Contracting frictions and inefficient layoffs over the life-cycle

Martin Kerndler*

July 13, 2018

Abstract

Low employment rates above age 55 are a major policy concern in many European countries. This paper analyzes the role of layoffs for the employment the elderly. Realistic frictions in wage contracting are introduced into an age-structured directed search model of the labor market. It turns out that although the contracting friction generates inefficiently high layoff rates at all ages, it particularly depresses the employment rate of the elderly. Moreover, the friction lowers 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.

Keywords: contracting frictions, layoffs, elderly workers, early retirement

JEL classification: J14, J31, J41, J63, J68

1 Introduction

For its *Employment Outlook 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."*

*TU Wien, Institute of Statistics and Mathematical Methods in Economics, Wiedner Hauptstraße 8/105-3, 1040 Vienna, Austria; martin.kerndler@econ.tuwien.ac.at. For valuable comments and suggestions, I thank Michael Reiter and Támas Papp, as well as participants at various seminars and conferences. Funding for this project was partially provided by the Vienna Graduate School of Economics.

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 policymakers 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 (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

[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.

[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).

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, the present paper 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, 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 paper is structured as follows. Section 2 briefly summarizes the literature on inefficient layoffs and motivates the particular friction considered in this paper. Section 3 introduces the model. Section 4 derives the equilibrium and comparative static effects. The analytical results are complemented by a numerical assessment in Section 6, which illustrates the role of the friction when an early retirement reform is enacted, and investigates complementary labor market reforms. Section 7 concludes. Appendix A contains an overview of all defined functions, variables, and parameters. All proofs and additional lemmas are delegated to Appendix B.

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 fric-

[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.

tions, 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.

The present paper embeds 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.

As pointed out by Boeri et al. (2017) this set of assumptions can be rationalized by asymmetric information, where the productivity draw is private knowledge of the firm. Alternative microfoundations for the absence of renegotiation may include employer's considerations about motivation, fairness, and the use of wage contracts as a screening device for new hires.

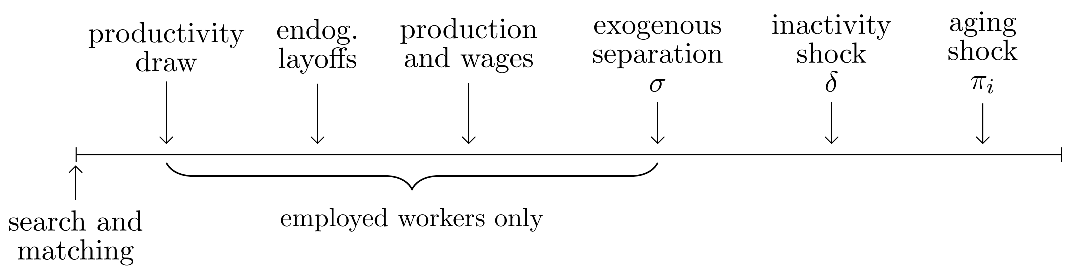

Figure 1: Timing within a period

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.

3 Model setup

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 1. At the beginning of a period, unemployed workers apply to vacancies that offer some wage contract ω_i. With probability $p(\theta_i)$

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

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 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.

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 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 $F_i(y) = F\big(\frac{y-\mu_i}{s_i}\big)^{\alpha_i}$ 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 illustrates how the density function and cumulative distribution function are affected by changes in α_i, taking the standard normal distribution as reference

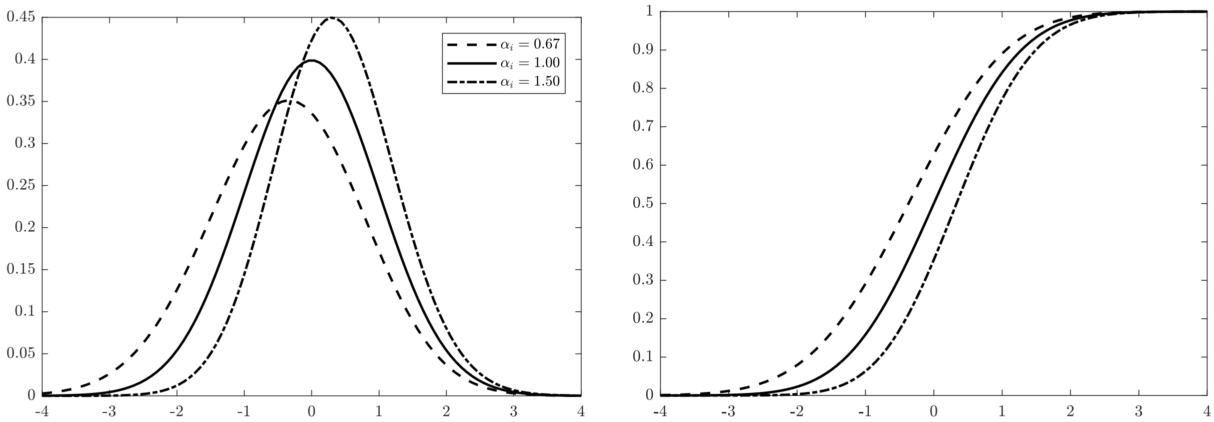

Figure 2: Density and distribution function of the normal distribution with $\mu_i = 0$, $s_i = 1$, and different levels of α_i.

($F = \Phi$). 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 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 Appendix B.1.

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 $q(\theta_i) = \frac{M(JS_i,V_i)}{V_i} = M\big(\frac{1}{\theta_i}, 1\big)$, 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, $\omega_o = (w_o)$.

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 6 I allow for additional government spending and/or revenue from labor market policies.

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 4.1 considers the labor market equilibrium of old job seekers, before I turn to prime-age job seekers in Section 4.2. Section 4.3 defines aggregate economic measures and the equilibrium tax level. The analysis proceeds under the following functional restrictions:

Assumption 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 the paper 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 3. *Assume that $\alpha_i = 1$ for all $i \in \{m, s, o\}$.*

Under Assumption 3, the monotonicity properties of the hazard rate h demanded by Assumption 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 $i \in \{m, s, o\}$.

