The Aggregate Effects of Labor Market Frictions

Michael W. L. Elsby   Ryan Michaels   David Ratner[1]

June 27, 2016

Abstract

Labor market frictions distort the path of aggregate employment by impeding the flow of labor across firms. For a canonical class of frictions, we show how observable measures of such flows can be used to formulate a diagnostic for the effect of frictions on aggregate employment dynamics. Application of this diagnostic to establishment microdata for the United States suggests that canonical frictions induce only a modest distortion to the path of aggregate employment relative to its frictionless counterpart: the sluggish dynamics of employment would have looked similar even in the absence of such frictions.

Preliminary and incomplete

JEL codes: E32, J63, J64.

Keywords: Adjustment costs, search frictions, firm dynamics.

[1] Elsby: University of Edinburgh (mike.elsby@ed.ac.uk). Michaels: Federal Reserve Bank of Philadelphia (ryan.michaels@phil.frb.org). Ratner: Federal Reserve Board (david.d.ratner@frb.gov).

We thank seminar participants at numerous institutions for helpful comments. All errors are our own. This research was conducted with restricted access to Bureau of Labor Statistics (BLS) data. The views expressed here do not necessarily reflect the views of the BLS, the Federal Reserve Bank of Philadelphia, the staff and members of the Federal Reserve Board, or the Federal Reserve System as a whole. Elsby and Michaels gratefully acknowledge financial support from the UK Economic and Social Research Council (ESRC), Award reference ES/L009633/1.

What are the effects of labor market frictions on aggregate employment dynamics?

In this paper, we present a new approach to this question for a canonical class of employment adjustment frictions. A compelling feature of this class is its ability to capture a key stylized fact of establishment dynamics, namely the empirical prevalence of inaction in employment adjustment. Notable examples include standard models of fixed adjustment costs that induce intermittent, discrete adjustments (Caballero, Engel and Haltiwanger 1997); per-worker hiring and firing costs that induce further distortions to the magnitude of adjustments (Bentolila and Bertola 1990; Hopenhayn and Rogerson 1993); and search and matching frictions (Pissarides 1985; Mortensen and Pissarides 1994). This class thus encompasses a large body of modern research on aggregate labor markets.

Our point of departure in section 1 is the observation that the mechanism through which, in theory, canonical frictions propagate the path of aggregate employment is by impeding the flow of labor across the firm-size distribution. Intermittent adjustment implies that only a fraction of desired, frictionless adjustments are implemented. Consequently, flows of labor to and from each firm size are retarded relative to an economy without frictions. In addition, distortions to the magnitude of adjustments induced by per-worker or search frictions further divert inflows away from their frictionless destination. By obstructing these firm-size flows, labor market frictions distort aggregate employment, since the latter is reflected by the mean of the firm-size distribution.

Establishment panel data allow several components of the forgoing cross-sectional employment flows to be measured; specifically total inflows to, and the probability of outflow from, each employment level in the firm-size distribution. In turn, these allow measurement of a notional firm-size distribution that is associated with flow balance, defined as the distribution that equates inflows to outflows at each employment level. A key contribution of this paper is to show how the latter can be used to infer a diagnostic for the aggregate effects of a prominent class of frictions.

A robust implication of canonical models is that the firm-size distribution implied by flow balance, and specifically its mean, exhibit an overshooting property relative to their frictionless counterparts. Aggregate flow-balance employment is predicted to rise more than frictionless employment in aggregate expansions, and decline more in recessions. Accordingly, since actual employment will move less than its frictionless counterpart, we can place a bound on the effect of canonical frictions by comparing the paths of flow-balance and actual employment.

This overshooting property is shaped by two economic forces: a partial equilibrium effect that holds in the absence of adjustment of wages; and a further equilibrium effect induced by such wage adjustment. In partial equilibrium, we show that the firm-size distribution associated with flow balance captures, first, a rightward shift in the distribution of desired, frictionless employment. In addition, it reflects an increased propensity of firms to adjust to versus from high employment levels; the elasticity of these cross-sectional flows is a critical component of the model's dynamics. Consequently, mean employment responds at least as much as its frictionless counterpart to aggregate shocks.

Equilibrium wage adjustment reinforces this property. To the extent that labor market frictions attenuate the response of labor demand to aggregate labor productivity, the associated response of equilibrium wages also will be attenuated relative to the frictionless case, amplifying the equilibrium employment response under flow balance.

Our analysis clarifies, and provides a testable implication for, the propagation mechanism embodied in this class of models. Frictions in this class may induce a sluggish response of equilibrium employment, but only by virtue of their ability to restrain the flow of labor across firms. However, while frictions dampen the level of the cross sectional flows, these flows are predicted to be highly elastic to aggregate shocks. A consequence is that employment under flow balance responds to such shocks even more aggressively than its frictionless counterpart. The question, then, is whether the data support such a stark contrast between the dynamics of (actual) equilibrium employment and employment under flow balance.

In section 2, we show how to test the model's implications on rich establishment microdata and provide several assessments of the empirical relevance of the propagation mechanism in this class of models. The results provide little empirical support for the model's proposed propagation mechanism. As a result, the data suggest that these frictions can account for only a modest distortion to the path of aggregate employment.

Our data are derived from the U.S. Quarterly Census of Employment and Wages for the period 1992Q1 through to 2014Q2. These data enable us to observe the outflows from, and inflows to, each position in the firm size distribution. Accordingly, we can derive a measure of aggregate employment implied by flow balance and compare its path to actual, observed employment.

A striking feature of these results, and the basis of our main conclusion, is that the empirical time series for our diagnostic tracks closely the dynamics of actual aggregate employment. Aside from a brief period around the trough of the Great Recession, the two series are remarkably similar: The median (mean) absolute deviation between the two series is just 0.5 (0.8) of a log point over the sample period. When viewed through the lens of canonical models, our diagnostic provides a bound for the path of frictionless employment. The inference is therefore that the aggregate employment effects of this class of frictions are modest—the path of aggregate employment would have looked similar even in the absence of these frictions. A corollary of this result is that the sources of propagation of aggregate employment dynamics captured by popular models of labor market frictions appear not to be a quantitatively important force in the data.

In the remainder of the paper, we inquire into the sources of this failure of canonical models. We suggest that a critical feature of this class of frictions is that, although they diminish the average levels of cross-sectional flows of labor across firms, these flows are nonetheless very responsive to aggregate shocks. The latter of course underlies the overshooting property of these models that we highlight. But we also find that it is counterfactual. We show that the empirical dynamics of these flows are both more sluggish, and an order of magnitude less volatile than implied by canonical models. Accordingly, we offer suggestions for future research that explores the potential role of these sluggish cross-sectional flows in an account of persistent aggregate employment dynamics.

1. Labor market frictions and firm size dynamics

In this section we describe the economic channels through which labor market frictions shape aggregate employment in a canonical class of models. A key observation is that these frictions affect aggregate employment by impeding the flow of firms across different firm sizes. We show that this class models in turn implies a specific structure to these firm size dynamics. We use this structure to motivate a diagnostic for the aggregate effects of these frictions. A virtue of this diagnostic that we take up in later sections of this paper is that it can be measured directly from establishment microdata.

1.1 Fixed costs

A leading model of labor market frictions postulates the presence of a fixed cost of adjusting employment, independent of the scale of adjustment. The early work of Hamermesh (1989) suggested that such a friction could account for important features of establishment employment dynamics, an observation that informed the later influential empirical analyses of Caballero and Engel (1993) and Caballero, Engel and Haltiwanger (1997).[2] In what follows we review the well-understood distortions of firms' labor demand policies induced by this friction. More importantly for our purposes, we use this to infer the implications for firm size flows, and thereby aggregate employment.

With regard to the structure of labor demand, the key implication of a fixed cost is that employment will be adjusted only intermittently and, upon adjustment, discretely—adjustment will be "lumpy." Thus, labor demand takes the form of a threshold "*Ss*" policy, illustrated in Figure 1A:

$$n = \begin{cases} n^* & \text{if} \quad n^* > U(n_{-1}), \\ n_{-1} & \text{if} \quad n^* \in [L(n_{-1}), U(n_{-1})], \\ n^* & \text{if} \quad n^* < L(n_{-1}). \end{cases} \tag{1}$$

Here n^* is the level of employment that a firm chooses if it adjusts. Under the *Ss* policy, a firm's current employment n is adjusted away from its past level n_{-1} whenever n^* deviates sufficiently from n_{-1}, as dictated by the adjustment triggers $L(n_{-1}) < n_{-1} < U(n_{-1})$.

[2] See also King and Thomas (2006), Cooper, Haltiwanger and Willis (2007, 2015), and Bachmann (2012).

Caballero, Engel and Haltiwanger (1995, 1997) refer to n^* as *mandated* employment, interpreted as the level of employment the firm would choose if the friction were suspended for the current period. In principle, the latter is distinct from *frictionless* employment, which emerges if the fixed cost is suspended *indefinitely*. For reasonably calibrated models within this canonical class, however, the dynamics of mandated and frictionless employment are very similar.[3] Henceforth, then, we shall refer to n^* as frictionless, or desired, employment.

The dynamics of *aggregate* employment implied by the firm behavior in equation (1) can be inferred from its implications for firm size flows. Imagine the economy enters the period with a density of past employment, $h_{-1}(\cdot)$, and that realizations of idiosyncratic and aggregate shocks induce a density of desired employment $h^*(\cdot)$. Our strategy is to infer a law of motion for the current-period density $h(\cdot)$ implied by equation (1). This in turn will imply a path for aggregate employment in the economy, which we denote by N, since the latter is captured by the mean of the density, $N \equiv \int mh(m)dm$.

The adjustment policy in Figure 1A suggests a straightforward approach to constructing a law of motion for the firm-size density $h(\cdot)$. Consider first the outflow of mass from some employment level m. Among the $h_{-1}(m)$ mass of firms that enter the period with m workers, only the fraction whose desired employment n^* lies outside the inaction region $[L(m), U(m)]$ will choose to incur the adjustment cost and leave the mass. Symmetrically, now consider the inflow of mass to employment level m. Among the $h^*(m)$ mass of firms whose desired employment is equal to m, only the fraction whose inherited employment n_{-1} lies outside of the inverse inaction region $[U^{-1}(m), L^{-1}(m)]$ will choose to incur the adjustment cost and flow to m. Thus, the change in the mass at employment level m follows the law of motion

$$\Delta h(m) = \tau(m)h^*(m) - \phi(m)h_{-1}(m), \tag{2}$$

where $\tau(m)$ and $\phi(m)$ are respectively the probabilities of adjusting to and from an employment level m,

$$\begin{aligned} \tau(m) &= \Pr(n_{-1} \notin [U^{-1}(m), L^{-1}(m)] | n^* = m)\text{, and} \\ \phi(m) &= \Pr(n^* \notin [L(m), U(m)] | n_{-1} = m). \end{aligned} \tag{3}$$

Formal derivations of equations (2) and (3) are provided in the Appendix.

1.2 An empirical diagnostic

[3] This has been proven analytically for the case of a plausibly small fixed adjustment cost (Gertler and Leahy, 2008; Elsby and Michaels, 2014). In the Appendix, we also verify numerically that the distinction between frictionless and mandated employment is quantitatively inconsequential for the results we report below.

With this theoretical law of motion in hand, our next step is to consider which of its components can be *measured* empirically using available data. As we shall see, establishment-level panel data allow one to observe much of equation (2): One can measure the mass at each employment level at each point in time, $h_{-1}(m)$ and $h(m)$; one can also observe the fraction of establishments at each employment level that adjusts away, $\phi(m)$. Crucially, however, we observe only the *total* inflow, $\tau(m)h^*(m)$, rather than its constituent parts. Of course, this is precisely the identification problem that we are attempting to surmount: If we could measure both $\tau(m)$ and $h^*(m)$, the latter would allow us to infer a measure of aggregate *frictionless* employment $N^* \equiv \int mh^*(m)dm$. Comparison of N^* with the observed path of actual aggregate employment N would then indicate the wedge between these two induced by the adjustment friction.

Our approach is instead to ask whether we can infer useful information about the path of desired aggregate employment N^* without observing directly the density of frictionless employment $h^*(m)$. Our point of departure is to note that, for fixed adjustment rates $\tau(m)$ and $\phi(m)$, the firm size density will converge to a position where the inflow of mass to each m is balanced by outflows from that point. This *flow balance* condition implies a density

$$\hat{h}(m) \equiv \frac{\tau(m)}{\phi(m)} h^*(m). \tag{4}$$

$\hat{h}(m)$ is useful for several reasons. First, it can be measured straightforwardly, since it requires knowledge only of the total inflow, $\tau(m)h^*(m)$, and the probability of outflow $\phi(m)$, both of which are observed in establishment panel data.

Second, we argue in what follows that flow balance conveys important information on the path of frictionless employment, and thereby on the role of adjustment frictions in the dynamics of aggregate employment. Specifically, note that the aggregate employment level implied by flow balance, $\hat{N} \equiv \int m\hat{h}(m)dm$, can be written as

$$\hat{N} = N^* + cov_{h^*}\left(m, \frac{\tau(m)}{\phi(m)}\right), \tag{5}$$

where cov_{h^*} denotes a covariance taken with respect to the distribution of frictionless employment, $h^*(m)$.

Equation (5) reveals that aggregate employment under flow balance $\hat{N}$ will *overshoot* the path of aggregate frictionless employment N^* under a monotonicity condition—namely that firms on average are more likely to adjust to versus from high (low) employment levels following positive (negative) innovations to aggregate frictionless employment. This implies that, after a positive innovation, $\tau(m)/\phi(m)$ will decline for low m (since fewer firms adjust to versus from low m) and rise for high m (since more firms adjust to versus from high m). Thus, $\tau(m)/\phi(m)$ "tilts up" with respect to m, raising the covariance term in (5). Under this condition, $\widehat{N}$ will rise more than N^* when the latter rises, and fall more than N^* when it falls.

The monotonicity condition that underlies this intuition is closely related to the *selection effect* that has been emphasized in the literature on adjustment frictions (Caballero and Engel 2007; Golosov and Lucas 2007). This refers to a property shared by state-dependent models of adjustment whereby the firms that adjust tend to be those with the greatest desired adjustment. By the same token, firms in these models also will adjust in the *direction* of the desired adjustment.

The forgoing intuition can be formalized tractably in standard models of fixed adjustment frictions, such as that set out in Caballero and Engel (1999). In this environment, firms face an isoelastic production function $y = pxn^\alpha$ that is subject to idiosyncratic shocks x. Firms thus face the following decision problem

$$\Pi(n_{-1}, x) \equiv \max_n \{pxn^\alpha - wn - C^+\mathbb{I}[n > n_{-1}] - C^-\mathbb{I}[n < n_{-1}] + \beta\mathbb{E}[\Pi(n, x')|x]\}, \quad (6)$$

where p denotes (for now, fixed) aggregate productivity, w the wage, and $C^{+/-}$ the fixed costs of adjusting employment up and down.

Caballero and Engel show that, if idiosyncratic shocks follow a geometric random walk, $\ln x' = \ln x + \varepsilon'_x$, and the adjustment costs $C^{+/-}$ are scaled to be proportional to the firm's frictionless labor costs, the labor demand problem has a tractable homogeneity property. This has two useful implications: First, the adjustment triggers in (1) are linear and time invariant, $L(n_{-1}) = L \cdot n_{-1}$ and $U(n_{-1}) = U \cdot n_{-1}$ for constants $L < 1 < U$. Second, desired (log) employment adjustments, $\ln(n^*/n_{-1})$, are independent of initial firm size n_{-1}.[4]

Proposition 1 uses these properties of the canonical model to formalize the heuristic claim above that changes in aggregate employment under flow balance overshoot changes in aggregate frictionless employment. It assumes firms perceive aggregate productivity p as fixed, and characterizes comparative statics with respect to a (one-time) change in p. Because of the model's loglinear structure, the result is most simply derived in terms of aggregate *log* frictionless employment, which we shall denote by $\mathcal{N}^*$, and its counterpart under flow balance, $\widehat{\mathcal{N}}$.

Proposition 1 *Consider the model of fixed adjustment costs (6). Relative to a prior constant-$\mathcal{N}^*$ steady state, a small change in aggregate log frictionless employment, $\Delta\mathcal{N}^*$, induces on impact the following change in aggregate log employment under flow balance,*

$$\Delta\widehat{\mathcal{N}} \approx \frac{1 - \epsilon_w}{1 - \epsilon_{w^*}} \cdot (1 + \psi) \cdot \Delta\mathcal{N}^*, \quad (7)$$

[4] The Appendix provides a formal statement and proof of this result in Lemma 1.

where $\psi > 0$*, and* ϵ_w *and* ϵ_{w^*} *are the elasticities, respectively with and without frictions, of equilibrium wages to aggregate productivity* p*.*

In Proposition 1, the response of $\widehat{\mathcal{N}}$ overshoots the frictionless response of $\mathcal{N}^*$ for two reasons. The first is a partial equilibrium response that emerges for fixed wages. If $\epsilon_w = \epsilon_{w^*}$, the presence of the $(1+\psi) > 1$ term ensures that the change in aggregate log employment under flow balance strictly overshoots its frictionless counterpart. This is the formalization of the intuition surrounding equation (5) that increases in desired employment $\mathcal{N}^*$ are augmented in $\widehat{\mathcal{N}}$ by increases in the propensity to adjust toward higher employment levels.

In addition, however, Proposition 1 reveals how differential equilibrium wage responses reinforce this overshooting property still further. To the extent that adjustment frictions restrict the response of labor demand to an aggregate shock, they also will restrict the response of equilibrium wages for a given labor supply schedule, $\epsilon_w < \epsilon_{w^*}$. It follows that $(1-\epsilon_w)/(1-\epsilon_{w^*}) > 1$, further amplifying the equilibrium employment response under flow balance.

While Proposition 1 has a number of virtues—it holds irrespective of whether adjustment is symmetric ($C^+ = C^-$) or asymmetric ($C^+ \neq C^-$), for example—it also has limitations. It relies on the homogeneity of the canonical model implied by the random walk assumption on shocks. It is also a comparative statics result, describing the response of the economy to a change in aggregate labor demand, indexed by p, that is expected to occur with zero probability from the firms' perspectives. For these reasons, in the next subsection, we explore the robustness of the overshooting result in numerical simulations that relax these assumptions.

