Online Appendix for "Productivity Shocks, Long-Term Contracts and Earnings Dynamics"

Neele Balke*, Thibaut Lamadon†

September 25, 2020

*Kenneth C. Griffin Dept. of Economics, University of Chicago, nbalke@uchicago.edu.
†Kenneth C. Griffin Dept. of Economics, University of Chicago, lamadon@uchicago.edu, IFAU, IFS and NBER.

W1 Model web appendix

W1.1 Properties of equilibrium functions

In this section, we define the set $\mathbb{J}$ of profit functions of the firm and then, taking an arbitrary $J \in \mathbb{J}$ as given, derive properties of the market tightness, job finding probability, and search and effort policy functions in equilibrium.

Definition W1 (Definition of $\mathbb{J}$)**.** *Let $\mathbb{J}$ be defined as the set of firms' value functions $J : \mathbb{S} \times \mathbb{V} \to \mathbb{R}$ such that*

(J1) For all $(x, z) \in \mathbb{S}$ and all $V_1, V_2 \in \mathbb{V}$ with $V_1 \leq V_2$, the difference $J(x, z, V_2) - J(x, z, V_1)$ is bounded by $-\overline{B}_J(V_2 - V_1)$ and $-\underline{B}_J(V_2 - V_1)$ where $\overline{B}_J \geq \underline{B}_J > 0$ are some constants.

(J2) For all $(x, z, V) \in \mathbb{S} \times \mathbb{V}$, $J(x, z, V)$ is bounded in $[\underline{J}, \overline{J}]$ where $\overline{J} = \frac{f(\overline{x},\overline{z}) - u^{-1}\left(\underline{v} + c(\underline{e}) - \beta\overline{v}\right)}{1-\beta}$ and $\underline{J} = \frac{f(\underline{x},\underline{z}) - u^{-1}\left(\overline{v} + c(\overline{e}) - \beta\underline{v}\right)}{1-\beta}$.

(J3) For all $(x, z) \in \mathbb{S}$, $J(x, z, V)$ is concave in V.

(J4) For all $(x, z) \in \mathbb{S}$, $J(x, z, V)$ is differentiable in V.

Lemma W1 (Uniqueness of θ)**.** *The market tightness function $\theta(x, v)$ is unique in equilibrium.*

Proof of Lemma W1. Consider the firm value function, rewritten as:

$$J(x, z, V) = \max_{\pi_i, w_i, W_i} \sum_{i=1,2} \pi_i \Big(f(x, z) - w_i + \beta\tilde{p}(x, W_i) M(x, z, W_i) \Big)$$

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

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

where

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

Free entry can then be expressed as:

$$\Pi(x,v) = q(\theta(x,v))M(x,z_0,v) - k \leq 0.$$

From the assumption on $q(\theta)$ and its invertibility, as well as free entry, it follows that $\theta(x,v) = q^{-1}(k/M(x,z_0,v))$ for $M(x,z_0,v) \geq k$ (or equivalently for $v \leq \tilde{v}(x)$, where $\tilde{v}(x)$ is the solution to $k = M(x,z_0,v)$ with respect to v) and is bounded between 0 and $\overline{\theta} \equiv q^{-1}(k/\overline{J})$. Otherwise $\theta(x,v) = 0$. Hence, the market tightness function is unique. □

Lemma W2 (θ is decreasing and continuous in v). *For all $x \in \mathbb{X}$, the market tightness function, $\theta(x,v)$, is such that*

$$\frac{\overline{B}_J}{q'\left(\overline{\theta}\right)k}(v_2 - v_1) \leq \theta(x,v_2) - \theta(x,v_1) \leq \frac{\underline{B}_J\, k}{q'(0)\,\overline{J}^2}(v_2 - v_1), \quad \text{if } v_1 \leq v_2 \leq \tilde{v}(x),$$

$$\frac{\overline{B}_J}{q'\left(\overline{\theta}\right)k}(v_2 - v_1) \leq \theta(x,v_2) - \theta(x,v_1) \leq 0, \quad \text{if } v_1 \leq \tilde{v}(x) \leq v_2,$$

$$\theta(x,v_2) - \theta(x,v_1) = 0, \quad \text{if } \tilde{v}(x) \leq v_1 \leq v_2,$$

where $\underline{B}_J$ and $\overline{B}_J$ are the bi-Lipschitz bounds on all functions in $\mathbb{J}$.

Proof of Lemma W2. We suppress the dependence of various functions on x and z to improve readability. Let x be an arbitrary point in $\mathbb{X}$, and let v_1, v_2 be two points in $\mathbb{V}$ with $v_1 \leq v_2$. First, consider the case in which $v_1 \leq v_2 \leq \tilde{v}$. In this case, the difference $\theta(x,v_2) - \theta(x,v_1)$ is equal to

$$\theta(x,v_2) - \theta(x,v_1) = q^{-1}(k/M(v_2)) - q^{-1}(k/M(v_1)) = \int_{k/M(v_1)}^{k/M(v_2)} (q^{-1})'(t)\, dt,$$

where the first equality uses Lemma W1, and the second equality uses the fact that M is decreasing in v and $M(v_1) \geq M(v_2) \geq k > 0$. For all $v \in [\underline{v}, \tilde{v}]$, the derivative of the inverse function $q^{-1}(\cdot)$ evaluated at $k/M(v)$ is equal to $1/q'(\theta(x,v)) \in \left[1/q'(\overline{\theta}), 1/q'(0)\right]$, where $1/q'(\overline{\theta}) \leq 1/q'(0) < 0$. Therefore the last term in the previous equation satisfies:

$$\frac{1}{q'(\overline{\theta})}\left(\frac{k}{M(v_2)} - \frac{k}{M(v_1)}\right) \leq \int\limits_{k/M(v_1)}^{k/M(v_2)} (q^{-1})'(t)dt \leq \frac{1}{q'(0)}\left(\frac{k}{M(v_2)} - \frac{k}{M(v_1)}\right),$$

where

$$\frac{k}{M(v_2)} - \frac{k}{M(v_1)} = \int\limits_{M(v_2)}^{M(v_1)} \frac{k}{t^2}dt.$$

For all $v \in [\underline{v}, \tilde{v}]$, $M(v)$ is strictly decreasing in v and it is bounded between $\overline{J}$ and k. Therefore, setting t in the integral on the RHS above to be either k or $\overline{J}$ gives bounds such that

$$\int\limits_{M(v_2)}^{M(v_1)} \frac{k}{t^2}dt \leq \frac{1}{k}\left[M(v_1) - M(v_2)\right] \leq \frac{\overline{B}_J}{k}\left(v_2 - v_1\right),$$

$$\int\limits_{M(v_2)}^{M(v_1)} \frac{k}{t^2}dt \geq \frac{k}{\overline{J}^2}\left[M(v_1) - M(v_2)\right] \geq \frac{\underline{B}_J k}{\overline{J}^2}\left(v_2 - v_1\right),$$

where the latter inequalities use the fact that differences in J and hence also in M are bounded as in the definition of $\mathbb{J}$. Taken together, the difference $\theta(x, v_2) - \theta(x, v_1)$ is such that

$$\frac{\overline{B}_J}{q'(\overline{\theta})k}\left(v_2 - v_1\right) \leq \theta(x, v_2) - \theta(x, v_1) \leq \frac{\underline{B}_J k}{q'(0)\overline{J}^2}\left(v_2 - v_1\right).$$

Next, consider the case in which $v_1 \leq \tilde{v} \leq v_2$. Then the difference $\theta(x, v_2) -$

$\theta(x, v_1)$ satisfies:

$$\theta(x, v_2) - \theta(x, v_1) = \theta(x, \tilde{v}) - \theta(x, v_1) \leq \frac{\underline{B}_J k}{q'(0)\overline{J}^2}(\tilde{v} - v_1) \leq 0,$$
$$\theta(x, v_2) - \theta(x, v_1) = \theta(x, \tilde{v}) - \theta(x, v_1) \geq \frac{\overline{B}_J}{q'(\overline{\theta})k}(\tilde{v} - v_1) \geq \frac{\overline{B}_J}{q'(\overline{\theta})k}(v_2 - v_1),$$

where both lines use the bounds in the previous expression and the fact that $\theta(x, \tilde{v}) = \theta(x, v_2)$.

Finally, in the case where $\tilde{v} \leq v_1 \leq v_2$, Lemma W1 implies that $\theta(x, v_1) = \theta(x, v_2) = 0$.

□

Lemma W3 (p is strictly decreasing, strictly concave and continuous in v)**.** *For all $x \in \mathbb{X}$, and all $v \in [\underline{v}, \tilde{v}(x)]$, the composite function $p(\theta(x, v))$ is strictly decreasing and strictly concave in v.*

Proof of Lemma W3. The function $p(\theta)$ is strictly increasing in θ, and $\theta(x, v)$ is strictly decreasing in v for all $v \in [\underline{v}, \tilde{v}]$. Therefore, $p(\theta(x, v))$ is strictly decreasing in v for $v \in [\underline{v}, \tilde{v}]$. In order to prove that the composite function $p(\theta(x, v))$ is strictly concave in v for $v \in [\underline{v}, \tilde{v}]$, consider arbitrary $v_1, v_2 \in [\underline{v}, \tilde{v}]$, with $v_1 \neq v_2$, and an arbitrary number $\alpha \in (0, 1)$. Let $v_\alpha = \alpha v_1 + (1 - \alpha)v_2$.

The function $M(x, z, v)$ is continuous and concave in v, which follows from the Maximum Theorem under Convexity as the two conditions that $\mathbb{E}_{x'z'}[J(x', z', W_{x'z'})|x, z]$ is concave and that the constraint is a continuous correspondence with a convex graph are satisfied (see Sundaram et al. (1996), p. 238). So, since $M(v)$ is concave in v and the function k/v is strictly convex in v, we have

$$\frac{k}{M(v_\alpha)} \leq \frac{k}{\alpha M(v_1) + (1 - \alpha)M(v_2)} < \alpha \frac{k}{M(v_1)} + (1 - \alpha)\frac{k}{M(v_2)}.$$

Since $p(q^{-1}(\cdot))$ is strictly decreasing and weakly concave, the previous inequality implies that

$$\begin{aligned} p\left(q^{-1}\left(k/M(v_\alpha)\right)\right) >& p\left(q^{-1}\left(\alpha\frac{k}{M(v_1)}+(1-\alpha)\frac{k}{M(v_2)}\right)\right) \\ \geq& \alpha p\left(q^{-1}\left(\frac{k}{M(v_1)}\right)\right)+(1-\alpha)p\left(q^{-1}\left(\frac{k}{M(v_2)}\right)\right). \end{aligned}$$

Since $q^{-1}(k/M(v))$ is equal to $\theta(x,v)$ for all $v \in [\underline{v}, \tilde{v}]$, the last inequality can be rewritten as

$$p(\theta(x,v_\alpha)) > \alpha p(\theta(x,v_1)) + (1-\alpha)p(\theta(x,v_2))$$

which establishes that $p(\theta(x,v))$ is strictly concave in v for all $v \in [\underline{v}, \tilde{v}]$. Since every concave function is continuous, $p(\theta(x,v))$ is also continuous in this range.

□

We introduce the return to search $D(x,W) \equiv \max_{v' \in \mathbb{V}} d(x,v',W)$, where $d(x,v',W) \equiv p\left(\theta(x,v')\right)(v'-W)$, which is maximized by the search policy function $m(x,W)$ with $m : \mathbb{X} \times \mathbb{V} \to \mathbb{V}$, given the market tightness function θ. Lemma 3.1 in **?** establishes that $m(x,W)$ is unique such that:

$$m(x,W) = \begin{cases} \arg\max_{v' \in \mathbb{V}} d(x,v',W) & \text{if } W < \tilde{v}(x) \\ W & \text{otherwise.} \end{cases}$$

The next lemmas establish that $D(x,W)$ is a decreasing function in W and that $m(x,W)$ is increasing in W.

Lemma W4 (D is decreasing and continuous in W, m is increasing and continuous in W)**.** *For all $x \in \mathbb{X}$ and all $W_1, W_2 \in \mathbb{V}$ with $W_1 \leq W_2$, the return to search function, D, satisfies:*

$$-\left(W_2 - W_1\right) \leq D\left(x, W_2\right) - D\left(x, W_1\right) \leq 0$$

and the search policy function, m, is such that

$$0 \leq m\left(x, W_2\right) - m\left(x, W_1\right) \leq W_2 - W_1.$$

Proof of Lemma W4. Let $W_1 \leq W_2$ be two arbitrary points in $\mathbb{V}$. Then:

$$\begin{aligned} D\left(x, W_2\right) - D\left(x, W_1\right) &\leq d(x, m(x, W_2), W_2) - d(x, m(x, W_2), W_1) \\ &\leq -p(\theta(x, m(x, W_2))(W_2 - W_1) \leq 0, \\ D\left(x, W_2\right) - D\left(x, W_1\right) &\geq d(x, m(x, W_1), W_2) - d(x, m(x, W_1), W_1) \\ &\geq -p(\theta(x, m(x, W_1))(W_2 - W_1) \geq -(W_2 - W_1), \end{aligned}$$

where the first inequality in both lines uses the fact that $D(x, W_i)$ is equal to $d(x, m(x, W_i), W_i)$ and greater than $d(x, m(x, W_{-i}), W_i)$ where $-i \neq i$ and $i, -i = 1, 2$. Thus the first part of the lemma holds.

Next, if $W_1 \geq \tilde{v}(x)$, then $m(x, W_2) = W_2$ and $m(x, W_1) = W_1$. If $W_1 \leq \tilde{v}(x) \leq W_2$, then $m(x, W_2) = W_2$ and $m(x, W_1) \in (W_1, \tilde{v}(x))$. In both cases, the second claim clearly holds. Now, consider the remaining case where $W_1 \leq W_2 < \tilde{v}(x)$. Since $d(x, m(x, W_1), W_1) \geq d(x, m(x, W_2), W_1)$ and $d(x, m(x, W_2), W_2) \geq d(x, m(x, W_1), W_2)$ we have:

$$\begin{aligned} 0 &\geq d(x, m(x, W_2), W_1) - d(x, m(x, W_1), W_1) + d(x, m(x, W_1), W_2) - d(x, m(x, W_2), W_2) \\ &= p(\theta(x, m(x, W_2))(W_2 - W_1) - p(\theta(x, m(x, W_1))(W_2 - W_1) \\ &= [p(\theta(x, m(x, W_2)) - p(\theta(x, m(x, W_1))](W_2 - W_1). \end{aligned}$$

Since $p(\theta(x, v))$ is decreasing in v (see Lemma W3), this also implies that $m(x, W_2) \geq m(x, W_1)$. If it holds with equality, i.e. if $m(x, W_2) = m(x, W_1)$, the second part of the lemma holds as well. If instead $m(x, W_2) > m(x, W_1)$, consider the arbitrary real number $\Delta \in \left(0, \frac{m(x,W_2)-m(x,W_1)}{2}\right)$ so that

$$d(x,m(x,W_1),W_1) \geq d(x,m(x,W_1)+\Delta,W_1)$$
$$p(\theta(x,m(x,W_1)))(m(x,W_1)-W_1) \geq p(\theta(x,m(x,W_1)+\Delta))(m(x,W_1)+\Delta-W_1)$$
$$[p(\theta(x,m(x,W_1)))-p(\theta(x,m(x,W_1)+\Delta))](m(x,W_1)-W_1) \geq p(\theta(x,m(x,W_1)+\Delta))\Delta$$
$$m(x,W_1)-W_1 \geq \frac{p(\theta(x,m(x,W_1)+\Delta))\Delta}{p(\theta(x,m(x,W_1)))-p(\theta(x,m(x,W_1)+\Delta))}.$$

Similarly, because $d(x,m(x,W_2),W_2) \geq d(x,m(x,W_2)-\Delta,W_2)$, it holds that

$$m(x,W_2)-W_2 \leq \frac{p(\theta(x,m(x,W_2)-\Delta))\Delta}{p(\theta(x,m(x,W_2)-\Delta))-p(\theta(x,m(x,W_2)))}.$$

Recall that the function $p(\theta(x,v))$ is decreasing and concave in v for all $v \leq \tilde{v}(x)$. Since $m(x,W_1)+\Delta \leq m(x,W_2)-\Delta$, then $p(\theta(x,m(x,W_1)+\Delta)) \geq p(\theta(x,m(x,W_2)-\Delta))$. Similarly, since $m(x,W_1) < m(x,W_2)$, $p(\theta(x,m(x,W_1)))-p(\theta(x,m(x,W_1)+\Delta)) \leq p(\theta(x,m(x,W_2)-\Delta))-p(\theta(x,m(x,W_2)))$. From these observations and the inequalities above, it follows that $m(x,W_2)-m(x,W_1) \leq W_2-W_1$. Hence, the lemma holds. □

Lemma W5 ($\hat{p}$ is decreasing and continuous in W)**.** *For all $x \in \mathbb{X}$ and all $W_1, W_2 \in \mathbb{V}$ with $W_1 \leq W_2$, the quitting probability $\hat{p}(x,W) \equiv p(\theta(x,m(x,W)))$ is such that*

$$-\overline{B}_p\left(W_2-W_1\right) \leq \hat{p}\left(x,W_2\right)-\hat{p}\left(x,W_1\right) \leq -\underline{B}_p\left(W_2-W_1\right) \qquad (1)$$

where $\overline{B}_p = -\frac{p'(0)\overline{B}_J}{q'(\overline{\theta})k} > 0$ and $\underline{B}_p = 0$.

Proof of Lemma W5. Let x be an arbitrary point in $\mathbb{X}$, and let W_1, W_2 be points in $\mathbb{V}$ with $W_1 \leq W_2$. Recall from Lemma W4 that $0 \leq m\left(x,W_2\right)-m\left(x,W_1\right) \leq W_2-W_1$. From Lemma W2, it follows that the difference $\theta(x,m\left(x,W_2\right))-\theta(x,m\left(x,W_1\right))$ is greater than $\left(W_2-W_1\right)\overline{B}_J/[q'(\overline{\theta})k]$ and smaller than 0. Finally, given concavity of p in θ, the difference $p(\theta(x,m(x,W_2)))-p(\theta(x,m(x,W_1)))$

is such that

$$\frac{p'(0)\,\overline{B}_J}{q'(\overline{\theta})k}\,(W_2 - W_1) \leq p(\theta(x, m(x, W_2))) - p(\theta(x, m(x, W_1))) \leq 0,$$

which gives the bounds on $\hat{p}(x, W)$. □

Lemma W6 (Differentiability of $\hat{p}(x, W)$ in W)**.** *For all $x \in \mathbb{X}$ and all $W \in \mathbb{V}$ with , the quitting probability $\hat{p}(x, W) = p(x, m(x, W))$ is differentiable a.e. in W.*

Proof of Lemma W6. The proof evolves in two steps. In the first step, it is shown that $p(\theta(x, v))$ is differentiable in v, then the second step turns to showing that $\hat{p}(x, W)$ is differentiable (almost everywhere) in W.

