equilibrium values under incomplete information, while superscript FC refers to the full commitment equilibrium. The proof is delegated to Section 3.A.

[9] The possibility of imperfect verifiability has been considered by Galdón-Sánchez and Güell (2003) and Boeri et al. (2017). In both studies, risk neutral workers are in equilibrium indifferent between shirking and working honestly, irrespective of the detection probability of courts.

Proposition 3.6. *Let individuals be risk averse. Assume that the matching function is Cobb-Douglas, and that for $P = P^{II}$ the inequality $p(\theta^{FC}) \leq p(\theta^{II}) < 1$ holds. Then optimal severance pay satisfies $P^{II} > P^{FC} > 0$, and the associated welfare levels satisfy $v^{II} < v^{FC}$.*[10]

Proposition 3.6 establishes that in general a larger amount of severance pay is necessary to maximize welfare in the presence of incomplete information. The maximum attainable welfare level, however, remains lower. Hence the result of Proposition 3.2 that optimal severance pay can remove the labor market distortions of the contracting friction is not valid for risk averse agents. Optimal government policy should explicitly take into account the severity of contracting frictions in the private sector.

How do the above results concerning risk averse individuals compare to the existing literature? Blanchard and Tirole (2008) as well as Fella and Tyson (2013) study the same friction that I impose in the incomplete information scenario. Both authors find that optimal severance pay fully insures risk averse workers against job loss. This is an apparent contradiction to Proposition 3.5. Interestingly, the reasons behind this discrepancy are very different for the two papers. Blanchard and Tirole (2008) study a one period model, in which by design the costs of a separation are limited to the immediate utility loss. As highlighted before Proposition 3.4, a separation also decreases future expected utility if it takes time for the worker to find a new job. This option value of employment, however, is absent in a one period model. The same option value is also absent in my model when I set $p(\theta^*) = 1$, which resulted in Proposition 3.4. As Proposition 3.4 predicts full insurance, it agrees with the observation of Blanchard and Tirole (2008). The second comparable paper, Fella and Tyson (2013), studies a dynamic model which is in many respects identical to mine. Yet, the authors additionally allow workers to self-insure against job loss by buying and selling a risk-less bond. In contrast to Proposition 3.5, the authors find that optimal severance pay fully insures workers against job loss if individuals exhibit CARA utility and individuals face no borrowing constraints. While these assumptions buy analytical tractability, they imply that in equilibrium the receipt of severance pay hardly increases consumption within the same quarter. I discuss this point in more detail in Section 3.2.4 and argue that the insights of Proposition 3.5 and Proposition 3.6 remain valid in empirically plausible settings.

3.2.3 Numerical examples

To illustrate the theoretical results and provide some quantification, I conduct several numerical exercises. As a baseline, I rely largely on the parameterization of Menzio et al. (2016), who calibrate a similar model to the US economy at a monthly frequency. The discount factor reflects an annual interest rate of 4 percentage points, $\beta = 0.9967$. The probability for an exogenous separation δ is 0.002, and a new productivity draw occurs with probability 0.0094, which is roughly every 8.5 years. The productivity distribution is assumed to be normal with

[10]The assumption concerning the job-finding probability imposed in Proposition 3.6 is very likely to be satisfied in realistic calibrations, as evident from the numerical experiments of Section 3.2.3.

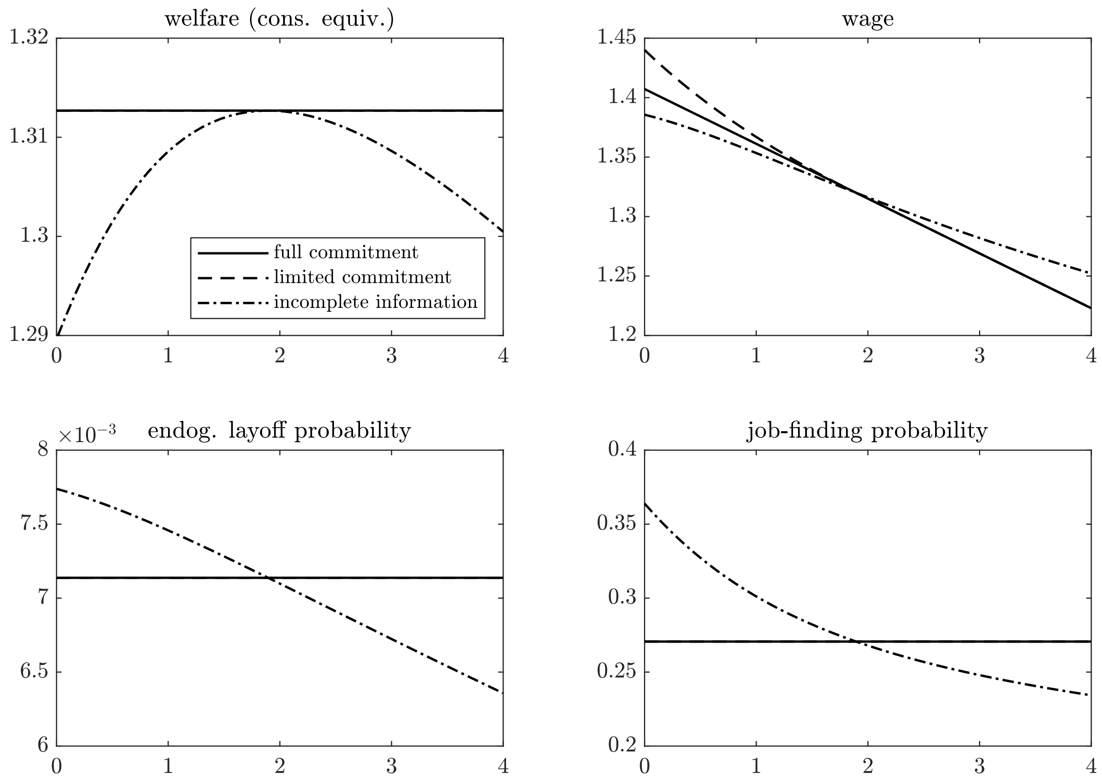


Figure 3.1. effect of severance pay (on the horizontal axis) with risk neutral individuals

mean 1. The standard deviation σ is 0.39 to reflect a productivity ratio between the 90th and the 10th percentile of the distribution of 3 (Menzio et al., 2016). The value of home production b is taken to be 0.8 and the vacancy posting cost c is 7. The matching function is Cobb-Douglas, $p(\theta) = \min\{A\theta^{1/2}, 1\}$ with $A = 1$. With risk neutrality and full commitment, the labor market equilibrium gives rise to an endogenous layoff probability of 0.71 percent and a job-finding probability of 27 percent.

Risk neutrality. Figure 3.1 demonstrates the neutrality of severance pay when individuals are risk neutral and information is symmetric. The welfare measure reported in the upper left pattern of the figure is the equivalent constant consumption flow that attains the same discounted lifetime utility $\mathcal{W}(P)$. It is calculated as $c(P) = u^{-1}\big((1-\beta)\mathcal{W}(P)\big)$. In line with Proposition 3.1, under full or limited commitment, welfare is independent of the level of severance pay. Furthermore, the figure shows that whether firms can commit to the separation rule neither affects welfare, nor layoff and job-finding probabilities. Under incomplete information, by contrast, severance pay does have a welfare effect. Without severance pay, the layoff rate is above the constrained efficient level. Also the job-finding probability is higher, because individuals react to the informational friction by contracting lower wages, which triggers more vacancy posting of the firm (see also Chapter 1 of this thesis). Severance pay reduces both margins at the same time. Welfare peaks at a level of $P^* = 1.89$, where it equals the welfare levels of full and limited commitment. This amounts to 1.44 monthly wages. Optimal severance pay

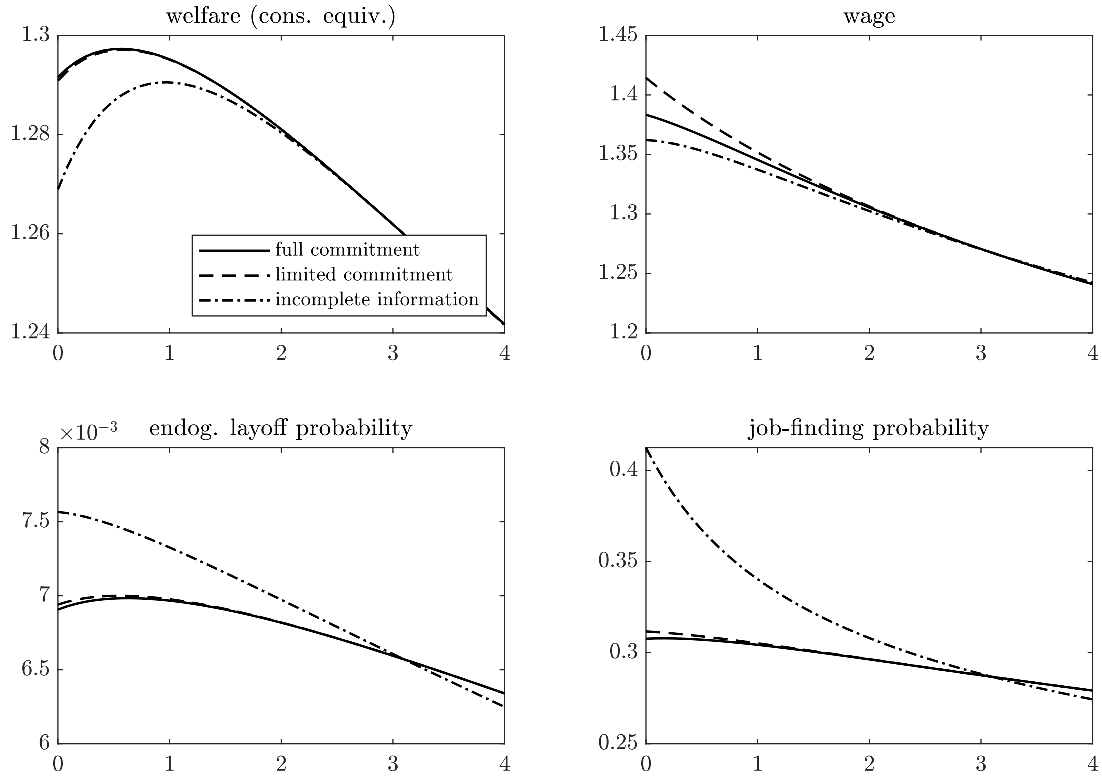


Figure 3.2. effect of severance pay (on the horizontal axis) with risk averse individuals

completely undoes the distortions in hiring and firing caused by the informational friction, as postulated by Proposition 3.2.

Risk aversion. Figure 3.2 repeats the exercise for risk averse individuals with logarithmic utility. Severance pay is not neutral in any of the three regimes. The welfare-maximizing levels are 0.56 under full commitment and 0.58 under limited commitment. Optimal severance pay with incomplete contracts is 0.96. In line with Proposition 3.6, optimal severance pay increases with the severity of the contractual friction. Moreover, Figure 3.2 shows that the welfare improvement due to severance pay is highest with incomplete information. Relative to $P = 0$, severance pay can increase welfare in consumption equivalents by 1.7% under incomplete information, although its optimal level corresponds to only 0.7 monthly wages. The attainable welfare gain is much lower under full commitment (0.44%) and limited commitment (0.48%).[11]

Comparing to Figure 3.1 suggests that optimal severance pay under incomplete information is decreasing in the degree of risk aversion. The reason is that higher risk aversion lowers the equilibrium wage. This reduces the layoff probability and raises the job-finding probability. Altogether, the individuals require less insurance against job loss in the form of severance pay.

[11] Alvarez and Veracierto (2001) also conduct a welfare analysis of severance pay in a calibrated model of the United States. They document that welfare continues to increase even if severance pay replaces a full year of wage income. However, their result is driven by decreasing search costs rather than increasing utility of consumption. The search cost channel is absent in my model since unemployed individuals search freely. Another difference to their model is that I assume hand-to-mouth consumers. This is relaxed in Section 3.2.4, resulting in a modest increase of optimal severance pay to 1.2 monthly wages.

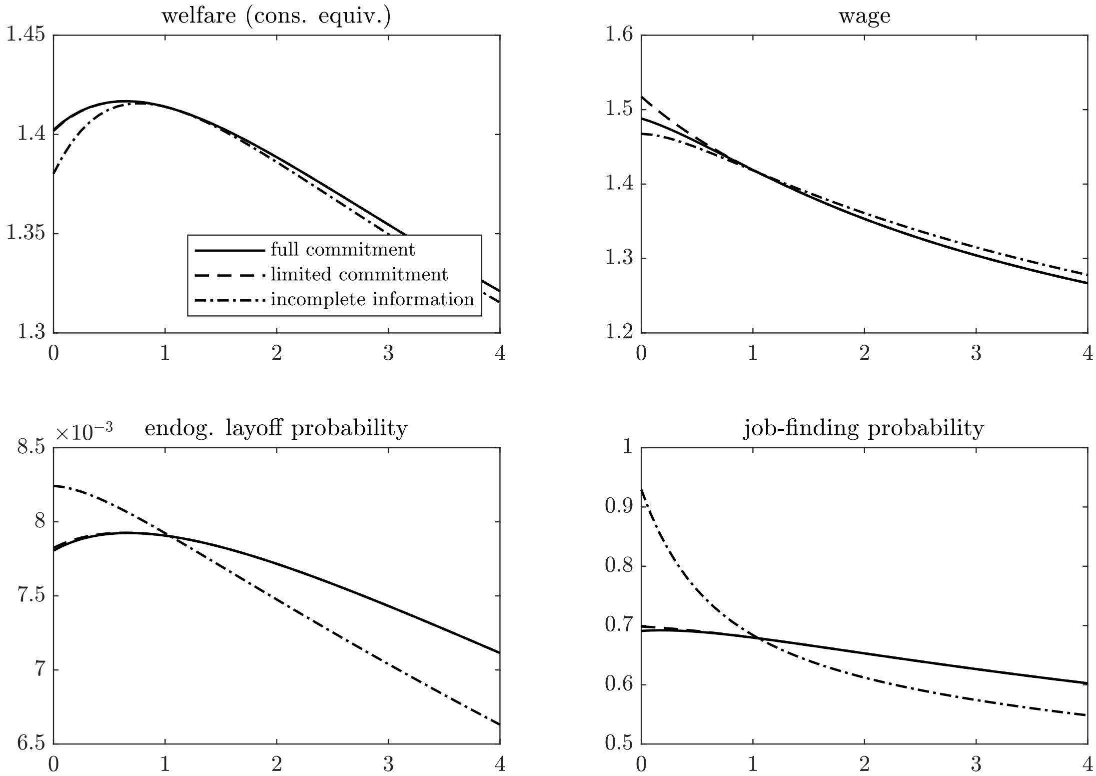


Figure 3.3. effect of severance pay (on the horizontal axis) with high matching productivity ($A = 2$)

It is important to note that with optimal severance pay, the layoff and the job-finding probabilities under incomplete information exceed the levels of the other two regimes. While reducing the employment distortions towards zero is possible, it is not optimal in terms of welfare. The reason is that with increasing severance pay, the consumption profile becomes more unbalanced.

Search frictions. Figure 3.3 and Figure 3.4 illustrate the finding of Proposition 3.4. For Figure 3.3, the matching technology is raised from $A = 1$ to 2. This more than doubles the job-finding probabilities relative to the baseline. Although the values are far below $p(\theta^*) = 1$, optimal severance pay under the three contractual regimes is very similar. The particular values are 0.65 for full and limited commitment, and 0.77 for incomplete information. The difference in welfare is virtually negligible. Further increasing A to 2.7 pushes the job-finding probabilities close to 1, see Figure 3.4. In line with Proposition 3.4, optimal severance pay then eliminates the distortions in hiring and firing caused by the informational friction. These observations suggest that the higher the job-finding probability, the less relevant is the contractual framework for optimal policy making. The longer displaced workers require to get back to employment, however, the more costly are excess layoffs caused by contracting frictions for society and the higher is the optimal policy intervention.

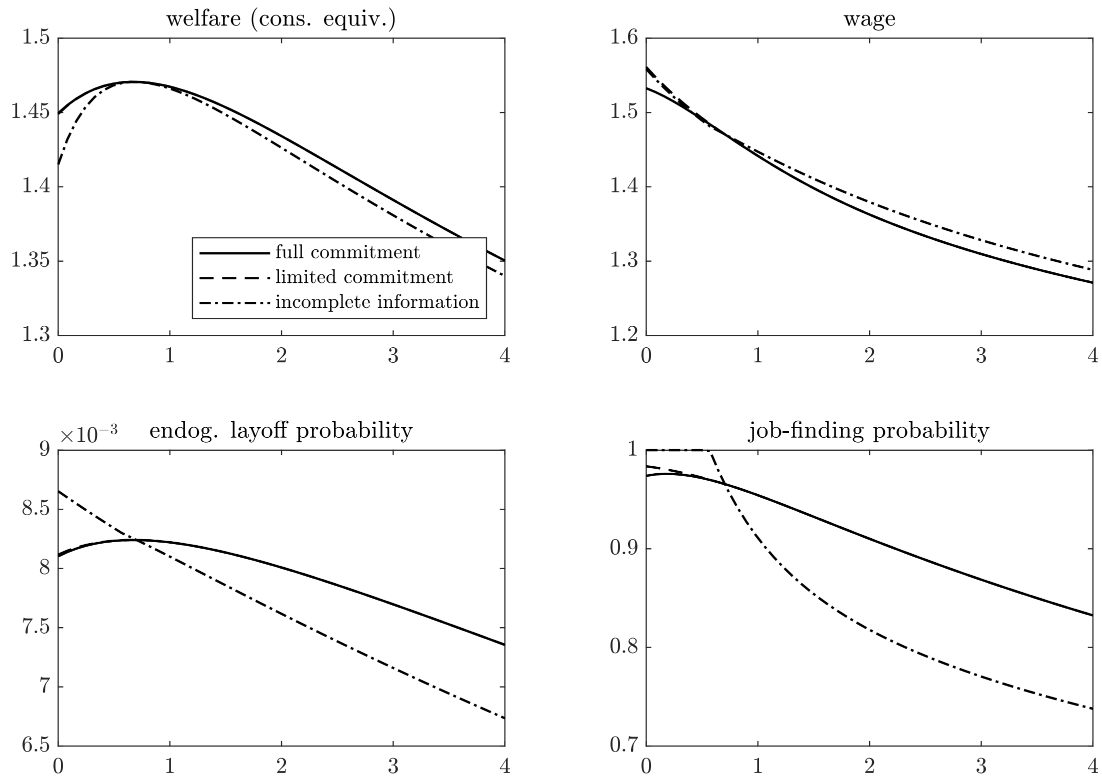


Figure 3.4. effect of severance pay (on the horizontal axis) with high matching productivity ($A = 2.7$)

Shock persistence. Finally, I explore the role of ϕ, which governs the persistence of productivity shocks. If shocks are very persistent, excessively high levels of severance pay lock workers into low-productive matches that should rather be destroyed. This possibility of inefficient retentions has already been recognized by Hall and Lazear (1984). In Figure 3.5, I set $\phi = 0.2$, which implies that a productivity draw lasts for 5 months on average. It is apparent that especially with incomplete information, implementing more than the optimal severance pay level $P^* = 0.69$ hurts the economy little in terms of welfare.[12] Hence a good estimate of the persistence of productivity shocks is crucial to properly quantify the welfare effects of severance pay, particularly the detrimental effects of excessive severance pay.

3.2.4 Extension: Smoothing severance pay

For analytical tractability, the analysis so far assumed that individuals consume all of their income in each period. While this is a common assumption in labor economics, in the present context it implies that laid off workers consume all their severance pay at once. For larger levels of severance pay this assumption might bear little realism and imply a very unbalanced consumption profile during unemployment, with consumption equal to $b + P$ in the first period of unemployment and consumption equal to b thereafter. Since increasing P makes consumption more unbalanced, the above analysis might exaggerate the welfare costs of severance pay for risk

[12]To remain in the empirically plausible range of layoff and job-finding probabilities, the standard deviation of the productivity distribution was changed to $\sigma = 0.21$ and the posting cost to $c = 3.5$.

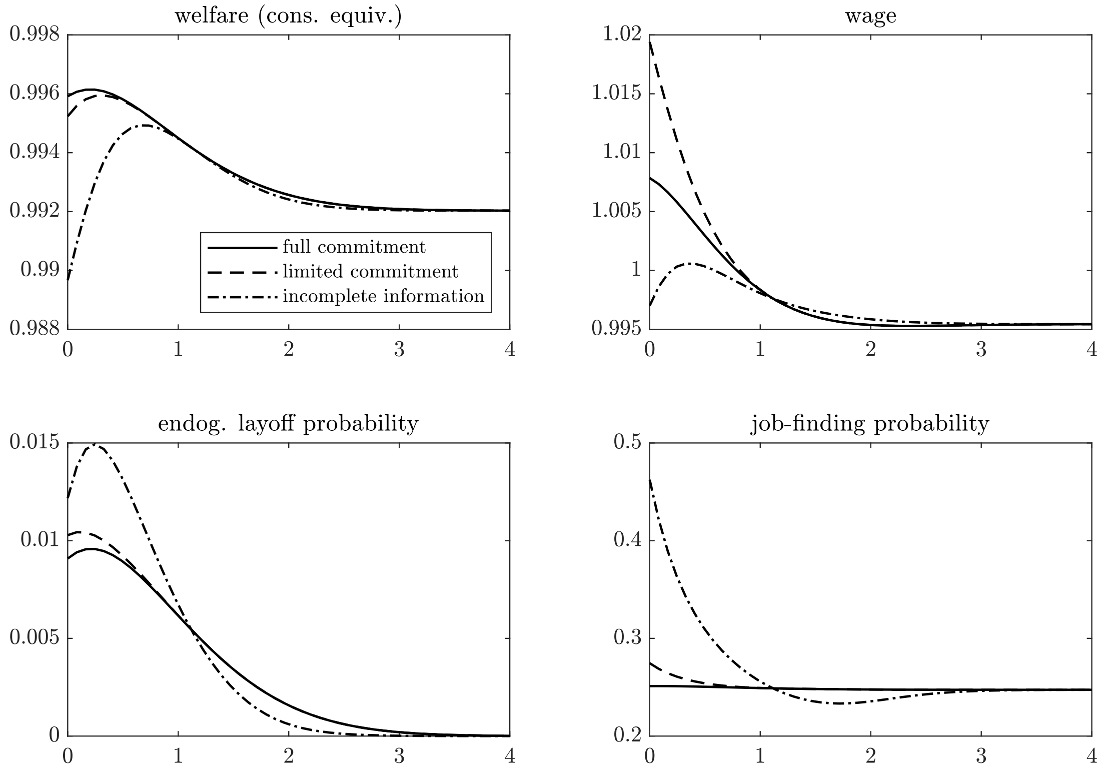


Figure 3.5. effect of severance pay (on the horizontal axis) with smaller shock persistence ($\phi = 0.2$)

averse workers. I therefore extend the model by allowing individuals to smooth the consumption of severance pay over several periods.

A natural way to achieve this is by considering an optimal consumption-saving decision at the individual level. This would make the model essentially equivalent to the model proposed in Fella and Tyson (2013). Assuming CARA utility, the authors find that optimizing individuals behave according to the permanent income hypothesis. A one-off increase in their stock of assets by P increases their consumption in all future periods by rP. The marginal propensity to consume out of severance pay is therefore equal to the interest rate r. In standard calibrations, this implies that only around 1% of severance pay is consumed within the first quarter of unemployment.

Jappelli and Pistaferri (2010) show that observed changes in consumption expenditures after income shocks substantially exceed predictions of standard models. While up to now no study seems to have estimated the marginal propensities to consume out of severance pay, the reaction to these kinds of payments might be most comparable to other one-off increases in household income. Well studied examples for such a temporary income gain are the US tax rebates of the previous two recessions. Parker et al. (2013) investigate the effects of the 2008 tax rebates on household expenditures. They estimate that on average households spent 12 to 30 percent of their stimulus payments on nondurable consumption goods within the same quarter. Including durable consumption, the figure increases to 50 to 90 percent. Households in the lower third of the income distribution (where unemployed households are likely to be overrepresented) are even found to spend 128% of the additional income on consumption, which is mostly due to durable goods such as vehicles. Johnson et al. (2006) conduct a similar analysis for the effects of the 2001 tax rebate and estimate that low income households spent close to 75% of the rebate on nondurable consumption goods within the first quarter after receipt.

For poor households, the high MPC (marginal propensity to consume) is typically explained by borrowing constraints. Yet, Johnson et al. (2006) find that even the average household had MPCs between 20% and 40%. Kaplan and Violante (2014) argue that this figure is too large to be explained by poor liquidity-constrained households alone. Drawing on further empirical insights of Misra and Surico (2014) concerning the 2001 tax rebate, Kaplan and Violante (2014) develop and calibrate a structural economic model with a liquid asset and an illiquid asset. The latter asset category captures housing and retirement accounts, for instance. They provide evidence that one third of US households can be classified as "wealthy hand-to-mouth", who essentially consume their income within a period and hold most of their wealth in illiquid assets. The authors show that the "wealthy hand-to-mouth" also have MPCs above the average and are essential to understand the high MPCs observed in the data.[13]

In the following, I allow individuals to smooth their consumption of severance pay in a way that tries to capture the empirical and theoretical insights from the literature on temporary income gains summarized above. Employed workers consume their wage income within the same period. Unemployed individuals, however, can put some money under the pillow. This stock of savings is denoted by a. In each period, unemployed individuals are assumed to spend their unemployment income b plus a constant fraction of severance pay λP as long as there is money left under the pillow, $a \geq 0$. Therefore, the individual stock of savings remains constant while employed, $a'(a) = a$, and evolves according to $a'(a) = \max\{a - \lambda P, 0\}$ while unemployed.[14]

Since savings affect the value of unemployment, optimal job search decisions and employment contracts depend on the worker's asset stock at the search stage. Formally, the equilibrium can be characterized by the constrained optimization problem

$$V(a, v) = \max_{(\theta, w, \underline{y})} p(\theta)\mathbb{E}W(w, \underline{y}, a, v) \quad \text{s.t.} \quad q(\theta)\mathbb{E}J(w, \underline{y}) = c,$$

together with the equilibrium condition $V(a, v) = v(a)$. The definition of expected firm surplus is unchanged by the introduction of severance pay smoothing. But the value of unemployment now evolves according to

$$U(a) = u(c^u(a)) + \tilde{\beta}[U(a'(a)) + v(a'(a))]$$

[13]One characteristic that distinguishes severance pay from the US tax rebates, however, is that the latter only amounts on average to 480 dollars per eligible household. Whereas optimal severance pay in Section 3.2.3 was up to 1.5 monthly wages. The empirical estimates can therefore not be taken directly to my model.