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

$$V_o := \max_{(\theta_o, w_o)} p(\theta_o)\mathbb{E}W_o^+(w_o) \quad \text{s.t.} \quad q(\theta_o)\mathbb{E}J_o^+(w_o) = c. \tag{1}$$

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

$$J_o(w_o; y) = y - w_o + \beta_o[\phi\mathbb{E}J_o^+(w_o) + (1-\phi)J_o(w_o; y)], \tag{2}$$

$$W_o(w_o) = u(w_o - \tau) - u(b_o - \tau) + \beta_o[\phi\mathbb{E}W_o^+(w_o) + (1-\phi)W_o(w_o) - V_o], \tag{3}$$

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), $J_o(w_o; y) = \frac{y - \underline{y}_o(w_o)}{1-\beta_o(1-\phi)}$, and therefore the layoff threshold solves

$$\underline{y}_o - w_o + \frac{\beta_o \phi}{1 - \beta_o(1-\phi)} \int_{\underline{y}_o}^{\infty} y - \underline{y}_o \, dF_o(y) = 0. \tag{4}$$

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

Proposition 1. *For any $w_o \in \mathbb{R}$, equation* (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 Appendix 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 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 (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]

4.1.1 Equilibrium conditions

The first order optimality conditions of problem (1) can be summarized as

$$u'(w_o^* - \tau) = \frac{1-\gamma}{\gamma} \frac{W_o(w_o^*)}{J_o(\underline{y}_o^*)} + (1 - \beta_o(1-\phi)) h_o(\underline{y}_o^*) \frac{\partial \underline{y}_o^*}{\partial w_o} W_o(w_o^*), \tag{5}$$

$$q(\theta_o^*) \mathbb{E} J_o^+(w_o^*) = c, \tag{6}$$

where $\underline{y}_o^* = y_o(w_o^*)$ is defined in (4). The left-hand side of equation (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

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

layoff threshold y_o, measured by the partial derivative $\frac{y_o^*}{\partial w_o} = \frac{1-\beta_o(1-\phi)}{1-\beta_o(1-\phi F_o(\underline{y}_o^*))} > 0$, 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 (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 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 (4)–(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. *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 $F_o(w_o^*) - F_o(b_o)$.

4.1.2 Comparative static effects

To obtain comparative static effects, I continue to assume that old age lasts for one period only, $\pi_o = 1$. Equation (5) then can be expressed as

$$\Phi(w_o^*) = u'(w_o^* - \tau) - \frac{1-\gamma}{\gamma}\frac{W_o(w_o^*)}{J_o(w_o^*)} - h_o(w_o^*)W_o(w_o^*) = 0, \tag{7}$$

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

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

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 (7) and triggers a *substitution effect (SE)*,

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

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

$$\frac{dF_o(w_o^*)}{d\xi} = \frac{\partial F_o(w_o^*)}{\partial \xi} + f_o(w_o^*)\frac{\partial w_o^*}{\partial \xi} = \underbrace{\frac{\partial F_o(w_o^*)}{\partial \xi} + f_o(w_o^*)\left(\frac{\partial w_o^*}{\partial \xi}\right)^{IE}}_{IE}\underbrace{+f_o(w_o^*)\left(\frac{\partial w_o^*}{\partial \xi}\right)^{SE}}_{SE}. \tag{8}$$

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 (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,

$$\begin{aligned}\frac{d\mathbb{E}J_o^+(w_o^*)}{d\xi} &= -\int_{w_o^*}^{\infty}\frac{\partial F_o(y)}{\partial \xi}\,dy - (1-F_o(w_o^*))\frac{\partial w_o^*}{\partial \xi} \\ &= \underbrace{-\int_{w_o^*}^{\infty}\frac{\partial F_o(y)}{\partial \xi}\,dy - (1-F_o(w_o^*))\left(\frac{\partial w_o^*}{\partial \xi}\right)^{IE}}_{IE}\underbrace{-(1-F_o(w_o^*))\left(\frac{\partial w_o^*}{\partial \xi}\right)^{SE}}_{SE}.\end{aligned} \tag{9}$$

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 B.1 and Lemma B.2 in Appendix 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 3 establishes that this negative direct effect dominates the positive wage effect in (8) and (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 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 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 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 (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 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 (9). If $\frac{\partial w_o^*}{\partial s_o} > 0$, the upper boundary on γ established by Proposition 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.

4.2 Labor market equilibrium of prime-age job seekers

After this detailed analysis of old job seekers, I turn to the search problem of prime-age job seekers who search 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

$$J_m(\omega_m; y) = y - w_m + \beta_m[\phi\mathbb{E}J_m^+(\omega_m) + (1-\phi)J_m(\omega_m; y)] + \beta\pi_m(1-\sigma)\mathbb{E}J_s^+(w_s), \tag{10}$$

$$\begin{aligned} W_m(\omega_m) = u(w_m - \tau) - u(b_m - \tau) + \beta_m[\phi\mathbb{E}W_m^+(\omega_m) + (1-\phi)W_m(\omega_m) - V_m] \\ + \beta\pi_m(1-\sigma)[\mathbb{E}W_s^+(w_s) - V_o]. \end{aligned} \tag{11}$$

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 π_m 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 (4). The layoff threshold of a prime-age worker is denoted by $\underline{y}_m(\omega_m)$ and characterized by the equation

$$\underline{y}_m - w_m + \frac{\beta_m\phi}{1-\beta_m(1-\phi)} \int_{\underline{y}_m}^{\infty} y - \underline{y}_m \, dF_m(y) + \beta\pi_m(1-\sigma)\mathbb{E}J_s^+(w_s) = 0. \tag{12}$$

Compared to equation (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 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 $W_m(\omega_m) = \frac{u(w_m-\tau)-u(b_m-\tau)-\beta_m V_m+\beta\pi_m(1-\sigma)[\mathbb{E}W_s^+(w_s)-V_o]}{1-\beta_m(1-\phi F_m(\underline{y}_m))}$.

