A1 Model appendix

A1.1 Existence of equilibrium

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

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

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

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

A1.2 Characterization of the optimal contract

Proof of Lemma 1.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Because of independence between x and z, it is inconsequential for $x(t; \omega)$ whether the firm starts in (x, z_1) or (x, z_2). However, $z_2(t; \omega) \geq z_1(t; \omega)$ because the transition function $g(\cdot,\cdot)$ is assumed to be monotonic and $z_2 > z_1$.

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

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

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

$$\begin{aligned} J(x, z_2, V) \geq J_2 &= \int \sum_{t=1}^{\infty} \beta^{t-1} \Big(f(x(t;\omega), z_2(t;\omega)) - w_1(t;\omega) \Big) \Lambda_1(t;\omega) \, \mathrm{d}\omega \\ &\geq \int \sum_{t=1}^{\infty} \beta^{t-1} \Big(f(x(t;\omega), z_1(t;\omega)) - w_1(t;\omega) \Big) \Lambda_1(t;\omega) \, \mathrm{d}\omega \\ &= J(x, z_1, V), \end{aligned}$$

which gives the result. See Dardanoni (1995) for more details on properties of monotonic Markov chains. □

Recall the expected profit equation:

$$\begin{aligned} J(x, z, V) = \max_{\pi_i, w_i, W_i, W_{ix'z'}} & \sum_{i=1,2} \pi_i \Big(f(x, z) - w_i + \beta \tilde{p}(x, W_i) \mathbb{E}_{x'z'}[J(x', z', W_{ix'z'}) | x, z] \Big) \\ s.t \quad (\lambda) \quad & V \leq \sum_{i=1,2} \pi_i \left(u(w_i) + \tilde{r}(x, W_i) \right) \\ (\omega_i) \quad & W_i = \mathbb{E}_{x'z'}[W_{ix'z'} | x, z], \\ & \sum_{i=1,2} \pi_i = 1, \end{aligned}$$

where λ and ω_i denote Lagrange multipliers on constraints.

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

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

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

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

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

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

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

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

Next, the first order condition with respect to W_i is

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

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

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

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

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

resulting in the following expression:

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

Focusing on $\tilde{p}(x, W_i) > 0$ and $\pi_i > 0$ (because otherwise the worker is leaving the current firm and the next period wage is irrelevant) we now rewrite:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

or in reordered form:

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

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

Overshooting w^*. The next period wage does not overshoot the target wage w^*, which we show by contradiction. Suppose the wage in lottery realization 1 w'_1 overshoots w^* from below, i.e. lottery outcome 1 yields a wage increase and next period's wage is higher than the target wage, and assume that in lottery realization 2 the wage increases exactly up to the target wage: $w'_1 > w'_2 = w^*$. The associated expected firm profits are then such that they are positive in the first realization and exactly zero in the other:

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

Using $w'_1 > w'_2$ and concavity of the utility function implies:

$$\begin{aligned} \frac{1}{u'(w'_1)} &> \frac{1}{u'(w'_2)} \\ -J'(x', z', W_{1x'z'}) &> -J'(x', z', W_{2x'z'}) \\ W_{1x'z'} &> W_{2x'z'} \\ \mathbb{E}_{x'z'}[J(x', z', W_{1x'z'})|x, z] &< \mathbb{E}_{x'z'}[J(x', z', W_{2x'z'})|x, z] = 0, \end{aligned}$$

where the transformations use Lemma A1 and the fact that J is a decreasing and concave function. The result contradicts our initial supposition and so wages do not overshoot the efficiency wage. Further, it is impossible that both w'_1 and w'_2 overshoot w^* from below because this would imply negative expected profits and contradicts that wages grow, so overshooting in both lotteries from below is also ruled out. Similarly, one can make a similar argument for overshooting from above.