[14]Grating workers access to the capital market such that they earn interest on their savings has little effect on the results as in equilibrium savings are run down very quickly. The behavior imposed here mirrors in many respects the equilibrium outcomes of Fella and Tyson (2013). There, employed workers do not save, while unemployed workers run down their savings. The crucial difference between the two models is the different magnitude of the implies marginal propensity to consume out of temporary income gains.

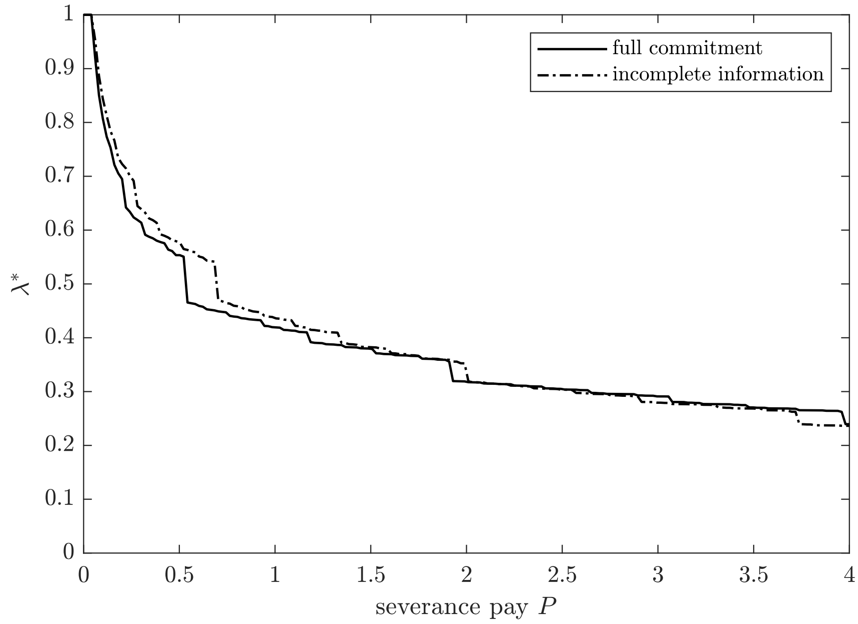


Figure 3.6. optimal MPC out of severance pay as a function of severance pay

where $c^u(a) = b + \min\{a, \lambda P\}$ and $a'(a) = \max\{a - \lambda P, 0\}$. The value of employment satisfies

$$N(w, y, a) = u(w(y)) + \tilde{\beta}[\phi \mathbb{E}N(w, \underline{y}, a) + (1 - \phi)N(w, y, a)],$$

where $\mathbb{E}N(w, \underline{y}, a) = \int_{\underline{y}(a)}^{\infty} N(w, y, a)\, dF(y) + F(\underline{y}(a))U(a+P)$. Expected surplus of employment over unemployment is $\mathbb{E}W(w, \underline{y}, a, v) = \mathbb{E}N(w, \underline{y}, a) - U(a)$.

Rather than imposing a certain value of λ, which measures the monthly marginal propensity to consume out of severance pay, the individual optimally chooses λ at the beginning of her lifetime, i.e. $\lambda^* = \max_\lambda V(0, v)$.[15] For given P and λ, the optimality conditions for the employment contract and the labor market tightness are identical to Section 3.2.1, although these objects now depend on the asset stock of the job seeker. The optimal values P^* and λ^* are determined by numerical optimization. The results of Section 3.2.3 obtain for $\lambda = 1$, when the total amount of severance pay is consumed immediately. Figure 3.6 shows that the optimal share of immediate consumption out of severance pay is monotonically decreasing in P. The presence of discrete jumps in λ^* is to due to the non-negativity constraint on assets and the discrete timing of the model. To avoid local optima, a global optimization method was used to determine λ^*.

Figure 3.7 shows that the welfare maximizing severance pay under full commitment is $P^* = 1.225$. This is associated with $\lambda^* = 0.39$, which means that 39% of severance pay are consumed in the first and second month after the separation, and the remaining 22% are consumed in the third month, provided that no new job is found before. The optimal consumption behavior therefore implies that a displaced worker who remains unemployed for a quarter consumes all of

[15] Note that if λ were re-optimized after every transition to unemployment, the asset stock at the beginning of the unemployment spell would be an additional state variable. If λ were instead re-optimized in each period of unemployment we would be back in the setting where the permanent income hypothesis holds and only a small fraction of P would be consumed each period.

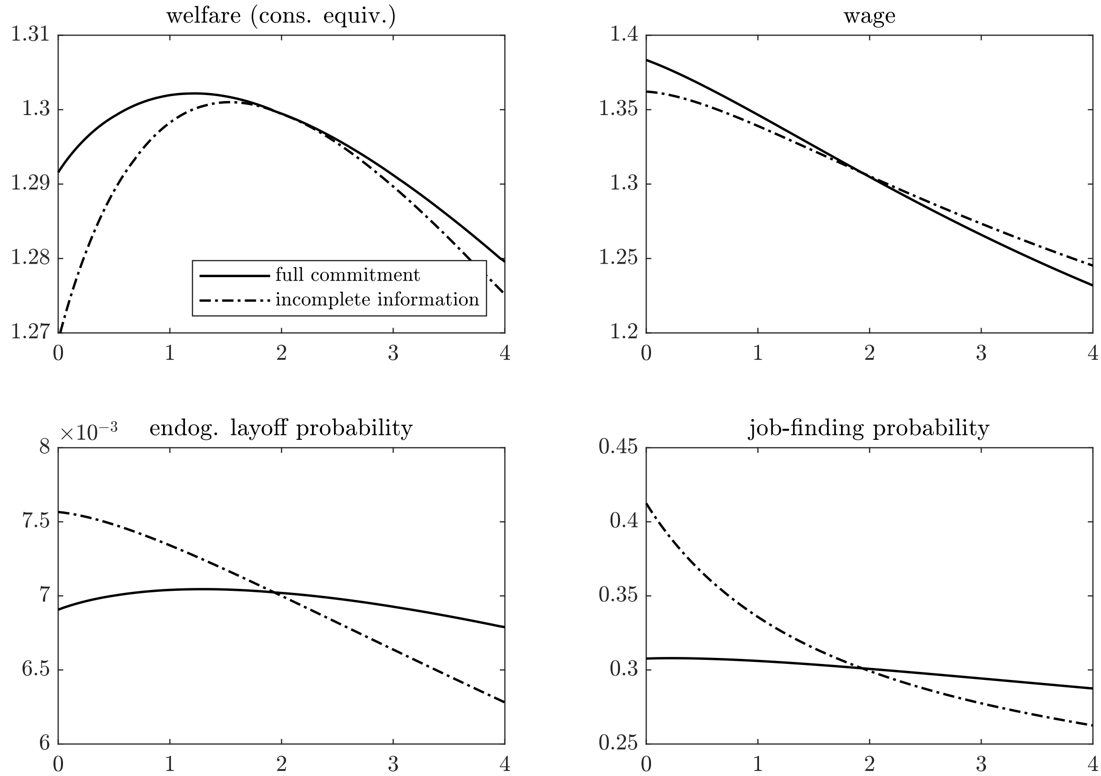


Figure 3.7. effect of severance pay (on the horizontal axis) with risk averse individuals and smoothing of severance pay

her severance pay during this period. In contrast to Section 3.2.2, optimal severance pay under full commitment does no longer insure workers against a drop in consumption in the first period after separation. Instead, displaced workers initially consume 1.278, while consumption during employment is $w^* = 1.337$. The utility loss caused by the initial drop in consumption is more than compensated by the more balanced consumption profile in subsequent periods.

Under incomplete information, optimal severance pay is $P^* = 1.551$, which exceeds the optimal level under full commitment by roughly 27%. The main insight of Proposition 3.6 therefore also holds in the extended model, at least for realistic calibrations. The MPC chosen by the individuals is $\lambda^* = 0.38$ and close to the full commitment case. It implies that optimal severance pay under incomplete information is such that workers enjoy a slight consumption increase after job loss. In the first months after a layoff, workers consume 1.389, which exceeds consumption during employment by about 5% ($w^* = 1.321$). This consumption pattern was already present in Section 3.2.2, where $w^* < b + P^*$. With smoothing of severance pay, consumption in the first period of unemployment decreases to $b + \lambda^* P^*$ but still remains above w^* in the chosen calibration. This prediction is in line with Fella (2007), where workers face the same friction and decide about optimal consumption and savings in each period. With decreasing absolute risk aversion, optimal severance pay implies that consumption increases when entering unemployment to compensate the dismissed worker for the higher uncertainty of future income.[16]

[16] Fella (2007) is in fact the working paper version of Fella and Tyson (2013). The published paper only

3.3 Dynamic contracts

The focus so far was on a stationary environment and the optimal *level* of severance pay. Wage contracts as well as severance pay schemes typically contain features that depend on the length of tenure in the firm. This section allows firms and workers to write dynamic contracts that can be contingent on tenure and on the history of productivity shocks (unless productivity is private information). The goal is to characterize the optimal *tenure profile* of severance pay. Throughout the section a non-explosive path of severance pay is assumed, $\lim_{T\to\infty} \beta^T P_T = 0$.

Only the full commitment regime and the incomplete information regime will be discussed. The intermediate case of limited commitment turns out to be less tractable and is therefore omitted. The reason is that for a given path $(P_t)_{t=0}^{\infty}$ the optimal dynamic contract under limited commitment is typically obtained by solving a sequence of recursive optimization problems where the worker's value W_t becomes an additional state variable (Thomas and Worrall, 1988; Rudanko, 2009). Since optimal severance pay can by assumption only depend on tenure but neither on W_t nor on y_t, it cannot be obtained by simply extending the set of private decision variables by P_t. As a consequence, the optimal policy cannot be characterized by a simple application of the envelope theorem as it was the case in Section 3.2.2 (compare also the discussion after Lemma 3.1).

Define the full history of productivity shocks that have occurred since the beginning of the match as $h_t = (y_0, y_1, \ldots, y_t)$. To save on notation, the productivity process introduced in Section 3.2 is represented by its conditional distribution function $G(y_t|y_{t-1}) = (1-\phi)\mathbf{1}\{y_{t-1} \geq y_t\} + \phi F(y_t)$ where $\mathbf{1}$ is the indicator function.[17] The value functions of firm and worker then evolve over tenure t according to

$$J_t(w_t, h_t) = y_t - w_t(h_t) + \tilde{\beta}\mathbb{E}[J_{t+1}(w_{t+1}, \underline{y}_{t+1})|h_t],$$
$$W_t(w_t, h_t, v) = u(w_t(h_t)) - u(b) + \tilde{\beta}\{\mathbb{E}[W_{t+1}(w_{t+1}, \underline{y}_{t+1}, v)|h_t] - v\}.$$

For $t \geq 1$, the expected surplus functions are

$$\mathbb{E}[J_t(w_t, \underline{y}_t)|h_{t-1}] = \int_{\underline{y}_t(h_{t-1})}^{\infty} J_t(w_t, h_t)\, dG(y_t|y_{t-1}) - G(\underline{y}_t(h_{t-1})|y_{t-1})P_t,$$
$$\mathbb{E}[W_t(w_t, \underline{y}_t, v)|h_{t-1}] = \int_{\underline{y}_t(h_{t-1})}^{\infty} W_t(w_t, h_t, v)\, dG(y_t|y_{t-1}) + G(\underline{y}_t(h_{t-1})|y_{t-1})\Delta_t,$$

where $\Delta_t := u(b + P_t) - u(b)$. To define an equilibrium, denote with $\underline{y} = (\underline{y}_0, (\underline{y}_t(h_{t-1}))_{t=1}^{\infty})$ the sequence of separation rules and with $w = (w_t(h_t))_{t=0}^{\infty}$ the sequence of wage schedules. Furthermore, define the expected surplus functions at the beginning of a match as $\mathbb{E}J_0(w, \underline{y}) := \int_{\underline{y}_0}^{\infty} J_0(w_0, y_0)\, dF(y_0) - F(\underline{y}_0)P_0$ and $\mathbb{E}W_0(w, \underline{y}, v) := \int_{\underline{y}_0}^{\infty} W_0(w_0, y_0, v)\, dF(y_0) + F(\underline{y}_0)\Delta_0$.

discusses the special case of constant absolute risk aversion, in which no consumption drop happens on impact.

[17] Note that $\int_a^{\infty} k(y)\, dG(y|y_{t-1}) = (1-\phi)k(y_{t-1})\mathbf{1}\{y_{t-1} \geq a\} + \phi \int_a^{\infty} k(y)\, dF(y)$ for any measurable function k.

Definition 3.3 (Dynamic Equilibrium)**.** *A dynamic equilibrium with severance pay schedule* $(P_t)_{t=0}^{\infty}$ *consists of a function* $\Theta^*(w, \underline{y}) \geq 0$*, separation thresholds* $\underline{y}^*$*, wage schedules* w^**, and a value* $v^* \geq 0$ *such that*

- *firms maximize profit under free entry,* $q(\Theta^*(w, \underline{y}))\mathbb{E}J_0(w, \underline{y}) \leq c$ *for all* $(w, \underline{y})$*, with equality for* $(w^*, \underline{y}^*)$*,*
- *job seekers apply optimally,* $v^* \geq p(\Theta^*(w, \underline{y}))\mathbb{E}W_0(w, \underline{y}, v^*)$ *for all* $(w, \underline{y})$ *and* $\Theta^*(w, \underline{y}) \geq 0$ *with complementary slackness, where* $v^* = p(\Theta^*(w^*, \underline{y}^*))\mathbb{E}W_0(w^*, \underline{y}^*, v^*)$*.*

3.3.1 Full commitment

The equilibrium can again be characterized as the solution to (3.3)–(3.4), substituting the above value functions. Differentiating the Lagrangian $\mathcal{L}^{FC} = p(\theta)\mathbb{E}W_0(w, \underline{y}, v) + \lambda[q(\theta)\mathbb{E}J_0(w, \underline{y}) - c]$ with respect to $w_t(h_t)$ reveals that the optimal dynamic contract with complete information implies full insurance across states of the world and over time, $w_t^*(h_t) = w^*$. The separation thresholds do not depend on the history of productivity shocks but may depend on tenure, $\underline{y}_t^*(h_{t-1}) = \underline{y}_t^*$. The necessary optimality conditions for w^*, $\underline{y}_t^*$, and θ^* are

$$\frac{1 - \varepsilon(\theta^*)}{\varepsilon(\theta^*)} \frac{\mathbb{E}W_0(w^*, \underline{y}^*, v)}{\mathbb{E}J_0(w^*, \underline{y}^*)} = u'(w^*), \tag{3.1}$$

$$\frac{W_t(w^*, \underline{y}_t^*, v) - \Delta_t}{u'(w^*)} + J(w^*, \underline{y}_t^*) + P_t = 0, \quad t \geq 0, \tag{3.2}$$

and the free entry condition $q(\theta^*)\mathbb{E}J_0(w^*, \underline{y}^*) = c$. Equation (3.1) has the same form and interpretation as (3.5). Equation (3.2) determines the separation threshold $\underline{y}_t^*$. It is easy to see that the separation threshold is constant over time if and only if severance pay is constant. The condition for optimal severance pay is

$$\frac{1 - \varepsilon(\theta^*)}{\varepsilon(\theta^*)} \frac{\mathbb{E}W_0(w^*, \underline{y}^*, v)}{\mathbb{E}J_0(w^*, \underline{y}^*)} = u'(b + P_t^*), \quad t \geq 0. \tag{3.3}$$

Since the left-hand side of the equation is independent of t, optimal severance pay is indeed tenure-independent. Conditions (3.1)–(3.3) are therefore equivalent to (3.5), (3.6), and (3.12). This demonstrates that with full commitment, focusing on stationary contracts from the outset constitutes no restriction in this model.

3.3.2 Incomplete information

If information is incomplete, contracts cannot depend on productivity, such that $w_t(h_t) = w_t$ and $\underline{y}_t(h_{t-1}) = \underline{y}_t$. Additionally, the separation thresholds must be self-enforcing in every period, $J_t(w_t, \underline{y}_t) + P_t = 0$. Incomplete information may give rise to strategic dynamic behavior. Workers could use the wage-tenure profile as a means to gradually learn the productivity state. For instance, if a certain wage does not lead to a layoff, the worker can infer from this a lower bound on the current match productivity. The lower bound could be refined in subsequent periods by increasing wages in a step-wise fashion—at the risk of eventually dropping out. An alternative strategy that seems equally plausible is that workers contract a high wage already in period 1 to turn down relatively low-productive matches straightaway and move to matches with higher productivity. The stochastic duration of probability draws complicates the analysis substantially, and I therefore only discuss the two edge cases $\phi = 1$ and $\phi = 0$ where the duration is deterministic. With $\phi = 1$, productivity shocks vanish within one period, which also excludes strategic considerations. With $\phi = 0$, the productivity drawn in the first period lasts for the whole duration of the match. For intermediate values $\phi \in (0, 1)$, the optimal contract and optimal severance pay are likely to combine features of the two solutions presented. Regardless of the value of ϕ, the Lagrangian under incomplete information is $\mathcal{L}^{II} = p(\theta)\mathbb{E}W_0(w, \underline{y}, v) + \lambda[q(\theta)\mathbb{E}J_0(w, \underline{y}) - c] + \sum_{t=0}^{\infty} \mu_t[J_t(w_t, \underline{y}_t) + P_t]$. The two cases merely differ in the definition of the value functions.

3.3.3 Incomplete information and purely transitory shocks

For $\phi = 1$, the value functions are the same as under full commitment but with $G(y_t|y_{t-1}) = F(y_t)$. It is easy to see that under incomplete information and purely transitory shocks worker surplus does not depend on the current productivity. To simplify notation, I thus write $W_t(w_t, v)$ in the following. The necessary first order conditions are

$$\frac{1 - \varepsilon(\theta^*)}{\varepsilon(\theta^*)} \frac{\mathbb{E}W_0(w^*, \underline{y}^*, v)}{\mathbb{E}J_0(w^*, \underline{y}^*)} = u'(w_0^*) - \frac{f(\underline{y}_0^*)}{1 - F(\underline{y}_0^*)}[W_0(w_0^*, v) - \Delta_0], \tag{3.4}$$

$$u'(w_{t-1}^*) = u'(w_t^*) - \frac{f(\underline{y}_t^*)}{1 - F(\underline{y}_t^*)}[W_t(w_t^*, v) - \Delta_t], \quad t \geq 1 \tag{3.5}$$

$$J_t(w_t^*, \underline{y}_t^*) + P_t = 0, \quad t \geq 0, \tag{3.6}$$

together with the free entry condition $q(\theta^*)\mathbb{E}J_0(w^*, \underline{y}^*) = c$. Equations (3.4) and (3.6) mirror conditions (3.9) and (3.10). Condition (3.5) determines the optimal wage-tenure profile. It reveals that the slope of the profile depends on the sign of $W_t(w_t^*, v) - \Delta_t$, which is the worker's valuation of employment relative to a separation with severance pay.

According to (3.5), the wage should increase between period $t - 1$ and period t if and only if the worker prefers a layoff with severance pay over work, $W_t(w_t^*, v) < \Delta_t$. The intuition behind this finding is that a high wage at tenure t increases the separation probability in all previous periods by the forward looking behavior of firms captured by (3.6). If a layoff is costly in utility terms, $W_t(w_t^*, v) > \Delta_t$, this gives workers an incentive to front-load wage income. If severance pay insures workers fully against job loss, $W_t(w_t^*, v) = \Delta_t$, the optimal wage profile is flat. Proposition 3.1 below shows that optimal severance pay does *not* provide perfect insurance against job loss and that wages remain tenure-dependent even with optimal severance pay.

Equations (3.4)–(3.5) also describe the optimal behavior if individuals were risk neutral. Since marginal utility is then always equal to one, it must hold that $W_t(w_t^*, v) = \Delta_t$ for $t \geq 1$, and the starting wage w_0^* is used to allocate the optimal share of match surplus to the worker. Apart from the very first period, workers are therefore indifferent between work and separation. This property of the optimal contract is also highlighted by Boeri et al. (2017). Proposition 3.1 below shows that it does not extend to risk averse individuals.

Optimal severance pay. The first order conditions for optimal severance pay can be summarized as

$$\frac{1-\varepsilon(\theta^*)}{\varepsilon(\theta^*)}\frac{\mathbb{E}W_0(w^*,\underline{y}^*,v^*)}{\mathbb{E}J_0(w^*,\underline{y}^*)} = MP_0^*, \tag{3.7}$$

$$MP_{t-1}^* = MP_t^* - \frac{f(\underline{y}_{t-1}^*)}{1-F(\underline{y}_{t-1}^*)}(W_{t-1}^*(w_{t-1}^*,v^*) - \Delta_{t-1}^*),\ t \geq 1. \tag{3.8}$$

Condition (3.7) is analogous to (3.14), while (3.8) determines the optimal tenure profile of severance pay. The marginal gains of higher severance pay in period t are $MP_t^* := u'(b + P_t^*) + \frac{f(\underline{y}_t^*)}{F(\underline{y}_t^*)}(W_t^*(w_t^*,v^*) - \Delta_t^*)$. The first term in MP_t^* measures the utility gain in case of a separation, while the second term captures that higher P_t decreases the separation threshold $\underline{y}_t$ because separating for the firm becomes more costly in period t. The last term on the right-hand side of equation eq. (3.8) represents an additional marginal cost in the trade-off between P_t and P_{t-1}. A higher P_t increases separation costs at tenure t, which leads forward-looking firms to increase separations at tenure $t-1$ due to a lower continuation value. More generally, protecting employment at tenure t by means of severance pay comes at the downside that separation probabilities increase in preceding periods. Due to these intertemporal considerations, a decreasing tenure profile of severance pay may actually be optimal. To see this, assume for the moment a logistic productivity distribution, $F(x) = [1 + e^{-(x-\mu)/s}]^{-1}$ with $\mu \in \mathbb{R}$ and $s > 0$.[18] Condition (3.8) can then be rewritten

$$u'(b + P_{t-1}^*) - u'(b + P_t^*) = (1 - F(\underline{y}_t^*))(W_t^*(w_t^*,v^*) - \Delta_t^*) - (W_{t-1}^*(w_{t-1}^*,^*) - \Delta_{t-1}^*).$$

A necessary condition for P_t^* to be increasing in tenure is therefore that $W_t^*(w_t^*,v^*) - \Delta_t^*$ is increasing in tenure. The more valuable the match becomes over time relative to a separation, the more the planner cares about low separation rates later on and is willing to accept higher separation rates early in the match. In the baseline version of the model, however, there is no explicit mechanism that makes the match more valuable over time. Section 3.3.3 below extends the model for moral hazard, which creates a rationale for back-loading worker surplus and can give rise to optimal severance pay that is increasing in tenure.

[18] The logistic distribution satisfies $f(x)/F(x) = (1 - F(x))/s$ and $f(x)/(1 - F(x)) = F(x)/s$.

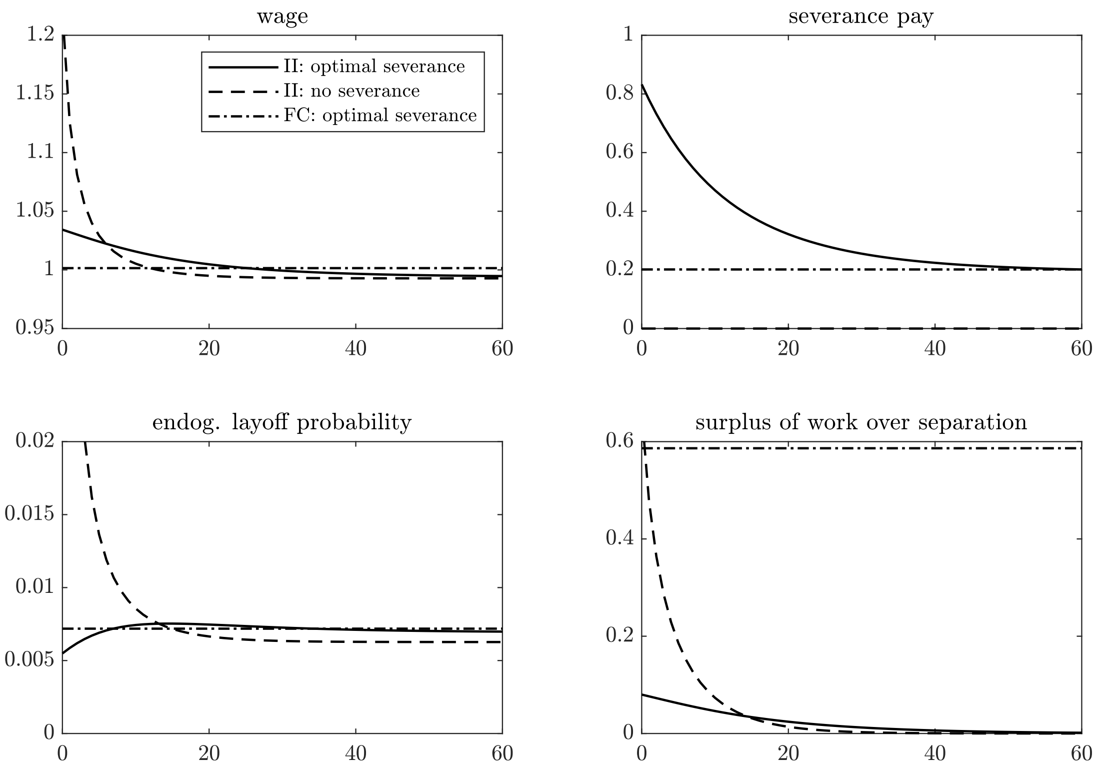


Figure 3.1. equilibrium with dynamic contracts, horizontal axis: tenure in months

Finally, the set of first order conditions for wages and severance pay can be combined into

$$u'(w_t^*) = u'(b + P_t^*) + \frac{f(\underline{y}_t^*)}{(1 - F(\underline{y}_t^*))F(\underline{y}_t^*)}(W_t^*(w_t^*, v^*) - \Delta_t^*), \quad t \geq 0, \tag{3.9}$$

such that an equilibrium with optimal severance pay is characterized by equations (3.4)–(3.6) and (3.9) together with the free entry condition. Further insights can only be derived by a numerical solution of the model. What can be shown analytically is that, unlike in the full commitment case, a flat wage-tenure profile is not optimal. Furthermore, optimal severance pay does not fully insure the worker against job loss. The main prediction of Proposition 3.5 therefore extends to tenure-dependent policies. The proof is by induction and given in Section 3.A,

Proposition 3.1. *Let individuals be risk averse. Consider an equilibrium with optimal severance pay and $p(\theta^*) < 1$. Then there does not exist a $T \geq 0$ such that w_t^* or $W_t^* - \Delta_t^*$ are constant for $t \geq T$.*

Numerical illustration. To illustrate the optimal contract and the optimal policy under incomplete information with fully transitory shocks, I use the baseline calibration of Section 3.2.3 but with $\phi = 1$. I set $\sigma = 0.65$ and $c = 3$ to implement the same layoff and job-finding probabilities under full commitment as with the original calibration.