First, $p(\theta(x, v))$ is strictly concave, strictly decreasing and continuous in v and Rademacher's Theorem states that every concave function is differentiable almost everywhere. Hence, it needs to be shown that p is differentiable everywhere. Observe that $M(x, z_0, v)$ is concave in v and so it is differentiable almost everywhere. To show it is differentiable everywhere, assume that at a specific point $\tilde{v}$ the function M is not differentiable, so there exists a point of non-differentiability at $M(x, z_0, \tilde{v})$. We show that this cannot be the case. Let a different function $\tilde{M}(x, z_0, v)$ be defined as

$$\tilde{M}(x, z_0, v) = \sum_{x' \neq x'_i} \sum_{z' \neq z'_i} P(x'|x) P(z'|z_0) J(x', z', W^*_{x'z'}(x, z_0, \tilde{v})) + \\ P(x'_i|x) P(z'_i|z_0) J\left(x'_i, z'_i, v - \sum_{x' \neq x'_i} \sum_{z' \neq z'_i} P(x'|x) P(z'|z_0) W^*_{x'z'}(x, z_0, \tilde{v})\right),$$

where $W^*_{x'z'}(x, z_0, \tilde{v}) = \arg\max M(x, z_0, \tilde{v})$. $\tilde{M}$ is similar to M, specifically, feasibility is imposed in both functions and they are equal at $\tilde{v}$, $\tilde{M}(x, z_0, \tilde{v}) = M(x, z_0, \tilde{v})$. However, $\tilde{M}$ uses the optimal strategy from point $\tilde{v}$ at a point v, such that it is always weakly below M, $\tilde{M}(x, z_0, v) \leq M(x, z_0, v)$. The function

$\tilde{M}$ is also concave and continuously differentiable in v because v only appears in the last term and J is concave and differentiable in v. Hence, the Benveniste-Scheinkman Lemma allows for the conclusion that the function $M(x, z_0, v)$ is differentiable in v. Since the right hand side of the free entry condition $q(\theta(x,v)) = \dfrac{k}{M(x, z_0, v)}$ is differentiable in v (and q is differentiable in θ), so is $\theta(x, v)$. Finally, from the assumption that $p(\theta)$ is C_2 it must be that $p(\theta(x, v))$ is differentiable in v.

Second, differentiability of $p(\theta(x, m))$ in m carries over to differentiability of $\hat{p}(x, W) \equiv p(\theta(x, m(x, W)))$ in W. The function $m(x, W)$ is increasing and continuous in W, see Lemma W4. Lebesgue's Theorem for the differentiability of monotone functions states that a monotone function is differentiable almost everywhere. For the points of Lebesgue measure zero that are not differentiable $\tilde{W}$, use either the left or right differential of m at $\tilde{W}$ or the Gâteaux derivative, which exists everywhere due to Lipschitz continuity of m. To conclude, $\hat{p}(x, W)$ is differentiable (almost everywhere) because p is differentiable in m, which in turn is differentiable (almost everywhere) in W. □

W1.2 Existence of equilibrium

We show the existence of a recursive search equilibrium with firm-level shocks, worker shocks and effort on the job, closely following Menzio and Shi (2010) and Tsuyuhara (2016). The procedure aims at showing that the Bellman operator maps the set of firms' value functions, $\mathbb{J}$, into itself.

Lemma W7 (Continuity of θ in J)**.** *Consider two arbitrary functions $J_m, J_n \in \mathbb{J}$. Let $\theta_j(x, v)$ be the market tightness function implied by J_j for $j = m, n$. For any $\rho > 0$, if $\|J_m - J_n\| < \rho$ then $\|\theta_m - \theta_n\| < \varepsilon_\theta \rho$ where $\varepsilon_\theta \equiv \frac{-\overline{B}_J}{q'(\overline{\theta})k\underline{B}_J}$.*

Proof of Lemma W7. Let v be an arbitrary point in $\mathbb{V}$. From the boundedness property (J1) of the set $\mathbb{J}$, it follows that

$$J_n\left(v+\underline{B}_J^{-1}\rho\right)-J_n(v)\leq-\rho\;\Rightarrow\;J_n(v)-\rho\geq J_n\left(v+\underline{B}_J^{-1}\rho\right).$$

The same property of $\mathbb{J}$ is exploited to show that

$$J_n\left(v\right)-J_n\left(v-\underline{B}_J^{-1}\rho\right)\leq-\rho\;\Rightarrow\;J_n(v)+\rho\leq J_n\left(v-\underline{B}_J^{-1}\rho\right).$$

These observations and $||J_m-J_n||<\rho$ imply

$$J_m(v)<J_n(v)+\rho\leq J_n\left(v-\underline{B}_J^{-1}\rho\right)$$
$$J_m(v)>J_n(v)-\rho\geq J_n\left(v+\underline{B}_J^{-1}\rho\right).$$

The definition of market tightness and the first line lead to $\theta_m(x,v)\leq\theta_n\left(x,v-\underline{B}_J^{-1}\rho\right)$. Similarly, from the second line in the above result, it follows that $\theta_m(x,v)\geq\theta_n\left(x,v+\underline{B}_J^{-1}\rho\right)$. Hence,

$$\theta_m(x,v)-\theta_n(x,v)<\theta_n\left(x,v-\underline{B}_J^{-1}\rho\right)-\theta_n(x,v)\leq\varepsilon_\theta\rho$$
$$\theta_m(x,v)-\theta_n(x,v)>\theta_n\left(x,v+\underline{B}_J^{-1}\rho\right)-\theta_n(x,v)\geq-\varepsilon_\theta\rho$$

with $\varepsilon_\theta\equiv\frac{-\overline{B}_J}{q'(\overline{\theta})k\underline{B}_J}$. Thus, $|\theta_m(x,v)-\theta_n(x,v)|<\varepsilon_\theta\rho$. Since this result holds for all $(x,z,v)\in\mathbb{S}\times\mathbb{V}$, we conclude that $\|\theta_m-\theta_n\|<\varepsilon_\theta\rho$. □

Recall the return to search $D_m(x,W)\equiv\max_{v'\in\mathbb{V}}p\left(\theta_m(x,v')\right)(v'-W)$, which is maximized by the unique search policy function $m(x,W)$ with $m:\mathbb{X}\times\mathbb{V}\to\mathbb{V}$. In Appendix W1.1 it was shown that D is a decreasing function in W and that m is increasing in W. Further, we use $\hat{p}(x,W)\equiv p(\theta(x,m(x,W)))$ as the composite job finding function.

Lemma W8 (Continuity of D in J)**.** *Consider $J_m,J_n\in\mathbb{J}$. Let $D_j(x,W)$ be the worker value of searching implied by J_j for $j=m,n$. If $||J_m-J_n||<\rho$ then $\|D_m-D_n\|<\varepsilon_D\rho$ where $\varepsilon_D\equiv\varepsilon_\theta p'(0)(\overline{v}-\underline{v})$.*

Proof of Lemma W8. Let $\rho\in\mathbb{R}_{++}$ be an arbitrary number. Consider arbitrary functions $J_m, J_n \in \mathbb{J}$ such that $\|J_m - J_n\| < \rho$, and an arbitrary point $(x, W) \in \mathbb{X} \times \mathbb{V}$. Accordingly, we can construct the distance between $D_m(x, W)$ and $D_n(x, W)$ using Lemma W7.

$$\begin{aligned} |D_m(x,W) - D_n(x,W)| &\leq \max_{v' \in \mathbb{V}} | \left[p\left(\theta_m(x,v')\right) - p\left(\theta_n(x,v')\right) \right] (v' - W) | \\ &\leq \left\{ \max_{v' \in \mathbb{V}} |p\left(\theta_m(x,v')\right) - p\left(\theta_n(x,v')\right)| \right\} \left\{ \max_{v' \in \mathbb{V}} |v' - W| \right\} \\ &\leq \left\{ \max_{v' \in \mathbb{V}} \left| \int_{\theta_m(x,v')}^{\theta_n(x,v')} p'(t) dt \right| \right\} (\overline{v} - \underline{v}) < p'(0)\varepsilon_\theta(\overline{v} - \underline{v})\rho, \end{aligned}$$

where θ_j denotes the market tightness function computed with J_j. Since this holds for all $(x, W) \in \mathbb{X} \times \mathbb{V}$, we can conclude that $\|D_m - D_n\| < \varepsilon_D \rho$ with $\varepsilon_D = p'(0)\varepsilon_\theta(\overline{v} - \underline{v})$. □

Lemma W9 (Continuity of $\hat{p}$ in J)**.** *Consider $J_m, J_n \in \mathbb{J}$. Let $\hat{p}_j(x, W)$ be the composite transition function implied by J_j for $j = m, n$. If $||J_m - J_n|| < \rho$ then $||\hat{p}_m - \hat{p}_n|| < \varepsilon_p(\rho)$ where $\varepsilon_p(\rho) = \max\{2\overline{B}_p \rho^{1/2} + p'(0)\varepsilon_\theta \rho, 2\varepsilon_D \rho^{1/2}\}$ and $\overline{B}_p = -p'(0)\overline{B}_j/(kq'(\overline{\theta}))$. In addition $\varepsilon_p(\rho) \to 0$ as $\rho \to 0$.*

Proof of Lemma W9. Let $\rho \in \mathbb{R}_{++}$ be an arbitrary number. Consider arbitrary functions $J_m, J_n \in \mathbb{J}$ such that $\|J_m - J_n\| < \rho$, and an arbitrary point $(x, W) \in \mathbb{X} \times \mathbb{V}$. Without loss of generality, assume that $m_n(x, W) \leq m_m(x, W)$, where m_j is computed with θ_j and associated with J_j. In the proof consider three mutually exclusive cases and drop the (x, W) arguments from m_n and m_m for brevity.

Case 1: $p(\theta_n(x, m_n)) \leq p(\theta_m(x, m_m))$.

The distance between $p(\theta_m(x, m_m))$ and $p(\theta_n(x, m_n))$ is such that

$$(0 \ \leq) \ p(\theta_m(x, m_m)) - p(\theta_n(x, m_n)) \leq p(\theta_m(x, m_n)) - p(\theta_n(x, m_n)) < p'(0)\varepsilon_\theta \rho,$$

which exploits that $p(\theta_m(x, v))$ is decreasing in v, $m_m \geq m_n$ and the bounds characterized in Lemma W7.

Case 2: $p(\theta_n(x,m_n)) > p(\theta_m(x,m_m))$ and $m_m - 2\rho^{1/2} \leq m_n \leq m_m$.

Then:

$$\begin{aligned}(0 \quad <) \;\; & p(\theta_n(x,m_n)) - p(\theta_m(x,m_m)) \\ = \;\; & p(\theta_n(x,m_n)) - p(\theta_n(x,m_m)) + p(\theta_n(x,m_m)) - p(\theta_m(x,m_m)) \;<\; 2\overline{B}_p\rho^{1/2} + p'(0)\varepsilon_\theta\rho,\end{aligned}$$

which exploits Lemmas W7 and W5.

Case 3: $p\left(\theta_n\left(x,m_n\right)\right) > p\left(\theta_m\left(x,m_m\right)\right)$ and $m_n < m_m - 2\rho^{1/2} < m_m$.

First, note that $m_n \geq W$, as

$$m_n \begin{cases} \in (W, \tilde{v}_n) & \text{if } W < \tilde{v}_n \\ = W & \text{otherwise.} \end{cases}$$

As a result, we can write that $m_m > W + \rho^{1/2}$. Otherwise, if $m_m \leq W + \rho^{1/2}$, then $m_n < W - \rho^{1/2} < W$, which is a contradiction. Similarly, $m_m > W$ implies that $m_m < \tilde{v}_m$.

Note that $p(\theta_m(x,m_m))(m_m - W) \geq p(\theta_m(x,m_m - \rho^{1/2}))(m_m - \rho^{1/2} - W)$, because m_m is the optimal search decision when $J = J_m$. Therefore, we have

$$\begin{aligned}p(\theta_m(x,m_m))\rho^{1/2} &\geq \left[p(\theta_m(x,m_m - \rho^{1/2})) - p(\theta_m(x,m_m))\right](m_m - \rho^{1/2} - W) \\ &\geq \left[p(\theta_m(x,m_n)) - p(\theta_m(x,m_n + \rho^{1/2}))\right](m_m - \rho^{1/2} - W) \\ &\geq \left[p(\theta_m(x,m_n)) - p(\theta_m(x,m_n + \rho^{1/2}))\right](m_n + \rho^{1/2} - W).\end{aligned}$$

To obtain the second inequality we use the facts that $p(\theta_m(x,v))$ is concave in v for all $v \in [\underline{v}, \tilde{v}_m]$, that $m_n + \rho^{1/2} < m_m < \tilde{v}_m$ and that $m_m - \rho^{1/2} - W > 0$. To obtain the third inequality, consider that $m_n + \rho^{1/2} < m_m - \rho^{1/2}$ and that $p(\theta_m(x,m_n)) - p(\theta_m(x,m_n + \rho^{1/2})) > 0$.

Further, note that $p(\theta_n(x,m_n))(m_n - W)$ is greater than $p(\theta_n(x,m_n + \rho^{1/2}))(m_n +$

$\rho^{1/2} - W)$. Then:

$$p(\theta_n(x,m_n))\rho^{1/2} \leq \left[p(\theta_n(x,m_n)) - p(\theta_n(x,m_n+\rho^{1/2}))\right](m_n+\rho^{1/2}-W).$$

Subtracting this inequality from the previous result, dividing by $\rho^{1/2}$, and then applying Lemma W8 gives:

$$\begin{aligned}
0 &< p(\theta_n(x,m_n)) - p(\theta_m(x,m_m)) \\
&\leq \rho^{-1/2}\left[p(\theta_n(x,m_n)) - p(\theta_m(x,m_n)) + p(\theta_m(x,m_n+\rho^{1/2})) - p(\theta_n(x,m_n+\rho^{1/2})\right] \\
&\qquad \times (m_n+\rho^{1/2}-W)) \\
&< 2p'(0)\varepsilon_\theta\rho^{1/2}(\overline{v}-\underline{v}) = 2\varepsilon_D\rho^{1/2}.
\end{aligned}$$

Therefore, it can be established that the distance between $p(\theta_n(x,m_n))$ and $p(\theta_m(x,m_m))$ is such that

$$|p(\theta_n(x,m_n)) - p(\theta_m(x,m_m))| < \max\left\{2\overline{B}_p\rho^{1/2} + p'(0)\varepsilon_\theta\rho, 2\varepsilon_D\rho^{1/2}\right\} \equiv \varepsilon_p(\rho).$$

The $\rho^{1/2}$ term implies that $\lim_{\rho\to 0}\varepsilon_p(\rho) = 0$. Since this result holds for all $(x,W) \in \mathbb{X}\times\mathbb{V}$, we conclude that $\|\hat{p}_m - \hat{p}_n\| < \varepsilon_p(\rho)$. □

Lemma W10 (Continuity of U in J)**.** *Consider $J_m, J_n \in \mathbb{J}$. Let U_j be the worker unemployment value function implied by J_j for $j = m, n$. If $||J_m - J_n|| < \rho$ then $||U_m - U_n|| < \varepsilon_U\rho$, where $\varepsilon_U \equiv \beta\varepsilon_D/(1-\beta)$.*

Proof of Lemma W10. Let $\rho \in \mathbb{R}_{++}$ be an arbitrary number. Consider arbitrary functions $J_m, J_n \in \mathbb{J}$ such that $\|J_m - J_n\| < \rho$. For an arbitrary point $x \in \mathbb{X}$, the distance between $U_m(x)$ and $U_n(x)$ is

$$\begin{aligned}|U_m(x) - U_n(x)| =& |[b(x) + \beta \mathbb{E}_{x'} U_m(x') + D_m(x', U_m(x'))] - \\ & [b(x) + \beta \mathbb{E}_{x'} U_n(x') + D_n(x', U_n(x'))]| \\ \leq & \beta \mathbb{E}_{x'} \Big\{ |\, [U_m(x') + D_m(x', U_m(x'))] - [U_n(x') + \max D_m(x', U_n(x'))] \,| \\ & + |D_m(x', U_n(x')) - D_n(x', U_n(x'))| \Big\} \\ < & \beta \, \|U_m - U_n\| + \beta \varepsilon_D \rho\end{aligned}$$

To obtain the second inequality, we use that the distance between $U_m + D_m(U_m)$ and $U_n + D_m(U_n)$ is smaller than the distance between U_m and U_n. Since this result holds for all $x \in \mathbb{X}$,

$$\|U_m - U_n\| < \beta \, \|U_m - U_n\| + \beta \varepsilon_D \rho \;\Rightarrow\; \|U_m - U_n\| < \frac{\beta}{1-\beta} \varepsilon_D \rho,$$

which delivers the result. □

Lemma W11 (Bounding worker effort 1: Continuity of Ω in J)**.** *Consider $J_m, J_n \in \mathbb{J}$. Let $\Omega_j(x, W) = W + \kappa D_j(x, W) - \mathbb{E}_{x'}[U_j(x')|x]$ be the function implied by J_j for $j = m, n$. If $||J_m - J_n|| < \rho$ then $||\Omega_m - \Omega_n|| < \varepsilon_\Omega \rho$, where $\varepsilon_\Omega \equiv \kappa \varepsilon_D + \varepsilon_U$.*

Proof of Lemma W11.

$$\begin{aligned}|\Omega_m(x, W) - \Omega_n(x, W)| &= |\kappa \left(D_m(x, W) - D_n(x, W)\right) - \mathbb{E}_{x'} \left(U_m(x') - U_n(x')\right)| \\ &= |\kappa \left(D_m(x, W) - D_n(x, W)\right) + \mathbb{E}_{x'} \left(U_n(x') - U_m(x')\right)| \\ &< (\kappa \varepsilon_D + \varepsilon_U) \rho \equiv \varepsilon_\Omega \rho\end{aligned}$$

which delivers the result. □

Lemma W12 (Bounding worker effort 2: Continuity of e^* in J)**.** *Consider $J_m, J_n \in \mathbb{J}$. Let $e^*_j(x, W) = \Delta(\Omega_j(x, W))$ be the worker optimal effort function* *implied by J_j for $j = m, n$. If $||J_m - J_n|| < \rho$ then $||e^*_m - e^*_n|| < \varepsilon_e \rho$, where $\varepsilon_e = \overline{\Delta}' \varepsilon_\Omega$.*

