

SUPPLEMENT TO “FIRM AND WORKER DYNAMICS IN A FRICTIONAL
LABOR MARKET”

(*Econometrica*, Vol. 90, No. 4, July 2022, 1425–1462)

ADRIEN BILAL

Department of Economics, Harvard University and NBER

NIKLAS ENGBOM

Stern School of Business, New York University, CEPR, and NBER

SIMON MONGEY

Federal Reserve Bank of Minneapolis, Department of Economics, University of Chicago, and NBER

GIOVANNI L. VIOLANTE

Department of Economics, Princeton University, CEBI, CEPR, IFS, IZA, and NBER

THIS APPENDIX IS ORGANIZED AS FOLLOWS. Section **A** provides intuition for how our assumptions (**A**) yield a tractable Bellman equation for joint value. Section **B** provides a characterization of the surplus function. Section **C** derives an alternative limit of the Bellman equation for the joint value as decreasing returns vanish. Section **D** derives the limiting behavior of our economy when frictions vanish. Section **E** illustrates identification of the model.

APPENDIX A: STATIC EXAMPLE

Set up. Consider a firm with decreasing returns to scale technology $y(z, n)$ such that $y(z, 0) = 0$. Suppose the firm starts with productivity z and $n = 1$ worker. The current contract between the firm and the incumbent specifies a wage $w_1 \in (b, y(z, 1))$, where $b = U$ is the value of unemployment. At this point, the incumbent worker does not have a credible threat to quit into unemployment nor the firm has a credible threat to fire the worker. Then, the labor market opens. For now, we also assume that the firm has sunk the cost of a vacancy c . Later we explicitly consider the decision to post a vacancy.

A.1. *UE Hire*

We describe how to obtain the “UE hire” term in (1). Assume the firm’s vacancy meets an unemployed worker. Four different cases can arise from the combination of hiring/not hiring and renegotiating/not renegotiating the wage with the incumbent. Our assumption on external negotiation (**A-EN**) requires that in all cases the take–leave wage offer of the firm to the outside worker is $w_2 = b$. Our internal negotiation assumption (**A-IN**) requires that the joint value with and without renegotiation is the same and simply equals output $y(z, n)$. Let w_1^* be the incumbent wage after the internal negotiation.

Adrien Bilal: adrienbilal@fas.harvard.edu

Niklas Engbom: ne466@nyu.edu

Simon Mongey: mongey@uchicago.edu

Giovanni L. Violante: violante@princeton.edu

If the firm hires the new worker, its profits are as follows:

$$\underbrace{y(z, 2) - w_1 - b}_{\text{Without renegotiation}}, \quad \underbrace{y(z, 2) - w_1^* - b}_{\text{With renegotiation}}$$

If the firm does not hire the new worker, its profits are

$$\underbrace{y(z, 1) - w_1}_{\text{Without renegotiation}}, \quad \underbrace{y(z, 1) - w_1^*}_{\text{With renegotiation}}$$

We now describe which case occurs. This requires understanding when our mutual consent assumption (**A-MC**) coupled with limited commitment on layoffs (**A-LC**) bind. In particular, the firm may obtain a credible threat to trigger renegotiation of w_1 . We focus first on when a hire occurs.

Hire. A hire *without* renegotiation occurs when the following two conditions hold:

$$\underbrace{y(z, 2) - w_1 - b \geq y(z, 1) - b}_{\text{No credible threat}}, \quad \underbrace{y(z, 2) - w_1 - b \geq y(z, 1) - w_1}_{\text{Optimal to hire w/o renegotiation}}. \quad (12)$$

The first condition illustrates that the threat to fire the incumbent worker is not credible, which under (**A-MC**) implies no renegotiation. Keeping the incumbent worker at w_1 and employing the outside worker at b delivers a higher value to the firm than the threat of “swapping”: firing worker one and hiring the unemployed worker in his place. Given no renegotiation, the second condition ensures hiring is privately optimal for the firm.

A hire *with* renegotiation occurs when the following two conditions hold:

$$\underbrace{y(z, 2) - w_1 - b < y(z, 1) - b}_{\text{Credible threat}}, \quad \underbrace{y(z, 2) - w_1^* - b > y(z, 1) - w_1^*}_{\text{Optimal to hire w/ renegotiation}}. \quad (13)$$

The firm has now a credible threat to fire the incumbent worker according to (**A-LC**). This is possible only under decreasing returns to scale: even though $w_1 < y(z, 1)$, the first inequality in (13) implies $w_1 > y(z, 2) - y(z, 1)$, that is, the incumbent wage is above its own marginal product. Employing the outside worker at b and keeping the incumbent worker at w_1 delivers a lower value than “firing and swapping.” The second condition is necessary for hiring to be optimal under the renegotiated wage w_1^* to the incumbent worker.

Under the zero sum game assumption (**A-IN**), the renegotiated wage w_1^* only redistributes value between the incumbent worker and the firm, but does not affect total value.⁴³ In addition, it must be individually rational, and so $w_1^* \in [b, y(z, 2) - y(z, 1)]$. Without further assumptions, we cannot say where exactly the new wage lies within this interval, but we can nonetheless pin down allocations.

Rearranging the optimal hiring conditions, we observe that both are satisfied as long as

$$y(z, 2) - y(z, 1) > b. \quad (14)$$

⁴³Two relevant cases that would violate this condition are (i) if worker’s effort depends on the wage and enters the production function, and (ii) concave utility.

Note that without internal renegotiation (**A-IN**), the hiring condition would differ in the two cases. If wages could not be cut and the firm had a credible threat, the incumbent worker would be fired and the firm would always hire the unemployed worker ($y(z, 1) > b$). As a result, to determine firm size next period, one would need to know the incumbent's wage to distinguish between the two cases (thus, in the general model with n workers, the whole wage distribution). Similarly, if a fraction of output were to be lost because of the internal negotiation, a violation of (**A-IN**), the hiring conditions in (12) and (13) would differ, and again one would need to know wages to determine whether a hire occurs.

We can write inequality (14) in terms of joint value. Workers' values are simply equal to their wage w_i for $i \in \{1, 2\}$. The firm's value is simply equal to its profits. The fact that wages are valued linearly by both worker and firm implies that the joint value $\Omega(z, n)$ is independent of wages:

$$\Omega(z, n) = \underbrace{y(z, n) - \sum_{i=1}^n w_i}_{\text{Firm value}} + \underbrace{\sum_{i=1}^n w_i}_{\text{Sum of workers' values}}, \quad \text{for any } (w_i)_{i=1}^n.$$

Using the definition of joint value, equation (14) characterizes when the UE hire occurs:

$$\Omega(z, 2) - \Omega(z, 1) > U. \quad (15)$$

Thus, the decision of hiring from unemployment does not depend on wages, but only on productivity, size, and the value of unemployment $U = b$.

No Hire. Consider now the cases where no hiring occurs. Recall that, once an unemployed worker is met, the firm has always a credible threat against the incumbent since $w_1 > b$. No hire with renegotiation therefore occurs when the following two conditions hold:

$$\underbrace{y(z, 1) - b > y(z, 1) - w_1}_{\text{Credible threat}}, \quad \underbrace{y(z, 1) - w_1^* \geq y(z, 2) - w_1^* - b}_{\text{Optimal to not hire}}. \quad (16)$$

After renegotiation, the incumbent wage is driven down to w_1^* . Since this outcome represents a redistribution of value between firm and worker then, consistent with (**A-IN**), the joint value remains $\Omega(z, 1)$.⁴⁴ Finally, the no-hiring condition in (16) can be rewritten as in (14) with the opposite inequality, $\Omega(z, 2) - \Omega(z, 1) \leq U$.

Combined. The firm hires from unemployment when the *marginal value* of the job-seeker exceeds the value of unemployment:

$$\frac{\Omega(z, 2) - \Omega(z, 1)}{2 - 1} > U. \quad (17)$$

In addition, the joint value of the firm and its workers rises by $\frac{\Omega(z, 2) - \Omega(z, 1)}{2 - 1} - U$ when the hire occurs. This is exactly the *UE hire* term in the HJB equation (1). In the case of a hire, incumbent wages may or may not be renegotiated but this has no impact on whether

⁴⁴The value before renegotiation was $\Omega(z, 1) = z - w_1 + w_1 = z$. The joint value after renegotiation is $\Omega(z, 1) = z - w_1^* + w_1^* = z$.

hiring occurs, or how the joint value changes. When this condition fails, the firm does not hire, wages are renegotiated, but the joint value remains constant. All decisions require knowledge of (z, n) only, but not of incumbents' wages.

We now generalize the *UE* hire case analyzed in the main text to a firm with multiple incumbents.

A.2. *UE Hire When the Internal Renegotiation Involves Multiple Workers*

It is sufficient to consider the case of two incumbent workers, $n = 2$. Without loss of generality, assume that the second worker is paid more than the first, $w_2 > w_1$. As in the approach taken earlier, suppose the firm has posted a vacancy that has met an unemployed worker. We have three cases to consider, which illustrate how the firm may use a worker outside the firm to sequentially renegotiate wages of workers inside the firm.

