Productivity Shocks, Long-Term Contracts and Earnings Dynamics[*]

Neele Balke[†], Thibaut Lamadon[‡]

September 26, 2020

Abstract

This paper examines how employer- and worker-specific productivity shocks transmit to earnings and employment in an economy with search frictions and firm commitment. We develop an equilibrium search model with worker and firm shocks and characterize the optimal contract offered by competing firms to attract and retain workers. In equilibrium, risk-neutral firms provide only partial insurance against shocks to risk-averse workers and offer contingent contracts, where payments are backloaded in good times and frontloaded in bad times. We prove that there exists a unique spot target wage, which serves as an attraction point for smooth wage adjustments. The structural model is estimated on matched employer-employee data from Sweden. The estimates indicate that firms absorb persistent worker and firm shocks, with respective passthrough values of 27 and 11%, but price permanent worker differences, a large contributor (32%) to variations in wages. A large share of the earnings growth variance can be attributed to job mobility, which interacts with productivity shocks. We evaluate the effects of redistributive policies and find that almost 40% of government-provided insurance is undone by crowding out firm-provided insurance.

1 Introduction

Firms play an influential role in determining the level of earnings and employment risk for their workers when designing job contracts (Knight (1921)). Empirical work suggests that they only partially insure their workforce against productivity shocks (Guiso, Pistaferri, and Schivardi (2005)). Understanding wage setting and contract formation is then central to key questions in economics, including the sources of earnings uncertainty, the implications of job mobility and the effectiveness of government insurance.

The theory of dynamic contracts provides a foundation for the use of long-term agreements between firms and workers (Harris and Holmstrom (1982), Thomas and Worrall (1988), Holmstrom and Milgrom (1991)). However, combining optimal contracts with equilibrium models of job mobility with search frictions is challenging, both theoretically and empirically. Such contracts do not always admit closed-form solutions (Abreu, Pearce, and Stacchetti (1986)), and history dependence makes identification difficult. In fact, the empirical literature on earnings and employment dynamics is often silent about whether and how firms endogenously set the level of risk for their workers.

In this paper, we characterize the optimal contract in an equilibrium search model with risk-averse workers who lack the ability to commit, and both individual and firm-level productivity shocks. We establish a tractable solution and conditions for the non-parametric identification of the productivity processes. The model is parameterized and estimated on Swedish matched employer-employee data. Using the estimates, we first examine different insurance channels and how they interact with incentives for workers. We then quantify the relative variance contribution of different sources of uncertainty to wages and how large the passthrough of productivity shocks to workers is. Finally, we measure the effectiveness of government policies aimed at reducing

earnings uncertainty by redistributing from high to low earners.

We develop a structural model with search frictions and one-sided lack of commitment on the worker side. Search is directed and risk-neutral, ex-ante homogeneous firms post dynamic contracts to attract workers in different submarkets.[1] Workers are heterogeneous in skills and risk-averse. They search in the submarket that offers the best trade-off between expected utility and job-finding probability. While firms can commit, workers cannot. In particular, firms are able to credibly promise to insure their workers even when making future losses. In contrast, the choices workers make over which market to search on the job and how much effort to put into retaining a job are unobservable and non-contractible (Shi (2009), Tsuyuhara (2016)). This creates a monitoring problem with decision margins on transitions to both unemployment and other jobs. As a result, dynamic contracts optimally balance the incentives for search and effort with the provision of insurance against shocks. In this model, firms flexibly specify wages for each future productivity path, rather than setting fixed piece rate wages or Nash bargaining contracts, which impose shock transmission by construction. The model thus offers a framework for studying the passthrough of individual- and firm-level productivity shocks to wages and mobility decisions.

We prove the existence of an equilibrium and show that the optimal contract features backloading of wages at the time of hiring (like in Stevens (2004), Burdett and Coles (2003), Shi (2009), Tsuyuhara (2016)) as well as frontloading of wages in bad times (similar to the optimal unemployment benefits in Hopenhayn and Nicolini (1997)). Workers do not internalize the full future surplus of the match when choosing how much effort to exert and where to search. The firm thus chooses to "tilt" the wage profile to appropriately incentivize the

---

[1]The pioneering work in directed search is due to Montgomery (1991), Peters (1991), Moen (1997), Shimer (2005), Burdett, Shi, and Wright (2001) as well as Menzio and Shi (2010), which this paper builds on.

worker. We prove that there exists a unique spot target wage, which serves as the attraction point for these smooth wage adjustments. The existence of the target wage has several implications. First, since the target wage can lie below the current wage level in equilibrium, wage cuts occur on the job despite firm commitment. Second, since the target wage is a function of worker and firm productivity, wages respond to both worker- and firm-specific shocks, in keeping with the empirical literature (Guiso, Pistaferri, and Schivardi (2005), Friedrich, Laun, Meghir, and Pistaferri (2019), Card, Cardoso, Heining, and Kline (2018)). This is an important departure from a competitive market, where earnings equal marginal productivity and do not respond to firm-level shocks.

We develop conditions for the identification of the productivity processes and the production function. Independence between worker and firm shocks enables us to separate them using coworkers' earnings, and the Markov property of the contract allows us to rely on results for hidden Markov chains (Hu and Shum (2012)). Solving for the optimal contract is challenging. To achieve tractability, we not only use promised utility as an additional state (Spear and Srivastava (1987)), but also show that it is sufficient to solve for a single promised marginal utility across all realizations of future productivity (Marcet and Marimon (2019)). The model is estimated in general equilibrium on administrative employer-employee data from Sweden. Using simulated methods of moments, we match transition rates, earnings dynamics and the comovement of firm output with both worker earnings and mobility.

Our impulse response analysis reveals that the contract provides partial insurance, especially against negative shocks, and strongly relies on job mobility to stabilize wages after negative shocks to firm productivity. We compare the baseline model to a first best contract, where firms offer perfectly flat wages, which highlights the fact that wage backloading in good times only happens in the presence of commitment and incentive issues.

To quantify different sources of wage uncertainty, we decompose the variance of earnings and earnings growth into structural components. We find that the time-varying worker and firm shocks contribute much more to variability in output than earnings, suggesting that firms provide substantial insurance to workers. In contrast, firms barely attenuate permanent productivity differences between workers, the largest contemporaneous contributor (32%) to variation in wages across workers. We also find that most (86%) of the variance in earnings growth is associated with worker transitions in and out of employment, which interact with productivity shocks.

We expand on the conventional definition of passthrough to account for three additional factors. First, we consider persistent shocks spanning a long time horizon in contrast to strictly transitory or fully permanent shocks. Second, since shocks alter job transitions, we account for the effects on workers after they change jobs. Third, in addition to looking at wage effects, we extend our analysis to consider the impacts on the lifetime utility of the worker. Based on our preferred specification, we estimate passthrough values of 27 and 11% in response to worker- and firm-level productivity shocks, respectively.

Finally, we look at how efforts by the government to provide public insurance to workers are eroded by firms, who optimally raise the passthrough of shocks to earnings in response to such policies. A revenue-neutral tax policy that redistributes income from high to low earners achieves only 60% of its direct insurance effect because it crowds out insurance provision by the firm. These findings demonstrate how accounting for firms' decisions to provide insurance can have important implications for the anticipated efficacy of public policies.

Related literature. There is an important theoretical literature on long-term contracts between firms and workers. Baily (1974) and Azariadis (1975) study long-term contracts with commitment and develop the insurance role of the firm. Harris and Holmstrom (1982) derive the optimal contract when workers

4

cannot commit and find that positive shocks pass through to worker earnings. Thomas and Worrall (1988) extend this work by looking at lack of commitment by both workers and firms in the presence of rents. MacLeod and Malcomson (1989) develop the implications of non-verifiable output. Building on this work, we include search frictions into the analysis of long-term contracts.

In the frictional search literature, Stevens (2004), Burdett and Coles (2003) and Shi (2009) derive optimal wage-tenure contracts and show the presence of backloading in the absence of shocks. Tsuyuhara (2016) introduces effort to control job destruction and derives similar conclusions. Lentz (2013) develops the optimal wage-tenure contract when firms can respond to outside offers. Menzio and Shi (2010) introduce aggregate and match shocks and develop the block-recursive properties of directed search equilibria, i.e. that the aggregate distribution does not enter the state space of the firm. Our paper complements theirs by theoretically and empirically characterizing the link between wages and productivity. Schaal (2017) develops the incentive-compatible contract in a directed search environment with risk-neutral workers. Rudanko (2009) derives and evaluates the optimal contract with two-sided lack of commitment and aggregate shocks in a model without on-the-job search or private actions. Ábrahám, Alvarez-Parra, and Forstner (2017) study a contract with moral hazard in production and its implications for cross-sectional wage dispersion. To our knowledge, the current paper is the first to characterize the optimal long-term contract offered in equilibrium by firms in an economy with search frictions, on-the-job search, firm and worker shocks and risk-averse workers.

This paper contributes to the growing empirical literature analyzing how firm shocks are transmitted to worker earnings. Guiso, Pistaferri, and Schivardi (2005) are the first to document the wage effects of permanent and transitory firm shocks, using Italian data. Carlsson, Messina, and Nordström Skans (2014) find similar estimates using detailed information on product prices in Sweden.

In Friedrich, Laun, Meghir, and Pistaferri (2019), the authors estimate a model with exogenous earnings dynamics, including firm-level shocks, and endogenous mobility decisions. Roys (2016) uses firm-level data to study the link between firm shocks, wage bills, and employment in a model where wages are set by Nash bargaining. We extend this literature by endogenizing the level of insurance that firms offer workers.

Our analysis fits into a wider empirical literature that uses search models to understand wage dispersion and mobility. Postel-Vinay and Robin (2004), Bagger, Fontaine, Postel-Vinay, and Robin (2014) and Lise, Meghir, and Robin (2016) estimate models of earnings with risk-neutral workers and sequential contracting. We extend this literature and use matched employer-employee data to estimate a search model with optimal contracting and risk-averse workers.

Two papers have empirically assessed the presence of optimal contracts. Chiappori, Salanie, and Valentin (1999) directly investigate the presence of downward rigidity in wages over time, while Lemieux, MacLeod, and Parent (2009) study the presence of performance pay contracts and their contribution to the rise in earnings inequality in the US. We complement this work by introducing search frictions.

Finally, this paper also relates to the large empirical literature on earnings and employment dynamics. MaCurdy (1982), Abowd and Card (1989), Meghir and Pistaferri (2004) and Altonji, Smith, and Vidangos (2013) document the importance of persistent shocks in the process of earnings and employment using longitudinal data. Hall and Mishkin (1980), Blundell, Pistaferri, and Preston (2008) and Low, Meghir, and Pistaferri (2010) evaluate how income shocks affect consumption and how these effects are mitigated by government transfers. In this paper, we extend this work by showing how the earnings process itself might change due to optimal responses by firms in contexts where governments insure workers against earnings risks.

Outline. Section 2 presents the equilibrium search model and Section 3 characterizes the optimal contract. In Section 4, we describe the model implementation, the data, moment selection and the estimation strategy. In Section 5, we analyze the results and report the effects of a redistributive tax policy.

2 The contracting model

This section presents an equilibrium search model that examines the extent to which risk-neutral firms use employment contracts to provide partial insurance for risk-averse workers in the presence of firm-level and worker-specific productivity shocks. In this model, workers can search on the job, search is directed, and firms compete to create vacancies and offer dynamic contracts to attract applicants. While firms can commit to contracts, workers cannot. In particular, workers' choices about which market to search and the level of effort to put into retaining their jobs are unobservable and hence non-contractible.

2.1 Environment

Agents and preferences. Time is discrete, indexed by t and continues forever. The economy is composed of a discrete uniform distribution of infinitely-lived workers with ability indexed by $x_t \in \mathbb{X}=\{x_1, x_2...x_{n_x}\}$ and evolving over time according to a Markov process. Workers enjoy utility over consumption, captured by the increasing and concave utility function $u : \mathbb{R} \to \mathbb{R}$, and pay a utility cost of effort $c : \mathbb{R} \to \mathbb{R}$, where $c(0)=0, c' \in [0, \bar{c}'], c'(0)=0, c''>0$. Lacking access to asset markets, employed workers consume their wages w_t and unemployed workers receive unemployment benefits $b(x_t)$.[2] Workers maximize the expected sum of utility from consumption less effort, discounted at a factor β.

---

[2]Although introducing worker savings would be an interesting additional feature, it drastically complicates the problem, both theoretically and numerically. Introducing hidden savings is an active area of research in principal-agent environments (see e.g. Ábrahám and Pavoni (2008), Attanasio and Pavoni (2011)).

The employer side of the market is composed of a uniform distribution of ex-ante identical firms indexed by $j \in [0,1]$. Firms produce output, pay wages and post job vacancies. A given firm employs a mass of workers $l_{jt}(x_t, z_t)$, where each worker has her own productivity x_t and job-specific match quality z_t. The firm produces total output Y_{jt} using a constant-returns-to-scale technology:

$$Y_{jt} = \sum_x \sum_z l_{jt}(x,z) f(x,z), \tag{1}$$

where $f(x_t, z_t)$ is the output of a worker of type x_t in a match of quality z_t.

The discrete match quality $z_t \in \mathbb{Z}$ evolves over time according to a monotonic transition rule $z_{t+1}=g(z_t, \nu_{jt})$ and is governed by firm-level productivity shocks ν_{jt}. In every period, a firm experiences only one firm-wide productivity shock ν_{jt} but has a distribution of match qualities and worker types. At each x_t and z_t the firm pays a distribution of wages w_t, which is set dynamically.

Each firm chooses how many vacancies to post at a unit cost k. The vacancies are viable for one period and become active jobs if they match with a worker. New hires start with a match quality $z_{t+1}=g(z_0, \nu_{jt})$ for some fixed z_0.[3] The objective of the firm is to maximize the present value of profits.

Labor market. The matching process between workers and firm vacancies is constrained by search frictions. The labor market is organized in a set of sub-markets indexed by $(x_t, v_t) \in \mathbb{X} \times \mathbb{V}$, where $v_t \in \mathbb{V}=[\underline{v}, \overline{v}]$ is the value promised to the worker in that submarket. The promised value v_t equals the expected lifetime utility for a worker of type x_t who matches with a firm in submarket (x_t, v_t). Firms choose the submarkets (x_t, v_t) where they open vacancies and workers direct their search to any v_t queue associated with their own type x_t.[4]

---

[3] A fixed z_0 guarantees that firms are ex-ante identical, which simplifies the solution.

[4] Theorem 3 in Menzio and Shi (2011) reveals that workers will separate by type in equilibrium if markets are indexed by the value that each type x_t gets in a particular submarket $\mathbf{v}=(v(x_1), v(x_2)...v(x_{n_x})) \in \mathbb{R}^{n_x}$, and workers can apply to any submarket. In equilibrium only a given type x_t visits a particular market. This market can then be represented directly by (x_t, v_t), as done in the current paper.

Figure 1: Within-period time line
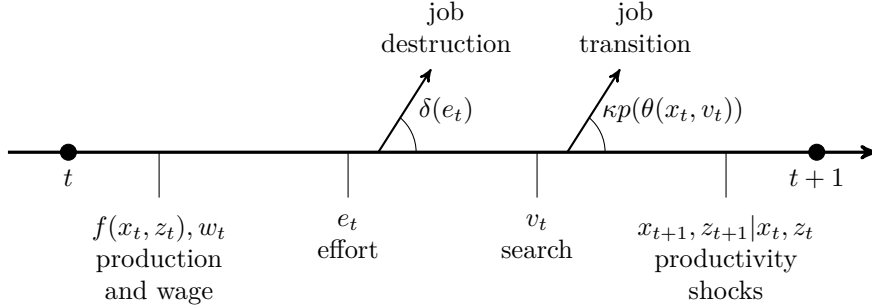
Within each queue the matching between firms and workers is random.

Each submarket displays a tightness represented by the function $\theta : \mathbb{X} \times \mathbb{V} \to \mathbb{R}_+$, which is the ratio of the number of vacancies to the number of workers. Tightness is queue-specific, which means that in queues with a high ratio of vacancies to workers it is harder for firms to hire, and different worker types find jobs at different rates. In queue (x_t, v_t) a worker of type x_t matches with probability $p(\theta(x_t, v_t))$ and a vacancy is filled with probability $q(\theta(x_t, v_t))$. The job finding probability $p(\theta) \in [0,1]$ is assumed to be twice continuously differentiable, strictly increasing and strictly concave, where $p(0)=0$ and $p'(0)<\infty$. Similarly, $q(\theta) \in [0,1]$ is a twice continuously differentiable, strictly decreasing and strictly convex function such that $q(\theta)=p(\theta)/\theta$, $q(0)=1$ and $p(q^{-1}(\cdot))$ is concave. When choosing which submarket to participate in, both firms and workers take into account the value as well as the probability of matching.

Timing of actions.   Each period is divided into four stages as illustrated in Figure 1. First, production takes place at given productivity levels x_t and z_t. For each match, the firm collects output and pays a wage w_t to the worker. Second, workers choose their effort level e_t by balancing effort costs $c(e_t)$ and job destruction probability $\delta(e_t)$, where $\delta(0)=1$, $\delta' \in [\underline{\delta}', 0)$ and $\delta'' \leq 0$. There is a decreasing probability of job loss in effort $\delta(e_t)$, which captures the idea that a negligent worker might lose a client or break a machine and so is more likely

to be fired, and that she can choose to become unemployed at zero cost since $c(0)=0$. Although the worker chooses $\delta(e_t)$, the firm can also induce separation by promising a sufficiently low future lifetime utility such that the worker quits. Third, in the search stage workers choose which submarket (x_t, v_t) to visit and get matched with probability $\kappa p(\theta(x_t, v_t))$, where $\kappa \in [0, 1]$ denotes their on-the-job search efficiency. If matched, the worker moves to a new job and the current job is destroyed. If they do not match, the job continues to the fourth stage, in which a new x_{t+1} is realized and the firm shock ν_t updates the match quality z_{t+1}.

Informational structure and contracts. A contract defines the wage and actions for a matched worker and firm for all future histories. Call $s_t \equiv (x_t, z_t) \in \mathbb{S} = \mathbb{X} \times \mathbb{Z}$ the productivity tuple of a match in period t and call $s^\tau \equiv (s_1...s_\tau) \in \mathbb{S}^\tau$ a given history of realizations between today and τ periods in the future. The future history of productivity for a match is common knowledge to both the worker and the firm and so is fully contractible. However, the worker's effort and search decisions are private information. The contract $\mathcal{C}$ offered by the firm to the worker is then represented by:

$$\mathcal{C} \equiv (\mathbf{w}, \boldsymbol{\zeta}) \quad \text{where} \quad \mathbf{w} \equiv \{w_\tau(s^\tau)\}_{\tau=0}^\infty; \quad \boldsymbol{\zeta} \equiv \{v_\tau(s^\tau), e_\tau(s^\tau)\}_{\tau=0}^\infty. \quad (2)$$

The first component of the contract $\mathbf{w}$ captures the firm's wage policy for each future history $w_\tau(s^\tau)$. The second component $\boldsymbol{\zeta}$ comprises the worker's responses and can be thought of as the unenforceable actions suggested by the contract. Specifically, it includes the effort level $e_\tau(s^\tau)$ and the submarket the employee applies to on the job $v_\tau(s^\tau)$ for each future history of productivity.[5] Although both of these actions are unobserved by the firm, we focus on contracts where the contractual recommendations are incentive-compatible. The

---

[5]Derivations will require that contracts specify simple probabilities over actions rather than actions themselves. This is left implicit at this point but will be clarified in the recursive formulation of the problem.

firm thus chooses both wages and workers' actions, subject to the incentive compatibility constraint that the actions match the workers' optimal responses.