4.2.1 Equilibrium conditions

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

$$u'(w_m^* - \tau) = \frac{1-\gamma}{\gamma}\frac{W_m(\omega_m^*)}{J_m(\underline{y}_m^*)} + (1-\beta_m(1-\phi))h_m(\underline{y}_m^*)\frac{\partial \underline{y}_m^*}{\partial w_m}W_m(\omega_m^*), \tag{13}$$

$$u'(w_s^* - \tau) = u'(w_m^* - \tau) + (1-\beta_o(1-\phi))h_s(\underline{y}_s^*)\frac{\partial \underline{y}_s^*}{\partial w_s}W_s(w_s^*), \tag{14}$$

$$q(\theta_m^*)\mathbb{E}J_m^+(\omega_m^*) = c, \tag{15}$$

where the layoff threshold $\underline{y}_m^* = \underline{y}_m(\omega_m^*)$ is defined in (12) and $\underline{y}_s^* = \underline{y}_s(w_s^*)$ is defined analogous to (4). Condition (13) resembles equation (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 (13) determines the present value that the worker receives in optimum, condition (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 (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 (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 3 visualizes the two equations (13)–(14) in the (w_m, w_s)-space. Condition (13) defines a decreasing curve, which I refer to as the surplus sharing (SS) curve in Figure 3. It connects all wage combinations that implement the optimal surplus sharing rule. Condition (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 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 $\overline{b}_o > b_m$, such that for $b_o \in [b_m, \overline{b}_o)$ the wage contract is interior and the wage level is decreasing with age, $w_m^* > w_s^* >$*

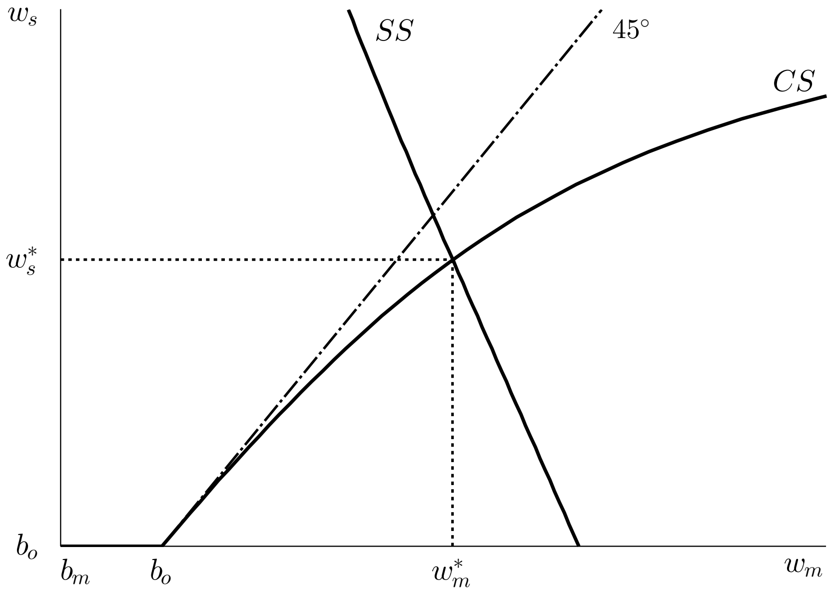

Figure 3: Wage determination of prime-age job seekers.

b_o. For $b_o \geq \bar{b}_o$, the optimal contract satisfies $w_m^* \leq w_s^* = b_o$.

Proposition 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 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 grants a 30% higher unemployment income to senior workers compared to prime-age workers, the optimal contract is still interior, as can be seen from Table 3.

4.2.2 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 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 4.1.2. From the first order conditions (13)–(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 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 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 3 and Proposition 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 (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 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 4.1.2 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 (13) is in general ambiguous. Under an additional assumption, however, the effect on firm surplus dominates.

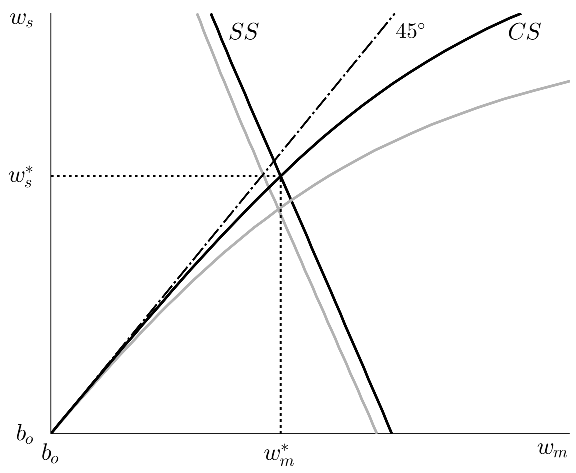

Figure 4: Wage response to an increase in μ_s.

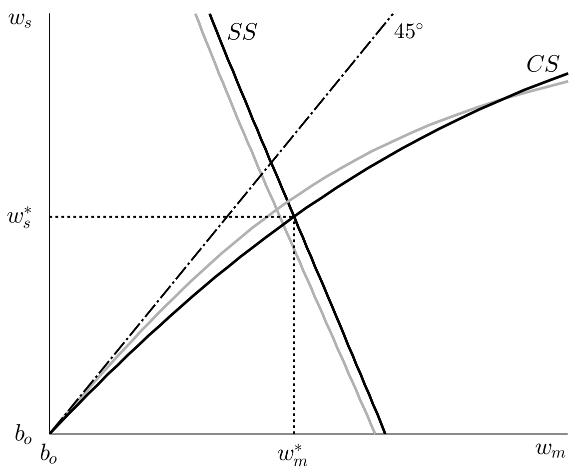

Figure 5: Wage response to an increase in s_s.

Proposition 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 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 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 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 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 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 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 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 $1 - F_s(w_s^*)$.

[7] Note that $u'(w_m^* - \tau) \leq \frac{1}{\gamma}\frac{W_m(\omega_m^*)}{J_m(\underline{y}_m^*)}$ by (13) and Lemma B.1(i). Therefore the assumption is satisfied if w_s^* is not substantially lower than w_m^*.