First, the firm hires *without* renegotiation if:

$$\underbrace{y(z, 3) - w_1 - w_2 - b > y(z, 2) - w_1 - b,}_{\text{No credible threat to } w_2}$$

$$\underbrace{y(z, 3) - w_1 - w_2 - b > y(z, 2) - w_1 - w_2.}_{\text{Optimal to hire under } (w_1, w_2)}$$

Hiring with current wages is preferred to replacing the most expensive incumbent—there is no credible threat—and given no renegotiation, hiring is optimal. Since $w_2 > w_1$, no credible threat to worker 2 implies no credible threat to worker 1.

Second, the firm hires *with* renegotiation with worker 2 if:

$$\underbrace{y(z, 2) - w_1 - b > y(z, 3) - w_1 - w_2 - b > y(z, 2) - w_2 - b,}_{\text{Credible threat for worker 2 only}}$$

$$\underbrace{y(z, 3) - w_1 - w_2^* - b > y(z, 2) - w_1 - w_2^*}_{\text{Optimal to hire under } (w_1, w_2^*)}$$

The threat is credible for worker 2, but is not for worker 1, and, conditional on renegotiating to (w_1, w_2^*) , hiring is optimal.

Third, the firm hires *with* renegotiation with *both* workers if:

$$\underbrace{y(z, 2) - w_1 - b > y(z, 2) - w_2 - b > y(z, 3) - w_1 - w_2 - b,}_{\text{Credible threat for both workers}}$$

$$\underbrace{y(z, 3) - w_1^* - w_2^* - b > y(z, 2) - w_1^* - w_2^*}_{\text{Optimal to hire under } (w_1^*, w_2^*)}$$

In all three cases, the optimal hiring condition can be written in terms of joint value as

$$\frac{\Omega(z, 3) - \Omega(z, 2)}{3 - 2} > U. \quad (18)$$

This last inequality does not depend on the order of the internal negotiation between firm and workers. In conclusion, the distribution of wages among incumbents again determines the patterns of wage renegotiation, but is immaterial for the sufficient condition for hiring.

Hiring occurs whenever the marginal value of adding a worker to the coalition exceeds the value of unemployment.

Assumption **(A-LC-c)** that was not present in the one worker example plays a role here. Suppose that the renegotiated wage for worker 2 is pushed all the way down to b , making her indifferent between staying and quitting. Worker 1 could transfer a negligible amount to worker 2 in exchange of her quitting, which would raise the firm's marginal product and, possibly, remove its own threat. This is problematic for the representation because in this latter case the hiring condition becomes $y(z, 2) - w_1 - b > y(z, 1) - w_1$, distinct from (18). Thus, to know whether a firm hires or not, one would need to know the wage distribution inside the firm. **(A-LC-c)** is sufficient to rule out transfers among workers and to prevent this scenario from happening.

Note that this transfer scheme between workers occurring during the internal negotiation changes the joint value, and hence one can think of **(A-LC-c)** as being subsumed into **(A-IN)** already.

In what follows, we return to the case where the firm has only one worker.

A.3. *EE Hire*

Now suppose that the worker matched with the firm's vacancy is currently employed at another firm with productivity z' and a single worker $n' = 1$. The situation is not that different from *UE hire*, except that the potential hire may have a better outside option in the form of the retention offer made to her by her current employer under **(A-EN)**. To see the similarity for now, we fix this wage offer at \bar{w} . The same four cases described in Section A.1 can arise, except with \bar{w} playing the role of b .⁴⁵ We can therefore reason as before and jump to the result that hiring will occur if and only if the following counterpart to (15) holds:

$$\Omega(z, 2) - \Omega(z, 1) > \bar{w}.$$

We now determine the poached worker's outside option \bar{w} . The poached firm's willingness to pay is a wage \tilde{w} that makes it indifferent between retaining and releasing the worker: $y(z', 1) - \tilde{w} = 0$. Hence the contacted worker switches to the new employer as long as the poaching firm offers $\bar{w} \geq \tilde{w} = y(z', 1)$. Bertrand competition between the two firms implies that the poaching firm offers $\bar{w} = y(z', 1)$, which is exactly the marginal value of the worker at the poached firm. As in the case of *UE hire*, whether *EE hire* occurs can be summarized by joint values:

$$\frac{\Omega(z, 2) - \Omega(z, 1)}{2 - 1} > \frac{\Omega(z', 1) - \Omega(z', 0)}{1 - 0}. \quad (19)$$

The *EE hire* decision is entirely characterized by knowledge of the pair (z, n) for the two firms.⁴⁶ The value gain to the firm and its workers is the difference between the left-hand side and right-hand side of equation (19). This comparison of marginal values is precisely the *EE hire* term in the HJB equation (1).

⁴⁵Renegotiation will happen for different values of w_1 in the no hire case. Indeed, to establish the presence of a credible threat w_1 must be compared to \bar{w} instead of b , but this has no allocative implications for the hiring decisions.

⁴⁶The case when the firm meets a worker at a firm with $(z', 2, w_1, w_2)$ is similar. Suppose the firm meets worker 1. The poached firm has the additional option of cutting w_2 , but this is inconsequential for the argument because it only redistributes value within the poached-from firm.

Finally, this exercise explains the absence of a *EE quit* term in (1). The payment received by its poached worker is equal to the poached coalition’s willingness to pay, which is in turn exactly equal to the worker’s marginal value to the coalition. The joint value of the poached coalition therefore does not change as it loses its worker. *EE quit* events play an important role in the dynamics of employment at the firm, but no role in the dynamics of $\Omega(z, n)$.

A.4. Vacancy Posting

Up to this point, we assumed that a meeting between a hiring firm and a job-seeker had already occurred. We now turn to the vacancy posting decision and explain why **(A-VP)** is crucial for tractability.

Recall that in the hiring scenarios just analyzed, two cases arise when the firm can credibly force a wage cut: (i) when it hires and the incumbent wage is above the post-hire new marginal product; (ii) when hiring is not profitable, but the firm can credibly “fire and swap,” that is, as long as the reservation wage of the external worker met through search is below the incumbent wage. The firm has therefore incentives to spend resources on vacancy posting for the sole purpose of transferring value between agents, a privately inefficient outcome. The amount spent would depend on the incumbent’s wage, breaking the tractability of our representation. Private efficiency reinstates tractability.

We start with the firm’s preferred vacancy policy. Without loss of generality, suppose firms only meet unemployed workers (hence, upon a meeting, the “fire and swap” threat is always credible). The generalization to the case where the worker contacted can be either unemployed or employed is straightforward. Let v be the number of vacancies posted, $c(v)$ the associated cost, and qv the probability a single vacancy meets a single worker. If no meeting occurs, then as per **(A-MC)**, w_1 does not change so the value of the firm does not change. The firm maximizing the expected return from vacancy posting net of costs is

$$\max_v -c(v) + qv \left[\underbrace{\max\{y(z, 2) - w'_1 - b, y(z, 1) - b\}}_{\text{Hire (cases 1\&2)}} - (y(z, 1) - w_1) \right].$$

Following a meeting, three cases may occur. In **Case 1**, the firm hires and there is no renegotiation, $w'_1 = w_1$. This case arises when the wage of the incumbent worker is low enough. Then, adding a second worker does not reduce the marginal product of labor down to the point where the firm has a credible layoff threat. In **Case 2**, the firm hires but the wage of the incumbent is renegotiated down to $w'_1 = w_1^*$. In this case, diminishing marginal returns drive the marginal product of labor with two workers below the incumbent’s initial wage. In **Case 3**, the firm is better off not hiring, but under the threat of swapping out the incumbent, renegotiates w_1 down to b . The firm’s preferred vacancy policy v^f then equates marginal cost to marginal expected return:

$$c_v(v^f) = q[\max\{y(z, 2) - w'_1 - b, y(z, 1) - b\} - (y(z, 1) - w_1)]. \quad (20)$$

The first-order condition (20) highlights that the firm’s preferred vacancy policy depends on the incumbent’s wage w_1 because this wage determines the gains from forcing a renegotiation through vacancy posting. This dependence is a source of intractability because, in the general model with n workers, (20) would depend on the entire wage distribution inside the firm.

Our assumption **(A-VP)** ensures that firms do not post v^f , but instead post the privately efficient amount of vacancies, which does not depend on worker wages. We now show how our micro-foundation **(A-VPI)** implements **(A-VP)**.

Case 1—Hire Without Renegotiation. In this case, the outcome is already *privately efficient*. The worker's value does not decrease ($w'_1 = w_1$), and by the fact that a hire occurs, the firm's value must increase. We can also write the expected return as $qv[\Omega(z, 2) - \Omega(z, 1) - U]$. Since the return is independent of w_1 , then the efficient vacancy policy is independent of w_1 . The firm is choosing vacancies as if it were maximizing the joint surplus without having to appeal to additional assumptions.

In cases 2 and 3, the outcome is *privately inefficient* because the firm may profit from vacancies that, if met by a job-seeker, deliver a credible threat to cut the incumbent's wage to $w'_1 < w_1$.

Our assumption (A-VPI) allows the worker to correct for this overposting by conceding a single pay cut in exchange for an alternative level of vacancies.⁴⁷ The firm will accept this wage cut and choose the worker's preferred vacancies if it delivers at least the value obtained under the firm's preferred vacancies v^f . We show that the worker's preferred package satisfying incentive compatibility restores efficiency in vacancy posting.