Importantly, the contract space is completely flexible in how wages respond to tenure and productivity histories. In particular, unlike piece rate contracts or Nash bargaining, it does not impose any restrictions on how firms set wages in response to productivity shocks, which is the central question of this paper. Since firms choose the amount of passthrough of a shock endogenously, the model helps to rationalize the firms' contribution to workers' uncertainty about their earnings. By taking the model to the data, we are able to quantify the level of risk absorption by the firm and the effectiveness of government-provided insurance in mitigating the risk to workers.

2.2 Worker's problem

Consider an unemployed worker of type x_t who receives benefits $b(x_t)$ in the current period and wants to search on the submarket that offers the best trade-off between promised future utility and job finding probability. Dropping all time subscripts and focusing on a stationary equilibrium, the value of being unemployed $U(x)$ can be written as:

$$U(x) = \sup_{v_0} u(b(x)) + \beta p(\theta(x, v_0))v_0 + \beta \left(1 - p(\theta(x, v_0))\right) \mathbb{E}_{x'}[U(x')|x], \quad \text{(BE-U)}$$

where v_0 denotes the expected lifetime utility in the submarket chosen by the unemployed worker.

Next, let an employed worker be in a job with productivity levels (x, z) and a current promised value V. Suppose the firm pays a wage w in this period and promises an expected lifetime utility W from tomorrow into the future. Then the employed worker faces the following problem when making her current search

and effort decisions (v_1, e):

$$\max_{v_1, e} u(w) - c(e) + \beta\delta(e)\mathbb{E}_{x'}[U(x')|x] + \beta(1 - \delta(e))\kappa p(\theta(x, v_1))v_1$$

$$+ \beta(1 - \delta(e))(1 - \kappa p(\theta(x, v_1)))W. \quad \text{(EQ-W)}$$

The optimal worker policies $v_1^* : \mathbb{X} \times \mathbb{V} \to [\underline{v}, \overline{v}]$ and $e^* : \mathbb{X} \times \mathbb{V} \to [0, \overline{e}]$ depend on the worker's current x and the promised expected utility for next period W, but not on the match quality z, the wage w or the current promised value V.

Definition 1. *We define the composite retention probability $\tilde{p} : \mathbb{X} \times \mathbb{V} \to [0, 1]$ and the utility return to the worker $\tilde{r} : \mathbb{X} \times \mathbb{V} \to \mathbb{R}$ as functions of x and the promised utility W using short-hand $e^* = e^*(x, W)$ and $v_1^* = v_1^*(x, W)$:*

$$\tilde{p}(x, W) \equiv (1 - \delta(e^*))(1 - \kappa p(\theta(x, v_1^*)))$$

$$\tilde{r}(x, W) \equiv - c(e^*) + \beta(1 - \delta(e^*))\kappa p(\theta(x, v_1^*))(v_1^* - W)$$

$$+ \beta\delta(e^*)\mathbb{E}_{x'}[U(x')|x] + \beta(1 - \delta(e^*))W.$$

The functions $\tilde{p}(x, W)$ and $\tilde{r}(x, W)$ capture how a firm can incentivize a worker when setting the wage dynamically.

2.3 Firm's problem

Setting up the firm's problem involves resolving two difficulties. First, each firm has infinitely many employment relationships that need to be optimized. This can be addressed by relying on a constant-returns-to-scale technology combined with linear vacancy costs, which allows us to separate the firm's problem into individual jobs. Second, the contract a firm offers to a new worker specifies the wage for every future sequence of productivity shocks. Solving this problem for all possible contingencies is intractable because the dimensionality grows to infinity. However, following Spear and Srivastava (1987), we can rewrite the firm's problem recursively by augmenting the productivity state space (x, z) with the previously promised utility to the worker V.

The firm's expected profit from a match $J(x, z, V)$ can be expressed recursively as:

$$J(x, z, V) = \max_{\pi_i, w_i, W_i, W_{ix'z'}} \sum_{i=1,2} \pi_i \Big(f(x, z) - w_i + \beta \tilde{p}(x, W_i) \mathbb{E}_{x'z'}[J(x', z', W_{ix'z'})|x, z] \Big)$$

$$s.t \quad V \leq \sum_{i=1,2} \pi_i \left(u(w_i) + \tilde{r}(x, W_i) \right), \quad \text{(BE-F)}$$

$$W_i = \mathbb{E}_{x'z'}[W_{ix'z'}|x, z], \quad \sum_{i=1,2} \pi_i = 1.$$

At each state (x, z, V), the firm chooses the randomization $\pi_i : \mathbb{S} \times \mathbb{V} \to [0, 1]$ over a two-point lottery, whose only purpose is to ensure concavity of the expected profit function. Although the underlying utility and productivity functions are concave, incentive constraints can, in general, create additional gains from randomization for firms and workers because the allocation space may become non-convex. Lotteries are a standard way to convexify the feasible set and so retain concavity of $J(x, z, V)$, which is useful to establish the existence of an equilibrium (Prescott and Townsend (1984)). All other choices, including the wage $w_i : \mathbb{S} \times \mathbb{V} \to \mathbb{R}_+$, promised future expected utilities $W_i : \mathbb{S} \times \mathbb{V} \to \mathbb{V}$ and productivity-specific promised utilities $W_{ix'z'} : \mathbb{S} \times \mathbb{V} \to \mathbb{V}$, are made contingent on the lottery realization i at the beginning of each period. The values $W_{ix'z'}$ reflect that we augmented the state space with V. These values enter the continuation value and are set optimally contingent on the realized productivity in the next period (x', z').

When choosing the control variables to maximize expected returns, the firm has to respect previous commitments and incentive compatibility. The promise-keeping constraint ensures that the firm honors last period's promise to deliver the value V to the worker (see the left hand side of the first constraint). Incentive compatibility of the worker is embedded in $\tilde{r}(\cdot, \cdot)$ and $\tilde{p}(\cdot, \cdot)$.

Next, consider the return to opening a vacancy in market (x, v), given va-

13

cancy creation cost k and initial match quality z_0:

$$\Pi(x, v) = \max_{W_{0x'z'}} q(\theta(x, v))\mathbb{E}_{x'z'}[J(x', z', W_{0x'z'})|x, z_0] - k$$

$$\text{s.t.} \quad v = \mathbb{E}_{x'z'}[W_{0x'z'}|x, z_0]. \tag{BE-V}$$

Finally, firms open vacancies in a given market if and only if expected returns to the vacancy are positive.

2.4 Equilibrium

Free entry. We impose a free entry condition on each submarket. Free entry together with linear vacancy costs imply that as long as the return to a vacancy $\Pi(x, v)$ is positive, firms post more vacancies to make profits. However, the more vacancies are posted in a market, the tighter the market gets, lowering the vacancy filling rate and consequently each vacancy's expected value. This means that firms open vacancies in each market until expected profits become non-positive:

$$\forall(x, v) \in \mathbb{X} \times \mathbb{V}: \quad \Pi(x, v) \leq 0. \tag{EQ1}$$

This condition pins down the equilibrium tightness $\theta(x, v)$ and the mass of vacancies $\phi(x, v)$ in each active submarket (x, v). We consider a symmetric equilibrium where all firms open an identical number of vacancies in each queue.

Market clearing. Let $h(x, z, V)$ be the stationary distribution of workers with productivity x in matches of quality z with promised utility V. Let $\mu(x)$ represent the mass of unemployed workers of type x. The distributions $h(x, z, V)$, $\mu(x)$ and $\phi(x, v)$ represent the equilibrium allocation in the labor market. In order for labor markets to clear, these equilibrium distributions must be generated by equilibrium decisions.

To derive the law of motion for the stationary distribution $h(x, z, V)$ in the economy, we initially focus on the firm-side problem. A firm is large in the sense

14

that it hires a continuum of workers, and at the same time it is infinitesimal in the economy because it cannot hire a significant share of the entire work force. For each firm, the distribution of match quality, worker types and promised values is pinned down by its entire history of shocks ν^∞. The infinite shock history matters because a shock in one period can alter the distribution of jobs at the firm level beyond that period. For instance, if a firm experienced a very bad shock last period, this will affect the distribution of workers beyond current productivity levels. However, conditional on the history ν^∞, the identity of the firm is redundant. In other words, the notion of a firm in this context is captured by a shared history of shocks. Hence, it is sufficient to characterize the joint distribution over (x, z, V) at the firm level, denoted by $h(x, z, V | \nu^\infty)$, where ν^∞ is exogenous. The law of motion for this distribution is:

$$h\big(x, z, V | [\nu; \nu^\infty]\big) = \int_{V'} \sum_{x'} \phi(x', V') q(\theta(x', V')) P(x|x') \times$$

$$\mathbf{1}\{z = g(z_0, \nu)\} \mathbf{1}\{V = W_{0xz}(x', z_0, V')\} \, \mathrm{d}V'$$

$$+ \int_{V'} \sum_{x'} \sum_{z'} \sum_{i=1,2} \pi_i(x', z', V') \tilde{p}(x', W_i(x', z', V')) h(x', z', V' | \nu^\infty) P(x|x') \times$$

$$\mathbf{1}\{z = g(z', \nu)\} \mathbf{1}\{V = W_{ixz}(x', z', V')\} \, \mathrm{d}V'.$$

Importantly, conditional on the shock history, it is unnecessary to keep track of which worker was in which firm, a direct result of having search markets. Since firms are large and symmetric, the overall state of the economy can be characterized by the integral of the joint distribution over shock histories at the firm level ν^∞. This pins down the aggregate stationary distribution $h(x, z, V)$:

$$h(x, z, V) = \int h(x, z, V | \nu^\infty) P(\nu^\infty) \, \mathrm{d}\nu^\infty. \qquad \text{(EQ2)}$$

Next, the mass of unemployed workers $\mu(x)$ evolves according to:

$$\mu(x') = \sum_x \mu(x)P(x'|x)[1 - p(\theta(x, v_0^*(x)))]+$$

$$\int_V \sum_x \sum_z \sum_{i=1,2} (1 - \mu(x))P(x'|x)[1 - \tilde{p}(x, W_i(x, z, V))]h(x, z, V) \, \mathrm{d}V. \quad \text{(EQ3)}$$

Finally, the vacancy mass $\phi(x, v)$ must satisfy $\forall (x, v)$:

$$\phi(x, v) = \theta(x, v)\Big[\mu(x)\mathbf{1}\{v = v_0^*(x)\}$$

$$+ \kappa \int_V \sum_z \sum_{i=1,2} \pi_i(x, z, V)\mathbf{1}\{v = v_{1i}^*(x, W_i(x, z, V))\}h(x, z, V) \, \mathrm{d}V\Big]. \quad \text{(EQ4)}$$

Based on these conditions, we define the equilibrium.

Definition 2. *A **recursive search equilibrium** consists of firm value functions $J(x, z, V)$ and $\Pi(x, V)$, an unemployment value function $U(x)$, a job retention probability $\tilde{p}(x, W_i)$, a worker return function $\tilde{r}(x, W_i)$, optimal contract policy functions $\xi = \{\pi_i, w_i, e_i, v_{1i}, v_0, W_i, W_{ix'z'}, W_{0x'z'}\}_{i=1,2}$, a market tightness function $\theta(x, v) \in \mathbb{R}_+$, an active aggregate job distribution $h(x, z, V)$, a mass of unemployed workers $\mu(x)$ over types x, and a mass of vacancies $\phi(x, v)$ across submarkets (x, v), such that:*

(a) $J(x, z, V), \Pi(x, V)$, and $U(x)$ satisfy BE-F, BE-V and BE-U for all (x, z, V),

(b) $\tilde{p}(x, W_i(x, z, V))$ and $\tilde{r}(x, W_i(x, z, V))$ satisfy EQ-W for all (x, z, V),

(c) ξ contains the associated policy functions,

(d) $\theta(x, v)$ and $\phi(x, v)$ satisfy the free entry condition EQ1 for all (x, v),

(e) $h(x, z, V)$ is generated by $\phi(x, v), \mu(x)$ and ξ as in EQ2, and

(f) $\mu(x)$ and $\phi(x, v)$ clear the market in accordance with EQ3 and EQ4.

In equilibrium, neither workers nor firms have an incentive to deviate from their contracts. Solving for an equilibrium, which involves large-dimensional distributions of workers over employment states and queues, is generally very difficult. The approach we adopt relies on the properties of equilibrium, es-

pecially the fact that the unique tightness function $\theta(x, v)$ is not a function of the aggregate state of the economy.[6] Intuitively, since the labor market is divided into submarkets, the tightness in each submarket does not depend on the distribution of workers over other submarkets. As a result, given $\theta(x, v)$ in a specific submarket, a worker who queues in that submarket has a probability of finding a job that is independent of other markets. Consequently, a worker's return from searching and retention probability are independent of the distribution and hence the firm problem is distribution-free, i.e. the Bellman equations do not include $h(x, z, V)$ as an additional state. This allows for a two-step computational procedure. In a first step, one can solve for the value functions, which depend only on (x, z, V); and in a second step, simulate them in order to recover the associated distributions.

The separation of expected profits from aggregate distributions also ensures the tractability of block recursive equilibria (Menzio and Shi (2010)). Building on the work of these authors and of Tsuyuhara (2016), we can formally prove the existence of equilibrium.

Proposition 1 (Existence of Equilibrium). *There exists a recursive search equilibrium, in which the expected profit function $J(x, z, V)$ is continuously differentiable as well as strictly decreasing and concave with respect to V.*

Proof. See Appendix A1.1. □

3 Contract characterization

3.1 Incentive vs. insurance trade-off

A worker is less likely to leave a firm when she is promised a higher future value. Intuitively, the worker searches submarkets with a higher v as the expected future utility in her current job increases. However, submarkets that offer a

---

[6]All equilibrium properties are derived in the supplementary Appendix W1.1.

higher lifetime value have a lower job finding rate and so the worker is less likely to find a new match. This is summarized in Lemma 1:

Lemma 1 (Pareto Properties). *Given* (x, V, W_i),

(a) $v_1^*(x, W_i)$ *and* $e^*(x, W_i)$ *are uniquely determined,*

(b) $\tilde{p}(x, W_i)$ *is continuous, differentiable a.e. and increasing in* W_i,

(c) $\tilde{r}(x, W_i)$ *is increasing and differentiable a.e. in* W_i *and* $\tilde{r}'(x, W_i) = \beta \tilde{p}(x, W_i)$,

(d) *expected profits* $J(x, z, V)$ *are increasing in match quality* z.

Proof. See Appendix A1.2. □

Lemma 1 sheds light on the firm's trade-off between providing dynamic incentives and insurance to the worker when choosing promised values and wages. By increasing future promised values, a firm incentivizes its workers to exert more effort, and so increases the probability that each match continues to exist. Dynamically incentivizing workers thus makes a case for backloading wages, i.e. for increasing promised values and wages over time. However, higher promised values also lower a firm's profits and worker risk aversion implies that the cheapest way to deliver a given lifetime utility to workers is by keeping the wage constant over time. This intertemporal insurance motive thus calls for a flat wage profile and no backloading.

To understand how firms resolve this trade-off between incentives and insurance, it is useful to analyze how firms decide to compensate workers over time. The following proposition provides a clear prediction for how wages evolve as a function of the current state of the match:

Proposition 2 (Optimal Wage Growth). *For any current state* (x, z, V), *within each lottery realization i, the following relationship between wage growth and expected firm profits holds:*

$$\eta(x, W_i) \cdot \mathbb{E}_{x'z'}[J(x', z', W_{ix'z'})|x, z] = \frac{1}{u'(w_i')} - \frac{1}{u'(w)}, \qquad \text{(FOC)}$$

18

where $\eta(x, W) = \frac{\partial}{\partial W} \log \tilde{p}(x, W) > 0$ *is the derivative of the log probability of retention with respect to the future value promised to the worker W, w is the current wage and w'_i is next period's wage in lottery realization i.*

Proof. See Appendix A1.2. □

The optimal balance between insurance and incentive motives ties together all the features of the model and the FOC allows us to uncover where each of the features enters. For instance, the right hand side only involves worker preferences and reveals how much the worker likes intertemporal wage smoothing. The left hand side reflects the presence of search frictions and decisions about how much effort to exert, as captured by $\eta(x, W)$. One can also directly observe that the optimal contract exhibits full insurance across future states as (x', z') drop out of the FOC. The intertemporal dimension of insurance thus competes against incentives when optimally setting wage contracts. The benefit of intertemporal insurance (the right hand side) grows with the amount of risk aversion, and since the curvature of utility typically decreases in consumption levels, is higher at low wages. The extent of insurance offered by the optimal contract thus depends endogenously on where the worker is in the wage and value distribution. The flip side of smoothing wages is that it lowers the returns to dynamic incentives, as measured by the left hand side. The term $\eta(x, W)$ shows how elastic the retention probability is with respect to future promised values W. Effort contributes to this term through the endogenous job destruction probability $\delta(e^*(x, W))$, and search frictions enter through the optimal job finding rate $p(\theta(x, v_1^*(x, W)))$. While we know that in equilibrium effort is increasing in W and job finding rates are decreasing in W, this does not allow a conclusion about how $\eta(x, W)$ changes with W. The answer depends on the second derivative of the retention probability $\tilde{p}(x, W)$. The qualitative effect of changes in W on the level of insurance therefore remains ambiguous despite the

fact that expected profits decrease in W. However, in the limit without search frictions and effort decisions, $\eta(x,W)$ goes to zero because the endogenous separation probability becomes insensitive to local changes in W. As a result, the model approaches a full insurance economy and the incentive constraint starts resembling a participation constraint, creating a discontinuity where it binds. This means that absent the incentive issues due to search and effort decisions, workers enjoy complete wage smoothing until a better outside option arrives.

The proposition also reveals that the worker's wage growth has the same sign as the expected profit of the firm under the optimal contract. Whenever the firm expects positive profits, it is optimal to increase wages, and whenever profits are expected to be negative, it is optimal to lower wages. This means that across all shock histories, a change in wages is positive if and only if the expected profit for the firm is also positive. This optimal strategy sacrifices some of the worker's utility, since she prefers flat wages, for a higher probability of keeping the match alive. Intuitively, inducing less separation is a way for the firm to benefit longer from a profitable match in a context with workers who value only their own share of the match surplus.

The trade-off between insurance and incentives is thus solved by wage backloading when the firm expects positive profits. To formalize this argument further, we define:

Definition 3. *The **target wage** $w^*(x,z)$ given (x,z) is defined as the wage associated with the promised value $W^*(x,z)$ for which $M(x,z,W^*(x,z))=0$, where*

$$M(x,z,W) = \max_{W_{x'z'}} \mathbb{E}_{x'z'}[J(x',z',W_{x'z'})|x,z]$$

$$s.t \quad W = \mathbb{E}_{x'z'}[W_{x'z'}|x,z],$$

and, with $W^*_{x'z'}(x, z)$ denoting the argmax of $M(x, z, W^*(x, z))$, is equal to:

$$w^*(x, z) = u'^{-1}\left(-\frac{1}{J'(x', z', W^*_{x'z'}(x, z))}\right).$$

The target wage captures the wage level at which expected firm profits are zero and the worker collects the entire value of the match. The optimal contract can be further characterized using this target wage.