Proof of Lemma W12. The optimization problem for the worker EQ-W leads to the first order condition for effort given by

$$-c'(e^*(x, W_i)) - \beta\delta'(e^*(x, W_i))\Omega(x, W_i) = 0.$$

Using the implicit function theorem, it follows that the derivative of e^*_i with respect to W_i is

$$\begin{aligned} e^{*\prime}(x, W_i) &= -\frac{-\beta\delta'(e^*_i)\Omega'(x, W_i)}{-c''(e^*_i) - \beta\delta''(e^*_i)\Omega(x, W_i)} \\ &= \underbrace{\frac{-\beta(\delta'(e^*_i))^2}{c''(e^*_i)\delta'(e^*_i) + c'(e^*_i)\delta''(e^*_i)}}_{\equiv \Delta'} \Omega'(x, W_i), \end{aligned}$$

where $\Delta(\Omega(x, W_i)))$ is the implicitly defined function for optimal effort. From the assumptions $\delta' \in [\underline{\delta}', 0), \delta(\cdot)'' \leq 0$ and $c' \in [0, \overline{c}']$ and the fact that $c(\cdot)$ is convex, the numerator is negative and bounded and the denominator is negative. Therefore, Δ' is positive and bounded by $\overline{\Delta}' \equiv |\sup \Delta'(\cdot)|$. Now, continuity of effort e^*_i in J can be established as follows:

$$\begin{aligned} |e^*_m(x, W_i) - e^*_n(x, W_i)| &= |\Delta\left(\Omega_m(x, W_i)\right) - \Delta\left(\Omega_n(x, W_i)\right)| \\ &\leq \overline{\Delta}' \left|\Omega_m(x, W_i) - \Omega_n(x, W_i)\right| \\ &< \overline{\Delta}' \varepsilon_\Omega \rho \equiv \varepsilon_e \rho. \end{aligned}$$

Since this holds for all $x \in \mathbb{X}$ it can be concluded that $||e^*_m - e^*_n|| < \varepsilon_e \rho$. □

Moving forward we define $\tilde{J}$, an update of the firm's value function J, as:

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

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

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

It can also be expressed as $\tilde{J}(x,z,V) = (TJ)(x,z,V)$ using the operator T. Next, let $F(\gamma,x,z,V)$ be the objective function of the reduced problem:

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

$$\text{s.t. } w_i = \begin{cases} u^{-1}\left(\dfrac{V - \pi_j [u(w_j) + \tilde{r}(x,W_j)]}{\pi_i} - \tilde{r}(x,W_i) \right) & \text{if } \pi_i \neq 0 \\ u^{-1}\left(V - \tilde{r}(x,W_i) \right) & \text{else,} \end{cases}$$

where $\gamma \in \Gamma$ denotes the tuple $(\{\pi_i, W_i, W_{ix'z'}\}_{i=1,2})$ and Γ is defined as the set of γ's such that $\pi_i \in [0,1]$, $\pi_1 + \pi_2 = 1, W_i \in \mathbb{V}, W_{ix'z'} : \mathbb{S} \to \mathbb{V}$, and $W_i = \mathbb{E}_{x'z'} \{W_{ix'z'} | x,z\}$. Finally, $\gamma^*(x,z,V)$ is the optimal solution such that

$$\tilde{J}(x,z,V) = \max_{\gamma \in \Gamma} F(\gamma(x,z,V),x,z,V) = F(\gamma^*,x,z,V).$$

Lemma W13 (Operator is self-mapping). *The Bellman operator is self-mapping, i.e. the image $\tilde{J}$ of $J \in \mathbb{J}$ also belongs to the set $\mathbb{J}$.*

Proof of Lemma W13. We need to show that the image through the Bellman operator satisfies the 4 properties of $\mathbb{J}$. Denote $F'(\gamma,x,z,V)$ as the derivative of $F(\gamma,x,z,V)$ with respect to V. It is straightforward to show that

$$F'(\gamma,x,z,V) = -\frac{1}{u'(w_i)} \in \left[-\frac{1}{\underline{u}'}, -\frac{1}{\overline{u}'} \right].$$

Condition 1: $\tilde{J}$ is bi-Lipschitz continuous in V.

Let (x, z) be an arbitrary point in $\mathbb{S}$ and let $V_1, V_2 \in \mathbb{V}$ be two arbitrary points with $V_1 \leq V_2$.

$$\begin{aligned} \left|\tilde{J}\left(x, z, V_2\right) - \tilde{J}\left(x, z, V_1\right)\right| &\leq \max_{\gamma\in\Gamma} \left|F\left(\gamma, x, z, V_2\right) - F\left(\gamma, x, z, V_1\right)\right| \\ &\leq \max_{\gamma\in\Gamma} \left|\int_{V_1}^{V_2} F'(\gamma, x, z, t)\, \mathrm{d}t\right| \\ &\leq \max_{\gamma\in\Gamma} \int_{V_1}^{V_2} \left|F'(\gamma, x, z, t)\right| \mathrm{d}t \;\leq\; \left|V_2 - V_1\right| / \underline{u}' \end{aligned}$$

The first inequality uses the fact that one could potentially find another γ that increases the distance. The expression implies that the function $\tilde{J}$ is Lipschitz continuous in V and differentiable almost everywhere. The function F is differentiable with respect to V. Therefore, at any point of differentiability, the derivative of $\tilde{J}$ with respect to V is equal to $F'(\gamma^*(x, z, V), x, z, V)$. From these properties of $\tilde{J}$, it follows that

$$\tilde{J}\left(x, z, V_2\right) - \tilde{J}\left(x, z, V_1\right) = \int_{V_1}^{V_2} F'\left(\gamma^*(x, z, t), x, z, t\right) \mathrm{d}t \in \left[-\frac{V_2 - V_1}{\underline{u}'}, -\frac{V_2 - V_1}{\overline{u}'}\right]$$

Therefore, $\tilde{J}$ is bi-Lipschitz continuous.

Condition 2: $\tilde{J}$ is bounded.

Let (x, z, V) be an arbitrary point in $\mathbb{S} \times \mathbb{V}$. The value $\tilde{J}(x, z, V)$ is such that

$$\tilde{J}(x, z, V) \leq f(\overline{x}, \overline{z}) - u^{-1}(\underline{v} + c(\underline{e}) - \beta\overline{v}) + \beta\overline{J} \leq \overline{J},$$

where we simply use the bounds on each of the terms. For the lower bound, let γ_0 denote the tuple $(\{\pi_{i,0}, W_{i,0}, W_{ix'z',0}\}_{i=1,2})$ such that $\pi_{1,0} = 0, \pi_{2,0} = 1, W_{i,0} = W_{ix'z',0} = \underline{v}$, and observe that

$$\tilde{J}(x, z, V) \geq F(\gamma_0, x, z, V) \geq f(\underline{x}, \underline{z}) - u^{-1}(\overline{v} + c(\overline{e}) - \beta\underline{v}) + \beta\underline{J} \geq \underline{J},$$

where the first inequality makes use of the fact that $\gamma_0 \in \Gamma$, and the second inequality makes use of the bounds on x, z, v, e and J.

Condition 3: $\tilde{J}$ is concave.

This is a direct implication of the presence of the lottery. Let V_1 and V_2 be two arbitrary values in $[\underline{v}, \overline{v}]$, and let $V_\alpha = \alpha V_1 + (1-\alpha)V_2$, where $\alpha \in (0,1)$. One can show that $J(V_\alpha) \geq \alpha J(V_1) + (1-\alpha) J(V_2)$.

Condition 4: $\tilde{J}$ is differentiable.

From above, $\tilde{J}(x, z, V)$ is decreasing in V because an increase in V tightens the promise-keeping constraint, concave with respect to V by construction because of the two-point lottery over promised expected values, continuous and differentiable almost everywhere. To show that $\tilde{J}$ is differentiable everywhere, we adapt the derivation steps presented in Koeppl (2006)[1] to the one-sided commitment model of this paper. Suppose for a fixed (x, z), there is a point $\tilde{V}$ where $\tilde{J}(x, z, \tilde{V})$ is not differentiable and call $(\tilde{\pi}_i, \tilde{w}, \tilde{W}_i, \tilde{W}_{ix'z'})$ the firm's optimal action at that point. This action is by definition feasible and delivers $\tilde{V}$ to the worker. Next, consider a strategy that delivers any V around $\tilde{V}$ by changing the wage to $w^*(V) \equiv u^{-1}(V - \tilde{V} + u(\tilde{w}))$ while the remaining actions $(\tilde{\pi}_i, \tilde{W}_i, \tilde{W}_{ix'z'})$ stay the same. We define the function $\hat{J}(x, z, V)$ as the value that uses strategy $(\tilde{\pi}_i, w^*(V), \tilde{W}_{ix'z'}, \tilde{W}_i)$, which is also feasible by construction. Then, by definition of $\tilde{J}$ it must be that $\hat{J}(x, z, V) \leq \tilde{J}(x, z, V)$ and $\hat{J}(x, z, \tilde{V}) = \tilde{J}(x, z, \tilde{V})$.

Next, since $u(\cdot)$ is concave, increasing and twice differentiable, $-u^{-1}(\cdot)$ is also concave and twice differentiable. Moreover, V enters $\hat{J}(x, z, V)$ only through $-w^*(V)$ and so $\hat{J}(x, z, V)$ inherits concavity and differentiability from the utility function at any point V, including $\tilde{V}$. Finally, since $\hat{J}$ is a function that is concave, continuously differentiable, lower than $\tilde{J}$ and equal to $\tilde{J}$ at $\tilde{V}$ we can

[1] Koeppl (2006) shows that with two-sided limited commitment it is sufficient to have one state realization where neither participation constraint binds to achieve differentiability of the Pareto frontier.

apply Lemma 1 from Benveniste and Scheinkman (1979), which reveals that $\tilde{J}(x,z,V)$ is differentiable at $\tilde{V}$. Consequently, $\tilde{J}$ is differentiable everywhere.

□

Lemma W14 (Continuity of the operator)**.** *Consider $J_m, J_n \in \mathbb{J}$. Let $\tilde{J}_j(x,z,V)$ be the firm's value mapping implied by J_j for $j = m,n$. If $||J_m - J_n|| < \rho$, then $||\tilde{J}_m - \tilde{J}_n|| < \varepsilon_T(\rho)$.*

Proof of Lemma W14. Let $F_j(\gamma_j, x, z, V)$ be the objective function of the firm's optimal contracting problem implied by J_j. Consider $J_m, J_n \in \mathbb{J}$ such that $||J_m - J_n|| < \rho$. Take $V \in \mathbb{V}$ such that $\tilde{J}_m(x,z,V) - \tilde{J}_n(x,z,V) > 0$. Let $\gamma_j^*(x,z,V)$ be the maximizer of $F_j(\gamma_j,x,z,V)$ and $w_j(\gamma)$ be the wage function given by J_j. Then, dropping the arguments of γ_j^* for brevity:

$$
\begin{aligned}
0 &\leq |\tilde{J}_m(x,z,V) - \tilde{J}_n(x,z,V)| \\
&= |F_m(\gamma_m^*, x, z, V) - F_n(\gamma_n^*, x, z, V)| \\
&\leq |F_m(\gamma_m^*, x, z, V) - F_n(\gamma_m^*, x, z, V)| \\
&\leq | - w_m(\gamma_m^*) + \sum_{i=1,2} \pi_{i,m}\{f(x,z) + \beta\tilde{p}_m(x, W_{i,m}))\mathbb{E}_{x'z'}[J_m(x',z',W_{ix'z',m})|x,z]\} \\
&\quad + w_n(\gamma_m^*) - \sum_{i=1,2} \pi_{i,m}\{f(x,z) + \beta\tilde{p}_n(x, W_{i,m}))\mathbb{E}_{x'z'}[J_n(x',z',W_{ix'z',m})|x,z]\}| \\
&\leq |w_m(\gamma_m^*) - w_n(\gamma_m^*)| \\
&\quad + \sum_{i=1,2} \pi_{i,m}|\{f(x,z) + \beta\tilde{p}_m(x, W_{i,m}))\mathbb{E}_{x'z'}[J_m(x',z',W_{ix'z',m})|x,z]\} \\
&\qquad\qquad - \{f(x,z) + \beta\tilde{p}_n(x, W_{i,m}))\mathbb{E}_{x'z'}[J_n(x',z',W_{ix'z',m})|x,z]\}|.
\end{aligned}
$$

The objective is to estimate a bound for $|\tilde{J}_m(x,z,V) - \tilde{J}_n(x,z,V)|$ by looking at each part of the last expression separately as follows.

(1) Consider $|w_m(\gamma_m^*) - w_n(\gamma_m^*)|$ first. Since utility u is a strictly concave function, for any w_1 and w_2, $|w_1 - w_2|u' < |u(w_1) - u(w_2)|$ where u' is the smaller

of $u'(w_1)$ and $u'(w_2)$. By definition,

$$\begin{aligned}u(w_m(\gamma_m^*)) =& V - \sum_{i=1,2} \pi_i \tilde{r}(x, W_{i,m}) \\ =& V - \beta \mathbb{E}_{x'}[U_m(x')|x] \\ &- \sum_{i=1,2} \pi_{i,m}[-c(e_m(x, W_{i,m})) + \beta(1 - \delta(e_m(x, W_{i,m})))\Omega_m(x, W_{i,m})]\end{aligned}$$

and

$$\begin{aligned}u(w_n(\gamma_m^*)) =& V - \beta \mathbb{E}_{x'}[U_n(x')|x] \\ &- \sum_{i=1,2} \pi_{i,m}[-c(e_n(x, W_{i,m})) + \beta(1 - \delta(e_n(x, W_{i,m})))\Omega_n(x, W_{i,m})].\end{aligned}$$

Therefore, we can express the distance as

$$\begin{aligned}&|u(w_m(\gamma_m^*)) - u(w_n(\gamma_m^*))| \\ \leq& \beta|U_m - U_n| + \sum_{i=1,2} \pi_{i,m}[|c(e_m(x, W_{i,m})) - c(e_m(x, W_{i,m}))| \\ &+ \beta|(1 - \delta(e_m(x, W_{i,m})))\Omega_m(x, W_{i,m}) - (1 - \delta(e_n(x, W_{i,m})))\Omega_n(x, W_{i,m})|].\end{aligned}$$

$|U_m - U_n|$ is bounded by ε_U. The last term is also bounded due to:

$$\begin{aligned}&|(1 - \delta(e_m(x, W_{i,m})))\Omega_m(x, W_{i,m}) - (1 - \delta(e_n(x, W_{i,m})))\Omega_n(x, W_{i,m})| \\ \leq& |(1 - \delta(e_m(x, W_{i,m})))\Omega_m(x, W_{i,m}) - (1 - \delta(e_n(x, W_{i,m})))\Omega_m(x, W_{i,m})| \\ &+ |(1 - \delta(e_n(x, W_{i,m})))\Omega_m(x, W_{i,m}) - (1 - \delta(e_n(x, W_{i,m})))\Omega_n(x, W_{i,m})| \\ \leq& |(1 - \delta(e_m(x, W_{i,m}))) - (1 - \delta(e_n(x, W_{i,m})))|\overline{v} \\ &+ (1 - \delta(e_n(x, W_{i,m})))|\Omega_m(x, W_{i,m}) - \Omega_n(x, W_{i,m})| \\ \leq& -\underline{\delta}'|e_m(x, W_{i,m}) - e_n(x, W_{i,m})|\overline{v} + |\Omega_m(x, W_{i,m}) - \Omega_n(x, W_{i,m})| \\ \leq& (-\underline{\delta}'\varepsilon_e\overline{v} + \varepsilon_\Omega)\rho,\end{aligned}$$

using the fact that $\Omega(\cdot)$ cannot exceed $\overline{v}$. Collecting bounds yields:

$$|u(w_m(\gamma_m^*)) - u(w_n(\gamma_m^*))| \leq (\beta\varepsilon_U + \overline{c}'\varepsilon_e + \beta(-\underline{\delta}'\varepsilon_e\overline{v} + \varepsilon_\Omega))\rho.$$

So, from the property of concave functions, the first term is bounded by:

$$|w_m(\gamma_m^*) - w_n(\gamma_m^*)| \leq u'^{-1} \cdot (\beta\varepsilon_U + \overline{c}'\varepsilon_e + \beta(-\underline{\delta}'\varepsilon_e\overline{v} + \varepsilon_\Omega))\rho.$$

(2) Next, consider the following term:

$$\sum_{i=1,2} \pi_{i,m} |\{f(x,z) + \beta\tilde{p}_m(x, W_{i,m}))\mathbb{E}_{x'z'}[J_m(x', z', W_{ix'z',m})|x,z]\}$$
$$- \{f(x,z) + \beta\tilde{p}_n(x, W_{i,m}))\mathbb{E}_{x'z'}[J_n(x', z', W_{ix'z',m})|x,z]\}|$$

This expression can be divided into two sub-components stemming from substituting in $\tilde{p}$. Similarly to above, the bound for each sub-component can be found as follows. The first subcomponent can be bounded directly:

$$\begin{aligned}
&|(1-\delta(e_m(x, W_{i,m})))J_m(W_{i,m}) - (1-\delta(e_n(x, W_{i,m})))J_n(W_{i,m})| \\
&\leq |(1-\delta(e_m(x, W_{i,m})))J_m(W_{i,m}) - (1-\delta(e_n(x, W_{i,m})))J_m(W_{i,m})| \\
&\quad + |(1-\delta(e_n(x, W_{i,m})))J_m(W_{i,m}) - (1-\delta(e_n(x, W_{i,m})))J_n(W_{i,m})| \\
&= |(1-\delta(e_m(x, W_{i,m}))) - (1-\delta(e_n(x, W_{i,m})))|J_m(W_{i,m}) \\
&\quad + (1-\delta(e_n(x, W_{i,m})))|J_m(W_{i,m}) - J_n(W_{i,m})| \\
&\leq -\underline{\delta}'|e_m(x, W_{i,m}) - e_n(x, W_{i,m})|\overline{J} + |J_m(W_{i,m}) - J_n(W_{i,m})| \\
&\leq (-\underline{\delta}'\varepsilon_e\overline{J} + 1)\rho.
\end{aligned}$$