Case 2—Hire With Renegotiation. In this case, the incumbent's wage w_1 is high enough that the firm finds it profitable to raise the contact probability with an unemployed worker beyond what would be efficient. Although the hiring outcome is efficient *ex post*, too many resources are spent on vacancies *ex ante*. Let w_1^* be the renegotiated wage after a meeting. The worker chooses a package of vacancies and a wage cut in all states (v^w, x) that solves

$$\max_{v^w, x} qv^w(w_1^* - w_1) - x \quad (21)$$

subject to

$$\begin{aligned} & qv^w[(y(z, 2) - (w_1 - x) - b) - (y(z, 1) - w_1)] - c(v^w) \\ & \geq qv^f[(y(z, 2) - w_1^* - b) - (y(z, 1) - w_1)] - c(v^f). \end{aligned} \quad (IC)$$

The worker anticipates that after a meeting their wage will be renegotiated to $w_1^* < w_1$. Given this wage cut, the worker seeks to limit the probability of this event by cutting back on vacancies. Incentive compatibility (IC) requires that as the worker cuts vacancies it also cuts its wage so that the firm accepts the proposed policy v^w over v^f .

The Pareto problem (21) yields the result that vacancy posting is independent of w_1 . First, given the linear objective function, (IC) holds with equality. Thus, we can substitute out x . Second, the zero-sum game assumption (A-IN) implies that w_1^* is a renegotiated wage that only redistributes value, and hence drops out. Third, all terms that do not depend on (x, v^w) are irrelevant to the worker's decision. Adopting the value notation, this leaves the following objective function:

$$\max_{v^w} qv^w[(\Omega(z, 2) - U) - \Omega(z, 1)] - c(v^w).$$

The decision can therefore be characterized by the *privately efficient return*, which is the change in joint value net of the cost of the new hire, $\Omega(z, 2) - \Omega(z, 1) - U$.

⁴⁷A pay cut regardless of the outcome of the search for a new worker maps exactly into a transfer from worker to firm, which is how we approach the proof. Promising *state-contingent* wage cuts that depend on who the firm meets or whether a meeting occurs is not possible given our assumption of what is verifiable and contractible. Even if these states were verifiable, the result would only be for the worker to offer a menu of wage cuts across states. This would increase worker value but not change allocations; hence for consistency with the rest of our assumptions, we assume a single wage cut.

Case 3—No Hire With Renegotiation. In this case, the “fire and swap” threat is credible. The incumbent’s wage w_1 is high enough and the marginal product of an additional worker is below b . Replacing the return to hiring by the wage cut for the incumbent worker, the previous logic delivers

$$\max_{v^w} qv^w [\Omega(z, 1) - \Omega(z, 1)] - c(v^w) \implies v^w = 0.$$

Absent the transfer from worker to firm, the firm would post positive vacancies v^f even if the return from hiring is negative, that is, $\Omega(z, 2) - \Omega(z, 1) < U$ to induce a wage cut, and v^f would depend on w_1 . Under **(A-VPI)**, the worker takes a preemptive wage cut, and vacancies are zero, the efficient amount in this case.

Combined. Combining all three cases, privately efficient vacancies solve

$$\max_v qv \left[\max \left\{ \frac{\Omega(z, 2) - \Omega(z, 1)}{2 - 1} - U, 0 \right\} \right] - c(v).$$

Note three properties of this solution. First, the firm always hires when it meets an unemployed worker. Second, optimal vacancy posting equates the marginal gain in joint value to the marginal cost of a vacancy, and it only depends on (z, n) . Third, this condition is the flip-side of the separation frontier. Expression (1) states that if $\Omega_n(z, n) > U$, then the firm will not separate with workers. The terms inside the max expression say that if this is true, then the firm will post vacancies.⁴⁸

We conclude that under **(A-VPI)**, the joint value is sufficient to characterize the vacancy decision. The distribution of wages in the firm is immaterial.

Multiple Incumbents. When the firm employs more than one worker, the efficient transfer scheme can be implemented by randomly selecting a worker under threat to offer a package of wage cuts and vacancies. In exchange, the firm posts the efficient number of vacancies. Under such a scheme, the initiating worker is strictly better off while the firm and the other workers are indifferent. We establish this case in detail in Appendix II.

A.5. Layoffs, Quits, Exit, Entry

Having described most of the terms in the HJB (1), we conclude with the boundary conditions for exit, layoffs, and the free-entry condition.

Layoffs. Consider now a firm with $n = 2$ workers paid (w_1, w_2) , and assume that $w_1 < y(z, 1)$ such that worker 1 is never under threat of layoff. The firm has a credible threat to fire worker 2 if

$$y(z, 1) - w_1 > y(z, 2) - w_1 - w_2.$$

Such a situation may occur if, for example, productivity has just declined. The firm has a credible threat to negotiate down to a wage level w_2^* such that $y(z, 1) - w_1 = y(z, 2) - w_1 - w_2^*$ and keep worker 2 employed. From the worker’s perspective, it is individually rational

⁴⁸It is possible to determine the optimal wage cut x that delivers the efficient policy, but throughout the paper we focus on allocations only.

to accept any wage w_2^* above b . Worker 2 is laid off if $y(z, 1) - w_1 > y(z, 2) - w_1 - b$. In terms of joint value, this can be written in exactly the form of the layoff frontier (2):

$$\frac{\Omega(z, 2) - \Omega(z, 1)}{2 - 1} < U.$$

The firm lays off workers until the marginal joint value of the worker is equal to the value of unemployment.⁴⁹ As noted earlier, this is the complement to the condition for posting vacancies. The special case with $n = 1$ of this scenario also arises in the one worker–one firm model with productivity shocks of Postel-Vinay and Turon (2010).

Quits to Unemployment. Since in this static model workers will accept a renegotiated wage down to $w_i^* = b$, they will only quit at the point where the firm has a credible threat to lower wages below b . This is exactly the point at which the marginal value is equal to the value of unemployment. In this sense, *layoffs* as described above are indistinguishable from quits to unemployment, as in any model with privately efficient separations. For ease of language, all *endogenous UE* transitions are referred to as *layoffs*, and we use *quits* to refer only to *EE* transitions.

Finally, recall that in the dynamic model unemployed job-seekers are promised a wage that implements a value U to them. If events occur in the firm that reduce the continuation value to that worker below U (e.g., a negative productivity shock), the incumbent may have a credible threat to quit and renegotiate her wage to restore its value at U , or above it, depending on the details of the internal negotiation. However, such renegotiation is, again, only a transfer of value within the firm. Separations remain privately efficient even in the dynamic model.

Exit. Now consider the exit decision of a firm with one worker. The private value of exit to the firm is the scrap value $\vartheta > 0$. The firm therefore exits if and only if $y(z, 1) - w_1^* < \vartheta$, where w_1^* is a possibly renegotiated wage contingent on the firm remaining in operation. If the profit from operating at the lowest possible renegotiated wage $w_1^* = b$ is greater than ϑ , then the firm will continue to operate. Hence the firm exits if $y(z, 1) - b < \vartheta$, and the renegotiated wage only affects the distribution of value.⁵⁰ The exit condition can be written as $\Omega(z, 1) - U < \vartheta$, and in the general case of n workers is exactly the boundary condition in (1): $\Omega(z, n) - nU < \vartheta$.

Entry. Upon entry, the firm has n_0 workers hired from unemployment. The private entry cost of the firm is c_0 , so entry requires $\int y(z, n_0) d\Pi_0(z) - n_0b > c_0$. Using $\Omega(n, z) = y(z, n)$ and $U = b$, this requires $\int \Omega(z, n_0) d\Pi_0(z) > c_0 + n_0U$.

A.6. From Static to Dynamic

This static example showcases how to obtain every component of (1) from our set of assumptions. Appendix II generalizes this proof to the dynamic case. Two insights assist

⁴⁹Note that, when both workers are under threat, the particular order in which values of workers are reduced is immaterial to the condition $\Omega(z, 2) - \Omega(z, 1) < U$. One could, for example, lower the wages of both workers proportionally, increasing the value of the firm, but a worker *must* be fired if $y(z, 2) - w_1^* - b < y(z, 1) - w_1^*$ for any $w_1^* \geq b$.

⁵⁰The firm has no credible threat to reduce w_1 if $y(z, 1) - w_1 > \vartheta$. The firm can credibly threaten exit if $\vartheta \in (y(z, 1) - w_1, y(z, 1) - b)$, but in this case w_1 can be reduced to a point where this threat is no longer credible.

us. First, the proof begins with a discrete workforce. Here, we are helped by continuous time, which removes complicated binomial probabilities of one, two, three, etc. incumbent workers meeting a competitor's vacancy. Second, we take the continuous workforce limit of the discrete workforce HJB equation. This limit delivers the joint value representation (1) in terms of the derivative of the joint value function rather than differences of values which, when moving up or down by one worker, are symmetric due to continuous differentiability.

APPENDIX B: CHARACTERIZATION OF SURPLUS FUNCTION

Here, we prove the comparative statics on the surplus function $S(n, z)$ discussed in the main text.