Figure 3.1 plots the optimal path of wages, severance pay, and the layoff probability over the first five years of a match. Additionally, the surplus of work over a separation with severance pay, $W_t^* - \Delta_t^*$, is depicted. As discussed above in Section 3.3.1, under full commitment all variables are independent of tenure, which corresponds to the dash-dotted lines in Figure 3.1. With incomplete information and no severance pay, the optimal contract is represented by the dashed line. The optimal wage is strongly decreasing during the first year of the match. The wage in the initial month is 1.22 which quickly drops to a long-run level of 0.993. The probability for an endogenous layoff follows a similar path. It starts off at 8.14 percent and decreases to 0.63 percent within the first two years. Optimal severance pay (solid line) essentially smooths wages and layoff probabilities over time. Compared to the *laissez-faire*, the initial wage is much lower (1.034) while the long-run level is slightly higher (0.995). The high layoff probability at the beginning of the match disappears completely. Instead, the layoff probability fluctuates slightly around its long-run level of 0.7 percent. To reduce the initial wage and layoff probabilities that individuals would choose in a *laissez-faire* economy, a high severance pay is necessary at the beginning of a match. It equals 0.833 and therefore 0.8 monthly wages in the first period and decays slowly to 0.201 (0.2 monthly wages) in the long-run. The bottom right panel shows that the surplus of work over a separation with severance pay is decreasing in tenure. If severance pay spurs moral hazard, this should become more important at higher tenure durations because the surplus gradually decreases to zero—although by Proposition 3.1 it never actually reaches zero.

The quantity by which optimal severance pay drops over time may seem surprising, especially because governments either mandate flat or tenure-increasing profiles. There are two main reasons for this deviation between the model and real world. First, the analysis above imposes $\phi = 1$, which rules out any persistent productivity differences between matches. With $\phi > 1$, optimal severance pay is likely to be lower in the initial periods since otherwise there may be a large mass of low productivity matches that only survive due to severance pay. This conjecture is corroborated by the findings in Section 3.3.4 below which suggest that a constant or increasing profile of severance pay is optimal if productivity is fully persistent ($\phi = 0$). Second, the model does not take into account moral hazard. Boeri et al. (2017) succeed to generate a tenure-increasing profile of severance pay in a similar model with risk neutral agents, where they additionally consider moral hazard on the worker's side. The authors find that wages increase with tenure if investing into the work relation becomes more cumbersome over time. Due to the same informational friction that is considered here, tenure-increasing wages then lead to an increasing mass of excess layoffs. To counteract, severance pay must increase in tenure as well.

The mechanism of Boeri et al. (2017) could be easily built into my model by including the additional no-shirking conditions $\mathbb{E}W_t(w_t, \underline{y}_t, v) - (1-\gamma)\Delta_t \geq C_{t-1}$ where C_{t-1} is the cost of effort provision and γ is the probability with which the government is able to correctly identify a disciplinary layoff due to shirking. Depending on the size of investment costs and how strongly they increase over time, optimal severance pay may either be decreasing, U-shaped, or increasing with tenure. Below I take a different approach and assume that workers choose a probability with which they shirk instead of giving them a binary decision. That way, I can rationalize tenure-increasing severance pay without the need to assume that working honestly becomes more cumbersome with increasing tenure.

Extension: Endogenous effort provision

As argued above, risk aversion and incomplete information seem unable to explain why most countries mandate severance pay that is increasing in tenure.[19] Along the lines of Holzmann et al. (2012) and Boeri et al. (2017) I consider endogenous effort provision of the worker as an additional channel. In every period, the worker privately chooses how much effort e_{t+1} to provide in period $t+1$.[20] Work effort e_{t+1} is normalized to the interval $[0,1]$, and $1-e_{t+1}$ is the probability with which the worker is shirking. To keep things simple, a shirking worker is always caught by the employer and loses her job for disciplinary reasons without receiving severance pay.[21] The surplus functions of worker and firm satisfy, respectively,

$$\begin{aligned} W_t(w_t, y_t, v) &= \max_{e_{t+1}\in[0,1]} \{u(w_t) - u(b) + \tilde{\beta}[e_{t+1}\mathbb{E}W_{t+1}(w_{t+1}, \underline{y}_{t+1}, v) - D(e_{t+1}) - v]\}, \\ J_t(w_t, y_t) &= y - w_t + \tilde{\beta}e_{t+1}\mathbb{E}J_{t+1}(w_{t+1}, \underline{y}_{t+1}), \end{aligned}$$

where $D(e)$ captures the disutility of providing e units of effort. If effort provision is costless, $D \equiv 0$, the model reduces to the version studied above. I instead assume that the disutility function is strictly increasing and convex with $D(0) = 0$, $D'(e) > 0$, and $D''(e) > 0$. To avoid case distinctions, I further impose $D'(0) = 0$ and $\lim_{e\to 1} D'(e) = \infty$.

The first order condition for optimal effort provision is $D'(e^*_{t+1}) = \mathbb{E}W_{t+1}$ such that $e^*_{t+1} = \varphi(\mathbb{E}W_{t+1})$ where $\varphi := (D')^{-1}$. The assumptions on D guarantee that $\varphi(\mathbb{E}W_{t+1})$ is always between 0 and 1. The first order conditions for wages and separation thresholds are identical to (3.4)–(3.6), where condition (3.5) is replaced by

$$u'(w^*_{t-1}) = \frac{u'(w^*_t) - \frac{f(\underline{y}^*_t)}{1-F(\underline{y}^*_t)}(W^*_t - \Delta_t)}{1 - \frac{\varphi'(\mathbb{E}W^*_t)}{\varphi(\mathbb{E}W^*_t)}\mathbb{E}J^*_t[u'(w^*_t) - \frac{f(\underline{y}^*_t)}{1-F(\underline{y}^*_t)}(W^*_t - \Delta_t)]}, \quad t \geq 1. \tag{3.10}$$

The difference between (3.10) and (3.5) stems from the denominator, which is less than 1 because $\varphi'(\mathbb{E}W_{t+1}) = 1/D''(e^*_{t+1}) > 0$. As a result, $u'(w^*_{t-1}) > u'(w^*_t) - \frac{f(\underline{y}^*_t)}{1-F(\underline{y}^*_t)}(W^*_t - \Delta_t)$ and the wage-tenure profile becomes more back-loaded relative to the baseline model. Optimal severance pay in the extended model is still characterized by equation (3.9) and is also likely to become more back-loaded.

[19] Annex B of Holzmann et al. (2012) reports that of the countries that have mandated severance pay systems, the number of compensated weekly wages is increasing in tenure in 125 countries and flat in 24 countries.

[20] The timing convention follows Boeri et al. (2017) but it is not crucial to the results. Workers could equally decide on e_{t+1} in the beginning of period $t+1$.

[21] This assumes that institutions can perfectly distinguish between economic layoffs due to low match productivity and disciplinary layoffs due to worker misconduct. Extensions to imperfect observability in the spirit of Galdón-Sánchez and Güell (2003) and Boeri et al. (2017) are possible but not relevant to illustrate the mechanism.

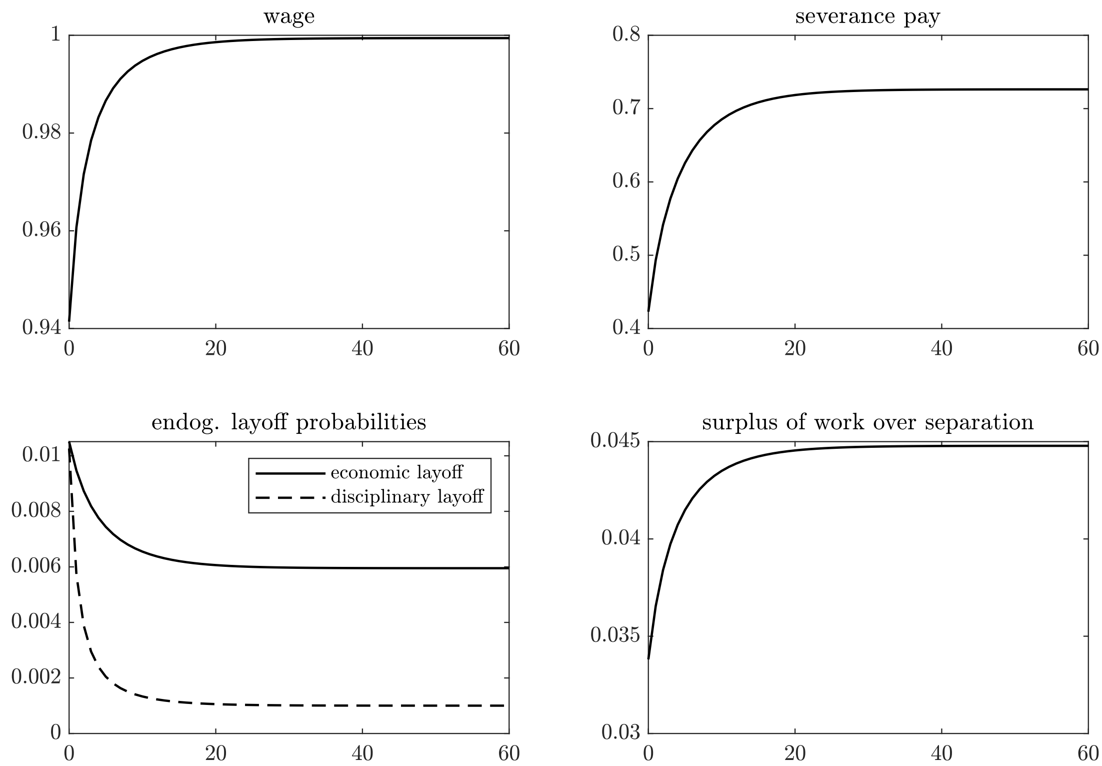


Figure 3.2. equilibrium with incomplete information and moral hazard, horizontal axis: tenure in months

Numerical illustration. For illustration purposes, the disutility function D is chosen such that $\varphi(x) = 1 - \exp(-\frac{x}{s})$. This implies $D'(e) = -s\ln(1-e)$ and $D(e) = s[e + (1-e)\ln(1-e)]$. It is easy to see that this disutility function satisfies all properties demanded above. I set $s = 0.1$ and adjust $\sigma = 0.37$ and $c = 1.65$ to achieve layoff and job-finding probabilities similar to the baseline results. The equilibrium paths are represented by the solid lines in Figure 3.2. The qualitative patterns are very different from Figure 3.1. The incentive to back-load wages to ensure high effort provision dominates the incentive to front-load wages that stems from the informational friction. Both wages and severance pay are concave increasing in tenure. At the same time, the probability of economic layoffs (due to an unfavorable productivity draw) and disciplinary layoffs (due to shirking) reduce with tenure. Both are equal to 1 percent in the first period of the match. Especially the probability for a disciplinary layoff $1 - e^*_{t+1}$ reduces very quickly during the first months. In the long run, disciplinary layoffs occur with a probability of 0.1 percent, while the probability for an economic layoff converges to 0.6 percent.

Altogether, the observations of this section suggest that among the three rationales for severance pay identified by Holzmann et al. (2012), the human capital channel is essential to understand why severance pay is increasing in tenure in most countries. To generate upwards sloping severance pay, Boeri et al. (2017) require tenure-increasing costs of effort provision, while the extended model presented here does without this assumption. In the calibration studied above, workers optimally increase effort over time, which endogenously gives rise to investment costs $D(e^*_t)$ that increase in tenure. Figure 3.2 also shows that in every period workers prefer work over a separation with severance pay. In the model of Boeri et al. (2017), by contrast, this applies only in the first period of the match, and workers are indifferent between work and layoff thereafter.

3.3.4 Incomplete information and fully persistent draws

Finally, I discuss the optimal contract and optimal severance pay if information is incomplete and productivity draws are fully persistent. The central observation of this section is that with fully persistent productivity draws, an optimal contract features a constant wage and a constant separation threshold, such that all separations happen in the first period of the match. While the exact proof only applies to risk neutral individuals, numerical investigations with logarithmic utility suggest that this finding is more general.

With fully persistent productivity draws, all uncertainty is revealed at beginning of the match. It is sufficient to limit attention to contracts that specify separation thresholds $\underline{y}_t$ that (weakly) increase in tenure. The reason is that if $\underline{y}_{t+1} < \underline{y}_t$, no layoffs occur in period $t+1$ since all workers with productivity below $\underline{y}_t$ have already been separated earlier. Hence a contract with $\underline{y}_{t+1} = \underline{y}_t$ gives rise to the same labor market outcomes.

Assuming weak monotonicity, a layoff occurs in period $t \geq 1$ if y falls into the interval $[\underline{y}_{t-1}, \underline{y}_t)$. For the separation threshold to be self-enforcing, it must be the case that $J_t(\underline{y}_t) + P_t = 0$ whenever $\underline{y}_t > \underline{y}_{t-1}$ and $J_t(\underline{y}_t) + P_t \geq 0$ whenever $\underline{y}_t = \underline{y}_{t-1}$, where firm surplus evolves over time as $J_t(y) = y - w_t + \tilde{\beta} \max\{J_{t+1}(y), -P_{t+1}\}$. Due to the maximum operator, a large number of case distinctions would be necessary to derive the first order conditions, which would obscure analytical insights. I therefore proceed under the more stringent Assumption 3.2.

Assumption 3.2. *The optimal contract is such that the separation threshold $\underline{y}_t^*$ is strictly increasing until period $T \geq 0$ and constant thereafter. $T = \infty$ is allowed.*

Strict monotonicity implies that the worker learns more and more accurate lower bounds on her productivity state, until she eventually loses her job. It is shown in Section 3.B that expected firm surplus at the search stage can be written as

$$\mathbb{E}J_0(w, \underline{y}^*) = \sum_{s=0}^{T-1} \tilde{\beta}^s \int_{\underline{y}_s^*}^{\infty} y - \underline{y}_s^* \, dF(y) + \tilde{\beta}^T \int_{\underline{y}_T^*}^{\infty} \frac{y - \underline{y}_T^*}{1 - \tilde{\beta}} \, dF(y) - P_0, \tag{3.11}$$

where the separation thresholds equal $\underline{y}_s^* = w_s - P_s + \tilde{\beta} P_{s+1}$ for $s < T$ and $\underline{y}_T^* = (1 - \tilde{\beta})\left[\sum_{s=T}^{\infty} \tilde{\beta}^{s-T} w_s - P_T\right]$. For $T = \infty$, the last term in (3.11) is absent. Analogously, expected worker surplus can be written

$$\mathbb{E}W_0(w, \underline{y}^*, v) = \sum_{s=0}^{T} \tilde{\beta}^s (1 - F(\underline{y}_s^*))(W_s - \Delta_s) + \Delta_0,$$

where $W_s = u(w_s) - u(b) - \tilde{\beta} v + \tilde{\beta} \Delta_{s+1}$ for $s < T$ and $W_T = \sum_{s=T}^{\infty} \beta^{s-T}(u(w_s) - u(b) - \tilde{\beta} v)$ if

$T < \infty$. The first order conditions with respect to wages reveal $w_s^* = w_T^*$ for $s \geq T$ as well as

$$\frac{1-\varepsilon(\theta^*)}{\varepsilon(\theta^*)} \frac{\mathbb{E}W_0(w^*, \underline{y}^*, v)}{\mathbb{E}J_0(w^*, \underline{y}^*)} = u'(w_s^*) - \frac{f(\underline{y}_s^*)}{1-F(\underline{y}_s^*)}(W_s^* - \Delta_s), \quad s \leq T. \tag{3.12}$$

If productivity is constant throughout the match, w_s^* only affects the separation threshold in period s, and (3.12) captures the trade-off between wage and retention probability. The difference to (3.4)–(3.5) stems from the fact that with transitory shocks, w_s^* affects firm surplus in all preceding periods, not just in the current one. Optimal severance pay must satisfy

$$\frac{1-\varepsilon(\theta^*)}{\varepsilon(\theta^*)} \frac{\mathbb{E}W_0(w^*, \underline{y}^*, v^*)}{\mathbb{E}J_0(w^*, \underline{y}^*)} = u'(b + P_0^*) + \frac{f(\underline{y}_0^*)}{F(\underline{y}_0^*)}(W_0^* - \Delta_0^*),$$

$$(1 - F(\underline{y}_{s-1}^*))[u'(w_{s-1}^*) - u'(b + P_s^*)] = (1 - F(\underline{y}_s^*))[u'(w_s^*) - u'(b + P_s^*)], \quad 1 \leq s \leq T.$$

Severance pay after period T does not affect the labor market equilibrium. To ensure $J_t(\underline{y}_T^*) + P_t^* \geq 0$ for $s \geq T$, however, it must be greater than or equal to P_T^*.

It is easy to see from the above conditions that constant paths for w_t^*, $\underline{y}_t^*$, and P_t^* satisfy all the necessary first order conditions. At least if risk aversion becomes negligible, it can be formally shown that the optimal solution is indeed time-constant.

Proposition 3.2. *Let workers be risk neutral ($\eta \to 0$) and let the hazard function of the productivity distribution, $h := f/(1-F)$, be increasing. Any optimal contract that satisfies Assumption 3.2 features a constant wage and separation threshold.*[22]

Proof. The proof is by contradiction and only uses condition (3.12) and the expressions for separation thresholds $\underline{y}_t^*$ and worker surplus W_t^* given above. Assume that $T \geq 1$. Condition (3.12) implies $u'(w_0^*) - h(\underline{y}_0^*)W_0^* = u'(w_t^*) - h(\underline{y}_t^*)W_t^*$. Substituting $\underline{y}_t^*$ and W_t^*, noting $u(x) = x$, this reduces to $h(\underline{y}_0^*)(\underline{y}_0^* - b - \tilde{\beta}v) = h(\underline{y}_t^*)(\underline{y}_t^* - b - \tilde{\beta}v)$. Because the hazard function is increasing, it must hold that $\underline{y}_0^* = \underline{y}_t^*$ for $1 \leq t \leq T$. This violates Assumption 3.2 by which $\underline{y}_0^* < \underline{y}_t^*$ for $1 \leq t \leq T$. Because this argument holds for any $T \geq 1$, an optimal contract requires $T = 0$. Wages and separation thresholds are then constant for the entire duration of the match. □

The above proposition establishes that a flat contract is optimal if workers are risk neutral. It is therefore likely to translate to risk averse workers who have an explicit preference for smooth consumption. As pointed out above, a constant contract along with constant severance pay indeed solves the first order conditions for any degree of risk aversion. There is, however, no analytical proof that this is the only solution to the equations. Numerical experiments with logarithmic utility suggest that the constant solution may indeed be unique for plausible calibrations.

For $\phi \in (0, 1)$, the optimal contract and the optimal policy are likely to interpolate the results obtained for the edge cases $\phi = 0$ and $\phi = 1$. Since the above analysis yields constant

[22]Many standard distributions have increasing hazard functions, for instance the normal distribution, the logistic distribution, the Gumbel distribution, or the Weibull distribution with shape parameter greater than or equal to 1.

solutions, I conjecture that the qualitative behavior of the intermediate cases is similar to what has been found in Section 3.3.3. Obtaining an explicit solution under incomplete information is complicated by the stochastic duration of the probability draws and is left for future research.

3.4 Conclusion

This essay has investigated the link between welfare-maximizing severance pay and bilateral contracting frictions on the labor market. Three contractual regimes are considered. The benchmark is the full commitment regime with possibly productivity-contingent wages and commitment to both the wage schedule and a separation rule. In the limited commitment regime, firm and worker commit to the wage schedule but employment is at will. In the incomplete information regime, employment is at will and wages cannot depend on productivity.

In a stationary environment, optimal severance pay does not depend on the severity of the contracting frictions if individuals are risk neutral. Moreover, optimal severance pay removes the distortions in hiring and firing caused by contracting frictions such that there is no welfare loss relative to full commitment. The same holds for risk averse individuals, provided that the search frictions on the worker's side of the labor market are negligible. Otherwise, the optimal size of severance pay is higher the more severe the contracting frictions, and severance pay can no longer undo the hiring and firing distortions of the contracting frictions. The numerical investigation suggests that in presence of incomplete information the welfare gain of the optimal policy relative to the *laissez-faire* can be sizeable and that a large part of the welfare loss caused by the contracting friction is compensated.

Extending the analysis to dynamic contracts reveals that risk aversion and contracting frictions together cannot explain why most countries around the world mandate severance pay levels that increase in tenure. Considering moral hazard of the workforce, for instance in the form of private effort provision, seems crucial to explain this empirical regularity and should receive more attention in future research. The model also suggests a close connection between the tenure component of severance pay and the steepness of individual wage-tenure profiles, which calls for further empirical investigation.

3.A Proofs omitted in the text

Proof of Lemma 3.1. To save on notation, define $z^*(v,P) = (\lambda^*(v,P), w^*(v,P), \underline{y}^*(v,P))$. The envelope theorem implies $\frac{\partial V(v,P)}{\partial P} = \frac{d\mathcal{L}(z^*(v,P),v,P)}{dP} = \frac{\partial \mathcal{L}(z^*(v,P),v,P)}{\partial P}$. The same holds for the partial derivative with respect to v, which amounts to $\frac{\partial V(v,P)}{\partial v} = -p(\theta^*)\frac{\tilde{\beta}(1-F(\underline{y}^*))}{1-\tilde{\beta}(1-\phi F(\underline{y}^*))} < 0$. Since the denominator on the right-hand side of (3.11) is positive, $\frac{dv^*(P)}{dP}$ and $\frac{\partial \mathcal{L}}{\partial P}$ have the same sign. The last statement of the proposition follows directly. □

Proof of Proposition 3.1. Suppose that $(\theta^*, w_P^*, \underline{y}^*, v^*)$ is an equilibrium with severance pay P. It is easy to verify that under full commitment, $(\theta^*, w_0^*, \underline{y}^*, v^*)$ is an equilibrium for $P = 0$, where $w_0^* = w_P^* + \frac{F(\underline{y}^*)}{1-F(\underline{y}^*)}P$. Therefore, all equilibrium objects apart from the wage level coincide. To show neutrality of severance pay under limited commitment, a case distinction is necessary. If $J(w_P^*, \underline{y}^*) + P > 0$, the equilibrium under limited commitment is also an equilibrium under full commitment and the above argument applies. If $J(w_P^*, \underline{y}^*) + P = 0$ then $(\theta^*, w_0^*, \underline{y}^*, v^*)$ is an equilibrium for $P = 0$, where $w_0^* = w_P^* - y_P^+ + y_0^+ - P$ and y_P^+ denotes the profitability threshold for severance pay level P, that is the productivity level at which the layoff constraint becomes binding. □

Proof of Proposition 3.3. The line of argumentation is very similar in all three cases. Since by the free entry condition every equilibrium must feature $\mathbb{E}J^* > 0$, the first order conditions imply that $\mathbb{E}W^* > 0$ must hold in an equilibrium with optimal severance pay. Now assume that $P^* \leq 0$. Under full commitment, combining conditions (3.5) and (3.12) yields $w^* = b + P^*$ and therefore $w^* \leq b$. By (3.2) expected worker surplus is then non-positive, which is a contradiction to $\mathbb{E}W^* > 0$. Under limited commitment, combining conditions (3.7) and (3.13) yields $w^* \leq b + P^*$. If $P^* \leq 0$, this implies $w^* \leq b$ and leads to the same contraction as before. With incomplete information, I first demonstrate that $W^* - \Delta^* > 0$. Since $P^* \leq 0$ implies $\Delta^* \geq 0$ and $\mathbb{E}W^* = (1 - F(\underline{y}^*))W^* + F(\underline{y}^*)\Delta > 0$, it must hold that $W^* > 0$. Taking all together, we have $W^* - \Delta^* \geq W^* > 0$. Now, combining (3.9) and (3.14) yields $w^* < b + P^*$. Since $P^* \geq 0$ this implies $w^* < b$ and again leads to a contradiction with $\mathbb{E}W^* > 0$. □

Proof of Proposition 3.6. Cobb-Douglas matching technology implies a constant elasticity of the vacancy-filling probability, $\varepsilon(\theta) \equiv \varepsilon$. Let symbols of the form $x^i(P)$ refer to the equilibrium value of x in scenario i when the level of severance pay is P. By (3.11), the derivative of $v^{FC}(P)$ evaluated at P^{II} satisfies

$$\left.\frac{dv^{FC}(P)}{dP}\right|_{P^{II}} \propto \left.\frac{\partial \mathcal{L}^{FC}}{\partial P}\right|_{P^{II}} \propto u'(b + P^{II}) - \frac{1-\varepsilon}{\varepsilon}\frac{\mathbb{E}W^{FC}(P^{II})}{\mathbb{E}J^{FC}(P^{II})}$$

Substituting the objective function and the free-entry condition allows to rewrite $\frac{\mathbb{E}W^i(P)}{\mathbb{E}J^i(P)} = \frac{V^i(P)}{c\theta^i(P)}$ for $i \in \{FC, II\}$ and any P. Since incomplete information adds one constraint the optimization problem, $V^{II}(P) \leq V^{FC}(P)$ for all P. Furthermore, the stipulated inequality on the job-finding probability implies $\theta^{FC}(P^{II}) \leq \theta^{II}(P^{II})$. Together, this reveals

$$u'(b+P^{II}) - \frac{1-\varepsilon}{\varepsilon}\frac{\mathbb{E}W^{FC}(P^{II})}{\mathbb{E}J^{FC}(P^{II})} \leq u'(b+P^{II}) - \frac{1-\varepsilon}{\varepsilon}\frac{\mathbb{E}W^{II}(P^{II})}{\mathbb{E}J^{II}(P^{II})} < 0,$$

where the sign follows from (3.14) and the fact that $W^{II}(P^{II}) > \Delta^{II}$ by Proposition 3.5. Hence v^{FC} is strictly increasing at $P = P^{II}$. At $P = 0$, observe

$$\left.\frac{dv^{FC}(P)}{dP}\right|_{P=0} \propto \left.\frac{\partial \mathcal{L}^{FC}}{\partial P}\right|_{P=0} \propto u'(b) - u'(w^{FC}(0)) > 0.$$

The sign follows from $w^{FC}(0) > b$ as otherwise expected worker surplus is zero, which violates the first order condition (3.5). By continuity, the function $v^{FC}(P)$ must attain a local maximum P^{FC} in the interval $(0, P^{II})$. The existence of a further local maximum $P_2^{FC} > P^{II}$ is ruled out by the fact that this would require a local minimum in (P^{II}, P_2^{FC}) because of continuity. However, any solution to the necessary first order conditions (3.5), (3.6), and (3.12) constitutes a local maximum of $v^{FC}(P)$ as postulated by the Lemma below. Combining the above insights shows that optimal welfare satisfies $v^{II}(P^{II}) \leq v^{FC}(P^{II}) < v^{FC}(P^{FC})$. □