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

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

A1.3 Solving the model

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

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

where

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

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

Proof. We have the following recursive formulation for J:

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

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

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

Consider the Pareto problem

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

and combine the terms to get:

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

The final step is to split the sup

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

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

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

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

$$\mathcal{P}_{01}(x,z,\rho) = \sup_{\pi_i,w_i,W_2>\underline{W}(x),W_{ix'z'}} \pi_1\Big(f(x,z) - w_1 + \beta\tilde{p}(x,\underline{W}(x))\mathbb{E}_{x'z'}[J(x',z',W_{1x'z'})|x,z]$$
$$+ \rho\Big(u(w_1) + \tilde{r}(x,\underline{W}(x))\Big)\Big)$$
$$+ \pi_2\Big(f(x,z) - w_2 + \beta\tilde{p}(x,W_2)\mathbb{E}_{x'z'}[J(x',z',W_{2x'z'})|x,z]$$
$$+ \rho\Big(u(w_2) + \tilde{r}(x,W_2)\Big)\Big)$$
$$s.t \quad (\omega_i) \quad W_i = \mathbb{E}_{x'z'}[W_{ix'z'}|x,z],$$

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

$$\mathcal{P}_{01}(x,z,\rho) = \sup_{\pi_i,w_i,W_2>\underline{W}(x),W_{ix'z'}} \pi_1\Big(f(x,z) - w_1 + \rho\Big(u(w_1) + \beta\mathbb{E}_{x'}[U(x')|x]\Big)\Big)$$
$$+ \pi_2\Big(f(x,z) - w_2 + \beta\tilde{p}(x,W_2)\mathbb{E}_{x'z'}[J(x',z',W_{2x'z'})|x,z]$$
$$+ \rho\Big(u(w_2) + \tilde{r}(x,W_2)\Big)\Big)$$
$$s.t \quad (\omega_i) \quad W_i = \mathbb{E}_{x'z'}[W_{ix'z'}|x,z],$$

where the choice variables W_1 and $W_{1x'z'}$ disappear and so does the constraint associated with ω_1. We apply the same treatment as in the case of $\mathcal{P}_{11}$ to get:

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

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

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

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

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

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

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

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

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

□

A2 Data appendix

A2.1 Data sources

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

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

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

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

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

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

yearly time dummies.

Table A1: Construction of value added: FORBRUKNINGSVARDE

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

A2.2 Sample construction

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

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

Table A2: Data description

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

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

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

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

A3 Counterfactuals appendix

A3.1 First best

We start with the original firm problem:

$$\begin{aligned}
J(x,z,V) = \max_{\pi_i, w_i, W_i, W_{ix'z'}} \sum_{i=1,2} \pi_i \Big(f(x,z) - w_i + \beta \tilde{p}(x, W_i) \mathbb{E}_{x'z'}[J(x', z', W_{ix'z'}) | x, z] \Big) \\
s.t \quad (\lambda) \quad V = \sum_{i=1,2} \pi_i \Big(u(w_i) + \tilde{r}(x, W_i) \Big), \\
(\omega_i) \quad W_i = \mathbb{E}_{x'z'}[W_{ix'z'} | x, z].
\end{aligned}$$

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

$$\begin{aligned}
J^{fb}(x,z,V) = \max_{\pi_i, w_i, W_i, W_{ix'z'}, e_i, v_i} \sum_{i=1,2} \pi_i \Big(f(x,z) - w_i \\
+ \beta (1 - \delta(e_i))(1 - \kappa p(\theta(x, v_i))) \mathbb{E}_{x'z'}[J^{fb}(x', z', W_{ix'z'}) | x, z] \Big) \\
s.t \quad (\lambda) \quad V = \sum_{i=1,2} \pi_i \Big(u(w_i) - c(e_i) + \delta(e_i) \beta \mathbb{E}_{x'}[U(x') | x] \\
+ \beta \kappa (1 - \delta(e_i)) p(\theta(x, v_i))(v_i - W_i) + \beta (1 - \delta(e_i)) W_i \Big), \\
(\omega_i) \quad W_i = \mathbb{E}_{x'z'}[W_{ix'z'} | x, z].
\end{aligned}$$

We define the first best Pareto problem:

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

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

$$
\begin{aligned}
\mathcal{P}^{fb}(x,z,\rho) = \inf_{\omega_i} \sup_{\pi_i, w_i, W_i \geq \underline{W}(x), v_i, e_i} \sum_{i=1,2} \pi_i \Big(& f(x,z) - w_i \\
& + \rho \Big(u(w_i) - c(e_i) + \delta(e_i)\beta \mathbb{E}_{x'}[U(x')|x] \\
& + \beta\kappa(1-\delta(e_i))p(\theta(x,v_i))(v_i - W_i) + \beta(1-\delta(e_i))W_i \Big) \\
& - \beta\omega_i(1-\delta(e_i))(1-\kappa p(\theta(x,v_i)))W_i \\
& + \beta(1-\delta(e_i))(1-\kappa p(\theta(x,v_i)))\mathbb{E}_{x'z'}[\mathcal{P}^{fb}(x',z',\omega_i)|x,z] \Big).
\end{aligned}
$$