B.1. *Hamilton–Jacobi–Bellman Variational Inequality for Total Value*

Before characterizing these conditions, we note that the joint value representation (1) and smooth-pasting boundary conditions that define the exit and layoff boundaries (2) are derived from solving the following Hamilton–Jacobi–Bellman variational inequality (see [Pham \(2009\)](#)), which we present here for completeness. Its general formulation in terms of optimal switching between three regimes (operation, layoffs, exit) on the entire positive quadrant, can be written as the following system:

$$\begin{aligned}
 & \max \left\{ -\rho\Omega(z, n) + \max_{v \geq 0} -\delta n [\Omega_n(z, n) - U] + qv\phi [\Omega_n(z, n) - U] \right. \\
 & \quad + qv(1 - \phi) \int \max[\Omega_n(z, n) - \Omega_n(z', n'), 0] d\tilde{H}_n(z', n') \\
 & \quad + \mu(z)\Omega_z(z, n) + \frac{\sigma(z)^2}{2}\Omega_{zz}(z, n); \\
 & \quad \left. \underbrace{\vartheta + nU - \Omega(z, n)}_{\text{Exit}}; \underbrace{\max_{k \in [0, n]} \Omega(z, k) + (n - k)U - \Omega(z, n)}_{\text{Layoff}} \right\} \\
 & = 0, \quad \forall (z, n) \in \mathbb{R}_+^2. \tag{22}
 \end{aligned}$$

The HJBVI implies necessary “value-matching” and “smooth-pasting” boundary conditions; see [Brekke and Oksendal \(1990\)](#), [Peskir and Shiryaev \(2006\)](#) and [Stokey \(2009\)](#).

Value matching conditions are standard, and simply state that the value function must be continuous at the exit and separation boundaries. Smooth pasting conditions obtain only when coalitions are actually crossing the exit or layoff boundaries. Intuitively, coalitions can then take an interior first-order optimality condition when they choose the stopping boundary. Thus, smooth pasting obtains either when there is volatility, or when the drift pushes coalitions outside of the continuation region.

Combining these observations, for exit, we have a value matching condition that holds for the entire boundary $n_E^*(z)$, and a smooth pasting condition in the n direction that holds only where the drift is negative and firms actually exit. We have a smooth pasting condition in the z direction that holds for the entire boundary $n_E^*(z)$ because there is volatility in the z direction. We collect these conditions in Conditions (iii) in Section 4.1.

B.2. S Is Increasing in n

The no-endogenous-separations condition $S_n \geq 0$ implies that the surplus is increasing in n .

B.3. S Is Increasing in z

Rewrite the problem in terms of $x = \log z$. Denote with a slight abuse of notation $y(x, n) = y(e^x, n)$.

$$\begin{aligned} \rho S(x, n) &= \max_{v \geq 0} y(x, n) - nb - c(v) \\ &\quad + [q\phi v - \delta n]S_n(x, n) + q(1 - \phi)v\mathcal{H}(S_n(x, n)) \\ &\quad + \mu S_x(x, n) + \frac{\sigma^2}{2} S_{xx}(x, n), \end{aligned}$$

where we integrated by parts, and denoted $\mathcal{H}(s) = \int_0^s H_n(r) dr$. Denote $\zeta(x, n) = S_x(x, n)$. Differentiate the Bellman equation w.r.t. x and use the envelope theorem to obtain

$$\begin{aligned} \rho \zeta(x, n) &= y_x(x, n) \\ &\quad + \{[q\phi + q(1 - \phi)H_n(S_n(x, n))]v^*(x, n) - \delta n\} \zeta_n(x, n) \\ &\quad + \mu \zeta_x(x, n) + \frac{\sigma^2}{2} \zeta_{xx}(x, n). \end{aligned}$$

Now consider the stochastic process defined by

$$\begin{aligned} dx_t &= \mu dt + \sigma dW_t, \\ dn_t &= \{[q\phi(1 - \phi)H_n(S_n(x_t, n_t)) + q\phi]v^*(x_t, n_t) - \delta n_t\} dt. \end{aligned} \quad (23)$$

This corresponds to the true stochastic process for productivity, but a hypothetical process for employment, that in general differs from the realized one. We can now use the Feynman–Kac formula (Pham (2009)) to go back to the sequential formulation:

$$\zeta(x, n) = \mathbb{E} \left[\int_0^T e^{-\rho t} y_x(x_t, n_t) + e^{-\rho T} \zeta(x_T, n_T) | x_0 = x, n_0 = n, \{x_t, n_t\} \text{ follows (23)} \right]$$

and where T is the hitting time of either the separation or exit region. By assumption, $y_x > 0$, so the contribution of the first part is always positive. On the exit region, smooth-pasting requires that $\zeta = 0$. In the interior of the separation region, $\zeta = 0$. Under our regularity assumption, we thus get $\zeta = 0$ on the layoff boundary. Thus,

$$\zeta(x, n) = \mathbb{E} \left[\int_0^T e^{-\rho t} y_x(x_t, n_t) dt | x_0 = x, n_0 = n, \{x_t, n_t\} \text{ follows (23)} \right] > 0,$$

which concludes the proof.

B.4. S Is Concave in n

Denote $s(z, n) = S_n(z, n)$. Differentiate the Bellman equation w.r.t. n on the interior of the domain, use the envelope theorem and integrate by parts to obtain

$$\begin{aligned} (\rho + \delta)s(z, n) &= y_n(z, n) - b \\ &\quad + \{[q\phi + q(1 - \phi)H_n(s(z, n))]v^*(z, n) - \delta n\}s_n(z, n) \\ &\quad + \mu(z)s_z(z, n) + \frac{\sigma^2(z)}{2}s_{zz}(z, n). \end{aligned}$$

Recall that

$$(1 + \gamma)\bar{c}[v^*(z, n)]^\gamma = q\phi s(z, n) + q(1 - \phi)\mathcal{H}(s(z, n)). \quad (24)$$

In particular, differentiating w.r.t. n ,

$$\gamma(1 + \gamma)\bar{c}[v^*(z, n)]^{\gamma-1}v_n^*(z, n) = [q\phi + q(1 - \phi)H_n(s(z, n))]s_n(z, n)$$

and so

$$\gamma \frac{v_n^*(z, n)}{v^*(z, n)} = \frac{\phi + (1 - \phi)H_n(s(z, n))}{\phi + (1 - \phi)\bar{H}(s(z, n))} \frac{s_n(z, n)}{s(z, n)},$$

where $\bar{H}(s) = \frac{\mathcal{H}(s)}{s} \leq 1$. Now denote $\zeta(z, n) = s_n(z, n) = S_{nn}(z, n)$. Differentiate the recursion for s w.r.t. n to obtain

$$\begin{aligned} &(\rho + 2\delta - q(1 - \phi)H'_n(s(z, n))v^*(z, n)s_n(z, n) \\ &\quad - q[\phi + (1 - \phi)H_n(s(z, n))v_n^*(z, n)]\zeta(z, n) \\ &= y_{nn}(z, n) \\ &\quad + \{[\lambda\phi + \lambda(1 - \phi)H_n(s(z, n))]v^*(z, n) - \delta n\}\zeta_n(z, n) \\ &\quad + \mu(z)\zeta_z(z, n) + \frac{\sigma^2(z)}{2}\zeta_{zz}(z, n). \end{aligned}$$

Now define the “effective discount rate”

$$\begin{aligned} R(z, n, s_n(z, n)) &= \rho + 2\delta - q(1 - \phi)H'_n(s(z, n))v^*(z, n)s_n(z, n) - q[\phi + (1 - \phi)H_n(s(z, n))]v_n^*(z, n) \\ &= \rho + 2\delta - qv^*(z, n)s_n(z, n) \\ &\quad \times \underbrace{\left\{ (1 - \phi)H'_n(s(z, n)) + \frac{\phi + (1 - \phi)H_n(s(z, n))}{\gamma s(z, n)} \frac{\phi + (1 - \phi)H_n(s(z, n))}{\phi + (1 - \phi)\bar{H}(s(z, n))} \right\}}_{\equiv P(z, n) > 0}, \end{aligned}$$

where the second equality uses the expression for v_n^* derived above. Define the stochastic process

$$\begin{aligned} dz_t &= \mu(z_t) dt + \sigma(z_t) dW_t, \\ dn_t &= \{[q(1 - \phi)H_n(S_n(z_t, n_t)) + q\phi]v^*(z_t, n_t) - \delta n_t\} dt. \end{aligned} \quad (25)$$

As before, we can use the Feynman–Kac formula to obtain

$$\zeta(z, n) = \mathbb{E} \left[\int_0^T e^{-\int_0^t R(z_\tau, n_\tau, \zeta(z_\tau, n_\tau)) d\tau} y_{nn}(z_t, n_t) dt + e^{-\int_0^T R(z_\tau, n_\tau, \zeta(z_\tau, n_\tau)) d\tau} \zeta(z_T, n_T) \right. \\ \left. \middle| z_0 = z, n_0 = n, \{z_t, n_t\} \text{ follows (25)} \right]$$

for T the first hitting time of the exit/separation region. The contribution of the first term is always negative. Note that ζ enters in the effective discount rate. Inside the separation region and in the exit regions, $\zeta = 0$. We restrict attention to twice continuously differentiable functions, so $\zeta = 0$ on the exit and separation frontiers. Then

$$\zeta(z, n) = \mathbb{E} \left[\int_0^T e^{-\int_0^t R(z_\tau, n_\tau, \zeta(z_\tau, n_\tau)) d\tau} y_{nn}(z_t, n_t) dt \middle| z_0 = z, n_0 = n, \{z_t, n_t\} \text{ follows (25)} \right] < 0,$$

which concludes the proof.