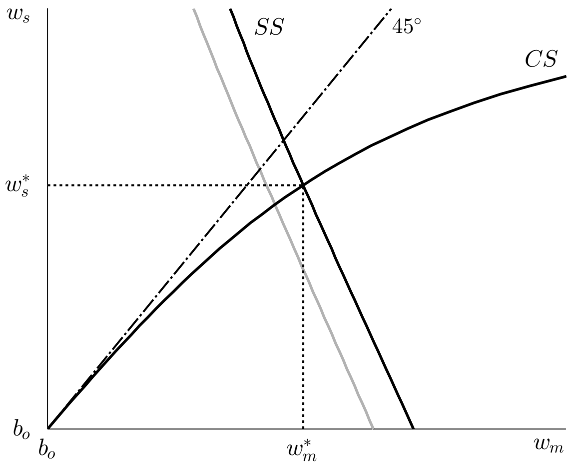

Figure 6: Wage response to an increase in b_m.

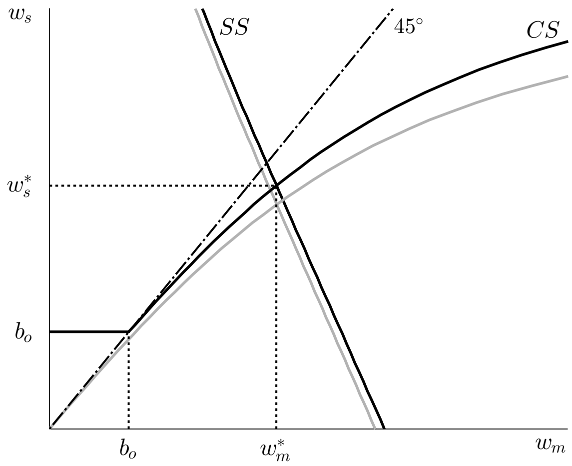

Figure 7: Wage response to an increase in b_o.

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 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 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, 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 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.

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 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 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 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,

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

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 $x = u^{-1}(\mathcal{W}/N)$.

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 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. Therefore, layoffs are bilaterally efficient, and the layoff threshold of the firm becomes the reservation productivity y_i^r implicitly defined by $W_i(w_i; y_i^r) = J_i(w_i; y_i^r) = 0$.

5.1 Labor market equilibrium of old job seekers

Firm and worker surplus at the production stage satisfy equations (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

$$u\big(y_o^r + \beta_o \phi \mathbb{E}J_o^+(w_o) - \tau\big) - u(b_o - \tau) + \beta_o \phi \mathbb{E}W_o^+(w_o) - \beta_o V_o = 0. \tag{16}$$

The equilibrium on the labor market for old job seekers is characterized as in (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]

$$w_o^\bullet(y) = \min\{\overline{w}_o^\bullet, y + \beta_o \phi \mathbb{E}J_o^+(w_o^\bullet)\} \text{ for } y \geq y_o^r, \tag{17}$$
$$u'(\overline{w}_o^\bullet - \tau) = \frac{1-\gamma}{\gamma} \frac{\mathbb{E}W_o^+(w_o^\bullet)}{\mathbb{E}J_o^+(w_o^\bullet)} + \frac{\beta_o \phi}{1 - \beta_o(1 - \phi F_o(y_o^r))} \Delta_o, \tag{18}$$
$$q(\theta_o^\bullet) \mathbb{E}J_o^+(w_o^\bullet) = c, \tag{19}$$

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 (4). According to condition (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 (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 8 where the thick solid line corresponds to the wage schedule $w_o^\bullet(y)$.

[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 $*$.

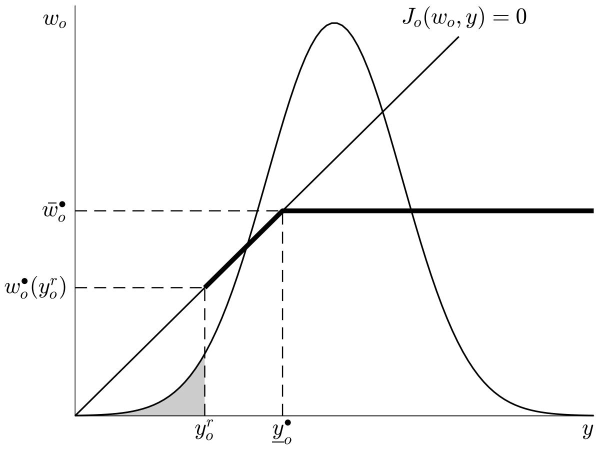

Figure 8: Labor market equilibrium of old job-seekers without the contracting friction.

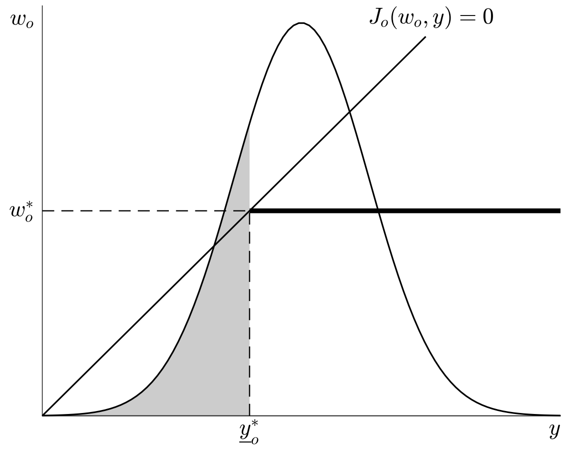

Figure 9: Labor market equilibrium of old job-seekers with the contracting friction.

Condition (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 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.

Proposition 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. In the general case, however, this condition seems necessary to ensure that the layoff probability is indeed higher with the friction.

Perhaps surprisingly, Proposition 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 8 and Figure 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 9. Without

[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.

the friction, layoffs only occur below the reservation productivity y_r as illustrated in Figure 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 (6) and (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 more vacancies. Proposition 9 implies that the contracting friction increases labor turnover, while its effect on equilibrium employment is ambiguous.