1.3 Quantitative illustrations

We illustrate the dynamics of fixed costs models that resemble the canonical model described above, but with two differences. First, we relax the random walk assumption on idiosyncratic shocks, which we allow to follow a geometric AR(1),

$$\ln x' = \rho_x \ln x + \varepsilon'_x, \text{ where } \varepsilon'_x \sim N(0, \sigma_x^2). \tag{8}$$

Second, we allow for the presence of aggregate productivity shocks, and for their stochastic process to be known to firms in the model. The evolution of these aggregate shocks also is assumed to follow a geometric AR(1),

$$\ln p' = \rho_p \ln p + \varepsilon'_p, \text{ where } \varepsilon'_p \sim N\big(0, \sigma_p^2\big). \tag{9}$$

To mirror the timing of the data we use later in the paper, a period is taken to be one quarter. Based on this, we set the discount factor β to 0.99, consistent with an annual interest rate of around 4 percent. To parameterize the remainder of the model, we appeal to the empirical literature that estimates closely related models of firm dynamics.

The returns to scale parameter α is set to 0.64, as in the estimates of Cooper, Haltiwanger, and Willis (2007, 2015).

The choice of parameters of the idiosyncratic productivity shock process (8) is informed by the estimates of Abraham and White (2006). They estimate a quarterly persistence parameter ρ_x of approximately 0.7, which we implement. Our choice of the standard deviation of the idiosyncratic innovation ε_x' of around $\sigma_x = 0.15$ is set a little higher than Abraham and White's estimate of 0.10, since the latter lies at the lower end of the range of estimates in the literature.[5]

The parameters of the process for aggregate technology in (9) are chosen so that aggregate frictionless employment in the model exhibits a persistence and volatility comparable to aggregate employment in U.S. data. This yields $\rho_p = 0.95$ and $\sigma_p = 0.0026$. Although frictions augment persistence, and dampen volatility, the intent is for the model environment to resemble broadly the U.S. labor market with respect to these *un*conditional moments. Importantly, the approach does not build in any persistence in employment *conditional* on technology.

Finally, with respect to the adjustment cost, here we report results for the case of symmetric frictions, $C^+ = C^-$; results for asymmetric costs are provided in the Appendix. We explore three parameterizations that successively raise the friction. The first is chosen to replicate the observed fraction of firms that do not adjust from quarter to quarter. In the data used later in the paper, this inaction rate averages 52.5 percent. However, the latter inaction rate is implied by a fixed cost equal to 1.3 percent of quarterly revenue, which is at the lower end of available estimates (Bloom 2009; Cooper, Haltiwanger, and Willis 2007, 2015). For this reason, we also explore fixed costs that imply quarterly inaction rates of 67 percent and 80 percent, corresponding to adjustment costs of 2.7 percent and 5.8 percent of quarterly revenue, respectively.

We solve the labor demand problem for 250,000 firms via value function iteration on an integer-valued employment grid, $n \in \{1,2,3 \ldots\}$. The latter mirrors the integer constraint in the data, allowing one to construct the density $\hat{h}(\cdot)$ in the simulated data in the same way as we later implement in the real data.

To simulate equilibrium wage responses, we impose an aggregate labor supply schedule that is assumed to prevail in the model with and without the adjustment friction. Based on the estimates of Chetty (2012) and Chetty et al. (2012), we parameterize the labor supply function to have a (constant) Frisch elasticity of 0.5.[6] To solve the model, we implement the bounded rationality algorithm of Krusell and Smith (1998), whereby

[5] The Appendix derives these parameters and contrasts them with other estimates reported in related literature. It also shows that the implied dynamics of aggregate employment implied by reasonable changes in these parameters are qualitatively similar to those described here.

[6] Using survey questions about the long-run response to hypothetical wealth windfalls, Kimball and Shapiro (2010) estimate a median Frisch elasticity of 0.6 and a mean of 1. Consistent with Proposition 1, we have verified that aggregate employment under flow balance overshoots its frictionless counterpart even in the latter parameterization. Results are available on request.

firms condition their labor demands on a linear forecast rule that relates the log aggregate employment to its lag and aggregate productivity. We then iterate on the coefficients of this forecast rule until the firms' simulated choices are consistent with the rule.[7]

Figure 2 plots simulated impulse responses of aggregate employment N, together with its frictionless and flow-balance counterparts, N^* and $\widehat{N}$ respectively. The overshooting result anticipated in Proposition 1 is clearly visible in the model dynamics. For all three parameterizations of the adjustment cost our proposed diagnostic, $\widehat{N}$, responds more aggressively to the aggregate shock than frictionless employment N^*. Moreover, the magnitude of the overshooting of $\widehat{N}$ relative to N^* is substantial in the model, responding on impact around twice as much to the impulse.

These results provide a first example of how observable measures of firm size flows can be used to infer properties of the path of aggregate employment in the absence of a canonical friction. The next subsections extend this insight to two other popular models of labor market frictions.

1.4 Linear costs

Prominent alternative models of labor market frictions appeal instead to linear costs of adjustment in which the friction is discrete *at the margin*, and rises with the scale of adjustment. This class encompasses models of per-worker hiring and firing costs, including the influential contributions of Oi (1962), Nickell (1978), Bentolila and Bertola (1990) and Hopenhayn and Rogerson (1993).

Relative to the fixed costs case examined above, linear frictions alter the structure of both labor demand and firm size dynamics. Although labor demand will continue to feature intermittent adjustment, a key difference is that, conditional on adjusting, firms will no longer discretely set employment to their frictionless target n^*. Rather, they will reduce the magnitude of hires and separations, shedding fewer workers when they shrink, and hiring fewer workers when they expand. Formally, the policy rule for separations, which we shall denote by $l(\cdot)$, will differ from the policy rule used for hiring, denoted $u(\cdot)$, inducing the *continuous Ss* policy illustrated in Figure 1B,

$$n = \begin{cases} u^{-1}(n^*) & \text{if} \quad n^* > u(n_{-1}), \\ n_{-1} & \text{if} \quad n^* \in [l(n_{-1}), u(n_{-1})], \\ l^{-1}(n^*) & \text{if} \quad n^* < l(n_{-1}), \end{cases} \tag{10}$$

where $l(n_{-1}) < n_{-1} < u(n_{-1})$ for all n_{-1}.

The key distinction, that the *direction* of adjustment must be taken into account in the presence of linear costs, also leaves its imprint on the law of motion for the firm size distribution. As before, the labor demand policy in Figure 1B motivates the form of this

[7] In the Appendix we also report results for the fixed-wage results. These are qualitatively consistent with those shown here.

law of motion. This reveals that the structure of outflows is qualitatively unchanged—of the $h_{-1}(m)$ density of firms currently at employment level m, only those with frictionless employment outside the inaction region $[l(m), u(m)]$ will adjust away. But inflows are now differentiated by the direction of adjustment. The inflow of mass adjusting down to m is comprised of firms whose past employment n_{-1} is greater than m, and whose mandated employment n^* is equal to $l(m) < m$. Likewise, the inflow of mass flowing up to m consists of firms with $n_{-1} < m$ and $n^* = u(m) > m$.

Piecing this logic together yields the following law of motion for the firm size density,

$$\Delta h(m) = \tau_l(m) h_l^*(m) + \tau_u(m) h_u^*(m) - \phi(m) h_{-1}(m). \tag{11}$$

Extending the interpretation of the fixed costs case above, here $h_l^*(m) = l'(m) h^*\big(l(m)\big)$ and $h_u^*(m) = u'(m) h^*\big(u(m)\big)$ are the densities of employment that would emerge if all firms adjusted, respectively, according to the separation rule, $l(m)$, and hiring rule, $u(m)$. However, only a fraction of firms will in fact adjust. The adjustment probabilities take the form

$$\begin{aligned} \tau_l(m) &= \Pr\big(n_{-1} > m | n^* = l(m)\big), \\ \tau_u(m) &= \Pr\big(n_{-1} < m | n^* = u(m)\big), \text{ and} \\ \phi(m) &= \Pr(n^* \notin [l(m), u(m)] | n_{-1} = m), \end{aligned} \tag{12}$$

where $\tau_l(m)$ is the probability that a firm adjusts *down to* m, while $\tau_u(m)$ is the probability that a firm adjusts *up to* m.

To construct the density under flow balance for the linear costs case note that, for fixed adjustment rates $\tau_l(m)$, $\tau_u(m)$ and $\phi(m)$, the law of motion (11) implies that the firm size density will converge to

$$\hat{h}(m) \equiv \frac{\tau_l(m)}{\phi(m)} h_l^*(m) + \frac{\tau_u(m)}{\phi(m)} h_u^*(m). \tag{13}$$

Once again, the behavior of aggregate employment under flow balance can be formalized most straightforwardly in a canonical linear cost model in which firms face isoelastic production $y = pxn^\alpha$, and idiosyncratic shocks that follow a geometric random walk. The key difference is that the adjustment friction is now scaled by the magnitude of adjustment, so that firms face the decision problem:[8]

$$\Pi(n_{-1}, x) \equiv \max_n \{pxn^\alpha - wn - c^+ \Delta n^+ + c^- \Delta n^- + \beta \mathbb{E}[\Pi(n, x') | x]\}. \tag{14}$$

A simple extension of Caballero and Engel's (1999) homogeneity results for the fixed cost model can be used to show that if idiosyncratic shocks follow a geometric random walk, and if per-worker hiring and firing costs are proportional to wages, the adjustment

[8] Nickell (1978, 1986) first formalized the linear-cost model in the context of a labor demand model. Bentolila and Bertola (1990) introduced uncertainty into Nickell's continuous-time formulation. Equation (14) is a discrete-time analogue to Bentolila and Bertola's model (although the shocks need not be Gaussian).

triggers in (10) are linear and time invariant, $l(n) = l \cdot n$ and $u(n) = u \cdot n$ for constants $l < 1 < u$, and that desired (log) employment adjustments, $\ln(n^*/n_{-1})$, are independent of initial firm size n_{-1}.[9]

As in Proposition 1 above for the fixed costs case, the latter properties allow one to relate the response of aggregate flow-balance log employment $\widehat{\mathcal{N}}$ to the response of aggregate frictionless log employment $\mathcal{N}^*$ following a change in aggregate productivity.

Proposition 2 *Consider the model of linear adjustment costs (14). Relative to a prior constant-$\mathcal{N}^*$ steady state, a small change in aggregate log frictionless employment, $\Delta\mathcal{N}^*$, induces on impact the following change in aggregate log employment under flow balance,*

$$\Delta\widehat{\mathcal{N}} \approx \frac{1-\epsilon_w}{1-\epsilon_{w^*}} \cdot \Delta\mathcal{N}^*, \tag{15}$$

where ϵ_w and ϵ_{w^*} are the elasticities, respectively with and without frictions, of equilibrium wages to aggregate productivity p.

Just as in the model of fixed costs, the response of $\widehat{\mathcal{N}}$ relative to $\mathcal{N}^*$ is shown to be mediated by the wage elasticities ϵ_w and ϵ_{w^*}, and is qualitatively independent of any asymmetries in the frictions $c^+ \neq c^-$. In contrast to the fixed costs case, though, the extent to which $\widehat{\mathcal{N}}$ overshoots the frictionless response of $\mathcal{N}^*$ now depends entirely on the response of equilibrium wages.

For fixed wages, the response of $\widehat{\mathcal{N}}$ no longer overshoots that of $\mathcal{N}^*$, but is approximately *equal* to it. The key difference is that firms adjust only *partially* toward their frictionless employment under linear frictions. A rise in $\mathcal{N}^*$ places more firms on the hiring margin, where employment is set below its frictionless counterpart, and fewer firms on the separation margin, where employment exceeds its frictionless level. Both forces serve to attenuate the response of $\widehat{\mathcal{N}}$ relative to the fixed costs case. Proposition 2 shows that, to a first order, this attenuation offsets exactly the partial equilibrium overshooting of the diagnostic $\widehat{\mathcal{N}}$ in the fixed costs case.

The effects of differential equilibrium wage responses remain as before, however. Sluggish frictional responses of labor demand to an aggregate shock will induce sluggish equilibrium wage responses under frictions, such that $\epsilon_w < \epsilon_{w^*}$. This again gives rise to overshooting, as shown in Proposition 2.

Figures 3 and 4 show that the result of Proposition 2 is mirrored in numerical simulations of models that allow the presence of more plausible idiosyncratic shocks x, and fully stochastic aggregate shocks p. These numerical results are based on the same methods and baseline parameterization described in section 1.3.

[9] Again, the Appendix provides a formal statement and proof of this result in Lemma 1.

Figure 3 illustrates impulse responses of actual, frictionless and flow-balance aggregate employment in the presence of *symmetric* linear frictions where $c^{+} = c^{-}$. As before, each panel of Figure 3 successively raises the friction to produce increasingly higher average rates of inaction in employment adjustment. Note that the response of actual employment becomes progressively more sluggish as the friction rises, which dampens the response of the wage. As foreshadowed by Proposition 2, the response of flow-balance employment therefore increasingly overshoots the frictionless path.

Figure 4 in turn reveals that this result is unimpaired by the presence of *asymmetric* frictions, as suggested by Proposition 2. Its first three panels report results for successively higher hiring costs, $c^{+} > 0$ and $c^{-} = 0$; the latter three panels do the same for firing costs, $c^{-} > 0$ and $c^{+} = 0$. Strikingly, it is hard to discern differences between the impulse responses for hiring and firing costs, and between these and the impulse response for the symmetric case in Figure 3.

The message of Figures 3 and 4, then, is that the insight of Proposition 2 is robust to empirically reasonable parameterizations of canonical models of linear frictions. This reinforces the message of section 1.3 that the flow balance is indeed a useful diagnostic for the path of frictionless employment, and thereby for the aggregate effects of canonical frictions.

However, Proposition 2 does not allow the adjustment triggers to vary, since these are independent of $\Delta\mathcal{N}^*$ under the time-invariant linear frictions we have consider thus far. This is a key distinction with respect to models of search frictions, to which we now turn.

1.5 Search costs

The canonical Diamond-Mortensen-Pissarides (DMP) model of search frictions, in which a single firm matches with a single worker, can be extended to a setting with "large" firms that operate a decreasing-returns-to-scale production technology (Acemoglu and Hawkins 2014; Elsby and Michaels 2013). The presence of search frictions implies two modifications to the canonical linear cost model studied above.

First, search frictions induce a *time-varying* per-worker hiring cost. Hiring is mediated through vacancies, each of which is subject to a flow cost c, and is filled with a probability q that depends on the aggregate state of the labor market. Under a law of large numbers, the effective per-worker hiring cost is thus c/q, which varies over time with variation in the vacancy-filling rate q. The typical firm's problem therefore takes the form:

$$\Pi(n_{-1}, x) \equiv \max_{n} \left\{ pxn^{\alpha} - w(n, x)n - \frac{c}{q}\Delta n^{+} + \beta\mathbb{E}[\Pi(n, x')|x] \right\}. \tag{16}$$

Second, search frictions induce *ex post* rents to employment relationships over which a firm and its workers may bargain. In an extension of the bilateral Nash sharing rule invoked in standard one-worker-one-firm search models, Elsby and Michaels (2013) show that a marginal surplus-sharing rule proposed by Stole and Zwiebel (1996) implies a wage equation of the form

$$w(n,x) = \eta \frac{px\alpha n^{\alpha-1}}{1-\eta(1-\alpha)} + (1-\eta)\omega. \tag{17}$$

Here $\eta \in [0,1]$ indexes worker bargaining power, and ω is the annuitized value of workers' threat point. Bruegemann, Gautier and Menzio (2015) show that the marginal surplus-sharing rule underlying (17) can be derived from an alternating-offers bargaining game between a firm and its many workers in which the strategic position of each worker in the firm is symmetric.

As before, we consider first a version of the search model with a tractable homogeneity property. Specifically, we study the case in which the friction, embodied in the vacancy cost, is proportional to the workers' outside option, $c \propto \gamma\omega$.[10] Under these assumptions, the Appendix shows that the homogeneity properties used for the models discussed in previous subsections continue to hold, with one exception: although the adjustment triggers remain linear, they no longer are invariant to shifts in aggregate productivity, for the simple reason that the friction varies with the aggregate state.

Proposition 3 reveals that the result of Proposition 2 extends to search frictions, under a few restrictions.

Proposition 3 *Consider the model of search costs in (16) and (17). Assume (i) firms are patient, $\beta \approx 1$; (ii) frictions are small, $\gamma^2 \approx 0$; and (iii) the distribution of ε_x is symmetric. Then, relative to a prior constant-$\mathcal{N}^*$ steady state, a small change in aggregate log frictionless employment, $\Delta\mathcal{N}^*$, induces on impact the following change in aggregate log employment under flow balance,*

$$\Delta\widehat{\mathcal{N}} \approx \frac{1-\epsilon_\omega}{1-\epsilon_{w^*}}\Delta\mathcal{N}^*, \tag{18}$$

where ϵ_ω and ϵ_{w^*} are the elasticities of ω and frictionless wages w^* to aggregate productivity p.

As in earlier results, Proposition 3 suggests that the responses of $\widehat{\mathcal{N}}$ and $\mathcal{N}^*$ are shaped by both partial equilibrium and equilibrium forces, which we consider in turn.

In partial equilibrium, Proposition 3 shows that the response of aggregate employment under flow balance $\widehat{\mathcal{N}}$ still approximates the response of aggregate log frictionless

[10] This can be motivated through the presence of a dual labor market in which recruitment is performed by workers hired in a competitive market, who are paid according to the annuitized value of unemployment ω.

employment $\mathcal{N}^*$, but under a few additional restrictions. We argue in what follows that these restrictions are plausible.

The first two restrictions—that firms are patient, and that frictions are small—are quantitative. We address their plausibility by examining results from a numerical model that does not impose these restrictions. As above, this model sets the discount factor β to match an annual interest rate of 4 percent, and sets the vacancy cost c to replicate a set of average probabilities of employment adjustment. The numerical results will thus address the extent to which β is close enough to one, and the friction sufficiently small, for the insight of Proposition 3 to hold.

The third restriction concerns the symmetry of the distribution of idiosyncratic shocks. This can be justified along two grounds. First, it is conventional to implement shock processes with symmetrically distributed—typically Normal—innovations. Second, it is also consistent with the observed pattern of employment adjustment, which is close to symmetric (see Davis and Haltiwanger 1992, and Elsby and Michaels 2013, among others).

These three restrictions aid the proof of Proposition 3, which is based on symmetry. If the firm is sufficiently patient ($\beta \approx 1$), the cost of hiring in the current period implies an equal cost of firing in the subsequent period. As a result, one can show that the optimal policy is symmetric, to a first-order approximation around $\gamma = 0$, as long as the driving force ε_x is symmetric. In terms of the notation of the policy rules, this means the upper and lower adjustment triggers, $u(n_{-1}) = u \cdot n_{-1}$ and $l(n_{-1}) = l \cdot n_{-1}$, satisfy $\ln u \approx -\ln l$, and move by approximately the same amount in response to a shift in aggregate productivity.