B.5. S Is Supermodular in $(\log z, n)$

Denote again $s(x, n) = S_n(x, n)$, where $x = \log z$. Recall that

$$(\rho + \delta)s(x, n) = y_n(x, n) - b \\ + \{ [q\phi + q(1 - \phi)H_n(s(x, n))]v^*(x, n) - \delta n\} s_n(x, n) \\ + \mu s_x(x, n) + \frac{\sigma^2}{2} s_{xx}(x, n)$$

and that

$$(1 + \gamma)\bar{c}[v^*(x, n)]^\gamma = q\phi s(x, n) + q(1 - \phi)\mathcal{H}(s(x, n)).$$

In particular, differentiating w.r.t. x ,

$$\gamma \frac{v_x^*(x, n)}{v^*(x, n)} = \frac{\phi + (1 - \phi)H_n(s(x, n))}{\phi + (1 - \phi)\bar{H}(s(x, n))} \frac{s_x(x, n)}{s(x, n)}.$$

Now denote $\zeta(x, n) = s_x(x, n) = S_{xn}(x, n)$. Differentiate the recursion for $s(x, n)$ w.r.t. x to obtain

$$(\rho + \delta - q(1 - \phi)H'_n(s(x, n)))v^*(x, n)s_x(x, n) \\ - q[\phi + (1 - \phi)H_n(s(x, n))]v_x^*(x, n)\zeta(x, n) \\ = y_{nx}(x, n) \\ + \{ [\lambda\phi + \lambda(1 - \phi)H_n(s(x, n))]v^*(x, n) - \delta n\} \zeta_n(x, n) \\ + \mu \zeta_x(x, n) + \frac{\sigma^2}{2} \zeta_{xx}(x, n).$$

As before, define the “effective discount rate”

$$\begin{aligned}
R(x, n, s_x(x, n)) &= \rho + \delta - q(1 - \phi)H'_n(s(x, n))v^*(x, n)s_x(x, n) - q[\phi + (1 - \phi)H_n(s(x, n))]v^*_x(x, n) \\
&= \rho + \delta - qv^*(x, n)s_x(x, n) \\
&\quad \times \underbrace{\left\{ (1 - \phi)H'_n(s(x, n)) + \frac{\phi + (1 - \phi)H_n(s(x, n))}{\gamma s(x, n)} \frac{\phi + (1 - \phi)H_n(s(x, n))}{\phi + (1 - \phi)\bar{H}(s(x, n))} \right\}}_{\equiv P(x, n) > 0},
\end{aligned}$$

where the second equality uses the expression for v_n^* derived above. As before, define the stochastic process

$$\begin{aligned}
dx_t &= \mu dt + \sigma dW_t, \\
dn_t &= \left\{ [q(1 - \phi)H_n(S_n(e^{x_t}, n_t)) + q\phi]v^*(x_t, n_t) - \delta n_t \right\} dt.
\end{aligned} \tag{26}$$

As before, we can use the Feynman–Kac formula to obtain

$$\begin{aligned}
\zeta(x, n) &= \mathbb{E} \left[\int_0^T e^{-\int_0^t R(x_\tau, n_\tau, \zeta(x_\tau, n_\tau)) d\tau} y_{nx}(x_t, n_t) dt + e^{-\int_0^T R(x_\tau, n_\tau, \zeta(x_\tau, n_\tau)) d\tau} \zeta(x_T, n_T) \right. \\
&\quad \left. \middle| x_0 = z, n_0 = n, \{x_t, n_t\} \text{ follows (26)} \right]
\end{aligned}$$

for T the first hitting time of the exit/separation region. The contribution of the first term is always positive. Inside the separation region and in the exit regions, $\zeta = 0$. We restrict attention to twice continuously differentiable functions, so $\zeta = 0$ on the exit and separation frontiers. Then

$$\begin{aligned}
\zeta(x, n) &= \mathbb{E} \left[\int_0^T e^{-\int_0^t R(x_\tau, n_\tau, \zeta(x_\tau, n_\tau)) d\tau} y_{nx}(x_t, n_t) dt \right. \\
&\quad \left. \middle| x_0 = z, n_0 = n, \{x_t, n_t\} \text{ follows (26)} \right] > 0,
\end{aligned}$$

which concludes the proof.

B.6. Net Employment Growth

Denote again $s(z, n) = S_n(z, n)$. Net employment growth in the continuation region is

$$\frac{dn_t}{dt} = q[\phi + (1 - \phi)H_n(s(z, n))]v^*(z, n) - \lambda^E(1 - H_v(s(z, n)))n - \delta n \equiv g(z, n).$$

Using the expression the optimal vacancy condition $v^*(z, n)$ in (24),

$$\begin{aligned}
g(z, n) &= \frac{q^{1+1/\gamma}}{[(1 + \gamma)\bar{c}]^{1/\gamma}} (\phi + (1 - \phi)H_n(s(z, n))) (\phi s(z, n) + (1 - \phi)\mathcal{H}(s(z, n)))^{1/\gamma} \\
&\quad - \lambda^E(1 - H_v(s(z, n)))n - \delta n.
\end{aligned}$$

From the previous comparative statics on $S(z, n)$, it is straightforward to see that $g(z, n)$ is increasing in $\log z$ and decreasing in n .

APPENDIX C: ALTERNATIVE CRS LIMIT

Consider the surplus equation (4). Assume $\alpha = 1$, $\vartheta = 0$ and vacancy costs homogeneous of degree one in (v, n) . Let $\nu = v/n$. In this case, the joint surplus is linear in n , $S(z, n) = \hat{S}(z)n$, where

$$\begin{aligned} (\rho + \delta)\hat{S}(z) &= \max_{\nu \geq 0} y(z) - b + q(\theta)\nu \left[\phi \bar{S}(z) + (1 - \phi) \int_0^{\hat{S}(z)} (\hat{S}(z) - S') dH_n(S') \right] - c(\nu) \\ &\quad + \mu(z)\hat{S}_z(z) + \frac{\sigma^2(z)}{2}\hat{S}_{zz}(z), \end{aligned} \quad (27)$$

where, once again, $H_n(S') = H(z)$ and the *marginal* surplus still depends only on exogenous productivity z . The model continues to behave like a one-worker–one-firm model for all *firm decisions*, up to rescaling vacancies by size. This economy, however, produces different worker dynamics from the limiting one described in Section 4.3.1 since gross hires now depend on firm size.

APPENDIX D: FRICTIONLESS LIMITS

D.1. Setup

Frictional Problem. Start by recalling the Bellman equation for the joint surplus in the frictional case:

$$\begin{aligned} \rho S(z, n) &= \max_{\nu} y(z, n) - nb - c(\nu) - \delta n S_n(z, n) \\ &\quad + q(\theta)\nu \left\{ \phi S_n + (1 - \phi) \int_0^{S_n} H_n(s) ds \right\} \\ &\quad + (\mathbb{L}S)(z, n) \\ \text{s.t. } S(z, n) &\geq \vartheta, \quad S_n(z, n) \geq 0, \end{aligned} \quad (28)$$

where H_n is the employment-weighted cumulative distribution function of marginal surpluses. \mathbb{L} is the differential operator that encodes the continuation value from productivity shocks, $(\mathbb{L}S)(z, n) = \mu(z)S_z(z, n) + \frac{\sigma(z)^2}{2}S_{zz}(z, n)$. Recall that $\phi = \frac{u}{u + \xi(1-u)}$ is the probability that a vacancy meets an unemployed worker, and q is the vacancy meeting rate.

Inside the continuation region, the density function $h(z, n)$ of the distribution of firms by productivity and size is determined by the stationary KFE,

$$0 = -\frac{\partial}{\partial n}(h(z, n)g(z, n)) + (\mathbb{L}^*h)(z, n),$$

where \mathbb{L}^* is the formal adjoint of the operator \mathbb{L} , and $g(z, n)$ is the growth rate of employment

$$g(z, n) = q(\theta)v^*(z, n) \left[\phi + (1 - \phi)H_n(S_n(z, n)) \right] - \xi \lambda^U n \left[1 - H_v(S_n(z, n)) \right], \quad (29)$$

where λ^U is the meeting rate from unemployment, and ξ the relative search efficiency of the employed.

The mass of entrant firms m_0 is determined by the free-entry condition

$$c_0 = \mathbb{E}_0[S(z, n_0)], \quad (30)$$

where n_0 is initial employment, which is a parameter, and \mathbb{E}_0 is the expectation operator under the productivity distribution for entrants $\Pi_0(z)$. Note that the surplus is a function of m_0 through the vacancy meeting rate $q(\theta)$, since θ is increasing in m_0 .