5.2 Labor market equilibrium of prime-age job seekers

Firm and worker surplus at the production stage satisfy equations (10)–(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

$$u'(\overline{w}_m^\bullet - \tau) = \frac{1-\gamma}{\gamma}\frac{\mathbb{E}W_m^+(\omega_m^\bullet)}{\mathbb{E}J_m^+(\omega_m^\bullet)} + \frac{\beta_m\phi}{1-\beta_m(1-\phi F_m(y_m^r))}\Delta_m, \tag{20}$$

$$u'(\overline{w}_s^\bullet - \tau) = \mathbb{E}[u'(w_m^\bullet - \tau)|y \geq y_m^r] + \frac{\beta_o\phi}{1-\beta_o(1-\phi F_s(y_s^r))}\Delta_s, \tag{21}$$

where y_i^r is the reservation productivity of a type i worker. As in (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 (20). The optimal age profile of wages is determined by condition (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 (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 (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.

5.3 Economic aggregates

With productivity-contingent contracts, all demographic and aggregate economic variables are defined as in Section 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

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

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.

6 Numerical illustration and policy implications

To assess the quantitative importance of the contracting friction, I solve the model outlined in Section 4 numerically and compare it to the counterfactual economy without the friction described in Section 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.

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).

[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.

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 1: Parameters set directly

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 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 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%. Since up to age 60 only unemployment benefits can be collected, g_o is set to reflect a replacement rate of $\frac{0.531+0.932\cdot 0.94}{2} = 0.704$.

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

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: Calibrated parameter values and calibration targets

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.

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_2N_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. 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

[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.

(a) with contracting friction

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	

(b) without contracting friction

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 $\mathbb{E}[w_i^\bullet \vert y \geq y_i^r]$	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 3: Equilibrium for the baseline economy

surplus calibration of Hagedorn and Manovskii (2008) [0.955].

6.2 Equilibrium

Panel (a) of Table 3 shows the equilibrium of the calibrated model. In line with Proposition 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 the old job seekers (type o in Table 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 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 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 as given. The corresponding equilibrium is given in panel (b) of Table 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 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 3 reveals that the friction reduces aggregate welfare by 2.9% in consumption equivalents, while output is depressed by 2.7%. If individuals were naive about the link between wages and layoff risk, the aggregate costs of the friction would be even higher, see also Section 6.5.

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 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 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 Table 3(a), since this group of unemployed did not have access to ER benefits anyway.

Comparing Table 4(a) to Table 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).

(a) with contracting friction

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

(b) without contracting friction

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 4: Equilibrium after the early retirement (ER) reform

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 4(b) to Table 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. As a result, the aggregate costs of the friction are higher after the early retirement reform than before.

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

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 (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 (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 4.1.2. 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 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 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.

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 4).[12] The goal of this exercise is *not* to design 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 sever-

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

ance pay. To compare the potentials and caveats of each of these labor market programs, the analysis takes the post-reform economy of Table 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]

6.4.1 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 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 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.

[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 paper.

[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.

(a) training program with symmetric returns

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

(b) training program with asymmetric returns

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 6: Equilibrium after the ER reform and implementation of a training program

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, 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 6(b) shows that while wages are lower with asymmetric returns, the macroeconomic effects of the two scenarios are almost identical.

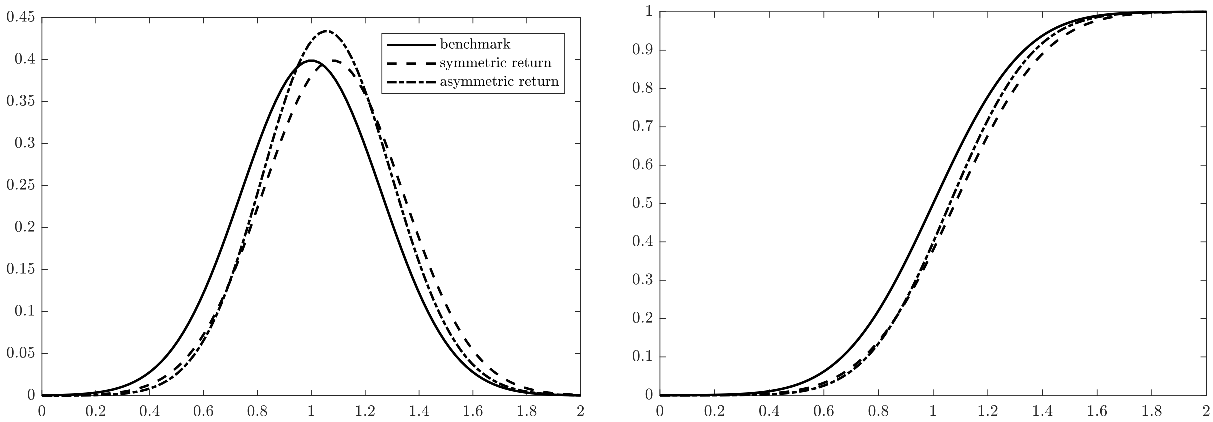

Figure 10: Density and distribution function of Y_s in the baseline ($\mu_s = \alpha_s = 1$), after the training program with symmetric returns ($\mu_s = 1.086, \alpha_s = 1$), and the program with asymmetric returns ($\mu_s = 1, \alpha_s = 1.337$).

Figure 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.

6.4.2 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 threshold, firm surplus, and equilibrium conditions can be seen from Appendix 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 by about 2.3%. The resulting increase in old age firm surplus makes a wage cost subsidy for prime-age workers almost unnecessary.

[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 7: Equilibrium after the ER reform and introduction of wage cost subsidies

Comparing Table 7 to Table 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 7 to Table 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 7 is paid to 53% of the older population.[16]

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 8: Equilibrium after the ER reform and introduction of layoff taxes

6.4.3 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 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 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%.