Then note that:

$$\begin{aligned}
&|\hat{p}_m(x, W_{i,m})J_m(W_{i,m}) - \hat{p}_n(x, W_{i,m})J_n(W_{i,m})| \\
&\leq |\hat{p}_m(x, W_{i,m})J_m(W_{i,m}) - \hat{p}_n(x, W_{i,m})J_m(W_{i,m})| \\
&\quad + |\hat{p}_n(x, W_{i,m})J_m(W_{i,m}) - \hat{p}_n(x, W_{i,m})J_n(W_{i,m})| \\
&= |\hat{p}_m(x, W_{i,m}) - \hat{p}_n(x, W_{i,m})|J_m(W_{i,m}) + \hat{p}_n(x, W_{i,m})|J_m(W_{i,m}) - J_n(W_{i,m})| \\
&\leq \varepsilon_p(\rho)\bar{J} + \rho,
\end{aligned}$$

which is used to find the bounds of the second sub-component:

$$
\begin{aligned}
&|(1-\delta(e_m(x,W_{i,m})))\hat{p}_m(x,W_{i,m})J_m(W_{i,m})-(1-\delta(e_n(x,W_{i,m})))\hat{p}_n(x,W_{i,m})J_n(W_{i,m})| \\
\leq\ & |(1-\delta(e_m(x,W_{i,m})))\hat{p}_m(x,W_{i,m})J_m(W_{i,m})-(1-\delta(e_n(x,W_{i,m})))\hat{p}_m(x,W_{i,m})J_m(W_{i,m})| \\
&+|(1-\delta(e_n(x,W_{i,m})))\hat{p}_m(x,W_{i,m})J_m(W_{i,m})-(1-\delta(e_n(x,W_{i,m})))\hat{p}_n(x,W_{i,m})J_n(W_{i,m})| \\
=\ & |(1-\delta(e_m(x,W_{i,m})))-(1-\delta(e_n(x,W_{i,m})))|\hat{p}_m(x,W_{i,m})J_m(W_{i,m}) \\
&+(1-\delta(e_n(x,W_{i,m})))|\hat{p}_m(x,W_{i,m})J_m(W_{i,m})-\hat{p}_n(x,W_{i,m})J_n(W_{i,m})| \\
\leq\ & (-\underline{\delta}'\varepsilon_e\overline{J}+1)\rho+\varepsilon_p(\rho)\overline{J}.
\end{aligned}
$$

Collecting the inequalities from (1) and (2), the overall bound is given by:

$$
\begin{aligned}
&|\tilde{J}_m(x,z,V)-\tilde{J}_n(x,z,V)| \\
\leq\ & u'^{-1}\cdot(\beta\varepsilon_U+\overline{c}'\varepsilon_e+\beta(\varepsilon_\Omega-\underline{\delta}'\varepsilon_e\overline{v}))\rho \\
&+\beta(1+\kappa)(1-\underline{\delta}'\varepsilon_e\overline{J})\rho+\beta\kappa\varepsilon_p(p)\bar{J} \\
\equiv\ & \varepsilon_T(\rho).
\end{aligned}
$$

Hence, the operator T is continuous. □

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.

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. □

W2 Identification web appendix

In this supplementary appendix, we show how properties of the theoretical model map into conditional independence restrictions that can be used to develop a non-parametric identification argument. In a nutshell, there are four important features that can be used. First, coworker trajectories are independent of each other conditional on the firm shock. Second, the way that workers' lifetime utility and the productivity processes evolve together form a Markov-switching model as described in Hu and Shum (2012). Third, in the absence of flat regions in the Pareto frontier, the value of the worker maps into the wage one-for-one. Finally, monotonicity of the target wage allows labeling the unobserved states of firm productivity.

Our strategy consists of the following steps. To start, we describe the model's data counterpart. Next, we show how the restrictions of the model help identify the law of motion of the wage as well as the laws of motion of firm and worker productivities. The conditional choice probabilities together with the Bellman equation allow us to then recover the structural parameters of the model. This procedure has the flavor of the two-step approach of Hotz and Miller (1993), recovering the conditional choice probabilities before finding the structural parameters. All proofs are deferred to the last subsection.

W2.1 Data

Consider a worker i observed over T periods. Call $X_{it} \in \{1, ..., n_x\}$ his unobservable ability and $Z_{it} \in \{1, ..., n_z\}$ his firm level match quality if employed, with Z_{it}=0 if not employed. Y_{it} denotes the wage and is set to Y_{it}=0 for an unemployed worker. We call M_{it} the mobility realization between $t-1$ and t, where M_{it}=0 if the worker stays in the same firm, M_{it}=1 if the worker moves to a new firm, M_{it}=2 for transitions into unemployment, M_{it}=3 for transitions out of unemployment, and finally, M_{it}=4 if an unemployed worker remains unemployed. Note that the timing implies that a separation in the current period is reflected in M_{t+1}, not M_t, which is natural given the timing in the model where the wage is collected before separation.

We supplement data on individual i with information about K coworkers who joined the firm at the same time as worker i (potentially multiple periods in the past) and index them by $k(i,t)$. Their wages are denoted Y^c_{ikt}, where Y^c_{ikt}=0 if the coworker became unemployed. For an unemployed worker i we consider the coworkers of the last employer. Our data is then formed from a random sample of sequences of the form $\{Y_{it}, M_{it}, Y^c_{i1t}, ..., Y^c_{iKt}\}_{i=1,...,N,t=1,...,T}$.

W2.2 Identifying the choice probabilities

The first goal is to show that the structure of the model can be used to identify $Pr[Y_{it+1}, M_{it+1}|Y_{it}, \tilde{X}_{it}, Z_{it}]$ as well as $Pr[Z_{it+1}|Z_{it}]$ and $Pr[\tilde{X}_{it+1}|\tilde{X}_{it}]$ from the observed joint density of $\{Y_{it}, M_{it}, Y^c_{i1t}, ..., Y^c_{iKt}\}_{i=1,...,N,t=1,...,T}$. To improve readability, we denote $S_{it} \equiv (Y_{it}, M_{it})$ and think of Y_{it} as a discrete outcome.[2]

Let K=2 and T=4, which is sufficient for identification. Consider individuals who joined an employer in period 1 and stay there for at least two periods, i.e. condition on the mobility set $\bar{M}_i \equiv \mathbf{1}\{M_{i1} \in \{1, 3\}\} \cdot \mathbf{1}\{M_{i2}=0\}$. Furthermore, we introduce $H_{i2} \equiv (Z_{i1}, Z_{i2})$ and $H_{i3} \equiv (Z_{i1}, Z_{i2}, Z_{i3})$ as the sequence of realized Z_{it} in the firm that worker i and all coworkers joined at $t = 1$. $\tilde{H}_{it}$ denotes the same sequences up to a permutation.

In Lemma W15 we recover individual-specific wage and mobility distributions jointly with the sequence of firm shocks captured by $\tilde{H}_{i3}$, $Pr[S_{i1}, S_{i2}, S_{i3}, S_{i4}, \tilde{H}_{i3}|\bar{M}_i{=}1]$. The proof relies on the property of the model that conditional on the sequence of shocks $\tilde{H}_{i3}$, the realizations of wages of all coworkers are independent of each other because all common shocks must be firm shocks.[3] This conditional independence structure allows us to apply the result for discrete mixtures in Hall and Zhou (2003).

Lemma W16 uses the Markovian property of the contract to recover $Pr[S_{i3}|S_{i2}, \tilde{X}_{i2}, \tilde{H}_{i3}, \bar{M}_i{=}1]$ from $Pr[S_{i1}, S_{i2}, S_{i3}, S_{i4}, \tilde{H}_{i3}|\bar{M}_i{=}1]$ for a permutation $\tilde{X}_{i2}$ of X_{i2}, which is (s_3, s_2, h_3)-specific. The proof closely follows Hu and Shum (2012) on the identification of a Markov-switching model. Since the productivity process is independent of the wage process and match quality realization, the condition of "limited feedback" required in the original paper is satisfied. Additionally, we adopt a non-primitive rank condition on the law of motion of wages.

Lemma W17 and Lemma W18 provide rank conditions sufficient to label X_{i2} across values of (s_3, s_2, h_3). These conditions require sufficient variation in S_{i4}

[2]Using continuous outcomes requires changing the rank condition into a linear independence requirement of the marginal distributions, see Allman, Matias, and Rhodes (2009) Theorem 8.

[3]Here we can use the Z_{it} sequence directly, rather than the ν_t sequence, since different coworkers started in the same period, and hence share the exact same Z_{it} history.

and S_{i1} across values of X_{i2}. Once X_{i2} is consistently labeled, monotonicity of the target wage $w^*(x,z)$ in z can be used to label and order the values of Z_{i2} in each $\tilde{H}_{i3}$ history, see Lemma W19. In addition, under the assumption of diagonal dominance of the transition matrix, we recover $Pr[Z_{i3}|Z_{i2}]$. Lemma W20 uses the identified $Pr[S_{i3}|S_{i2},\tilde{X}_{i2},\tilde{H}_{i3},\bar{M}_i{=}1]$ to recover $Pr[\tilde{X}_{i3}|\tilde{X}_{i2}]$. At this point, we know $Pr[Y_{i3},M_{i3}|Y_{i2},\tilde{X}_{i2},Z_{i2}]$, $Pr[Z_{i3}|Z_{i2}]$ and $Pr[\tilde{X}_{i3}|\tilde{X}_{i2}]$ for a common permutation of X_{i2} and $Y_{i2}{>}0$. Since the model is stationary, this delivers $Pr[Y_{i,t+1},M_{i,t+1}|Y_{it},\tilde{X}_{it},Z_{it}]$ for $Y_{it}{>}0$ and the laws of motion $Pr[Z_{i,t+1}|Z_{it}]$ and $Pr[\tilde{X}_{i,t+1}|\tilde{X}_{it}]$ in Lemma W21.

W2.3 Identifying the model parameters

After identifying the transition probabilities in the previous section, we are interested in recovering the structural parameters of the model, in particular $f(x,z)$. Lemma W22 shows that the present value of the worker $V(x,z,w)$ at each state (x,z,w) is uniquely defined from the transition probabilities. One complication of reconstructing this present value is to express the continuation value at job losses because we don't want to assume that the flow value of unemployment $b(x)$ is observed. To overcome this, we use the fact that workers who are indifferent between working and not working exert zero effort, and so their probability of quitting approaches one. Conditioning on $\delta^* \simeq 1$, a worker's continuation value at the job is thus identical to the value of being unemployed. Another difficulty is to reconstruct the value v_1^* that the worker gets after a J2J transition. This however can be addressed by using the present value conditional on moving.

Recovering the production function $f(x,z)$ is achieved in Lemma W23 based on the property of the optimal contract that $J'(x,z,V) = \frac{1}{u'(w)}$. Using $V(x,z,w)$ from Lemma W22, we can integrate the first order condition to get $J(x,z,V)$ up to a (x,z)-specific constant. This intercept is pinned down by the residual claimant wage $w^*(x,z)$, for which the expected profit of the firm equals zero.

One could ask if additional information would be able to discipline the two functions $u(\cdot)$ and $c(\cdot)$. We show that even in the case where $c(\cdot)$ is not known, $V(x,z,w)$ can take the form of a Volterra integral equation of the second kind with a unique solution under very mild conditions. As for the utility function $u(\cdot)$, we note that an overall measure of passthrough from productivity to earnings or an analysis similar to Guiso, Pistaferri, and Schivardi (2005) could help measure the amount of risk aversion. We leave this for future research.

W2.4 Proofs

Lemma W15 (Firm shock history h)**.** *$Pr[S_{i1},S_{i2},S_{i3},S_{i4},\tilde{H}_{i3}|\bar{M}_i{=}1]$ is identified from the joint probability $Pr[S_{i1},S_{i2},S_{i3},S_{i4},Y^c_{i11},...,Y^c_{i24}\,|\,\bar{M}_i{=}1]$, where $\tilde{H}_{i3}=\sigma(H_{i3})$ for some permutation σ, under the assumptions of the structural model and the following conditions:*

i) $Pr[S_{i1},S_{i2},S_{i3},S_{i4}|H_{i3},\bar{M}_i{=}1]$ and $Pr[Y^c_{i,1,1},Y^c_{i,1,2},Y^c_{i,1,3},Y^c_{i,1,4}|H_{i3},\bar{M}_i{=}1]$ have rank $n_h{=}n_z^3$.

ii) There exists (y_1,y_2,y_3,y_4) and (y'_1,y'_2,y'_3,y'_4) such that for all values h_3 of H_{i3} the following quantities are different:

$$\frac{Pr[H_{i3}{=}h_3,Y^c_{i21}{=}y_1,Y^c_{i22}{=}y_2,Y^c_{i23}{=}y_3,Y^c_{i24}{=}y_4|\bar{M}_i{=}1]}{Pr[H_{i3}{=}h_3,Y^c_{i21}{=}y'_1,Y^c_{i22}{=}y'_2,Y^c_{i23}{=}y'_3,Y^c_{i24}{=}y'_4|\bar{M}_i{=}1]}.$$

Proof. We apply the identification result of mixtures, which depends on conditional independence. In the model, the wage path of a given worker is a function of the worker's own shock sequence, but given the firm shock history, individual-specific shocks are independent across coworkers. Hence, conditional independence holds as long as we go far enough back to condition on the full firm shock history shared between coworkers. For this reason we look at workers who enter in period 1 and write:

$$
\begin{aligned}
Pr[S_{i1}, ..., S_{i4}, Y^c_{i11}, ..., Y^c_{i24} | \bar{M}_i{=}1] \\
&= \sum_{H_{i3}} Pr[H_{i3} | \bar{M}_i{=}1] \cdot Pr[S_{i1}, ..., S_{i4}, Y^c_{i11}, ..., Y^c_{i24} | H_{i3}, \bar{M}_i{=}1] \\
&= \sum_{H_{i3}} Pr[H_{i3} | \bar{M}_i{=}1] \cdot Pr[S_{i1}, ..., S_{i4} | H_{i3}, \bar{M}_i{=}1] \\
&\qquad \times \Big(\prod_k Pr[Y^c_{ik1}, ..., Y^c_{ik4} | H_{i3}, \bar{M}_i{=}1] \Big).
\end{aligned}
$$

The objects of interest are $Pr[H_{i3}|\bar{M}_i{=}1]$ and $Pr[S_{i1}, ..., S_{i4}|H_{i3}, \bar{M}_i{=}1]$. With only two coworker observations we receive three independent measures of the income sequence conditional on the sequence $H_{i3} = (Z_{i1}, Z_{i2}, Z_{i3})$.

For convenience we write $\boldsymbol{y} = (y_1, y_2, y_3, y_4)$ and $\boldsymbol{s} = (s_1, s_2, s_3, s_4)$ with respective supports of size $n_{\boldsymbol{y}}$ and $n_{\boldsymbol{s}}$ and construct a matrix $B(\boldsymbol{y})$, defined for a fixed value $\boldsymbol{y}$, with the following elements:

$$
\Big[B(\boldsymbol{y})\Big]_{pq} = Pr[(S_{i1}, ..., S_{i4}) = \boldsymbol{s}_p, (Y^c_{i11}, ..., Y^c_{i14}) = \boldsymbol{y}_q, (Y^c_{i21}, ..., Y^c_{i24}) = \boldsymbol{y} | \bar{M}_i{=}1].
$$

We further define the following matrices of interest:

$$
\begin{aligned}
\Big[L_{\boldsymbol{S}|H_3}\Big]_{pq} &= Pr[(S_{i1}, ..., S_{i4}) = \boldsymbol{s}_p | H_{i3} = h_q, \bar{M}_i{=}1] \\
\Big[L_{\mathbf{Y}^c_1|H_3}\Big]_{pq} &= Pr[(Y^c_{i11}, ..., Y^c_{i14}) = \boldsymbol{y}_p | H_{i3} = h_q, \bar{M}_i{=}1] \\
\Big[D_{\mathbf{Y}^c_2, H_3}(\boldsymbol{y})\Big]_{pq} &= \mathbf{1}\{p{=}q\} \cdot Pr[(Y^c_{i21}, ..., Y^c_{i24}) = \boldsymbol{y}, H_{i3} = h_q | \bar{M}_i{=}1].
\end{aligned}
$$

Note that conditional mean independence gives:

$$
B(\boldsymbol{y}) = L_{\boldsymbol{S}|H_3} D_{\mathbf{Y}^c_2, H_3}(\boldsymbol{y}) L'_{\mathbf{Y}^c_1|H_3}.
$$

We then compute a singular value decomposition $B(\boldsymbol{y}') = USV'$ where S is a diagonal matrix with non-negative values of size $n_h \times n_h$, U and V are of size $n_{\boldsymbol{s}} \times n_h$ and $n_{\boldsymbol{y}} \times n_h$. In addition $U'U$ and $V'V$ are the identity matrix of size n_h. This gives us that $U'B(\boldsymbol{y}')V$, $U'L_{\boldsymbol{S}|H_3}$, and $L'_{\mathbf{Y}^c_1|H_3}V$ are full rank. We construct:

$$U'B(\boldsymbol{y})V\Big(U'B(\boldsymbol{y}')V\Big)^{-1} = U'L_{\boldsymbol{S}|H_3}D_{\mathbf{Y}_2^c,H_3}(\boldsymbol{y})L'_{\mathbf{Y}_1^c|H_3}V \times \Big(U'L_{\boldsymbol{S}|H_3}D_{\mathbf{Y}_2^c,H_3}(\boldsymbol{y}')L'_{\mathbf{Y}_1^c|H_3}V\Big)^{-1}$$
$$= U'L_{\boldsymbol{S}|H_3}D_{\mathbf{Y}_2^c,H_3}(\boldsymbol{y})D_{\mathbf{Y}_2^c,H_3}(\boldsymbol{y}')^{-1}\Big(U'L_{\boldsymbol{S}|H_3}\Big)^{-1}$$

So, the eigenvalue decomposition of $U'B(\boldsymbol{y})V\Big(U'B(\boldsymbol{y}')V\Big)^{-1}$ delivers $U'L_{\boldsymbol{S}|H_3}$ as the eigenvectors. Since U is known from the SVD decomposition and condition ii) guarantees that the eigenvalues are different, we find a unique $L_{\boldsymbol{S}|H_3}$ up to a normalization and a permutation. The notation $\tilde{H}_{i3}$ captures the permutation. The normalization is pinned down by the fact that $L_{\boldsymbol{S}|H_3}$ is a density and hence needs to sum to one. This gives us $Pr[S_{i1}, ..., S_{i4}|\tilde{H}_{i3}, \bar{M}_i{=}1]$.