Lemma. *Denote with $v^*(P)$ be the equilibrium value of search under full commitment let the matching technology be Cobb-Douglas. Then $v^*(P)$ is strictly concave at any optimum $P = P^*$.*

Proof of the Lemma. Differentiating (3.11) and noting that $\frac{dv^*(P^*)}{dP} = \frac{\partial V(v^*(P^*),P^*)}{\partial P^2} = 0$ reveals that $\frac{d^2v^*(P)}{dP^2} = \frac{\partial^2 V(v^*(P),P)}{\partial P^2}/[1 - \frac{\partial V(v^*(P),P)}{\partial v}]$ at any optimum $P = P^*$. Since $\frac{\partial V(v^*(P),P)}{\partial v} < 0$, $v^*(P)$ is concave at $P = P^*$ if and only if $\frac{\partial^2 V(v^*(P),P)}{\partial P^2} < 0$. To save on notation, I omit v and use $z^*(P) = (\lambda^*(P), w^*(P), \underline{y}^*(P))$. Remember that $V(P) = \mathcal{L}(z^*(P), P)$ where $z^*(P)$ satisfies the first order conditions $\mathcal{L}_z(z^*(P), P) = 0$. Differentiating $V(P)$ yields $V'(P) = \mathcal{L}_z(z^*(P), P)z_P^*(P) + \mathcal{L}_P(z^*(P), P) = \mathcal{L}_P(z^*(P), P)$. The second derivative is

$$V''(P) = \mathcal{L}_{Pz}(z^*(P), P)z_P^*(P) + \mathcal{L}_{PP}(z^*(P), P). \tag{3.A.1}$$

To obtain the derivative of the policy functions $z_P^*(P)$, differentiate the first order condition with respect to P, which yields $\mathcal{L}_{zz}(z^*(P), P)z_P^*(P) + \mathcal{L}_{zP}(z^*(P), P) = 0$. Assuming that $\mathcal{L}_{zz}$ is invertible (which will be verified later on) yields $z_P^*(P) = -[\mathcal{L}_{zz}(z^*(P), P)]^{-1}\mathcal{L}_{zP}(z^*(P), P)$. Substituting this into (3.A.1) yields $V''(P) = \mathcal{L}_{PP} - \mathcal{L}_{Pz}\mathcal{L}_{zz}^{-1}\mathcal{L}_{zP}$ where all terms are evaluated in $(z^*(P), P)$. Since $\mathcal{L}_{zz}$ is a 4×4 matrix it is not recommendable to analytically calculate the inverse. Instead, consider the block matrix

$$\mathcal{L}_{(z,P)(z,P)} = \begin{pmatrix} \mathcal{L}_{zz} & \mathcal{L}_{Pz} \\ \mathcal{L}_{zP} & \mathcal{L}_{PP} \end{pmatrix}$$

Applying the formula for the determinant of a block matrix, $\det \mathcal{L}_{(z,P)(z,P)} = \det \mathcal{L}_{zz} \cdot \det[\mathcal{L}_{PP} - \mathcal{L}_{Pz}\mathcal{L}_{zz}^{-1}\mathcal{L}_{zP}] = \det \mathcal{L}_{zz} \cdot V''(P)$ since the term in square brackets is equal to the scalar $V''(P)$

by the considerations above. As a result, we arrive at the more tractable expression

$$V''(P) = \frac{\det \mathcal{L}_{(z,P)(z,P)}}{\det \mathcal{L}_{zz}} \tag{3.A.2}$$

where both determinants are evaluated in $(z^*(v,P), v, P)$. The remainder of the proof verifies that $V''(P^*) > 0$. In a point $(z^*(v,P), v, P)$ the Hessian of $\mathcal{L}$ with respect to z can be written

$$\mathcal{L}_{zz} = \begin{pmatrix} 0 & & & \\ -\frac{(1-F(\underline{y}))q(\theta)}{1-\tilde{\beta}(1-\phi F(\underline{y}))} & \frac{(1-F(\underline{y}))p(\theta)u''(w)}{1-\tilde{\beta}(1-\phi F(\underline{y}))} & & \\ -\frac{f(\underline{y})(1-\tilde{\beta}(1-\phi))(J+P)q(\theta)}{1-\tilde{\beta}(1-\phi F(\underline{y}))} & 0 & -\frac{f(\underline{y})\lambda q(\theta)}{1-\tilde{\beta}(1-\phi F(\underline{y}))} & \\ q'(\theta)\mathbb{E}J & \frac{(1-F(\underline{y}))q(\theta)u'(w)}{1-\tilde{\beta}(1-\phi F(\underline{y}))} & \frac{f(\underline{y})(1-\tilde{\beta}(1-\phi))q(\theta)u'(w)(J+P)}{1-\tilde{\beta}(1-\phi F(\underline{y}))} & -q'(\theta)u'(w)\mathbb{E}J \end{pmatrix}$$

where $J := J(w, \underline{y})$ and for simplicity I omit all asterisks. The determinant is

$$\begin{aligned} \det \mathcal{L}_{zz} = & \frac{f(\underline{y})(1-F(\underline{y}))^2\lambda q(\theta)^3 u'(w) q'(\theta)\mathbb{E}J}{[1-\tilde{\beta}(1-\phi F(\underline{y}))]^3} + \frac{f(\underline{y})(1-F(\underline{y}))p(\theta)u''(w)q'(\theta)^2\mathbb{E}J^2\lambda q(\theta)}{[1-\tilde{\beta}(1-\phi F(\underline{y}))]^2} \\ & - \frac{f(\underline{y})^2(1-F(\underline{y}))(1-\tilde{\beta}(1-\phi))^2 p(\theta)q(\theta)^2 u''(w)q'(\theta)\mathbb{E}Ju'(w)(J+P)^2}{[1-\tilde{\beta}(1-\phi F(\underline{y}))]^3}. \end{aligned}$$

Since $u''(w) < 0$ and $q'(\theta) < 0$ each of the first two terms in the sum are positive, while the third term is negative. Therefore, $\det \mathcal{L}_{zz} < 0$, which also confirms that $\mathcal{L}_{zz}$ is indeed invertible in equilibrium, as repeatedly assumed before.

To set up the block matrix $\mathcal{L}_{(z,P)(z,P)}$ note that $\mathcal{L}_{PP} = \frac{F(\underline{y})(1-\tilde{\beta}(1-\phi))p(\theta)u''(w)}{1-\tilde{\beta}(1-\phi F(\underline{y}))}$ and

$$\mathcal{L}_{zP} = \begin{pmatrix} -\frac{F(\underline{y})(1-\tilde{\beta}(1-\phi))q(\theta)}{1-\tilde{\beta}(1-\phi F(\underline{y}))} & 0 & 0 & \frac{F(\underline{y})(1-\tilde{\beta}(1-\phi))q(\theta)u'(w)}{1-\tilde{\beta}(1-\phi F(\underline{y}))} \end{pmatrix}.$$

At an optimum $P = P^*$, the determinant of the block matrix $\mathcal{L}_{(z,P)(z,P)}$ can be written

$$\begin{aligned} \det \mathcal{L}_{(z,P)(z,P)} = & \frac{F(\underline{y})(1-\tilde{\beta}(1-\phi))p(\theta)u''(w)\det \mathcal{L}_{zz}}{1-\tilde{\beta}(1-\phi F(\underline{y}))} \\ & + \frac{F(\underline{y})^2 f(\underline{y})(1-F(\underline{y}))(1-\tilde{\beta}(1-\phi))^2\lambda q(\theta)^3 u'(w)p(\theta)u''(w)q'(\theta)\mathbb{E}J}{[1-\tilde{\beta}(1-\phi F(\underline{y}))]^4}. \end{aligned}$$

Since $u''(w) < 0$, $q'(\theta) < 0$, and $\det \mathcal{L}_{zz} < 0$, the determinant is strictly positive. Equation (3.A.2) then implies $V''(P^*) < 0$ and therefore $\frac{d^2 v^*(P)}{dP^2} < 0$ at $P = P^*$. □

Proof of Proposition 3.1. Assume that $W_t^* = \Delta_t^*$ for all $t \geq T$. By equation 3.5, this is equivalent to $w_t^* = w^*$ for $t \geq T-1$. This results in constant severance pay $P_t^* = P^*$ and $\Delta_t^* = \Delta^*$ for $t \geq T$ by (3.8). Furthermore, $W_t^* = \Delta^*$ and $\mathbb{E}W_t^* = \Delta^*$ for $t \geq T$. Worker surplus in period T is $W_T^* = u(w^*) - u(b) + \tilde{\beta}[\mathbb{E}W_{T+1}^* - v] = \Delta^* + \tilde{\beta}[\Delta^* - v]$ since $w^* = b + P^*$ by (3.9). Since $W_T^* = \Delta^*$ it must hold that $v = \Delta^*$. Worker surplus in period $T-1$ is $W_{T-1}^* = u(w^*) - u(b) + \tilde{\beta}[\mathbb{E}W_T^* - v] = \Delta^*$.

Therefore, $W_t^* = \Delta^*$ for all $t \geq T-1$. By induction, this line of arguments can be used to show that $W_t^* = \Delta^*$ for all $t \geq 0$. As a result, $\mathbb{E}W_0^* = \Delta^*$. In equilibrium, $v = p(\theta^*)\mathbb{E}W_0^* = p(\theta^*)\Delta^*$. Since $v = \Delta^*$ and $p(\theta^*) < 1$ it must hold that $\mathbb{E}W_0^* = 0$. The latter contradicts (3.4) as $\mathbb{E}J^* > 0$ by the free entry condition. □

3.B Deriving the value functions with fully persistent draws

Assumption 3.2 implies that for every $t < T$ a match with productivity $\underline{y}_t^*$ is dissolved in the next period. Period t firm surplus is therefore $J_t(\underline{y}_t^*) = \underline{y}_t^* - w_t - \tilde{\beta}P_{t+1}$ and the separation threshold must satisfy $\underline{y}_t^* = w_t - P_t + \tilde{\beta}P_{t+1}$. Matches with $y \geq \sup_t \underline{y}_t^*$ last forever. If $T < \infty$, the supremum is attained in period T, such that $\sup_t \underline{y}_t^* = \underline{y}_T^*$. This implies $J_T(\underline{y}_T^*) = \sum_{s=T}^{\infty} \tilde{\beta}^{s-T}(\underline{y}_T^* - w_s)$. The separation threshold then satisfies $\underline{y}_T^* = (1-\tilde{\beta})\big[\sum_{s=T}^{\infty} \tilde{\beta}^{s-T} w_s - P_T\big]$. Expected firm surplus in period 0 can be decomposed as

$$\mathbb{E}J_0(w, \underline{y}^*) = \int_{-\infty}^{\infty} J_0(w,y)\, dF(y) = \sum_{t=0}^{T} \int_{\underline{y}_{t-1}^*}^{\underline{y}_t^*} J_0(w,y)\, dF(y) + \int_{\sup_t \underline{y}_t^*}^{\infty} J_0(w,y)\, dF(y) \tag{3.B.1}$$

where $\underline{y}_{-1}^* := -\infty$. The expressions for firm surplus are different in each of the integral terms in equation (3.B.1) due to different employment durations. Matches with productivity in $[\underline{y}_{t-1}^*, \underline{y}_t^*)$ are dissolved in period t. Discounted to period 0, firm surplus is $J_0(w,y) = \sum_{s=0}^{t-1} \tilde{\beta}^s(y - w_s) - \tilde{\beta}^t P_t = \sum_{s=0}^{t-1} \tilde{\beta}^s(y - \underline{y}_s^*) - P_0$. If $y \geq \sup_t \underline{y}_t^*$, the match only ends for exogenous reasons, and surplus at the beginning of the match is $J_0(w,y) = \sum_{s=0}^{\infty} \tilde{\beta}^s(y - w_s) = \sum_{s=0}^{T-1} \tilde{\beta}^s(y - \underline{y}_s^*) + \tilde{\beta}^T \frac{y - \underline{y}_T^*}{1-\tilde{\beta}} - P_0$. Substituting the expressions for firm surplus into the right-hand side of (3.B.1) and exchanging the order of summation yields (3.11). Expected worker surplus is derived in the same way.

4 Size and persistence matters: Wage and employment insurance at the micro level

4.1 Introduction

It has long been known that wages of job stayers fluctuate relatively little with economic conditions. This applies both at the macro level with respect to the business cycle (Bils, 1985; Devereux, 2001; Haefke et al., 2013) and at the micro level with respect to idiosyncratic firm-specific performance (Bronars and Famulari, 2001; Guiso et al., 2005). This study contributes to the latter line of literature by jointly analyzing the role that persistence and size of idiosyncratic productivity shocks have on the ability of firms to insure their workers against wage and employment fluctuations.

The separate effect of shock persistence and shock size on wage insurance is relatively well understood. The seminal study of Guiso et al. (2005) uses time-series based methods to distinguish between temporary and permanent changes in productivity, and estimates micro wage elasticities separately for each type of shock. The authors find that workers' wages are fully insured against temporary productivity shocks, and only permanent shocks affect wages. Yet, the estimated wage elasticity of 0.0686 indicates a considerable degree of smoothing of productivity fluctuations.[1] A complementary view on wage flexibility at the micro level is given by the International Wage Flexibility Project (Dickens et al., 2007), who stress asymmetries in wage changes. To this purpose, Dickens and Goette (2006) develop a histogram based approach that compares the observed distribution of wage changes to a hypothetical symmetric distribution. The more right-skewed the observed distribution, the more pronounced is downwards wage rigidity. Cross-country comparisons show that downwards wage rigidity is a general property of employment relations. Whether it applies to the nominal wage or to the real wage depends on labor market institutions and country-specific wage setting practices (Messina et al., 2010). The extent of downwards wage rigidity varies with firm and worker characteristics (Du Caju et al., 2007).

The goal of this essay is to combine the insights of these two strands of literature. It analyses jointly the role played by the persistence of a shock and the direction of a shock (or more

[1] Replication studies for Portugal (Cardoso and Portela, 2009) and Germany (Guertzgen, 2014) reach similar conclusions. Kátay (2016) finds imperfect insurance with respect to transitory shocks in Hungary.

generally the size of a shock) in wage insurance at the firm level. Since firms can also adjust to negative productivity shocks through downsizing, the effect on layoffs is investigated as well. The estimation strategies of the above mentioned papers cannot be used to address this question. The histogram approach of Dickens and Goette (2006) can characterize asymmetries in the cross-section but does not explicitly link them to firm-specific shocks. The time-series approach of Guiso et al. (2005), on the other hand, allows to estimate wage elasticities that vary by shock persistence, but identification hinges on the assumption that wages react linearly to productivity shocks. The estimation strategy used in this essay draws on their methodology. In a first step, I perform a productivity regression similar to theirs and use the same time-series based arguments to identify transitory and permanent productivity changes. The difference to Guiso et al. (2005) comes in the second step at which individual wage elasticities are estimated. A neat property of their approach is that the two wage elasticities with respect to transitory and permanent shocks are separately identified by the change in *total* productivity and certain orthogonality conditions. It is therefore not necessary to explicitly decompose the total productivity change into its unobserved transitory and permanent components. For this reason, I refer to the method of Guiso et al. (2005) as the *indirect method* in the following. Identification of the wage elasticities fails if productivity affects wages in a nonlinear way or if the required orthogonality conditions do not hold. The alternative approach proposed in this essay is the *direct method*. The observed total productivity change is decomposed into its unobserved transitory and permanent components using a linear Kalman smoother. The predicted components can then be explicitly included in wage and/or layoff regressions. This allows to estimate flexible functional relations between wage and productivity changes, and even to perform semiparametric estimation.

Using rich matched employer-employee data from Germany, I find that the direct method delivers similar results as the indirect method if wage responses are assumed to be linear and the indirect method is able to identify the effects. The data, however, indicates the presence of nonlinearities: the elasticity of wages depends on the size of the productivity shock. In particular, I detect stronger wage rigidity for tail events, i.e. shocks in the lowest and highest decile of the shock distribution. With respect to permanent shocks, wages react largely symmetrically to positive and negative shocks. Between the 10th and the 90th percentile the wage elasticity is constant at 0.11. Transitory productivity shocks lead to asymmetric wage responses. Negative shocks tend to reduce wages, while positive shocks are fully captured by the firm. These general patterns hide important heterogeneity at the worker level. The downward wage flexibility with respect to both types of shocks is in fact limited to blue-collar workers. Whereas wages of white-collar workers do not respond to negative shocks and appear to be perfectly downwards rigid.

Firms also adjust to shocks by dismissing workers, but only in response to negative permanent shocks. Linear probability regressions at the worker level reveal an elasticity of 1.44 between layoff probability and shock size. This increase in layoff probability is again limited to blue-collar workers, while white-collar workers enjoy perfect employment insurance.

The essay proceeds as follows. Section 4.2 summarizes theoretical results on wage and employment insurance. Section 4.3 gives an overview of the data. The econometric analysis is conducted in Section 4.4. It introduces the new estimation method and applies it to wage and layoff data. Section 4.5 concludes.

4.2 Theoretical considerations

This section reviews important theoretical results considering wage and employment responses to idiosyncratic shocks at the firm level. In a frictionless labor market, workers would be perfectly insured again idiosyncratic shocks. The reason is that firms employ workers up to the point where the marginal product of labor equals the market wage. Any deviation from the market wage leads to immediate worker relocation, such that in equilibrium the marginal product of labor (MPL) is equalized across firms. Since exogenous idiosyncratic productivity shocks do not affect the market wage, employment adjusts to keep MPL constant. Hence even if idiosyncratic productivity is very volatile, neither MPL nor wages should change over time, and all adjustment is via employment. The empirical facts, in particular the observation of large and persistent fluctuations in firm-specific MPL over time (see Guiso et al., 2005, and others), indicate the importance of reallocation frictions.

With search and matching frictions in the spirit of Pissarides (1990), wage responses to idiosyncratic productivity shocks depend on the particular wage-setting mechanism. In Germany, as in most European countries, unions play an important role in wage-setting through collective bargaining. Although strictly speaking, a collective bargaining agreement (CBA) only applies between members of the negotiating parties—the employer association and the labor union—, collectively bargained wages are generally extended to non-unionized labor in covered firms as well (Guertzgen, 2009). CBAs typically set a wage floor as well as a minimum wage increase for all covered workers, which is likely to generate downwards wage rigidity. Since collective bargained wage increases typically compensate for (expected) inflation, Dickens et al. (2007) and Babecký et al. (2010) find that real wage rigidity is more pronounced than nominal wage rigidity in countries with more centralized bargaining. Yet, even if hourly wages are rigid downwards, firms can adjust their wage cost at other margins, such as overtime hours or bonus payments. Additionally, so called opening clauses allow covered firms in Germany to pay below the CBA level under certain conditions, which brings further wage flexibility and helps to avoid layoffs (Brändle and Heinbach, 2013).

Even without institutional constraints, contracted wages are likely to be rigid due to other considerations. The literature on implicit contracts evolves around the idea that risk averse workers do not have access to the capital market, and risk neutral firms insure them against income fluctuations by paying a constant wage (Baily, 1974; Azariadis, 1975). Financing constraints of the firm might limit the scope of insurance that the employer can provide, such that sufficiently large productivity shocks may make wage adjustments necessary. However, since work effort and job search behavior of employees are hardly observed, the feasibility of downwards wage adjustments is limited by their adverse effects on motivation and quits.[2] A firm survey conducted by Du Caju et al. (2015) confirms that employers indeed worry about the motivational impact of wage cuts as well as their effect on quit rates of productive workers. Therefore, an optimal response to a big negative productivity shock might combine a relatively small decrease in wages of stayers and a shrinking of the workforce through layoffs.[3]

But not all workers might be affected by wage cuts and layoffs to the same extent. Along the lines of Shapiro and Stiglitz (1984), wage cuts increase the incentive to shirk, which is more important in occupations where employee effort is hard to monitor. Since these are typically white-collar occupations, white-collar workers might enjoy more wage insurance against negative shocks than blue-collar workers. In certain occupations, blue-collar workers may even earn piece-rates instead of an hourly wage. Wage cuts also make quits more likely, which requires hiring and training of new workers (Stiglitz, 1974). Since training costs are usually higher for white-collar workers and it may take longer to find an adequately skilled replacement, this provides another rationale for more downwards wage rigidity of white-collar workers. For the same reason, white-collar workers may be less likely to be laid off if the firm experiences trouble. Additionally, Oi (1962) argues that some groups of workers are more complementary in the production process than others. Employment adjusts mainly through hiring and firing of workers that are relatively easy to substitute by other fixed production factors. Along these lines, blue-collar workers performing manual tasks may be more substitutable to capital than white-collar workers who perform cognitive tasks.

4.3 Data and sample selection

The study uses the longitudinal version of the linked employer-employee data of the IAB (LIAB), see Alda et al. (2005). The data set is administered by the Institute for Employment Research (IAB) and allows for simultaneous analysis of the supply and demand side of the German labor market from 1993 until 2010. On the employer side, the LIAB uses the representative annual survey data of the IAB establishment panel. This panel entails questions on sales, investment, employment, and industrial relations. The individual level uses official data of the employment register. Information on wages, occupation, qualification, gender, tenure, experience, and age are linked to the employer data by a common identifier. The unit of observation is an establishment, which mostly corresponds to a plant or a branch. Since it is unknown which establishments belong to the same firm, the econometric analysis is at the establishment level.

The last wave considered in the analysis is 2009, which contains retrospective information on investment and sales in 2008. I do not use the latest available wave as in 2009 many establishments implemented special employment and wage policies to tackle the Great Recession,

[2] Classical papers in this vein include Weiss (1980), Akerlof (1982), Shapiro and Stiglitz (1984), Lindbeck and Snower (1989), and Akerlof and Yellen (1990).

[3] Menzio and Moen (2010) present a model which gives rise to a similar optimal firm policy. Rather than considering unobserved work or search effort, they impose the constraint that firms should never have an incentive to replace an incumbent worker with a newly hired one. The incentive to attract new workers with high starting wages makes wages of stayers downward rigid, such that employment adjusts more strongly to negative shocks.

such as short-time work (Brenke et al., 2013). At a smaller scale, short-time work (STW) is also used during normal times, compare Balleer et al. (2016). To the extent that STW provides an additional facility for troubled firms to overcome severe idiosyncratic shocks at conditions that vary little over time, including observations of establishments that adopt short-time work yields a more complete picture of wage and employment insurance in Germany.[4] To avoid bias, however, periods during which increased take-up of STW is mainly due to discretionary changes in the STW policy should be excluded. As argued by Balleer et al. (2016), the Great Recession was such a period.[5]

Only privately-owned establishments in the private, non-financial sector are included in the analysis. The financial industry has to be excluded because no sensible productivity measure is available. I exclude very small establishments that in some year report less than 5 employees. Establishments with consistently missing information on sales, investment or employment are ignored. Because I perform a dynamic panel regressions on the establishment data, at least three consecutive observations are required per establishment. Altogether, the establishment-level regressions are based on 2697 establishments, see the first column of Table 4.A.1.

On the worker side, only male employees up to age 59 are considered due to a spike in separation rates at age 60. Women are excluded because the LIAB does not provide information about the nature of a separation (voluntary quit or involuntary layoff). As a workaround, Section 4.4.3 uses transitions from employment to non-employment to proxy employer-induced layoffs, in line with Boockmann and Steffes (2010). While this appears to be a reasonable proxy for men because of their high attachment to the labor market, it is less convincing for women. Compared to men, female transitions to non-employment are more often driven by personal or family-related reasons, such as labor supply of the spouse, child care, or informal care for a relative. Since neither of these variables is observed in the LIAB, separations for family-related reasons would be incorrectly labeled as employer-induced layoffs. Including women in the wage regression is less problematic and presented as a robustness check.[6]

Since no information on hours worked is available, the analysis is restricted to full-time employment. This restriction could bias my estimates if in response to productivity shocks workers switch back and forth between full-time to part-time employment. Over the whole sample period, however, less than 5% of male workers are observed to switch between these two employment states. The majority of these switches occur after age 50 and result in a permanent reduction of working time.

By nature of the data, wages are top-coded at the social security threshold. This applies to 16% of the observations. Observations with censored wages are excluded from the wage regressions, but are included in the layoff regressions. The respective sample statistics can be

[4]Moreover, establishment-level information on STW take-up is only available in few waves of the IAB panel.

[5]Balleer et al. (2016) also report the frequent use of STW at the beginning of the 1990s in East Germany. Since East German establishment are contained in the IAB panel only from 1996 onwards, the effect on the estimation results is likely to be small. Additionally, by construction of the LIAB the bulk of observations stems from the period 2000 to 2008, see also Table 1 in Klosterhuber et al. (2013).

[6]If women are nevertheless included in the layoff regressions, coefficient estimates in Table 4.6 change little, while standard errors increase substantially.

found in Table 4.A.1. Since all regressions are in first differences, only workers with at least two consecutive observations at the same establishment are considered. Nominal variables were deflated using the consumer price index with base year 2010.[7]

4.4 Econometric analysis

The econometric analysis proceeds in three steps. The first step uses establishment-level data to identify idiosyncratic shocks to productivity and describe their statistical properties. This closely follows Guiso et al. (2005) and Guertzgen (2014). The second step applies wage regressions at the worker level to estimate wage elasticities. After a short review of the indirect method, the direct method is introduced. The two methods are then compared to each other assuming linear wage responses to productivity shocks. Thereafter, nonparametric and piecewise linear relations are considered. In a third step, the analysis is extended to layoffs.

4.4.1 Productivity regressions

Productivity is measured in terms of sales per worker $y_{jt} = Y_{jt}/L_{jt}$, where Y_{jt} refers to the value of sales in year t, and L_{jt} is the stock of employees at June 30 of year t. Both figures are taken from the IAB establishment panel. Therefore, L_{jt} measures the total workforce of an establishment and not only those workers that satisfy the sample selection criteria outlined in Section 4.3. From a theoretical point of view, using value added instead of sales would be preferable because it better captures establishment-level quasi-rents. The LIAB allows to construct value added by multiplying the value of sales with the reported share of material costs in total sales. However, Addison et al. (2006, p.260) argue that "unlike the sales measures, these share-in-sales values seem to be little more than 'informed guesstimates.'" This is because the majority of values take multiples of 5 percent, and there is unrealistically high variation in these shares over time. For this reason, but also since previous studies have found little difference between using value added and sales for estimating wage insurance at the firm level, my productivity measure is based on sales.[8]

To isolate idiosyncratic shocks, establishment productivity is regressed on a set of dummies that capture the aggregate cycle as well as industry- and region-specific effects. To ensure that the unexplained changes in sales per worker stem from exogenous shocks rather than variation of factor inputs, I additionally control for the capital-labor ratio $k_{jt} = K_{jt}/L_{jt}$. A proxy for the capital stock K_{jt} of an establishment is calculated from reported investment data as explained

[7]Missing wage information is sometimes imputed using a hypothetical model for wage determination, which is mostly a Mincerian wage equation. The goal of this essay, however, is exactly to come up with a model that explains wage formation by additionally taking into account firm performance.