For ease of exposition, and without loss of generality, we make three additional assumptions. First, we consider isoelastic vacancy cost functions

$$c(v) = \frac{\bar{c}}{1 + \gamma} v^{1+\gamma},$$

and normalize $\bar{c} = 1$, but the result does not depend on the particular functional form nor on the normalization. Also, we specialize to a Cobb–Douglas matching function $m(s, v) = As^\beta v^{1-\beta}$, where A is match efficiency, a proxy for labor market frictions. Third, we set to zero exogenous separations to unemployment $\delta = 0$, but endogenous separations when $S(n, z) < 0$ still occur, and we denote by Δ the aggregate endogenous separation rate.

To ease notation, we write $B \approx C$ for a first-order Taylor expansion. We also denote $\|S_n\| = \mathbb{E}^{SS}[S_n^{1/\gamma}]^\gamma$, where \mathbb{E}^{SS} denotes the expectation under the steady-state distribution of marginal surpluses. This is also the Lebesgue $(1/\gamma)$ -norm of S_n under the steady-state probability measure.

Finally, we note that in characterizing the limits we make use of the fact that both m_0 and v must remain finite: infinite entry and vacancy costs would violate the economy's resource constraint.

Comparative Statics. We describe behavior of the economy in the limit when match efficiency $A \rightarrow \infty$. We do so for two different configurations of the economy:

1. No on-the-job search: $\xi = 0$
2. On-the-job search: $\xi > 0$

D.2. No on-the-Job Search

Since $\xi = 0$, $\phi = 1$. From (28), the first order condition for vacancies gives

$$v^*(z, n) = (qS_n)^{1/\gamma}. \quad (31)$$

Using this optimality condition in the value function of hiring firms,

$$\begin{aligned} \rho S(z, n) &= y(z, n) - nb + \frac{\gamma}{1 + \gamma} \cdot q(\theta)^{\frac{1}{1+\gamma}} S_n^{\frac{1}{1+\gamma}} + (\mathbb{L}S)(z, n) \\ \text{s.t. } S(z, n) &\geq \vartheta, \quad S_n(z, n) \geq 0, \end{aligned}$$

which now only depends on $q(\theta)$ as the sole aggregate. Hence, free-entry (30) uniquely pins down $q(\theta)$ to the same number no matter what value A takes. Therefore, the value function always satisfies the same Bellman equation, irrespective of A . Hence, throughout the state space, at any given (n, z) , marginal surpluses $S_n(z, n)$ remain the same as A

varies. Moreover, since the value $S(z, n)$ is independent from A , so are all the decisions by firms. As a result, the endogenous separation rate Δ always remains the same, and in particular, finite.

Aggregates in the Limit. We now study how aggregates v, u, θ evolve along this limiting path. Given the matching function, these determine all other equilibrium objects: q, λ^U, λ^E .

Integrating both sides of the FOC for vacancies under the firm distribution, and using the matching function which implies that $q = A\theta^{-\beta}$, aggregate vacancies are

$$v = m_0 q^{\frac{1}{\gamma}} \|S_n\|^{\frac{1}{\gamma}} = m_0 A^{\frac{1}{\gamma}} \theta^{-\frac{\beta}{\gamma}} \|S_n\|^{\frac{1}{\gamma}}.$$

Since q remains constant, and v and m_0 are finite in the limit, then the first equality implies that $\|S_n\|$ remains finite in the limit.

In the limit, the unemployment rate is $u \approx \frac{A}{\lambda^U}$. The matching function implies $\lambda^U = A\theta^{1-\beta}$. Combined, the unemployment rate is $u \approx \Delta A^{-1} \theta^{-(1-\beta)}$. Combining these expressions with the expression for aggregate vacancies v , tightness satisfies

$$\theta = \frac{v}{u} \approx \frac{m_0 A^{\frac{1}{\gamma}} \theta^{-\frac{\beta}{\gamma}} \|S_n\|^{\frac{1}{\gamma}}}{\Delta A \theta^{1-\beta}}$$

so that

$$\theta^{\beta \frac{1+\gamma}{\gamma}} \approx \left(\frac{m_0}{\Delta} \right) \|S_n\|^{\frac{1}{\gamma}} A^{\frac{1+\gamma}{\gamma}}.$$

Since m_0, Δ , and $\|S_n\|$ are finite, θ diverges with A . Therefore, λ_U diverges as well. On the worker side, since λ_U diverges to infinity, u goes to zero. On the firm side, m_0 remains finite, but changes such that q remains constant and vacancies remain finite.

Invariant Distribution of Marginal Surpluses. We now turn to the invariant distribution $h(z, n)$. After substituting optimal vacancies into (29) evaluated at $\xi = 1 - \phi = 0$, one obtains that the growth of employment in the hiring region is

$$g(z, n) = q(qS_n(z, n))^{\frac{1}{\gamma}}.$$

Since $S_n(z, n)$ remains constant throughout the state space, then employment growth in the hiring region remains constant throughout the state space. The firm loses no workers to employment because there is no on-the-job search. Since $S_n(z, n)$ and $U = b/\rho$ both stay unchanged, then the employment outflows to unemployment are still unchanged. Since $S(z, n)$ is unchanged, then the exit decision is also unchanged.

Hence the law of motion of employment is independent of A and the steady-state distribution $h(z, n)$ is also independent from A . Therefore, the values of firms $S(z, n)$ are the same across the state space and the relative mass of firms at each (z, n) is unchanged, despite higher but finite mass of entrants m_0 .

D.3. On-the-Job Search

We now turn to the case in which on-the-job search remains positive at some fixed value $\xi > 0$. We follow the same logic as before, with some additional steps due to on-the-job search.

Consider (28) written in terms of the return on a vacancy $R(S_n)$:

$$\begin{aligned} \rho S(z, n) &= \max_v y(z, n) - nb - c(v) + q(\theta)vR(S_n) + (\mathbb{L}S)(z, n) \\ \text{s.t. } S(z, n) &\geq \vartheta, \quad S_n(z, n) \geq 0, \end{aligned}$$

where

$$R(S_n) = \phi S_n + (1 - \phi) \int_0^{S_n} H_n(s) ds. \quad (32)$$

The growth of employment is

$$g(z, n) = qv[\phi + (1 - \phi)H_n(S_n(z, n))] - \xi\lambda^U n[1 - H_v(S_n(z, n))]. \quad (33)$$

Aggregates in the Limit. We restrict attention to the economically meaningful case in which (1) output and aggregate vacancies remains finite and strictly positive in the limit, and (2) the rate at which workers separate into unemployment remains finite in the limit. These restrictions are equivalent to a guess and verify strategy, in which we guess that (1–2) hold and then verify those conditions.

Consider first meeting rates. Because some measure n of employed job-seekers are always present regardless of A , effective search effort $s = u + \xi n$ remains finite and positive even if u goes to zero. By (1), vacancies also remain finite. Combined, these imply that market tightness $\theta = v/s$ remains finite. Since $q = A\theta^{-\beta}$ and $\lambda^U = A\theta^{1-\beta}$, then both meeting rates diverge to infinity at the same rate as A .⁵¹

Consider unemployment and aggregate vacancies. (2) requires that the rate at which workers separate into unemployment is a positive constant Δ in the limit. Since $u \approx \frac{\Delta}{\lambda^U}$, and λ^U diverges, then the unemployment rate converges to zero. Since the unemployment rate converges to zero, then ϕ also converges to zero, and thus $s = \xi$. Firm level and aggregate vacancies are given by

$$v = q^{\frac{1}{\gamma}} R(S_n)^{\frac{1}{\gamma}}, \quad v = m_0 q^{\frac{1}{\gamma}} \|R(S_n)\|^{\frac{1}{\gamma}}. \quad (34)$$

(1) implies that both aggregate vacancies v and the mass of entering firms m_0 remain finite. Since v is finite and m_0 is finite, while q diverges at the same rate as A , then $\gamma > 0$ requires $\|R(S_n)\|$ must go to zero at the same rate as A goes to infinity.

Invariant Distribution of Marginal Surpluses. We now show that the distribution of marginal surpluses degenerates to a single value on the support of the invariant distribution.

First, we use (34) to express firm level vacancies as a share of aggregate vacancies, where that share is determined by the firms' return on a vacancy relative to the average return:

$$v = \frac{1}{m_0} \left(\frac{R(S_n)}{\|R(S_n)\|} \right)^{\frac{1}{\gamma}} v = \frac{1}{m_0} \left(\frac{R(S_n)}{\|R(S_n)\|} \right)^{\frac{1}{\gamma}} \left(\frac{\lambda^U \xi}{q} \right), \quad (35)$$

⁵¹Strictly speaking, free-entry then ensures that θ is pinned down to a strictly positive value. This proof is more lengthy but does not require any additional assumptions and is available upon request.

where the second equality uses $q = A(\varpi/\xi)^{-\beta}$, and $\lambda^U = A(\varpi/\xi)^{1-\beta}$, which jointly imply that $\varpi = \lambda^U \xi/q$. Now consider the expression for growth of employment inside the continuation region (33), under the limiting case of $\phi = 0$:

$$g(z, n) \approx qvH_n(S_n(z, n)) - \xi\lambda^U n[1 - H_v(S_n(z, n))].$$

Substituting in the expression for firm vacancies (35) and collecting $\lambda^U \xi$ terms,

$$g(z, n) \approx \lambda^U \xi \left\{ \frac{1}{m_0} \left(\frac{R(S_n)}{\|R(S_n)\|} \right)^{\frac{1}{\gamma}} H_n(S_n) - n[1 - H_v(S_n)] \right\}.$$

Since λ^U diverges but growth must remain finite on the support of the invariant distribution, the term in braces must be equal to zero in the limit:

$$\frac{1}{m_0} \left(\frac{R(S_n)}{\|R(S_n)\|} \right)^{\frac{1}{\gamma}} H_n(S_n) = n[1 - H_v(S_n)]. \quad (36)$$

Using this relation, we can show that the distribution of marginal surplus converges pointwise to a degenerate limiting distribution H_n^∞ , that is, for every z there is a unique $n^*(z)$.