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

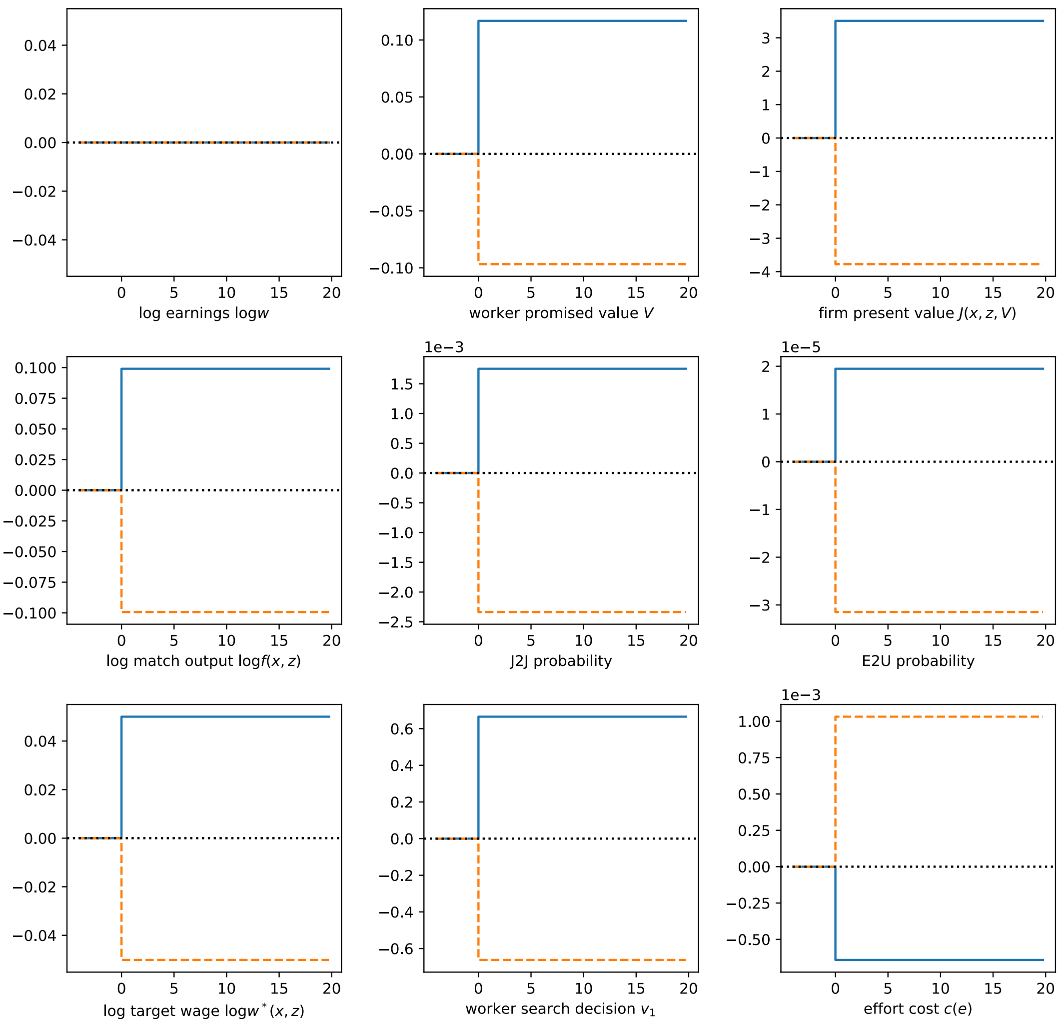
$$
\begin{aligned}
\mathcal{P}^{fb}(x,z,\rho) = \sup_{\pi_i, w_i, v_i, e_i} \sum_{i=1,2} \pi_i \Big(& f(x,z) - w_i \\
& + \rho \Big(u(w_i) - c(e_i) + \delta(e_i)\beta \mathbb{E}_{x'}[U(x')|x] + \beta\kappa(1-\delta(e_i))p(\theta(x,v_i))v_i \Big) \\
& + \beta(1-\delta(e_i))(1-\kappa p(\theta(x,v_i)))\mathbb{E}_{x'z'}[\mathcal{P}^{fb}(x',z',\rho)|x,z] \Big).
\end{aligned}
$$