Proposition 3 (Target Wage Tracking). *For each viable match (x, z) there exists a unique target wage $w^*(x, z)$, which is increasing in z. The wage specified in the optimal contract tracks this target wage for each lottery realization i:*

$$w \leq w^*(x, z) \quad \Rightarrow \quad w \leq w'_i \leq w^*(x, z) \quad \text{incentive to stay,}$$

$$w \geq w^*(x, z) \quad \Rightarrow \quad w^*(x, z) \leq w'_i \leq w \quad \text{incentive to leave.}$$

Proof. See Appendix A1.2. □

At any state (x, z, V), the wage adjusts towards $w^*(x, z)$. If the current wage w is below the target wage, the firm will increase wages in the following period, but not so much as to exceed $w^*(x, z)$. The prospective wage increase induces the worker to exert more effort to retain her job, leading to a higher level of insurance against job loss. Since the wage will grow, the higher level of job loss insurance is achieved at the expense of having less than full wage insurance. Conversely, if the firm currently pays more than $w^*(x, z)$, it lowers future wages towards the target wage and reduces the probability of retaining the employee. Here, the firm provides less wage insurance by allowing the wage to fall and at the same time less insurance against job loss by incentivizing lower levels of effort. Hence, wage insurance and job loss insurance go in the same or opposite directions depending on the current wage level relative to the target wage. Finally, the wage stays constant when $w=w^*(x, z)$ because the left hand side of the FOC is zero. Intuitively, since the worker collects the entire value of the match, her incentives are aligned with the firm and the

incentive problem vanishes. Only the insurance motive prevails and the contract prescribes constant wages.

3.2   Backloading

Propositions 2 and 3 amount to a formal backloading result at the time of hiring. Although the firm is able to commit, it chooses to "tilt" the wage profile instead of setting a perfectly flat consumption path for the worker even in the absence of shocks. This is optimal because the worker makes effort and search decisions e and v that affect the probability that the match continues to exist. When some of the match surplus goes to the firm, the worker does not internalize the full future value of the match when making these decisions (unless she is at $w^*(x, z)$). It is then optimal for the firm to frontload some profits and backload wages. Importantly, this happens every time a new match forms due to the free entry condition. At that moment, expected profits must be positive to recover vacancy costs and so wages increase in accordance with Proposition 2. The model thus exhibits, on average, an increasing wage profile after every new hire, irrespective of the possibility of wage decreases later on.

Backloading is a well-known property of long-term contracts with lack of commitment on the worker side. Stevens (2004) demonstrated this property in a search environment with risk-neutral agents and a minimum wage constraint. In our paper, workers are risk-averse and so the contract must optimally balance the incentive problem with the desire for consumption smoothing as shown in Burdett and Coles (2003), Shi (2009) and Tsuyuhara (2016).

Another property that hinges on backloading is that workers can experience wage cuts during value-enhancing moves to other firms even if their own productivity is unchanged. As in Postel-Vinay and Robin (2002), skill losses are not necessary to rationalize the empirically prevalent wage cuts after job transitions, because higher promised values can be delivered through backloaded

22

wages. The heterogeneity in match quality z is crucial for this property to arise in our setup. In the absence of shocks, as for instance in Shi (2009), Tsuyuhara (2016) or Burdett and Coles (2003), wages would not go down during J2J transitions. Homogeneous firms share the same profit function, whose concavity in the promised value would dictate that the higher value of the new job must be associated with a higher wage.

Since the firm can only imperfectly monitor the worker's actions, the firm faces an incentive constraint, capturing that the worker chooses a specific hidden action only when appropriately incentivized.[7] Here, the worker's effort and search decisions affect the duration of the match and the availability of outside options. The common role of effort and search in the incentive problem for the firm is emphasized by the fact that they enter the retention probability $\tilde{p}(x, W_i)$ together. These two hidden actions have been studied separately in Menzio and Shi (2010), where directed search is unobservable by the firm, and in Tsuyuhara (2016), where effort affects the job finding probability. However, we believe this paper is the first to combine the two in a study looking at the dynamic transmission of shocks to wages under the optimal contract. We turn to these dynamics next.

3.3   Transmission of productivity shocks

Proposition 3 reveals that the realized wage smoothly tracks the reference wage $w^*(x, z)$, which depends on productivity x and is increasing in match quality z. This implies that wages respond to both positive and negative productivity shocks, from both the firm and worker side. To better illustrate the rich features of this wage setting, Figure 2 draws the qualitative wage path under different contract arrangements in response to positive and negative worker- and firm-level shocks. The solid red line shows the wage dynamics under the optimal

---

[7]This incentive problem resembles the principal-agent problem outlined in Rogerson (1985) where a risk-averse worker chooses a hidden effort level.

Figure 2: Wages and productivity shocks

(a) positive x shock

(b) positive z shock

(c) negative x shock

(d) negative z shock
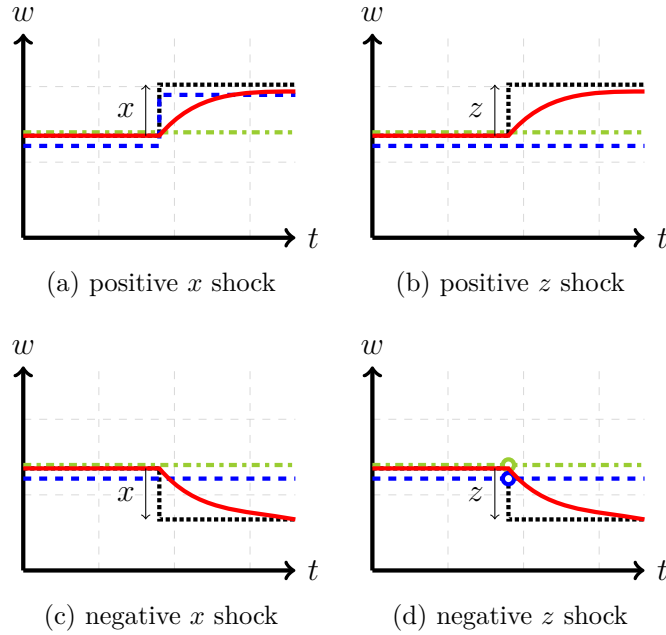
Notes: This figure plots wages under different contracting environments in response to productivity shocks. The dash-dotted green line represents a two-sided commitment contract, the dashed blue line shows firm commitment without imperfect monitoring, and the solid red line captures firm commitment and incentive issues. The dotted black lines depict the productivities.

contract in this paper, where one-sided lack of commitment and moral hazard act as the two key drivers. For comparison, we plot the wage dynamics under the optimal contract against the wage dynamics with these two elements removed. The dashed blue line depicts a similar one-sided limited commitment model but without a monitoring problem, and the dash-dotted green line represents a full commitment contract that also lacks hidden actions.

Commitment. To understand the differences between contract arrangements, it is useful to analyze the commitment problem separately from the moral hazard issue. With regard to the former, the differences between the dash-dotted green line and the dashed blue line are exclusively due to different assumptions on worker commitment. While risk-neutral firms can commit to insure risk-averse

workers in both environments, the worker is free to choose a better outside option only in the model depicted by the dashed blue line.

In the presence of full commitment, firms optimally insure workers against all shocks, resulting in flat dash-dotted green lines in all cases. As in Menzio and Shi (2011), the equilibrium is efficient and retains the social planner's allocation in response to shocks. Despite fluctuations in x, the firm pays the same wage to the worker. The worker commits not to quit at high x, while the firm commits to not dissolve the match at low x. However, when quality z falls, the best way to achieve insurance is through dissolving the match and letting the worker start a new job. Separation, indicated by a circle in Figure 2d, benefits both employers and workers who are able to form new, higher quality z_0 matches. This is in line with Azariadis (1975) who finds that enforceable contracts can lead to an equilibrium in which fluctuations in wages are eliminated but job changes occur.

If instead only the firm is able commit to a contract, the worker cannot commit to ignoring better spot market offers. A positive x shock then improves the worker's outside options, resulting in an immediate pay raise. This is represented by a jump in the dashed blue line following a positive x shock in Figure 2a. In contrast, a negative x shock is fully absorbed by the employer who has committed to an unresponsive wage ex-ante as part of the optimal contract. Since outside offers are worse, the worker's participation constraint is slack and the match continues, as in Harris and Holmstrom (1982). Despite being absent in Harris and Holmstrom (1982), one can extend the one-sided commitment framework with firm-specific shocks to match quality z. A positive z shock exclusively affects the existing match but is inconsequential for wages at other jobs with the same x, so neither employment relationships nor wages change. Finally, when match productivity falls below z_0, the worker moves to another firm where she receives the same compensation as before.

Therefore, commitment plays a crucial role in shaping wage dynamics. As soon as workers cannot commit to forgo better outside options, the firm stops offering full insurance, which differs from the two-sided commitment case. Two further insights become apparent when considering one-sided commitment. First, the fact that in Harris and Holmstrom (1982) wages increase but never decrease might suggest that one-sided commitment generally implies downward wage rigidity. However, a simultaneous drop in x and z can indeed generate a wage decline. If it becomes too expensive for the current firm to perfectly insure an unproductive worker in a low quality match, the firm will cut the wage or induce separation. If the worker moves to a new firm, the match quality will be higher but the new employer will set a lower spot wage due to the worker's lower productivity. Therefore, a combination of negative z and x shock can trigger a wage decline, even when firms exhibit the ability to commit. The downward wage rigidity in Harris and Holmstrom (1982) is precipitated by the assumption that all jobs are identical, while heterogeneity in jobs in our setup allows for decreasing wage paths. Second, it is inefficient to sustain any match with a quality below z_0 if new matches can be formed in frictionless markets. Without separation, the opportunity for higher quality in new matches remains unused by the firm and the worker. However, the worker may not move if she has worse outside options from a simultaneous x drop. The firm is then stuck with a low x worker in a low z match. In this case, a social planner would prefer to terminate the employment relationship to avoid the dead-weight loss of $z < z_0$ associated with inefficient immobility.

Imperfect monitoring. We consider imperfect monitoring, which accounts for the differences between the dashed blue line and the solid red line in Figure 2. Both the blue and red lines represent situations with one-sided commitment, i.e. firms can commit to dynamic contracts whereas workers cannot, but only the solid red line includes a hidden action dimension. Here, we consider unob-

servable effort and search decisions that affect the likelihood of separation. This comparison reveals two core findings. First, hidden actions make wages responsive to both positive and negative shifts in expected profits. In contrast, without moral hazard wages only respond to positive x shocks. Second, conditional on a wage change, the imperfect monitoring model exhibits much smoother wage adjustments.

The responsiveness of wages to all shocks (solid red line) is consistent with the FOC. For instance, in the presence of imperfect monitoring, adverse shocks to z lead to two inefficiencies absent a wage drop. The worker exerts too much effort to keep the less productive match alive, but also searches markets with too high lifetime utility, simply because she does not internalize the reduction in joint surplus. To counteract these inefficiencies, the firm lowers the wage (Figure 2d). The reverse holds true when the worker does not fully internalize the benefits of continuing a high z match. In this situation, the firm adds incentives to maintain the match by increasing wages, whereas in a world without monitoring problem search and effort would be dictated and wages would remain constant (Figure 2b). A similar argument holds when firm profits are affected by x shocks (Figure 2a and Figure 2c).

Differences in the smoothness of wage adjustments in Figure 2a are rooted in how a worker is retained over time. In the blue dashed case, an increase in productivity x increases the worker's outside options, which the firm must match immediately to keep the worker, resulting in a wage jump. With imperfect monitoring, the firm backloads payments to increase the likelihood of retaining the worker, as established above. This happens smoothly due to the worker's risk aversion, and so wages increase more slowly in the model represented by the red solid line. The gradual passthrough hinges on the incentive structure and probabilistic nature of separation, and hence on the presence of search frictions, but is not limited to directed search. Random search, as in Burdett

27

and Coles (2003), also produces smooth wage adjustments because the worker's reservation wage in light of non-contractible outside offers is unobserved.

From a technical point of view, the worker's participation constraint in the model represented by the dashed blue line is replaced by an incentive constraint in the case of the red solid line. Under a participation constraint, a better outside option can be obtained with certainty, so the firm must lift the wage instantly or the worker leaves. In contrast, with an incentive constraint, obtaining the outside option is probabilistic and mobility frictions allow for a slow wage increase. Of course, one can reinterpret a participation constraint as a stark incentive problem. The key is then to restrict the incentive problem to always bind at a corner solution where the worker either leaves or stays with certainty.

One may wonder why a firm would choose to lower wages smoothly in bad times rather than fully renegotiating them as in Thomas and Worrall (1988). This is because it is optimal for a firm to use its commitment power to promise insurance to a worker at the time of hiring. The fact that after a bad shock a firm is ex-post bound to the contract does not mean that commitment is ex-ante undesirable; to the contrary, the firm benefits from its commitment ability. The wage goes down in Thomas and Worrall (1988) because the firm participation constraint binds, whereas we allow the firm to optimize without this additional constraint, which results in an ex-ante preferable outcome.

4 Estimating the contracting model

There are two important challenges when taking the model to the data. The first difficulty concerns identification. Individual productivity is not directly observable, and the way it translates into earnings and labor participation is highly non-linear. The second obstacle is tractability. Solving directly for promised utilities in each future state is infeasible.

Non-parametric identification. We provide sufficient conditions for the

identification of the production function $f(\cdot)$ and the productivity processes for x and z using a five-year panel dataset on workers' and coworkers' earnings and participation. We assume the preference functions $u(\cdot)$ and $c(\cdot)$ are known.[8] First, we show that the independence between coworker trajectories after conditioning on firm-level productivity histories allows us to isolate firm shocks from worker shocks. Next, we use the Markovian properties of the optimal contract to identify the wage and participation process conditional on the latent variables x and z, drawing on the identification result for hidden Markov chains in Hu and Shum (2012). Finally, we show that the present value of the worker, the firm profit function, and the production function can be recovered from the conditional choice probabilities together with the Bellman equation.

Tractability. Solving BE-F directly would require optimizing over the promised utilities $W_{ix'z'}$ for every (x, z, V) in each future state of the world (x', z'). This becomes unmanageable as soon as reasonable supports for $\mathbb{X}$ and $\mathbb{Z}$ are considered. However, the first order condition with respect to $W_{iz'x'}$ reveals that the promised utilities at different future states are chosen optimally by the firm in order to equalize marginal utilities across those states. The solution can thus be characterized by a single promised marginal utility. In Appendix A1.3, we rewrite the contracting problem recursively using marginal utilities[9] and rely on this formulation in the numerical implementation.

4.1 Model specification

Each period corresponds to one quarter and the discount factor β=0.987 is set to match an annual interest rate of 5%. For the estimation, we use the utility function $u(w) = \frac{w^{1-\sigma}-1}{1-\sigma}$, where the parameter of constant relative risk aversion

---

[8]Supplementary Appendix W2 includes the formal results.

[9]This is known as the recursive Lagrangian approach developed by Kocherlakota (1996), Marcet and Marimon (2019), Messner, Pavoni, and Sleet (2012) and Cole and Kubler (2012). Farhi and Werning (2013) uses promised marginal utility directly.

is set to a standard value $\sigma=1.5$. These values are taken from Low, Meghir, and Pistaferri (2010). Work effort controls separation through $\delta(e)=1-e$ and incurs utility cost

$$c(e) = \frac{\gamma_0}{\gamma_1 - 1} + \gamma_0(1 - e) - \frac{\gamma_0}{1 - \frac{1}{\gamma_1}}(1 - e)^{1 - \frac{1}{\gamma_1}},$$

where the condition $c(0)=0$ is satisfied at any values of (γ_0, γ_1). The flow value of being unemployed is set to $u(b)$ for a fixed b.

Let worker productivity x be characterized by two subcomponents: one permanent x_0 and one persistent x_1. The sets of values for these subcomponents, as well as for match quality z, are drawn from discretized log normal distributions with a log mean of zero and log variances $\sigma^2_{x_0}$, $\sigma^2_{x_1}$ and σ^2_z, respectively. We set the initial match quality to its median value $z_0=1$. Since x_0 is fixed, the transition matrix for x is block-diagonal such that each block includes the transition probabilities for x_1. With a reset probability $(1-\lambda_x)$, productivity x_1 is redrawn uniformly from its prior distribution, independently of previous levels of x_1. Otherwise, with probability λ_x, x_1 persists. The transition matrix of z follows a similar Markov structure with persistence parameter λ_z. This determines the law of motion for match quality, $z' = g(z, \nu)$, where $g(z, \nu)$ is the conditional quantile function of the transition matrix and ν is the firm-level shock. The function $g(z, \nu)$ is monotonic by construction and implicitly assumes a uniform distribution of firm shocks ν that allows for correlated match qualities at the firm level. It does not, however, imply that all workers at one firm have the same z, even if they share the same shocks ν.

The output of a match is given by $f(x_0, x_1, z) = a \cdot x_0 \cdot x_1 \cdot z$, where total factor productivity a is normalized such that average utility is one. We employ the matching function $q(\theta) = [\alpha^\varphi/(\alpha^\varphi + \theta^\varphi)]^{\frac{1}{\varphi}}$ and estimate the model in equilibrium such that the job finding rate of workers is derived from the free entry condition EQ1. We use a common estimate of the curvature $\varphi=0.8$, normalize the vacancy

cost to $k=1$, and estimate the matching efficiency $\alpha \in (0,1)$. Alternatively, one could fix α and estimate the vacancy cost k. Finally, we also estimate the on-the-job search efficiency κ.

We use a measure of value added per worker at the firm level.[10] To remain consistent with the theory, we use a constant-returns-to-scale production function when aggregating match outputs. We sum across workers over four consecutive quarters and construct annualized earnings w^a, match output f^a and value added per worker y^a to match the frequency of value added and earnings in the data. Finally, we add measurement error to log earnings and to log value added per worker, with respective standard deviations m_w and m_y. In total, this specification requires estimating the parameter set $\vartheta = \{\gamma_0, \gamma_1, b, \sigma_{x_0}, \sigma_{x_1}, \sigma_z, \lambda_x, \lambda_z, \alpha, \kappa, m_w, m_y\}$.

4.2 Data

The sample is built on employer-employee matched data prepared in Friedrich, Laun, Meghir, and Pistaferri (2019), which links three Swedish administrative data sets: the Longitudinal Database on Education, Income and Employment (LOUISE), which provides data on demographic and socioeconomic variables for the entire working age population; the Register-Based Labor Market Statistics (RAMS), which tracks worker earnings and employment spells and includes firm identifiers; and the Structural Business Statistics (SBS), which includes a measure of value added and employment size for each firm and year.

We use five consecutive years (2001 to 2006) for our analysis. We adjust monetary variables for inflation. On the worker side, all self-employed individuals are removed from the sample. To abstract from labor supply decisions on the extensive margin, we focus on males between 25 and 50 years of age, which

---

[10]Since firms are ex-ante identical, vacancy costs are homogeneous and the equilibrium is symmetric, the model is less suitable for mapping the firm size distribution directly into the data. Therefore, we focus on value added per worker in the estimation.

removes the labor participation decisions of females during child-bearing years and decisions related to retirement. However, we include active and non-active job seekers in order to account for mobility in and out of work. We limit the employer side to firms with a positive value added. Still, our final sample includes almost 1.2 million unique workers and over 70,000 unique firms. Appendix A2 provides further details on the data sources and sample construction.