[16]The model generates a cost-benefit link of the wage subsidy that is empirically plausible. Albanese and Cockx (2015) 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 m	s	old job seekers 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 9: Equilibrium after the ER reform and introduction of severance pay

In general, using layoff taxes to correct the employment distortions caused by the contracting friction 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.

6.4.4 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 Appendix 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 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 9 compared to all previously considered labor market policies.[17]

[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 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.

Boeri et al. (2017) postulate the same contracting friction as this paper 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 9 to Table 4(b) reveals that while severance pay can restore the equilibrium employment levels, the labor market 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.

As with the layoff tax, the net employment effects of severance pay on old job seekers crucially depend on the response of equilibrium wages. A combination of severance pay and training might be a more robust policy and also turns out to be 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.

6.5 Bounded rationality

In Tables 3 and 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 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 C.3. Figure 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 might be 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

individual variables	prime-age job seekers m	s	old job seekers 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 10: Equilibrium after the ER reform with boundedly rational prime-age job seekers.

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 (14). An alternative interpretation of this experiment is to change s_s to zero and interpret the resulting substitution effect.

As evident from condition (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 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 3(a). Table 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 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.

7 Conclusion

In this paper, 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 sensibly 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.

References

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

A. Albanese and B. Cockx. Permanent Wage Cost Subsidies for Older Workers: An Effective Tool for Increasing Working Time and Postponing Early Retirement? IZA Discussion Paper No. 8988, 2015.

F. Alvarez and M. Veracierto. Severance payments in an economy with frictions. *Journal of Monetary Economics*, 47(3):477–498, 2001.

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.

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*, 189:1–10, September 2015.

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.

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.

C. Dauth and O.-S. Toomet. On government-subsidized training programs for older workers. *Labour*, 30:371–392, 2016.

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.

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.

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.

C. Göbel and T. Zwick. Are personnel measures effective in increasing productivity of old workers? *Labour Economics*, 22:80–93, 2013.

P. Guimarães, F. Martins, and P. Portugal. Upward Nominal Wage Rigidity. IZA Discussion Paper No. 10510, 2017.

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.

M. Hashimoto. Human Capital as a Shared Investment. *American Economic Review*, 71(3): 475–482, 1981.

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.

J. Kennan. Private Information, Wage Bargaining and Employment Fluctuations. *The Review of Economic Studies*, 77(2):633–664, 2010.

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. P. Lazear. Why is there mandatory retirement? *The Journal of Political Economy*, 87(6): 1261–1284, 1979.

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, I. A. Telyukova, and L. Visschers. Directed search over the life cycle. *Review of Economic Dynamics*, 19:38–62, 2016.

E. R. Moen. Competitive Search Equilibrium. *Journal of Political Economy*, 105(2):385–411, 1997.

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. LFS by sex and age, 2018. URL https://stats.oecd.org/Index.aspx?DataSetCode=LFS_SEXAGE_I_R. Date accessed: Feb. 11, 2018.

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.

G. Ramey and J. Watson. Contractual fragility, job destruction, and business cycles. *The Quarterly Journal of Economics*, 112(3):873–911, 1997.

M. R. Sampford. Some Inequalities on Mill's Ratio and Related Functions. *The Annals of Mathematical Statistics*, 24(1):130–132, 1953.

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.

A. Weiss. Job Queues and Layoffs in Labor Markets with Flexible Wages. *Journal of Political Economy*, 88(3):526–538, 1980.

R. Winter-Ebmer. Benefit duration and unemployment entry: A quasi-experiment in Austria. *European Economic Review*, 47:259–273, 2003.

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 4), period wage schedule (Section 5)
$\overline{w}$	base wage of the wage schedule $w(y)$ (Section 5)
$u(.)$	utility function, defined on (d, ∞)
$\underline{y}(\omega)$	layoff threshold (Section 4), profitability threshold (Section 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 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 A.2: Overview of model parameters

B Mathematical appendix

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 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 written $h(z) = e^{-z^2/2} \big[\int_z^\infty e^{-t^2/2}\, dt \big]^{-1}$. 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.

B.2 Additional lemmas

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

Lemma 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 1 and Assumption 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 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 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 1 and 3, the following properties hold:*

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

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

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

(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

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

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 (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 (B.1) and collecting terms yields

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

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 1(ii) since $\mathbb{E}Y_i = \mu_i + s_i\mathbb{E}Z$ for $\alpha_i = 1$. □

B.3 Proofs omitted in the text

Proof of Proposition 1. Define the function on the left-hand side of (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 B.2(i) and Assumption 1(ii). As a result, $\underline{y}_o$ is decreasing in both parameters. □

Proof of Proposition 2. Under $\pi_o = 1$, the equilibrium wage must satisfy $\Phi(w_o^*) = 0$, where Φ is given in (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 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 B.2(iii) implies

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

In the limit $\lim_{w\to\infty} \Phi(w) < 0$ since $u'(w)$ vanishes asymptotically by Assumption 2. 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 $\theta_o^* = \left[A \mathbb{E} J_o^+(w_o^*)/c\right]^{1/\gamma}$, and $V_o = p(\theta_o^*) \mathbb{E} W_o^+(w_o^*)$. □

Proof of Proposition 3. The increase in w_o^* follows immediately from Lemma B.1 and Lemma 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, and that $\frac{\partial J_o(w_o)}{\partial \mu_o} = -J_o'(w_o)$ and $\frac{\partial h_o(w_o)}{\partial \mu_o} = -h_o'(w_o)$. □