A similar eigenvalue decomposition for $\Big(U'B(\boldsymbol{y}')V\Big)^{-1}U'B(\boldsymbol{y})V$ yields $L_{\mathbf{Y}_1^c|H_3}$ and consequently $D_{\mathbf{Y}_2^c,H_3}(\boldsymbol{y})$. From there we compute:

$$Pr[\tilde{H}_{i3} = h_q|\bar{M}_i{=}1] = \frac{Pr[(Y_{i21}^c, ..., Y_{i24}^c)' = \boldsymbol{y}, \tilde{H}_{i3} = h_q, \bar{M}_i{=}1]}{Pr[(Y_{i21}^c, ..., Y_{i24}^c)' = \boldsymbol{y}|\tilde{H}_{i3} = h_q, \bar{M}_i{=}1]}$$
$$= \frac{Pr[(Y_{i21}^c, ..., Y_{i24}^c)' = \boldsymbol{y}, \tilde{H}_{i3} = h_q, \bar{M}_i{=}1]}{Pr[(Y_{i11}^c, ..., Y_{i14}^c)' = \boldsymbol{y}|\tilde{H}_{i3} = h_q, \bar{M}_i{=}1]},$$

where the second equality uses the fact that coworkers are interchangeable. □

Lemma W16 (Law of motion of s)**.** *Under the assumptions of the structural model and in the absence of flat regions in the Pareto frontier,* $Pr[S_{i3}{=}s_3|S_{i2}{=}s_2, \tilde{X}_{i2}{=}x, \tilde{H}_{i3}{=}h_3, \bar{M}_i{=}1]$, $Pr[S_{i4}{=}s_4|S_{i3}{=}s_3, \tilde{X}_{i2}{=}x, \tilde{H}_{i3}{=}h_3, \bar{M}_i{=}1]$ *and* $Pr[\tilde{X}_{i2}|S_{i2}, S_{i1}, \tilde{H}_{i3}, \bar{M}_i{=}1]$ *are identified for each* (s_3, s_2, h_3) *with* $s_2{=}(y_2, 0)$ *and all values* (s_4, x)*, where* $\tilde{X}_{i2}{=}\sigma_{s_3s_2h_3}(X_{i2})$ *for an unknown permutation* $\sigma_{s_3s_2h_3}$*, if:*

i) The matrix $A(s_2, s_3, h_3)$ *has rank* n_x*, where each element is defined as* $a_{pq} = Pr[S_{i1}{=}s_p, S_{i2}{=}s_2, S_{i3}{=}s_3, S_{i4}{=}s_q|\tilde{H}_{i3}{=}h_3]$.

ii) There exists (s_2', s_3') *such that for all* x_2 *and* $x_2'{\neq}x_2$ *we have*

$\lambda_{s_2,s_2',s_3,s_3'}(x_2) \neq \lambda_{s_2,s_2',s_3,s_3'}(x_2')$, *where* $\lambda_{s_2,s_2',s_3,s_3'}(x_2)$ *is defined as:*

$$\lambda_{s_2,s_2',s_3,s_3'}(x_2) = \frac{Pr[S_{i3}{=}s_3|S_{i2}{=}s_2, X_{i2}{=}x_2, \tilde{H}_{i3}{=}h_3, \bar{M}_i{=}1]}{Pr[S_{i3}{=}s_3'|S_{i2}{=}s_2, X_{i2}{=}x_2, \tilde{H}_{i3}{=}h_3, \bar{M}_i{=}1]} \times \frac{Pr[S_{i3}{=}s_3'|S_{i2}{=}s_2', X_{i2}{=}x_2, \tilde{H}_{i3}{=}h_3, \bar{M}_i{=}1]}{Pr[S_{i3}{=}s_3|S_{i2}{=}s_2', X_{i2}{=}x_2, \tilde{H}_{i3}{=}h_3, \bar{M}_i{=}1]}.$$

Proof. An implication of Lemma W15 is that $Pr[S_{i4}, S_{i3}|S_{i2}, S_{i1}, \tilde{H}_{i3}{=}h_3, \bar{M}_i{=}1]$ is identified. Below, we drop all i subscripts and the conditioning on $\tilde{H}_{i3}{=}h_3$ and $\bar{M}_i{=}1$ to increase readability. We also focus on the case where the number of points of support in S_{it} is the same as the number of points of support in X_{it}. This can be extended to allow for larger support for S_{it} by adding a singular value decomposition as in Lemma W15. Such an extension, while being straightforward, makes the notation more cumbersome and hence we omit it.

The first step is to manipulate $Pr[S_4, S_3|S_2, S_1]$, following Hu and Shum

(2012):

$$
\begin{aligned}
&Pr[S_4, S_3|S_2, S_1] \\
&\quad = \sum_{X_2}\sum_{X_1} Pr[S_4, S_3, X_2, X_1|S_2, S_1] \\
&\quad = \sum_{X_2}\sum_{X_1} Pr[S_4|S_3, S_2, S_1, X_2, X_1] \cdot Pr[S_3, X_2|S_2, S_1, X_1] \cdot Pr[X_1|S_2, S_1] \\
&\quad = \sum_{X_2}\sum_{X_1} Pr[S_4|S_3, X_2] \cdot Pr[S_3, X_2|S_2, S_1, X_1] \cdot Pr[X_1|S_2, S_1] \\
&\quad = \sum_{X_2}\sum_{X_1} Pr[S_4|S_3, X_2] \cdot Pr[S_3|S_2, S_1, X_2, X_1] \cdot Pr[X_2|S_2, S_1, X_1] \cdot Pr[X_1|S_2, S_1] \\
&\quad = \sum_{X_2}\sum_{X_3} Pr[S_4|S_3, X_2] \cdot Pr[S_3|S_2, X_2, X_1] \cdot Pr[X_2|S_2, S_1, X_1] \cdot Pr[X_1|S_2, S_1] \\
&\quad = \sum_{X_2}\sum_{X_1} Pr[S_4|S_3, X_2] \cdot Pr[S_3|S_2, X_2, X_1] \cdot Pr[X_2, X_1|S_2, S_1] \\
&\quad = \sum_{X_2}\sum_{X_1} Pr[S_4|S_3, X_2] \cdot Pr[S_3|S_2, X_2] \cdot Pr[X_2, X_1|S_2, S_1] \\
&\quad = \sum_{X_2} Pr[S_4|S_3, X_2] \cdot Pr[S_3|S_2, X_2] \cdot \sum_{X_1} Pr[X_2, X_1|S_2, S_1] \\
&\quad = \sum_{X_2} Pr[S_4|S_3, X_2] \cdot Pr[S_3|S_2, X_2] \cdot Pr[X_2|S_2, S_1].
\end{aligned}
$$

This manipulation relies on $Pr[S_4|S_3, S_2, S_1, X_2, X_1] = Pr[S_4|S_3, X_2]$, which follows from the Markovian property of the contract where w_{t+1} is determined by (x_t, z_t, V_t) together with the fact that, in the absence of flat regions of $J(x, z, V)$, conditioning on the wage w_t is equivalent to conditioning on V_t since $\frac{1}{u'(w_t)} = -J'(x_t, z_t, V_t)$. Having the sequence of firm shocks in the conditioning set is essential because without it, the contract would lose its Markovian structure. The wage process is still Markovian if a worker moves between periods 2 and 3 or 3 and 4 because the underlying match quality is reset to z_0 and hence there is still no time dependence. The same argument applies to $Pr[S_3|S_2, S_1, X_2, X_1] = Pr[S_3|S_2, X_2, X_1]$ and to the limited feedback property that allows us to use $Pr[S_3|S_2, X_2, X_1] = Pr[S_3|S_2, X_2]$.[4]

[4]See Hu and Shum (2012) for a precise definition of limited feedback.

In a second step, we continue by defining the following matrices:

$$\begin{aligned}\left[L_{S_4,s_3|s_2,S_1}\right]_{pq} &= \Pr[S_4{=}s_p, S_3{=}s_3|S_2{=}s_2, S_1{=}s_q]\\ \left[L_{S_4|s_3,X_2}\right]_{pq} &= \Pr[S_4{=}s_p|S_3{=}s_3, X_2{=}s_q]\\ \left[L_{X_2|s_2,S_1}\right]_{pq} &= \Pr[X_2{=}x_p|S_2{=}s_2, S_1{=}s_q],\end{aligned}$$

as well as a diagonal matrix $D_{s_3|s_2,X_2}$ with elements:

$$\left[D_{s_3|s_2,X_2}\right]_{pq} = \mathbf{1}\{p{=}q\} \cdot \Pr[S_3{=}s_3|S_2{=}s_2, X_2{=}x_p].$$

The result of the first step in terms of these matrices for (s_1, s_2, s_1', s_2') is:

$$\begin{aligned}L_{S_4,s_3|s_2,S_1} &= L_{S_4|s_3,X_2} D_{s_3|s_2,X_2} L_{X_2|s_2,S_1} \qquad (2)\\ L_{S_4,s_3'|s_2,S_1} &= L_{S_4|s_3',X_2} D_{s_3'|s_2,X_2} L_{X_2|s_2,S_1}\\ L_{S_4,s_3'|s_2',S_1} &= L_{S_4|s_3',X_2} D_{s_3'|s_2',X_2} L_{X_2|s_2',S_1}\\ L_{S_4,s_3|s_2',S_1} &= L_{S_4|s_3,X_2} D_{s_3|s_2',X_2} L_{X_2|s_2',S_1}.\end{aligned}$$

Since assumption i) ensures that these matrices are invertible, we compute:

$$L_{S_4,s_3|y_2,Y_1} L^{-1}_{S_4,s_3'|y_2,Y_1} \left(L_{S_4,s_3'|y_2',Y_1} L^{-1}_{S_4,s_3|y_2',Y_1} \right) = L_{S_4|s_3,X_2} \tilde{D} L^{-1}_{S_4|s_3,X_2},$$

where

$$\tilde{D} = D_{s_3|y_2,X_2} D^{-1}_{s_3'|y_2,X_2} D_{s_3'|y_2',X_2} D^{-1}_{s_3|y_2',X_2}.$$

Hence, as long as the diagonal elements of $\tilde{D}$ are distinct, as guaranteed by condition ii), the eigenvalue decomposition of the left hand side identifies $\Pr[S_4{=}s_4|S_3{=}s_3, \tilde{X}_2{=}x_2]$ as eigenvectors up to a permutation of the values x_2 that are specific to $(s_3, s_3', s_2, s_2', \tilde{H}_{i3})$ and a scaling factor. The scaling factor is pinned down by the fact that the probabilities have to sum to 1. In addition, noting

$$\left(L^{-1}_{S_4,s_3'|s_2',S_1} L_{S_4,s_3'|s_2,S_1} \right)^{-1} L^{-1}_{S_4,s_3|s_2',S_1} L_{S_4,s_3|s_2,S_1} = L^{-1}_{X_2|s_2,S_1} \tilde{D} L_{X_2|s_2,S_1},$$

this shows that the same ordering of eigenvalues delivers $L_{X_2|s_2,S_1}$ with the same permutation of X_2, i.e. $Pr[\tilde{X}_2|S_2,S_1]$. Combining these as in equation (2) gives $D_{s_3|s_2,X_2}$, which is our third object of interest $Pr[S_3|S_2,\tilde{X}_2]$. □

Lemma W17 (Labeling x within h)**.** *For each history h_3, we can align the $\sigma_{s_3 s_2 h_3}(\cdot)$ permutations of X_{i2} across values of (s_3, s_2) if:*

i) For each history h_3 and any x_2, $x_2' \neq x_2$ and s_3, there exists s_4 such that

$$Pr[S_{i4}{=}s_4|S_{i3}{=}s_3, X_{i2}{=}x_2, \tilde{H}_{i3}{=}h_3, \bar{M}_i{=}1] \\ \neq Pr[S_{i4}{=}s_4|S_{i3}{=}s_3, X_{i2}{=}x_2', \tilde{H}_{i3}{=}h_3, \bar{M}_i{=}1].$$

ii) For each history h_3 and any x_2, $x_2' \neq x_2$ and s_2, there exists s_1 such that

$$Pr[X_{i2}{=}x_2|S_{i2}{=}s_2, S_{i1}{=}s_1, \tilde{H}_{i3}{=}h_3, \bar{M}_i{=}1] \\ \neq Pr[X_{i2}{=}x_2'|S_{i2}{=}s_2, S_{i1}{=}s_1, \tilde{H}_{i3}{=}h_3, \bar{M}_i{=}1].$$

Proof. We start matching the labeling of X_{i2} within values of (s_3, h_3) by using the identified $\Pr[S_{i4}{=}s_4|S_{i3}{=}s_3, \tilde{X}_{i2}{=}x_2, \tilde{H}_{i3}{=}h_3, \bar{M}_i{=}1]$ from Lemma W16. Taking two values $s_2 \neq s_2'$ for a given s_3 and h_3, we can now pair vectors using condition i), which guarantees that only the same x_2 will be equal in $\Pr[S_{i4}{=}s_4|S_{i3}{=}s_3, X_{i2}{=}x_2, H_{i3}{=}h_3, \bar{M}_i{=}1]$ at all s_4. This resolves the labeling of X_{i2} across s_2 within values s_3.

Next, we turn to X_{i2} permutations across s_3 values. For this, we use $Pr[\tilde{X}_{i2}|S_{i2}, S_{i1}, \tilde{H}_{i3}, \bar{M}_i{=}1]$ from Lemma W16 and fix a common (s_2, h_3). For two different $s_3 \neq s_3'$, condition ii) allows us to match the permutation over x_2 values because it ensures that for any two values $x_2 \neq x_2'$, the corresponding vectors $Pr[\tilde{X}_{i2}{=}x_2|S_{i2}{=}s_2, S_{i1}{=}s_1, \tilde{H}_{i3}{=}h_3, \bar{M}_i{=}1]$ and $Pr[\tilde{X}_{i2}{=}x_2'|S_{i2}{=}s_2, S_{i1}{=}s_1, \tilde{H}_{i3}{=}h_3, \bar{M}_i{=}1]$ differ in at least one s_1 value. This means that all permutations of X_{i2} across different s_3 are labeled. □

Lemma W18 (Labeling x across h)**.** *We can align the $\sigma_{s_3 s_2 h_3}(\cdot)$ permutations of X_{i2} across h_3 if:*

i) For any (z_1, z_2) and $(z_1', z_2') \neq (z_1, z_2)$, there exists (s_3, s_2, s_1) such that

$$Pr[S_{i3}{=}s_3 | S_{i2}{=}s_2, S_{i1}{=}s_1, Z_{i1}{=}z_1, Z_{i2}{=}z_2, \bar{M}_i{=}1] \\ \neq Pr[S_{i3}{=}s_3 | S_{i2}{=}s_2, S_{i1}{=}s_1, Z_{i1}{=}z_1', Z_{i2}{=}z_2', \bar{M}_i{=}1].$$

ii) For any x_2, $x_2' \neq x_2$ and (z_1, z_2), there exists (s_1, s_2) such that

$$Pr[\tilde{X}_{i2}{=}x_2 | S_{i2}{=}s_2, S_{i1}{=}s_1, Z_{i1}{=}z_1, Z_{i2}{=}z_2, \bar{M}_i{=}1] \\ \neq Pr[\tilde{X}_{i2}{=}x_2' | S_{i2}{=}s_2, S_{i1}{=}s_1, Z_{i1}{=}z_1, Z_{i2}{=}z_2, \bar{M}_i{=}1].$$

iii) For any z_3 and $z_3' \neq z_3$, there exists (s_3, s_4) such that

$$\sum_{x_2} Pr[S_{i4}{=}s_4 | S_{i3}{=}s_3, \tilde{X}_{i2}{=}x_2, Z_{i3}{=}z_3, \bar{M}_i{=}1] \\ \neq \sum_{x_2} Pr[S_{i4}{=}s_4 | S_{i3}{=}s_3, \tilde{X}_{i2}{=}x_2, Z_{i3}{=}z_3', \bar{M}_i{=}1].$$

iv) For any x_2, $x_2 \neq x_2'$ and z_3, there exists (s_3, s_4) such that

$$Pr[S_{i4}{=}s_4 | S_{i3}{=}s_3, \tilde{X}_{i2}{=}x_2, Z_{i3}{=}z_3, \bar{M}_i{=}1] \\ \neq Pr[S_{i4}{=}s_4 | S_{i3}{=}s_3, \tilde{X}_{i2}{=}x_2', Z_{i3}{=}z_3, \bar{M}_i{=}1].$$

Proof. First, we want to align the labeling of X_{i2} across z_3 for fixed values

(z_1, z_2). To tell which h_3 histories share the same (z_1, z_2), we construct:

$$\sum_{x_2} Pr[S_{i3}|S_{i2}, \tilde{X}_{i2}{=}x_2, \tilde{H}_{i3}, \bar{M}_i{=}1] \cdot Pr[\tilde{X}_{i2}{=}x_2|S_{i2}, S_{i1}, \tilde{H}_{i3}, \bar{M}_i{=}1]$$
$$= \sum_{x_2} Pr[S_{i3}|S_{i2}, \tilde{X}_{i2}{=}x_2, Z_{i1}, Z_{i2}, \bar{M}_i{=}1] \cdot \frac{Pr[Z_{i3}|S_{i2}, S_{i1}, \tilde{X}_{i2} = x_2, Z_{i1}, Z_{i2}, \bar{M}_i{=}1]}{Pr[Z_{i3}|S_{i2}, S_{i1}, Z_{i1}, Z_{i2}, \bar{M}_i{=}1]}$$
$$\times Pr[\tilde{X}_{i2} = x_2|S_{i2}, S_{i1}, Z_{i1}, Z_{i2}, \bar{M}_i{=}1]$$
$$= \sum_{x_2} Pr[S_{i3}|S_{i2}, \tilde{X}_{i2}{=}x_2, Z_{i1}, Z_{i2}, \bar{M}_i{=}1] \cdot \frac{Pr[Z_{i3}|Z_{i2}, \bar{M}_i{=}1]}{Pr[Z_{i3}|Z_{i2}, \bar{M}_i{=}1]}$$
$$\times Pr[\tilde{X}_{i2} = x_2|S_{i2}, S_{i1}, Z_{i1}, Z_{i2}, \bar{M}_i{=}1]$$
$$= \sum_{x_2} Pr[S_{i3}|S_{i2}, \tilde{X}_{i2}{=}x_2, Z_{i1}, Z_{i2}, \bar{M}_i{=}1] \cdot Pr[\tilde{X}_{i2}{=}x_2|S_{i2}, S_{i1}, Z_{i1}, Z_{i2}, \bar{M}_i{=}1]$$
$$= Pr[S_{i3}|S_{i2}, S_{i1}, Z_{i1}, Z_{i2}, \bar{M}_i{=}1],$$

where all probabilities in the first line have already been identified. Condition i) states that $Pr[S_{i3}|S_{i2}, S_{i1}, Z_{i1}, Z_{i2}, \bar{M}_i{=}1]$ is separable, hence we can partition the h_3 histories into subgroups with identical (z_1, z_2) without knowing the actual values of the pair (z_1, z_2).