4.3   Moments of interest

To capture the level of risk sharing between firms and workers together with transition dynamics in the labor market, we select three sets of moments to be matched: transition probabilities, individual earnings dynamics and moments relating firm output to both worker earnings and mobility.

First, using information on monthly spell data, we follow Friedrich, Laun, Meghir, and Pistaferri (2019) and construct quarterly transition probabilities of starting a job (U2E), losing a job (E2U) and changing jobs (J2J). As reported in Table 1, the numbers we obtain are 0.17 for U2E, 0.022 for E2U and 0.026 for J2J rates. The J2J probability is notably lower than in the US and other comparable European countries.

Second, we generate moments of individual earnings dynamics. Yearly earnings data may imprecisely measure the duration of partial employment spells within the year. To alleviate this concern, we only use observations that report 12 months of work. In this group, the overall variance of log earnings in the data is 0.14. Earnings grow by 2.5% on average for all fully employed workers over two consecutive years, and have a variance of 0.025. We extract the autocovariance of log earnings growth from the data at −0.007. Job switchers — individuals who are fully employed in one year, switch in a second year to a new employer and continue working there for the entire third year — experience an average wage growth of 6.4% over these two years. This is over two times the

32

Table 1: Moments and model fit

	data	model
Pr^{U2E}	0.17 (0.00035)	0.16
Pr^{J2J}	0.026 (4.1e-05)	0.026
Pr^{E2U}	0.022 (4.4e-05)	0.020
$\text{Var}_{S^{\text{E}}}\left[\log w_{it}\right]$	0.14 (0.00031)	0.15
$\mathbb{E}_{S^{\text{EE}}}\left[\Delta \log w_{it}\right]$	0.025 (0.00010)	0.028
$\text{Var}_{S^{\text{EE}}}\left[\Delta \log w_{it}\right]$	0.025 (0.00012)	0.024
$\text{Cov}_{S^{\text{EEE}}}\left[\Delta \log w_{it}, \Delta \log w_{it-1}\right]$	-0.0068 (6.4e-05)	-0.0083
$\mathbb{E}_{S^{\text{J2J}}}\left[\log w_{it} - \log w_{it-2}\right]$	0.064 (0.00055)	0.061
$\mathbb{E}_{S^{\text{E}}}\left[\log w_{it}\right] - \mathbb{E}_{S^{\text{U2E}}}\left[\log w_{it}\right]$	0.55 (0.0038)	0.50
$\text{Cov}_{S^{\text{UEUE}}}\left[\log w_{i,\tau_i(1)}, \log w_{i,\tau_i(2)}\right]$	0.092 (0.00035)	0.090
$\text{Var}_{S^{\text{EE}}}\left[\Delta \log y_{it}\right]$	0.098 (0.0070)	0.091
$\text{Cov}_{S^{\text{EEE}}}\left[\Delta \log y_{it}, \Delta \log y_{it-1}\right]$	-0.034 (0.0026)	-0.036
$\text{Cov}_{S^{\text{EE}}}\left[\Delta \log w_{it}, \Delta \log y_{it}\right]$	0.00099 (0.00013)	0.0010
$\text{Cov}_{S^{\text{EE}}}\left[\Delta \log(1 - \tilde{p}_{it}), \Delta \log y_{it}\right]$	-0.013 (0.0040)	-0.014

Notes: Transition probabilities are quarterly, all other quantities are at the yearly level. Subscripts S capture different sets of observations for the corresponding empirical averages. Standard errors are computed using bootstrap. See the supplementary Appendix W3.2 for a full description.

average gain, suggesting a positive earnings return associated with job transitions. To estimate the next two moments, we use unemployment spells. The first moment relates to the mean-to-minimum wage ratio of Hornstein, Krusell, and Violante (2011) and captures the log difference between the average yearly earnings in the entire economy and the average yearly earnings for hires coming

out of unemployment. This gap is 0.55, meaning that the average wage is about twice what new hires earn. The second moment is the within-individual covariance of earnings for workers coming out of unemployment twice, an indicator of their persistent heterogeneity. At 0.092, this covariance is higher than half of the overall variance in earnings, suggesting a strong role for persistent worker heterogeneity in the data.

The final set of moments links firm performance to earnings dynamics and mobility. To start, we retrieve the growth of log value added per worker at the firm level. The cross-sectional growth variance is 0.1, while the first order autocovariance is -0.034. The covariance between the growth rates of value added per worker and the earnings of those who stay with the firm is 0.001. This speaks to one of the core theoretical findings from the analysis of the model: wages move with the performance of the firm.[11] Finally, to understand how firm performance affects the transition probabilities of workers, we include a moment that captures how worker retention varies with output. To this end, we compute the covariance between changes in the log probability of workers leaving the firm and the log value added per worker. We find a value of -0.013, indicating that when firm performance goes up workers are less likely to leave the firm. In total, there are 12 parameters to estimate in order to match 14 moments in the data.

4.4 Model fit

We estimate the model via indirect inference.[12] The model fit is good, as shown in Table 1. All three transition probabilities are matched well. Although the model gives a central role to insurance, it can generate a large variance in log

---

[11]Assuming a unit root process with i.i.d. measurement error, one can combine these three moments to compute a passthrough measure of permanent shocks. This results in a passthrough value of 4.5% which is in keeping with 7.8% in Guiso, Pistaferri, and Schivardi (2005), see Appendix A3.2.

[12]See supplementary Appendix W4 for the estimation procedure as well as a sensitivity analysis following Andrews, Gentzkow, and Shapiro (2017).

earnings with a small autocovariance of wage growth, as in the data. The fit of the wage growth variance is also good. Average earnings growth is slightly overstated unconditionally and slightly understated conditional on J2J moves and U2E transitions. Overall, the model captures the earnings process well, including the joint process of firm performance and worker earnings. With respect to firm performance, the total variance of permanent value added growth is estimated to lie within the standard deviation, and its lagged covariance and covariance with wage growth come close to the data. Lastly, value added growth covaries with changes in separation probabilities similarly in both the model and the data.

The corresponding parameter values are presented in Table 2. Worker productivity and match quality are highly persistent with parameter estimates at 0.91 and 0.95, respectively. They are also very dispersed, especially compared to the permanent worker productivity x_0, allowing for large cross-sectional productivity differences. Utility costs of effort increase fast with a curvature parameter of 0.37. The match efficiency parameter is 0.19, and workers search about 47% less efficiently on the job than when unemployed. Finally, the measurement errors on earnings and value added are around 0.20.

5 Results analysis

5.1 Understanding the optimal contract

Impulse response analysis. We report the impulse responses in the model to permanent positive and negative innovation shocks to x and z, scaled to generate a 10% output change. In practice, we simulate the histories of a cross-section of matches, and compare a treatment group that receives the permanent productivity shock with a control group that does not. Figures 3 and 4 report the differences in variables of interest between the control and treatment groups

Table 2: Parameter estimates

persistence for worker productivity	λ_x	0.91 (1.8e-05)
persistence for match quality	λ_z	0.95 (4.3e-06)
dispersion for worker permanent productivity	σ_{x_0}	0.33 (5.2e-05)
dispersion for worker transitory productivity	σ_{x_1}	0.70 (0.00023)
dispersion for match quality	σ_z	0.49 (0.00018)
effort cost parameter	γ_0	0.00064 (1.6e-07)
effort cost curvature	γ_1	0.37 (3.8e-05)
flow payment while unemployed	b	0.11 (5.4e-05)
efficiency of the matching function	α	0.19 (8.2e-06)
on-the-job search efficiency	κ	0.53 (8.7e-05)
measurement error on earnings	m_w	0.20 (0.00088)
measurement error on value added per worker	m_y	0.19 (4.0e-05)

Notes: Standard errors in brackets are computed according to the asymptotic variance formula derived in supplementary Appendix W4.4.

around the event at time 0.

After a positive shock to worker productivity x, output and the target wage immediately jump upward by 10 and 5%, respectively (see solid blue lines in Figure 3). At the same time, the plot shows how the positive responses in earnings and promised value are delayed, conforming with the theoretical result of wage smoothing in the optimal contract. This means that during the transition towards the new target wage, the worker is underpaid relative to her productivity, and so work effort goes down and labor market search is shifted
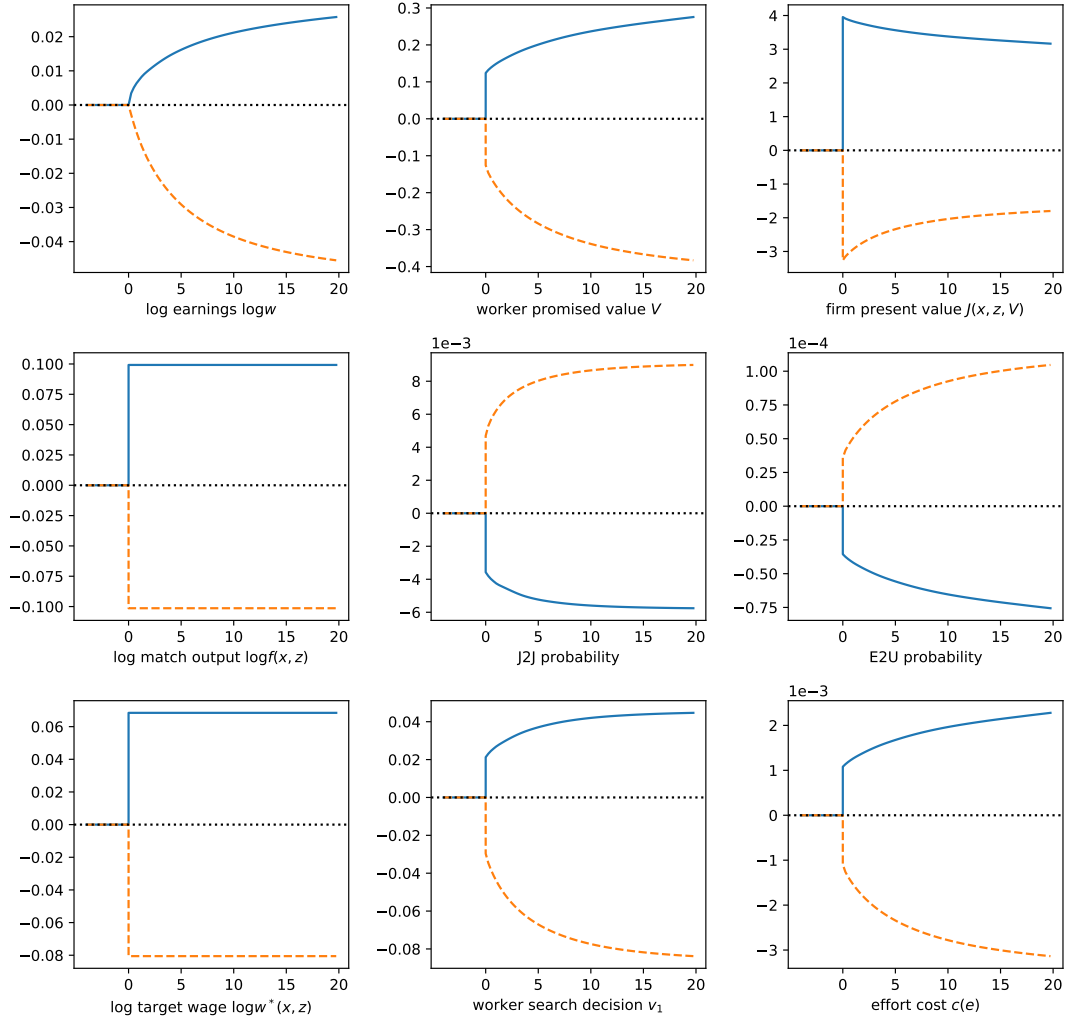
Figure 3: Average impulse response to x change at estimated parameters
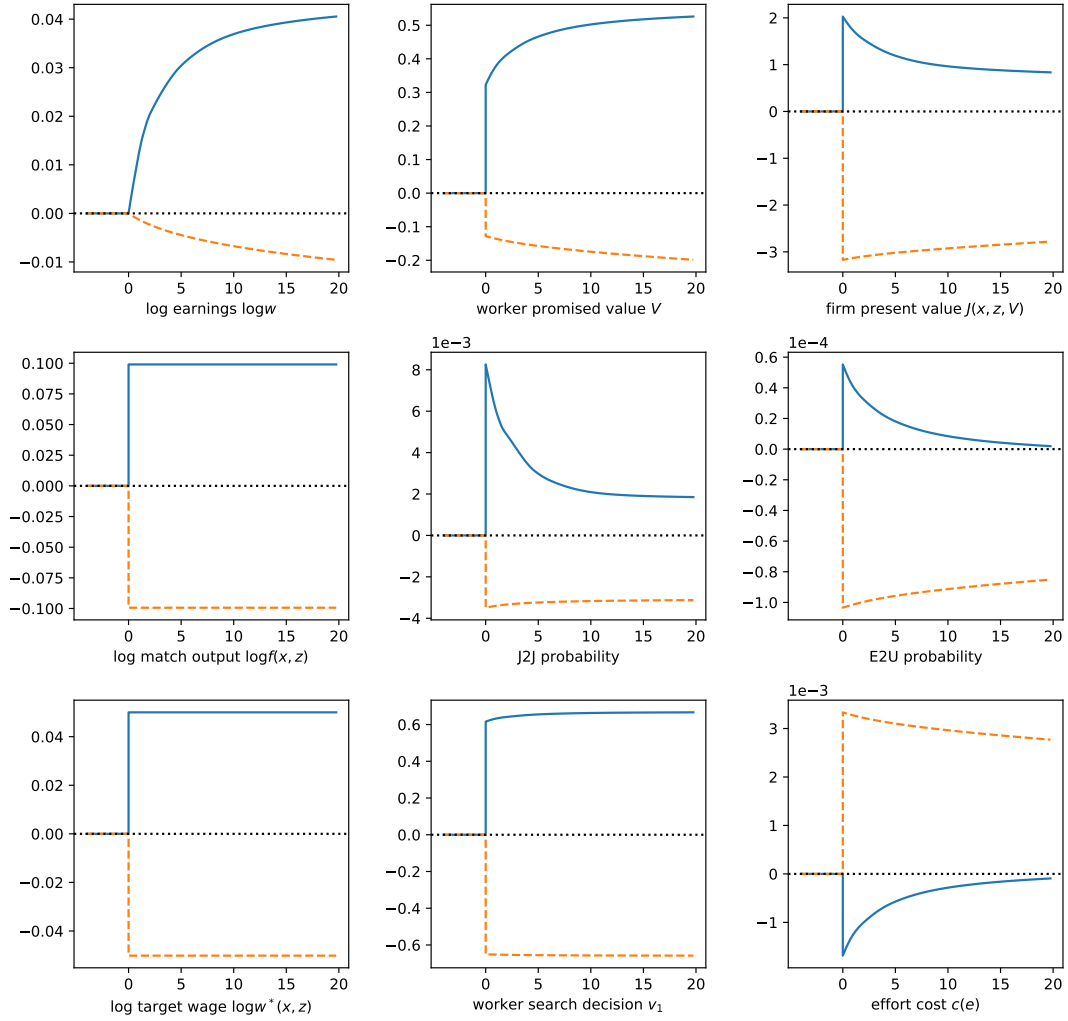
Notes: Effect of a positive (solid blue) and negative (dashed red) permanent x shock over time (years). Starting (x, z) values are drawn from the stationary distribution. Initial wages are target wages. Separation is ruled out.

towards jobs with higher lifetime utility v. Low effort and modified search drive up separation and the worker is more likely to transition into other jobs or (to a lesser extent) into unemployment. With time, as wages and promised utility approach their new long-run targets, effort levels improve and transition rates fall again. As the wage converges to the target wage, the firm's expected profit shrinks to zero and hence the worker extracts all the benefit of the positive pro-

Figure 4: Average impulse response to z change at estimated parameters

Notes: Effect of a positive (solid blue) and negative (dashed red) permanent z shock over time (years). Starting (x, z) values are drawn from the stationary distribution. Initial wages are target wages. Separation is ruled out.

ductivity shock. Search is permanently directed towards better jobs reflecting the permanent nature of the productivity shock.

Unsurprisingly, a negative x shock triggers adjustments of the opposite sign (see dashed red lines). However, the simulation reveals crucial asymmetries between positive and negative shocks, where the latter trigger earnings changes of smaller magnitude. The asymmetry reflects the fact that the firm provides

the worker with partial insurance tailored towards negative x shocks. Specifically, despite a target wage drop of similar magnitude, wages and promised values decrease much more slowly following a negative shock than a positive one. The worker has a stronger incentive to keep her relatively well-paid job and so chooses to exert more effort, which dampens the difference in J2J transitions to about one third relative to the positive shock. At the same time, firm profits take a bigger hit.

Figure 4 plots the impulse response functions of a permanent shock to match quality z. As in the previous graph, when a positive z shock hits, output and the target wage shoot up on impact. But there are three key differences compared to x shocks. First, the target wage reacts more strongly relative to output. Second, lifetime utility and wages rise more slowly, resulting in larger firm profits for an extended period. Third, changes in J2J and E2U transition probabilities are both smaller in magnitude and of opposite sign. The reason for all three points is that in contrast to x shocks, the productivity boost here is job-specific. This implies that the employee creates a higher value for the employer. However, unlike before, the worker's outside options have not changed and so the firm backloads wage payments longer without increased risk of separation. The effort and search decisions respond according to the worker's growing promised value of the match, lowering separation rates.

When comparing negative to positive z shocks, all adjustments between treatment and control groups reverse direction. Moreover, the main difference between the two shocks is that the J2J transition rates are about twice as large in magnitude. This asymmetry captures the fact that separation is used as an adjustment margin under the optimal contract in response to negative shocks to match productivity because both the firm and the worker can gain by forming a new match. Consequently, the firm offers lower values to its workers, who in turn reduce their effort and look for jobs with lower lifetime utility. As a

result, separation rates are high when production is low, which is in keeping with the negative correlation between changes in output and separation in the data. Since output and separation move in the same direction under x shocks, z shocks seem to be more important for successfully matching this moment.

To conclude, we find that the transmission effects of productivity shocks are delayed and there is a significant level of backloading of wages. The worker enjoys partial wage insurance, which insulates her particularly well against negative x shocks.

First best comparison. We now compare the impulse responses to shocks between the preceding baseline model and a first best contract, in which firms can observe the search and effort decisions of workers. This comparison is useful to understand how the optimal baseline contract accounts for incentive and commitment problems. For conciseness, we only point out the main differences here but we include the full first best responses in Figures A1 and A2 in Appendix A3.1, together with the solution technique.

Under the first best contract, wages stay perfectly flat while all other variables jump immediately to new long-run levels in response to productivity shocks. Although workers enjoy full wage insurance, their lifetime utility rises after a positive x or negative z shock through decreases in costly effort and applications to higher value jobs, which are also found more easily. Note that searching markets with higher v does not prevent job-to-job transition rates from simultaneously rising. This is because the tightness in each (x, v) submarket is not only determined by v but also by the higher productivity x. Increasing match qualities z imposes a utility loss on the worker ex-post, in contrast to the baseline model, although the first best contract delivers a higher value ex-ante to both firms and workers. In fact, at every value of firm profits J, the worker achieves a higher expected utility V in the contract without incentive and commitment problems.