The effects of equilibrium adjustment in wages also mirror earlier results, and depend on the flexibility of frictionless wages w^* relative to workers' outside option in the presence of frictions ω, summarized in the elasticities ϵ_{w^*} and ϵ_ω, respectively.

We parameterize these elasticities as follows. As before, in the frictionless case ϵ_{w^*} is related to the Frisch elasticity of labor supply, which we again set to 0.5. This implies $\epsilon_{w^*} = 2/(3-\alpha) \approx 0.848$ when α is set to equal 0.64.[11]

The counterpart to ϵ_{w^*} in the search model, ϵ_ω, depends on the structure of the worker's threat point ω. Elsby and Michaels (2013) show that a "large-firm" extension of the wage bargain implemented in the canonical search model implies

$$\omega = \frac{\eta}{1-\eta} c\theta + b, \tag{19}$$

where θ is labor market tightness, the ratio of aggregate vacancies to unemployment, and b is the flow payoff to unemployment. Thus, in this standard case, a key determinant of

[11] ϵ_{w^*} is not observed in a real world with search frictions. However, longer-run labor supply responses, e.g., Hicksian elasticities, are arguably less influenced by such frictions. Chetty (2012) argues that, under balanced growth preferences, estimated Hicksian elasticites imply a Frisch elasticity of no more than 0.5.

the persistence of wages, and thereby of labor demand, are the dynamics of the vacancy-unemployment ratio θ. Empirically, the latter are considerably sluggish. While carefully calibrated large-firm search models are able to generate some of this persistence in θ, they cannot account for its entirety.

To ensure that firms' labor demand choices reflect an empirically realistic process for ω, we proceed by treating the process for θ as a known function of aggregate state variables. Using observed time series, we estimate the least-squares regression of log tightness on log aggregate employment and log aggregate productivity. This estimated function, $\theta(N,p)$, is then inserted into ω, and is used to calculate the cost to hire c/q in (16). As is standard, we take q to be a log-linear function of θ with elasticity 0.6 (Petrongolo and Pissarides, 2001).[12]

It remains to choose worker bargaining power, η. We pin this down based on evidence from microdata on wages. Taking account of the shifting composition of employment over the business cycle, microdata-based estimates suggest that real wages are about as cyclical as employment (Solon, Barsky, and Parker 1994; Elsby, Solon, and Shin 2016). Accordingly, we set η to match an elasticity of average real wages with respect to aggregate employment approximately equal to one.

The implied magnitudes for ϵ_ω vary somewhat across the different parameterizations of the search friction. We find that ϵ_ω lies between 0.58 (when the search friction is set to induce an inaction rate of 52.5 percent per quarter) and 0.51 (when the search friction induces an inaction rate of 80 percent per quarter).

Proposition 3 implies that the response of aggregate employment under flow balance should overshoot that of frictionless employment under these parameterizations, since $(1-\epsilon_\omega)/(1-\epsilon_{w^*})$ falls between 2.76 and 3.22. Figure 5 shows that this prediction of Proposition 3 is visible in numerical simulations of the model based on the same methods and baseline parameterization described in section 1.3—that is, with stationary idiosyncratic shocks x, and fully stochastic aggregate shocks p. The impulse responses in Figure 5 suggest that aggregate employment under flow balance reacts on impact of the aggregate shock more than twofold its frictionless counterpart.

2. Empirical implementation

The previous section gave a theoretical rationale for an empirical diagnostic, flow-balance employment $\widehat{N}$, which provides a bound for the aggregate effects of a canonical class of labor market frictions. A key virtue of this diagnostic is that it can be measured with access to establishment panel data on employment. In this section, we apply these results to a rich source of microdata from the United States.

[12] The form of the forecast rule for log aggregate employment is the same as described in earlier sections of this paper, and Krusell and Smith's algorithm is used to solve the model. The Appendix reports coefficients of this forecast rule, as well as the estimates of $\theta(N,p)$ referenced in the main text.

2.1 Data

The data we use are taken from the Quarterly Census of Employment and Wages (QCEW). The QCEW is compiled by the Bureau of Labor Statistics (BLS) in concert with State Employment Security Agencies. The latter collect data from all employers in a state that are subject to the state's Unemployment Insurance (UI) laws. Firms file quarterly UI Contribution Reports to the state agency, which provide payroll counts of employment in each month. These are then aggregated by the BLS, which defines employment as the total number of workers on the establishment's payroll during the pay period that includes the 12^{th} day of each month. Following BLS procedure, we define quarterly employment as the level of employment in the third month of each quarter.[13]

From the cross-sectional QCEW data, the BLS constructs the Longitudinal Database of Establishments (LDE), which we use in what follows. Although data are available for the period 1990Q1 to 2014Q2, we restrict attention to data from 1992Q1 due to difficulty in matching establishments in the first two years of the sample.

Sample restrictions. The QCEW data are a near-complete census of workers in the United States, covering approximately 98 percent of employees on non-farm payrolls. The dotted line in Figure 6 plots the time series of log aggregate employment in private establishments in the full QCEW sample. Relative to this full sample we apply three further sample restrictions, illustrated by the successive lines in Figure 6.

First, access for outside researchers to QCEW/LDE microdata is restricted to a subset of forty states that approve access onsite at the BLS for external research projects. As a result, our sample excludes data for Florida, Illinois, Massachusetts, Michigan, Mississippi, New Hampshire, New York, Oregon, Pennsylvania, Wisconsin, and Wyoming.

Second, we restrict our sample to continuing establishments with positive employment in consecutive quarters. Specifically, we construct a set of overlapping quarter-to-quarter balanced panels that exclude births and deaths of establishments *within* the quarter. Note that we do not balance *across* quarters, so births in a given panel will appear as incumbents in the subsequent panel (if they survive). We focus on continuing establishments because the canonical models of adjustment frictions analyzed above are intended to describe adjustment patterns among incumbent firms.[14]

Our final sample restriction is to exclude establishments with more than 1000 employees in consecutive quarters. We do this for practical reasons. To measure the flow-

[13] The count of workers includes all those receiving any pay during the pay period, including part-time workers and those on paid leave.

[14] An additional, more practical reason for focusing on continuers is that the matching of establishments over time is more subject to error when identifying births and deaths in the QCEW. Among other steps, the latter involves identification of predecessor and successor establishments to detect spurious births or deaths. By excluding establishment births and deaths, our sample excludes establishments that require linkage using predecessor or successor information. The latter is a very small subset of establishments, around 0.1 percent in 2014Q2.

balance employment distribution in equations (4) and (13), and hence the diagnostic suggested by the theory, we require measures of establishment flows between points in the firm size distribution—specifically, inflows of mass to each employment level, and the probability of outflow. To measure the latter with sufficient precision requires sufficient sample sizes at all points in the distribution. Since establishments with more than 1000 employees comprise a very small fraction of U.S. establishments—for example, less than 0.1 percent in 2014Q2—sample sizes become impracticably thin beyond 1000 employees, inducing substantial noise in implied estimates of our diagnostic.

Recall that our goal is to understand whether canonical models of labor market frictions can account for the dynamics of aggregate employment. A worry, then, is that the forgoing sample restrictions might alter the dynamics of aggregate employment in our sample relative to the full sample.

Figure 6 reveals that this is not the case. In terms of levels, the largest loss of sample size occurs because we are unable to access data for all states, accounting for around 30 percent of total employment in the United States. The further exclusion of non-continuing establishments and large establishments accounts, respectively, for around 2 percent and 10 percent of employment. However, Figure 6 shows that the path of aggregate employment in our sample resembles, in both trend and cycle, the path of aggregate employment in the full QCEW sample. The correlation between log aggregate employment in the published QCEW series for all states and that in our final microdata sample is 0.99.

Measurement. To estimate our diagnostic, we require first an estimate of the distribution of employment under flow balance, $\hat{h}(m)$. Rearranging equations (4) and (13), we can write the latter as

$$\hat{h}_t(m) = h_{t-1}(m) + \frac{\Delta h_t(m)}{\phi_t(m)}, \tag{20}$$

where t indexes quarters, $h_{t-1}(m)$ is the previous quarter's mass of establishments with employment m, $\Delta h_t(m) \equiv h_t(m) - h_{t-1}(m)$ is the quarterly change in that mass, and $\phi_t(m)$ is the fraction of establishments that adjusts away from an employment level of m in quarter t. Thus, estimation of $\hat{h}_t(m)$ requires only an estimate of the outflow adjustment probability $\phi_t(m)$, in addition to measures of the evolution of the firm size distribution $h_t(m)$.

The simplest approach to measuring $\phi_t(m)$ is to use our microdata to compute the fraction of establishments with m workers in quarter t that reports employment different from m in quarter $t+1$. As alluded to above in motivating our sample restrictions, however, a practical issue that arises is that sample sizes become small as m gets large, inducing sampling variation in estimates of $\phi_t(m)$.

We further address this issue by discretizing the employment distribution at large m. An advantage of the substantial sample sizes in the QCEW/LDE microdata is that we can be relatively conservative in this regard. In particular, we allow individual bins for each integer employment level up to 250 workers. In excess of 99 percent of establishments lie in this range, and so sample sizes in each bin are large, between about 100 and 1.3 million establishments. For establishment sizes of 250 through 500 workers we use bins of length five, allowing us to maintain sample sizes above about 80 establishments in each quarter. Further up the distribution, of course, sample sizes get smaller, so we extend our bin length to ten for employment levels between 500 and 999 workers. In this range, sample sizes are at least 15 establishments in each quarter.

Denoting these bins by b, we estimate the firm size mass and the outflow probability as

$$h_t(b) = \sum_i \mathbb{I}[n_{it} \in b], \text{ and } \phi_t(b) = \frac{\sum_i \mathbb{I}[n_{it} \notin b | n_{it-1} \in b]}{\sum_i \mathbb{I}[n_{it-1} \in b]}, \tag{21}$$

where i indexes establishments. We use these measures to compute the flow-balance mass in each bin according to equation (20) as $\hat{h}_t(b) = h_{t-1}(b) + [\Delta h_t(b)/\phi_t(b)]$. Finally, we compute aggregate employment and its flow-balance counterpart by taking the inner product of h_t and $\hat{h}_t$ with the midpoints of each bin, denoted m_b,

$$N_t = \sum_b m_b h_t(b), \text{ and } \hat{N}_t = \sum_b m_b \hat{h}_t(b). \tag{22}$$

2.2 Inferring the aggregate effects of frictions

With this estimate of flow-balance aggregate employment $\hat{N}_t$ in hand, we can now contrast its dynamics with the path of *actual* aggregate employment N_t. Recall the predictions from the canonical models we have analyzed: Figures 2 through 5 reveal that flow-balance employment overshoots its frictionless counterpart on impact, and never dips much below it over the full horizon of the impulse response. The quantitative importance of canonical frictions can then be inferred by how much flow-balance employment overshoots *actual* employment in a labor market upswing, and *vice versa* in a downswing. If the gap between the two series is small in the data, canonical models imply that the actual and frictionless paths must nearly coincide. The inference is that frictions have a modest effect on aggregate employment dynamics.

A first look at the data. Figure 7 plots the time series of N_t and $\hat{N}_t$ derived from application of equation (22) to the QCEW/LDE microdata. Both series are expressed in log deviations from a quadratic trend.[15] Figure 7 reveals that $\hat{N}_t$ is a leading indicator of

[15] To be consistent with the method of detrending we later use to estimate impulse responses, we use a quadratic trend rather than the HP filter. As Ashley and Verbrugge (2006) have noted, a drawback of the HP filter is that it uses *future* realizations to derive the cyclical component in the present period. As a

actual employment N_t, and is also more volatile. Specifically, the standard deviation of N_t is 0.025, whereas the standard deviation of $\widehat{N}_t$ is 0.031.

On the whole, however, the differences between the two series are modest. The median (mean) absolute difference between the series is just 0.5 (0.8) log points. Indeed, there is remarkably little daylight between the two series between 1992 and 2008. Even in the 2001 recession, flow-balance employment very closely tracks the drop in actual employment. The only substantial difference between the series emerges in the Great Recession. For instance, in the five quarters that bracket the trough of the recession, 2008Q4 to 2009Q4, the mean difference between the series is about 3 log points. However, this difference is short-lived. Since 2010, the two series have moved in tandem: employment has increased 11.6 log points, whereas flow-balance employment has increased 11.9 log points.

Together these observations give a first suggestion that, when viewed through the lens of canonical models of labor market frictions, the aggregate effects of such frictions account for a modest part of aggregate employment dynamics.

Time series matching. To contrast the data with the models' predictions more precisely, we undertake a simulation exercise devised by King and Rebelo (1999) and Bachmann (2012). They show that it is possible to find a sequence of aggregate shocks that generates a path for aggregate model-generated outcomes—in our case employment—that matches an empirical analogue. In what follows, we use this technique to contrast the time series of flow-balance employment in model and data when the path of aggregate employment in each is constructed to be the same.

The procedure relies on the ability to summarize the dynamics of aggregate employment implied by the model using a simple aggregate law of motion. In a related adjustment cost model, Bachmann shows that an AR(1) specification that relates log aggregate employment to its own lag and current labor productivity does an excellent job of summarizing these dynamics. We find that the same property holds for our model.

Figures 2 through 5 suggest that linear cost models are especially capable of generating persistence in actual aggregate employment. We therefore initiate an algorithm with a variant of the linear cost model.[16] To give the model the best chance of generating substantial persistence in aggregate employment, we use the parameterization that yields an inaction rate of 80 percent per quarter (which substantially exceeds the empirical rate of 52.5 percent).

---

result, lags of HP deviations included as regressors in a VAR are not predetermined with respect to the outcome variable.

[16] We assume $c^+ = c^-$ and do not pursue the effects of asymmetries in the adjustment costs here: the theoretical results of Propositions 1 and 2, and the quantitative results of Figure 4, suggest that any such asymmetries affect neither the dynamics of aggregate employment, nor its flow-balance counterpart.

Following Bachmann, in a first step we use this model to generate 89 quarters of simulated data (the same time span as in the data). We then estimate via OLS the following AR(1) process that relates log aggregate employment to its lag and current total factor productivity p_t,

$$\ln N_t = \hat{v}_0 + \hat{v}_1 \ln N_{t-1} + \hat{v}_2 \ln p_t. \tag{23}$$

With estimates of equation (23) in hand, it is possible to back out a series for productivity $\{p_t\}$ that replicates the path of log aggregate employment $\{\ln N_t\}$. Since the resultant sequence $\{p_t\}$ may not be consistent with the data generating process assumed, in a final step we re-parameterize the productivity process and re-initialize the model with this updated process. These steps are repeated until the moments of the implied productivity series are consistent with the parameterization assumed. In practice, the AR(1) specification in (23) fits the data closely (the R-squared of the regression is 0.9985), and so the algorithm converges quite quickly, after just a few iterations.[17]

Figure 8 illustrates the results. To smooth out high frequency noise, we use the above algorithm to replicate the four-quarter moving average of log aggregate employment in the data. The standard deviation of this smoothed series for actual employment $\ln N$ is 0.023. The model yields a notably more variable path for aggregate flow-balance employment, $\widehat{N}$. The model-implied standard deviation of $\ln \widehat{N}$ is 0.038, much larger than its empirical counterpart of 0.028. Similarly, by the NBER-dated troughs of the last two recessions, the model-generated series for $\widehat{N}$ has fallen an additional 5 log points. Aggregate flow-balance employment also recovers more quickly in the wake of these downturns. In the six to eight quarters after the Great Recession, for instance, the model's flow-balance employment rises almost 10 log points. Its empirical counterpart increases by less than half that amount over the same period.

Measuring persistence. A final way of visualizing the difference between the data and the models' predictions is to contrast the response of flow-balance employment to estimated shifts in the aggregate driving force. Rather than attempting to use the data to identify structural shocks, which is prone to controversy, we instead undertake a descriptive analysis of the dynamic properties of aggregate employment. A commonly used gauge for the latter is a comparison of the dynamics of employment relative to output-per-worker. In what follows, we interpret unforecastable movements in output-per-worker as being indicative of innovations to the (latent) driving force, and estimate the reaction of flow-balance employment, in model and data, to these forecast errors. This serves as a simple way of summarizing the persistence of flow-balance employment.

[17] The i.i.d. innovations of the implied productivity process have standard deviation 0.0062, which is more than twice that used in the process that underlies Figures 2 through 5. But, across those figures, our goal was to hold the aggregate productivity process fixed, and a standard deviation of 0.0062 would yield much larger fluctuations in model-generated employment than ever seen in the data.

Formally, we proceed as follows. Denote log output-per-worker by y_t. In a first stage, we estimate innovations in y_t that are unforecastable conditional on lags of y, and lags of log aggregate employment $\ln N$. Specifically, we use quarterly data on output-per-worker in the nonfarm business sector from the BLS Productivity and Costs release and our measure of actual employment from the QCEW to estimate the following AR(L) specification:

$$y_t = \alpha^y + \sum_{s=1}^{L} \beta_s^y y_{t-s} + \sum_{s=1}^{L} \gamma_s^y \ln N_{t-s} + \delta_1^y t + \delta_2^y t^2 + \varepsilon_t^y. \tag{24}$$

Within the context of the models considered in section 1, lags of output per worker y can be interpreted as proxies for lags of the driving force p. More broadly, they can viewed as proxies for past realizations of business cycle driving forces. Note that secular trends are captured using a quadratic time trend.

The estimated residuals from this first-stage regression, $\hat{\varepsilon}_t^y$, are then used as the innovations to output-per-worker from which we derive impulse responses of actual and flow-balance employment in a second stage,

$$\begin{aligned} \ln N_t &= \alpha^N + \sum_{s=0}^{L-1} \beta_s^N \hat{\varepsilon}_{t-s}^y + \sum_{s=1}^{L} \gamma_s^N \ln N_{t-s} + \delta_1^N t + \delta_2^N t^2 + \varepsilon_t^N, \text{ and} \\ \ln \widehat{N}_t &= \alpha^{\widehat{N}} + \sum_{s=0}^{L-1} \beta_s^{\widehat{N}} \hat{\varepsilon}_{t-s}^y + \sum_{s=1}^{L} \gamma_s^{\widehat{N}} \ln \widehat{N}_{t-s} + \delta_1^{\widehat{N}} t + \delta_2^{\widehat{N}} t^2 + \varepsilon_t^{\widehat{N}}. \end{aligned} \tag{25}$$

Note that the timing in the lag structure of innovations to output-per-worker permits a *contemporaneous* relationship between these innovations and employment, as suggested by the model-based impulse responses described in section 1.