Proof of Proposition 4. The wage effect follows from Lemma B.1 and Lemma 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 is $(\frac{\partial w_o^*}{\partial s_o})^{SE} = \Phi'(w_o^*)^{-1} \frac{\partial h_o(w_o^*)}{\partial s_o} W_o(w_o^*)$, 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 $\Phi'(w_o^*) < -h_o'(w_o^*) W_o(w_o^*) < 0$, we have that $(\frac{\partial w_o^*}{\partial s_o})^{SE} > \Phi'(w_o^*)^{-1} h_o'(w_o^*) \frac{\mu_o - w_o^*}{s_o} W_o(w_o^*) > -\frac{\mu_o - w_o^*}{s_o}$. 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 $\frac{\partial \Phi(w_o^*)}{\partial s_o} \leq \frac{u'(w_o^* - \tau)}{s_o} \left[1 + \frac{1 - J_o(w_o^*) h_o(w_o^*)}{J_o(w_o^*)} (w_o^* - \mu_o)\right]$. The latter bound is only valid if $w_o^* \leq \mu_o$. Combining the two inequalities yields $\frac{\partial w_o^*}{\partial s_o} < \gamma \left\{ \frac{J_o(w_o^*)}{s_o} + [1 - J_o(w_o^*) h_o(w_o^*)] \frac{w_o^* - \mu_o}{s_o} \right\} = \gamma \frac{\partial J_o(w_o^*)}{\partial s_o}$. The direct effect in (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 B.2(i) and Assumption 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 $J_o(w_o^*) + w_o^* - \mu_o - s_o \frac{\partial w_o^*}{\partial s_o} > J_o(w_o^*) + w_o^* - \mu_o - \gamma \left\{ J_o(w_o^*) + [1 - J_o(w_o^*) h_o(w_o^*)](w_o^* - \mu_o) \right\}$. The right-hand side is nonnegative 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 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. 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 B.2 and Assumption 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 $\overline{b}_o$ as postulated by the proposition. □

Proof of Proposition 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 $\left(\frac{\partial w_m^*}{\partial \xi}, \frac{\partial w_s^*}{\partial \xi}\right)' = \left(-\frac{\partial \Psi}{\partial w_s}, \frac{\partial \Psi}{\partial w_m}\right)' \frac{\partial \Phi}{\partial \xi} D^{-1}$ 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 B.1 and Lemma 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 (8)–(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

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

Straightforward differentiation reveals that in optimum $\frac{\partial \Phi}{\partial w_s} = \beta(1-\sigma)(1 - F_s(w_s^*))\left[\frac{\partial \Phi}{\partial w_m} - u''(w_m^* - \tau)\right]$. 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 (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 3 and Proposition 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 3 and Proposition 4. □

Proof of Proposition 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 $\frac{\partial \Phi(\omega_m^*)}{\partial \mu_s} > -\beta(1-\sigma)(1-F_s(w_s^*))u'(w_m^* - \tau)\left[\frac{h_s(w_s^*)W_s(w_s^*)}{W_m(\omega_m^*)} + \frac{J_m'(\underline{y}_m^*)}{J_m(\underline{y}_m^*)}\right]$. 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 7 therefore ensures $\frac{\partial \Phi(\omega_m^*)}{\partial \mu_s} > 0$. □

Proof of Proposition 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 (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 (6) and (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 $\mathbb{E}J_o^+(w_o^*) > \mathbb{E}J_o^+(w_o^\bullet)$. □

C Equilibrium with labor market policies

C.1 Surplus functions and optimality conditions

To study different labor market policies in Section 6.4, the model of Section 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

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

The first order condition (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 $\frac{\underline{y}_o^*}{\partial w_o} = \frac{1-\beta_o(1-\phi)}{1-\beta_o(1-\phi F_o(\underline{y}_o^*))}$. 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,

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

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

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

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

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

Severance pay directly affects welfare, which is updated to

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

C.2 Quantitative effects

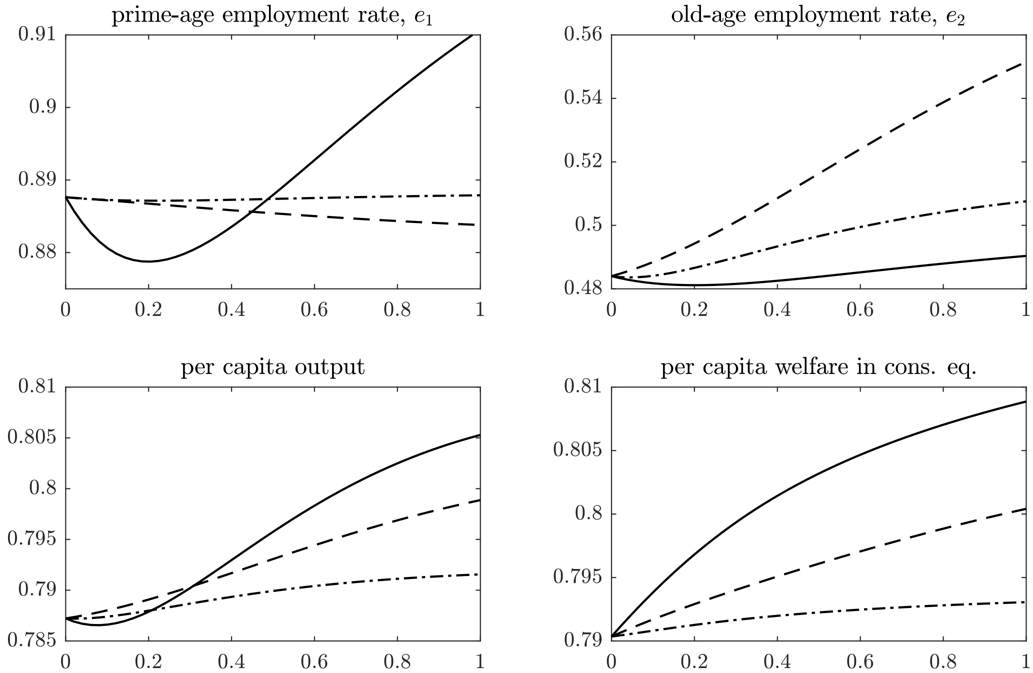

Figure C.1: effect of severance pay on employment, output and welfare, relative to Table 4(a); only one variable is altered at a time; solid line: P_m, dashed line: P_s, dash-dotted line: P_o