Table 3: Level variance decompositions

	total	x_0	x_1	z	other
			overall		
match output f_{it}^a	0.31	11%	61%	24%	5%
target wage w_{it}^{*a}	0.14	32%	33%	31%	4%
earnings w_{it}^a	0.14	32%	7%	9%	53%
	within individual, over time				
match output f_{it}^a	0.08	0%	18%	6%	2%
target wage w_{it}^{*a}	0.03	0%	10%	7%	1%
earnings w_{it}^a	0.01	0%	3%	3%	5%

Notes: Using simulated data from the model, we run linear regressions with dummies for each level of (x_0, x_1, z) and report the variances as percentages of the total cross-sectional variances. All terms are annualized, in logs and exclude measurement error.

Decomposing earnings variances. To identify the sources of wage uncertainty, we use the model to decompose the variances of earnings and earnings growth into structural components. We also include match output to quantify the extent of insurance provided under the optimal contract against different types of shocks.

In Table 3, we show that the time-varying component of worker heterogeneity, x_1, contributes to most (61%) of the variations in annualized match output, but to only a small part of the variation in earnings (7%). Instead, the largest contemporaneous contributor (32%) to cross-sectional variations in wages is the permanent component of worker heterogeneity, x_0. This suggests that firms provide significant insurance to workers and pass through only a small share of time-varying worker shocks. The insurance against firm shocks appears smaller because their variance contribution to earnings (9%) is comparable to that of x_1 shocks, but the latter account for about two and a half times more of the variance in output.

Notably, a large part of the earnings variance is not captured by current

Table 4: Growth variance decompositions

	Total	U2E/E2U	J2J	x_1	z
$\mathrm{Var}(\Delta \log f_{it}^a)$	0.108	-8.0%	-3.6%	78.9%	18.3%
$\mathrm{Var}(\Delta \log w_{it}^a)$	0.005	86.2%	13.9%	40.0%	13.7%

Notes: Using simulated data from the model, we individually remove transitions in and out of employment, J2J transitions, x_1 and z shocks, keeping policies ξ fixed. All terms are annualized, in logs and exclude measurement error.

productivity levels, but instead can be attributed to the dispersion created by the path dependency of the contract, i.e. by differences in the promised value V.[13] For instance, two workers at the same current (x_0, x_1, z) may well have different values V and receive different wages because they differ in their employment histories and therefore in how far they are from the common target wage. We thus also conduct a within-individual variance decomposition over time. Our results consistently suggest that workers enjoy considerable insurance against productivity fluctuations, especially against x_1 shocks. Shocks to x_1 and z have a bigger impact on the variation of the target wage than on actual earnings, which is indicative of the wage smoothing embedded in the contract.

Next, we decompose the variances in growth of both yearly output and earnings into endogenous mobility and exogenous shock components using simulated data. To quantify the contribution of different features of the model, we remove one element at a time. Table 4 shows the results of this exercise. Shutting down transitional dynamics associated with moving in and out of employment lowers the wage variations by 86.2%. For comparison, the reduction from worker-specific shocks is about half this size (40.0%) and the contribution from z shocks is even smaller (13.7%). The fact that the terms sum to more than 100% reflects the presence of interactions between mobility and productivity shocks. Intuitively, to the extent that negative productivity shocks affect

---

[13]The residual remains large even when including interaction terms between x_0, x_1 and z.

earnings less than positive shocks, as seen in the impulse response analysis, a worker who is paid less than the target wage can expect stronger wage adjustments in response to changes in $w^*(x, z)$. Combining this observation with the fact that a worker exiting unemployment tends to start her new job below $w^*(x, z)$ illustrates how mobility can interact with shocks.

As in the level decomposition, x_1 captures the largest (78.9%) share of the variation in match output but contributes less to the wage growth variance (40.0%), reflecting the insurance against time-varying worker shocks embedded in the contract. Interestingly, J2J transitions affect the growth variances of output and wages in opposite directions. While a positive contribution to the wage growth variance can be expected when considering that moving allows workers to extract higher wages faster in response to positive x shocks, the option to separate after adverse z shocks also curtails drops in output. J2J transitions thus limit the effects of shocks on production and show up as a negative contributor to the output growth variance.

Passthrough analysis. Typically, passthrough analyses assess how much of a persistent and transitory shock to firm output is transmitted to worker earnings. We extend this definition to account for three difficulties. First, the passthrough in our setup needs to consider a shock with a given level of persistence, rather than purely transitory or fully permanent shocks. In order to capture the full effect of a given shock, we thus evaluate its impact on match output and on worker outcomes in terms of expected present values (EPV). And to study both worker and firm shocks we use match output rather than an aggregate measure such as value added. Second, a shock in our model alters the incentive structure and hence job transitions — its overall effect is no longer limited to a single firm-worker relationship. We thus compute the EPV of wages and output both for the case where workers and employers are forced to stay together, as well as for the case where workers can choose to move to new employers or into

Table 5: Passthrough analysis

	mobility	x_1 shock	z shock
utility passthrough	outcome only	0.27	0.11
	yes	0.25	0.18
wage passthrough	yes	0.33	0.42
	no	0.24	0.67

Notes: Passthroughs are computed as the elasticity of the discounted present value of worker outcomes with respect to match output. Mobility is removed in row 4 for wages, and in rows 1 and 4 for productivity. See Appendix A3.2.

unemployment. Third, one can go beyond focusing on wages and instead look at the impacts on the total lifetime utility of the worker, which accounts for effort and risk aversion. We therefore contrast the wage passthrough, based on the present discounted value of wages, with a utility passthrough, computed as the change in the constant consumption equivalent of the worker's promised value with and without the shock. We compute passthroughs as an elasticity at the individual level and average over the stationary distribution. In Appendix A3.2, we show how this approach can map onto previous definitions of passthrough in the literature.

The results of our passthrough analysis are presented in Table 5. In our preferred specification (row 1), we report the passthrough of an innovation shock in x_1 and z to a worker's total lifetime utility and find 0.27 and 0.11, respectively. In other words, a productivity shock, which raises the EPV of output in a given firm-worker match by 10%, translates into an increase in the worker's lifetime utility amounting to a 2.7% and 1.1% higher constant consumption equivalent, respectively. Here, the worker outcome incorporates the present value of utility, including wages at future employers. The measure of the shock in terms of match output is not affected by the contractual environment or endogenous mobility and hence is exogenous. This allows us to interpret the passthrough as reflecting the contractual decision by the firm about how much of an exogenous

shock to transmit to the employee. Alternatively, we can allow for labor mobility in production (row 2), and find similar values (0.25 and 0.18). In order to isolate how much of the overall utility effect is due to the changes in wages, we compute wage passthroughs and find elasticities of 0.33 for worker shocks and 0.42 for firm shocks (row 3). These measures are very sensitive to the assumption that no separation occurs, which results in an overstated passthrough value for z shocks (0.67) and a lower elasticity for x_1 shocks (0.24). In conclusion, these findings suggest a significant amount of shock passthrough. They also show that both wages and mobility are important components of firm insurance.

5.2   Policy evaluation

The previous results show that while workers are partially insured by their employers, they still face non-negligible earnings uncertainty. Here, we consider how a policy maker might attempt to mitigate this uncertainty through tax policies that alter the level of earnings inequality faced by individuals. For instance, a government may wish to tax individuals with high incomes in order to pay transfers to low earners. Such a progressive tax schedule may contribute to earnings stabilization.

We analyze the effect of a revenue-neutral tax policy that redistributes income from high to low wage earners. Budget neutrality respects differences in total benefits paid due to changes in the unemployment rate. We follow Heathcote, Storesletten, and Violante (2014) and parameterize the policy as $w_{\text{net}} = \tau_0 w_{\text{gross}}^{\tau_1}$, pick a set of values for τ_1 and solve for the revenue-neutral τ_0 in the estimated model. Lowering τ_1 reduces the progressivity of the tax schedule and dictates a higher τ_0, whereas increasing τ_1 has the opposite effect. In Table 6, we report the effect of the tax policies $\tau_1 = (0.9, 1.1)$ on the cross-sectional variance of annualized earnings and earnings growth, net of measurement error. The first two columns report the policy effects on gross and net wages, holding

Table 6: Effect of revenue-neutral policies

		baseline contracts		reoptimized contracts	
		gross	net	gross	net
More progressive	$Var(\log w^a)$	0.146	0.117 (-19.7%)	0.151 $(+3.8\%)$	0.123 (-15.6%)
$(\tau_1 = 0.9, \tau_0 = 1.15)$	$Var(\Delta \log w^a)$	0.005	0.004 (-22.6%)	0.005 $(+4.0\%)$	0.004 (-14.3%)
Less progressive	$Var(\log w^a)$	0.146	0.175 $(+19.9\%)$	0.140 (-3.7%)	0.172 $(+17.9\%)$
$(\tau_1 = 1.1, \tau_0 = 0.87)$	$Var(\Delta \log w^a)$	0.005	0.006 $(+15.5\%)$	0.004 (-8.4%)	0.005 $(+9.6\%)$

Notes: Policies take the form $w_{\text{net}} = \tau_0 w_{\text{gross}}^{\tau_1}$. Baseline contracts are fixed, while reoptimized contracts include equilibrium responses of workers and firms to the policy. Percentage changes are relative to the baseline estimation. Outcomes are annualized and exclude measurement error.

the decisions of workers and firms fixed. These numbers can be interpreted as the intended policy effect. However, labor market contracts are generally free to change in response to new tax policies. Hence, the last two columns show the same effects under full equilibrium reoptimization of contracts.

We find that if a government redistributes from high to low earners, almost 40% of the protection against variation in earnings growth is undone by a reduction in firm-provided insurance. In other words, firm insurance is crowded out by public insurance. To see this, note that the more progressive policy achieves a 22.6% reduction in wage growth variation under baseline contracts, but falls well short of this when allowing for contract reoptimization. The reasoning is as follows: firms are aware that the government is providing insurance to their workers and so they choose to pass on more productivity risk to them. The new equilibrium contracts thus stipulate higher $(+4.0\%)$ pre-tax variation, which attenuates the effect of the transfer policy to a 14.3% reduction in wage growth variance. A less progressive tax policy, due to a higher τ_1, increases the net wage growth variance $(+15.5\%)$ without contract adjustments. Since this reduces workers' utility, firms counteract the policy by setting less (-8.4%) variable gross wages. But as the firms' remedial actions are incomplete, the final wage variation is still more volatile $(+9.6\%)$ than under the baseline sce-

nario. A similar crowding-out effect of $10-20\%$ dampens the policy impact on wage variances. This exercise demonstrates how, at realistic parameter values, accounting for firms' decisions to provide insurance can have important implications for the effectiveness of government-provided insurance.

6 Conclusion

In this paper, we study different sources of earnings uncertainty faced by workers in the labor market. We develop an equilibrium model with search frictions and show that the optimal contract between firms and workers prescribes a wage that smoothly tracks both positive and negative worker and firm shocks, deviating from the prediction of a perfectly competitive model. We provide conditions for identification and estimate a parameterized version of the model using Swedish matched employer-employee data. Our main findings are that shocks are only partially transmitted to worker earnings and that firms insure their workers extensively against non-permanent worker shocks. We extend prior definitions of passthrough in order to incorporate job mobility. Finally, we show that publicly provided insurance substantially crowds out insurance provision by firms.

Future work could extend the current model in several ways. Including savings, which would allow individuals to self-insure, is an important next step despite being computationally and theoretically difficult. Other work could build on the block recursivity of the equilibrium in order to study aggregate shocks and evaluate the cost of business cycles in an environment with firm insurance. Another promising avenue for future work is to include heterogeneity across firms and examine how the provision of insurance links to firm size by allowing for permanent heterogeneity, amenities at the firm level or heterogeneous vacancy costs. One could also link the question of passthrough to commonly found collective agreements in the labor market, such as guaranteed minimum

wage increases, limitations to write long-term contracts or common pay components. These extensions, which are constraints on the contracting space, could build on the results developed in this paper.

Finally, we believe that the insurance framework developed here will be useful for other research agendas where relational long-term contracts and insurance provision, together with search frictions, are important. For example, this framework may prove useful for those working on relational banking, insurance markets and repeated buyer-seller transactions.

References

ABOWD, J. M., AND D. CARD (1989): "On the covariance structure of earnings and hours changes," Econometrica, 57(2), 411–445.

ÁBRAHÁM, A., F. ALVAREZ-PARRA, AND S. FORSTNER (2017): "The effects of moral hazard on wage inequality in a frictional labor market," Discussion paper.

ÁBRAHÁM, A., AND N. PAVONI (2008): "Efficient Allocations with Moral Hazard and Hidden Borrowing and Lending: A Recursive Formulation," Review of Economic Dynamics, 11(4), 781–803.

ABREU, D., D. PEARCE, AND E. STACCHETTI (1986): "Optimal cartel equilibria with imperfect monitoring," Journal of Economic Theory, 39(1), 251–269.

ALTONJI, J. G., A. SMITH, AND I. VIDANGOS (2013): "Modeling earnings dynamics," Econometrica, 81(4), 1395–1454.

ANDREWS, I., M. GENTZKOW, AND J. M. SHAPIRO (2017): "Measuring the Sensitivity of Parameter Estimates to Estimation Moments," Quarterly Journal of Economics, 132(4), 1553–1592.

ATTANASIO, O. P., AND N. PAVONI (2011): "Risk Sharing in Private Information Models With Asset Accumulation: Explaining the Excess Smoothness of Consumption," Econometrica, 79(4), 1027–1068.

AZARIADIS, C. (1975): "Implicit Contracts and Underemployment Equilibria," Journal of Political Economy, 83(6), 1183–1202.

BAGGER, J., F. FONTAINE, F. POSTEL-VINAY, AND J.-M. ROBIN (2014): "Tenure, Experience, Human Capital, and Wages: A Tractable Equilibrium

Search Model of Wage Dynamics," American Economic Review, 104(6), 1551–96.

Baily, M. N. (1974): "Wages and employment under uncertain demand," The Review of Economic Studies, 41(1), 37–50.

Blundell, R., L. Pistaferri, and I. Preston (2008): "Consumption Inequality and Partial Insurance," American Economic Review, 98(5), 1887–1921.

Burdett, K., and M. Coles (2003): "Equilibrium Wage-Tenure Contracts," Econometrica, 71(5), 1377–1404.

Burdett, K., S. Shi, and R. Wright (2001): "Pricing and Matching with Frictions," Journal of Political Economy, 109(5), 1060–1085.

Card, D., A. R. Cardoso, J. Heining, and P. Kline (2018): "Firms and labor market inequality: Evidence and some theory," Journal of Labor Economics, 36(S1), S13–S70.

Carlsson, M., J. Messina, and O. Nordström Skans (2014): "Firm-level shocks and labor adjustments," Discussion Paper 2014:28.

Chiappori, P., B. Salanie, and J. Valentin (1999): "Early Starters versus Late Beginners," Journal of Political Economy, 107(4), 731–760.

Cole, H., and F. Kubler (2012): "Recursive contracts, lotteries and weakly concave Pareto sets," Rev. Econ. Dyn., 15(4), 479–500.

Dardanoni, V. (1995): "Income distribution dynamics: monotone Markov chains make light work," Soc. Choice Welfare, 12(2), 181–192.

Farhi, E., and I. Werning (2013): "Insurance and taxation over the life cycle," Review of Economic Studies, 80(2), 596–635.

Friedrich, B., L. Laun, C. Meghir, and L. Pistaferri (2019): "Earnings dynamics and firm-level shocks," Discussion paper, National Bureau of Economic Research.

Guiso, L., L. Pistaferri, and F. Schivardi (2005): "Insurance within the Firm," Journal of Political Economy, 113(5), 1054–1087.

Hall, R. E., and F. S. Mishkin (1980): "The sensitivity of consumption to transitory income: estimates from panel data on households," Discussion paper, National Bureau of Economic Research.

Harris, M., and B. Holmstrom (1982): "A Theory of Wage Dynamics," The Review of Economic Studies, 49(3), 315–333.

HEATHCOTE, J., K. STORESLETTEN, AND G. L. VIOLANTE (2014): "Consumption and labor supply with partial insurance: An analytical framework," American Economic Review, 104(7), 2075–2126.

HOLMSTROM, B., AND P. MILGROM (1991): "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design," Journal of Law, Economics, & Organization, 7, 24–52.

HOPENHAYN, H. A., AND J. P. NICOLINI (1997): "Optimal Unemployment Insurance," Journal of Political Economy, 105(2), 412–438.

HORNSTEIN, A., P. KRUSELL, AND G. L. VIOLANTE (2011): "Frictional wage dispersion in search models: A quantitative assessment," American Economic Review, 101(7), 2873–98.

HU, Y., AND M. SHUM (2012): "Nonparametric identification of dynamic models with unobserved state variables," Journal of Econometrics, 171(1), 32–44.

KNIGHT, F. H. (1921): Risk, uncertainty and profit, vol. 31. Houghton Mifflin.

KOCHERLAKOTA, N. R. (1996): "Implications of Efficient Risk Sharing without Commitment," The Review of Economic Studies, 63(4), 595–609.

LEMIEUX, T., W. B. MACLEOD, AND D. PARENT (2009): "Performance pay and wage inequality," The Quarterly Journal of Economics, 124(1), 1–49.

LENTZ, R. (2013): "Optimal Wage-Tenure Contracts without Search Intensity Commitment," Discussion paper, Citeseer.

LISE, J., C. MEGHIR, AND J.-M. ROBIN (2016): "Matching, sorting and wages," Review of Economic Dynamics, 19, 63–87.

LOW, H., C. MEGHIR, AND L. PISTAFERRI (2010): "Wage Risk and Employment Risk over the Life Cycle," American Economic Review, 100(4), 1432–1467.

MACLEOD, W. B., AND J. M. MALCOMSON (1989): "Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment," Econometrica, 57(2), 447–480.

MACURDY, T. E. (1982): "The use of time series processes to model the error structure of earnings in a longitudinal data analysis," Journal of Econometrics, 18(1), 83–114.

MARCET, A., AND R. MARIMON (2019): "Recursive contracts," Econometrica, 87(5), 1589–1631.

MEGHIR, C., AND L. PISTAFERRI (2004): "Income Variance Dynamics and Heterogeneity," Econometrica, 72(1), 1–32.

MENZIO, G., AND S. SHI (2010): "Block recursive equilibria for stochastic models of search on the job," Journal of Economic Theory, 145(4), 1453–1494.

———— (2011): "Efficient Search on the Job and the Business Cycle," Journal of Political Economy, 119(3), 468–510.

MESSNER, M., N. PAVONI, AND C. SLEET (2012): "Recursive methods for incentive problems," Rev. Econ. Dyn., 15(4), 501–525.

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

MONTGOMERY, J. D. (1991): "Equilibrium Wage Dispersion and Interindustry Wage Differentials," The Quarterly Journal of Economics, 106(1), 163–179.

PETERS, M. (1991): "Ex Ante Price Offers in Matching Games Non-Steady States," Econometrica, 59(5), 1425–1454.