We proceed by contradiction. Suppose that H_n converges to a limiting distribution H_n^∞ that is non-degenerate.⁵² Consider a firm at the top of the distribution, such that $1 - H_v(S_n) = 0$. The probability that the firm loses a worker is zero, so the right-hand side is zero. However, by the supposition that H_n is nondegenerate, then $R(S_n)$ in (32) converges to a nonzero value, since the firm can increase its value by poaching from workers below it on the ladder. Since there is some $R(S_n)$ that is nonzero, then $\|R(S_n)\|$ also converges to a nonzero value. Therefore, the right-hand side of (36) is zero, but the left-hand side is positive, which violates the above equality, a contradiction. Hence, in the limit H_n^∞ must be degenerate, and marginal surpluses of firms converge to a common limit, which we denote S_n^* .

Since the limiting distribution H_n^∞ is degenerate at every z , the invariant joint distribution of employment and productivity lines up along a strip $\{z, n^*(z)\}$ implicitly defined by $S_n(n^*(z), z) = S_n^*$. Since $S_{nn} < 0$ and $S_{zn} > 0$, $n^*(z)$ is strictly increasing.

Unique Value for S_n^ on the Limiting Strip.* We now show that the unique equilibrium value of the marginal surplus, S_n^* , is zero. The first step of the proof is to express the marginal surplus as the present discounted value of flow marginal products $y_n - b$. Second, we show that these marginal products would be equal to $-b$ should the marginal surplus S_n^* be any strictly positive value. We then conclude that $S_n^* = 0$.

For the first step, we start by maximizing out vacancies in the HJB in (28) to obtain

$$\begin{aligned} \rho S(z, n) = & y(z, n) - bn + \frac{\gamma}{1 + \gamma} \left(q\phi S_n(z, n) + q(1 - \phi) \int_0^{S_n(z, n)} H_n(s) ds \right)^{\frac{1+\gamma}{\gamma}} \\ & + (\mathbb{L}S)(z, n). \end{aligned} \quad (37)$$

⁵²So the probability measure of S_n in the cross-section would converge in distribution to a nondegenerate limit.

Differentiate (37) with respect to n to obtain a HJB for the marginal surplus,

$$\begin{aligned} \rho S_n(z, n) &= y_n(z, n) - b + \left(q\phi S_n(z, n) + q(1 - \phi) \int_0^{S_n(z, n)} H_n(s) ds \right)^{\frac{1}{\gamma}} \\ &\quad \times (q\phi + q(1 - \phi)H_n(S_n(z, n)))S_{nn}(z, n) \\ &\quad + (\mathbb{L}S_n)(z, n), \end{aligned} \quad (38)$$

where the size process associated with marginal surplus dynamics is

$$\begin{aligned} \hat{g}(z, n) &= \left(q\phi S_n(z, n) + q(1 - \phi) \int_0^{S_n(z, n)} H_n(s) ds \right)^{\frac{1}{\gamma}} \\ &\quad \times (q\phi + q(1 - \phi)H_n(S_n(z, n))). \end{aligned} \quad (39)$$

Note that $\hat{g}(z, n) \geq 0$, with equality if and only if $S_n(z, n) = 0$.

The HJB (38) therefore rewrites as

$$\rho S_n(z, n) = y_n(z, n) - b + \hat{g}(z, n)S_{nn}(z, n) + (\mathbb{L}S_n)(z, n). \quad (40)$$

Using the Feynman–Kac formula, we obtain a sequential representation of (40):

$$S_n(z, n) = \mathbb{E}_0 \left[\int_0^\infty e^{-\rho t} (y_n(z_t, \hat{n}_t) - b) dt \mid z_0 = z, \hat{n}_0 = n \right], \quad (41)$$

where z_t follows the actual productivity process, and \hat{n}_t follows (39).

We are now ready for the second step of the proof. Suppose for a contradiction that $S_n^* > 0$. Then $\hat{g}(z, n^*(z)) = +\infty$ because the first parenthesis in (39) is strictly positive ($q\phi \rightarrow \frac{\Delta}{\theta\xi} \in (0, +\infty)$), and the second parenthesis in (39) is infinite since $H_n(S_n^*) = 1$ and $q \rightarrow +\infty$. Similarly, for any $n \geq n^*(z)$ such that $S_n(z, n) > 0$, $\hat{g}(z, n) = +\infty$. Together, these observations imply that $\hat{n}_t = +\infty$ for any $t > 0$, starting from $n^*(z)$ at $t = 0$. Intuitively, if the marginal surplus from hiring were to remain always strictly positive, given the infinite meeting rate firms would keep growing without bound, a contradiction.

Under our Inada condition, we obtain, $y_n(z_t, \hat{n}_t) - b = -b < 0$ for any $t > 0$. Using (41), we obtain $S_n(z, n^*(z)) < 0$ a contradiction. Thus, it must be that $S_n^* = 0$.

Optimal Size. Our goal is now to characterize the optimal size $n^*(z)$. Our strategy is to leverage that the marginal surplus is zero $S_n^* = 0$. We connect the marginal surplus to the static net marginal product of labor $y_n - b$ using the sequential representation (41). This expression relates the marginal surplus to the present discounted value of all future net marginal products. To operationalize this idea, we split the time integral in (41) in several components: a component that captures the immediate future, and a continuation value. We define our candidate optimal size, $\bar{n}(z)$, to be such that $y_n(z, \bar{n}(z)) = b$. To show that $n^*(z) = \bar{n}(z)$, we proceed by contradiction.

Suppose first that $n^*(z) < \bar{n}(z)$, and so $y_n(z, n^*(z)) - b > 0$. Let $\epsilon > 0$, and rewrite (41) as

$$\begin{aligned} S_n(z, n^*(z)) &= \mathbb{E}_0 \left[\int_0^\epsilon e^{-\rho t} (y_n(z_t, \hat{n}_t) - b) dt \mid z_0 = z, \hat{n}_0 = n^*(z) \right] \\ &\quad + \mathbb{E}_0 [e^{-\rho\epsilon} S_n(z_\epsilon, \hat{n}_\epsilon) \mid z_0 = z, \hat{n}_0 = n^*(z)]. \end{aligned} \quad (42)$$

The first component is just the integral of the marginal product in a small time interval $[0, \epsilon]$. Recall that the marginal product is positive at $t = 0$ by assumption. Thus, by continuity, it must be that the marginal product $y_n(z_t, \hat{n}_t) - b$ is positive for all $t \leq \epsilon$ when ϵ is small enough. Therefore, the first component of (42) is strictly positive. The second component is also always positive because $S_n \geq 0$. Thus, (42) implies $S_n^* = S_n(z, n^*(z)) > 0$, a contradiction with $S_n^* = 0$. Therefore, we obtain that $n^*(z) \geq \bar{n}(z)$.

Suppose next for a contradiction that $n^*(z) > \bar{n}(z)$, and so $y_n(z, n^*(z)) - b < 0$. Our strategy for this inequality mirrors our previous one. We must however split the integral into three components rather than two to deal with continuation values.

Set $\hat{n}_0 = n^*(z)$, and let $\epsilon > 0$ be small enough. We define the stopping time

$$T = \inf\{t \geq \epsilon : \hat{n}_t = n^*(z)\}.$$

Recall that $\hat{g} \geq 0$. Therefore, $\hat{n}_t \geq n^*(z) > \bar{n}(z)$ for any $t > 0$. In addition, by definition of the stopping time T , we also have that $\hat{n}_t > n^*(z) > \bar{n}(z)$ for all $\epsilon \leq t \leq T$.

Then return to (41) evaluated at $(z, n^*(z))$. Write

$$\begin{aligned} S_n(z, n^*(z)) &= \mathbb{E}_0 \left[\int_0^\epsilon e^{-\rho t} (y_n(z_t, \hat{n}_t) - b) dt \middle| z_0 = z, \hat{n}_0 = n^*(z) \right] \\ &\quad + \mathbb{E}_0 \left[\int_\epsilon^T e^{-\rho t} (y_n(z_t, \hat{n}_t) - b) dt \middle| z_0 = z, \hat{n}_0 = n^*(z) \right] \\ &\quad + \mathbb{E}_0 [e^{-\rho T} S_n(z_T, n^*(z_T)) | z_0 = z, \hat{n}_0 = n^*(z)]. \end{aligned}$$

Similar to our previous argument, the first component is strictly negative when ϵ is small enough, by continuity. The second component is strictly negative because, by definition of T , $\hat{n}_t > n^*(z) > \bar{n}(z)$ for all $\epsilon \leq t \leq T$. The third component is zero since $S_n(z_T, n^*(z_T)) = S_n^* = 0$. Therefore, we obtain that $0 = S_n^* < 0$, a contradiction.