We label the values x_2 across z_3 following the same procedure used across s_3 values in Lemma W17. For two different histories h_3 and h_3' with the same (z_1, z_2), we compute $Pr[\tilde{X}_{i2}|S_{i2}, S_{i1}, Z_{i1}{=}z_1, Z_{i2}{=}z_2, \bar{M}_i{=}1]$. Taking a given value x_2 in h_3, condition ii) ensures that there is only one value for $\tilde{X}_{i2}$ in h_3' with identical $Pr[\tilde{X}_{i2}{=}x_2|S_{i2}{=}s_2, S_{i1}{=}s_1, Z_{i1}{=}z_1, Z_{i2}{=}z_2, \bar{M}_i{=}1]$ for all (s_1, s_2), and this value is the same x_2. Hence we have now aligned the values x_2 across different z_3 for each value pair (z_1, z_2).

Next, we observe:

$$
\begin{aligned}
&Pr[S_{i4}|S_{i3}, \tilde{X}_{i2}, \tilde{H}_{i3}, \bar{M}_i{=}1] \\
&\quad = \sum_{x_3} Pr[S_{i4}, \tilde{X}_{i3}{=}x_3|S_{i3}, \tilde{X}_{i2}, \tilde{H}_{i3}, \bar{M}_i{=}1] \\
&\quad = \sum_{x_3} Pr[S_{i4}|S_{i3}, \tilde{X}_{i3}{=}x_3, \tilde{X}_{i2}, \tilde{H}_{i3}, \bar{M}_i{=}1] \cdot Pr[\tilde{X}_{i3}{=}x_3|S_{i3}, \tilde{X}_{i2}, \tilde{H}_{i3}, \bar{M}_i{=}1] \\
&\quad = \sum_{x_3} Pr[S_{i4}|S_{i3}, \tilde{X}_{i3}{=}x_3, \tilde{X}_{i2}, Z_{i3}, \bar{M}_i{=}1] \cdot Pr[\tilde{X}_{i3}{=}x_3|S_{i3}, \tilde{X}_{i2}, Z_{i3}, \bar{M}_i{=}1] \\
&\quad = Pr[S_{i4}|S_{i3}, \tilde{X}_{i2}, Z_{i3}, \bar{M}_i{=}1].
\end{aligned}
$$

For each h_3 we can construct $\sum_{x_2} Pr[S_{i4}{=}s_4|S_{i3}{=}s_3, \tilde{X}_{i2}{=}x_2, Z_{i3}{=}z_3, \bar{M}_i{=}1]$. This quantity has two important properties. On the one hand, it does not depend on the ordering of x_2 values, and on the other hand, it does not depend on (z_1, z_2). Then condition iii) allows us to partition the h_3 histories into groups with common z_3 values by looking across values of (S_{i4}, S_{i3}).

With this in hand, we take two histories h_3 and h_3' with identical z_3, and compute for a given x_2 in h_3 the associated $Pr[S_{i4}{=}s_4|S_{i3}{=}s_3, X_{i2}{=}x_2, Z_{i3}{=}z_3, \bar{M}_i{=}1]$. Condition iv) guarantees that only one value of $\tilde{X}_{i2}$ in h_3', i.e. the same x_2, will have the exact same $Pr[S_{i4}{=}s_4|S_{i3}{=}s_3, \tilde{X}_{i2}{=}x_2, Z_{i3}{=}z_3, \bar{M}_i{=}1]$ for all (s_4, s_3). This allows us to align the values x_2 across h_3 within the same z_3.

Finally, we conclude that aligning all h_3 with common (z_1, z_2) as well as all h_3 with common z_3 in fact aligns the $\sigma_{s_3 s_2 h_3}(\cdot)$ permutations across all h_3. □

Lemma W19 (Labeling z)**.** *We can identify the values z_2 for each history h_3 if there exists x_2 such that for all (z_2, z_2') the target wages $w^*(x_2, z_2)$ and $w^*(x_2, z_2')$ lie at different y_2 values. In addition, we can identify the values of z_3, and hence $Pr[Z_{i3}|Z_{i2}]$ under the assumption of diagonal dominance.*

Proof. We rely on the monotonicity property of $w^*(x, z)$ in z within a given value x. If the current wage y_2 is below $w^*(x_2, z_2)$, the wage will increase between periods 2 and 3, and if it is above $w^*(x_2, z_2)$, the wage will decrease. Hence for a fixed value x_2 and for each h_3, we can get the bin of y_2 that includes $w^*(x_2, z_2)$ by computing

$$y_2^*(h_3, x_2) = \max y_2$$
$$\text{s.t. } Pr[Y_{i3} < Y_{i2} | \tilde{X}_{i2}{=}x_2, Y_{i2}{=}y_2, \tilde{H}_{i3}{=}h_3, M_{i3}{=}0, \bar{M}_i{=}1]{=}0.$$

For any history h_3 we have thus recovered the associated value of the target wage. As long as there is a value x_2 for which $w^*(x_2, z_2)$ and $w^*(x_2, z_2')$ are in different y_2 cells, we can order the $y_2^*(h_3, x_2)$ values across values of h_3, and given the monotonicity of the target wage in match quality, this gives us the values of z_2 for each history h_3. Simply put, calling $z_2(h_3)$ the value of z_2 in h_3 and for the particular x_2 from the assumption, we get that $z_2(h_3) = \frac{n_z}{n_h} \sum_{h_3'} 1[y_2^*(h_3', x_2) < y_2^*(h_3, x_2)]$. Here, it is key to be able to correctly label the values x_2 across histories h_3.

From Lemma W18 we already know which histories h_3 have a common z_3. Take such a set of histories that share a given z_3. Find the h_3 in that set such that $Pr[\tilde{H}_{i3}{=}h_3 | Z_{i2}{=}z_2(h_3)] > Pr[\tilde{H}_{i3}{=}h_3' | Z_{i2}{=}z_2(h_3)]$ for all possible h_3' unconditionally of all other variables. For this particular h_3 we know from diagonal dominance that $z_3(h_3){=}z_2(h_3)$. This pins down the z_3 value for the whole set. Given that we now know z_2 and z_3 for all h_3, we can construct the transition matrix $Pr[Z_{i3}|Z_{i2}]$. □

Lemma W20 (Law of motion of x)**.** *Using the identified $Pr[S_{i4}|S_{i3}, \tilde{X}_{i2}, \tilde{H}_{i3}, \bar{M}_i{=}1]$ and $Pr[S_{i3}|S_{i2}, \tilde{X}_{i2}, \tilde{H}_{i3}, \bar{M}_i{=}1]$, we recover $Pr[X_{i3}|X_{i2}]$ up to a common labeling if the matrix of $Pr[S_{i3}|S_{i2}, \tilde{X}_{i2}, \tilde{H}_{i3}, \bar{M}_i{=}1]$ with $(S_{i3}, S_{i2}, \tilde{H}_{i3})$ in rows and $\tilde{X}_{i2}$ in columns has full column rank.*

Proof. Using the identified $Pr[S_{i4}|S_{i3}, \tilde{X}_{i2}, \tilde{H}_{i3}, \bar{M}_i{=}1]$ we write for all (s_4, h_3, x_2)

and all $s_3 = (y_3, m_3)$ with $m_3{=}0$:

$$
\begin{aligned}
&Pr[S_{i4}{=}s_4|S_{i3}{=}s_3, \tilde{X}_{i2}{=}x_2, \tilde{H}_{i3}{=}h_3, \bar{M}_i{=}1] \\
&\quad = \sum_{x_3} Pr[S_{i4}{=}s_4|S_{i3}{=}s_3, \tilde{X}_{i3}{=}x_3, \tilde{X}_{i2}{=}x_2, \tilde{H}_{i3}{=}h_3, \bar{M}_i{=}1] \\
&\qquad\qquad \times Pr[\tilde{X}_{i3}{=}x_3|S_{i3}{=}s_3, \tilde{X}_{i2}{=}x_2, \tilde{H}_{i3}{=}h_3, \bar{M}_i{=}1] \\
&\quad = \sum_{x_3} Pr[S_{i3}{=}s_4|S_{i2}{=}s_3, \tilde{X}_{i2}{=}x_3, \tilde{H}_{i3}{=}h_3, \bar{M}_i{=}1] \cdot Pr[\tilde{X}_{i3}{=}x_3|\tilde{X}_{i2}{=}x_2],
\end{aligned}
$$

where the last line is derived from the following two considerations. First, we can manipulate $Pr[S_{i4}{=}s_4|S_{i3}{=}s_3, \tilde{X}_{i3}{=}x_3, \tilde{X}_{i2}{=}x_2, \tilde{H}_{i3}{=}h_3, \bar{M}_i{=}1]$:

$$
\begin{aligned}
&Pr[S_{i4}{=}s_4|S_{i3}{=}s_3, \tilde{X}_{i3}{=}x_3, \tilde{X}_{i2}{=}x_2, \tilde{H}_{i3}{=}h_3, \bar{M}_i{=}1] \\
&\quad = Pr[S_{i4}{=}s_4|Y_{i3}{=}y_3, M_{i3}{=}0, \tilde{X}_{i3}{=}x_3, \tilde{X}_{i2}{=}x_2, \tilde{H}_{i3}{=}h_3, M_{i2}{=}0, M_{i1}{\in}\{0,3\}] \\
&\quad = Pr[S_{i4}{=}s_4|Y_{i3}{=}y_3, M_{i3}{=}0, \tilde{X}_{i3}{=}x_3, Z_{i3}{=}z_3(h_3)] \\
&\quad = Pr[S_{i3}{=}s_4|Y_{i2}{=}y_3, M_{i2}{=}0, \tilde{X}_{i2}{=}x_3, Z_{i2}{=}z_3(h_3)] \\
&\quad = Pr[S_{i3}{=}s_4|S_{i2}{=}s_3, \tilde{X}_{i2}{=}x_3, Z_{i2}{=}z_3(h_3), \bar{M}_i{=}1] \\
&\quad = Pr[S_{i3}{=}s_4|S_{i2}{=}s_3, \tilde{X}_{i2}{=}x_3, \tilde{H}_{i3}{=}h_3, \bar{M}_i{=}1],
\end{aligned}
$$

where the Markovian property of the contract gives us that $Pr[S_{i4}|S_{i3}, X_{i3}, H_{i3}, \bar{M}_i{=}1] = Pr[S_{i4}|S_{i3}, X_{i3}, Z_{i3}]$ and stationarity of the environment insures that $Pr[S_{i4}|S_{i3}, X_{i3}, Z_{i3}]$ and $Pr[S_{i3}|S_{i2}, X_{i2}, Z_{i2}]$ are the same. Second, $Pr[\tilde{X}_{i3}{=}x_3|S_{i3}{=}s_3, \tilde{X}_{i2}{=}x_2, \tilde{H}_{i3}{=}h_3, \bar{M}_i{=}1]$ simplifies to:

$$
\begin{aligned}
&Pr[\tilde{X}_{i3}{=}x_3|S_{i3}{=}s_3, \tilde{X}_{i2}{=}x_2, \tilde{H}_{i3}{=}h_3, \bar{M}_i{=}1] \\
&\quad = \frac{Pr[\tilde{X}_{i3}{=}x_3, S_{i3}{=}s_3|\tilde{X}_{i2}{=}x_2, \tilde{H}_{i3}{=}h_3, \bar{M}_i{=}1]}{Pr[S_{i3}{=}s_3|\tilde{X}_{i2}{=}x_2, \tilde{H}_{i3}{=}h_3, \bar{M}_i{=}1]} \\
&\quad = \sum_{s_2} \frac{Pr[\tilde{X}_{i3}{=}x_3, S_{i3}{=}s_3, S_{i2}{=}s_2|\tilde{X}_{i2}{=}x_2, \tilde{H}_{i3}{=}h_3, \bar{M}_i{=}1]}{Pr[S_{i3}{=}s_3|\tilde{X}_{i2}{=}x_2, \tilde{H}_{i3}{=}h_3, \bar{M}_i{=}1]} \\
&\quad = \sum_{s_2} Pr[\tilde{X}_{i3}{=}x_3, S_{i3}{=}s_3|S_{i2}{=}s_2, \tilde{X}_{i2}{=}x_2, \tilde{H}_{i3}{=}h_3, \bar{M}_i{=}1] \cdot \frac{Pr[S_{i2}{=}s_2|\tilde{X}_{i2}{=}x_2, \tilde{H}_{i3}{=}h_3, \bar{M}_i{=}1]}{Pr[S_{i3}{=}s_3|\tilde{X}_{i2}{=}x_2, \tilde{H}_{i3}{=}h_3, \bar{M}_i{=}1]} \\
&\quad = \sum_{s_2} Pr[\tilde{X}_{i3}{=}x_3|S_{i2}{=}s_2, \tilde{X}_{i2}{=}x_2, \tilde{H}_{i3}{=}h_3, \bar{M}_i{=}1] \cdot Pr[S_{i3}{=}s_3|S_{i2}{=}s_2, \tilde{X}_{i2}{=}x_2, \tilde{H}_{i3}{=}h_3, \bar{M}_i{=}1] \\
&\qquad\qquad \times \frac{Pr[S_{i2}{=}s_2|\tilde{X}_{i2}{=}x_2, \tilde{H}_{i3}{=}h_3, \bar{M}_i{=}1]}{Pr[S_{i3}{=}s_3|\tilde{X}_{i2}{=}x_2, \tilde{H}_{i3}{=}h_3, \bar{M}_i{=}1]} \\
&\quad = \sum_{s_2} Pr[\tilde{X}_{i3}{=}x_3|S_{i2}{=}s_2, \tilde{X}_{i2}{=}x_2, Z_{i2}{=}z_2(h_3)] \cdot \frac{Pr[S_{i3}{=}s_3, S_{i2}{=}s_2|\tilde{X}_{i2}{=}x_2, \tilde{H}_{i3}{=}h_3, \bar{M}_i{=}1]}{Pr[S_{i3}{=}s_3|\tilde{X}_{i2}{=}x_2, \tilde{H}_{i3}{=}h_3, \bar{M}_i{=}1]} \\
&\quad = \sum_{s_2} Pr[\tilde{X}_{i3}{=}x_3|\tilde{X}_{i2}{=}x_2] \cdot \frac{Pr[S_{i3}{=}s_3, S_{i2}{=}s_2|\tilde{X}_{i2}{=}x_2, \tilde{H}_{i3}{=}h_3, \bar{M}_i{=}1]}{Pr[S_{i3}{=}s_3|\tilde{X}_{i2}{=}x_2, \tilde{H}_{i3}{=}h_3, \bar{M}_i{=}1]} \\
&\quad = Pr[\tilde{X}_{i3}{=}x_3|\tilde{X}_{i2}{=}x_2],
\end{aligned}
$$

where X_{i3} is independent of $\bar{M}_i$ due to its Markovianity. Furthermore, $\tilde{X}_{i3}$ and S_{i3} are independent of each other given (S_{i2}, X_{i2}, Z_{i2}) in the optimal contract.

Hence, we get a linear system in $Pr[\tilde{X}_{i3}|\tilde{X}_{i2}]$ and the linear independence of $Pr[S_{i3}|S_{i2}, \tilde{X}_{i2}, \tilde{H}_{i3}, \bar{M}_i{=}1]$ across values of $\tilde{X}_{i2}$ is sufficient to guarantee recovering $Pr[X_{i3}|X_{i2}]$ up to a common permutation. □

Lemma W21 (Stationary laws of motion)**.** *From Lemmas W16, W19 and W20 we identify* $Pr[Y_{i,t+1}, M_{i,t+1}|Y_{it}, \tilde{X}_{it}, Z_{it}]$ *for* $Y_{it}{\neq}0$, $Pr[Z_{i,t+1}|Z_{it}]$ *and* $Pr[\tilde{X}_{i,t+1}|\tilde{X}_{it}]$.

Proof. The identified $Pr[S_{i3}|S_{i2}, \tilde{X}_{i2}, \tilde{H}_{i3}, \bar{M}_i{=}1]$ is equal to $Pr[S_{i3}|S_{i2}, \tilde{X}_{i2}, Z_{i2}]$ with $M_{i2}{=}0$ due to the Markovian structure of the model, see Lemma W20. For any $Y_{i2}{>}0$, the contract is also independent of M_{i2} given (S_{i2}, X_{i2}, Z_{i1}). Since this delivers $Pr[S_{i3}|Y_{i2}, \tilde{X}_{i2}, Z_{i2}]$ with $Y_{i2}{>}0$, stationarity of the model allows us to generalize and hence we identify $Pr[Y_{i,t+1}, M_{i,t+1}|Y_{it}, \tilde{X}_{it}, Z_{it}]$ for $Y_{it}{>}0$. Similarly, stationarity of the structural model also allows us to conclude that $Pr[Z_{i3}|Z_{i2}] = Pr[Z_{i,t+1}|Z_{it}]$ and $Pr[\tilde{X}_{i3}|\tilde{X}_{i2}] = Pr[\tilde{X}_{i,t+1}|\tilde{X}_{it}]$, recovered in Lemmas W19 and W20. □

Lemma W22 (Worker expected present value)**.** *With known utility function $u(\cdot)$, cost function $c(\cdot)$ and discount factor β, and in the absence of flat regions in the Pareto frontier, we show that the present value of the worker, $V(x,z,w)$, is uniquely defined from the transition probabilities of Lemma W21.*

Proof. In the absence of flat regions in the Pareto frontier, we can use the wage, w, as a state instead of the promised value, V, and thus express the expected worker value as $V(x,z,w)$, which we aim to identify at any given state (x,z,w). With $w'(x,z,w)$ denoting the wage function, recall from the model section:

$$\begin{aligned} V(x,z,w) &= \sup_{v_1,e} u(w) - c(e) + \beta\delta(e)\mathbb{E}_{x'}[U(x')|x] + \beta(1-\delta(e))\kappa p(\theta(x,v_1))v_1 \\ &\qquad + \beta(1-\delta(e))(1-\kappa p(\theta(x,v_1)))\mathbb{E}_{x'z'}[V(x',z',w'(x,z,w))|x,z] \\ &= u(w) - c(e^*) + \beta\delta^*\mathbb{E}_{x'}[U(x')|x] + (1-\delta^*)\beta\kappa p_1^* \cdot v_1(x,z,w) \\ &\qquad + \beta(1-\delta^*)(1-\kappa p_1^*)\mathbb{E}_{x'z'}[V(x',z',w'(x,z,w))|x,z], \end{aligned}$$

where we abstract from the lottery and substitute in the optimal policy (δ^*, e^*, p_1^*).