The estimates from the regressions in equations (24) and (25) allow us to trace out the dynamic relationship between each measure of log aggregate employment and a one-log-point innovation in output-per-worker. In practice, we use a lag order of $L = 4$ in both stages, (24) and (25).[18] Given the availability of our QCEW data, we estimate these regressions over the period, 1992Q2 to 2014Q2.

Panel A of Figure 9 plots the results. The dynamic response of aggregate employment takes a familiar shape, rising slowly after the innovation with a peak response of around 1 log point after five quarters. These hump-shaped dynamics mirror similar results found using different methods elsewhere in the literature (Blanchard and Diamond 1989; Fujita and Ramey 2007; Hagedorn and Manovskii 2011). This is one representation of the persistence of aggregate employment.

As suggested by the time series in Figure 7, the dynamics of the flow-balance diagnostic $\widehat{N}$ share many of these properties. Although its peak response occurs earlier—

[18] Experiments with different lag orders suggest that, although the peak of the hump-shaped impulse responses varies slightly across different lag lengths, Figure 9 is representative of results across these specifications.

after three quarters—reinforcing the impression of Figure 7 that $\widehat{N}$ is a leading indicator of the path of N, it exhibits a similar volatility, and a clear hump-shape.

To contrast the empirical dynamics illustrated in Figure 9A with those implied by canonical models of frictions, we rerun the regressions in equations (24) and (25) using model-generated data. Following our preceding discussion, we use the model with symmetric linear costs. But we now choose the linear cost to minimize the (sum of squares) distance between the empirical dynamics of actual employment N in Figure 9A and those implied by the model. We obtain a best fit for a linear cost that generates a quarterly inaction rate of 85 percent. Again, we are being generous to the model in allowing it to violate the inaction rate observed in the data.

Panel B of Figure 9 reveals that this parameterization of the model is able to generate a dynamic relationship between *actual* employment and output-per-worker that is comparable to the data. Although the model overstates the impact response, the amplitude and persistence of employment are similar to their empirical counterparts.

A key result of Figure 9B, however, is that the model-implied dynamics of flow-balance employment are profoundly different from those seen in the data. Confirming the impression of the theoretical impulse responses in Figure 3, $\widehat{N}$ *jumps* in response to innovations in output-per-worker in the model, with an initial response *five times* larger than that of actual employment N. In marked contrast, the empirical dynamics of $\widehat{N}$ in Figure 9A are much more sluggish, bearing a closer resemblance to the empirical path of actual employment than its model-implied counterpart.

The substantial discrepancy between the implied and observed dynamics of flow-balance employment is an important failure of canonical models of frictions, in the sense that the sources of propagation captured by these models appears not to be a quantitatively important force in the data.

2.3 Understanding the failure of canonical models

To examine the origins of this failure of canonical models, recall that the link between our diagnostic flow-balance employment $\widehat{N}$ and frictionless employment N^* is mediated through the behavior of firm size flows—the τs and ϕs of equations (4) and (13)—and that canonical frictions have strong predictions regarding the dynamics of these flows by establishment size.

As we have emphasized, a key benefit of the data is that we can measure aspects of these flows using the longitudinal dimension of the QCEW microdata—specifically the total inflow to, and the probability of outflow from, each employment level. Our next exercise, therefore, is to contrast the dynamics of the firm size distribution in the data to those implied by canonical models of frictions.

To do this, we first split establishments in the data into three size classes. We choose these to correspond to the lower quartile (fewer than 15 employees), interquartile range

(16 to 170 employees), and upper quartile (171 employees and greater) of establishment sizes. We then estimate descriptive impulse responses that mirror equations (24) and (25) for the total inflow to, and probability of outflow from, each size class.[19] As in our previous analysis of the dynamics of aggregate employment, we repeat these same steps using data simulated from the model underlying Figure 9B that is calibrated to match as closely as possible the empirical dynamics of aggregate employment.

Panels C through F of Figure 9 illustrate the results of this exercise. The empirical and model-implied dynamics share a qualitative property, namely that positive aggregate shocks render small (large) establishments more (less) likely to adjust away from their current employment, and induce fewer (more) establishments to adjust to low (high) employment levels.

Aside from this broad qualitative similarity, the quantitative dynamics reveal striking contrasts. The empirical behavior of firm size flows exhibits an inertia not only in the sense that their *levels* are retarded relative to a frictionless environment, but also in the sluggishness of their *responses* to aggregate disturbances.

We highlight three manifestations of this general observation. First, note that the empirical responses of the firm size flows in Figures 9C and 9E are an order of magnitude smaller than their theoretical counterparts in Figures 9D and 9F. Second, the dynamics of the flows in the data are much more sluggish than implied by canonical frictions. Firm size dynamics in the model respond aggressively on impact of the aggregate shock. In the data, the response is mild and delayed. Third, the empirical dynamics reveal an establishment size gradient in the magnitude of the response of firm size flows: Flows to and from smaller establishments respond less than their counterparts for larger establishments.

The upshot of this exercise is that canonical models of labor market frictions do a poor job of capturing the empirical dynamics of the firm size distribution. Since the latter is the key channel through which canonical frictions are supposed to impede aggregate employment dynamics, this is an important limitation of this class of model.

3. Summary and discussion

In this paper, we have explored the propagation mechanism embodied in a canonical class of labor market frictions. In postulating several forms of non-convex adjustment frictions, this class has the virtue of being able to reproduce the conspicuous degree of inaction observed in establishment employment dynamics. We further show that (some of) these labor market frictions are in turn able to generate at least part of the observed sluggishness in aggregate employment dynamics.

[19] To aggregate within a quartile range, we take a weighted average across establishment sizes, where the weight is the size's share of all establishments in the range.

However, canonical frictions have strong implications for the source of this propagation, for which we do not find empirical support. In this class of models, deviations of aggregate employment from its frictionless path arise because frictions retard the flow of labor across firms. Since the latter induces pent-up demand for adjustment, these firm size flows are predicted to respond rapidly to aggregate shocks. We use this to motivate an empirical diagnostic that is predicted to bound the path of frictionless employment under canonical frictions, and can be measured with access to establishment panel data.

We find that empirical measures of this diagnostic display only mild departures from the empirical path of actual employment, exhibit much more sluggish dynamics than implied by canonical frictions, and that the source of this tension can be traced to a failure of canonical models to capture the empirical persistence of firm size flows.

We highlight two possible conclusions in the light of these findings. The first is that labor market frictions that induce inaction are indeed unimportant for aggregate employment dynamics, and that the empirical sluggishness of aggregate employment is simply a reflection of sluggishness in firms' desired employment. The latter could in turn be accounted for by a return to the older literature on convex adjustment costs (as in, for example, Sargent 1978).

A drawback of such a conclusion, however, is that the presence of inaction is perhaps the most prominent stylized fact of microeconomic employment adjustment. In acknowledgement of this fact, a second, alternative conclusion is that future work should explore the possibility that such inaction might *interact* with other frictions. A constructive input into such an exploration is our finding that the flow of labor across firms is itself quite sluggish in establishment microdata. An interesting possibility, then, is that this persistence in firm size dynamics is itself a source of persistence in aggregate employment.

We provide one example of this possibility, based on an interaction of labor market and information frictions. Intuitively, if firms do not have full information on aggregate disturbances, it is possible that this information friction will attenuate the reaction of firms' hiring and firing policy rules, dampening the response of flows of labor across firms.

To illustrate, suppose aggregate productivity is the sum of transitory and permanent components. Firms observe aggregate productivity but not its constituent parts.[20] In the absence of the labor market frictions, firms' labor demand is the outcome of a simple static optimization problem, for which only knowledge of aggregate productivity is required. Thus, absent labor frictions, the information friction has no bite.

In the presence of employment adjustment frictions, however, firms must forecast the path of the aggregate state, which requires a judgment of the degree to which an aggregate disturbance is permanent. Hiring and firing decisions are thus based on perceptions of the persistent component of productivity. Standard signal extraction arguments will imply

[20] For early applications of this information structure in macroeconomics, see Brunner, Cukierman, and Meltzer (1980) and Gertler (1982). More recently, see Erceg and Levin (2003).

that such perceptions are a slow-moving state variable. Accordingly, hiring and firing decisions respond less aggressively to aggregate shocks on impact. This can lead, qualitatively, to the drawn-out dynamics of the labor market flows we observe in the data. Critically, this persistence in hiring and firing policies will in turn contribute to persistent aggregate employment dynamics.

The quantitative success of such a model will hinge on the rate at which firms update their assessment of the persistent component of aggregate productivity, and the extent to which such persistence can be reconciled with the large cyclical volatility of employment. Nonetheless, we suspect that an interaction of labor market frictions with a notion of imperfect information provides a promising avenue of further research that seeks to understand aggregate employment persistence.[21]

References

Abraham, Arpad and Kirk White. 2006. "The Dynamics of Plant-level Productivity in U.S. Manufacturing." Center for Economic Studies WP 06-20.

Acemoglu, Daron and William Hawkins. 2014. "Search with Multi-worker Firms." *Theoretical Economics*, 9(3): 583-628.

Ashley, Richard and Randal J. Verbrugge. 2009. "Nonlinear Dynamic Relationships in Time Series." *Econometric Reviews*, 28(1–3): 4-20.

Bachmann, Ruediger. 2012. "Understanding Jobless Recoveries after 1991 and 2001." Mimeo, University of Notre Dame.

Bentolila, Samuel and Giuseppe Bertola. 1990. "Firing Costs and Labour Demand: How Bad is Eurosclerosis?" *Review of Economic Studies,* 57(3): 381-402.

Blanchard, Olivier Jean, and Peter Diamond. 1989. "The Beveridge Curve." *Brookings Papers on Economic Activity* 1: 1–60.

Bloom, Nicholas. 2009. "The Impact of Uncertainty Shocks." *Econometrica* 77(3): 623-685.

Bruegemann, Bjoern, Pieter Gautier and Guido Menzio. 2015. "Intra Firm Bargaining and Shapley Values." Mimeo, Vrije Universiteit Amsterdam.

Brunner, Karl, Alex Cukierman, and Allan Meltzer. 1980. "Stagflation, persistent unemployment, and the permanence of economic shocks." *Journal of Monetary Economics*, 6: 467-492.

Caballero, Ricardo and Eduardo M. R. A. Engel. 1993. "Microeconomic Adjustment Hazards and Aggregate Dynamics." *Quarterly Journal of Economics,* 433(2): 359-383.

[21] Interestingly, though information frictions in macro have been revived in recent literature in monetary economics (see Mankiw and Reis' 2011 survey), they have been used much more sparingly in understanding of labor dynamics. (For an exception, see Venkateswaran (2013).)

Caballero, Ricardo and Eduardo M. R. A. Engel. 1999. "Explaining Investment Dynamics in U.S. Manufacturing: A Generalized *(S,s)* Approach." *Econometrica,* 67(4): 783-826.

Caballero, Ricardo and Eduardo M. R. A. Engel. 2007. "Price Stickiness in Ss Models: New Interpretations of Old Results." *Journal of Monetary Economics,* 54(S1): 100-121.

Caballero, Ricardo, Eduardo M. R. A. Engel, and John Haltiwanger. 1995. "Plant Level Adjustment and Aggregate Investment Dynamics." *Brookings Papers on Economic Activity,* 2*:* 1-54.

Caballero, Ricardo, Eduardo M. R. A. Engel, and John Haltiwanger. 1997. "Aggregate Employment Dynamics: Building from Microeconomics." *American Economic Review*, 87(1), 115-137.

Chetty, Raj. 2012. "Bounds on Elasticities with Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply." *Econometrica* 80(3): 969-1018.

Chetty, Raj, Adam Guren, Day Manoli, and Andrea Weber. 2012. "Does Indivisible Labor Explain the Difference between Micro and Macro Elasticities? A Meta-Analysis of Extensive Margin Elasticities." In *NBER Macroeconomics Annual 2012*, edited by D. Acemoglu, J. Parker, and M. Woodford, University of Chicago Press, 1-56.

Cooper, Russell, John Haltiwanger, Jonathan L. Willis. 2007. "Search frictions: Matching aggregate and establishment observations." *Journal of Monetary Economics* 54(1): 56-78.

Cooper, Russell, John Haltiwanger, Jonathan L. Willis. 2015. "Dynamics of labor demand: Evidence from plant-level observations and aggregate implications." *Research in Economics* 6(1): 37-50.

Davis, Steven J., and John Haltiwanger. 1992. "Gross Job Creation, Gross Job Destruction, and Employment Reallocation." *Quarterly Journal of Economics* 107(3): 819-864.

Elsby, Michael W. L. and Ryan Michaels. 2013. "Marginal Jobs, Heterogeneous Firms, and Unemployment Flows." *American Economic Journal: Macroeconomics,* 5(1): 1-48.

Elsby, Michael W. L. and Ryan Michaels. 2014. "Fixed Adjustment Costs and Aggregate Fluctuations," Mimeo, University of Edinburgh.

Elsby, Michael W. L., Donggyun Shin and Gary Solon. 2016. "Wage Adjustment in the Great Recession and other Downturns: Evidence from the United States and Great Britain." *Journal of Labor Economics* 34 (S1): S249-S291.

Erceg, Christopher and Andrew Levin. 2003. "Imperfect credibility and inflation persistence." *Journal of Monetary Economics*, 90: 915-944.

Fujita, Shigeru, and Garey Ramey. 2007. "Job Matching and Propagation." *Journal of Economic Dynamics and Control* 31 (11): 3671–98.

Gertler, Mark. 1982. "Imperfect information and wage inertia in the business cycle." *Journal of Political Economy*, 90(5): 967-987.

Gertler, Mark and John Leahy. 2008. "A Phillips Curve with an *Ss* Foundation." *Journal of Political Economy,* 116(3): 533-572.

Golosov, Mikhail and Robert E. Lucas. 2007. "Menu Costs and Phillips Curves." *Journal of Political Economy* 115(2): 171-199.

Hagedorn, Marcus, and Iourii Manovskii. 2011. "Productivity and the Labor Market: Comovement over the Business Cycle." *International Economic Review* 52 (3): 603–19.

Hamermesh, Daniel. 1989. "Labor Demand and the Structure of Adjustment Costs." *American Economic Review,* 79(4): 674-689.

Hopenhayn, Hugo and Richard Rogerson. 1993. "Job Turnover and Policy Evaluation: A General Equilibrium Analysis." *Journal of Political Economy,* 101(5): 915-938.

Kimball, Miles and Matthew Shapiro. 2010. "Labor Supply: Are Income and Substitution Effects Both Large or Both Small?" Mimeo, University of Michigan.

King, Robert and Sergio Rebelo. 1999. "Resuscitating Real Business Cycles." *Handbook of Macroeconomics*, Volume 1, chapter 14: 927-1007.

King, Robert and Julia Thomas. 2006. "Partial Adjustment Without Apology." *International Economic Review* 47(3): 779-809.

Krusell, Per, and Anthony Smith. 1998. "Income and Wealth Heterogeneity in the Macroeconomy." *Journal of Political Economy* 106(5): 867-896.

Mankiw, Gregory and Ricardo Reis. 2011. "Imperfect information and aggregate supply." *Handbook of Monetary Economics.*

Mortensen, Dale T., and Christopher A. Pissarides. 1994. "Job Creation and Job Destruction in the Theory of Unemployment." *Review of Economic Studies* 61 (3): 397–415.

Nickell, Stephen. 1978. "Fixed Costs, Employment and Labour Demand Over the Cycle." *Economica*, 45(180): 329-345.

Nickell, Stephen. 1986. "Dynamic Models of Labour Demand", Chapter 8, in Ashenfelter, O. and Layard, R. (eds.), *Handbook of Labor Economics*, North Holland.

Oi, Walter Y. 1962. "Labor as a Quasi-Fixed Factor." *Journal of Political Economy*, 70(6), 538-555.

Petrongolo, Barbara, and Christopher A. Pissarides. 2001. "Looking into the Black Box: A Survey of the Matching Function." *Journal of Economic Literature* 39 (2): 390–431.

Pissarides, Christopher A. 1985. "Short-Run Equilibrium Dynamics of Unemployment, Vacancies, and Real Wages." *American Economic Review* 75 (4):676–90.

Solon, Gary, Robert Barsky, and Jonathan Parker. 1994. "Measuring the Cyclicality of Real Wages: How Important is Composition Bias?" *Quarterly Journal of Economics*, 109(1): 1-25.

Stole, Lars A., and Jeffrey Zwiebel. 1996. "Intra-Firm Bargaining under Non-Binding Contracts." *Review of Economic Studies*, 63(3), 375-410.

Venkateswaran, Venky. 2013. "Heterogeneous information and labor market fluctuations." Mimeo.