We conclude that $n^*(z) = \bar{n}(z) = \arg \max_n y(z, n) - bn$.

Limiting Value Function. Return to the surplus equation (37). Evaluating at $(z, n^*(z))$, the vacancy return component is equal to zero because $S_n^* = 0$. Therefore,

$$\begin{aligned} \rho S(z, n^*(z)) &= y(z, n^*(z)) - n^*(z)b + \mu(z)S_z(z, n^*(z)) + \frac{\sigma(z)^2}{2}S_{zz}(z, n^*(z)) \\ \text{s.t. } S(z, n^*(z)) &\geq \vartheta. \end{aligned}$$

To arrive at our representation in the main text, we must show that the partial derivatives, for example, $\frac{\partial S}{\partial z}(z, n^*(z))$, are equal to the total derivatives $\frac{dS(z, n^*(z))}{dz}$. To that end, note that, in the limit

$$\begin{aligned} \frac{dS(z, n^*(z))}{dz} &= \frac{\partial S}{\partial z}(z, n^*(z)) + \frac{\partial S}{\partial n}(z, n^*(z)) \frac{dn^*(z)}{dz} \\ &= \frac{\partial S}{\partial z}(z, n^*(z)) + S_n^* \frac{dn^*(z)}{dz}. \end{aligned}$$

Because $S_n^* = 0$, we obtain

$$\frac{dS(z, n^*(z))}{dz} = \frac{\partial S}{\partial z}(z, n^*(z)).$$

Therefore, in the limit, with a slight abuse of notation, the surplus can be described by the value function evaluated on the strip, $S(z) := S(z, n^*(z))$ which evolves according to

$$\rho S(z) = y(z, n^*(z)) - n^*(z)b + \mu(z)S_z(z) + \frac{\sigma(z)^2}{2}S_{zz}(z)$$

and an exit cut-off determined by $S(z) = \vartheta$. This proves equation (10) in the main text.

Finally, the free-entry condition $\int S(z, n_0)\Pi_0(z) dz = c_0$ pins down the mass of firms m_0 , and thus market tightness θ along the limit.

APPENDIX E: IDENTIFICATION

To illustrate the identification of the model's parameters more formally, we conduct two exercises. First, we show how the minimum distance of the objective function changes as we move each parameter ψ_i in steps in a wide range around ψ_i^* , letting the other parameters ψ_{-i}^* adjust to minimize the distance criterion function. We argue that the model is identified if $\mathcal{G}(\psi_i, \psi_{-i}^*)$ plotted as a function of ψ_i , traces a steep “U” with a minimum at ψ_i^* . Figure E.1 plots this exercise and gives us confidence that our parameter vector is well identified.

Second, in the main text we discussed how each parameter is especially informed by a particular moment, despite the model being jointly identified. To support this argument, Figure E.2 plots each of the eight moments as a function of the corresponding parameter in panel C of Table 1, keeping all other parameters at their estimated values. All panels in the figure show significant variation in the moment of interest as a function of its respective parameter.

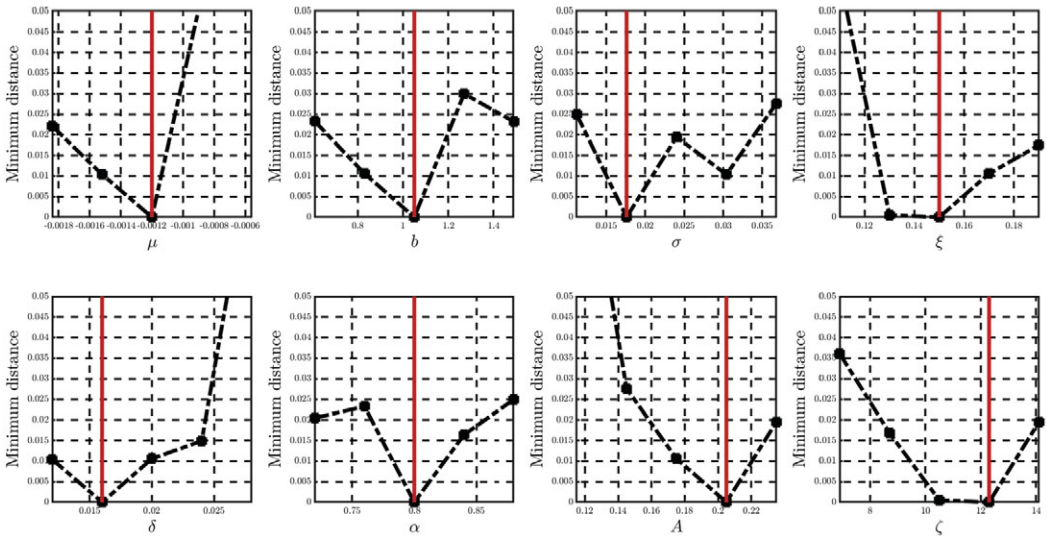


FIGURE E.1.—Minimum distance as function of each parameter. *Notes:* For each parameter $\psi_i \in \{\mu, \dots, b\}$, the line plots the minimum distance function $\psi_i \mapsto \mathcal{G}(\psi_i, \psi_{-i}^*)$, where ψ_{-i}^* adjusts to minimize the distance criterion. The vertical line marks the estimated value ψ_i^* listed in Table 1.

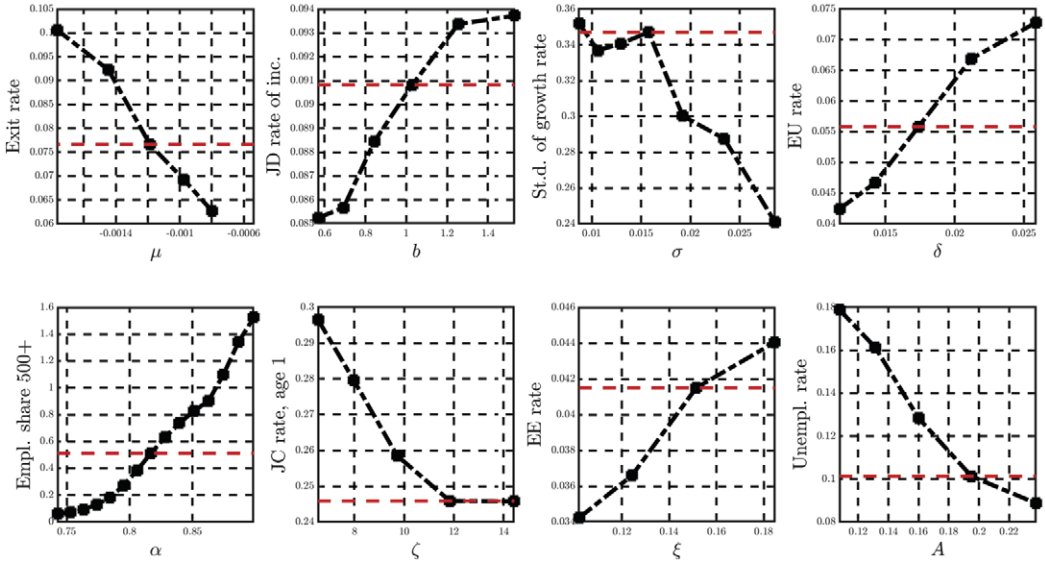


FIGURE E.2.—Each targeted moment against each parameter. *Notes:* This figure plots the relationship between each parameter $\psi_i \in \{\mu, \dots, b\}$ and the moment aligned with the parameter in Table 1. For each panel, the x -axis plots alternative values of the parameter. The y -axis plots the change in the corresponding moment in the steady state of the model obtained when all other parameters are as in Table 1.

REFERENCES

BREKKE, KJELL, AND BERNT OKSENDAL (1990): “The High Contact Principle as a Sufficiency Condition for Optimal Stopping,” Preprint series: Pure mathematics. [10]
 PESKIR, GORAN, AND ALBERT SHIRYAEV (2006): *Optimal Stopping and Free Boundary Problems*. Basel: Birkhauser. [10]
 PHAM HUYÈN (2009): *Continuous-Time Stochastic Control and Optimization With Financial Applications*. Berlin, Heidelberg: Springer. [10,11]
 POSTEL-VINAY, FABIEN, AND HÉLÈNE TURON (2010): “On-the-Job Search, Productivity Shocks, and the Individual Earnings Process,” *International Economic Review*, 51 (3), 599–629. [9]
 STOKEY, NANCY (2009): *The Economics of Inaction: Stochastic Control Models With Fixed Costs*. Princeton, NJ: Princeton University Press. [10]

Co-editor Charles I. Jones handled this manuscript.

Manuscript received 20 December, 2019; final version accepted 4 December, 2021; available online 25 January, 2022.