[8]Guiso et al. (2005) and Kátay (2016) use value added or value added per worker, while Cardoso and Portela (2009) use sales. Guertzgen (2014) conducts a similar analysis with an earlier version of the LIAB, using value added as performance indicator. The estimated variance ratio between transitory shocks and permanent shocks is about seven times higher than in the comparable studies of Guiso et al. (2005), Cardoso and Portela (2009) or Kátay (2016). Along the lines of Addison et al. (2006), most of the excess volatility in transitory shocks may be due to measurement error in the share of material costs.

	coefficient	std. err.
$\ln y_{jt-1}$	0.2101***	0.0376
$\ln k_{jt}$	0.3173***	0.0285
	χ^2-statistic	p-value
year dummies	95.72***	0.000
industry dummies	39.83***	0.000
regional dummies	10.54	0.837
	statistic	p-value
AR(2) test	1.32	0.186
AR(3) test	−0.83	0.407
AR(4) test	1.11	0.267
Hansen J test	39.23	0.415
establishments (observations)	2697 (17407)	

two-step difference GMM, corrected standard errors clustered at the establishment level, significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.1. Productivity regression

in Guertzgen (2014). To capture predictable dynamics such as precommited sales, the estimated model specification is autoregressive,

$$\ln y_{jt} = \rho \ln y_{jt-1} + \alpha \ln k_{jt} + Z'_{jt}\gamma + \varphi_j + \varepsilon_{jt}. \tag{4.1}$$

where ϕ_j is an establishment-specific intercept. The matrix Z_{jt} contains linear time trends interacted with year dummies (14), industry dummies (15), and regional dummies (18). These become regular dummy variables in the first differenced equation. It is important for the interpretation of the wage elasticities that the residuals of the productivity regression indeed capture exogenous productivity variation. For this reason, equation (4.1) is based on theoretical considerations and derived form a Cobb-Douglas production function at the establishment level. As demonstrated in Section 4.B, the exact representation includes the level of employment L_{jt} as additional explanatory variable, unless the production function features constant returns to scale. Table 4.B.1 shows that constant returns to scale cannot be rejected. The level of employment is therefore omitted from the productivity regression altogether.

Equation (4.1) is estimated in first differences,

$$\Delta \ln y_{jt} = \rho \Delta \ln y_{jt-1} + \alpha \Delta \ln k_{jt} + \Delta Z'_{jt}\gamma + \Delta\varepsilon_{jt},$$

using the Arellano and Bond (1991) GMM estimator, where $\Delta \ln y_{jt-1}$ is instrumented with lags 2 to 4 of $\ln y_{jt}$. Table 4.1 presents the two-step GMM estimates that account for clustering at the establishment level.[9] The point estimate of the autoregressive coefficient is 0.21, and the

[9]These were obtained using the user-written `xtabond2` command in Stata (Roodman, 2009b). Reported standard errors use Windmeijer's (2005) correction and are clustered at the establishment level.

order (k)	$\mathbb{E}[\Delta\hat{\varepsilon}_{jt}\Delta\hat{\varepsilon}_{jt-k}]$	std. err.
0	0.0795^{***}	0.0038
1	-0.0344^{***}	0.0024
2	0.0018	0.0012
3	-0.0009	0.0011
4	0.0013	0.0013
5	-0.0018	0.0018

standard errors bootstrapped with clustering at the establishment level, significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.2. Residual autocovariance structure of GMM regression of sales per worker

coefficient of the capital-labor ratio is 0.32. Both values are highly significant and within the credible range. The AR tests and the Hansen J test confirm that the second to fourth lags are valid instruments. Furthermore, a difference-in-Hansen test (not reported) verifies that $\ln k_{jt}$ can be treated as exogenous, which proofs to be robust to different choices of instrument sets (fewer instrument lags and/or collapsed instruments) as recommended by Roodman (2009a). This confirms that the residuals of the GMM estimation can indeed be regarded as idiosyncratic shocks to productivity that are exogenous to the establishment.

The autocovariance structure of the first differenced GMM residuals $\Delta\hat{\varepsilon}_{jt}$ is given in Table 4.2. Similar to Guiso et al. (2005) this information can be used to identify the stochastic process that generates the productivity shocks. Because the covariance estimates at lags greater than one are close to zero and statistically insignificant, Table 4.2 suggests the following error process,

$$\begin{aligned} \varepsilon_{jt} &= \zeta_{jt} + \tilde{v}_{jt}, \\ \zeta_{jt} &= \zeta_{jt-1} + \tilde{u}_{jt}, \end{aligned} \tag{4.2}$$

where $\tilde{v}_{jt}$ and $\tilde{u}_{jt}$ are mutually uncorrelated white noise processes with variances $\mathbb{E}\tilde{v}_{jt}^2 = \sigma_{\tilde{v}}^2$ and $\mathbb{E}\tilde{u}_{jt}^2 = \sigma_{\tilde{u}}^2$. The structural equations imply $\mathbb{E}[\Delta\varepsilon_{jt}\Delta\varepsilon_{jt-1}] = -\mathbb{E}\tilde{v}_{jt}^2 = -\sigma_{\tilde{v}}^2$ and $\mathbb{E}[\Delta\varepsilon_{jt}(\Delta\varepsilon_{jt-1} + \Delta\varepsilon_{jt} + \Delta\varepsilon_{jt+1})] = \sigma_{\tilde{u}}^2$. Computing the respective sample moments from the data yields variance estimates $\hat{\sigma}_{\tilde{v}}^2 = 0.0344$ and $\hat{\sigma}_{\tilde{u}}^2 = 0.0088$, which are both significantly different from zero at the 1% level. In line with previous literature, shocks to establishment productivity have a transitory and a permanent component.[10]

By virtue of (4.2), establishment productivity can be decomposed into a deterministic component D_{jt}, a non-stationary stochastic component P_{jt}, and a stationary stochastic component T_{jt},

$$\ln y_{jt} = D_{jt} + P_{jt} + T_{jt}$$

[10] Guertzgen (2009) identifies the same error process and obtains the variance estimates $\hat{\sigma}_{\tilde{v}}^2 = 0.1464$ and $\hat{\sigma}_{\tilde{u}}^2 = 0.010$. While the variance of the permanent innovation is similar, the variance of the transitory term is more than four times higher. This suggests that using value added instead of sales results in substantial measurement error (compare footnote 8).

where $D_{jt} := (1-\rho L)^{-1}(Z'_{jt}\gamma + \varphi_j)$, $P_{jt} := (1-\rho)^{-1}\zeta_{jt}$, $T_{jt} := (1-\rho L)^{-1}[\tilde{v}_{jt} - (1-\rho)^{-1}\rho\tilde{u}_{jt}]$, and L denotes the lag operator. This is an application of the Granger representation theorem, compare Guiso et al. (2005). The definitions of P_{jt} and T_{jt} imply

$$\begin{aligned} \Delta P_{jt} &= (1-\rho)^{-1}\tilde{u}_{jt} =: u_{jt}, \\ \Delta T_{jt} &= (1-\rho L)^{-1}\Delta v_{jt}, \end{aligned} \tag{4.3}$$

where u_{jt} and $v_{jt} := \tilde{v}_{jt} - \rho u_{jt}$ are the innovations to the permanent and transitory productivity component, respectively. Note that the year-on-year change in the stochastic productivity components is related to the total productivity shock by $\Delta P_{jt} + \Delta T_{jt} = (1-\rho L)^{-1}\Delta\varepsilon_{jt}$.

4.4.2 Wage responses

This section relates variation in wages that is unexplained by other observables to productivity shocks of the employer. Section 4.4.2 assumes that wages respond linearly to productivity shocks. After reviewing the indirect method, I introduce the direct method to estimate wage elasticities. The linearity assumption allows to compare the results of the two methods. The analysis is then generalized in Section 4.4.2 to nonlinear relationships between productivity shocks to wage innovations.

Linear wage responses

Guiso et al. (2005) propose a wage equation of the form

$$\ln w_{ijt} = X'_{ijt}\delta + \alpha P_{jt} + \beta T_{jt} + \phi_{ij} + \psi_{ijt}, \tag{4.4}$$

where w_{ijt} refers to the annual average wage income that worker i earns from establishment j in year t. In the LIAB data, this income measure includes all bonus payments that a worker receives on top of the base wage. The matrix X_{ijt} contains the same dummy variables as the productivity regression, as well as dummies for industrial relations, educational dummies, a white-collar dummy, a cubic polynomial in age, and a cubic polynomial in tenure.[11] The intercept ϕ_{ij} captures an unobserved fixed effect at the establishment, worker, or match level. First differencing of (4.4) yields

$$\Delta \ln w_{ijt} = \Delta X'_{ijt}\delta + \alpha\Delta P_{jt} + \beta\Delta T_{jt} + \Delta\psi_{ijt} \tag{4.5}$$

where ΔP_{jt} and ΔT_{jt} are unobserved but known to satisfy the structural equations (4.2)–(4.3). The first differencing implies that only wages of job stayers can be analyzed.

[11] The capital-labor ratio and the level of employment are excluded from (4.4) due to endogeneity problems. However, since $\Delta\varepsilon_{jt}$ is by construction orthogonal to these variables, their appearance in the regression has almost no effect on the estimates of α and β.

The indirect method. To identify the wage elasticities α and β in (4.5), Guiso et al. (2005) use an indirect approach that avoids determining the unknown productivity components ΔP_{jt} and ΔT_{jt}. They proceed in two steps: First, wage changes $\Delta \ln w_{ijt}$ are regressed on the set of observed characteristics ΔX_{ijt}, i.e. $\Delta \ln w_{ijt} = \Delta X_{ijt}'\delta + \Delta\omega_{ijt}$. By equation (4.5), the error term in this regression satisfies $\Delta\omega_{ijt} = \alpha\Delta P_{jt} + \beta\Delta T_{jt} + \Delta\psi_{ijt}$. Substituting (4.3) and applying the operator $(1-\rho L)$ on both sides yields

$$\Delta\tilde{\omega}_{ijt} := (1-\rho L)\Delta\omega_{ijt} = \alpha(1-\rho L)u_{jt} + \beta\Delta v_{jt} + (1-\rho L)\Delta\psi_{ijt}. \tag{4.6}$$

Since $\mathbb{E}[(\Delta\tilde{\omega}_{ijt} - \beta\Delta v_{jt})\Delta\varepsilon_{jt+1}] = 0$ and $\mathbb{E}[\Delta\varepsilon_{jt}\Delta\varepsilon_{jt+1}] = -\sigma_{\tilde{v}}^2$, the wage elasticity with respect to a transitory shock, β, can be identified using an IV regression of $\Delta\tilde{\omega}_{ijt}$ on $\Delta\varepsilon_{jt}$, using $\Delta\varepsilon_{jt+1}$ as instrument. Likewise, it can be shown that IV regression of $\Delta\tilde{\omega}_{ijt}$ on $\Delta\varepsilon_{jt}$, instrumented by $\Delta\varepsilon_{jt-1}+\Delta\varepsilon_{jt}+\Delta\varepsilon_{jt+1}$ identifies the wage elasticity with respect to a permanent shock α.

This method has two limitations. First, it is not possible to allow ΔP_{jt} and ΔT_{jt} to enter equation (4.5) nonlinearly because the exclusion restrictions of the IV no longer holds. There is no straightforward way of extending the methodology to a nonlinear setting. The second limitation is more subtle. As with any IV regression, the resulting estimates for α and β are biased in finite samples. This bias may be substantial if instruments are weak, even for samples of the size considered in this essay (see also the discussion starting on page 104).

The direct method. To overcome the linearity restriction as well as the potential identification problem, I use a more direct route. Exploiting (4.2)–(4.3), the residuals of the productivity regression are used to predict ΔT_{jt} and ΔP_{jt} by a linear Kalman smoother.[12] These predictions can be substituted into (4.5), from which the wage elasticities α and β can be estimated by OLS. Moreover, once the predictions for the stochastic components of productivity ΔT_{jt} and ΔP_{jt} have been obtained, any functional relation between wage changes and productivity shocks can be estimated.

To apply Kalman smoothing, the non-stationary process (4.2) is first differenced and written in state-space form for $2 \le t \le T_j$:

$$\Delta\varepsilon_{jt} = \begin{pmatrix} 1 & -1 \end{pmatrix} z_{jt} + \tilde{u}_{jt},$$

$$z_{jt} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} z_{jt-1} + \begin{pmatrix} \tilde{v}_{jt} \\ 0 \end{pmatrix},$$

where $z_{jt} := (\tilde{v}_{jt}, \tilde{v}_{jt-1})'$ is the unobserved state vector and T_j denotes the number of years between the first and the last observation of establishment j.[13] If the innovation variances $\sigma_{\tilde{u}}^2$

[12] While the Kalman filter uses past information to form optimal predictions about the future, the Kalman smoother uses all available information to form predictions. Hamilton (1994) presents an overview of these and other state-space methods.

[13] Gaps can easily be handled by the Kalman smoothing algorithm. In periods with missing $\Delta\varepsilon_{jt}$ only the state

and $\sigma_{\tilde{v}}^2$ are known, the best linear predictions for $\tilde{u}_{jt}$ and z_{jt} given $(\Delta\varepsilon_{j2}, \ldots, \Delta\varepsilon_{jT_j})$ can be obtained by Kalman smoothing, irrespective of the actual error distribution (Hamilton, 1994). Since the true errors $\Delta\varepsilon_{jt}$ are unobserved, they are replaced with the residuals of the productivity regression. Applying the Kalman smoother separately for every establishment then yields predictions for the fundamental shocks $(\tilde{u}_{jt})_{t=2}^{T_j}$ and $(\tilde{v}_{jt})_{t=1}^{T_j}$. Feeding them into (4.3) yields the time series $(\Delta P_{jt})_{t=2}^{T_j}$ and $(\Delta T_{jt})_{t=2}^{T_j}$.[14]

Apart from the state-space equations, the Kalman smoother requires knowledge of the shock variances $\sigma_{\tilde{u}}^2$ and $\sigma_{\tilde{v}}^2$. These are unknown and have to be estimated from the data. Hamilton (1994) suggests to estimate the variances by maximum likelihood, assuming that innovations are normally distributed. This is convenient since the log-likelihood is easy to evaluate if a Kalman filter has already been computed.[15] The accuracy of the Kalman smoothed time series hinges on accurate variance estimates. Assuming that all establishments draw their productivity innovations from the same distribution might be too restrictive. Guertzgen (2014) finds that the variance of permanent shocks tends to increase with establishment size, while the variance of transitory shocks decreases. I consider heteroscedasticity of the form

$$\ln \sigma_{\tilde{u}j}^2 = D_j' \lambda_{\tilde{u}}, \quad \ln \sigma_{\tilde{v}j}^2 = D_j' \lambda_{\tilde{v}}, \tag{4.7}$$

where D_j is a matrix of time-independent and exogenous establishment characteristics. In the baseline estimations, D_j contains dummies for the establishment size in the first period of observation. The parameter vectors $\lambda_{\tilde{u}}$ and $\lambda_{\tilde{v}}$ are estimated by Gaussian maximum likelihood following Hamilton (1994). As a robustness check, I obtain method-of-moments based variance estimates that do not require the normality assumption.

Comparison. Panel (a) of Table 4.3 compares the estimates for α and β obtained by the indirect method and the direct method. The indirect method uses the instruments described above, together with their 2nd, 3rd, and 4th power as proposed by Guiso et al. (2005). To account for heteroscedasticity at the establishment level, the coefficients were estimated by two-step efficient GMM. Two test statistics are reported in Table 4.3(a). Column F reports the Kleibergen-Paap Wald F statistic for detecting weak identification (Kleibergen and Paap, 2006). Column J reports the p-value of the Hansen test of overidentifying restrictions.[16] The estimated wage elasticity with respect to a permanent shock is 0.05 and strongly significant, while a transitory shock triggers no significant wage response.[17] The high values of the test statistics indicate that the instruments are valid and strong enough to identify the coefficients. The point estimates are in line with studies that apply the same methodology in other countries

equation is used for prediction until the next observation arrives.

[14]Note that both with the direct and the indirect approach one observation is lost, which is due to ΔT_{jt} being an AR(1) process.

[15]Computationally, the Kalman smoother is obtained by running a Kalman filter followed by a backwards pass, see (Hamilton, 1994, Section 2.4).

[16]Estimates and test statistics were obtained by the `ivreg2` command developed by Baum et al. (2007).

[17]All reported standard errors are based on 1000 bootstrap replications clustered at the establishment level. The bootstrap takes into account the uncertainty at each step of the multistage estimation procedure.

(a) baseline

	indirect method				direct method (ML)		direct method (MM)	
	coefficient	std. err.	F	J	coefficient	std. err.	coefficient	std. err.
α	0.0500***	0.0173	*23.14*	*0.747*	0.0625***	0.0143	0.0617***	0.0145
β	0.0162	0.0128	*28.13*	*0.523*	0.0189*	0.0102	0.0192*	0.0105

(b) using LIAB sampling weights

	indirect method				direct method (ML)		direct method (MM)	
	coefficient	std. err.	F	J	coefficient	std. err.	coefficient	std. err.
α	0.1046*	0.0626	*5.89*	*0.398*	0.0507***	0.0159	0.0490***	0.0147
β	0.0000	0.0096	*90.35*	*0.327*	0.0056	0.0075	0.0068	0.0079

F is the Kleibergen-Papp Wald F statistic, J is the p-value of the Hansen test, bootstrapped standard errors clustered at the establishment level, coefficient significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.3. Comparison of wage elasticity estimates

(Guiso et al., 2005; Cardoso and Portela, 2009; Kátay, 2016).[18]

For the direct method, two sets of results are reported. The innovation variances (or likewise the auxiliary parameters $\lambda_{\tilde{u}}$ and $\lambda_{\tilde{v}}$) are once estimated by Gaussian maximum likelihood (column ML) and once by method of moments (column MM).[19] Figure 4.1 shows the standard deviations estimated by the two methods if shocks are heteroscedastic with respect to establishment size, captured by four establishment size categories. In both cases the standard devision of permanent shocks increases with establishment size, while the smallest establishments experience the strongest transitory fluctuations. Apart from the smallest category, the estimated standard deviations are virtually identical. The wage elasticity estimates in Table 4.3(a) are therefore also very close to each other. The estimate for α is 0.062 and the estimate for β is 0.019. The estimated wage elasticities are therefore slightly higher than those obtained by the indirect method, but the difference is within one standard error. At the same time, the direct method yields smaller standard errors, which renders the estimate on β weakly significant.

The estimates obtained by the direct method are robust to alternative variance patterns, which can be seen from Table 4.D.2. Column (a) assumes that the variance of productivity shocks is the same across establishments. Column (b) allows for additional heteroscedasticity by considering size and industry dummies. In both cases, the point estimates are very similar to the baseline, while α is estimated less precisely.

Panel (b) of Table 4.3 repeats the same analysis, taking into account the sampling weights provided with the LIAB. The indirect method yields a point estimate for α that is more than twice as high as in the unweighted regression. However, this estimate is involved with a substantial standard error. The low F test statistic, which is well below the rule of thumb value of 10,

[18]The coefficient on permanent shocks is substantially higher than Guertzgen (2014) who reports $\alpha = -0.0307$ using an earlier version of the LIAB. See also footnote 8

[19]The method of moment estimates for $\sigma^2_{\tilde{v}j}$ and $\sigma^2_{\tilde{u}j}$ are based on the theoretical moment conditions $\sigma^2_{\tilde{v}j} = -\mathbb{E}[\Delta\varepsilon_{jt}\Delta\varepsilon_{jt-1}]$ and $\sigma^2_{\tilde{u}j} = \mathbb{E}[\Delta\varepsilon_{jt}(\Delta\varepsilon_{jt-1} + \Delta\varepsilon_{jt} + \Delta\varepsilon_{jt+1})]$, where the expected value is replaced by the sample average of all establishments in the same size category.

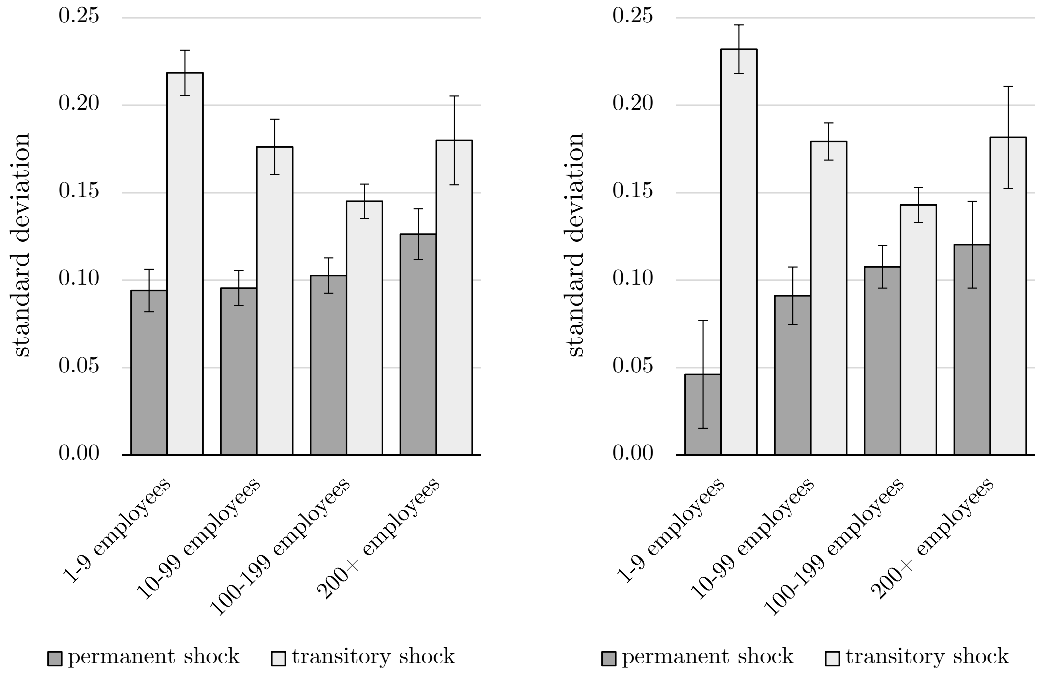


Figure 4.1. estimated standard deviation of the productivity shocks by establishment size: ML estimates (left), MM estimates (right), error bars indicate bootstrapped standard errors clustered at the establishment level

indicates weak instruments in the IV regression. Along the lines of Wooldridge (2002, p.108), this instrument weakness seems to generate substantial finite sample bias. The point estimate for α therefore bears little credibility. By contrast, the direct method continues to give plausible results. The point estimates are slightly lower compared to the unweighted estimates, while the standard errors are very similar.

That the indirect method may fail to identify the parameters of interest has not been documented so far. Yet, it is even possible to investigate analytically when this is likely to happen. In general, weak identification occurs if the first stage F test statistic is small or likewise if the R^2 statistic of the first stage regression is small. Section 4.C analyses the two first stage regressions used by the indirect method. It demonstrates that the population equivalents of the two first stage R^2 statistics depend only on the variance ratio $\phi := \sigma^2_{\tilde{u}}/\hat{\sigma}^2_{\tilde{v}}$. In particular, $R^2_\alpha = \phi^2/[(2+\phi)(2+3\phi)]$ and $R^2_\beta = 1/(2+\phi)^2$. As shown in Figure 4.C.1, R^2_α is increasing in ϕ while R^2_β is decreasing. In the baseline case without weighting, the variance ratio estimated from the data is $\phi = 0.79$.[20] The population R^2 statistics are then $R^2_\alpha = 0.051$ and $R^2_\beta = 0.129$, which seems sufficient to identify both parameters in Table 4.3(a). The sampling weights provided with the LIAB correct for the oversampling of large establishments which occurs by design of the IAB establishment survey. Observations of smaller establishments receive higher weights, while observations of larger establishments are downweighted. By Figure 4.1, smaller estab-

[20]The IV regression takes place at the worker level, where each establishment is essentially weighted by the number of its employees. The sample variances therefore do not coincide with the values reported in Section 4.4.1. In particular, $\sigma^2_{\tilde{u}} = 0.0112$ and $\sigma^2_{\tilde{v}} = 0.0142$.

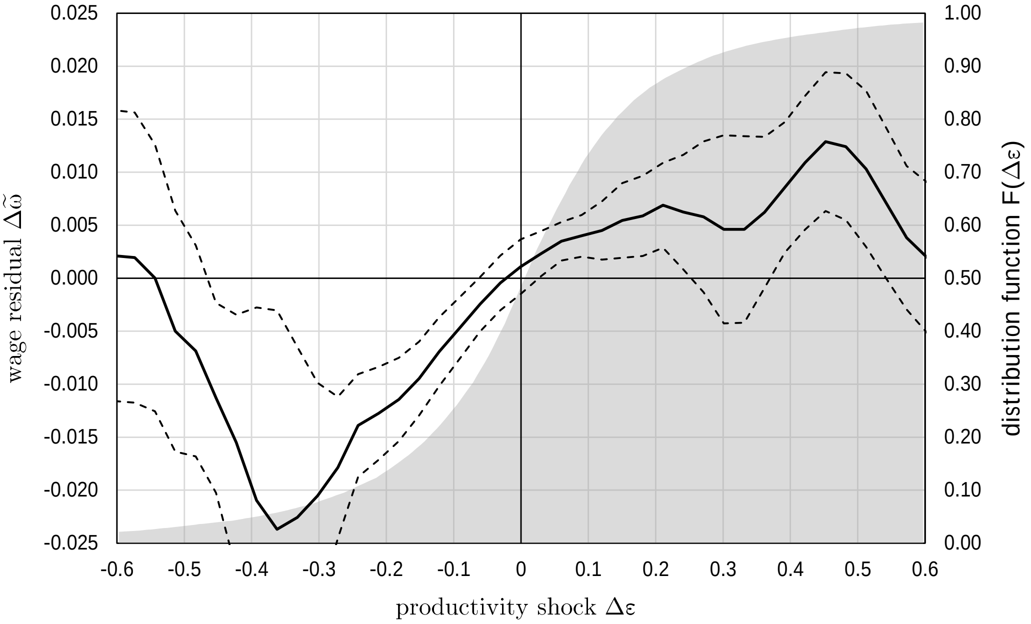
lishments experience a smaller variance of permanent shocks and larger variance of transitory shocks. Consequently, the variance ratio drops to $\phi = 0.29$ if weights are considered. This increases R^2_β to 0.191 while R^2_α reduces to 0.013, which is too low to identify the parameter in Table 4.3(b).