Figure 1. Ss policies in the presence of fixed, linear, and search adjustment frictions

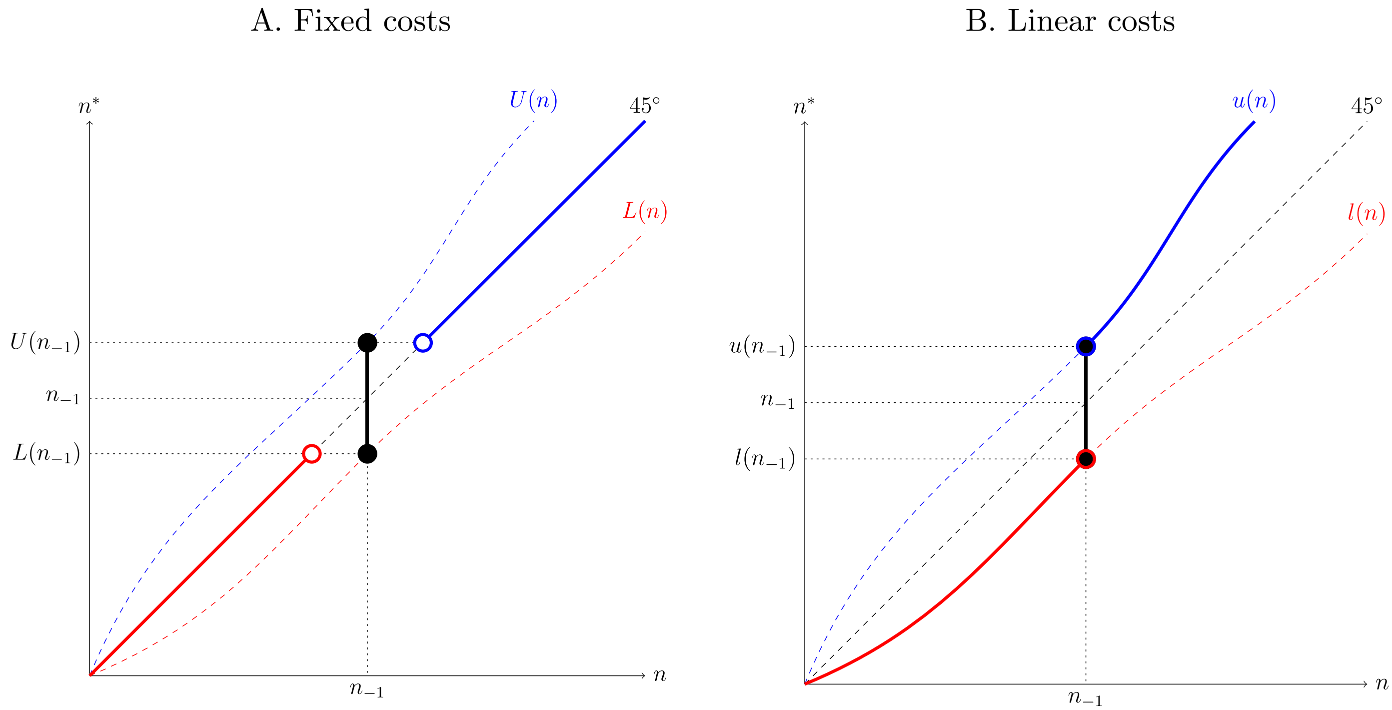


Figure 2. Impulse responses of aggregate employment: Fixed costs

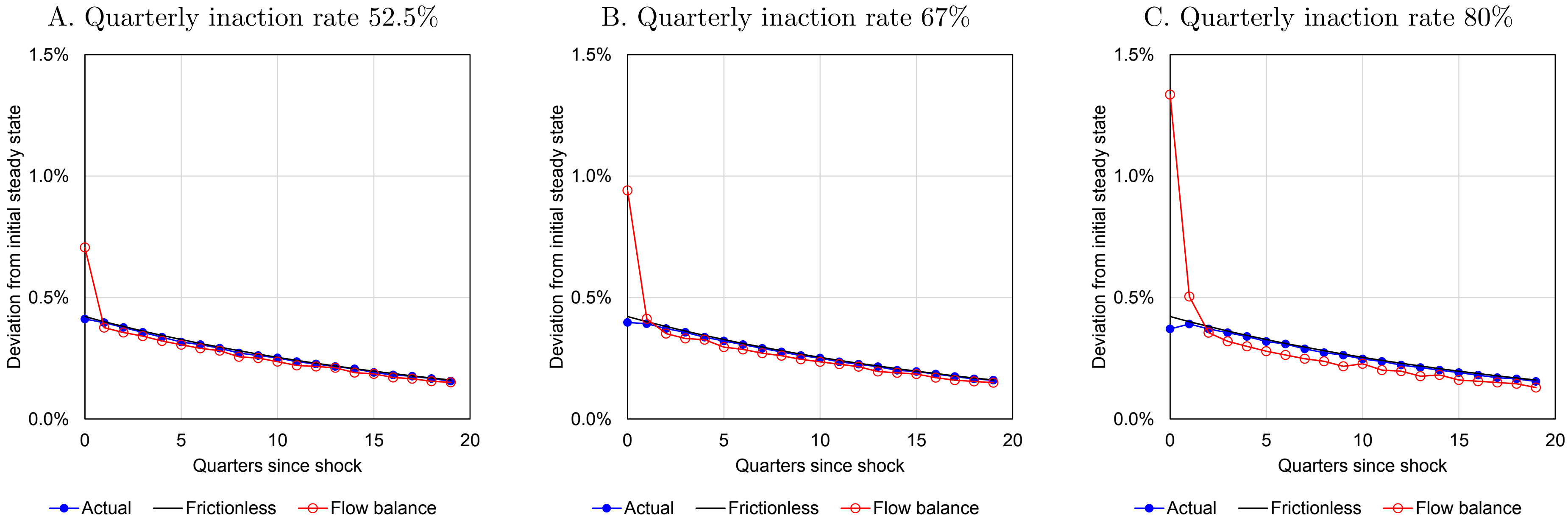


Figure 3. Impulse responses of aggregate employment: Linear costs

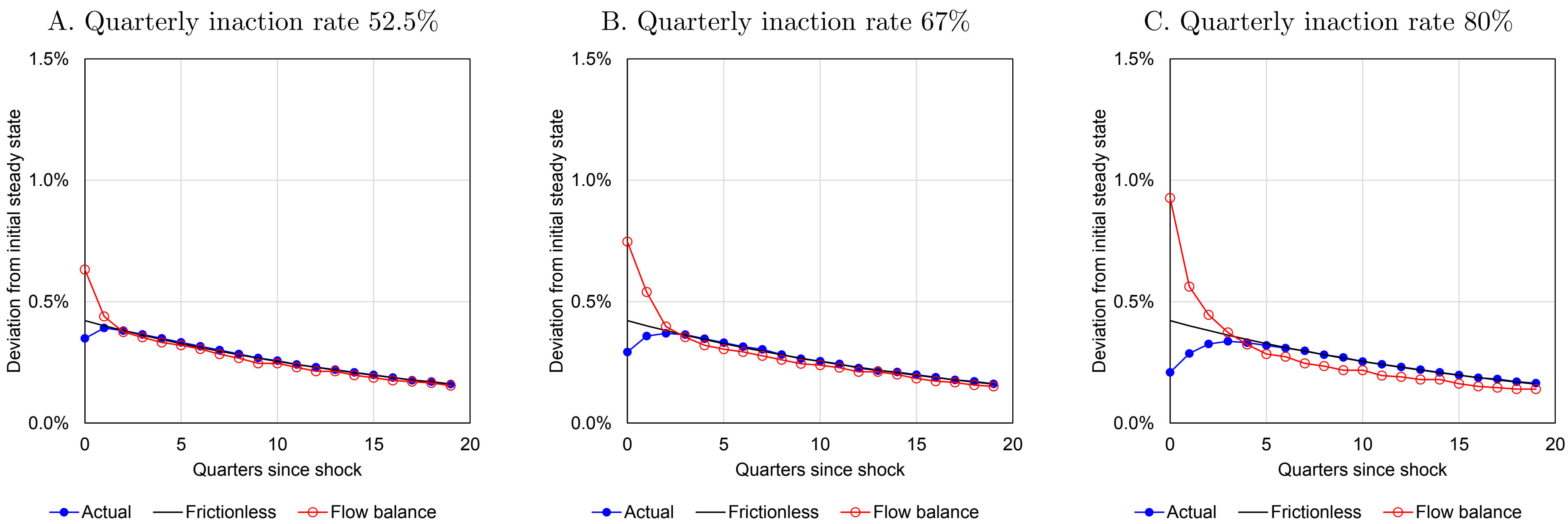
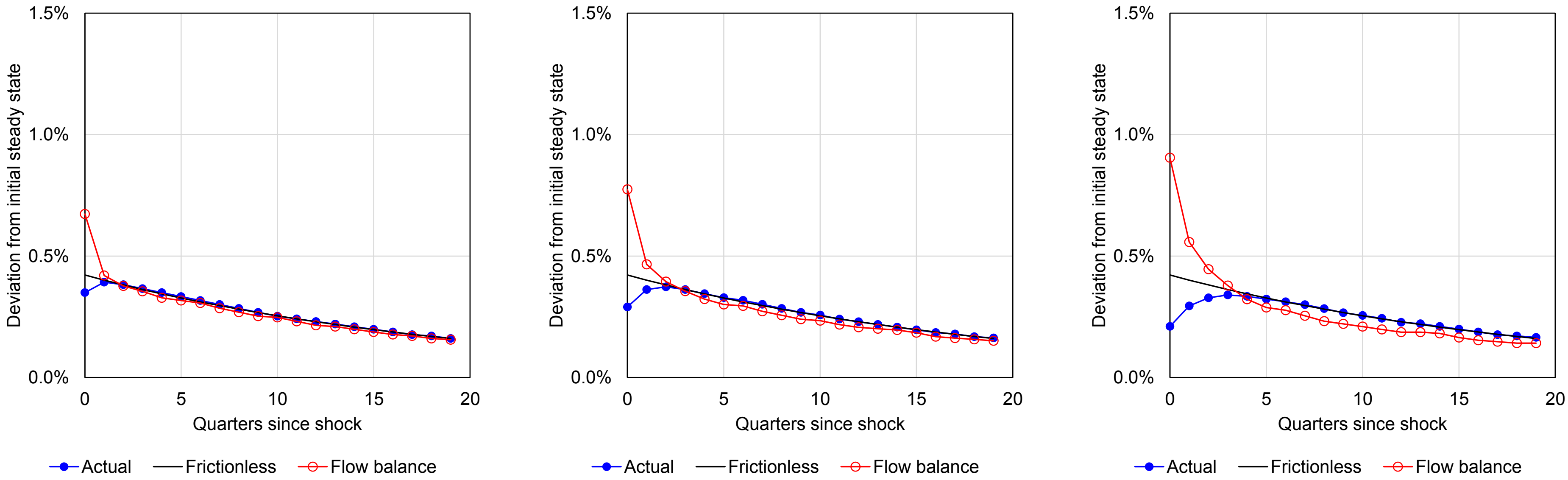