We now replace expectations and present values with empirical counterparts and construct a recursive expression for $V(x,z,w)$. Note that we can write $v_1(x,z,w)$ and $v_0(x)$ as functions of the empirical transitions from Lemma W21:

$$\begin{aligned} v_1(x,z,w) &= \mathbb{E}_{x'w'}[V(x',z_0,w')|X{=}x, Z{=}z, Y{=}w, M'{=}1] \\ v_0(x) &= \mathbb{E}_{x'w'}[V(x',z_0,w')|X{=}x, M'{=}3], \end{aligned}$$

where the expectations are taken with respect to $Pr[X_{i,t+1}|X_{it}]$ and

$Pr[S_{it}|S_{i,t-1}, X_{t-1}, Z_{t-1}]$. Replacing them in $V(x, z, w)$ gives:

$$\begin{aligned} V(x, z, w) = u(w) - c(e^*) + \beta\delta^*\mathbb{E}_{x'}[U(x')|x] \\ + \beta(1-\delta^*)\kappa p_1^*\mathbb{E}_{x'w'}[V(x', z_0, w')|x, z, w, M'{=}1] \\ + \beta(1-\delta^*)(1-\kappa p_1^*)\mathbb{E}_{x'z'w'}[V(x', z', w')|x, z, w, M'{=}0], \end{aligned}$$

where the unknowns are the functions U and V. To get $U(x)$ note that in the model, as effort approaches zero, the worker is indifferent between working and not working. With the previously imposed normalization $\delta(e) = 1 - e$, this point of indifference is where the job destruction probability δ approaches one. Let's then define:

$$\begin{aligned} \widetilde{V}(x, z, w) \equiv \kappa p_1^*\mathbb{E}_{x'w'}[V(x', z_0, w')|x, z, w, M'{=}1] \\ + (1-\kappa p_1^*)\mathbb{E}_{x'z'w'}[V(x', z', w')|x, z, w, M'{=}0], \end{aligned}$$

and call $\underline{w}(x, z)$ the wage such that:

$$\underline{w}(x, z) \equiv \arg\min_w \delta^*(x, z, w) \quad \text{s.t.} \quad \delta^*(x, z, w) < 1.$$

The first order condition $c'(e^*) = \beta\widetilde{V}(x, z, w) - \beta\mathbb{E}_{x'}[U(x')|x]$ together with $c'(0) = 0$ then implies $\mathbb{E}_{x'}[U(x')|x] = \widetilde{V}(x, z, \underline{w}(x, z))$, which we plug in:

$$\begin{aligned} V(x, z, w) = u(w) - c(1-\delta^*) \\ + \beta\delta^*\kappa p_1^*\mathbb{E}_{x'w'}[V(x', z_0, w'|x, z, \underline{w}(x, z), M'{=}1] \\ + \beta\delta^*(1-\kappa p_1^*)\mathbb{E}_{x'z'w'}[V(x', z', w')|x, z, \underline{w}(x, z), M'{=}0] \\ + \beta(1-\delta^*)\kappa p_1^*\mathbb{E}_{x'w'}[V(x', z_0, w')|x, z, w, M'{=}1] \\ + \beta(1-\delta^*)(1-\kappa p_1^*)\mathbb{E}_{x'z'w'}[V(x', z', w')|x, z, w, M'{=}0]. \end{aligned}$$

This mapping expresses $V(x, z, w)$ as an integral equation and satisfies the Blackwell-Boyd conditions of discounting and monotonicity. We thus establish uniqueness of the identified value function of the worker. □

Remark W1 (Identifying $c(\cdot)$)**.** *We can go one step further and find the function $c(\cdot)$ itself. Starting again from the effort decision,*

$$c'(1-\delta^*(x,z,w)) = \beta\widetilde{V}(x,z,w) - \beta\mathbb{E}_{x'}[U(x')|x],$$

we multiply both sides by $\delta^*_w(x,z,w)$, the derivative of $\delta^*(x,z,w)$ with respect to w, and integrate from $\underline{w}(x,z)$ to w. This gives:

$$\begin{aligned}-c(1-\delta^*(x,z,w)) &= \beta\int_{\underline{w}(x,z)}^{w} \delta^*_w(x,z,u)\Big(\widetilde{V}(x,z,u) - \mathbb{E}_{x'}[U(x')|x]\Big)\ du \\ &= \beta\int_{\underline{w}(x,z)}^{w} \delta^*_w(x,z,u)\Big(\widetilde{V}(x,z,u) - \widetilde{V}(x,z,\underline{w}(x,z))\Big)\ du,\end{aligned}$$

which can be substituted back into the main equation to get:

$$\begin{aligned}V(x,z,w) = u(w) &+ \beta\int_{\underline{w}(x,z)}^{w} \delta^*_w(x,z,u)\Big(\widetilde{V}(x,z,u) - \widetilde{V}(x,z,\underline{w}(x,z))\Big)\ du \\ &+ \beta\delta^*\widetilde{V}(x,z,\underline{w}(x,z)) \\ &+ \beta(1-\delta^*)\widetilde{V}(x,z,w),\end{aligned}$$

where $\delta^*_w(x,z,u) < 0$. This appears to have the form of a Volterra equation of the second kind. Existence and uniqueness is then guaranteed under very mild conditions, see Evans (1911) and Abdou, Soliman, and Abdel-Aty (2020), and $V(x,z,w)$ is uniquely identified.

Lemma W23. *$f(x,z)$ is identified from $V(x,z,w)$ and the properties of the model if the transition rules of X_{it} and Z_{it} are invertible.*

Proof. We use $V(x,z,w)$ from Lemma W22 and the property that $J'(x,z,V) = -\frac{1}{u'(w)}$, which is integrated to identify $J(x,z,V)$ up to a (x,z)-specific constant $a(x,z)$. Denoting as $w(x,z,V)$ the inverse function of $V(x,z,w)$, we have:

$$J(x,z,V) = a(x,z) - \int_{V(x,z,w^*(x,z))}^{V} \frac{1}{u'(w(x,z,\omega))}\ \mathrm{d}\omega.$$

At the target wage $w^*(x,z)$ wages stay constant, $w'(x,z,w^*(x,z))=w^*(x,z)$, and expected firm profits are zero, $\mathbb{E}_{x'z'}[J(x',z',V(x',z',w^*(x,z))|x,z]=0$, so we get the following linear system for the intercepts $a(x,z)$:

$$\begin{aligned} 0 &= \sum_x \sum_z \mathbb{E}_{x'z'}[J(x',z',V(x',z',w^*(x,z))|x,z] \\ &= \sum_x \sum_z \mathbb{E}_{x'z'}[a(x',z')|x,z], \end{aligned}$$

where invertibility of the transition rules guarantees that all $a(x,z)$ are uniquely defined. This identifies the $J(x,z,V)$ function.

The final step is to use the Bellman equation of the firm's contracting problem to recover the production function:

$$f(x,z) = J(x,z,V(x,z,w^*(x,z)) + w^*(x,z).$$

□

W3 Data web appendix

W3.1 Institutional background

In this section we discuss the institutions associated with wage setting in Sweden during the years in the data. An important aspect of the Swedish labor market is the presence of Industrial Agreement (IA). In the 1990's many such agreements were put in place, specifying wage floors that were negotiated at the industry or firm level. A National Mediation Office was also established with the power to appoint mediators. Fredriksson and Topel (2010) presents a detailed picture of the different systems using sources from the Swedish Mediation Office Annual Report (2002). To get an overview, we briefly describe the different models of agreements with their share in the private sector:

1. Local bargain without restrictions (7%): wage increases are set fully locally between the employer and the employee.
2. Local bargain with a fallback (8%): wage increases are set locally, but if the parties cannot agree a central agreement specifies a general wage increase.

3. Local bargain with a fallback plus a guaranteed wage increase (16%): same as before, but with an additional minimum wage increase guaranteed by the central agreement.
4. Local wage frame without a guaranteed wage increase (12%): the local parties receive a total wage increase, but they can decide locally how this total increase is distributed across employees.
5. Local wage frame with guarantee or a fallback regulating the guarantee (28%): same as before, but with in addition either a guaranteed wage increase or at least a fallback in case an agreement cannot be reached.
6. General pay increase plus local wage frame (18%): a specified pay increase plus a total increase that can be split among employees in a way which is decided locally (as in model 4).
7. General pay increase (11%): a pay increase specified by the central agreement.

From these numbers, we note that at the two extremes, 11% of private sector workers are subject to a general pay increase and 7% bargain over their wages without any restrictions. The remaining 82% of agreements involve some level of local negotiation and hence the wage variation should reflect firm performance (please refer to Table 3.3 in Fredriksson and Topel (2010) for more details).

While all these institutions are in place, how flexible the realized wages are and how they relate to productivity remains an empirical question. Indeed Fredriksson and Topel (2010) says: "While the IA model may have delivered incentives for wage restraint at the aggregate level, it is reasonable to think that it has had a minor influence on the wage structure." Carlsson, Häkkinen Skans, and Nordstrom Skans (2019) find evidence of flexibility in response to local shocks, pointing to an ability of wages to adapt to local productivity. Taking stock, we believe the richness of the data in the Swedish economy, together with a significant level of wage adjustment at the employer level, provides a natural environment to study the contracting between employer and employee.

W3.2 Moments description

Based on the quarterly sample we compute the following transition rates:

$$Pr^{\text{U2E}} = \frac{\sum_i \sum_{q>1} 1\Big\{j_{iq} > 0 \text{ and } j_{iq-1} = 0\Big\}}{\sum_i \sum_{q>1} 1\Big\{j_{iq-1} = 0\Big\}}$$

$$Pr^{\text{J2J}} = \frac{\sum_i \sum_{q>1} 1\Big\{j_{iq} \neq j_{iq-1} \text{ and } j_{iq} > 0 \text{ and } j_{iq-1} > 0\Big\}}{\sum_i \sum_{q>1} 1\Big\{j_{iq} > 0 \text{ and } j_{iq-1} > 0\Big\}}$$

$$Pr^{\text{E2U}} = \frac{\sum_i \sum_{q>1} 1\Big\{j_{iq} = 0 \text{ and } j_{iq-1} > 0\Big\}}{\sum_i \sum_{q>1} 1\Big\{j_{iq-1} > 0\Big\}}.$$

Next, for convenience, we define the empirical mean, variance and covariances over a set S of observations, which are effectively conditional empirical expectations, for any random variables X_{it} and Y_{it}:

$$\mathbb{E}_S\left[X_{it}\right] = \frac{\sum_{(i,t)\in S} X_{it}}{\sum_{(i,t)\in S} 1}$$

$$\text{Var}_S\left[X_{it}\right] = \mathbb{E}_S\left[\left(X_{it} - \mathbb{E}_S\left[X_{it}\right]\right)^2\right]$$

$$\text{Cov}_S[X_{it}, Y_{it}] = \mathbb{E}_S\Big[(X_{it} - \mathbb{E}_S X_{it})(Y_{it} - \mathbb{E}_S Y_{it})\Big].$$

Let $\tau_i^{\text{U2E}}(1)$ and $\tau_i^{\text{U2E}}(2)$ be the first and last transition from unemployment to employment within the sample for worker i. Since we use earnings, we use the yearly data. This gives:

$$\tau_i(1) = \min\Big\{t > 0 \text{ s.t. } j_{it} > 0 \text{ and } j_{it-1} = 0\Big\}$$

$$\tau_i(2) = \max\Big\{t > 0 \text{ s.t. } j_{it} > 0 \text{ and } j_{it-1} = 0\Big\}.$$

We then define the following sets S:

$$
\begin{aligned}
S^{\mathrm{E}} &= \Big\{(i,t) \text{ s.t. } j_{it}{>}0\Big\} \\
S^{\mathrm{EE}} &= \Big\{(i,t) \text{ s.t. } j_{it}{>}0 \text{ and } j_{it-1}{>}0\Big\} \\
S^{\mathrm{EEE}} &= \Big\{(i,t) \text{ s.t. } j_{it}{>}0,\ j_{it-1}{>}0 \text{ and } j_{it-2}{>}0\Big\} \\
S^{\mathrm{EAE}} &= \Big\{(i,t) \text{ s.t. } j_{it}{>}0 \text{ and } j_{it-2}{>}0\Big\} \\
S^{\mathrm{U2E}} &= \Big\{(i,t) \text{ s.t. } j_{it}{>}0 \text{ and } j_{it-1}{=}0\Big\} \\
S^{\mathrm{J2J}} &= \Big\{(i,t) \text{ s.t. } j_{it}{>}0,\ j_{it-2}{>}0 \text{ and } j_{it}{\neq}j_{it-2}\Big\} \\
S^{\mathrm{S}} &= \Big\{(i,t) \text{ s.t. } j_{it}{>}0 \text{ and } j_{it}{=}j_{it-1}\Big\} \\
S^{\mathrm{SS}} &= \Big\{(i,t) \text{ s.t. } j_{it}{>}0 \text{ and } j_{it}{=}j_{it-1}{=}j_{it-2}\Big\} \\
S^{\mathrm{UEUE}} &= \Big\{(i) \text{ s.t. } \tau_i(1) > 0 \text{ and } \tau_i(2) > 0\Big\}.
\end{aligned}
$$

This directly defines all moments in Table 1, except for the last, for which we construct the retention probability

$$
\tilde{p}_{jt} = \frac{\sum_i \sum_t 1\{j_{it} = j \text{ and } j_{it-1} = j\}}{\sum_i \sum_t 1\{j_{it-1} = j\}}
$$

and compute $\mathrm{Cov}_{S^{\mathrm{EE}}}\Big[\Delta \log(1 - \tilde{p}_{j_{it},t}), \Delta \log w_{it}\Big]$.

Finally, to get standard errors of the moments we employ a bootstrap strategy with 100 replications. For all individual-specific moments, we bootstrap at the individual level, while for moments involving firm quantities such as y_{jt} and $\tilde{p}_{jt}$, we bootstrap at the firm level.

W4 Estimation web appendix

In this appendix we detail the estimation procedure. In order to obtain the parameters via indirect inference, we solve the model at each parameter value, simulate data and finally compute the moments.

W4.1 Numerical solution to the model

We choose $n_z = 7$, $n_{x_0} = 3$ and $n_{x_1} = 5$ points of support for the productivity types, which results in a total of 105 different productivity levels. The promised utility has 200 points of support and is linearly interpolated. For a good starting value in the iterative procedure, we initially solve a simpler model without on-the-job search and the agency problem by iterating over the firm's value, solving the tightness function and updating the worker's problem.

Solving for the optimal contract is a computationally difficult problem, hence we try to keep it tractable. Given that the solution to the search problem is needed many times, we parameterize the $\hat{p}(x, W)$ curve for each x as follows:

$$\hat{p}(x, W) = a(x) + b(x)(W - \overline{W}(x))^{c(x)}.$$

The fit of this function provides an R-square larger than 0.99. The benefit of this parameterization is that the optimal search decision, the probability to receive an offer and the return to search can all be computed in closed form. Similarly, we introduce a second functional approximation for the value to the firm and approximate it using a power decomposition:

$$J(x, z, V) = a(x, z) + \sum_{k=1}^{K} (V - \bar{v}_k(x, z))^{c_k(x,z)}.$$

Setting K=1 provides an R-square above 0.99.

Based on these two functional approximations, we look for a fixed point. To do this, we solve the firm problem in its recursive Lagrangian representation:

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

This requires finding the optimal ω_i at each state, which we obtain from the zeros in the first order condition. We then iterate over updating the firm problem, the tightness function and the worker's unemployment value. During our iterations towards the fixed point, we update the equilibrium condition at a decreasing rate to avoid oscillation around the solution. We stop this procedure when the mean square error (scaled by the total L_2 norm) is below 10^{-8} between two consecutive iterations for all value functions.

W4.2 Simulating moments

The challenge when computing moments is to simulate firms as bundles of workers, each sharing a history of shocks. We draw a sequence of ν_t shocks and construct the corresponding paths of match quality z_t. The ν_t sequence can be thought of as a circle on which workers evolve as part of a firm, and so represents an infinite sequence of shocks. Workers who move to a new job start at a randomly chosen new point on the circle and are assigned a $z_t = z_0$. They then follow the predetermined sequence of z from that point forward. All workers at a given point on the circle are coworkers.

In practice we use a circle of length 200 and simulate $20,000$ workers with random starting points. Discarding a burn-in period, we finally focus on the last 30 periods of data. When computing the simulated moments, we repeat the simulation 20 times. For each simulation we redraw everything, including the ν_t and z_t sequences, and take averages over the replications.

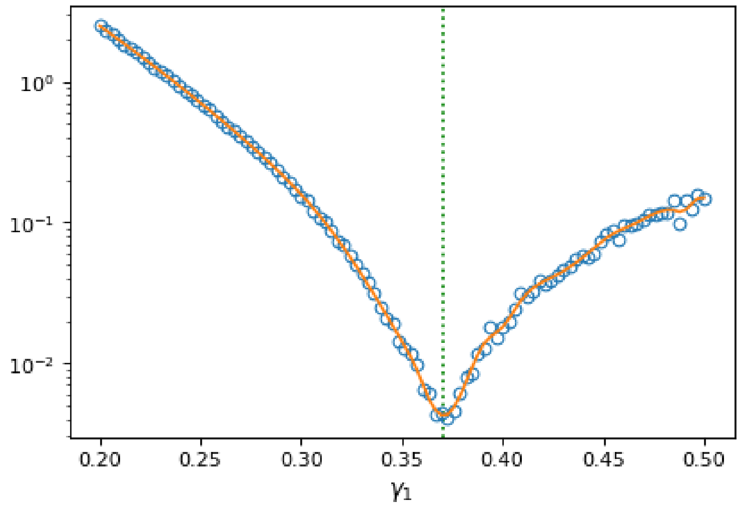
W4.3 Optimization

Our objective function is given by

$$\mathcal{O}(\theta) = (\hat{M} - M(\theta))'\mathcal{W}(\hat{M} - M(\theta)),$$

where $\hat{M}$ is the vector of moments from the data, $\mathcal{W}$ is a diagonal matrix of weights and $M(\theta)$ is the vector of moments simulated from the model. We weight all moments in the model by the inverse of their value in the data, with the exception that we scale the auto-covariances by their variances because the

Figure W1: Surrogate line search



Notes: This plots an example of our surrogate line search in the direction of the parameter γ_1 associated with effort cost. The blue dots are individual evaluations, the orange solid line is the fitted spline, the dashed line shows the previous value and the red vertical line is the updated number. Close to the end of the optimization the update is very small and minimizes the objective.

auto-covariances are often close to zero.