Altogether, the analysis above suggests that the direct method leads to similar point estimates than the indirect method, provided that the latter is able to identify the coefficients. If the ratio of variances $\sigma^2_{\tilde{u}}/\hat{\sigma}^2_{\tilde{v}}$ is too small, however, the indirect method may fail to identify the wage response to a permanent shock, while the direct method continues to perform well. In any case, the direct method yields lower standard errors.

Heterogeneity. Table 4.3 uses all establishments from the private sector except for the financial industry, for which no sensible productivity measure is available. Wage elasticities may differ across industries due to different production processes, incentive structures, and industrial relations. Therefore, separate regressions are performed for four broad industry categories: manufacturing, construction, sales, and services. The results of the direct method are reported in Table 4.D.3. No clear statistical pattern arises. Concerning permanent shocks, wages in the construction sector seem to be somewhat more flexible, while wages in the service sector are more rigid. Yet, the standard error on both estimates is relatively large. This is because most workers in the sample are in manufacturing.

As demonstrated by Figure 4.1, the variance of transitory and permanent shocks differs with establishment size. This could affect the degree of wage insurance that establishments can provide. Since larger establishments have to cope with more extreme permanent shocks, they might provide less insurance. At the same time, larger establishments may have a higher financial buffer that allows them to provide more insurance. Table 4.D.4 presents wage elasticity estimates from separate regressions in each of the four size categories. Differences across categories appear small and statistically insignificant, such that the two highlighted effects seem to cancel out on average.

Finally, Table 4.D.5 mirrors the analysis of Guertzgen (2014) who highlights the role of industrial relations for wage insurance at the establishment level. A key finding of the paper is that the presence of a works council is associated with higher wage rigidity. To identify the effect of industrial relations, several interaction terms are added to (4.5). In particular, the two productivity shocks are interacted with a dummy for the presence of a works council (WC), a dummy indicating an industry-wide collective bargaining agreement (CBA industry), and a dummy indicating whether the firm itself has negotiated a CBA with a trade union (CBA firm). The results are reported in Table 4.D.5. Concerning permanent shocks, the estimated wage elasticity in uncovered establishments (without WC and CBA) and establishments with industry-wide CBA are similar to those of Table 8 in Guertzgen (2014). By contrast, my results cannot confirm the central finding of her study that the presence of a works council makes wages more rigid. This discrepancy remains if the estimation is performed separately for every



left axis: local linear kernel regression, 95% confidence band based on bootstrapped standard errors clustered at the establishment level; right axis: empirical cdf (shaded area)

Figure 4.2. Nonparametric wage regression

establishment size category, or if the indirect estimation method is applied (Table 4.D.6).[21]

The previous results are only based on male employees. If women are included as well, the wage elasticity with respect to permanent shocks reduces. Table 4.D.7 allows wage responses to differ by gender. To also account for differences in wage levels, the underlying wage regression is extended by gender dummies. Compared to men, wages of female employees adjust less to producitivity shocks irrespective of persistence. This also holds within industries. Disaggregated analysis suggests that the lower female wage elasticities are related to collective bargaining. Table 4.D.8 only considers employees in the manufacturing sector. While coverage by an industry level CBA does not affect wage elasticities of men, it reduces the wage elasticity of women with respect to permanent shocks down to zero. A possible explanation is that the compensation package of women less frequently contains bonus payments, as reported by Geddes and Heywood (2003) for the US. If the base wage is rigid, as in the case of an industry-wide collective bargaining agreement, total compensation can respond less to firm-specific economic conditions.

Nonlinear wage responses

The main advantage of the direct method is the possibility to estimate nonlinear relations between wage changes and productivity shocks. Figure 4.2 illustrates that the linearity assumption

[21] The coefficient estimates of Guertzgen (2014) even indicate that wages do not respond to permanent shocks at all if a works council is present. Although this finding seems to be statistically robust, the author herself admits that "full insurance against permanent shocks under works councils is clearly at variance with other studies" (p.366).

imposed in Section 4.4.2 might indeed be too restrictive. The transformed wage residuals $\Delta\tilde{\omega}_{ijt}$ defined in (4.6) are regressed on the productivity residuals $\hat{\Delta}\varepsilon_{jt}$ by nonparametric regression. This implicitly assumes that permanent and transitory shocks have the same effect on wages. While this is at odds with the evidence presented above, it serves as a useful benchmark. It allows to apply a standard univariate kernel smoother and does not require the Kalman smoothed residuals. The solid line in Figure 4.2 is obtained by local linear regression with an Epanechnikov kernel and the rule-of-thumb bandwidth estimator. The dashed lines indicate the bootstrapped 95% confidence band, accounting for clustering at the establishment level. To illustrate the support of $\Delta\varepsilon_{jt}$, the shaded area illustrates the distribution function of the productivity residuals (scale on the right axis).[22]

Apart from the shock realizations in the bottom tail of the distribution, the relation between productivity shocks and wage changes is better described by a concave function rather than a linear one. Better productivity realizations lead to higher wages, but at a decreasing rate. The concave relationship vanishes for shocks below the 10th percentile. In this region, Figure 4.2 actually suggests that the wage elasticity turns negative. The more detrimental the shock, the less it is passed on to wages. The complementary analysis of the individual layoff probability in Figure 4.5 reveals that the establishments facing a shock in the 10th percentile start to dismiss workers instead of cutting wages more severely. In both figures, however, the standard error gets very large at the tails, such that this observation should not be overemphasized.

To account for the nonlinearities uncovered in Figure 4.2, equation (4.5) is generalized to

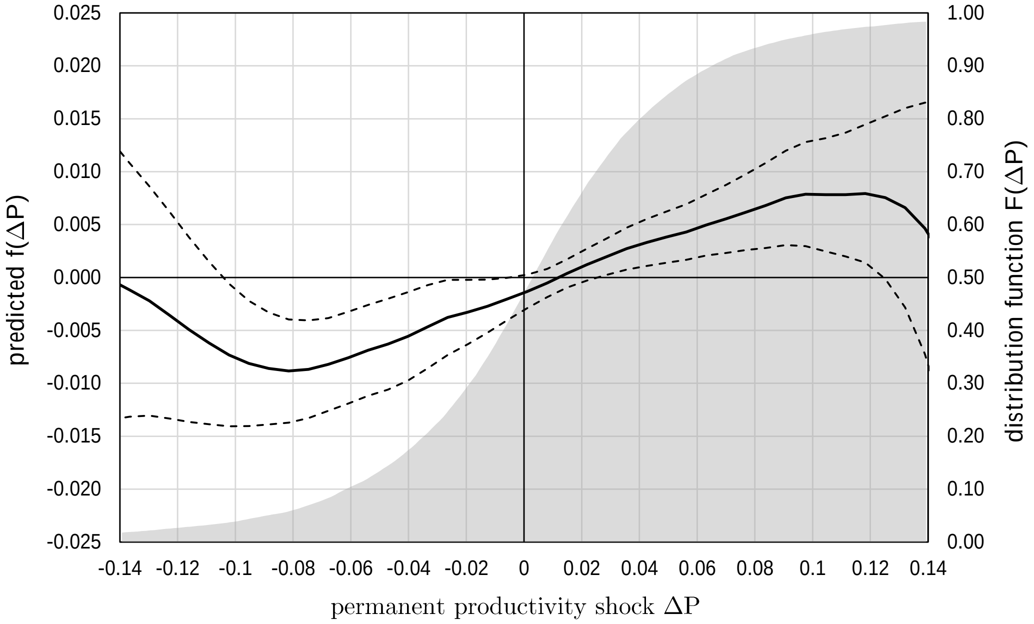
$$\Delta \ln w_{ijt} = \Delta X'_{ijt}\delta + f(\Delta P_{jt}) + g(\Delta T_{jt}) + \Delta\psi_{ijt}, \tag{4.8}$$

where f and g can be parametric or nonparametric functions. In estimation, the unknown productivity changes are replaced by the predictions $\hat{\Delta}P_{jt}$ and $\hat{\Delta}T_{jt}$ of the Kalman smoother. To find appropriate parametric forms of f and g, equation (4.8) is first estimated semiparametrically using an iterative backfitting procedure (Härdle et al., 2004, p.214-215). Starting with initial functional guesses $\hat{f}_0$ and $\hat{g}_0$, a first estimate of the parametric part, $\hat{\delta}_0$, is obtained by OLS. The partial residual $\Delta \ln w_{ijt} - \Delta X'_{ijt}\hat{\delta}_0 - \hat{f}_0(\hat{\Delta}P_{jt})$ is then regressed on $\hat{\Delta}T_{jt}$ by local linear kernel regression to form predictions $\hat{g}_1(\hat{\Delta}T_{jt})$. Likewise, the partial residual $\Delta \ln w_{ijt} - \Delta X'_{ijt}\hat{\delta}_0 - \hat{g}_0(\hat{\Delta}T_{jt})$ is nonlinearly regressed on $\hat{\Delta}P_{jt}$ to form predictions $\hat{g}_1(\hat{\Delta}P_{jt})$. The procedure continues until convergence. Note that the estimated function f and g are only identified up to an additive constant. In Figure 4.3 and Figure 4.4, the predicted functions are normalized to have zero mean.[23]

Figure 4.3 illustrates the predicted wage response with respect to a permanent productivity

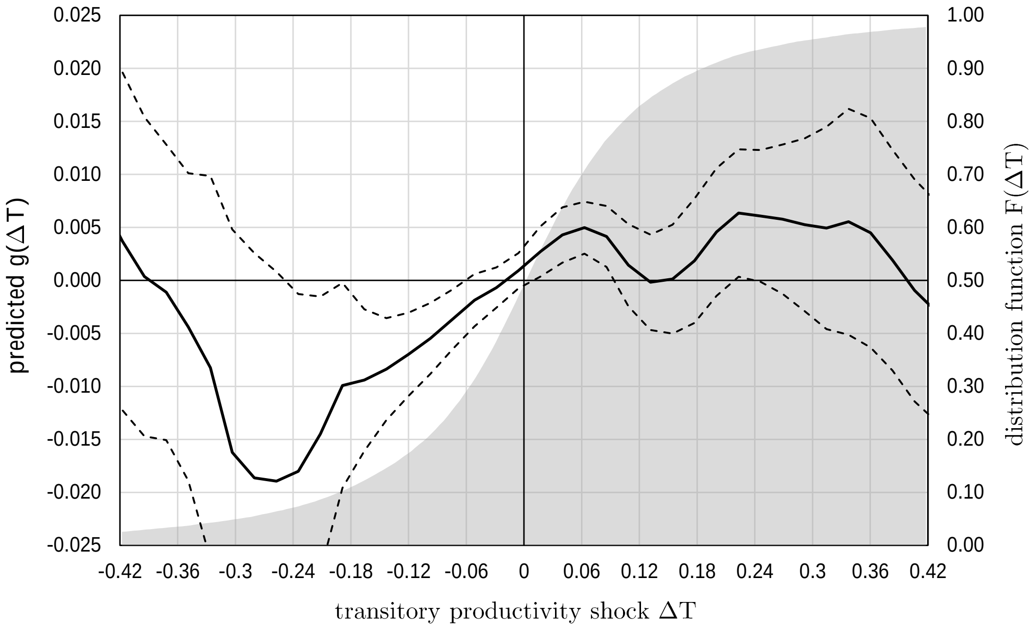
[22] The nonparametric estimation was done with Stata's `lpoly` command. Details on kernel and bandwidth selection can be found in the manual. The bandwidth was held constant for all bootstrap replications.

[23] Drawing on Figure 4.2, a cubic spline with breakpoints at the 10th, 50th, and 90th percentile of the distributions is used for the initial guess. Only one iteration of the backfitting algorithm is executed since no relevant changes in the predictions can be observed with more iterations. Using a linear starting function leads to visually identical predictions. Local linear kernel regressions use an Epanechnikov kernel with the rule-of-thumb bandwidth.



left axis: local linear kernel regression, 95% confidence band based on bootstrapped standard errors clustered at the establishment level; right axis: empirical cdf (shaded area)

Figure 4.3. Semiparametrically estimated wage response to permanent productivity shocks



left axis: local linear kernel regression, 95% confidence band based on bootstrapped standard errors clustered at the establishment level; right axis: empirical cdf (shaded area)

Figure 4.4. Semiparametrically estimated wage response to transitory productivity shocks

shock, $f(\Delta P_{jt})$. The wage response is approximately linear above the 10th percentile of the shock distribution. As the shock becomes more detrimental, wages reduce less and less, as already observed in Figure 4.2. Figure 4.4 illustrates the predicted wage response with respect to a transitory shock, $g(\Delta T_{jt})$. It is qualitatively very similar to Figure 4.2, which is due to the fact that most of the variation in productivity is transitory. Wages hardly respond to transitory positive shocks, while negative transitory shocks above the 10th percentile are linearly passed on to wages. Below the 10th percentile, establishments again seem reluctant to further cut wages, although the confidence band becomes extremely wide.

The above analysis yields important qualitative insights into the transmission of productivity shocks to wages. To compare to the literature, however, I also estimate local wage elasticities, which correspond to the slope of the wage response functions. To this purpose, I assume that f and g are piecewise linear functions. Furthermore, I choose a set of K intervals $\mathcal{I} = \{I_1, \ldots, I_K\}$ that form a partition of the real line, $\bigcup_{k=1}^{K} I_k = \mathbb{R}$. The function f and g are assumed to be linear on each of the intervals and continuous everywhere. They can hence be characterized by the following property,

$$
\begin{aligned}
f'(\Delta P_{jt}) &= \alpha_k \quad \text{if} \quad \Delta P_{jt} \in I_k, \\
g'(\Delta T_{jt}) &= \beta_k \quad \text{if} \quad \Delta T_{jt} \in I_k,
\end{aligned}
$$

where the coefficients α_k and β_k are the local wage elasticities. I consider three choices for the partition $\mathcal{I}$. The baseline is $\mathcal{I}_1 = \mathbb{R}$, which corresponds to a globally linear wage response and replicates the results from Section 4.4.2. The second partition $\mathcal{I}_2$ features one breakpoint at $\Delta P = \Delta T = 0$. This allows to estimate separate wage elasticities with respect to positive and negative productivity shocks. Finally, a third partition $\mathcal{I}_3$ uses the qualitative insights from Figures 4.3 and 4.4 to choose the breakpoints. I set them at the 10th, 50th, and 90th percentile of the respective shock distributions. The specific values are reported in Table 4.D.1.

Table 4.4 reports the coefficient estimates obtained from OLS estimation of (4.8) with piecewise linear functions f and g. The unobserved productivity shocks ΔP_{jt} and ΔT_{jt} are again replaced by the predictions of the Kalman smoother. Let me discuss the wage elasticities with respect to permanent shocks first. The first line in Table 4.4 assumes a globally linear relation between productivity and wage shocks, and therefore coincides with the estimate reported in the second group of columns in Table 4.3(a). Differentiating between positive and negative productivity shocks suggests severe downwards rigidity of real wages. While the wage elasticity with respect to a positive permanent shocks is 0.1065 and highly significant, the wage response to negative permanent shocks is insignificant and close to zero. Concluding that wages are completely rigid downwards, however, is erroneous. The last block of results in Table 4.4 reveals that the elasticity of 0 hides considerable heterogeneity. Above the 10th percentile of the shock distribution, wages are actually as elastic as in the positive domain. Whereas below the 10th percentile, the wage elasticity is weakly significantly negative. The wage elasticity estimates therefore corroborate the findings from the qualitative pattern in Figure 4.6. There

	interval I_k	permanent shock, ΔP_{jt} coefficient	std. err.	transitory shock, ΔT_{jt} coefficient	std. err.
$\mathcal{I}_1$	$\mathbb{R}$	0.0625***	0.0143	0.0189*	0.0102
$\mathcal{I}_2$	$(-\infty, 0)$	−0.0056	0.0263	0.0462***	0.0182
	$[0, +\infty)$	0.1121***	0.0270	−0.0067	0.0098
$\mathcal{I}_3$	$(-\infty, q_{10})$	−0.0638*	0.0354	0.0003	0.0181
	$[q_{10}, q_{50})$	0.1082**	0.0524	0.0821**	0.0325
	$[q_{50}, q_{90})$	0.1149**	0.0498	0.0043	0.0220
	$[q_{90}, +\infty)$	0.0641	0.0516	−0.0285*	0.0168

q_r refers to the rth percentile of the respective distribution; bootstrapped standard errors clustered at the establishment level, coefficient significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.4. Local wage elasticities for different partitions of $\mathbb{R}$

are no indications for downwards wage rigidity apart from the first decile of the distribution. If a very bad shock arrives, establishments keep wage cuts moderate and rather lay off workers (compare Table 4.6). In the middle of the distribution, between the 10th and 90th percentile, the estimated wage elasticity is around 0.11 and therefore almost twice as high as the estimate of 0.0625 obtained by assuming a globally linear wage response.

Let me now turn to the wage elasticities with respect to transitory shocks in Table 4.4. Not accounting for nonlinearities, the estimate is the same as in Table 4.3(a) and weakly significant. Distinguishing between positive and negative shocks reveals upwards wage rigidity. The elasticity estimate with respect to a positive transitory shock is insignificant and close to zero. Positive transitory shocks therefore purely increase firm rents. By contrast, the elasticity of 0.0459 estimated for negative shocks is strongly significant. Putting additional breakpoints confirms the upwards rigidity, while downward rigidity is observed in the lower tail of the distribution. For negative transitory shocks above the 10th percentile, the wage elasticity is 0.0821 and close to the estimate for permanent shocks of the same size. More detrimental shocks do not lead to further wage cuts.

Heterogeneity. Repeating the nonlinear analysis for particular industries or establishment size categories does not yield additional insights since standard errors become very high. Broadly, the findings of Table 4.4 seem to apply also on a more disaggregated level. It is feasible, however, to allow for worker heterogeneity. In particular, I interact the shocks with a dummy that indicates whether the worker is officially registered as a blue-collar or a white-collar worker. Firms may be reluctant to cut wages of white-collar workers because of agency and turnover considerations. First, their effort is more difficult to monitor such that a wage cut might result in shirking. Second, they are more expensive to replace if they shirk or decide to quit the firm since white-collar work requires more firm-specific human capital. Along these lines, white-collar workers are expected to be better insured against negative shocks than blue-collar workers. The estimation results are reported in Table 4.5. To reduce selection effects, the analysis is

	interaction × interval I_k	permanent shock, ΔP_{jt} coefficient	std. err.	transitory shock, ΔT_{jt} coefficient	std. err.
$\mathcal{I}_1$	blue-collar × $\mathbb{R}$	0.0613***	0.0173	0.0244*	0.0134
	white-collar × $\mathbb{R}$	0.0651***	0.0186	−0.0015	0.0088
$\mathcal{I}_2$	blue-collar × $(-\infty, 0)$	−0.0228	0.0330	0.0634**	0.0250
	blue-collar × $[0, +\infty)$	0.1198***	0.0331	−0.0129	0.0121
	white-collar × $(-\infty, 0)$	−0.0007	0.0257	0.0117	0.0131
	white-collar × $[0, +\infty)$	0.1224***	0.0277	−0.0139	0.0126
$\mathcal{I}_3$	blue-collar × (q_{10}, q_{50})	0.0920	0.0611	0.1117***	0.0429
	blue-collar × $[q_{50}, q_{90})$	0.1088**	0.0555	−0.0059	0.0285
	white-collar × (q_{10}, q_{50})	0.0453	0.0555	0.0136	0.0196
	white-collar × $[q_{50}, q_{90})$	0.1819***	0.0526	0.0054	0.0234

establishments in the manufacturing sector only; bootstrapped standard errors clustered at the establishment level, coefficient significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.5. Wage elasticity by worker type in the manufacturing sector

constrained to the manufacturing sector. If linear wage responses are estimated, wages of blue-collar and white-collar workers react identically to permanent shocks. Transitory shocks, by contrast, only affect wages of blue-collar workers. Accounting for nonlinearities reveals that all the downwards wage flexibility apparent in Table 4.4 stems from wages of blue-collar workers. Wages of white-collar workers, by contrast, do not react at all to negative shocks, irrespective of their persistence. That downwards wage rigidity is stronger for white-collar workers is in line with previous empirical evidence of Du Caju et al. (2007) for Belgium and Campbell (1997) for the US, for example.[24]

4.4.3 Layoff responses

Firms may adjust to negative shocks not only by lowering wages but also by dismissing workers. It is hard to statistically distinguish an employer-initiated layoff from an employee-initiated quit. Following Boockmann and Steffes (2010), a layoff is defined as a transition from employment to non-employment where (a) the non-employment spell lasts for at least 60 days and (b) the next employment spell is not with the same employer.[25]

The layoff regressions are based on the following linear probability model,[26]

$$lay_{ijt} = X'_{ijt}\delta + \alpha P_{jt} + \beta T_{jt} + \phi_{ij} + \psi_{ijt},$$

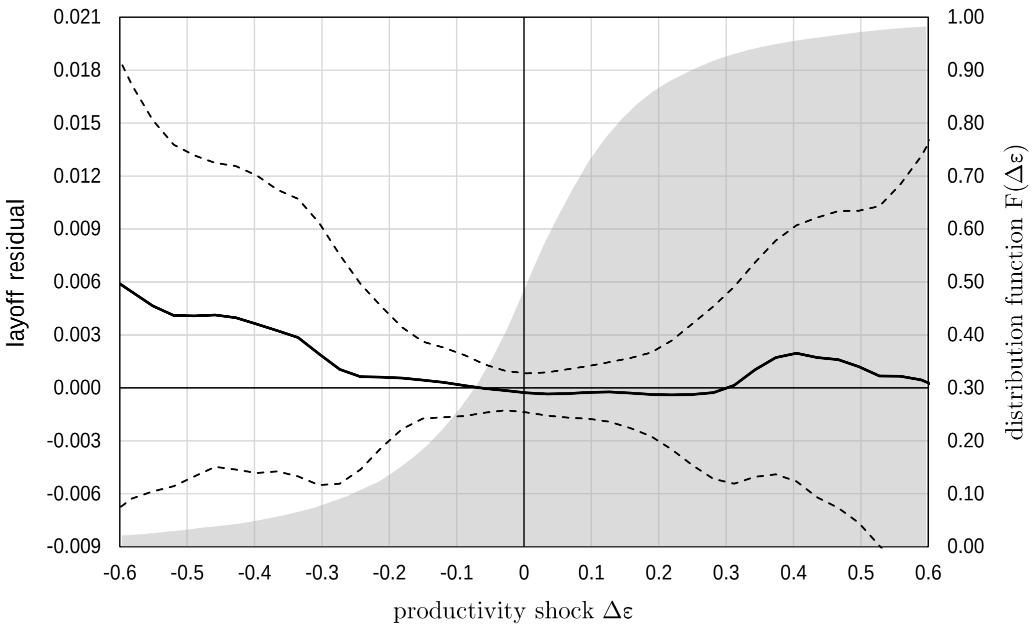
[24] Part of the observed downwards wage flexibility of blue-collar workers might also come from a more flexible hours margin. Nevertheless, especially the wage response with respect to transitory shocks is unlikely to be explained by changes in working time alone. The asymmetric response would imply that working hours fall after a negative transitory shock, but do not increase again after the shock vanishes.

[25] Changing the threshold to 30 days hardly affects the results.

[26] Alternative approaches such as a correlated random effects probit model (Wooldridge, 2002, p.616)

$$P(lay_{ijt} = 1 | X_{ijt}, P_{jt}, T_{jt}) = \Phi(X'_{ijt}\delta + \alpha P_{jt} + \beta T_{jt} + \overline{X}'_{ij}\xi_1 + \xi_2 \overline{P}_j + \xi_3 \overline{T}_j)$$

did not lead to credible results, possibly due to remaining unobserved heterogeneity.



left axis: local linear kernel regression, 95% confidence band based on bootstrapped standard errors clustered at the establishment level; right axis: empirical cdf (shaded area)

Figure 4.5. Nonparametric layoff regression

where the layoff dummy lay_{ijt} equals one if worker i is laid off by establishment j in year t and equals zero otherwise. This specification mirrors (4.4) and assumes a linear relationship between productivity shocks and layoff probabilities. First differencing sweeps out the fixed effect,

$$\Delta lay_{ijt} = \Delta X'_{ijt}\delta + \alpha \Delta P_{jt} + \beta \Delta T_{jt} + \Delta \psi_{ijt}. \tag{4.9}$$

Replacing ΔP_{jt} and ΔT_{jt} with the predictions of the Kalman smoother and estimation by OLS gives the first line of results in Table 4.6. Layoff probabilities do not seem to react to productivity shocks altogether.[27] Although Germany has stringent employment protection legislation that makes dismissals more complicated than in other countries, it is unlikely that troubled firms do not use this margin at all. Yet, the legal framework may lead firms to fire workers only if there is no other way to remain profitable. This is most likely to be the case after a negative permanent productivity shock. Therefore differentiating both between the shock persistence and the size of the shock might be crucial to obtain sensible estimates.

I explore first the role of nonlinearities in the relation between productivity shocks and variations in the layoff probability by fitting a nonparametric regression. Figure 4.5 is generated in the same way as Figure 4.2. The estimated nonparametric function suggests that only productivity shocks in the first decile of the shock distribution have an effect on layoffs, but there is considerable uncertainty involved in this statement. Furthermore, the estimated increase in the

[27] Estimating (4.9) with the indirect method fails to identify α and β. The Hansen J test rejects validity of the instrument sets at the 1% level. This is most likely due to the binary nature of the dependent variable.

		permanent shock, ΔP_{jt}		transitory shock, ΔT_{jt}	
	interval I_k	coefficient	std. err.	coefficient	std. err.
$\mathcal{I}_1$	$\mathbb{R}$	-0.0276	0.0213	0.0021	0.0095
$\mathcal{I}_2$	$(-\infty, 0)$	-0.0986^{**}	0.0462	0.0048	0.0237
	$[0, +\infty)$	0.0257	0.0291	-0.0001	0.0175
$\mathcal{I}_3$	$(-\infty, q_{10})$	-0.1165	0.0696	-0.0245	0.0388
	$[q_{10}, q_{50})$	-0.0980	0.0748	0.0179	0.0337
	$[q_{50}, q_{90})$	0.0580	0.0574	-0.0115	0.0261
	$[q_{90}, +\infty)$	-0.0121	0.0681	0.0184	0.0404

q_r refers to the rth percentile of the respective distribution; bootstrapped standard errors clustered at the establishment level, coefficient significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.6. Local semi-elasticity of the layoff probability for different partitions of $\mathbb{R}$

layoff probability for extremely bad draws is quantitatively small relative to the annual layoff rate of 6.9 percent.