Figure 4. Impulse responses of aggregate employment: Asymmetric linear costs

A. Quarterly inaction rate 52.5% B. Quarterly inaction rate 67% C. Quarterly inaction rate 80%

i. Pure hiring cost



ii. Pure firing cost

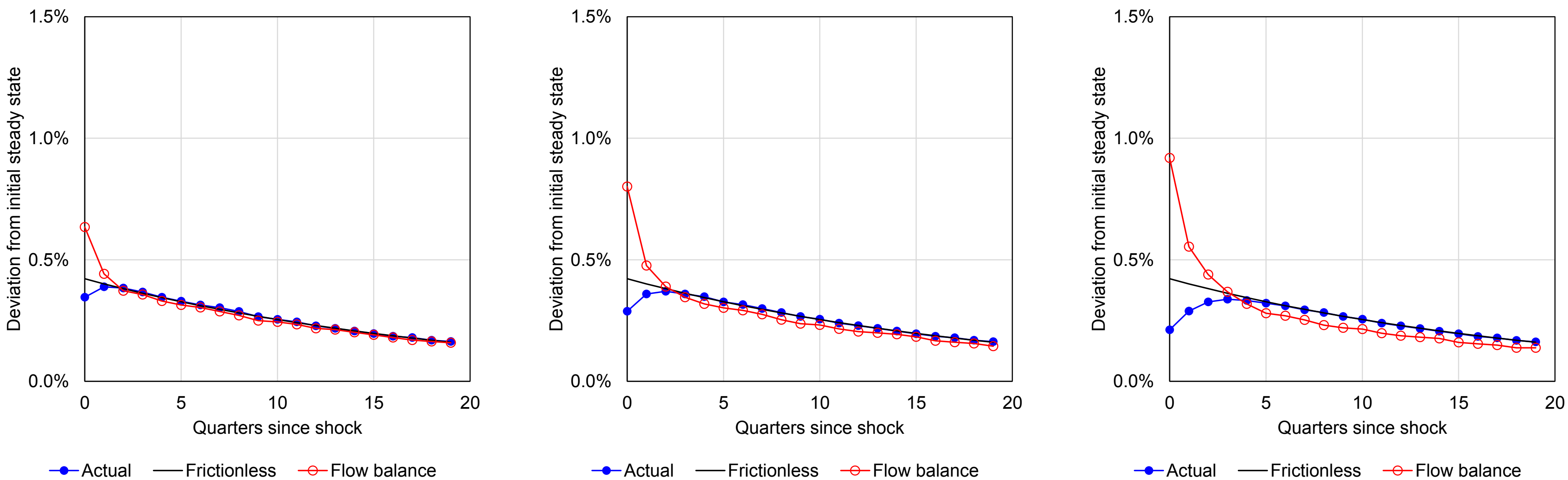


Figure 5. Impulse responses of aggregate employment: Search costs

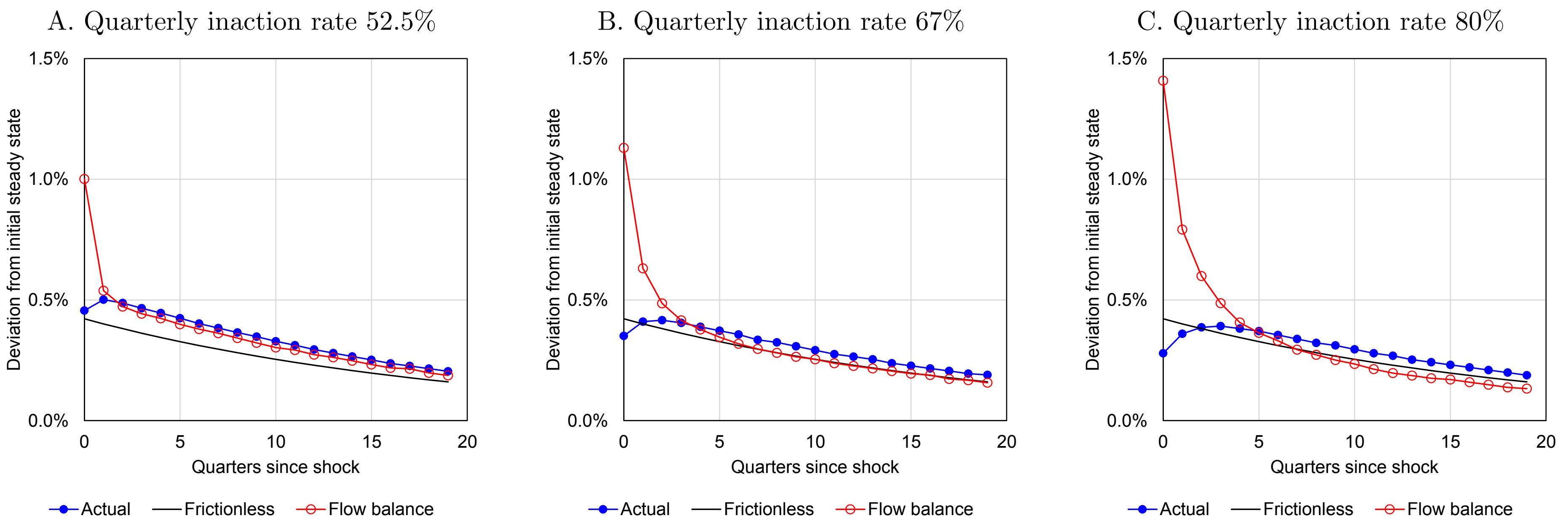


Figure 6. Aggregate employment in the QCEW by sample restriction

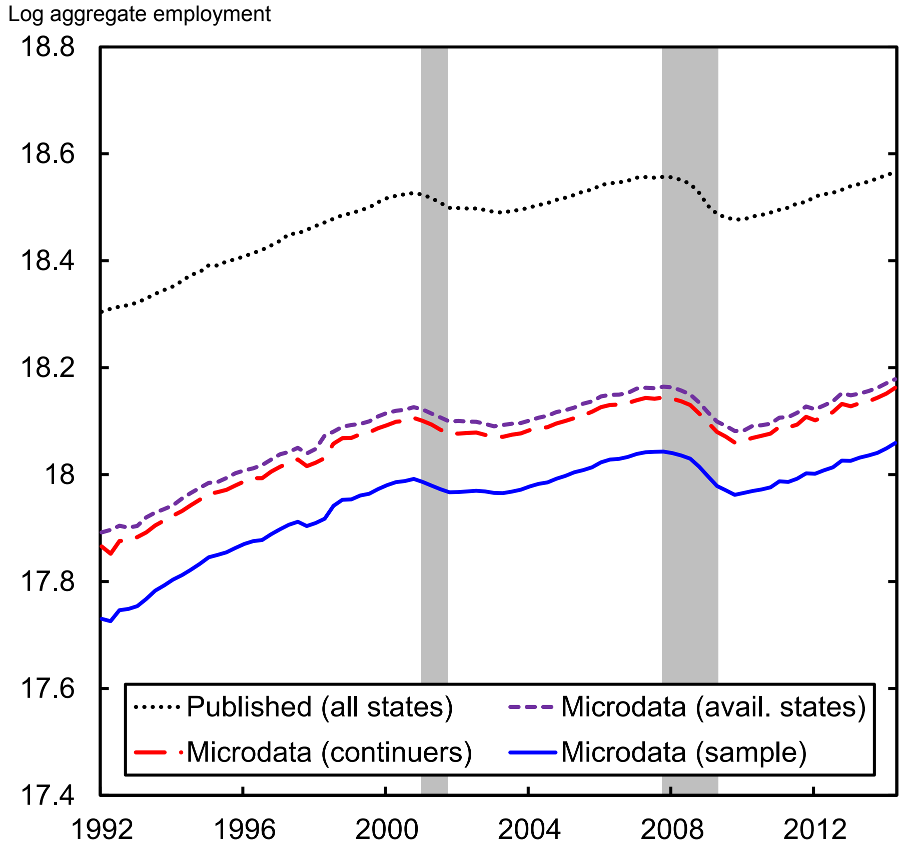


Figure 7. Actual and flow-balance log aggregate employment

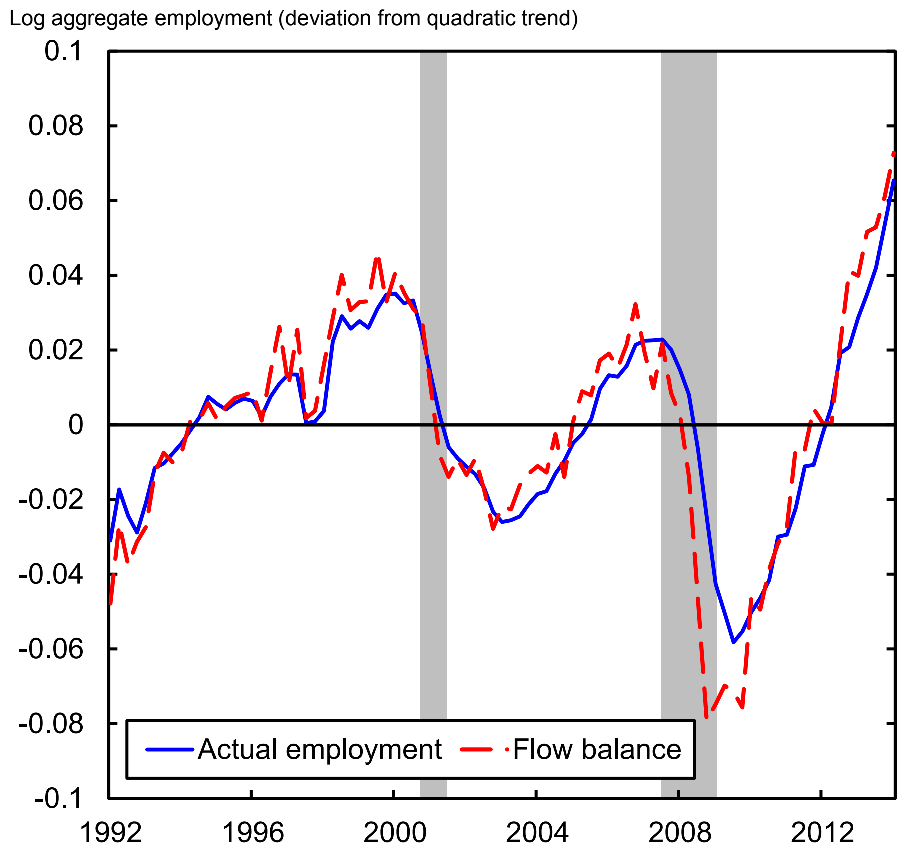


Figure 8. Model-implied time series for flow-balance log aggregate employment

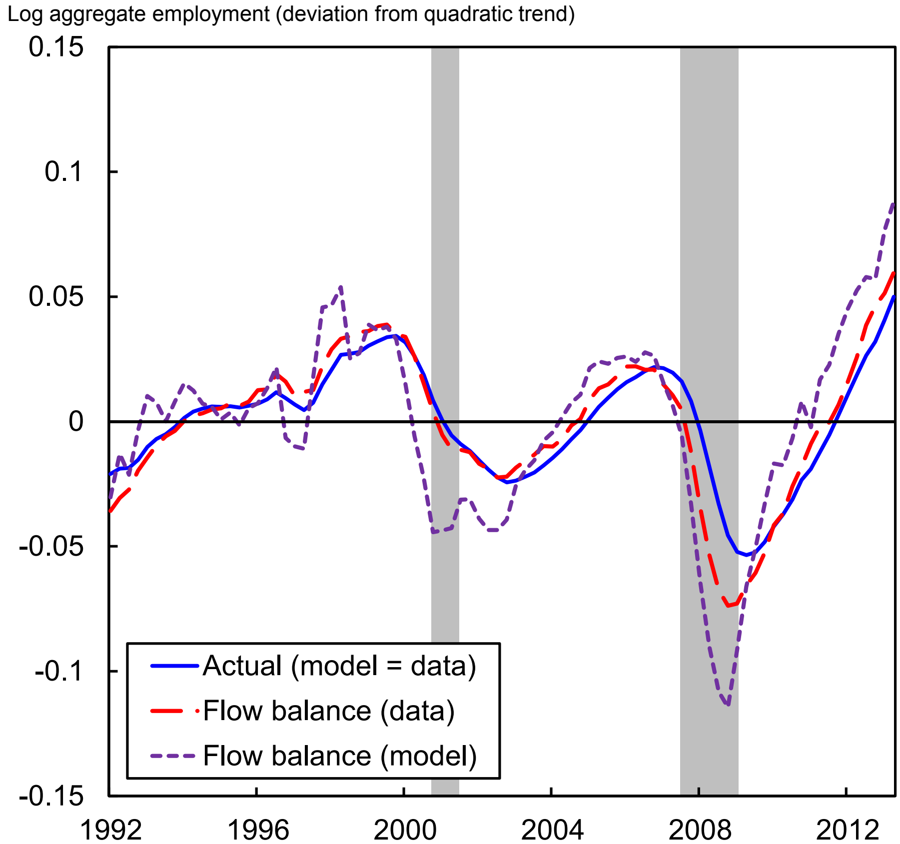
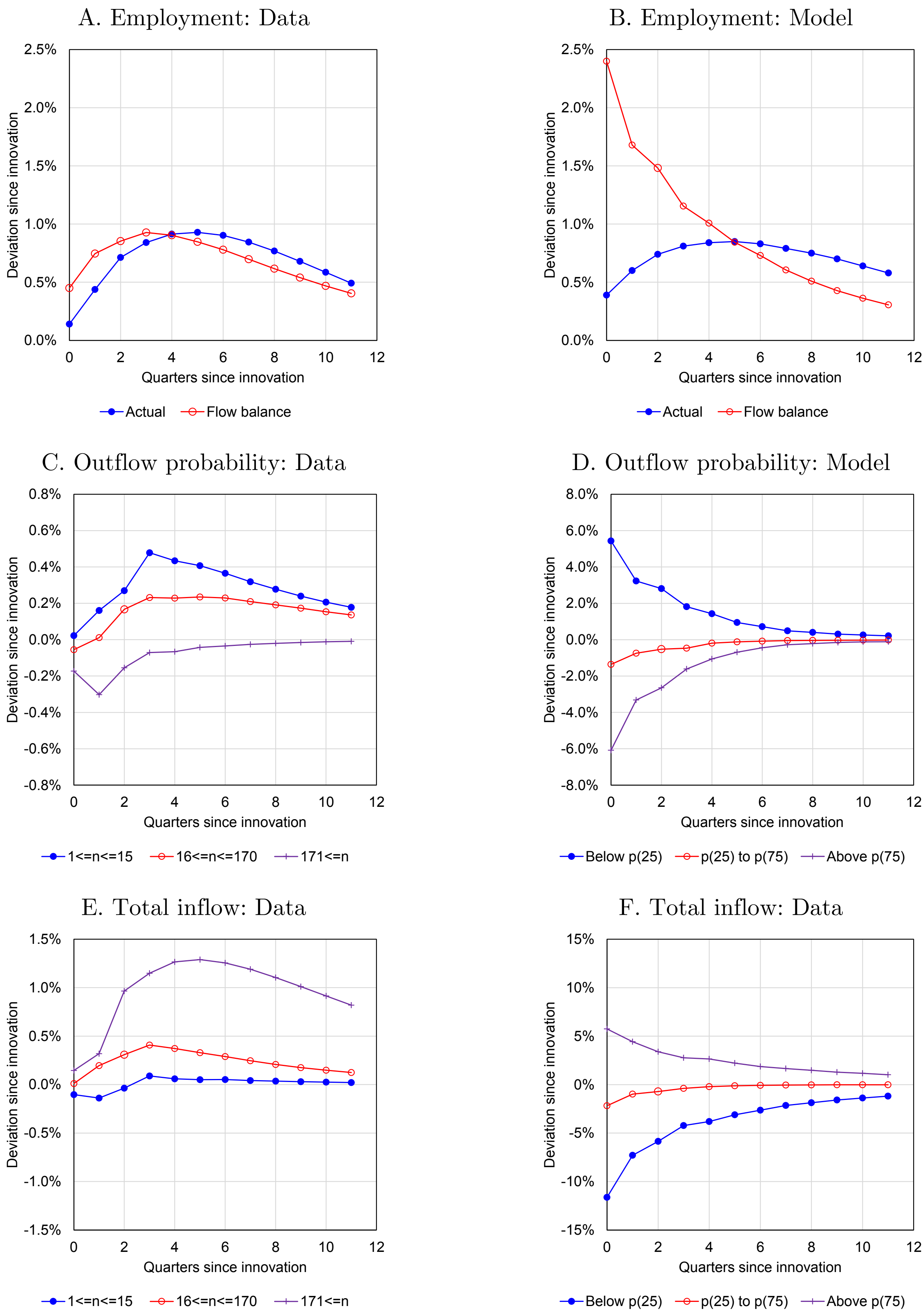


Figure 9. Descriptive impulse responses of employment and firm size flows: Data versus model



Appendices (incomplete)

A. Laws of motion for the firm size distribution

To derive the laws of motion for the density of employment across firms stated in the main text for the fixed and linear cost models, we require notation for several distributions. As in the main text, we denote the densities of employment, lagged employment and mandated employment by h, h_{-1} and h^*. Conventionally, we will refer to their respective distribution functions by analogous upper-case letters, H, H_{-1} and H^*. In addition, however, we require notation for the distributions of mandated employment conditional on lagged employment, which we denote by $\mathcal{H}^*(\xi|\nu) = \Pr(n^* < \xi | n_{-1} = \nu)$, and the distribution of lagged employment conditional on mandated employment, denoted by $\mathcal{H}(\nu|\xi) = \Pr(n_{-1} < \nu | n^* = \xi)$. Of course, the latter are related by Bayes' rule, $\mathcal{h}(\nu|\xi)h^*(\xi) = \mathcal{h}^*(\xi|\nu)h_{-1}(\nu)$, where lower-case script letters denote associated density functions. However, we preserve separate notation where it aids clarity.

With this notation in hand, we can use the labor demand policy rules—(1) for the fixed costs case, (10) for the linear costs case—to construct a law of motion for the distribution function of actual employment $H(n)$ implied by each type of friction. We then show how these imply the laws of motion for the density $h(n)$ stated in equations (2) and (11) in the main text.

Fixed costs. Consider a point m in the domain of the employment distribution. We wish to derive the flows in and out of the mass $H(m)$. To do this, we first derive the flows for a given lagged employment level n_{-1}. Then inflows into $H(m)$ are summarized as follows:

1) If $m < L(n_{-1})$, or equivalently $n_{-1} > L^{-1}(m)$, then the inflow is equal to $\mathcal{H}^*(m|n_{-1})$.
2) If $m \in [L(n_{-1}), n_{-1})$, or equivalently $n_{-1} \in (m, L^{-1}(m)]$, then the inflow is equal to $\mathcal{H}^*(L(n_{-1})|n_{-1})$.

Likewise, the outflows from $H(m)$ for a given n_{-1} can be evaluated as:

3) If $m \in (n_{-1}, U(n_{-1})]$, or equivalently $n_{-1} \in [U^{-1}(m), m)$, then the outflow is equal to $1 - \mathcal{H}^*(U(n_{-1})|n_{-1})$.
4) If $m > U(n_{-1})$, or equivalently $n_{-1} < U^{-1}(m)$, then the outflows is equal to $1 - \mathcal{H}^*(m|n_{-1})$.

The latter are the flows of mass for a *given* lagged employment, n_{-1}. Integrating with respect to the distribution of lagged employment $H_{-1}(n_{-1})$ recovers the aggregate flows and thereby the law of motion for $H(m)$,

$$\Delta H(m) = \int_{L^{-1}(m)} \mathcal{H}^*(m|n_{-1}) dH_{-1}(n_{-1}) + \int_m^{L^{-1}(m)} \mathcal{H}^*(L(n_{-1})|n_{-1}) dH_{-1}(n_{-1})$$
$$- \int_{U^{-1}(m)}^{m} [1 - \mathcal{H}^*(U(n_{-1})|n_{-1})] dH_{-1}(n_{-1}) \tag{26}$$
$$- \int^{U^{-1}(m)} [1 - \mathcal{H}^*(m|n_{-1})] dH_{-1}(n_{-1}).$$

Linear costs. Likewise, one can use the adjustment rule for the linear costs case, (10), to construct an analogous law of motion for H under this friction. Again, we first fix a given level of lagged employment, n_{-1}, and evaluate inflows to, and outflows from, $H(m)$.

These flows are simpler in the linear costs case. Inflows are given by the following case:

1) If $m < n_{-1}$, or equivalently $n_{-1} > m$, then the inflow is equal to $\mathcal{H}^*(l(m)|n_{-1})$.

Similarly, outflows are given by:

2) If $m > n_{-1}$, or equivalently $n_{-1} < m$, then the outflow is equal to $1 - \mathcal{H}^*(u(m)|n_{-1})$.

Following the same logic as in the fixed costs case above, the law of motion for $H(m)$ is thus given by

$$\Delta H(m) = \int_m \mathcal{H}^*(l(m)|n_{-1}) dH_{-1}(n_{-1}) - \int^m [1 - \mathcal{H}^*(u(m)|n_{-1})] dH_{-1}(n_{-1}). \tag{27}$$

Laws of motion for $h(n)$. Differentiating (26) and (27) with respect to m, cancelling terms, and using Bayes' rule to note that $\int_0^\nu \hbar^*(\xi|n_{-1}) h_{-1}(n_{-1}) dn_{-1} = \int_0^\nu \hbar(n_{-1}|\xi) h^*(\xi) dn_{-1}$ yields the simpler laws of motion for the density of employment $h(n)$, equations (2) and (11) in the main text.

B. Proofs of Propositions 1 and 2

To establish Propositions 1 and 2 in the main text, it is convenient first to define a notion of *quasi-frictionless* employment. Lemma 1 then shows that the firm's problem can be normalized with respect to quasi-frictionless employment to establish some useful homogeneity properties. Using this homogeneous problem, we are able to relate the change in aggregate log flow steady-state employment $\Delta\widehat{\mathcal{N}}$ to the change in aggregate log quasi-frictionless employment. In a final step, we link the change in aggregate log quasi-frictionless employment to the change in aggregate log *frictionless* employment.

Definition *(i) Quasi-frictionless employment* n^* *solves* $px\alpha {n^*}^{\alpha-1} \equiv w$, *where* w *is the frictional equilibrium wage; and (ii) frictionless employment* n^{**} *solves* $px\alpha {n^{**}}^{\alpha-1} \equiv w^*$, *where* w^* *is the frictionless equilibrium wage.*

Remark *The change in aggregate log quasi-frictionless employment* $\Delta\mathcal{N}^*$ *induced by a change in aggregate productivity* $\Delta\ln p$ *is related to the change in aggregate log frictionless employment* $\Delta\mathcal{N}^{**}$ *according to*

$$\Delta\mathcal{N}^* = \frac{1-\epsilon_w}{1-\epsilon_{w^*}}\Delta\mathcal{N}^{**}, \tag{28}$$

where ϵ_w *and* ϵ_{w^*} *denote the elasticities of the equilibrium wage to aggregate productivity* p*, respectively with and without frictions.*

Lemma 1 (Caballero and Engel 1999) *Consider the firm's problem,*

$$\Pi(n_{-1},x) \equiv \max_n\{pxn^\alpha - wn - C^+\mathbb{I}[n > n_{-1}] - C^-\mathbb{I}[n < n_{-1}] - c^+\Delta n^+ + c^-\Delta n^- + \beta\mathbb{E}[\Pi(n,x')|x]\}. \tag{29}$$

If (i) $\ln x' = \ln x + \varepsilon_x'$ *with* ε_x' *i.i.d., and (ii)* $C^{-/+} = \Gamma^{-/+}wn^*$ *and* $c^{-/+} = \gamma^{-/+}w$*, then (a) the adjustment policy take the form*

$$n = \begin{cases} n^*/u & \text{if} \quad n^* > U\cdot n_{-1}, \\ n_{-1} & \text{if} \quad n^* \in [L\cdot n_{-1}, U\cdot n_{-1}], \\ n^*/l & \text{if} \quad n^* < L\cdot n_{-1}. \end{cases} \tag{30}$$

for constants $L \le l < 1 < u \le U$ *; and (b) desired (log) employment adjustments,* $\ln(n^*/n_{-1})$*, are independent of initial firm size* n_{-1}*.*

Proof of Lemma 1. Since idiosyncratic shocks follow a geometric random walk, $\ln x' = \ln x + \varepsilon_x'$, so does frictionless employment, $\ln n^{*\prime} = \ln n^* + \varepsilon_{n^*}'$ *where* $\varepsilon_{n^*}' = \varepsilon_x'/(1-\alpha)$. Defining $z = n/n^*$ and $\zeta = n_{-1}/n^*$, a conjecture that $\Pi(n_{-1},x) = wn^*\widetilde{\Pi}(\zeta)$ implies

$$\widetilde{\Pi}(\zeta) = \max_z\left\{\frac{z^\alpha}{\alpha} - z - \Gamma^+\mathbb{I}[z > \zeta] - \Gamma^-\mathbb{I}[z < \zeta] - \gamma^+(z-\zeta)^+ + \gamma^-(z-\zeta)^- + \beta\mathbb{E}\left[e^{\varepsilon_{n^*}'}\widetilde{\Pi}\left(e^{-\varepsilon_{n^*}'}z\right)\right]\right\}. \tag{31}$$

We highlight two aspects of (31). First, the expectation over the forward value is no longer conditional, since it is taken over ε_{n^*}', which is i.i.d. Second, the firm's problem is simplified to the choice of a number $z = n/n^*$ for each realization of the single state variable $\zeta = n_{-1}/n^*$.

An *Ss* policy will thus stipulate that $z = \zeta$ for intermediate values of $\zeta \in [1/U, 1/L]$, and will set $z = 1/u$ whenever $\zeta < 1/U$, and $z = 1/l$ whenever $\zeta > 1/L$. Mapping back into employment terms implies the adjustment policy in (30), establishing part a) of the result. Note that the case of pure fixed costs implies $u = l$, while pure linear costs imply $l = L < U = u$.

To establish part b), note that the probability of a desired log employment adjustment of size less than δ can be written, in general, as

$$\Pr(\ln(n^{*\prime}/n) < \delta|n) = \Pr(\varepsilon_{n^*}' < \delta + \ln z\,|n) = \int \Pr(\varepsilon_{n^*}' < \delta + \ln z\,|n,z)\,dZ(z|n), \tag{32}$$

where $Z(z|n)$ denotes the distribution function of z given n. In the context of the canonical model, however, (32) simplifies. First, ε_{n^*}' is independent of n since the former is i.i.d. Second, z is also independent of n. To see this, note first that if a firm adjusts this period, its choice of z is uninformed by n—it sets $z = 1/u$ or $z = 1/l$. If the firm sets $n = n_{-1}$ but adjusted *last* period, then it sets $\ln z = \ln n_{-1} - \ln n^* = \ln z_{-1} - \varepsilon_{n^*}$ and z_{-1} is $1/u$ or $1/l$. Thus, z is again independent of n. More generally, suppose the firm last adjusted T periods ago, that is, $n = n_{-1} = \cdots = n_{-T}$ and $z_{-T} = 1/u$ or $1/l$. Then, $\ln z = \ln n_{-T} - \ln n^* = \ln z_{-T} - \sum_{t=0}^{T-1} \varepsilon_{n_{-t}^*}$. Each term here is independent of $n = n_{-T}$. Equation (32) therefore collapses to

$$\Pr(\ln(n^{*\prime}/n) < \delta|n) = \int \Pr(\varepsilon_{n^*}' < \delta + \ln z\,|z)\, dZ(z), \tag{33}$$

which does not depend on n.