POSTEL-VINAY, F., AND J.-M. ROBIN (2002): "Equilibrium wage dispersion with worker and employer heterogeneity," Econometrica, 70(6), 2295–2350.

———— (2004): "To match or not to match?: Optimal wage policy with endogenous worker search intensity," Review of Economic Dynamics, 7(2), 297–330.

PRESCOTT, E. C., AND R. M. TOWNSEND (1984): "General competitive analysis in an economy with private information," International Economic Review, pp. 1–20.

ROGERSON, W. P. (1985): "Repeated Moral Hazard," Econometrica, 53(1), 69–76.

ROYS, N. (2016): "Persistence of shocks and the reallocation of labor," Review of Economic Dynamics, 22, 109–130.

RUDANKO, L. (2009): "Labor market dynamics under long-term wage contracting," Journal of Monetary Economics, 56(2), 170–183.

SCHAAL, E. (2017): "Uncertainty and unemployment," Econometrica, 85(6), 1675–1721.

SHI, S. (2009): "Directed Search for Equilibrium Wage-Tenure Contracts," Econometrica, 77(2), 561–584.

SHIMER, R. (2005): "The assignment of workers to jobs in an economy with coordination frictions," Journal of Political Economy, 113(5), 996–1025.

SPEAR, S. E., AND S. SRIVASTAVA (1987): "On Repeated Moral Hazard with Discounting," The Review of Economic Studies, 54(4), 599–617.

STEVENS, M. (2004): "Wage-tenure contracts in a frictional labour market: Firms' strategies for recruitment and retention," The Review of Economic Studies, 71(2), 535–551.

STOKEY, N. L., R. LUCAS, AND E. PRESCOTT (1989): "Recursive methods in dynamic economics," Cambridge, MA: Harvard University.

THOMAS, J., AND T. WORRALL (1988): "Self-enforcing wage contracts," The Review of Economic Studies, 55(4), 541–554.

TSUYUHARA, K. (2016): "Dynamic Contracts with Worker Mobility via Directed On-the-Job Search," International Economic Review, 57(4), 1405–1424.

A1 Model appendix

A1.1 Existence of equilibrium

Proof of Proposition 1. First, fix an arbitrary $\varepsilon \in \mathbb{R}_{++}$. Let ρ_ε be the unique positive solution for ρ of the equation

$$\varepsilon_T(\rho) = \varepsilon$$

$\forall \, J_m, J_n \in \mathbb{J}$ such that $\|J_m - J_n\| < \rho_\varepsilon$. Lemma W14 implies that $\|TJ_m - TJ_n\| < \varepsilon$, which means that the equilibrium operator T is continuous. Next, let ρ_x and ρ_z denote the minimum distance between distinct elements associated with the sets $\mathbb{X}$ and $\mathbb{Z}$, respectively. Also, let $\|\cdot\|_E$ denote the standard norm on the Euclidean space $\mathbb{S} \times \mathbb{V}$. Let $\tilde{\rho}_\varepsilon = \min\{\underline{u}'\varepsilon, \rho_x, \rho_z\}$. For all $(x_1, z_1, V_1), (x_2, z_2, V_2) \in \mathbb{S} \times \mathbb{V}$ such that $\|(x_2, z_2, V_2) - (x_1, z_1, V_1)\|_E < \tilde{\rho}_\varepsilon$ and for all $J \in \mathbb{J}$, Lemma W13 implies that TJ satisfies the property (J1) of the set $\mathbb{J}$ and, consequently, $|(TJ)(x_2, z_2, V_2) - (TJ)(x_1, z_1, V_1)| < \varepsilon$. Hence, the family of functions $T(\mathbb{J})$ is equicontinuous. The lemma also implies that the Bellman operator is self-mapping. The lemmas can be found in the web Appendix W1.2.

From these properties, it follows that the equilibrium operator T satisfies the conditions of Schauder's fixed point theorem (Stokey, Lucas, and Prescott (1989), Theorem 17.4). Therefore, there exists a value function $J^* \in \mathbb{J}$ for the firm such that $TJ^* = J^*$. Let θ^* denote the market tightness function computed with J^*, which then gives rise to vacancy value and mass functions Π^* and ϕ^*, respectively. J^* and θ^* pin down the active job distribution h^*, a worker retention probability $\tilde{p}^*$ and a search return function denoted by $\tilde{r}^*$. Denote as U^* the unemployment value function computed with θ^* and let μ^* be the associated mass of unemployed workers. Let ξ^* denote the contract policy function computed with $J^*, \theta^*, \tilde{p}^*$ and U^*. The functions $\{J^*, \theta^*, \tilde{p}^*, \tilde{r}^*, U^*, \Pi^*, h^*, \phi^*, \mu^*, \xi^*\}$ satisfy the conditions in the definition of the recursive search equilibrium. $\qquad \square$

A1.2 Characterization of the optimal contract

Proof of Lemma 1.

Uniqueness of v_1^* and e^*. Recall that given (x, W_i) the policies v_1^* and e^* solve:

$$\max_{v_1, e} u(w) - c(e) + \beta\delta(e)\mathbb{E}_{x'}[U(x')|x] + \beta(1 - \delta(e))\kappa p(\theta(x, v_1))v_1$$
$$+ \beta(1 - \delta(e))(1 - \kappa p(\theta(x, v_1)))W_i.$$

Note that $v_1^*(x, W_i)$ can be determined independently of the effort choice and is equal to $m(x, W_i)$ and thus inherits its uniqueness, monotonicity and continuity in W_i. Next, we normalize $\delta(e) = 1 - e$ (or equivalently redefine c such that $c(e) = c(\delta^{-1}(1 - e))$). Then the first order condition for effort

$$c'(e) = \beta\kappa p(\theta(x, v_1^*(x, W_i)))\left(v_1^*(x, W_i) - W_i\right) + \beta W_i - \beta\mathbb{E}_{x'}[U(x')|x]$$

reveals that under the assumption that $c(\cdot)$ is strictly convex and twice differentiable, the effort policy $e^*(x, W_i)$ is also uniquely determined. Furthermore, the effort policy function inherits continuity and differentiability a.e. from $\hat{p}(x, W_i)$ and $D(x, W_i)$.

$\tilde{p}(x, W_i)$ **is continuous, differentiable a.e. and increasing in W_i.** Now, consider the composite transition probability, rewritten as

$$\tilde{p}(x, W_i) = e^*(x, W_i)\left(1 - \kappa\hat{p}(x, W_i)\right),$$

which is continuous and differentiable a.e. because the right hand side exhibits these properties. We take the derivative with respect to W_i

$$\tilde{p}'(x, W_i) = e^{*\prime}(x, W_i)\left(1 - \kappa\hat{p}(x, W_i)\right) - \kappa e^*(x, W_i)\hat{p}'(x, W_i) > 0,$$

where the inequality uses the fact that $e^*(x, W_i)$ is increasing in W_i and that $\hat{p}(x, W_i)$ is decreasing in W_i as shown in Lemmas W12 and W5, respectively.

$\tilde{r}(x, W_i)$ **is increasing and differentiable a.e. in** W_i **and** $\tilde{r}'(x, W_i) = \beta\tilde{p}(x, W_i)$. Finally, we use the envelope condition to compute the derivative of $\tilde{r}(x, W_i)$ with respect to W_i as

$$\tilde{r}'(x, W_i) = \beta e^*(x, W_i)(1 - \kappa p(\theta(x, v_1^*(x, W_i)))) = \beta\tilde{p}(x, W_i),$$

which proves that $\tilde{r}(x, W_i)$ is continuous and differentiable a.e.

Monotonicity of J **in** z. Let's consider two different match qualities $z_1 < z_2$ where $z_1, z_2 \in \mathbb{Z}$. The intuition guiding the following proof is that a firm starting at (x, z_2) can mimic the strategy of a reference firm in state (x, z_1) and make more profits than its reference competitor. We then show that the mimicking strategy, albeit feasible, delivers lower profits than the firm's best strategy.

Let ξ_1 be the optimal history-contingent policy of a reference firm starting at (x, z_1, V) and let $h^t = (s^t, \varepsilon^t) \in \mathbb{S}^t \times [0, 1]^t$ denote the entire shock history of productivity, match quality and lottery realizations. Then expected profits are given by:

$$J(x, z_1, V) = \sum_{t=1}^{\infty} \sum_{h^t} \beta^{t-1} \Big(f(x_t, z_t) - w_{1,t}(h^t) \Big) \Lambda_{1,t}(h^t),$$

where $w_{1,t}(h^t)$, $e_{1,t}(h^t)$ and $v_{1,t}(h^t)$ are the contract policies implemented by ξ_1 and $\Lambda_{1,t}(h^t) = \prod_{\tau=0}^{t-1}(1 - \delta(e_{1,\tau}(h^\tau)))(1 - \kappa p(\theta(x_\tau, v_{1,\tau}(h^\tau)))$ is the composition of all separation probabilities on the path.

Next, we change indexing from histories h^t to realizations $(t; \omega) \in [0, 1]$ in the probability space by ordering the histories lexicographically (such that the rank is determined first by worker productivity x, next by lottery realization ε and last by match qualities z). This allows us to rewrite expected profits as:

$$J(x, z_1, V) = \int \sum_{t=1}^{\infty} \beta^{t-1} \Big(f(x(t; \omega), z_1(t; \omega)) - w_1(t; \omega) \Big) \Lambda_1(t; \omega) \, d\omega.$$

Because of independence between x and z, it is inconsequential for $x(t; \omega)$ whether the firm starts in (x, z_1) or (x, z_2). However, $z_2(t; \omega) \geq z_1(t; \omega)$ be-

A3

cause the transition function $g(\cdot, \cdot)$ is assumed to be monotonic and $z_2 > z_1$.

Consider now the following value of a job starting in (x, z_2, V):

$$J_2 = \int \sum_{t=1}^{\infty} \beta^{t-1} \Big(f(x(t; \omega), z_2(t; \omega)) - w_1(t; \omega) \Big) \Lambda_1(t; \omega) \, d\omega,$$

which delivers the same value V to the worker because all wages, all x realizations and all transitions are identical to the ones associated with ξ_1. Since this contract starts at z_2 while using the optimal strategy of the reference firm at z_1, it equals at most the value of its own optimal strategy, i.e. $J_2 \le J(x, z_2, V)$. Given that histories are constructed such that $\forall (t; \omega), z_2(t; \omega) \ge z_1(t; \omega)$, it holds that

$$J(x, z_2, V) \ge J_2 = \int \sum_{t=1}^{\infty} \beta^{t-1} \Big(f(x(t; \omega), z_2(t; \omega)) - w_1(t; \omega) \Big) \Lambda_1(t; \omega) \, d\omega$$

$$\ge \int \sum_{t=1}^{\infty} \beta^{t-1} \Big(f(x(t; \omega), z_1(t; \omega)) - w_1(t; \omega) \Big) \Lambda_1(t; \omega) \, d\omega$$

$$= J(x, z_1, V),$$

which gives the result. See Dardanoni (1995) for more details on properties of monotonic Markov chains. $\qquad \square$

Recall the expected profit equation:

$$J(x, z, V) = \max_{\pi_i, w_i, W_i, W_{ix'z'}} \sum_{i=1,2} \pi_i \Big(f(x, z) - w_i + \beta \tilde{p}(x, W_i) \mathbb{E}_{x'z'}[J(x', z', W_{ix'z'})|x, z] \Big)$$

$$s.t \quad (\lambda) \quad V \le \sum_{i=1,2} \pi_i \left(u(w_i) + \tilde{r}(x, W_i) \right)$$

$$(\omega_i) \quad W_i = \mathbb{E}_{x'z'}[W_{ix'z'}|x, z],$$

$$\sum_{i=1,2} \pi_i = 1,$$

where λ and ω_i denote Lagrange multipliers on constraints.

Lemma A1 (Wage and lifetime utility). *For a given (x, z), a higher wage always means higher lifetime utility.*

Proof of Lemma A1. This is a direct implication of the concavity of J, the envelope condition and the first order condition for the wage:

$$J'(x, z, V) = -\frac{1}{u'(w)}.$$

Note that the wage $w_i = w$ is constant due to $u'(w_i) = 1/\lambda$. The concavity of $u(\cdot)$ then implies that w and V always move in the same direction. □

Proof of Proposition 2. From Lemma A1, the wage in the current period is given by

$$i = 1, 2 \quad u'(w_i) = \frac{1}{\lambda} = -\frac{1}{J'(x, z, V)}$$

and rolling this expression forward shows the wage next period in state (x', z') satisfies:

$$\frac{1}{u'(w'_{ix'z'})} = -J'(x', z', W_{ix'z'}).$$

Next, the first order condition with respect to W_i is

$$\pi_i \beta \tilde{p}'(x, W_i) \mathbb{E}_{x'z'}[J(x', z', W_{ix'z'})|x, z] + \lambda \pi_i \tilde{r}'(x, W_i) + \omega_i = 0,$$

which can be rewritten after substituting $\tilde{r}'(x, W_i) = \beta \tilde{p}(x, W_i)$, derived in Lemma 1, as:

$$\pi_i \beta \tilde{p}'(x, W_i) \mathbb{E}_{x'z'}[J(x', z', W_{ix'z'})|x, z] + \lambda \pi_i \beta \tilde{p}(x, W_i) + \omega_i = 0.$$

To replace ω_i, we use the first order condition for $W_{ix'z'}$, which is

$$\pi_i \beta \tilde{p}(x, W_i) J'(x', z', W_{ix'z'}) - \omega_i = 0,$$

resulting in the following expression:

$$\tilde{p}'(x, W_i) \mathbb{E}_{x'z'}[J(x', z', W_{ix'z'})|x, z] + \lambda \tilde{p}(x, W_i) + \tilde{p}(x, W_i) J'(x', z', W_{ix'z'}) = 0.$$

Focusing on $\tilde{p}(x, W_i) > 0$ and $\pi_i > 0$ (because otherwise the worker is leaving

the current firm and the next period wage is irrelevant) we now rewrite:

$$\frac{\tilde{p}'(x, W_i)}{\tilde{p}(x, W_i)} \mathbb{E}_{x'z'}[J(x', z', W_{ix'z'})|x, z] = -J'(x', z', W_{ix'z'}) - \lambda.$$

We finally use the envelope condition to express the right hand side in terms of current and future wages:

$$\frac{\tilde{p}'(x, W_i)}{\tilde{p}(x, W_i)} \mathbb{E}_{x'z'}[J(x', z', W_{ix'z'})|x, z] = \frac{1}{u'(w'_{ix'z'})} - \frac{1}{u'(w)},$$

where since $\tilde{p}'(x, W_i) > 0$ the inverse marginal utility and consequently wages move according to the sign of the expected surplus of the firm. Introducing $\eta(x, W_i) \equiv \frac{\tilde{p}'(x,W_i)}{\tilde{p}(x,W_i)} = \frac{\partial \log \tilde{p}(x,W_i)}{\partial W_i} > 0$ and using the fact that the wage in the following period must be independent of the realization of x' and z', we can state the result:

$$\eta(x, W_i)\mathbb{E}_{x'z'}[J(x', z', W_{ix'z'})|x, z] = \frac{1}{u'(w'_i)} - \frac{1}{u'(w)},$$

which shows that within each realization of the lottery, the wage will move in line with expected profits. □

Proof of Proposition 3. This proof establishes existence and uniqueness of the target wage, before turning to the transition towards it and monotonicity.

Existence of w^*. To begin with, we show that there exists a $W^*(x, z)$ that renders the continuation value zero, $M(x, z, W^*(x, z)) = 0$. On the one hand, $M(x, z, W) \leq 0$ for very large W. Imagine that the firm promises more to the worker than it could possibly produce, even if it kept the highest possible productivity and match quality forever. Then its continuation value cannot be positive. On the other hand, $M(x, z, W) \geq 0$ for very small W. If the firm promised very low wages to the worker, then it could either make positive profits or the worker could leave the firm, leaving it with a zero continuation value. Hence, clearly there exists a promised value for which $M(x, z, W) \geq 0$.

Next, recall from Lemma W3 that $M(x, z, W)$ is continuous in W, based on the continuity of J in V and the Maximum Theorem. This means that $M(x, z, W)$ is continuous and there exists a promised value W^* where it crosses (or touches) zero.

Finally, since there exists a value $W^*(x, z)$ that satisfies $M(x, z, W^*(x, z)) = 0$, there exists a wage associated with this $W^*(x, z)$, given by:

$$w^*(x, z) = u'^{-1}\left(-\frac{1}{J'(x', z', W^*_{x'z'}(x, z))}\right),$$

where $W^*_{x'z'}(x, z) = \arg\max_{W_{x'z'}} \mathbb{E}_{x'z'}[J(x', z', W_{x'z'})|x, z]$ s.t. $W^*(x, z) = \mathbb{E}_{x'z'}[W_{x'z'}|x, z]$. This is the target wage $w^*(x, z)$, which shows that it exists.

Uniqueness of w^*. The next goal is to show that given a fixed (x, z) there is a unique $w^*(x, z)$ such that the continuation value of the firm is zero. Since $M(x, z, W)$ is a strictly decreasing function of W, the value $W^*(x, z)$ rendering the continuation value zero is unique. The fact that a unique $W^*(x, z)$ implies a unique $w^*(x, z)$ is proven by contradiction. Suppose that despite a unique value $W^*(x, z)$, there are two different target wages $w_1^*(x, z) < w_2^*(x, z)$. Then it must be that $J'(x', z', W^*_{1x'z'}(x, z)) > J'(x', z', W^*_{2x'z'}(x, z))$ from the target wage's definition.

Now, for both target wages the $J'(x', z', W^*_{ix'z'}(x, z))$ do not depend on (x', z'). To see this, consider the first order condition for $W_{x'z'}$ in the optimization of M:

$$J'(x', z', W^*_{x'z'}) = \omega,$$

where ω is the Lagrange multiplier on the constraint. Although the Lagrange multiplier itself could be non-unique, within each of the two potential solutions considered here, it must hold that $J'(x', z', W^*_{ix'z'}(x, z))$ does not depend on (x', z') because ω does not depend on (x', z'). As a consequence, the condition that $J'(x', z', W^*_{1x'z'}(x, z)) > J'(x', z', W^*_{2x'z'}(x, z))$ implies that

$W^*_{1x'z'}(x, z) < W^*_{2x'z'}(x, z)$ at each (x', z'). Taking expectations conditional on starting from the same (x, z) leads to $\mathbb{E}_{x'z'}[W_{1x'z'}|x, z] < \mathbb{E}_{x'z'}[W_{2x'z'}|x, z]$, which equals $W^*_1(x, z) < W^*_2(x, z)$. But this is a contradiction to the fact that $W^*(x, z)$ is unique, which establishes uniqueness of the target wage $w^*(x, z)$.