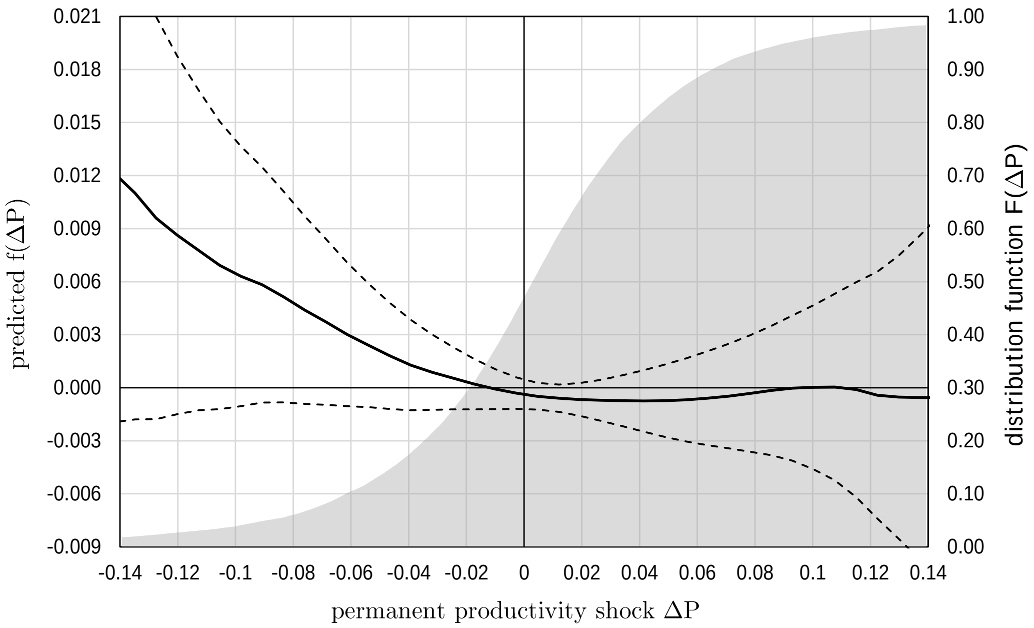
In a next step I distinguish again between permanent and transitory shocks,

$$\Delta lay_{ijt} = \Delta X'_{ijt}\delta + f(\Delta P_{jt}) + g(\Delta T_{jt}) + \Delta\psi_{ijt},$$

and perform a semiparametric regression. The explanatory variables are the same as in the wage regressions, and ΔP_{jt} and ΔT_{jt} are replaced by the predictions of the Kalman filter. Estimation uses the backfitting approach as explained in Section 4.4.2. The predictions for f and g are depicted in Figure 4.6 and Figure 4.7, respectively. Note that both functions are identified up to an additive constant. In the figures, their sample mean has been normalized to zero.

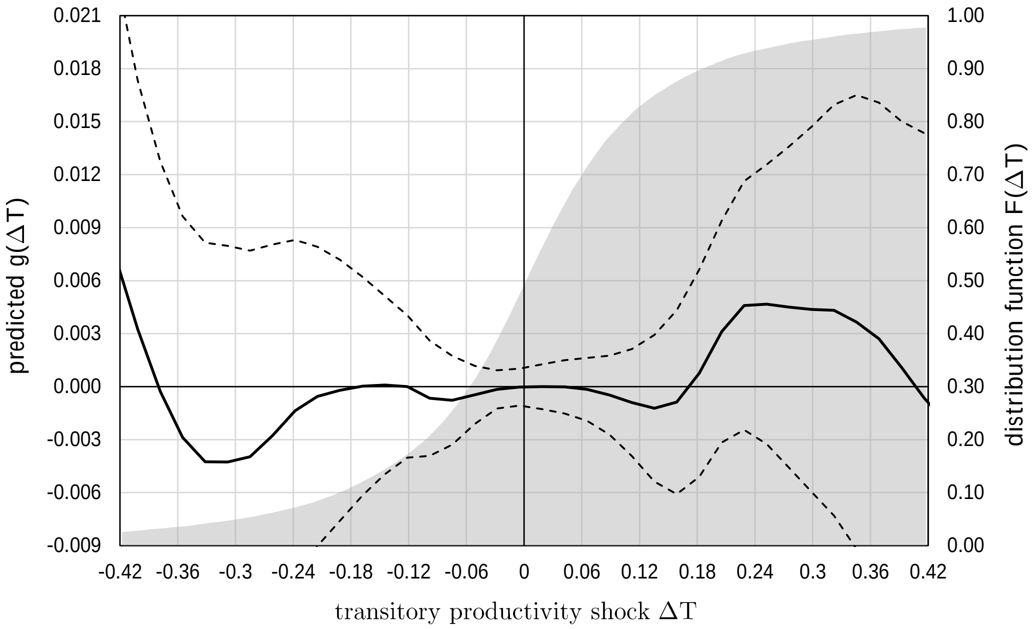
Figure 4.6 depicts the estimated layoff response to a permanent productivity shock. While positive shocks do not affect the layoff probability, in the negative domain the layoff probability increases with the severity of the shock. Confidence bands are still wide, but the increasing use of layoffs in response to bad events appears to be at least weakly statistically significant. By contrast, Figure 4.7 reveals that layoff responses to transitory shocks do not follow any specific pattern. In particular, there is no evidence for an increase in layoff rates after a negative transitory shock. Regarding the high employment protection legislation in Germany, the findings of Figures 4.6 and 4.7 are little surprising. In fact, they corroborate that the Kalman smoother is able to distinguish between the permanent and transitory shock component with reasonable accuracy.

Finally, I calculate local semi-elasticities of the layoff probability using linear piecewise specifications for f and g. I use the same partitions as in Section 4.4.2, and the estimated coefficients are reported in Table 4.6. Differentiating positive and negative shocks confirms the graphical insight that the layoff probability only reacts to negative permanent shocks. The estimated semi-elasticity is -0.0986 and significant at the 5% level. Considering that the average annual layoff rate in the sample is 6.9 percent, this implies an elasticity of $\frac{\Delta lay_{ijt} \,/\, \overline{lay}_{ijt}}{\Delta P_{ijt}} = -\frac{0.0986}{0.0687} = 1.44$ with respect to negative permanent shocks. Adding further breakpoints suggests that the layoff



left axis: local linear kernel regression, 95% confidence band based on bootstrapped standard errors clustered at the establishment level; right axis: empirical cdf (shaded area)

Figure 4.6. Nonparametrically estimated layoff response to permanent productivity shocks



left axis: local linear kernel regression, 95% confidence band based on bootstrapped standard errors clustered at the establishment level; right axis: empirical cdf (shaded area)

Figure 4.7. Nonparametrically estimated layoff response to transitory productivity shocks

	interaction × interval I_k	permanent shock, ΔP_{jt} coefficient	std. err.	transitory shock, ΔT_{jt} coefficient	std. err.
$\mathcal{I}_1$	blue-collar × $\mathbb{R}$	−0.0266	0.0229	−0.0029	0.0118
	white-collar × $\mathbb{R}$	0.0210	0.0229	−0.0003	0.0115
$\mathcal{I}_2$	blue-collar × $(-\infty, 0)$	−0.1015*	0.0528	0.0135	0.0278
	blue-collar × $[0, +\infty)$	0.0246	0.0321	−0.0184	0.0204
	white-collar × $(-\infty, 0)$	−0.0087	0.0549	0.0205	0.0299
	white-collar × $[0, +\infty)$	0.0452	0.0377	−0.0237	0.0208

establishments in the manufacturing sector only; bootstrapped standard errors clustered at the establishment level, coefficient significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.7. Semi-elasticity of the layoff probability by worker type in the manufacturing sector

elasticity above and below the 10th percentile of the shock distribution is nearly identical, such that conditioning on the sign of the shock is enough to adequately capture the nonlinearity of layoff responses.

Heterogeneity. Exploring heterogeneity in layoff responses across sector and establishment size categories does not yield any robust insights since standard errors get very large. It is feasible, however, to distinguish been white-collar and blue-collar employment as in Table 4.5. The observation there was that virtually all of the downwards flexibility in wages comes is explained by blue-collar workers, while wages of white-collar worker are unaffected by negative shocks. Does this downwards wage rigidity imply that white-collar workers more often lose their job after a bad productivity shock? Table 4.7 suggests the contrary. In fact, white-collar workers are perfectly insured against negative productivity shocks. The increase in layoff probabilities after a negative permanent shock apparent from Table 4.6 is limited to blue-collar workers. Firms may be reluctant to fire white-collar workers as they anticipate higher hiring and training costs compared to blue-collar workers once the economic situation improves. Additionally, white-collar workers may be more complementary to other production factors such as capital, while blue-collar workers are easier to substitute in the production process.

4.5 Conclusion

This essay explores how shock persistence and shock size affect the degree of wage and employment insurance that firms provide against idiosyncratic productivity shocks. The joint analysis of size and persistence requires a new econometric approach. I suggest a two step procedure. First, the stochastic properties of the productivity shock process are determined along the lines of Guiso et al. (2005). Second, a Kalman smoother is applied at the firm level to predict the permanent and transitory component of productivity shocks. These predicted time series can be included as explanatory variables in wage or layoff regressions. This allows to estimate arbitrary functional dependencies between productivity shocks and wage changes, whereas the original method of Guiso et al. (2005) is confined to linear dependencies.

Using rich matched employer-employee data from Germany, I find that both shock persistence and shock size matters. The wage elasticity with respect to a permanent productivity shock is constant between the 10th and 90th percentile of the shock distribution and becomes smaller at the tails. In response to extremely bad permanent shocks, firms seem to refrain from wage cuts altogether and adjust via layoffs, perhaps in an effort to reduce the quitting incentive for the remaining workers. Transitory productivity shocks lead to asymmetric wage responses. While negative shocks tend to reduce wages, positive transitory shocks are fully captured by the firm. Individual layoff probabilities do not respond to transitory productivity shocks.

The general patterns hide substantial heterogeneity at the worker level. The data suggests that wage cuts and employment loss after negative shocks are concentrated on blue-collar workers. Whereas white-collar workers enjoy full insurance against negative productivity shocks, irrespective of their size and persistence. That the adjustment to negative shocks goes primarily at the expense of blue-collar workers points to agency and turnover considerations of the employers: First, the effort of blue-collar workers may be easier to monitor which allows more downward wage flexibility without spurring unrighteous behavior. Second, blue-collar employment is usually less human capital intensive and therefore cheaper to replace due to lower hiring and training costs. Additionally, blue-collar employment may be easier to substitute by other production factors such as capital.

4.A Sample statistics

	establishment level						worker level			
	full sample		wage sample		layoff sample		wage sample		layoff sample	
	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.
sales per worker*	1.811	6.892	1.863	6.510	1.857	6.419	2.670	4.165	2.733	5.099
employment	181.3	772.1	204.0	828.3	198.7	813.9	2758.7	4814.4	3288.1	5477.5
capital-labor ratio*	0.947	5.913	0.950	2.751	0.952	2.808	1.409	2.172	1.426	2.182
works council	0.338		0.374		0.366		0.890		0.893	
industry CBA	0.464		0.483		0.479		0.795		0.804	
firm CBA	0.086		0.088		0.087		0.085		0.080	
1–9 employees	0.216		0.159		0.168		0.006		0.005	
10–99 employees	0.406		0.414		0.415		0.052		0.050	
100–199 employees	0.222		0.249		0.244		0.132		0.127	
200+ employees	0.156		0.177		0.173		0.810		0.818	
manufacturing	0.477		0.511		0.504		0.840		0.831	
construction	0.143		0.157		0.158		0.049		0.047	
sales	0.160		0.154		0.154		0.041		0.039	
services	0.220		0.178		0.185		0.070		0.083	
wage			85.23	25.35	87.13	29.13	107.23	27.28	116.29	39.09
tenure			9.312	4.407	8.506	4.398	12.234	7.393	11.578	7.823
age			41.98	4.974	41.93	4.771	41.459	8.702	41.817	8.762
white-collar			0.280		0.310		0.180		0.311	
no degree			0.090		0.086		0.163		0.136	
vocational degree			0.802		0.778		0.768		0.694	
high school			0.004		0.005		0.003		0.005	
voc. + high school			0.030		0.032		0.019		0.026	
applied university			0.034		0.042		0.027		0.073	
university			0.040		0.056		0.019		0.066	
establishments	2697		2531		2620					
individuals							216709		300667	

* measured in 100000 €

Table 4.A.1. Descriptive sample statistics

4.B Microfoundation of the productivity regression

Equation (4.1) can be motivated by a Cobb-Douglas production function $Y_{jt} = A_{jt}K_{jt}^{\alpha}L_{jt}^{\beta}$ at the establishment level. Diving by L_{jt} and taking logarithms yields

$$\ln y_{jt} = \alpha \ln k_{jt} + \delta \ln L_{jt} + Z'_{jt}\gamma + \varphi_j + \varepsilon_{jt}, \tag{4.B.1}$$

where $\delta := \alpha + \beta - 1$ and $\ln A_{jt} = Z'_{jt}\gamma + \varphi_j + \varepsilon_{jt}$. Estimating (4.B.1) in first differences generates an endogeneity problem because the change in employment, $\Delta \ln L_{jt}$, is correlated with $\Delta\varepsilon_{jt}$. In the literature, this is commonly resolved by using appropriate lags of the variable

	(a) static FE model		(b) dynamic FE model	
	coefficient	std. err.	coefficient	std. err.
$\ln y_{jt-1}$	——	——	0.2503***	0.0378
$\ln k_{jt}$	0.3205***	0.0289	0.3021***	0.0233
$\ln L_{jt}$	0.0234	0.0380	−0.0206	0.0318
	statistic	p-value	statistic	p-value
AR(2) test	−2.77	0.006	1.81	0.070
AR(3) test	−1.55	0.120	−0.69	0.493
AR(4) test	0.72	0.471	1.10	0.270
Hansen J test	44.70	0.211	80.66	0.252

two-step difference GMM, corrected standard errors clustered at the establishment level, significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.B.1. Productivity regressions including employment

in levels ($\ln L_{jt}$) as instruments for the difference $\Delta \ln L_{jt}$. This yields a set of moment conditions $\mathbb{E}[L_{jt-s}\Delta\varepsilon_{jt}] = 0$ that (together with the ones belonging to the exogenous regressors ΔZ_{jt}) are used to apply GMM estimation. However, difference-in-Hansen tests reveal that any lag of $\ln L_{jt}$ is itself correlated with $\Delta\varepsilon_{jt}$ and not a valid instrument.

An alternative option is to adopt ideas of Blundell and Bond (1998) and complement the first differenced equation by a level equation,

$$\Delta \ln y_{jt} = \alpha\Delta \ln k_{jt} + \Delta Z'_{jt}\gamma + \Delta\varepsilon_{jt}, \tag{4.B.2}$$

$$\ln y_{jt} = \delta \ln L_{jt} + Z'_{jt}\gamma + \varphi_j + \varepsilon_{jt},, \tag{4.B.3}$$

The difference equation (4.B.2) excludes the endogenous employment change, while the level equation (4.B.3) contains the employment variable and the strictly exogenous regressors Z_{jt}. The capital-labor ratio has to be excluded from the level equation since it is likely to be correlated with ϕ_j. Identification of α therefore only uses moments from the difference equation, where $\Delta \ln k_{jt}$ can be regarded as exogenous, as indicated by a series of difference-in-Hansen tests with different instrument choices. Identification of δ is still an open issue, since $\ln L_{jt}$ is likely to be correlated with the joint error term $\varphi_j + \varepsilon_{jt}$. Blundell and Bond (1998) suggest to instrument the variable in levels with its lagged first differences and to use moment conditions of the form $\mathbb{E}[(\varphi_j + \varepsilon_{jt})\Delta \ln L_{i,t-s}] = 0$. By (4.2), the time-varying part of the error term ε_{jt} accumulates permanent productivity shocks, $\varepsilon_{jt} = \zeta_{j1} + \sum_{s=1}^{t} \tilde{u}_{jt} + \tilde{v}_{jt}$. Since period t employment is likely to depend on the permanent innovation $\tilde{u}_{jt}$, lagged employment changes are unlikely to be orthogonal to ε_{jt}. However, future employment changes might be. Provided that ΔL_{jt} is uncorrelated with ϕ_j and future errors ε_{js} ($s > t+1$), the moment conditions $\mathbb{E}[(\varphi_j + \varepsilon_{jt})\Delta \ln L_{i,t+k}] = 0$ with $k > 1$ can be used to identify δ. In the estimation I use the 2nd, 3rd and 4th *leads* of $\Delta \ln L_{jt}$ as instruments for $\ln L_{jt}$.

The results are reported in Table 4.B.1(a). The estimate on the capital-labor ratio is close to the baseline (Table 4.1), while the coefficient on employment is insignificant and close to

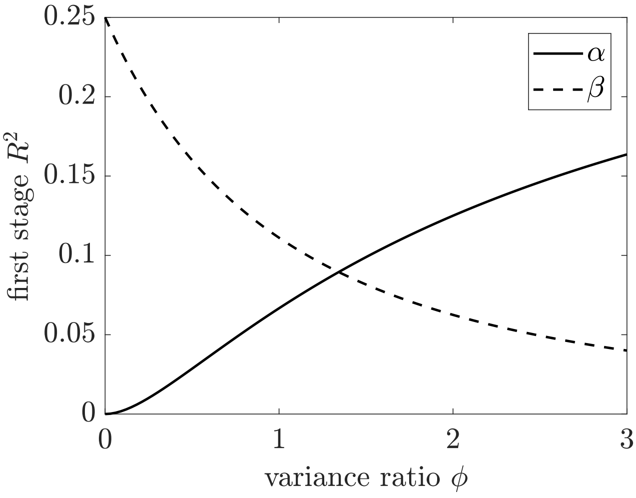


Figure 4.C.1. First stage R^2 for α and β as a function of the variance ratio $\phi = \sigma^2_{\tilde{u}}/\sigma^2_{\tilde{v}}$

zero. The Hansen test does not reject the validity of the overidentfying moment restrictions. These observations remain valid if the lagged dependent variable is included at the right-hand side of (4.B.1), see Table 4.B.1(b). Constant returns to scale at the establishment level, $\delta = 0$, cannot be rejected by a t-test. Since the point estimates of δ are also not significant in economic terms, I impose constant returns to scale form the outset, which boils down to the regression equation (4.1).

4.C When the indirect method fails

A common measure to detect weakness of an instrument is the first stage F statistic, which is a function of the R^2 statistic. In the linear regression model $x = \pi z + \xi$ the population coefficient of determination, R^2, equals the square of $\rho = \mathbb{E}[xz]/\sqrt{\mathbb{E}x^2\mathbb{E}z^2}$. Provided that $\mathbb{E}x = \mathbb{E}z = 0$, ρ is simply the population-equivalent of the correlation coefficient between x and z.

The first stage regression that is fit to identify β is $\Delta\varepsilon_{jt} = \pi\Delta\varepsilon_{jt+1} + \xi_{jt}$. Note that $\mathbb{E}\Delta\varepsilon^2_{jt} = \mathbb{E}\Delta\varepsilon^2_{jt+1} = \sigma^2_{\tilde{u}} + 2\sigma^2_{\tilde{v}}$ and $\mathbb{E}[\Delta\varepsilon_{jt}\Delta\varepsilon_{jt+1}] = -\sigma^2_{\tilde{v}}$. Therefore,

$$R^2_\beta = \frac{(\sigma^2_{\tilde{v}})^2}{(\sigma^2_{\tilde{u}} + 2\sigma^2_{\tilde{v}})^2} = \frac{1}{(2+\phi)^2}$$

where $\phi := \sigma^2_{\tilde{u}}/\sigma^2_{\tilde{v}}$. The first stage regression that is fit to identify α is $\Delta\varepsilon_{jt} = \pi\sum_{k=-1}^{1}\Delta\varepsilon_{jk+1} + \xi_{jt}$. Note that $\mathbb{E}[(\sum_{k=-1}^{1}\Delta\varepsilon_{jk+1})^2] = 3\sigma^2_{\tilde{u}} + 2\sigma^2_{\tilde{v}}$ and $\mathbb{E}[\Delta\varepsilon_{jt}\sum_{k=-1}^{1}\Delta\varepsilon_{jk+1}] = \sigma^2_{\tilde{u}}$. Therefore,

$$R^2_\alpha = \frac{(\sigma^2_{\tilde{u}})^2}{(\sigma^2_{\tilde{u}} + 2\sigma^2_{\tilde{v}})(3\sigma^2_{\tilde{u}} + 2\sigma^2_{\tilde{v}})} = \frac{\phi^2}{(2+\phi)(2+3\phi)}.$$

Figure 4.C.1 shows that R^2_α is increasing in ϕ while R^2_β is decreasing in ϕ. This suggests a trade-off between the accuracy of the estimated α and the accuracy of estimated β.

4.D Additional tables

Variance and distribution of productivity shocks

percentile	total shock, $\Delta\hat{\varepsilon}_{jt}$	permanent shock, $\Delta\hat{P}_{jt}$	transitory shock, $\Delta\hat{T}_{jt}$
5%	−0.3970	−0.0902	−0.3000
10%	−0.2568	−0.0621	−0.1925
25%	−0.1028	−0.0260	−0.0763
50%	0.0053	0.0030	0.0024
75%	0.1086	0.0316	0.0792
90%	0.2504	0.0653	0.1862
95%	0.3797	0.0911	0.2836

Table 4.D.1. Percentiles of the productivity shock distributions

Robustness of wage elasticity estimates

	(a) homoscedastic		(b) heteroscedastic: establishment size + industry	
	coefficient	std. err.	coefficient	std. err.
α	0.0701***	0.0162	0.0626***	0.0170
β	0.0201**	0.0091	0.0192*	0.0101

bootstrapped standard errors clustered at the establishment level, coefficient significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.D.2. Wage elasticity estimates for different variance structures

Heterogeneity by firm characteristics

	permanent shock, ΔP_{jt}		transitory shock, ΔT_{jt}	
industry	coefficient	std. err.	coefficient	std. err.
manufacturing	0.0615***	0.0162	0.0204*	0.0121
construction	0.0950***	0.0313	0.0113	0.0136
sales	0.0599**	0.0236	0.0015	0.0116
services	0.0228	0.0344	0.0259	0.0246
total	0.0625***	0.0143	0.0189*	0.0102

separate wage regressions by industry; bootstrapped standard clustered at the establishment level, coefficient significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.D.3. Wage elasticity by industry

	permanent shock, ΔP_{jt}		transitory shock, ΔT_{jt}	
size category	coefficient	std. err.	coefficient	std. err.
1–9 employees	0.0545	0.0407	0.0069	0.0097
10–99 employees	0.0681***	0.0177	0.0148***	0.0057
100–199 employees	0.0593***	0.0155	0.0110	0.0087
200+ employees	0.0588***	0.0178	0.0231	0.0146
total	0.0625***	0.0143	0.0189*	0.0102

separate wage regressions by size category; bootstrapped standard clustered at the establishment level, coefficient significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.D.4. Wage elasticity by establishment size

	permanent shock, ΔP_{jt}		transitory shock, ΔT_{jt}	
	coefficient	std. err.	coefficient	std. err.
ΔX_{jt}	0.0708*	0.0382	0.0179*	0.0105
ΔX_{jt} × CBA industry	−0.0142	0.0354	0.0035	0.0188
ΔX_{jt} × CBA firm	−0.0936	0.0801	0.0073	0.0291
ΔX_{jt} × WC	0.0104	0.0386	−0.0004	0.0209

establishments in the manufacturing sector only; bootstrapped standard clustered at the establishment level, coefficient significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.D.5. Wage elasticity by industrial relations (direct method)

	permanent shock, ΔP_{jt}		transitory shock, ΔT_{jt}	
	coefficient	std. err.	coefficient	std. err.
ΔX_{jt}	0.0741*	0.0397	0.0101	0.0188
ΔX_{jt} × CBA industry	−0.0257	0.0392	0.0133	0.0457
ΔX_{jt} × CBA firm	−0.0923	0.0879	0.0201	0.0489
ΔX_{jt} × WC	−0.0008	0.0458	−0.0121	0.0397
K-P Wald F stat.	1.688		4.620	
Hansen J stat. (p val.)	15.41 (0.212)		13.78 (0.315)	

establishments in the manufacturing sector only; bootstrapped standard clustered at the establishment level, coefficient significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.D.6. Wage elasticity by industrial relations (indirect method)

Heterogeneity by gender

	permanent shock, ΔP_{jt}		transitory shock, ΔT_{jt}	
	coefficient	std. err.	coefficient	std. err.
ΔX_{jt} × male	0.0628***	0.0143	0.0188*	0.0101
ΔX_{jt} × female	0.0443***	0.0170	0.0025	0.0072

bootstrapped standard clustered at the establishment level, coefficient significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.D.7. Wage elasticity by gender

	permanent shock, ΔP_{jt}		transitory shock, ΔT_{jt}	
	coefficient	std. err.	coefficient	std. err.
ΔX_{jt} × male	0.0722*	0.0376	0.0169	0.0106
ΔX_{jt} × female	0.0741**	0.0376	−0.0014	0.0146
ΔX_{jt} × CBA industry × male	−0.0132	0.0348	0.0041	0.0188
ΔX_{jt} × CBA industry × female	−0.0712*	0.0413	0.0244	0.0189
ΔX_{jt} × WC × male	0.0079	0.0381	0.0000	0.0206
ΔX_{jt} × WC × female	0.0371	0.0453	−0.0171	0.0199

establishments in the manufacturing sector only; bootstrapped standard clustered at the establishment level, coefficient significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.D.8. Wage elasticity by gender and industrial relations

Bibliography

D. Acemoglu and R. Shimer. Efficient Unemployment Insurance. *Journal of Political Economy*, 107(5):893–928, 1999.

J. T. Addison, T. Schank, C. Schnabel, and J. Wagner. Works Councils in the Production Process. *Schmollers Jahrbuch*, 126(812):251–283, 2006.

G. A. Akerlof. Labor Contracts as Partial Gift Exchange. *The Quarterly Journal of Economics*, 97(4):543–569, 1982.

G. A. Akerlof and J. L. Yellen. The Fair Wage-Effort Hypothesis and Unemployment. *The Quarterly Journal of Economics*, 105(2):255–283, 1990.

A. Albanese and B. Cockx. Permanent wage cost subsidies for older workers. An effective tool for employment retention and postponing early retirement? *Labour Economics*, 2018. doi: 10.1016/j.labeco.2018.01.005. In press.

H. Alda, S. Bender, and H. Gartner. The Linked Employer-Employee Dataset created from the IAB-Establishment Panel and the process-produced data of the IAB (LIAB). *Schmollers Jahrbuch*, 125:327–336, 2005.

F. Alvarez and M. Veracierto. Severance payments in an economy with frictions. *Journal of Monetary Economics*, 47(3):477–498, 2001.

M. Arellano and S. Bond. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies*, 58(2):277–297, 1991.

C. Azariadis. Implicit Contracts and Underemployment Equilibria. *Journal of Political Economy*, 83(6):1183–1202, 1975.

J. Babecký, P. Du Caju, T. Kosma, M. Lawless, J. Messina, and T. Rõõm. Downward Nominal and Real Wage Rigidity: Survey Evidence from European Firms. *Scandinavian Journal of Economics*, 112(4):884–910, 2010.