Our optimization procedure is a custom surrogate line search, i.e. we choose a direction in the parameter space and evaluate 100 points in that direction. We then fit a smoothing spline, picking the smoothing parameter to minimize the leave-one-out mean square prediction error. We finally pick our new parameter as the minimum of that smoothing spline. See Figure W1 for an example of such an approach.

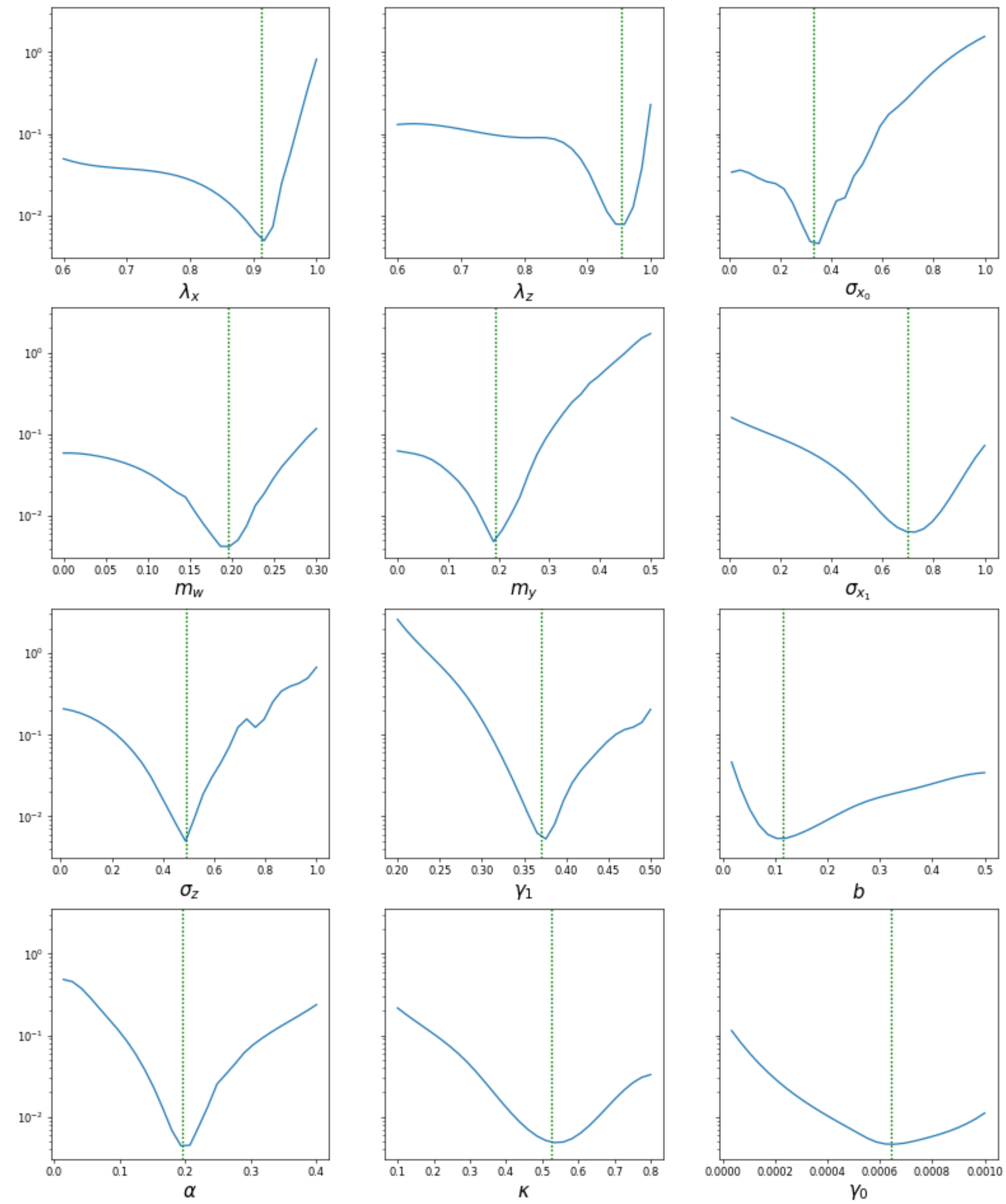
W4.4 Computing standard errors

To derive the standard errors we ignore simulation noise and employ the conventional sandwich formula using the diagonal variance covariance matrix for the moments $\hat{\Sigma}$:

$$\frac{1}{n_f}\left(\frac{\partial M(\theta)'}{\partial\theta}\mathcal{W}\frac{\partial M(\theta)}{\partial\theta}\right)^{-1}\frac{\partial M(\theta)'}{\partial\theta}\mathcal{W}\hat{\Sigma}\mathcal{W}\frac{\partial M(\theta)}{\partial\theta}\left(\frac{\partial M(\theta)'}{\partial\theta}\mathcal{W}\frac{\partial M(\theta)}{\partial\theta}\right)^{-1},$$

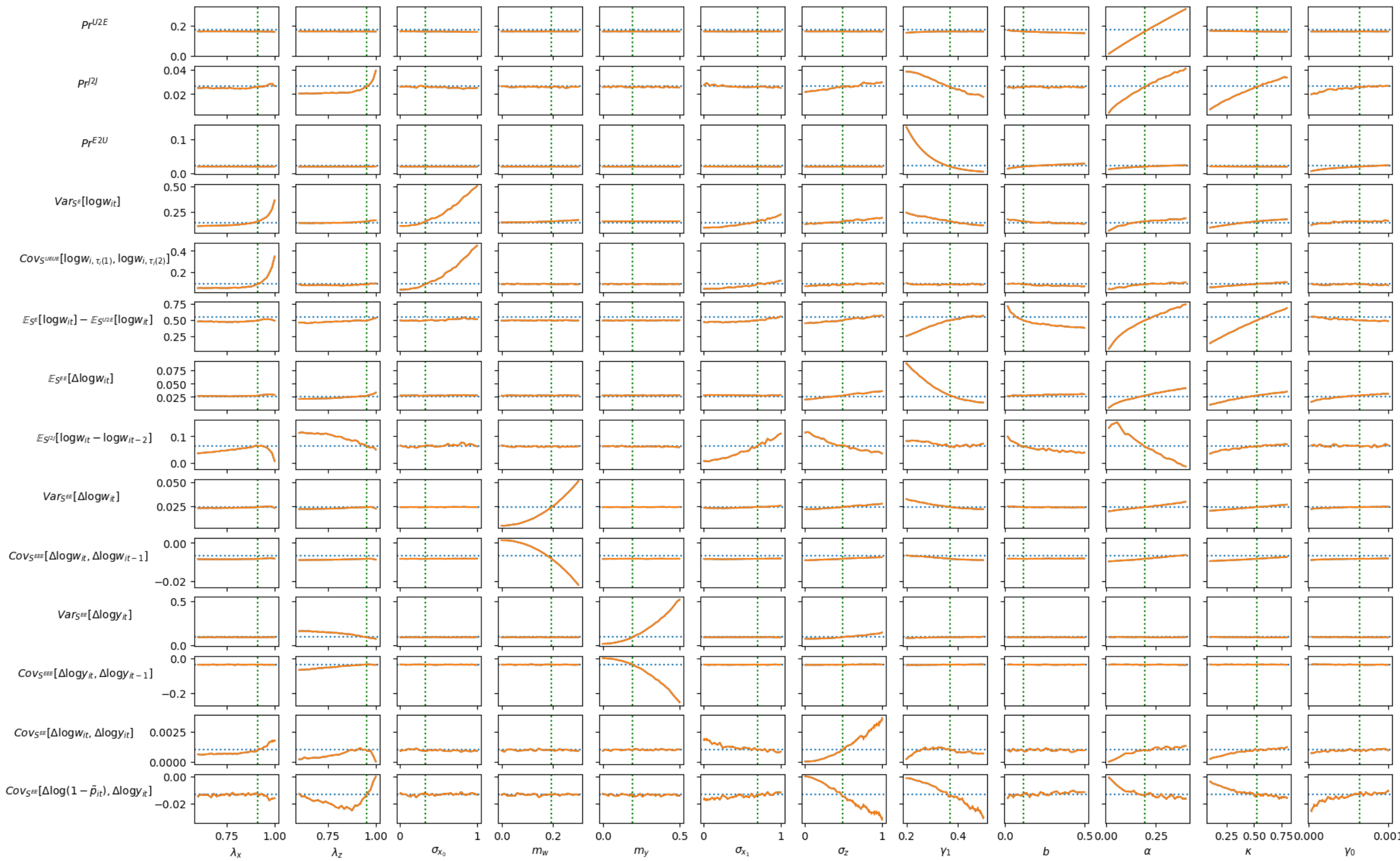
where $\hat{\Sigma}$ is estimated by bootstrap and where we set all off-diagonal terms to zero. We scale by the number of firms n_f because some moments are computed at the firm level and we consider an asymptotic with many firms.

Figure W2: Slices of the objective function



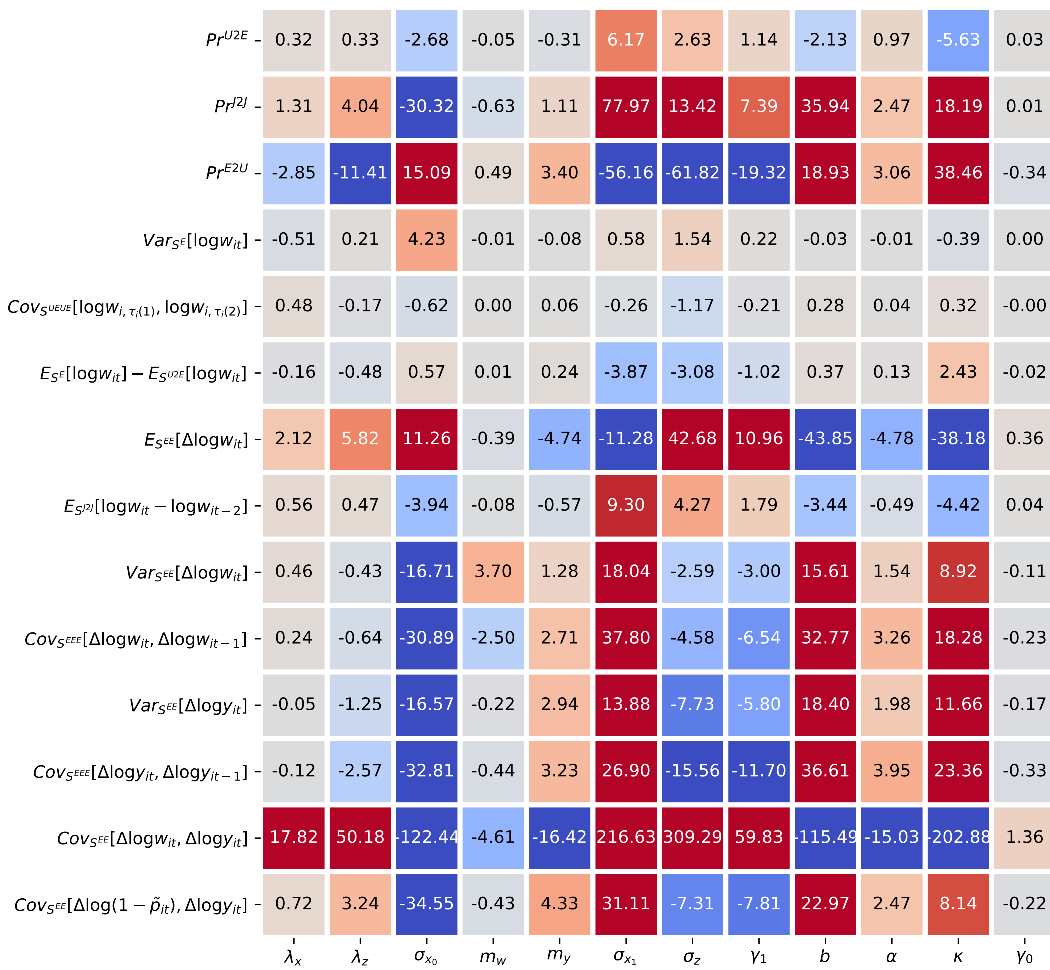
Notes: This plots the objective function against each parameter, away from the optimal parameter value. The y-axis is log-scaled.

Figure W3: All parameters and moments



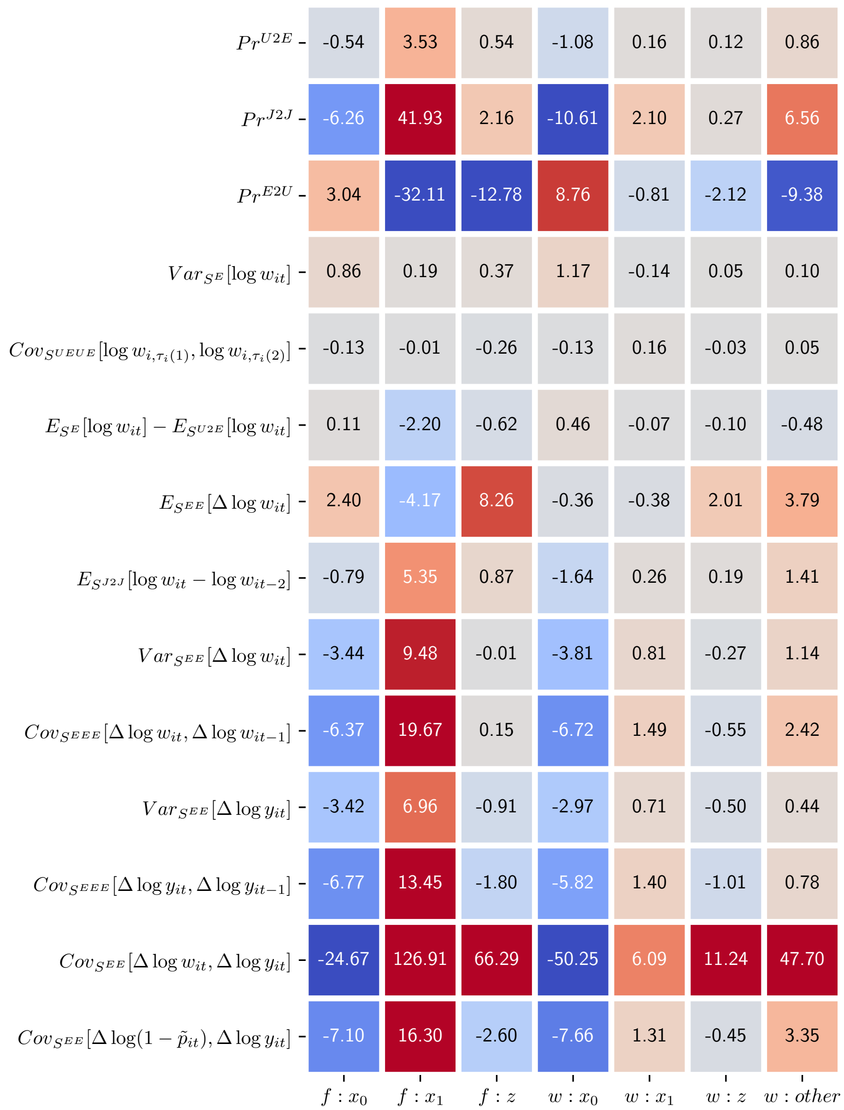
Notes: This plots each of the moments against each parameter, away from the optimal parameter value.

Figure W4: Sensitivity measure



Notes: Measure of sensitivity from Andrews, Gentzkow, and Shapiro (2017).

Figure W5: Sensitivity measure for level variance decompositions



Notes: Measure of sensitivity from Andrews, Gentzkow, and Shapiro (2017) for the counterfactual decomposition from Table 1. This reflects how changing one of the data moments by one unit would affect the counterfactual variance attributed to each of x_0, x_1, z in match output and in wages.

References

Abdou, M., A. Soliman, and M. Abdel-Aty (2020): "On a discussion of Volterra-Fredholm integral equation with discontinuous kernel," Journal of the Egyptian Mathematical Society, 28(1), 1–10.

Allman, E. S., C. Matias, and J. A. Rhodes (2009): "Identifiability of parameters in latent structure models with many observed variables," Ann. Stat., 37(6A), 3099–3132.

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.

Benveniste, L. M., and J. A. Scheinkman (1979): "On the Differentiability of the Value Function in Dynamic Models of Economics," Econometrica, 47(3), 727–732.

Carlsson, M., I. Häkkinen Skans, and O. Nordstrom Skans (2019): "Wage flexibility in a unionized economy with stable wage dispersion," Discussion paper, IZA.

Evans, G. C. (1911): "Volterra's integral equation of the second kind, with discontinuous kernel. II," Transactions of the American Mathematical Society, 12(4), 429–472.

Fredriksson, P., and R. H. Topel (2010): "Wage Determination and Employment in Sweden Since the Early 1990s: Wage Formation in a New Setting," in Reforming the Welfare State: Recovery and Beyond in Sweden, pp. 83–126. University of Chicago Press.

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

Hall, P., and X.-H. Zhou (2003): "Nonparametric Estimation of Component Distributions in a Multivariate Mixture," Ann. Stat., 31(1), 201–224.

Hotz, V. J., and R. A. Miller (1993): "Conditional choice probabilities and the estimation of dynamic models," The Review of Economic Studies, 60(3), 497–529.

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

Koeppl, T. (2006): "Differentiability of the efficient frontier when commitment to risk sharing is limited," Topics in Macroeconomics, 6(1), 1–6.

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.

Stokey, N. L., R. Lucas, and E. Prescott (1989): Recursive methods in dynamic economics. Harvard University Press.

Sundaram, R. K., et al. (1996): A first course in optimization theory. Cambridge University Press.

Swedish Mediation Office Annual Report (2002): "2002 Annual report," Discussion paper.

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