Randomization over increase and decrease. A firm never chooses a lottery to randomize over a wage increase and a wage decrease at the same time. If the lottery is degenerate the result holds directly, so we focus on non-degenerate lotteries. In that case the first order conditions with respect to π_i must be equal to zero (otherwise we are at a corner solution, which is the degenerate case). Combining the first order conditions for $i = 1, 2$ gives:

$$\beta\tilde{p}(x, W_1)\mathbb{E}_{x'z'}[J(x', z', W_{1x'z'})|x, z] + \lambda\tilde{r}(x, W_1) =$$
$$\beta\tilde{p}(x, W_2)\mathbb{E}_{x'z'}[J(x', z', W_{2x'z'})|x, z] + \lambda\tilde{r}(x, W_2),$$

or in reordered form:

$$\beta\tilde{p}(x, W_1)\mathbb{E}_{x'z'}[J(x', z', W_{1x'z'})|x, z] - \beta\tilde{p}(x, W_2)\mathbb{E}_{x'z'}[J(x', z', W_{2x'z'})|x, z] =$$
$$\lambda\left[\tilde{r}(x, W_2) - \tilde{r}(x, W_1)\right].$$

Now, suppose that the randomization yields two expected profits of opposite sign for the firm, i.e. in realization 1 expected profits are positive and in realization 2 negative. The left hand side in the above expression is then positive. For the right hand side to be positive, it must hold that $W_2 > W_1$ because $\tilde{r}$ is increasing in W_i and because the Lagrange multiplier λ is positive. However, Proposition A1 states that the wage will move according to the sign of the expected profit, and so $w'_1 > w'_2$. Lemma A1 further shows that at each (x', z') a higher wage implies a higher value $W_{1x'z'} > W_{2x'z'}$. Taking expectations it must be that $W_1 > W_2$, which is a contradiction. So, firms never randomize over wage increases and decreases at the same time. Instead, wages move according

to the sign of expected profits independent of the randomization.

Overshooting w^*. The next period wage does not overshoot the target wage w^*, which we show by contradiction. Suppose the wage in lottery realization 1 w_1' overshoots w^* from below, i.e. lottery outcome 1 yields a wage increase and next period's wage is higher than the target wage, and assume that in lottery realization 2 the wage increases exactly up to the target wage: $w_1' > w_2' = w^*$. The associated expected firm profits are then such that they are positive in the first realization and exactly zero in the other:

$$\mathbb{E}_{x'z'}[J(x', z', W_{1x'z'})|x, z] > \mathbb{E}_{x'z'}[J(x', z', W_{2x'z'})|x, z] = 0$$

Using $w_1' > w_2'$ and concavity of the utility function implies:

$$\frac{1}{u'(w_1')} > \frac{1}{u'(w_2')}$$

$$-J'(x', z', W_{1x'z'}) > -J'(x', z', W_{2x'z'})$$

$$W_{1x'z'} > W_{2x'z'}$$

$$\mathbb{E}_{x'z'}[J(x', z', W_{1x'z'})|x, z] < \mathbb{E}_{x'z'}[J(x', z', W_{2x'z'})|x, z] = 0,$$

where the transformations use Lemma A1 and the fact that J is a decreasing and concave function. The result contradicts our initial supposition and so wages do not overshoot the efficiency wage. Further, it is impossible that both w_1' and w_2' overshoot w^* from below because this would imply negative expected profits and contradicts that wages grow, so overshooting in both lotteries from below is also ruled out. Similarly, one can make a similar argument for overshooting from above.

This result has two additional implications: First, it proves that if one lottery outcome gives a higher continuation value to the firm it is accompanied by a lower wage increase. Second, if a firm pays the efficiency wage the continuation values in both lottery outcomes are zero.

Monotonicity in z. The final step is to show that the efficiency wage $w^*(x, z)$ increases in z. Recall that $J(x, z, V)$ is increasing in z and decreasing and concave in V but that the target wage for a firm currently in state (x, z, V) is not a function of V. Let's consider $z_1 < z_2$ such that $w^*(x, z_1) \leq w^*(x, z_2)$ needs to be shown. Call ξ_1 the optimal policy for $J(x, z_1, V_1)$ where V_1 delivers $w^*(x, z_1)$. Assume a firm in state (x, z_2, V_1) now adopts the same policy ξ_1, namely it pays $w^*(x, z_1)$ to a worker who receives V_1. The firm makes more profits than if it was at z_1 because $f(x, z)$ is increasing in z and its continuation value is larger as well due to the monotonicity in $g(z, \nu)$. However, the optimal policy at z_2 is to pay a higher wage $w_2' \geq w^*(x, z_1)$ to trade some output for a longer expected lifespan. At the same time, the optimal wage retains a positive continuation value such that $w_2' \leq w^*(x, z_2)$. This implies that $w^*(x, z_1) \leq w^*(x, z_2)$ and concludes. $\qquad\square$

A1.3   Solving the model

The main difficulty resides in solving the firm's problem because directly tackling BE-F requires finding the promised utilities $W_{ix'z'}$ in each state of the world for the next period. This becomes infeasible as soon as reasonable supports are considered for x and z. Therefore, instead of solving BE-F directly, we solve the following Pareto problem:

$$\mathcal{P}(x, z, \rho) = \inf_{\omega_i} \sup_{\pi_i, w_i, W_i \geq \underline{W}(x)} \sum_{i=1,2} \pi_i \Big(f(x, z) - w_i + \rho \big(u(w_i) + \tilde{r}(x, W_i) \big)$$
$$- \beta \omega_i \tilde{p}(x, W_i) W_i + \beta \tilde{p}(x, W_i) \mathbb{E}_{x'z'}[\mathcal{P}(x', z', \omega_i) | x, z] \Big)$$

where

$$\mathcal{P}(x, z, \rho) \equiv \sup_v J(x, z, v) + \rho v.$$

The following proof establishes its equivalence with the original problem. It exploits that the first order condition with respect to W_i reveals that the utilities promised in different future states are linked to each other.

Proof. We have the following recursive formulation for J:

$$J(x, z, V) = \max_{\pi_i, w_i, W_i, W_{ix'z'}} \sum_{i=1,2} \pi_i \Big(f(x, z) - w_i + \beta \tilde{p}(x, W_i) \mathbb{E}_{x'z'}[J(x', z', W_{ix'z'})|x, z] \Big)$$

$$s.t \quad (\lambda) \quad V = \sum_{i=1,2} \pi_i \Big(u(w_i) + \tilde{r}(x, W_i) \Big),$$

$$(\omega_i) \quad W_i = \mathbb{E}_{x'z'}[W_{ix'z'}|x, z].$$

Consider the Pareto problem

$$\mathcal{P}(x, z, \rho) = \sup_v J(x, z, v) + \rho v,$$

for which a recursive formulation can be constructed as follows. We initially substitute the definition of J together with its constraints into $\mathcal{P}$ and get:

$$\mathcal{P}(x, z, \rho) = \sup_{V, \pi_i, w_i, W_i, W_{ix'z'}} \sum_{i=1,2} \pi_i \Big(f(x, z) - w_i + \beta \tilde{p}(x, W_i) \mathbb{E}_{x'z'}[J(x', z', W_{ix'z'})|x, z] \Big) + \rho V$$

$$s.t \quad (\lambda) \quad V = \sum_{i=1,2} \pi_i \Big(u(w_i) + \tilde{r}(x, W_i) \Big),$$

$$(\omega_i) \quad W_i = \mathbb{E}_{x'z'}[W_{ix'z'}|x, z].$$

At this point we can substitute in the promise-keeping constraint:

$$\mathcal{P}(x, z, \rho) = \sup_{\pi_i, w_i, W_i, W_{ix'z'}} \sum_{i=1,2} \pi_i \Big(f(x, z) - w_i + \beta \tilde{p}(x, W_i) \mathbb{E}_{x'z'}[J(x', z', W_{ix'z'})|x, z]$$

$$+ \rho \Big(u(w_i) + \tilde{r}(x, W_i) \Big) \Big)$$

$$s.t \quad (\omega_i) \quad W_i = \mathbb{E}_{x'z'}[W_{ix'z'}|x, z].$$

For reasons that will become clear in the next step, we split the case where the worker potentially separates from the case where the match survives with certainty. To that end we define $\underline{W}(x)$ as the value such that $\tilde{p}(x, \underline{W}(x)) = 0$.

A11

As a consequence, the Pareto problem always delivers at least the value that promises $\underline{W}(x)$. We then rewrite the Pareto problem as:

$$\mathcal{P}(x, z, \rho) = \max\{\mathcal{P}_{01}(x, z, \rho), \mathcal{P}_{11}(x, z, \rho)\}$$

where $\mathcal{P}_{01}(x, z, \rho)$ uses $W_1 = \underline{W}(x)$ in the first outcome of the lottery but $W_2 > \underline{W}(x)$ in the second realization, while $\mathcal{P}_{11}(x, z, \rho)$ refers to promised values $W_1, W_2 > \underline{W}(x)$. The case in which the match discontinues with certainty is subsumed under $\mathcal{P}_{01}$ because the lottery can be assumed to be degenerate with $\pi_1 = 1$.

First, $\mathcal{P}_{11}$ can be written as:

$$\mathcal{P}_{11}(x, z, \rho) = \sup_{\pi_i, w_i, W_i > \underline{W}(x), W_{ix'z'}} \sum_{i=1,2} \pi_i \Big(f(x, z) - w_i$$
$$+ \beta \tilde{p}(x, W_i) \mathbb{E}_{x'z'}[J(x', z', W_{ix'z'})|x, z] + \rho\Big(u(w_i) + \tilde{r}(x, W_i)\Big)\Big)$$
$$s.t \quad (\omega_i) \quad W_i = \mathbb{E}_{x'z'}[W_{ix'z'}|x, z].$$

We introduce the ω_i-constraints in the optimization with weight $\beta \omega_i \tilde{p}(x, W_i)$ where $\tilde{p}(x, W_i) > 0$ since $W_i > \underline{W}(x)$:

$$\mathcal{P}_{11}(x, z, \rho) = \inf_{\omega_i} \sup_{\pi_i, w_i, W_i > \underline{W}(x), W_{ix'z'}} \sum_{i=1,2} \pi_i \Big(f(x, z) - w_i + \rho\Big(u(w_i) + \tilde{r}(x, W_i)\Big)\Big)$$
$$- \beta \omega_i \tilde{p}(x, W_i)(W_i - \mathbb{E}_{x'z'}[W_{ix'z'}|x, z])$$
$$+ \beta \tilde{p}(x, W_i) \mathbb{E}_{x'z'}[J(x', z', W_{ix'z'})|x, z]\Big),$$

and combine the terms to get:

$$\mathcal{P}_{11}(x, z, \rho) = \inf_{\omega_i} \sup_{\pi_i, w_i, W_i > \underline{W}(x), W_{ix'z'}} \sum_{i=1,2} \pi_i \Big(f(x, z) - w_i + \rho\Big(u(w_i) + \tilde{r}(x, W_i)\Big)\Big)$$
$$- \beta \omega_i \tilde{p}(x, W_i)W_i + \beta \tilde{p}(x, W_i) \mathbb{E}_{x'z'}[J(x', z', W_{ix'z'}) + \omega_i W_{ix'z'}|x, z]\Big).$$

A12

The final step is to split the sup

$$\mathcal{P}_{11}(x,z,\rho) = \inf_{\omega_i} \sup_{\pi_i,w_i,W_i > \underline{W}(x)} \sum_{i=1,2} \pi_i \Big(f(x,z) - w_i + \rho\big(u(w_i) + \tilde{r}(x,W_i)\big)$$

$$- \beta\omega_i\tilde{p}(x,W_i)W_i + \beta\tilde{p}(x,W_i)\mathbb{E}_{x'z'}\left[\sup_{W_{ix'z'}} J(x',z',W_{ix'z'}) + \omega_i W_{ix'z'}\Big|x,z\right]\Big),$$

and to use the definition for $\mathcal{P}$:

$$\mathcal{P}_{11}(x,z,\rho) = \inf_{\omega_i} \sup_{\pi_i,w_i,W_i > \underline{W}(x)} \sum_{i=1,2} \pi_i \Big(f(x,z) - w_i + \rho\big(u(w_i) + \tilde{r}(x,W_i)\big)$$

$$- \beta\omega_i\tilde{p}(x,W_i)W_i + \beta\tilde{p}(x,W_i)\mathbb{E}_{x'z'}[\mathcal{P}(x',z',\omega_i)|x,z]\Big).$$

Second, turning to $\mathcal{P}_{01}$:

$$\mathcal{P}_{01}(x,z,\rho) = \sup_{\pi_i,w_i,W_2 > \underline{W}(x),W_{ix'z'}} \pi_1 \Big(f(x,z) - w_1 + \beta\tilde{p}(x,\underline{W}(x))\mathbb{E}_{x'z'}[J(x',z',W_{1x'z'})|x,z]$$

$$+ \rho\big(u(w_1) + \tilde{r}(x,\underline{W}(x))\big)\Big)$$

$$+ \pi_2\Big(f(x,z) - w_2 + \beta\tilde{p}(x,W_2)\mathbb{E}_{x'z'}[J(x',z',W_{2x'z'})|x,z]$$

$$+ \rho\big(u(w_2) + \tilde{r}(x,W_2)\big)\Big)$$

$$s.t \quad (\omega_i) \quad W_i = \mathbb{E}_{x'z'}[W_{ix'z'}|x,z],$$

We can use $\tilde{p}(x,\underline{W}(x)) = 0$ and $\tilde{r}(x,\underline{W}(x)) = \beta\mathbb{E}_{x'}[U(x')|x]$. Hence:

$$\mathcal{P}_{01}(x,z,\rho) = \sup_{\pi_i,w_i,W_2 > \underline{W}(x),W_{ix'z'}} \pi_1 \Big(f(x,z) - w_1 + \rho\big(u(w_1) + \beta\mathbb{E}_{x'}[U(x')|x]\big)\Big)$$

$$+ \pi_2\Big(f(x,z) - w_2 + \beta\tilde{p}(x,W_2)\mathbb{E}_{x'z'}[J(x',z',W_{2x'z'})|x,z]$$

$$+ \rho\big(u(w_2) + \tilde{r}(x,W_2)\big)\Big)$$

$$s.t \quad (\omega_i) \quad W_i = \mathbb{E}_{x'z'}[W_{ix'z'}|x,z],$$

where the choice variables W_1 and $W_{1x'z'}$ disappear and so does the constraint

A13

associated with ω_1. We apply the same treatment as in the case of $\mathcal{P}_{11}$ to get:

$$\mathcal{P}_{01}(x, z, \rho) = \inf_{\omega_2} \sup_{w_i, W_2 > \underline{W}(x), W_{2x'z'}} \pi_1 \Big(f(x, z) - w_1 + \rho \Big(u(w_1) + \beta \mathbb{E}_{x'}[U(x')|x] \Big) \Big)$$
$$+ \pi_2 \Big(f(x, z) - w_2 + \rho \Big(u(w_2) + \tilde{r}(x, W_2) \Big) \Big)$$
$$- \beta \omega_2 \tilde{p}(x, W_2) W_2 + \beta \tilde{p}(x, W_2) \mathbb{E}_{x'z'}[\mathcal{P}(x', z', \omega_2)|x, z] \Big).$$

Finally, notice that using $W_1 = \underline{W}(x)$ is simply relaxing the strict constraint on W_1 in $\mathcal{P}_{11}$, in which case ω_1 becomes indeterminate but also irrelevant, and we can continue to minimize with respect to it. Combining the two options in a single expression yields:

$$\mathcal{P}(x, z, \rho) = \inf_{\omega_i} \sup_{\pi_i, w_i, W_i \geq \underline{W}(x)} \sum_{i=1,2} \pi_i \Big(f(x, z) - w_i + \rho \Big(u(w_i) + \tilde{r}(x, W_i) \Big) \Big)$$
$$- \beta \omega_i \tilde{p}(x, W_i) W_i + \beta \tilde{p}(x, W_i) \mathbb{E}_{x'z'}[\mathcal{P}(x', z', \omega_i)|x, z] \Big).$$

From the solution of this recursive problem we can reconstruct the lifetime utility of the worker V at a given (x, z, ρ) from the partial derivative of $\mathcal{P}$ with respect to ρ:

$$V(x, z, \rho) = \mathcal{P}'(x, z, \rho),$$

or inversely, at state (x, z, V) let $\rho^*(x, z, V)$ be the solution to the previous equation:

$$V = \mathcal{P}'(x, z, \rho^*(x, z, V)).$$

The profit function of the firm can then be expressed as:

$$J(x, z, V) \quad = \quad \mathcal{P}(x, z, \rho^*(x, z, V)) - \rho^*(x, z, V)V.$$

$\square$

A2 Data appendix

A2.1 Data sources

We rely on the raw matched employer-employee data set constructed in Friedrich, Laun, Meghir, and Pistaferri (2019) that combines information from three different data sources made available by The Institute for Evaluation of Labour Market and Education Policy (IFAU).[14] The reader can refer to their paper for additional information.

The first source is the Longitudinal Database on Education, Income and Employment (LOUISE) that provides data on demographic and socioeconomic variables for the entire working age population in Sweden from 1990 to today. We use information about age and gender.

The second data set is the Register-Based Labor Market Statistics (RAMS), which tracks the universe of employment spells in Sweden from 1985 onward. RAMS includes the gross yearly earnings and the first and last remunerated month for each employee-firm spell, as well as firm identifiers at the Corporate Registration Number level. On the firm side, RAMS records information about the industry and the type of legal entity for all firms with employees.

Finally, the Structural Business Statistics (SBS) is the third data source and contains accounting and balance sheet information for all non-financial corporations in Sweden from 1997 onward. Notably the data reports a measure of value added at the firm and year level in the variable named FORBRUKN-INGSVARDE as well as the reported employment size. See Table A1 for how this variable is constructed.

All monetary variables are adjusted for inflation (detrended with the CPI) and to construct firm productivity, we remove broad industries interacted with

---

[14]A special thanks to Benjamin Friedrich, Lisa Laun, Costas Meghir and Luigi Pistaferri for their help and to the IFAU for their continuous support.

yearly time dummies.

Table A1: Construction of value added: FORBRUKNINGSVARDE

sign	variable description and name
+	Raw materials (at VE level)
+	Other external costs (at the VE level)
+	Social costs and other costs
+	Other operating expenses
−	Losses on receivables
−	Other consumable equipment with a life expectancy of more than one year
−	Costs for travel and hotel mediated
−	Ground rent /RENT
−	Other costs in other external costs not counted as consumption
−	Severance pay
−	Pension payments
−	Received grants and allowances for staff
−	Compulsory social contributions
−	Wage Taxes
−	Other charges
−	Pension Provisions
−	Pension insurance premiums, etc.
−	Other costs in other operating costs not counted as consumption
−	Received contributions accounted for as cost reduction
−	Foreign exchange losses on claims and liabilities relating to operations
−	Profit / loss on disposal of tangible and intangible assets
−	Abandoned / redeemed shareholder contributions (Rest rorkost, v0139)
−	Group contribution (Rest rorkost, v0139)
−	Income Shares in partnerships and limited partnerships (Rest rorkost, v0139)
−	Profit / loss on disposal of shares (Rest rorkost, v0139)

A2.2   Sample construction

Our analysis focuses on the period 2001-2006. The sample includes all firms whose legal entity is either a limited partnership or limited company other than banking and insurance companies. We inherit two restrictions applied to the original data construction, namely that spells with monthly earnings below 3,416 in 2008 Swedish krona as well as spells spanning less than two months of employment (i.e. if the start is the same as the end month) are excluded.