Proof of Proposition 1. Denoting log employment by $\mathcal{n}$, the adjustment rules take the form $L(\mathcal{n}) = \mathcal{n} - \lambda$ and $U(\mathcal{n}) = \mathcal{n} + \upsilon$ for $\lambda > 0$ and $\upsilon > 0$. The density of log employment in flow balance is then defined by

$$\hat{h}(\mathcal{n}) \equiv \frac{1 - \mathcal{H}(\mathcal{n} + \lambda|\mathcal{n}) + \mathcal{H}(\mathcal{n} - \upsilon|\mathcal{n})}{1 - \mathcal{H}^*(\mathcal{n} + \upsilon|\mathcal{n}) + \mathcal{H}^*(\mathcal{n} - \lambda|\mathcal{n})} h^*(\mathcal{n}), \tag{34}$$

where $\mathcal{H}^*(\xi|\nu) \equiv \Pr(\mathcal{n}^* < \xi|\mathcal{n}_{-1} = \nu)$ and $\mathcal{H}(\nu|\xi) \equiv \Pr(\mathcal{n}_{-1} < \nu|\mathcal{n}^* = \xi)$. The property of the canonical model noted in result b) of Lemma 1, that $\mathcal{n}^* - \mathcal{n}_{-1}$ is independent of $\mathcal{n}_{-1}$, implies that

$$\mathcal{H}^*(\xi|\nu) = \Pr(\mathcal{n}^* - \mathcal{n}_{-1} < \xi - \nu) \equiv \widetilde{\mathcal{H}}^*(\xi - \nu). \tag{35}$$

This implies that the probability of adjusting away from $\mathcal{n}$ is independent of $\mathcal{n}$,

$$1 - \mathcal{H}^*(\mathcal{n} + \upsilon|\mathcal{n}) + \mathcal{H}^*(\mathcal{n} - \lambda|\mathcal{n}) = 1 - \int_{-\lambda}^{\upsilon} \tilde{h}^*(z)dz \equiv \phi. \tag{36}$$

Now consider the probability of adjusting to $\mathcal{n}$. Using Bayes' rule, we can write this as

$$\begin{aligned} 1 - \mathcal{H}(\mathcal{n} + \lambda|\mathcal{n}) + \mathcal{H}(\mathcal{n} - \upsilon|\mathcal{n}) &= 1 - \int_{\mathcal{n}-\upsilon}^{\mathcal{n}+\lambda} h^*(\mathcal{n}|\nu)\frac{h_{-1}(\nu)}{h^*(\mathcal{n})}d\nu \\ &= 1 - \int_{\mathcal{n}-\upsilon}^{\mathcal{n}+\lambda} \tilde{h}^*(\mathcal{n} - \nu)\frac{h_{-1}(\nu)}{h^*(\mathcal{n})}d\nu \\ &= 1 - \int_{-\lambda}^{\upsilon} \tilde{h}^*(z)\frac{h_{-1}(\mathcal{n} - z)}{h^*(\mathcal{n})}dz. \end{aligned} \tag{37}$$

Piecing this together, we have

$$\hat{h}(\mathcal{n}) = \frac{h^*(\mathcal{n}) - \int_{-\lambda}^{\upsilon} \tilde{h}^*(z)h_{-1}(\mathcal{n} - z)dz}{1 - \int_{-\lambda}^{\upsilon} \tilde{h}^*(z)dz}. \tag{38}$$

Multiplying both sides by $\mathcal{n}$ and integrating yields

$$\begin{aligned}
\widehat{\mathcal{N}} \equiv \int_{-\infty}^{\infty} n\hat{h}(n)dn &= \frac{\mathcal{N}^*}{\phi} - \frac{1}{\phi}\int_{-\infty}^{\infty}\int_{-\lambda}^{v} n\tilde{\hbar}^*(z)h_{-1}(n-z)dzdn \\
&= \frac{\mathcal{N}^*}{\phi} - \frac{1}{\phi}\int_{-\lambda}^{v} \tilde{\hbar}^*(z)\int_{-\infty}^{\infty} nh_{-1}(n-z)dn\,dz \\
&= \frac{\mathcal{N}^*}{\phi} - \frac{1}{\phi}\int_{-\lambda}^{v} \tilde{\hbar}^*(z)(\mathcal{N}_{-1}+z)dz \\
&= \frac{\mathcal{N}^*}{\phi} - \frac{1-\phi}{\phi}\mathcal{N}_{-1} - \frac{1}{\phi}\int_{-\lambda}^{v} z\tilde{\hbar}^*(z)dz.
\end{aligned} \tag{39}$$

Since there is a constant-$\mathcal{N}^*$ state prior to the aggregate shock, aggregate log employment is constant and equal to aggregate flow-balance employment, $\mathcal{N}_{-1} = \mathcal{N}_{-2} = \widehat{\mathcal{N}}_{-1}$. Imposing this and solving for $\widehat{\mathcal{N}}_{-1}$ yields

$$\widehat{\mathcal{N}}_{-1} = \mathcal{N}_{-1}^* - \int_{-\lambda}^{v} z\tilde{\hbar}_{-1}^*(z)dz. \tag{40}$$

Now consider a shock to aggregate log mandated employment, $\Delta\mathcal{N}^*$. On impact this will shift the mean of the distribution of desired employment adjustments, $\tilde{\hbar}^*(\cdot)$, by $\Delta\mathcal{N}^*$. Starting from the prior constant-$\mathcal{N}^*$ state, substitution of (40) into (39) implies

$$\Delta\widehat{\mathcal{N}} = \frac{\Delta\mathcal{N}^*}{\phi} - \frac{1}{\phi}\int_{-\lambda}^{v} z\Delta\tilde{\hbar}^*(z)dz. \tag{41}$$

To a first-order approximation around $\Delta\mathcal{N}^* = 0$,

$$\begin{aligned}
\int_{-\lambda}^{v} z\Delta\tilde{\hbar}^*(z)dz &= \int_{-\lambda}^{v} z\tilde{\hbar}_{-1}^*(z-\Delta\mathcal{N}^*)dz - \int_{-\lambda}^{v} z\tilde{\hbar}_{-1}^*(z)dz \\
&\approx -\int_{-\lambda}^{v} z\tilde{\hbar}_{-1}^{*\prime}(z)dz\,\Delta\mathcal{N}^* \\
&= \left[1-\phi-v\tilde{\hbar}_{-1}^*(v)-\lambda\tilde{\hbar}_{-1}^*(-\lambda)\right]\Delta\mathcal{N}^*.
\end{aligned} \tag{42}$$

The fact the latter is smaller in magnitude than $\Delta\mathcal{N}^*$, and equation (28), imply the stated result.

Proof of Proposition 2. The proof proceeds in the same way as the proof of Proposition 1 above. The adjustment rules again take the form $l(n) = n - \lambda$ and $u(n) = n + v$ for $\lambda > 0$ and $v > 0$. The density of log employment in flow balance is then defined by

$$\hat{h}(n) \equiv \frac{[1-\mathcal{H}(n|n-\lambda)]h^*(n-\lambda) + \mathcal{H}(n|n+v)h^*(n+v)}{1-\mathcal{H}^*(n+v|n)+\mathcal{H}^*(n-\lambda|n)}. \tag{43}$$

Since $\mathcal{H}^*(\xi|\nu) = \Pr(n^* - n_{-1} < \xi - \nu) \equiv \widetilde{\mathcal{H}}^*(\xi - \nu)$, the probability of adjusting away from n is again independent of n,

$$1-\mathcal{H}^*(n+v|n)+\mathcal{H}^*(n-\lambda|n) = 1 - \int_{-\lambda}^{v} \tilde{\hbar}^*(z)dz \equiv \phi. \tag{44}$$

Now consider the probabilities of adjusting down and up to n. Using Bayes' rule, we can write these as

$$
\begin{aligned}
1-\mathcal{H}(n|n-\lambda) &= \int_{n}^{\infty} \hbar(\nu|n-\lambda)d\nu \\
&= \int_{n}^{\infty} \hbar^{*}(n-\lambda|n)\frac{h_{-1}(\nu)}{h^{*}(n-\lambda)}d\nu \\
&= \int_{n}^{\infty} \tilde{\hbar}^{*}(n-\lambda-\nu)\frac{h_{-1}(\nu)}{h^{*}(n-\lambda)}d\nu \\
&= \int_{-\infty}^{-\lambda} \tilde{\hbar}^{*}(z)\frac{h_{-1}(n-\lambda-z)}{h^{*}(n-\lambda)}dz \text{, and}
\end{aligned} \tag{45}
$$

$$
\begin{aligned}
\mathcal{H}(n|n+\upsilon) &= \int_{-\infty}^{n} \hbar(\nu|n+\upsilon)d\nu \\
&= \int_{-\infty}^{n} \hbar^{*}(n+\upsilon|\nu)\frac{h_{-1}(\nu)}{h^{*}(n+\upsilon)}d\nu \\
&= \int_{-\infty}^{n} \tilde{\hbar}^{*}(n+\upsilon-\nu)\frac{h_{-1}(\nu)}{h^{*}(n+\upsilon)}d\nu \\
&= \int_{\upsilon}^{\infty} \tilde{\hbar}^{*}(z)\frac{h_{-1}(n+\upsilon-z)}{h^{*}(n+\upsilon)}dz.
\end{aligned} \tag{46}
$$

Piecing this together, we have

$$
\hat{h}(n) = \frac{\int_{-\infty}^{-\lambda} \tilde{\hbar}^{*}(z)h_{-1}(n-\lambda-z)dz + \int_{\upsilon}^{\infty} \tilde{\hbar}^{*}(z)h_{-1}(n+\upsilon-z)dz}{1-\int_{-\lambda}^{\upsilon} \tilde{\hbar}^{*}(z)dz}. \tag{47}
$$

Multiplying both sides by n and integrating yields

$$
\begin{aligned}
\widehat{\mathcal{N}} &= \frac{1}{\phi}\int_{-\infty}^{-\lambda} \tilde{\hbar}^{*}(z)\int_{-\infty}^{\infty} nh_{-1}(n-\lambda-z)dn\,dz + \frac{1}{\phi}\int_{\upsilon}^{\infty} \tilde{\hbar}^{*}(z)\int_{-\infty}^{\infty} nh_{-1}(n+\upsilon-z)dn\,dz \\
&= \frac{1}{\phi}\int_{-\infty}^{-\lambda} \tilde{\hbar}^{*}(z)(\mathcal{N}_{-1}+\lambda+z)dz + \frac{1}{\phi}\int_{\upsilon}^{\infty} \tilde{\hbar}^{*}(z)(\mathcal{N}_{-1}-\upsilon+z)dz \\
&= \mathcal{N}_{-1} + \frac{1}{\phi}\left[\lambda\int_{-\infty}^{-\lambda} \tilde{\hbar}^{*}(z)dz - \upsilon\int_{\upsilon}^{\infty} \tilde{\hbar}^{*}(z)dz\right] + \frac{1}{\phi}\int_{-\infty}^{\infty} z\tilde{\hbar}^{*}(z)dz - \frac{1}{\phi}\int_{-\lambda}^{\upsilon} z\tilde{\hbar}^{*}(z)dz \\
&= \frac{\mathcal{N}^{*}}{\phi} - \frac{1-\phi}{\phi}\mathcal{N}_{-1} + \frac{1}{\phi}\left[\lambda\int_{-\infty}^{-\lambda} \tilde{\hbar}^{*}(z)dz - \upsilon\int_{\upsilon}^{\infty} \tilde{\hbar}^{*}(z)dz\right] - \frac{1}{\phi}\int_{-\lambda}^{\upsilon} z\tilde{\hbar}^{*}(z)dz.
\end{aligned} \tag{48}
$$

Solving for $\widehat{\mathcal{N}}_{-1} = \mathcal{N}_{-1} = \mathcal{N}_{-2}$ in the prior constant-$\mathcal{N}^{*}$ state yields

$$
\widehat{\mathcal{N}}_{-1} = \mathcal{N}_{-1}^{*} + \lambda\int_{-\infty}^{-\lambda} \tilde{\hbar}_{-1}^{*}(z)dz - \upsilon\int_{\upsilon}^{\infty} \tilde{\hbar}_{-1}^{*}(z)dz - \int_{-\lambda}^{\upsilon} z\tilde{\hbar}_{-1}^{*}(z)dz. \tag{49}
$$

Substitution of (49) into (48) implies that a shock to aggregate log mandated employment that shifts the mean of $\tilde{\hbar}^{*}(\cdot)$ by $\Delta\mathcal{N}^{*}$ will induce a change in $\widehat{\mathcal{N}}$ relative to the prior constant-$\mathcal{N}^{*}$ state equal to

$$
\Delta\widehat{\mathcal{N}} = \frac{\Delta\mathcal{N}^{*}}{\phi} + \frac{1}{\phi}\left[\lambda\int_{-\infty}^{-\lambda} \Delta\tilde{\hbar}^{*}(z)dz - \upsilon\int_{\upsilon}^{\infty} \Delta\tilde{\hbar}^{*}(z)dz\right] - \frac{1}{\phi}\int_{-\lambda}^{\upsilon} z\Delta\tilde{\hbar}^{*}(z)dz. \tag{50}
$$

To a first-order approximation around $\Delta\mathcal{N}^{*} = 0$,

$$\int_{-\infty}^{-\lambda} \Delta\tilde{\hbar}^*(z)dz = \int_{-\infty}^{-\lambda} \tilde{\hbar}_{-1}^*(z-\Delta\mathcal{N}^*)dz - \int_{-\infty}^{-\lambda} \tilde{\hbar}_{-1}^*(z)dz \approx -\tilde{\hbar}_{-1}^*(-\lambda)\Delta\mathcal{N}^*,$$
$$\int_{v}^{\infty} \Delta\tilde{\hbar}^*(z)dz = \int_{v}^{\infty} \tilde{\hbar}_{-1}^*(z-\Delta\mathcal{N}^*)dz - \int_{v}^{\infty} \tilde{\hbar}_{-1}^*(z)dz \approx \tilde{\hbar}_{-1}^*(v)\Delta\mathcal{N}^*, \tag{51}$$

and, as in equation (42) above, $\int_{-\lambda}^{v} z\Delta\tilde{\hbar}^*(z)dz \approx \left[1-\phi - v\tilde{\hbar}_{-1}^*(v) - \lambda\tilde{\hbar}_{-1}^*(-\lambda)\right]\Delta\mathcal{N}^*$.

Substitution of these into equation (50) implies $\Delta\widehat{\mathcal{N}} \approx \Delta\mathcal{N}^*$. Combining with equation (28), yields the stated result.

C. Large-firm canonical search and matching model

The firm's problem for this model combines equations (16) and (17) in the main text to obtain:

$$\Pi(n_{-1},x) \equiv \max_n \left\{ Apxn^\alpha - (1-\eta)\omega n - \frac{c}{q(\theta)}\Delta n^+ + \beta\mathbb{E}[\Pi(n,x')|x] \right\},$$
$$\text{where } A \equiv \frac{1-\eta}{1-\eta(1-\alpha)}. \tag{52}$$

To establish Proposition 3 in the main text, we proceed as above.

Definition *(i) Quasi-frictionless employment n^* solves $Apx\alpha n^{*\alpha-1} \equiv (1-\eta)\omega$, where ω is the worker's outside option; and (ii) frictionless employment n^{**} solves $px\alpha n^{**\alpha-1} \equiv w^*$, where w^* is the frictionless equilibrium wage.*

Remark *The change in aggregate log quasi-frictionless employment $\Delta\mathcal{N}^*$ induced by a change in aggregate productivity $\Delta \ln p$ is related to the change in aggregate log frictionless employment $\Delta\mathcal{N}^{**}$ according to*

$$\Delta\mathcal{N}^* = \frac{1-\epsilon_\omega}{1-\epsilon_{w^*}}\Delta\mathcal{N}^{**}, \tag{53}$$

where ϵ_ω and ϵ_{w^*} respectively denote the elasticities of the worker's outside option ω and the frictionless wage w^* to aggregate productivity p.

Lemma 1′ *If (i) $\ln x' = \ln x + \varepsilon_x'$ with ε_x' i.i.d., and (ii) $c = \gamma(1-\eta)\omega$, then (a) the adjustment triggers take the form in (10), are linear, $l(n) = l\cdot n$ and $u(n) = u\cdot n$ for time-varying $l < 1 < u$; and (b) desired (log) employment adjustments, $\ln(n^*/n_{-1})$, are independent of initial firm size n_{-1}.*

Proof. Note that a conjecture that $\Pi(n_{-1},x) = (1-\eta)\omega n^*\tilde{\Pi}(\zeta)$ yields

$$\tilde{\Pi}(\zeta) \equiv \max_z \left\{ \frac{z^\alpha}{\alpha} - z - \frac{\gamma}{q(\theta)}(z-\zeta)^+ + \beta\mathbb{E}\left[e^{\varepsilon_{n^*}'}\tilde{\Pi}\left(e^{-\varepsilon_{n^*}'}z\right)\right]\right\}. \tag{54}$$

Results (a) and (b) follow from the proof to Lemma 1 above.

Lemma 2 *If (i) the adjustment triggers are symmetric,* $-\ln l = \ln u \equiv \mu$*, and (ii) the distribution of innovations* ε_{n^*} *is symmetric,* $\mathcal{E}(-\varepsilon_{n^*}) = 1 - \mathcal{E}(\varepsilon_{n^*})$*, then the distribution of desired (log) employment adjustments* $\ln(n^*/n_{-1})$ *is symmetric,* $\widetilde{\mathcal{H}}^*(-\varsigma) = 1 - \widetilde{\mathcal{H}}^*(\varsigma)$*.*

Proof. Note first that the distribution of the desired log change in employment, $n^* - n_{-1}$, conditional on last period's log gap, $z_{-1} = n_{-1} - n^*_{-1}$, takes the simple form $\Pr(n^* - n_{-1} < \varsigma | z_{-1}) = \mathcal{E}(\varsigma - z_{-1})$, since $\varepsilon_{n^*} \equiv n^* - n^*_{-1}$ is i.i.d. with distribution function $\mathcal{E}(\cdot)$. It follows that the unconditional distribution of $n^* - n_{-1}$ is

$$\widetilde{\mathcal{H}}^*(\varsigma) = \int_{-\mu}^{\mu} \Pr(n^* - n_{-1} < \varsigma | z_{-1})\, g(z_{-1}) \mathrm{d}z_{-1} = \int_{-\mu}^{\mu} \mathcal{E}(\varsigma - z_{-1}) g(z_{-1}) \mathrm{d}z_{-1}, \tag{55}$$

where $g(z_{-1})$ is the ergodic density of z_{-1}. It is simple to verify that $\mathcal{E}(-\varepsilon_{n^*}) = 1 - \mathcal{E}(\varepsilon_{n^*})$ implies $\widetilde{\mathcal{H}}^*(\varsigma) = 1 - \widetilde{\mathcal{H}}^*(-\varsigma)$, provided $g(\cdot)$ also is symmetric, which we now establish.