M. N. Baily. Wages and Employment under Uncertain Demand. *The Review of Economic Studies*, 41(1):37–50, 1974.

A. Balleer, B. Gehrke, W. Lechthaler, and C. Merkl. Does short-time work save jobs? A business cycle analysis. *European Economic Review*, 84:99–122, 2016.

C. F. Baum, M. E. Schaffer, and S. Stillman. Enhanced routines for instrumental variables/generalized method of moments estimation and testing. *Stata Journal*, 7(4):465–506, 2007.

L. Behaghel, B. Crépon, and B. Sédillot. The perverse effects of partial employment protection reform: The case of French older workers. *Journal of Public Economics*, 92(3-4):696–721, 2008.

S. Bentolila and G. Bertola. Firing Costs and Labour Demand: How Bad is Eurosclerosis? *The Review of Economic Studies*, 57(3):381–402, 1990.

M. J. Bils. Real Wages over the Business Cycle: Evidence from Panel Data. *Journal of Political Economy*, 93(4):666–689, 1985.

O. J. Blanchard and A. Landier. The perverse effects of partial labour market reform: fixed-term contracts in France. *The Economic Journal*, 112(480):F214–F244, 2002.

O. J. Blanchard and J. Tirole. The Joint Design of Unemployment Insurance and Employment Protection: A First Pass. *Journal of the European Economic Association*, 6(1):45–77, 2008.

R. Blundell and S. Bond. Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1):115–143, 1998.

T. Boeri and P. Garibaldi. Two Tier Reforms of Employment Protection: a Honeymoon Effect? *The Economic Journal*, 117(521):F357–F385, 2007.

T. Boeri, P. Garibaldi, and E. R. Moen. Inside severance pay. *Journal of Public Economics*, 145:211–225, 2017.

B. Boockmann. The effects of wage subsidies for older workers. *IZA World of Labor*, 189, September 2015.

B. Boockmann and S. Steffes. Workers, firms, or institutions: What determines job duration for male employees in Germany? *Industrial and Labor Relations Review*, 64(1):109–127, 2010.

A. Börsch-Supan and M. Weiss. Productivity and age: Evidence from work teams at the assembly line. *Journal of the Economics of Ageing*, 7:30–42, 2016.

T. Brändle and W. D. Heinbach. Opening clauses in collective bargaining agreements: More flexibility to save jobs? *Review of Economics*, 64(2):159–192, 2013.

K. Brenke, U. Rinne, and K. F. Zimmermann. Short-time work: The German answer to the Great Recession. *International Labour Review*, 152(2):287–305, 2013.

S. G. Bronars and M. Famulari. Shareholder Wealth and Wages: Evidence for White-Collar Workers. *Journal of Political Economy*, 109(2):328–354, 2001.

C. M. Campbell, III. The Variation in Wage Rigidity by Occupation and Union Status in the US. *Oxford Bulletin of Economics and Statistics*, 59(1):133–147, 1997.

A. R. Cardoso and M. Portela. Micro foundations for wage flexibility: Wage insurance at the firm level. *Scandinavian Journal of Economics*, 111(1):29–50, 2009.

A. Chéron, J.-O. Hairault, and F. Langot. Age-dependent employment protection. *The Economic Journal*, 121(557):1477–1504, 2011.

A. Chéron, J.-O. Hairault, and F. Langot. Life-cycle equilibrium unemployment. *Journal of Labor Economics*, 31(4):843–882, 2013.

J. S. Costain and M. Reiter. Business cycles, unemployment insurance, and the calibration of matching models. *Journal of Economic Dynamics and Control*, 32(4):1120–1155, 2008.

M. Cozzi and G. Fella. Job displacement risk and severance pay. *Journal of Monetary Economics*, 84:166–181, 2016.

C. Dauth and O.-S. Toomet. On government-subsidized training programs for older workers. *Labour*, 30:371–392, 2016.

P. J. Devereux. The cyclicality of real wages within employer-employee matches. *Industrial and Labor Relations Review*, 54(4):835–850, 2001.

W. T. Dickens and L. Goette. Estimating wage rigidity for the international wage flexibility project. mimeo, Brookings Institution, 2006.

W. T. Dickens, L. Goette, E. L. Groshen, S. Holden, J. Messina, M. E. Schweitzer, J. Turunen, and M. E. Ward. How Wages Change: Micro Evidence from the International Wage Flexibility Project. *Journal of Economic Perspectives*, 21(2):195–214, 2007.

L. Diéz-Catalán and E. Villanueva. Contract staggering and unemployment during the great recession: evidence from Spain. Banco de España Working Paper No. 1431, 2015.

D. Dorn and A. Sousa-Poza. 'Voluntary' and 'involuntary' early retirement: an international analysis. *Applied Economics*, 42(4):427–438, 2010.

P. Du Caju, C. Fuss, and L. Wintr. Downward Wage Rigidity for Different Workers and Firms: an evaluation for Belgium using the IWFP procedure. ECB Working Paper No. 840, 2007.

P. Du Caju, T. Kosma, M. Lawless, J. Messina, and T. Rõõm. Why Firms Avoid Cutting Wages: Survey Evidence from European Firms. *ILR Review*, 68(4):862–888, 2015.

C. Dustmann and U. Schönberg. Training and union wages. *The Review of Economics and Statistics*, 91(2):363–376, 2009.

Eurostat. Transition from work to retirement – evaluation of the 2012 labour force survey ad hoc module. `http://ec.europa.eu/eurostat/documents/1978984/6037334/Evaluation-Report-AHM-2012.pdf`, 2012.

Eurostat. Job vacancy statistics by occupation, NUTS 2 regions and NACE Rev. 1.1 activity - annual data (2000-2008), 2018. URL `http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=jvs_a_nace1`. Date accessed: Feb. 11, 2018.

G. Fella. Optimal Severance Pay in a Matching Model. Working Paper No. 598, Queen Mary University of London, School of Economics and Finance, 2007.

G. Fella. Matching, wage rigidities and efficient severance pay. *B.E. Journal of Economic Analysis and Policy*, 12(1), 2012.

G. Fella and C. J. Tyson. Privately optimal severance pay. *The B.E. Journal of Macroeconomics*, 13(1):415–453, 2013.

W. H. Fisher and C. Keuschnigg. Pension reform and labor market incentives. *Journal of Population Economics*, 23(2):769–803, 2008.

W. Frimmel, T. Horvath, M. Schnalzenberger, and R. Winter-Ebmer. Seniority wages and the role of firms in retirement. *Journal of Public Economics*, 164:19–32, 2018.

J. E. Galdón-Sánchez and M. Güell. Dismissal conflicts and unemployment. *European Economic Review*, 47(2):323–335, 2003.

L. A. Geddes and J. S. Heywood. Gender and Piece Rates, Commissions, and Bonuses. *Industrial Relations*, 42(3):419–444, 2003.

C. Göbel and T. Zwick. Are personnel measures effective in increasing productivity of old workers? *Labour Economics*, 22:80–93, 2013.

N. Guertzgen. Rent-sharing and Collective Bargaining Coverage: Evidence from Linked Employer-Employee Data. *Scandinavian Journal of Economics*, 111(2):323–349, 2009.

N. Guertzgen. Wage insurance within German firms: Do institutions matter? *Journal of the Royal Statistical Society. Series A: Statistics in Society*, 177(2):345–369, 2014.

P. Guimarães, F. Martins, and P. Portugal. Upward Nominal Wage Rigidity. IZA Discussion Paper No. 10510, 2017.

L. Guiso, L. Pistaferri, and F. Schivardi. Insurance within the Firm. *Journal of Political Economy*, 113(5):1054–1087, 2005.

C. Haefke, M. Sonntag, and T. van Rens. Wage rigidity and job creation. *Journal of Monetary Economics*, 60(8):887–899, 2013.

M. Hagedorn and I. Manovskii. The Cyclical Behavior of Equilibrium Unemployment and Vacancies Revisited. *American Economic Review*, 98(4):1692–1706, 2008.

J.-O. Hairault, F. Langot, and A. Zylberberg. Equilibrium unemployment and retirement. *European Economic Review*, 79:37–58, 2015.

R. E. Hall and E. P. Lazear. The excess sensitivity of layoffs and quits to demand. *Journal of Labor Economics*, 2(2):233–257, 1984.

J. D. Hamilton. State-space models. In R. Engle and D. McFadden, editors, *Handbook of Econometrics*, volume 4, chapter 50, pages 3039–3080. 1994.

W. Härdle, A. Werwatz, M. Müller, and S. Sperlich. *Nonparametric and Semiparametric Models.* Springer Series in Statistics. Springer Berlin Heidelberg, 2004.

M. Hashimoto. Human Capital as a Shared Investment. *American Economic Review*, 71(3): 475–482, 1981.

R. Holzmann, Y. Pouget, M. Vodopivec, and M. Weber. Severance Pay Programs around the World: History, Rationale, Status, and Reforms. In R. Holzmann and M. Vodopivec, editors, *Reforming Severance Pay: An International Perspective*, chapter 2, pages 17–120. The World Bank, 2012.

A. Hornstein, P. Krusell, and G. L. Violante. Unemployment and vacancy fluctuations in the matching model: inspecting the mechanism. *Economic Quarterly*, 91(3):19–50, 2005.

C. Jaag, C. Keuschnigg, and M. Keuschnigg. Pension reform, retirement, and life-cycle unemployment. *International Tax and Public Finance*, 17(5):556–585, 2010.

T. Jappelli and L. Pistaferri. The Consumption Response to Income Changes. *Annual Review of Economics*, 2(1):479–506, 2010.

D. S. Johnson, J. A. Parker, and N. S. Souleles. Household Expenditure and the Income Tax Rebates of 2001. *American Economic Review*, 96(5):1589–1610, 2006.

G. Kaplan and G. L. Violante. A Model of the Consumption Response to Fiscal Stimulus Payments. *Econometrica*, 82(4):1199–1239, 2014.

G. Kátay. Do Firms Provide Wage Insurance against Shocks? *Scandinavian Journal of Economics*, 118(1):105–128, 2016.

J. Kennan. Private Information, Wage Bargaining and Employment Fluctuations. *The Review of Economic Studies*, 77(2):633–664, 2010.

A. Kettemann, F. Kramarz, and J. Zweimüller. Job Mobility and Creative Destruction: Flexicurity in the Land of Schumpeter. CEPR Discussion Paper No. 12112, 2017.

F. Kleibergen and R. Paap. Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics*, 133(1):97–126, 2006.

W. Klosterhuber, J. Heining, and S. Seth. Linked-Employer-Employee-Data from the IAB: LIAB Longitudinal Model 1993-2010 (LIAB LM 9310). FDZ-Datenreport 08/2013, 2013.

M. Knell, W. Köhler-Töglhofer, and D. Prammer. Jüngste Pensionsreformen in Österreich und ihre Auswirkungen auf fiskalische Nachhaltigkeit und Pensionsleistungen. *Geldpolitik & Wirtschaft*, Q2:72–100, 2006.

R. Konle-Seidl. Retention and Re-integration of older workers into the labour market: What works? IAB Discussion Paper 17/2017, 2017.

E. Lalé. Labor-market frictions, incomplete insurance and severance payments. *Review of Economic Dynamics*, 2018. doi: 10.1016/j.red.2018.10.001. In press.

E. P. Lazear. Why is there mandatory retirement? *The Journal of Political Economy*, 87(6): 1261–1284, 1979.

E. P. Lazear. Job security provisions and employment. *The Quarterly Journal of Economics*, 105(3):699–726, 1990.

A. Lindbeck and D. J. Snower. *The Insider-Outsider Theory of Employment and Unemployment.* The MIT Press, Cambridge, MA, 1989.

M. Marmot, J. Banks, R. Blundell, C. Lessof, and J. Nazroo. *Health, wealth and lifestyles of the older population in England: ELSA 2002.* Institute for Fiscal Studies, 2003.

A. Mas and E. Moretti. Peers at Work. *American Economic Review*, 99(1):112–145, 2009.

G. Menzio and E. R. Moen. Worker replacement. *Journal of Monetary Economics*, 57(6): 623–636, 2010.

G. Menzio, I. A. Telyukova, and L. Visschers. Directed search over the life cycle. *Review of Economic Dynamics*, 19:38–62, 2016.

J. Messina, C. F. Duarte, M. Izquierdo, P. Caju, and N. L. Hansen. The Incidence of Nominal and Real Wage Rigidity: An Individual-Based Sectoral Approach. *Journal of the European Economic Association*, 8(2-3):487–496, 2010.

K. Misra and P. Surico. Consumption, Income Changes, and Heterogeneity: Evidence from Two Fiscal Stimulus Programs. *American Economic Journal: Macroeconomics*, 6(4):84–106, 2014.

E. R. Moen. Competitive Search Equilibrium. *Journal of Political Economy*, 105(2):385–411, 1997.

D. T. Mortensen and C. A. Pissarides. Job creation and job destruction in the theory of unemployment. *The Review of Economic Studies*, 61(3):397–415, 1994.

D. T. Mortensen and C. A. Pissarides. New developments in models of search in the labor market. In O. C. Ashenfelter and D. Card, editors, *Handbook of Labor Economics*, volume 3, part B, chapter 39, pages 2567–2627. Elsevier, 1999.

OECD. *Ageing and Employment Policies: Austria.* OECD Publishing, 2005.

OECD. *Live Longer, Work Longer.* Ageing and Employment Policies. OECD Publishing, 2006.

OECD. *OECD Employment Outlook 2013.* OECD Publishing, 2013.

OECD. *OECD Employment Outlook 2014.* OECD Publishing, 2014.

OECD. LFS by sex and age, 2018. URL `https://stats.oecd.org/Index.aspx?DataSetCode=LFS_SEXAGE_I_R`. Date accessed: Feb. 11, 2018.

W. Y. Oi. Labor as a Quasi-Fixed Factor. *Journal of Political Economy*, 70(6):538–555, 1962.

J. A. Parker, N. S. Souleles, D. S. Johnson, and R. McClelland. Consumer Spending and the Economic Stimulus Payments of 2008. *American Economic Review*, 103(6):2530–2553, 2013.

B. Petrongolo and C. Pissarides. Looking into the black box: a survey of the matching function. *Journal of Economic Literature*, 39(2):390–431, 2001.

C. A. Pissarides. *Equilibrium unemployment theory.* The MIT Press, 1990.

G. Ramey and J. Watson. Contractual fragility, job destruction, and business cycles. *The Quarterly Journal of Economics*, 112(3):873–911, 1997.

D. Roodman. Practitioners' corner: A note on the theme of too many instruments. *Oxford Bulletin of Economics and Statistics*, 71(1):135–158, 2009a.

D. Roodman. How to do xtabond2: An introduction to difference and system GMM in Stata. *The Stata Journal*, 9(1):86–136, 2009b.

L. Rudanko. Labor market dynamics under long-term wage contracting. *Journal of Monetary Economics*, 56(2):170–183, 2009.

M. R. Sampford. Some Inequalities on Mill's Ratio and Related Functions. *The Annals of Mathematical Statistics*, 24(1):130–132, 1953.

C. Shapiro and J. E. Stiglitz. Equilibrium unemployment as a worker discipline device. *American Economic Review*, 74(3):433–444, 1984.

R. Shimer. *Essays in search theory.* PhD Thesis, Massachusetts Institute of Technology, 1996.

V. Skirbekk. Age and Individual Productivity: A Literature Survey. *Vienna Yearbook of Population Research*, 1:133–154, 2004.

V. Skirbekk. Age and productivity capacity: descriptions, causes and policy options. *Ageing Horizons*, 8:4–12, 2008.

Statistik Austria. *Verdienststrukturerhebung 2002.* Verlag Österreich GmbH, 2006.

Statistik Austria. *Verdienststrukturerhebung 2006.* Verlag Österreich GmbH, 2009.

Statistik Austria. *Verdienststrukturerhebung 2014.* Verlag Österreich GmbH, 2017.

Statistik Austria. Arbeitslosenleistungen, 2018. URL http://statistik.at/web_de/statistiken/menschen_und_gesellschaft/soziales/sozialleistungen_auf_bundesebene/arbeitslosenleistungen/index.html. Date accessed: Feb. 11, 2018.

S. Staubli and J. Zweimüller. Does raising the early retirement age increase employment of older workers? *Journal of Public Economics*, 108:17–32, 2013.

J. E. Stiglitz. Alternative Theories of Wage Determination and Unemployment in LDC's: The Labor Turnover Model. *The Quarterly Journal of Economics*, 88(2):194–227, 1974.

J. Thomas and T. Worrall. Self-enforcing wage contracts. *The Review of Economic Studies*, 55 (4):541–554, 1988.

A. Weiss. Job Queues and Layoffs in Labor Markets with Flexible Wages. *Journal of Political Economy*, 88(3):526–538, 1980.

F. Windmeijer. A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of Econometrics*, 126(1):25–51, 2005.

R. Winter-Ebmer. Benefit duration and unemployment entry: A quasi-experiment in Austria. *European Economic Review*, 47:259–273, 2003.

J. M. Wooldridge. *Econometric Analysis of Cross Section and Panel Data.* The MIT Press, Cambridge, 2002.

Abstract

This dissertation consists of three essays relating to wage rigidities in ongoing employment relations. If employment is at will, efficient labor turnover requires occasional wage adjustments to make continuing employment mutually beneficial for the firm and the worker. In practice, the scope for wage adjustment may be restricted by institutional regulations and market failures. The first two essays study the implications of a specific form of wage rigidity within a search and matching framework. The third essay provides novel empirical insights.

The first essay analyzes an age-structured directed search model where wages cannot react to stochastic fluctuations in match productivity. I show that although this friction increases layoff rates at all ages, it particularly decreases the employment rate of elderly workers. Additionally, the contracting friction lowers the effectiveness of pension reforms. Restricting access to early retirement should therefore be complemented by labor market policies that improve firms' willingness to keep elderly workers employed.

One suitable policy in this regard is severance pay. Therefore, the second essay investigates how contractual flexibility and worker's risk attitudes shape the socially optimal design of severance pay. If workers are risk neutral or search frictions for workers are negligible, the optimal level of severance pay turns out to be independent of the severity of bilateral contracting frictions. Otherwise, severance pay should increase with the severity of the friction. Extending the analysis to dynamic contracts, I find that moral hazard considerations are central to understand why severance pay is often increasing in tenure.

The third essay empirically investigates wage rigidity at the micro level and explores how size and persistence of idiosyncratic firm-level productivity shocks affect individual wages and layoff probabilities. Using a novel estimation strategy, I document that wages respond largely symmetrically to permanent productivity shocks, while only negative transitory shocks affect wages. Layoff probabilities only respond to negative permanent productivity shocks. Interestingly, real wage cuts and employment loss after negative productivity shocks are limited to blue-collar workers, while white-collar workers appear to be fully insured against negative productivity shocks, both in terms of wages and in terms of employment.

Zusammenfassung

Die vorliegende Disseration besteht aus drei Aufsätzen zu Lohnrigiditäten in bestehenden Beschäftigungsverhältnissen. Wenn Arbeitsverträge einseitig aufgelöst werden können, sind gelegentlich Lohnanpassungen notwendig, damit eine Weiterbeschäftigung im Interesse von Arbeitergeber und Arbeitnehmer ist. In der Praxis sind solchen Lohnanpassungen Grenzen gesetzt, einerseits durch gesetzliche Regelungen und andererseits durch Marktimperfektionen. Die ersten beiden Aufsätze untersuchen die Auswirkungen einer speziellen Form von Lohnrigidität innerhalb eines Arbeitsmarktmodelles mit Such- und Matchingfriktionen. Der dritte Aufsatz präsentiert neue empirische Resultate.

Der erste Aufsatz analysiert ein altersstrukturiertes Arbeitsmarktmodell unter der Annahme, dass Löhne nicht auf stochastische Schwankungen in der Produktivität des Beschäftigungsverhältnisses reagieren können. Es zeigt sich, dass diese Restriktion besonders die Erwerbsquote im späteren Erwerbsalter reduziert. Darüber hinaus schmälert die Lohnrigidität die Beschäftigungszugewinne, welche sich durch Pensionsreformen erzielen lassen. Eine Reduzierung von Frühpensionsmöglichkeiten sollte demnach von Arbeitsmarktmaßnahmen begleitet werden, welche den Anreiz für Unternehmen erhöhen, ältere Arbeitnehmer länger zu beschäftigen.

Der zweite Aufsatz untersucht das gesellschaftlich optimale Design von Abfertigungssystemen in Abhängigkeit von der Risikopräferenz der Arbeitnehmer und von Restriktionen in der Gestaltung von Arbeitsverträgen. Falls Arbeitnehmer risikoneutral sind oder keine Suchfriktionen auf dem Arbeitsmarkt wahrnehmen, ist der optimale Abfertigungsbetrag unabhängig vom Ausmaß allfälliger Vertragsfriktionen. Andernfalls sollten Abfertigungszahlungen umso höher sein, je stärker eingeschränkt Arbeitnehmer und Arbeitgeber in der Vertragsgestaltung sind. Eine dynamische Analyse zeigt außerdem, dass Anreize für Fehlverhalten am Arbeitsplatz ausschlaggebend dafür sind, dass Abfertigungszahlungen oft mit der Verweildauer im Betrieb steigen.

Der dritte Aufsatz untersucht Lohnrigiditäten auf Mikroebene anhand von Betriebsdaten. Es wird aufgezeigt, wie das Ausmaß und die Persistenz von betriebsspezifischen Produktivitätsschocks die Löhne und Entlassungswahrscheinlichkeiten der Beschäftigten beeinflussen. Anhand einer neuartigen Schätzstrategie zeigt sich, dass Löhne großteils symmetrisch auf permanente Produktivitätsschocks reagieren, während nur negative transitorische Schocks die Lohnentwicklung beeinflussen. Die Entlassungswahrscheinlichkeiten reagieren lediglich auf negative permanente Produktivitätsschocks. Die aufgezeigten Effekte sind sehr heterogen. Reale Lohnkürzungen und Entlassungen treffen hauptsächlich Arbeiter, während Angestellte bei negativen Schocks perfekt abgesichert scheinen, sowohl gegen Lohnentgang als auch gegen Jobverlust.

Curriculum Vitae

Martin Kerndler
[redacted]

Personal Information
Year of birth: 1988
Citizenship: Austrian

Current position

University assistant (predoc) at Vienna University of Technology (TU Wien)
Institute of Statistics and Mathematical Methods in Economics
Research group Economics (since 10/2017)

Education

since 10/2013	**PhD candidate in Economics** Vienna Graduate School of Economics, University of Vienna expected graduation: January 2019
10/2011–09/2013	**MSc in "Mathematics in Economics"** TU Wien, Graduation with honors 02–06/2012 Erasmus semester at the University of Aarhus, Denmark
10/2008–09/2011	**BSc in "Statistics and Mathematics in Economics"** TU Wien, Graduation with honors
09/2002–05/2007	**Matura** at Bundeshandelsakademie Krems

Previous academic positions

09/2016–07/2017	**Lecturer** University of Vienna, Department of Economics
10/2013–09/2016	**Research assistant** University of Vienna, Department of Economics
10/2015–02/2016	**External lecturer** TU Wien, Institute of Mathematical Methods in Economics

07/2012–01/2004	**Research assistant**
10/2011–01/2012	TU Wien, Institute of Mathematical Methods in Economics *OeNB project "Industrial Diversity, Spatial Differentiation and Social Cohesion: Communicative Structures in the 'Housing Topos'"*

Publications

E. Deutsch and M. Kerndler. Produktivität durch Soziale Kohäsion: Der Beitrag der österreichischen Sozialmiete. In G. Biffl and N. Dimmel, editors, *Migrationsmanagement 2 – Wohnen im Zusammenwirken mit Migration und Integration*, omnium Verlag, ISBN 3390310186, 2016

Teaching

MSc level:

Instructor for *Dynamic Macroeconomics with Numerics* by Monika Gehrig-Merz
University of Vienna, summer term 2017

Instructor for *Dynamic Macroeconomics* by Alexia Fürnkranz-Prskawetz
TU Wien, winter term 2015/16

BSc level:

Lecturer for *Macroeconomics for Students of Mathematics in Economics*
TU Wien, winter term 2018/19

Instructor for *Principles of Microeconomics* by Timo Trimborn
TU Wien, summer term 2018

Instructor for *Macroeconomics for Students of Economics* by Monika Gehrig-Merz
University of Vienna, winter terms 2015/16 and 2016/17

Conference Presentations

11/2018	Workshop Arbeitsmarktökonomie, Vienna
05/2018	Annual Meeting of the Austrian Economic Association, Vienna
09/2017	Jahrestagung des Vereins für Socialpolitik, Vienna (poster)
08/2017	32nd Annual Congress of the European Economic Association, Lisbon
03/2017	RGS Doctoral Conference in Economics, Dortmund
12/2016	International Workshop on Heterogeneous Dynamic Models of Economic and Population Systems, Vienna (invited)
11/2016	Workshop Arbeitsmarktökonomie, Vienna
09/2016	Jahrestagung des Vereins für Socialpolitik, Augsburg
05/2016	Joint Annual Meeting of the Slovak Economic Association and the Austrian Economic Association, Bratislava

03/2016	6th Ifo Dresden Workshop on Labor Economics and Social Policy, Dresden
12/2015	8th Interdisciplinary Ph.D. Workshop "Perspectives on (Un-)Employment", Nuremberg
09/2015	9th European Workshop on Labour, Health and Education under Demographic Change, Vienna
05/2015	QED Jamboree, Cardiff
11/2013	International Workshop on Heterogenous Dynamic Models of Economic and Population Systems, Vienna
11/2013	International Conference on Health, Education and Retirement over the Prolonged Life-Cycle, Vienna (poster)
09/2013	8th European Workshop on Labour Markets and Demographic Change, Vienna

Further Conference Participations

09/2018	33rd Annual Congress of the European Economic Association, Cologne
08/2017	European Forum Alpbach, Alpbach, Austria
08/2016	European Forum Alpbach, Alpbach, Austria
08/2014	5th Lindau Meeting on Economic Sciences, Lindau

Awards and Fellowships

05/2018	Young Economist Award, Nationaloekonomische Gesellschaft (Austrian Economic Association)
08/2016, 08/2017	Scholarship from Club Alpbach Lower Austria for the European Forum Alpbach
09/2013–09/2016	Full Fellowship, Vienna Graduate School of Economics
08/2014	Fellow of the Austrian Federal Ministry of Science and Research at the 5th Lindau Meeting on Economic Sciences
2010–2013	Merit scholarship from TU Wien, Faculty of Mathematics and Geoinformation

Previous non-academic positions

08/2011	Raiffeisenlandesbank Niederösterreich-Wien, Department for balance sheet and business analysis
04/2008–09/2008	Raiffeisenbank Region St. Pölten, Controlling/Accounting