In order to abstract from labor force participation, we focus exclusively on men in the age range between 20 and 50. Indeed, both women in their 30s and men after age 50 appear to show participation and earnings shifts. All self-

Table A2: Data description

Number of year observations	5,599,375
Number of year observations with 12 months worked	3,463,405
Number of unique workers	1,158,954
Number of unique firms	72,767
Employment share	0.86
Mean log earnings among full-year observations	12.65
Variance of log earnings among full-year observations	0.14

employed workers are dropped from the original sample, but we include active and non-active job seekers to account better for mobility in and out of work. On the employer side, we restrict the sample to firms with positive reported value added. For these firms we construct a measure of value added per worker by dividing the value added measure (FORBRUKNINGSVARDE) by the reported firm size. We denote this variable as y_{jt} for employer j in year t.

Quarterly employment status. We aggregate the data to quarterly frequency in order to compute transition rates. For individuals with multiple jobs during a quarter we keep the main employment, defined as the employment that accounts for the largest share of quarterly earnings. We define a worker as employed if he is working at least 2 months for any employer during the quarter. The quarterly data has the set of columns (i, q, j_{iq}), where q counts time at the quarterly frequency and $j_{iq} = 0$ if individual i does not have any employment records in quarter q.

Full-year earnings. For all moments relying on earnings and value added, we further focus on full-year employment spells, i.e. spells for which the data reports 4 quarters of employment with the same firm. The earnings and value added data contain the set of columns $(i, t, j_{it}, w_{it}, y_{it})$, where t counts time at the yearly frequency, $j_{it} = 0$ if individual i does not have any full-year employment record in period t, w_{it} are earnings and $y_{it} = y_{j_{it},t}$.

A3 Counterfactuals appendix

A3.1 First best

We start with the original firm problem:

$$J(x, z, V) = \max_{\pi_i, w_i, W_i, W_{ix'z'}} \sum_{i=1,2} \pi_i \Big(f(x, z) - w_i + \beta \tilde{p}(x, W_i) \mathbb{E}_{x'z'}[J(x', z', W_{ix'z'})|x, z] \Big)$$

$$s.t \quad (\lambda) \quad V = \sum_{i=1,2} \pi_i \Big(u(w_i) + \tilde{r}(x, W_i) \Big),$$

$$(\omega_i) \quad W_i = \mathbb{E}_{x'z'}[W_{ix'z'}|x, z].$$

However, in the first best case, where the firm is no longer constrained by incentive constraints, it can dictate both the worker's effort and search decisions. So, we substitute out $\tilde{r}(\cdot)$ and $\tilde{p}(\cdot)$ and solve the following dynamic problem instead:

$$J^{fb}(x, z, V) = \max_{\pi_i, w_i, W_i, W_{ix'z'}, e_i, v_i} \sum_{i=1,2} \pi_i \Big(f(x, z) - w_i$$

$$+ \beta(1 - \delta(e_i))(1 - \kappa p(\theta(x, v_i))) \mathbb{E}_{x'z'}[J^{fb}(x', z', W_{ix'z'})|x, z] \Big)$$

$$s.t \quad (\lambda) \quad V = \sum_{i=1,2} \pi_i \Big(u(w_i) - c(e_i) + \delta(e_i)\beta \mathbb{E}_{x'}[U(x')|x]$$

$$+ \beta \kappa(1 - \delta(e_i))p(\theta(x, v_i))(v_i - W_i) + \beta(1 - \delta(e_i))W_i \Big),$$

$$(\omega_i) \quad W_i = \mathbb{E}_{x'z'}[W_{ix'z'}|x, z].$$

We define the first best Pareto problem:

$$\mathcal{P}^{fb}(x, z, \rho) = \sup_v J^{fb}(x, z, v) + \rho v,$$

and derive, analogously to the solution strategy of the baseline model (see Appendix A1.3), the following Bellman equation:

$$\mathcal{P}^{fb}(x,z,\rho) = \inf_{\omega_i} \sup_{\pi_i,w_i,W_i \geq \underline{W}(x),v_i,e_i} \sum_{i=1,2} \pi_i \Big(f(x,z) - w_i$$

$$+ \rho \Big(u(w_i) - c(e_i) + \delta(e_i)\beta\mathbb{E}_{x'}[U(x')|x]$$

$$+ \beta\kappa(1 - \delta(e_i))p(\theta(x,v_i))(v_i - W_i) + \beta(1 - \delta(e_i))W_i \Big)$$

$$- \beta\omega_i(1 - \delta(e_i))(1 - \kappa p(\theta(x,v_i)))W_i$$

$$+ \beta(1 - \delta(e_i))(1 - \kappa p(\theta(x,v_i)))\mathbb{E}_{x'z'}[\mathcal{P}^{fb}(x',z',\omega_i)|x,z] \Big).$$

Now note that the first order conditions for w_i and W_i, the envelope condition of J^{fb} and the definition of $\mathcal{P}^{fb}(x,z,\rho)$ deliver $\lambda = \rho = \omega_i$. The Pareto problem thus simplifies to:

$$\mathcal{P}^{fb}(x,z,\rho) = \sup_{\pi_i,w_i,v_i,e_i} \sum_{i=1,2} \pi_i \Big(f(x,z) - w_i$$

$$+ \rho \Big(u(w_i) - c(e_i) + \delta(e_i)\beta\mathbb{E}_{x'}[U(x')|x] + \beta\kappa(1 - \delta(e_i))p(\theta(x,v_i))v_i \Big)$$

$$+ \beta(1 - \delta(e_i))(1 - \kappa p(\theta(x,v_i)))\mathbb{E}_{x'z'}[\mathcal{P}^{fb}(x',z',\rho)|x,z] \Big).$$

Solving for $\mathcal{P}^{fb}(x,z,\rho)$ is easier than solving for $J^{fb}(x,z,V)$ directly because we can use an exogenous grid for all state variables, including ρ, and reduce the number of maximizers. Nevertheless, $J^{fb}(x,z,V)$ can be subsequently recovered from $\mathcal{P}^{fb}(x,z,\rho)$ through:

$$J^{fb}(x,z,V) \quad = \quad \mathcal{P}^{fb}(x,z,\rho^*(x,z,V)) - \rho^*(x,z,V)V,$$

where $\rho^*(x,z,V)$ is defined as the ρ that equates the partial derivative of $\mathcal{P}^{fb}$ with the promised value V.

Figure A1: Average impulse response to x change at first best
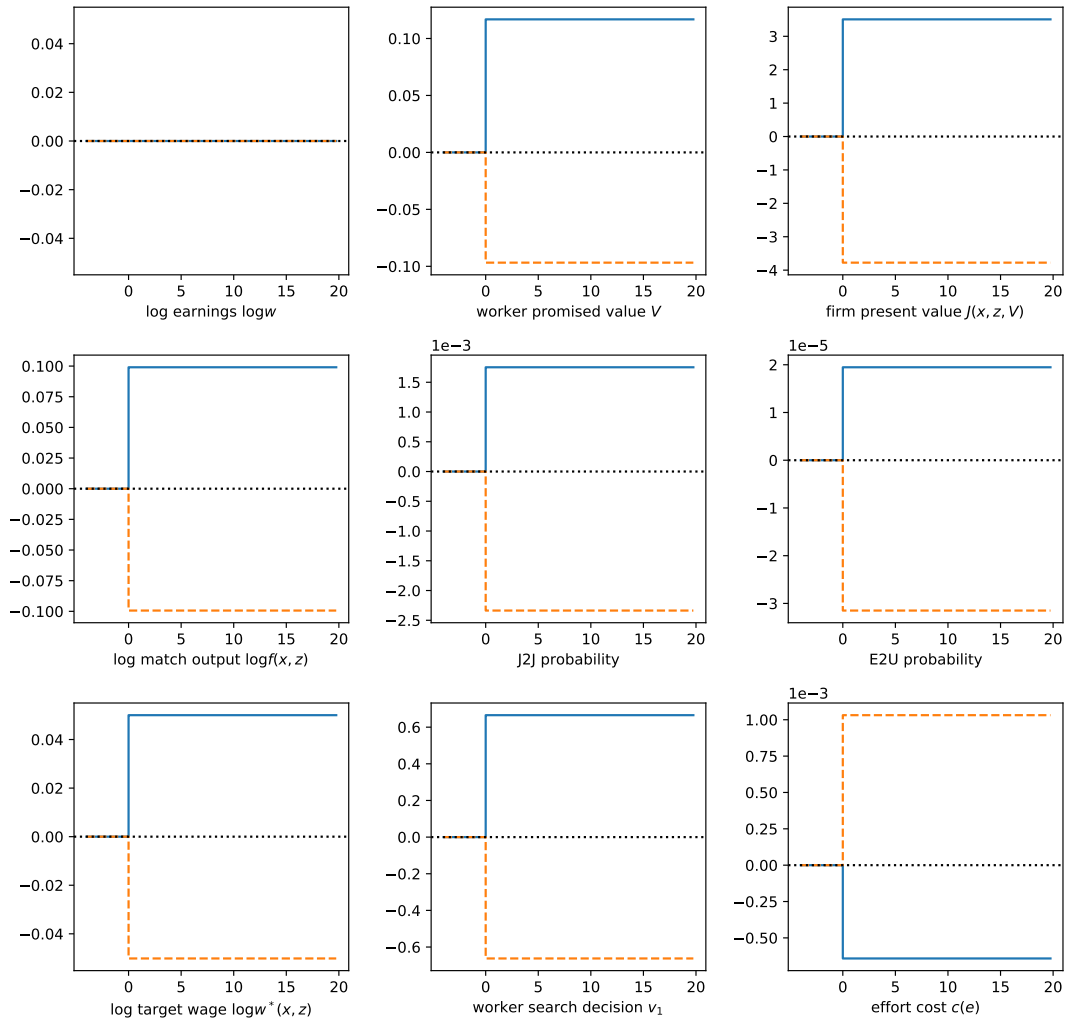
Notes: Effect of a positive (solid blue) and negative (dashed red) permanent x shock over time (years) at first best. Starting (x, z) values are drawn from the stationary distribution. Initial wages are target wages. Separation is ruled out.

Figure A2: Average impulse response to z change at first best
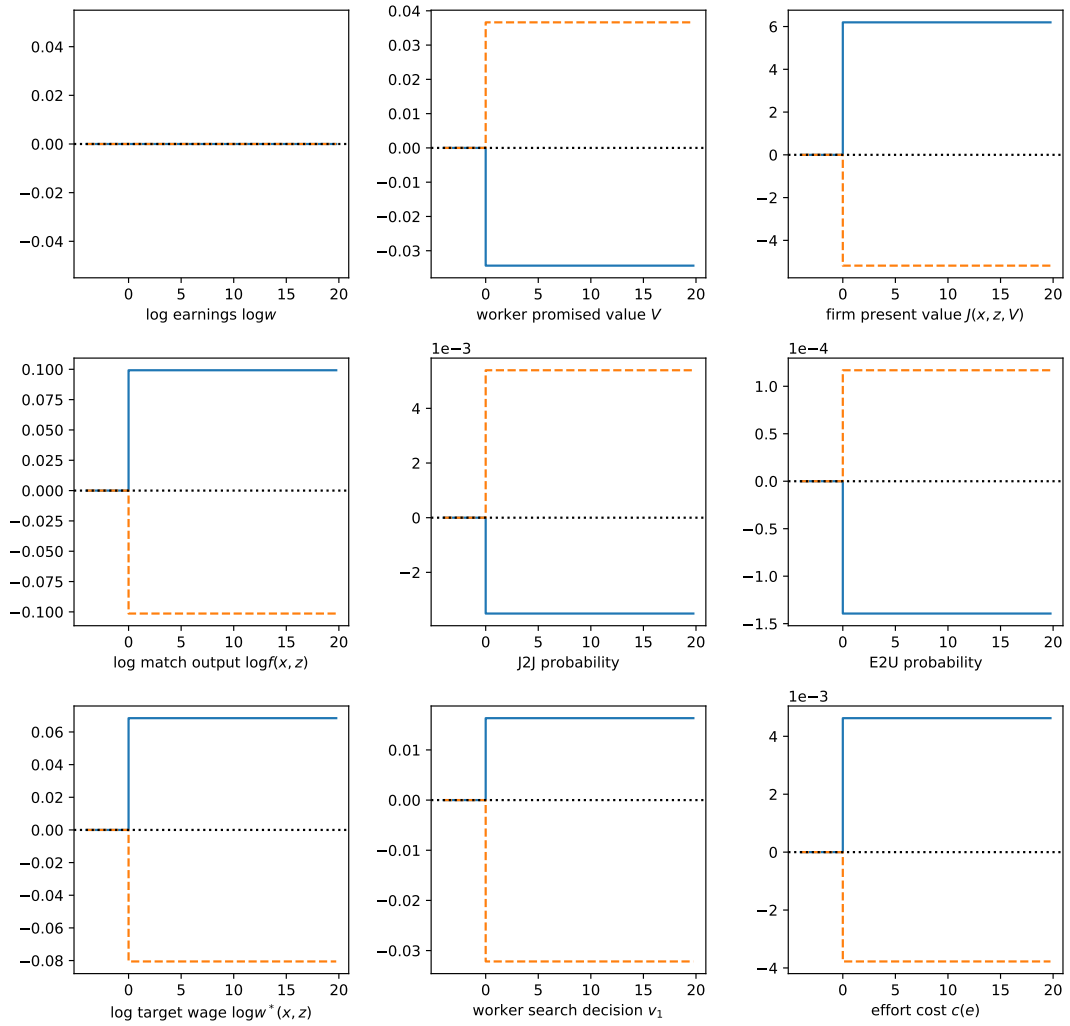
Notes: Effect of a positive (solid blue) and negative (dashed red) permanent z shock over time (years) at first best. Starting (x, z) values are drawn from the stationary distribution. Initial wages are target wages. Separation is ruled out.

A3.2 Passthrough analysis

Let us start describing the passthrough analysis by defining three outcome variables in each state (x, z, V). The first outcome of interest is a wage equivalent of the present value of the worker's utility, defined as:

$$w^{\mathrm{EQV}}(V) = u^{-1}((1 - \beta)V).$$

Next, we define an expected present value (EPV) of transfers, which includes all future wages w and benefits b paid to the worker. We do this for two sets of histories, one where we follow the worker, and one where we force the current match to continue to exist. To be precise, using the equilibrium policies $(w_i^*, e_i^*, v_0^*, v_{1i}^*, W_{ix'z'}^*) \in \xi$, we define our second outcome of interest as the solution to the following recursive equations:

$$
\begin{aligned}
b^{\mathrm{EPV}}(x) =& (1 - \beta)b + \beta p(\theta(x, v_0^*))\mathbb{E}_{x'}[w^{\mathrm{EPV}}(x', z_0, v_0^*)|x] \\
& + \beta(1 - p(\theta(x, v_0^*)))\mathbb{E}_{x'}[b^{\mathrm{EPV}}(x')|x] \\
w^{\mathrm{EPV}}(x, z, V) =& (1 - \beta)w_i^* + \beta\delta(e_i^*)\mathbb{E}_{x'}[b^{\mathrm{EPV}}(x')|x] \\
& + \beta(1 - \delta(e_i^*))p(\theta(x, v_{1i}^*))\mathbb{E}_{x'}[w^{\mathrm{EPV}}(x', z_0, v_{1i}^*)|x] \\
& + \beta(1 - \delta(e_i^*))(1 - p(\theta(x, v_{1i}^*)))\mathbb{E}_{x'z'}[w^{\mathrm{EPV}}(x', z', W_{ix'z'})|x, z],
\end{aligned}
$$

where transitions are guided by the equilibrium policies. Finally, the third outcome is defined in the same way, except that the match is forced to last:

$$
w^{\mathrm{EPV\text{-}EE}}(x, z, V) = (1 - \beta)w_i^* + \beta\mathbb{E}_{x'z'}[w^{\mathrm{EPV\text{-}EE}}(x', z', W_{ix'z'})|x, z].
$$

Turning to our two measures of output change from the underlying productivity shock, we define on the one hand $f^{\mathrm{EPV}}(x, z, V)$ exactly like $w^{\mathrm{EPV}}(x, z, V)$, and on the other hand $f^{\mathrm{EPV\text{-}EE}}(x, z, V)$ exactly like $w^{\mathrm{EPV\text{-}EE}}(x, z, V)$, where in both cases w_i^* is replaced with $f(x, z)$.

Consider a shock that shifts the current productivity of an individual in state

(x, z, V) to (x^1, z^1). We can then compute the difference between the values at $V^0 = W^*_{ix^0 z^0}(x, z, V)$ and $V^1 = W^*_{ix^1 z^1}(x, z, V)$, where $W^*_{ix' z'} \in \xi$ is evaluated at the initial state (x, z, V) and it holds that $(x^0, z^0) = (x, z)$. This allows us to compare the present value in both realizations of the shock, (x^0, z^0) and (x^1, z^1), precisely at the point where the firm provides insurance. Using V instead of $V^0 = W^*_{ix^0 z^0}(x, z, V)$ would include backloading, rather than strictly look at the effect of the shock. We report the passthrough as the average over individuals across states (x, z, V) taken from the stationary equilibrium in the economy.

Our preferred definition of the passthrough is the average of the ratio of the effect on the log wage equivalent $\log w^{\mathrm{EQV}}(V)$ to the change in the log productivity change in the match $\log f^{\mathrm{EPV\text{-}EE}}$. We write:

$$\mathbb{E}\left[\frac{\Delta \log w^{\mathrm{EQV}}(V)}{\Delta \log f^{\mathrm{EPV\text{-}EE}}(x, z, V)}\right]$$

Notably, in a simple unit root process with constant passthrough of a permanent shock and with log utility, this yields the same passthrough parameter reported in conventional decompositions. Too see this, consider $\log y_{it} = \log y_{it-1} + \mu_{it}$ and $\log w_{it} = \log w_{it-1} + \gamma \mu_{it} + u_{it}$. For simplicity, abstract from separations and let u_{it} and μ_{it} be i.i.d. random normal draws. The expected present value with log utility for a given value of μ satisfies $V(\mu) = V(0) + \frac{\gamma \mu}{1-\beta}$ and hence $\Delta \log w^{\mathrm{EQV}}(V) = \gamma \mu$. Similarly for the productivity we get that $\Delta \log f^{\mathrm{EPV\text{-}EE}} = \mu$, even in present value because the μ shock is permanent. This results in a passthrough value of γ and hence lines up with the conventional definition.

Note that the parameter γ in this simple joint process of $\log y_{it}$ and $\log w_{it}$ can be recovered by adapting the estimator of Guiso, Pistaferri, and Schivardi (2005) to the case where processes are unit root with i.i.d. measurement error.

It is given by:

$$\gamma = \frac{\text{Cov}_{S^{\text{EE}}}\left[\Delta \log w_{it}, \Delta \log y_{it}\right]}{\text{Var}_{S^{\text{EE}}}\left[\Delta \log y_{it}\right] + 2\text{Cov}_{S^{\text{EEE}}}\left[\Delta \log y_{it}, \Delta \log y_{it-1}\right]}.$$

Computing this ratio using the data moments in Table 1 gives an estimate for γ of 4.5%. This captures how much of a permanent shock to value added per worker is transmitted to worker's earnings. It appears to be of the same order of magnitude as the passthrough reported for the value added per worker in Guiso, Pistaferri, and Schivardi (2005) of 7.8% in column 7 of Table 8.

Our approach in terms of expected present values has the advantage of being independent of the functional form imposed on the process of w_{it} and y_{it}.