Our strategy is to conjecture that $g(\cdot)$ is symmetric and verify that this is implied. Consider a firm with an initial $z_{-1} = z - \varepsilon$ such that $z \in (-\mu, \mu)$ lies strictly inside the inaction range. Clearly, this firm migrates to z if it draws ε. Thus, the mass of firms at z this period is given by

$$g(z) = \int_{z-\mu}^{z+\mu} g(z - \varepsilon)\, \mathrm{d}\mathcal{E}(\varepsilon) = \int_{-\mu}^{\mu} g(y)\, \mathrm{d}\mathcal{E}(z - y), \tag{56}$$

where we have used the change of variable $y = z - \varepsilon$. Under the conjecture that $g(y) = g(-y)$, one can confirm $g(z) = g(-z)$. To see this, evaluate $g(\cdot)$ at $-z$, use symmetry of $\mathcal{E}(\cdot)$, a change of variable $\tilde{y} = -y$, and standard rules of calculus to obtain

$$\begin{aligned} g(-z) &= \int_{-\mu}^{\mu} g(y)\, \mathrm{d}\mathcal{E}(-z - y) = \int_{-\mu}^{\mu} g(y)\, \mathrm{d}\mathcal{E}(z + y) = -\int_{\mu}^{-\mu} g(-\tilde{y})\, \mathrm{d}\mathcal{E}(z - \tilde{y}) \\ &= \int_{-\mu}^{\mu} g(-y)\, \mathrm{d}\mathcal{E}(z - y). \end{aligned} \tag{57}$$

Now consider the mass at the lower adjustment barrier, $z = -\mu$. This is comprised of two parts: first, firms that begin at $-\mu$, draw a negative labor demand shock ($\varepsilon < 0$), and adjust to remain at $-\mu$; and second, firms that began away from $-\mu$ and then migrate there. Thus,

$$g(-\mu) = \mathcal{E}(0) g(-\mu) + \int_0^{2\mu} g(-\mu + \varepsilon) \mathrm{d}\mathcal{E}(-\varepsilon) = \frac{1}{\mathcal{E}(0)} \int_0^{2\mu} g(-\mu + \varepsilon) \mathrm{d}\mathcal{E}(\varepsilon), \tag{58}$$

where the second equality follows from symmetry of $\mathcal{E}(\cdot)$. A similar argument can be used to show that the mass at the upper adjustment barrier $z = \mu$ satisfies

$$g(\mu) = \frac{1}{\mathcal{E}(0)} \int_0^{2\mu} g(\mu - \varepsilon) \mathrm{d}\mathcal{E}(\varepsilon). \tag{59}$$

A conjecture of symmetry $g(-\mu + \varepsilon) = g(\mu - \varepsilon)$ is again confirmed, $g(-\mu) = g(\mu)$. It follows that $g(z) = g(-z)$ for all $z \in [-\mu, \mu]$, and symmetry of $\widetilde{\mathcal{H}}^*(\cdot)$ obtains.

Proof of Proposition 3. The first-order conditions that define the triggers for optimal adjustment $z \in \{1/u\,, 1/l\}$ are given by

$$\begin{aligned} l^{1-\alpha} + \beta D(1/l\,;\theta) &\equiv 1, \\ u^{1-\alpha} + \beta D(1/u\,;\theta) &\equiv 1 + \frac{\gamma}{q(\theta)}, \end{aligned} \tag{60}$$

where $D(z;\theta) \equiv \mathbb{E}\left[\widetilde{\Pi}'\left(e^{-\varepsilon'_{n^*}} z\right)\right]$. The latter satisfies the following recursion

$$\begin{aligned} D(z;\theta) = &\int_{\ln(lz)}^{\ln(uz)} \left[e^{(1-\alpha)\varepsilon'_{n^*}} z^{\alpha-1} - 1 + \beta D\left(e^{-\varepsilon'_{n^*}} z;\theta\right)\right] d\mathcal{E}(\varepsilon'_{n^*}) \\ &+ \frac{\gamma}{q(\theta)}[1 - \mathcal{E}(\ln(uz))]. \end{aligned} \tag{61}$$

We first consider a first-order approximation to the firm's optimal policies around $\gamma = 0$.[22] To this end, note first that

$$\begin{aligned} D_\gamma(z;\theta) &\approx \frac{1}{q(\theta)}[1 - \mathcal{E}(\ln(uz))] + \beta \int_{\ln(lz)}^{\ln(uz)} D_\gamma\left(e^{-\varepsilon'_{n^*}} z;\theta\right) d\mathcal{E}(\varepsilon'_{n^*}) \\ &= \frac{1}{q(\theta)}[1 - \mathcal{E}(\ln z)] \text{ when } \gamma = 0 \end{aligned} \tag{62}$$

Thus we can write $D(z;\theta) \approx \gamma[1 - \mathcal{E}(\ln z)]/q(\theta)$. Substituting into the first-order conditions and noting that $l = e^{-\lambda}$ and $u = e^{\upsilon}$ yields

$$\begin{aligned} e^{-(1-\alpha)\lambda} + \beta \frac{\gamma}{q(\theta)}[1 - \mathcal{E}(\lambda)] &\approx 1, \\ e^{(1-\alpha)\upsilon} + \beta \frac{\gamma}{q(\theta)}[1 - \mathcal{E}(-\upsilon)] &\approx 1 + \frac{\gamma}{q(\theta)}. \end{aligned} \tag{63}$$

Next, linearizing the leading terms around $\lambda = 0$ and $\upsilon = 0$, respectively, leads to

$$\begin{aligned} -(1-\alpha)\lambda + \beta \frac{\gamma}{q(\theta)}[1 - \mathcal{E}(\lambda)] &\approx 0, \\ (1-\alpha)\upsilon + \beta \frac{\gamma}{q(\theta)}[1 - \mathcal{E}(-\upsilon)] &\approx \frac{\gamma}{q(\theta)}. \end{aligned} \tag{64}$$

Imposing $\beta \approx 1$, and $\mathcal{E}(-\varepsilon) = 1 - \mathcal{E}(\varepsilon)$ yields $\lambda \approx \upsilon$.

Now return to the relationship between $\Delta\widehat{\mathcal{N}}$ and $\Delta\mathcal{N}^*$ in equation (50). Time-variation in the adjustment triggers alters the approximations around small aggregate shocks. Specifically, with $\lambda \approx \upsilon \approx \mu$, equation (50) becomes

[22] Equation (52) has the form $D(z) = \mathcal{C}(D,\gamma)(z)$, where $\mathcal{C}$ is a contraction map on the cross product of the space of bounded and continuous functions (where D "lives") and $[0,\Gamma]$, a closed subinterval of the nonnegative real line from which γ is drawn. By inspection, this map is continuously differentiable with respect to (w.r.t.) $\gamma \in [0,\Gamma]$. It then follows from Lemma 1 of Albrecht, Holmlund, and Lang (1991) that D is continuously differentiable w.r.t. γ and satisfies the recursion, $D_\gamma(z) = \mathcal{C}_\gamma(D,\gamma)(z) + \mathcal{C}_D(D_\gamma, D, \gamma)(z)$, where $\mathcal{C}_D$ is the Frechet derivative of $\mathcal{C}$. The right side of the latter expression defines a(nother) contraction map on a space of bounded and continuous functions. We have, then, that D_γ is bounded and continuous on $[0,\Gamma]$. Its calculation in (6262) follows.

$$\Delta\widehat{\mathcal{N}} \approx \frac{\Delta\mathcal{N}^*}{\phi} + \frac{1}{\phi}\left[\Delta\left(\mu\int_{-\infty}^{-\mu}\tilde{\hbar}^*(z)dz\right) - \Delta\left(\mu\int_{\mu}^{\infty}\tilde{\hbar}^*(z)dz\right)\right] - \frac{1}{\phi}\Delta\left(\int_{-\mu}^{\mu} z\tilde{\hbar}^*(z)dz\right). \quad (65)$$

Taking first-order approximations around $\Delta\mathcal{N}^* = 0$,

$$\begin{aligned}\Delta\left(\mu\int_{-\infty}^{-\mu}\tilde{\hbar}^*(z)dz\right) &= \mu\int_{-\infty}^{-\mu}\tilde{\hbar}^*_{-1}(z-\Delta\mathcal{N}^*)dz - \mu_{-1}\int_{-\infty}^{-\mu_{-1}}\tilde{\hbar}^*_{-1}(z)dz \\ &\approx \left[-\mu_{-1}\tilde{\hbar}^*_{-1}(-\mu_{-1}) + \left\{\widetilde{\mathcal{H}}^*_{-1}(-\mu_{-1}) - \mu_{-1}\tilde{\hbar}^*_{-1}(-\mu_{-1})\right\}\frac{\partial\mu}{\partial\Delta\mathcal{N}^*}\right]\Delta\mathcal{N}^*;\end{aligned} \quad (66)$$

similarly,

$$\begin{aligned}\Delta\left(\mu\int_{\mu}^{\infty}\tilde{\hbar}^*(z)dz\right) &= \mu\int_{\mu}^{\infty}\tilde{\hbar}^*_{-1}(z-\Delta\mathcal{N}^*)dz - \mu_{-1}\int_{\mu_{-1}}^{\infty}\tilde{\hbar}^*_{-1}(z)dz \\ &\approx \left[\mu_{-1}\tilde{\hbar}^*_{-1}(\mu_{-1}) + \left\{1-\widetilde{\mathcal{H}}^*_{-1}(\mu_{-1}) - \mu_{-1}\tilde{\hbar}^*_{-1}(\mu_{-1})\right\}\frac{\partial\mu}{\partial\Delta\mathcal{N}^*}\right]\Delta\mathcal{N}^*;\end{aligned} \quad (59)$$

and,

$$\begin{aligned}\Delta\left(\int_{-\mu}^{\mu} z\tilde{\hbar}^*(z)dz\right) &= \int_{-\mu}^{\mu} z\tilde{\hbar}^*_{-1}(z-\Delta\mathcal{N}^*)dz - \int_{-\mu_{-1}}^{\mu_{-1}} z\tilde{\hbar}^*_{-1}(z)dz \\ &\approx \Big[1-\phi-\mu_{-1}\tilde{\hbar}^*_{-1}(\mu_{-1}) - \mu_{-1}\tilde{\hbar}^*_{-1}(-\mu_{-1}) \\ &\quad - \left(\mu_{-1}\tilde{\hbar}^*_{-1}(-\mu_{-1})\frac{\partial\mu}{\partial\Delta\mathcal{N}^*}\right) + \left(\mu_{-1}\tilde{\hbar}^*_{-1}(\mu_{-1})\frac{\partial\mu}{\partial\Delta\mathcal{N}^*}\right)\Big]\Delta\mathcal{N}^*.\end{aligned} \quad (67)$$

Substituting these approximations back into (65) and cancelling yields

$$\Delta\widehat{\mathcal{N}} \approx \Delta\mathcal{N}^* + \frac{1}{\phi}\left\{\widetilde{\mathcal{H}}^*_{-1}(-\mu_{-1}) - \left[1-\widetilde{\mathcal{H}}^*_{-1}(\mu_{-1})\right]\right\}\frac{\partial\mu}{\partial\Delta\mathcal{N}^*}\Delta\mathcal{N}^*. \quad (68)$$

To complete the proof, note from Lemma 2 that symmetry of the adjustment barriers, and of $\mathcal{E}(\cdot)$, implies that $\widetilde{\mathcal{H}}^*(\cdot)$ is also symmetric. It follows that $\widetilde{\mathcal{H}}^*_{-1}(-\mu_{-1}) - \left[1-\widetilde{\mathcal{H}}^*_{-1}(\mu_{-1})\right] \approx 0$, and (68) collapses to $\Delta\widehat{\mathcal{N}} \approx \Delta\mathcal{N}^*$. The result then follows from equation (53).

D. Sensitivity analyses

This appendix describes our parameterization of the process for idiosyncratic shocks, contrasts it with alternatives in the literature, and performs sensitivity analyses on our main results to reasonable changes in model parameters.

Idiosyncratic shock process. Recall that we specify the process of idiosyncratic productivity as a geometric AR(1),

$$\ln x_{t+1} = \rho_x \ln x_t + \varepsilon_{xt+1}, \quad (69)$$

where $\varepsilon_{xt+1} \sim N(0, \sigma_x^2)$ and t denotes *quarters*. In what follows, we infer ρ_x and σ_x using estimates of the persistence and volatility of idiosyncratic productivity from data of *different* frequencies.

Annual data. Abraham and White (2006) use annual plant-level data from the U.S. manufacturing sector. Their measure of log idiosyncratic productivity would thus map most naturally to $\ln \sum_{t=1}^{4} x_t$, the log of the annual sum of quarterly realizations. For tractability, though, we will interpret the annual data as corresponding to $X_4 \equiv \sum_{t=1}^{4} \ln x_t$, and flag this violation of Jensen's inequality as a caveat to our results.

Abraham and White run least squares regressions of X_4 on $X_0 \equiv \sum_{t=-3}^{0} \ln x_t$. An average of the estimated slope coefficients from their weighted and unweighted regressions is

$$\frac{cov(X_4, X_0)}{var(X_0)} = 0.39. \tag{70}$$

Under the assumption that the annual data are generated from the quarterly process (69), one can relate the latter covariance and variance terms to ρ_x and σ_x as follows,

$$\begin{aligned} cov(X_4, X_0) &= cov\left(\sum_{t=1}^{4} \ln x_t, \sum_{t=-3}^{0} \ln x_t\right) \\ &= [\rho_x + 2\rho_x^2 + 3\rho_x^3 + 4\rho_x^4 + 3\rho_x^5 + 2\rho_x^6 + \rho_x^7] \frac{\sigma_x^2}{1-\rho_x^2}, \end{aligned} \tag{71}$$

and,

$$var(X_0) = [4 + 6\rho_x + 4\rho_x^2 + 2\rho_x^3] \frac{\sigma_x^2}{1-\rho_x^2}. \tag{72}$$

Solving for the implied quarterly persistence yields $\rho_x = 0.68$. To infer σ_x we use Abraham and White's estimates to infer that $var(X_0) = 0.21$, which in turn implies that $\sigma_x = 0.1034$.

Quarterly data. Cooper, Haltiwanger, and Willis (2015) estimate a dynamic labor demand model using quarterly data. Like us, they specify a geometric AR(1) for idiosyncratic productivity. They estimate a persistence parameter of $\rho_x = 0.4$ and an innovation standard deviation of $\sigma_x = 0.5$.

Monthly data. Cooper, Haltiwanger, and Willis (2007) estimate a search and matching model using monthly data. They specify an idiosyncratic productivity process that follows a geometric AR(1) at a *monthly* frequency. Their estimates of the monthly analogues to ρ_x and σ_x are $\rho_x^m = 0.395$ and $\sigma_x^m = 0.0449$.[23]

Under the assumption that quarterly data are generated by the monthly AR(1) estimated by Cooper, Haltiwanger and Willis, the above logic implies that ρ_x can be computed as

$$\rho_x = \frac{\rho_x^m + 2(\rho_x^m)^2 + 3(\rho_x^m)^3 + 2(\rho_x^m)^4 + (\rho_x^m)^5}{3 + 4\rho_x^m + 2(\rho_x^m)^2} = 0.194. \tag{73}$$

Likewise, the variance of the innovation to the quarterly process is

[23] These average Cooper et al.'s results for two specifications of adjustment costs that seem most germane to our application: (i) a fixed cost to adjust and linear cost to hire, and (ii) a fixed cost to adjust and linear cost to fire.

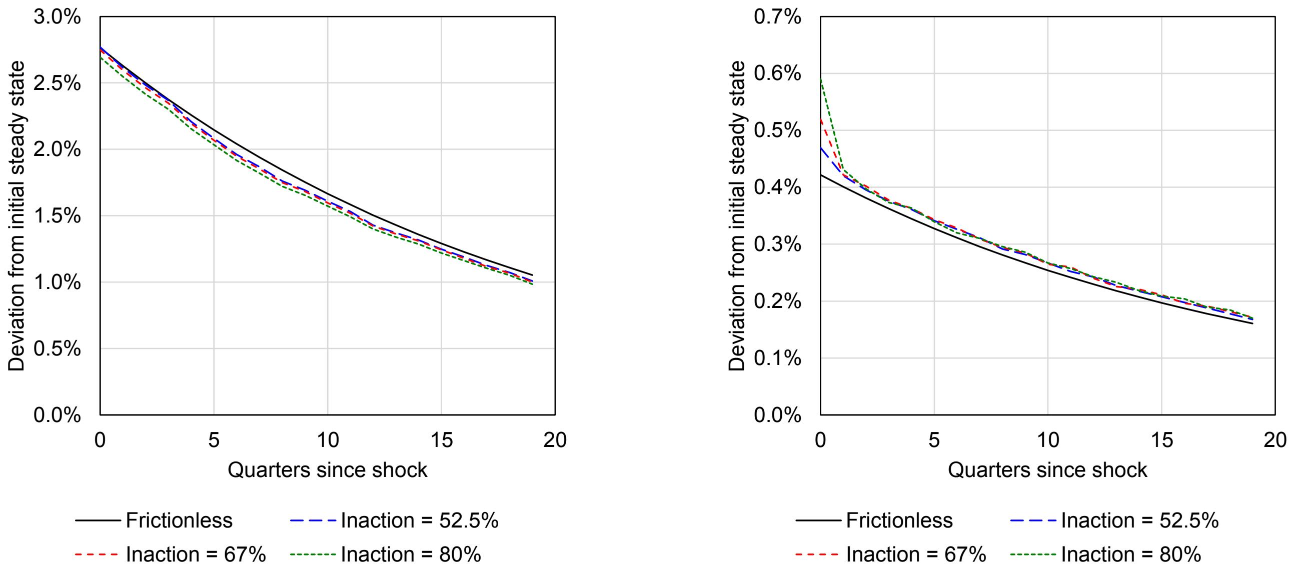
$$\sigma_x^2 = (1-\rho_x^2)\frac{3+4\rho_x^m+2(\rho_x^m)^2}{1-(\rho_x^m)^2}(\sigma_x^m)^2 = 0.251. \tag{74}$$

This implies a standard deviation $\sigma_x = 0.501$.

E. Additional figures

Figure A. Impulse responses of mandated and frictionless employment: Fixed costs

A. Fixed wages B. Equilibrium wage adjustment



Notes: Each panel plots the impulse response of aggregate frictionless employment in contrast to the impulse responses of mandated employment for each stated average quarterly inaction rate.