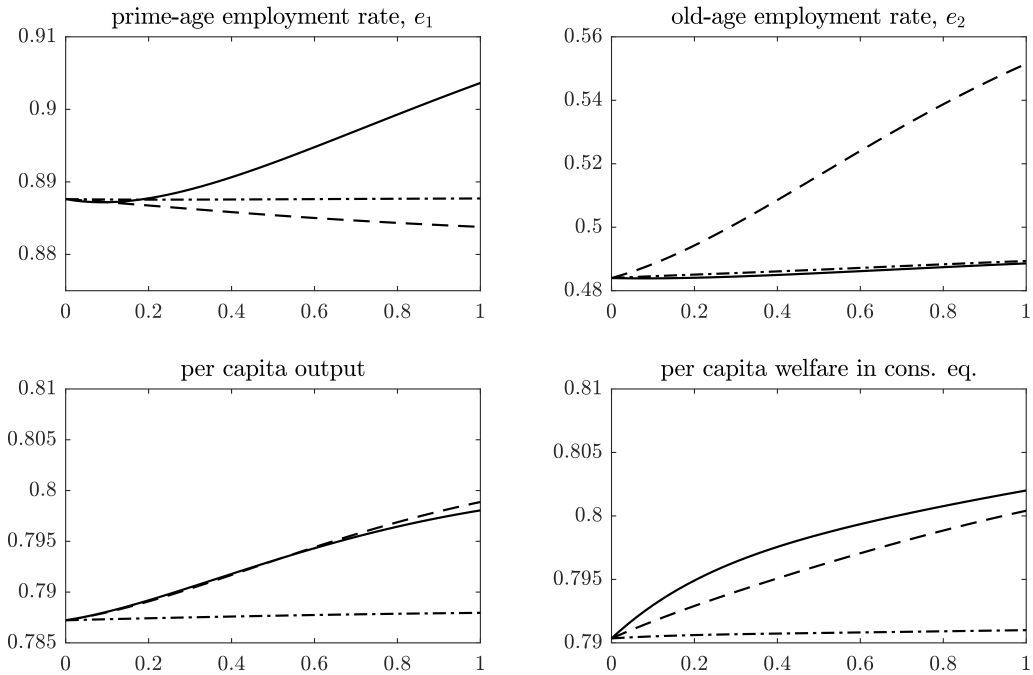

Figure C.2: effect of severance pay with a probation period on employment, output and welfare, relative to Table 4(a); only one variable is altered at a time; solid line: P_m, dashed line: P_s, dash-dotted line: P_o

C.3 Wage profiles in Austria

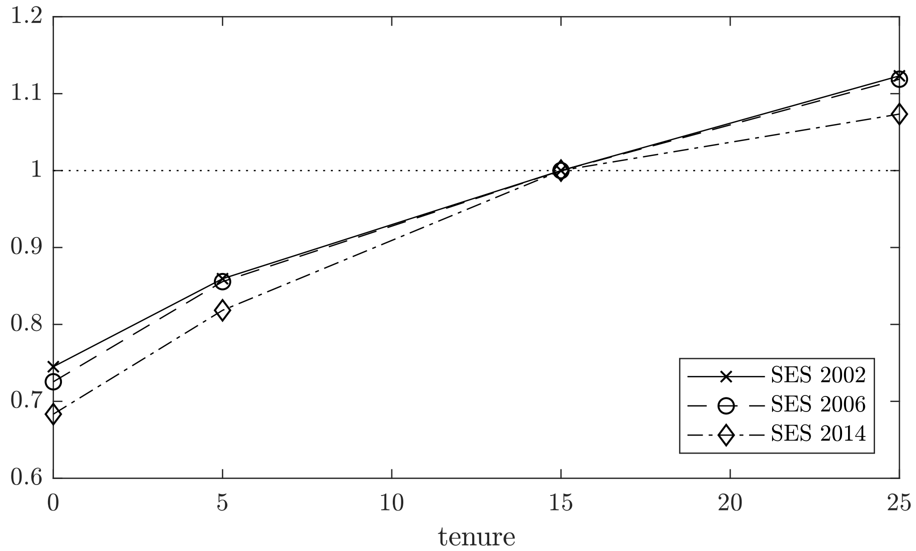

Figure C.3: hourly wage by tenure relative to tenure group 10–19, dependent employed males in the private sector in Austria, source: SES waves 2002, 2006, 2014 (Statistik Austria, 2006, 2009, 2017)

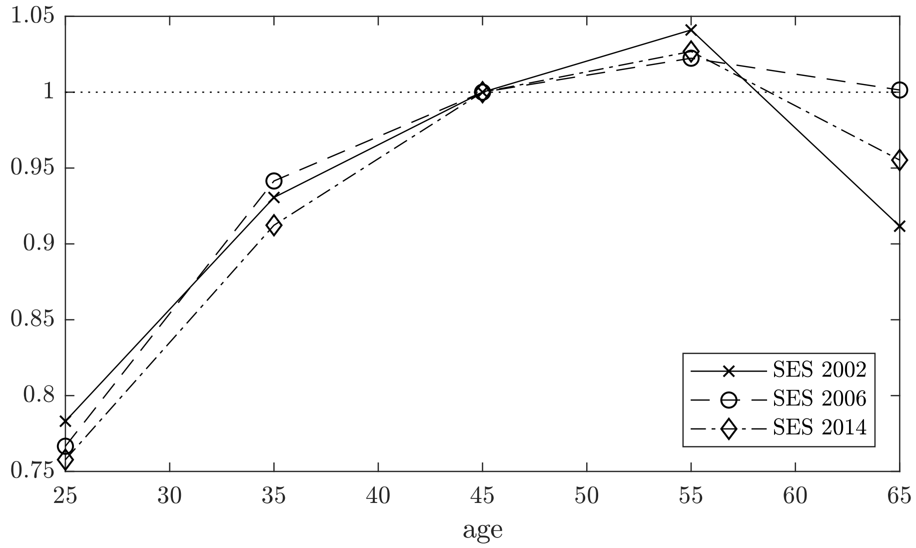

Figure C.4: hourly wage by age relative to age group 40–50, dependent employed males in the private sector in Austria, source: SES waves 2002, 2006, 2014 (Statistik Austria, 2006, 2009, 2017)