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

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

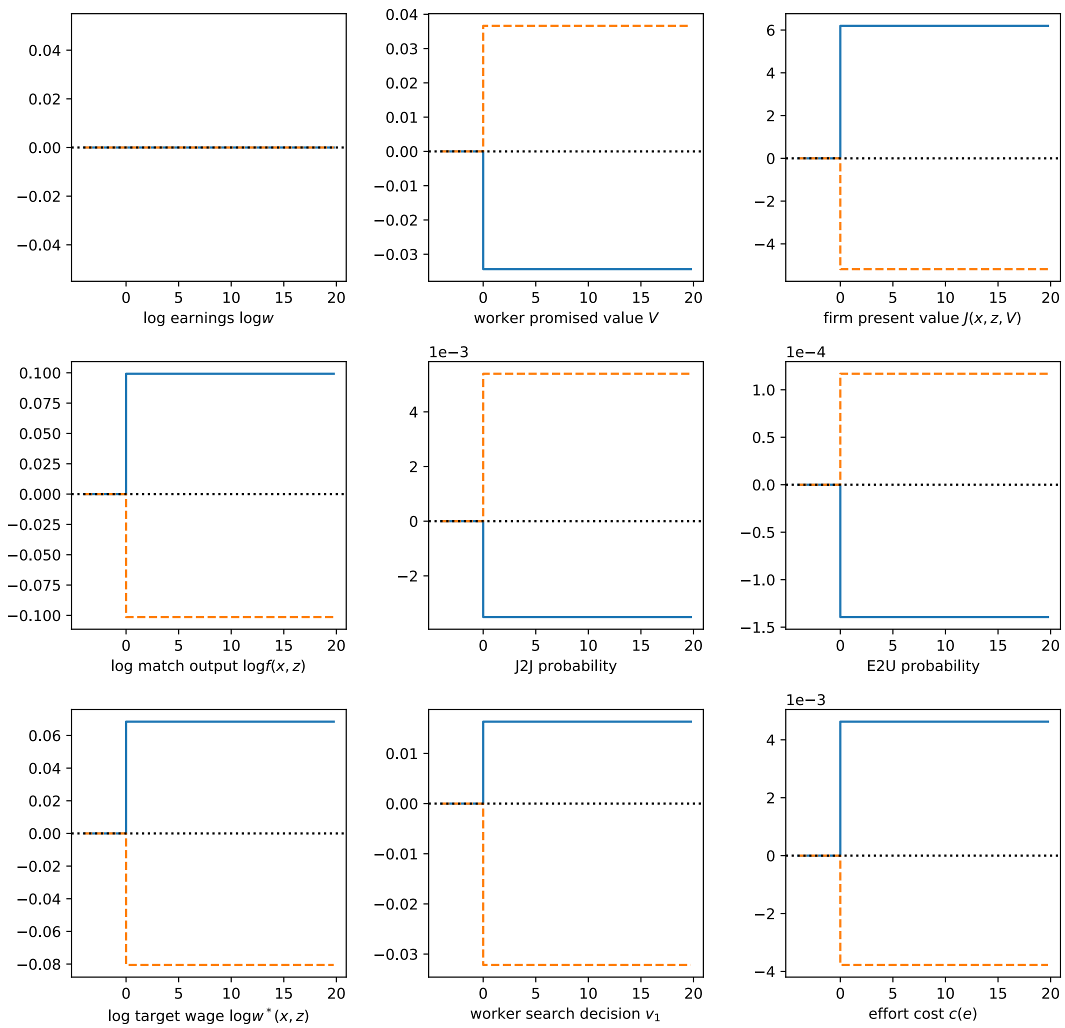
where $\rho^*(x,z,V)$ is defined as the ρ that equates the partial derivative of $\mathcal{P}^{fb}$ with the promised value V.

Figure A1: Average impulse response to x change at first best



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

Figure A2: Average impulse response to z change at first best



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

A3.2 Passthrough analysis

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

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

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

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

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

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

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

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

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

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

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

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

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

It is given by:

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

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

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