# The Macroeconomic Dynamics of Labor Market Policies*


Erik Hurst
Chicago Booth
Erik.Hurst@chicagobooth.edu

Patrick Kehoe
Stanford and Minneapolis Fed
pkehoe@stanford.edu

Elena Pastorino
Stanford and Hoover Institution
epastori@stanford.edu

Thomas Winberry
Wharton
twinb@wharton.upenn.edu


March 2025


**Abstract**

We develop a dynamic macroeconomic framework with worker heterogeneity, putty-clay adjustment frictions, and firm monopsony power to study the distributional impact of labor market policies over time. Our framework reconciles the well-known tension between low short-run and high long-run elasticities of substitution across inputs of production, especially among workers with different skills within a same education group. We use this framework to evaluate the effects of redistributive policies such as the minimum wage and the Earned Income Tax Credit. We argue that since these policies generate slow transition dynamics that can differ greatly in the short and long run, a serious assessment of their overall impact must take account of the entire time path of the responses they induce.



*We especially thank Mark Bils, Ricardo Caballero, John Campbell, Jeff Clemens, Jan Eeckhout, Eric French, Jonathan Heathcote, Dirk Krueger, Per Krusell, Marianna Kudlyak, Thomas Lemieux, Magne Mogstad, Simon Mongey, Lee Ohanian, Richard Rogerson, Ayşegul Şahin, Kjetil Storesletten, Gianluca Violante, and seminar participants at the Bank of Portugal, Federal Reserve Bank of Atlanta, Philadelphia, Richmond, and San Francisco, Berkeley, Chicago, Cornell, Harvard, MIT, Michigan, Minnesota Macro Summer 2022 Conference, NBER Summer Institute and Fall EFG 2022 Meetings, Northwestern, Princeton, NYU, Stanford, UCLA, UCSD, UC Santa Cruz, USC, Wharton, and Wisconsin. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System. A prior version of this paper circulated as "The Distributional Impact of the Minimum Wage in the Short and Long Run."

The high degree of wage inequality in the United States, coupled with growing evidence of firm monopsony power in labor markets, has renewed interest in redistributive labor market policies such as the minimum wage and the Earned Income Tax Credit (EITC). Although these policies are designed to support the income of low-wage workers, they also affect the relative prices of these workers' labor services by making them more or less expensive to employ. Hence, the ultimate effects of these policies on employment, income, and welfare critically depend on firms' ability to substitute across workers.

Our starting point is the observation that the elasticity of substitution across workers tends to be lower in the short run than in the long run. In particular, we are motivated by two regularities based on two different sources of variation in wages and employment. First, increases in the minimum wage alter the relative wages of affected workers. In the short run, these wage changes tend to produce small changes in employment, which suggests that the short-run elasticity of substitution across workers is low (see Neumark and Shirley (2022)'s survey). Second, changes in demographic structure or immigration flows alter the labor supply of some groups of workers relative to others. In the long run, these changes in the relative employment of different groups of workers tend to lead to small changes in their relative wages, which implies that the response of wages in the long run to large changes in employment is small. Hence, this finding suggests that the long-run elasticity of substitution across different groups of workers is large (see Card and Lemieux (2001) and Borjas and Katz (2007)). Taken together, these observations indicate that the ability of firms to substitute across different groups of workers is lower in the short run than in the long run.

Our goal is to develop a framework in which the elasticity of substitution across workers endogenously changes over time in order to study the dynamic effects of labor market policies on different groups of workers. We do so by building a model of the U.S. economy that incorporates key features including heterogeneous workers, monopsonistic labor markets, and a putty-clay technology such that in the long run capital can be combined with different workers in any ratios but these ratios are fixed once capital is installed. Given this technology, in the short run, firms do not adjust the mix of workers employed to produce output with existing capital, even for large changes in policies—as consistent with the evidence on the low short-run elasticity of substitution among workers in response to changes in the minimum wage. Over time, though, firms modify the composition of their workforce both by utilizing less of their existing capital and by investing in new capital that is less intensive in the use of the labor services of workers who have become *more* expensive to employ (workers affected by the minimum wage) or more intensive in the use of the labor services of workers who have become *less* expensive to employ (workers affected by the EITC). This pattern of adjustment is consistent with the measured high long-run elasticity of substitution among workers in response to changes in aggregate labor supply.

Our main result is that taking these slow transition dynamics into account is necessary to correctly assess the full impact of labor market policies. Fundamentally, the reason is that the slow adjustment in the mix of workers that firms employ delays the ultimate long-run effects of labor market policies. Hence, when policies increase employment, these dynamics delay their benefits, whereas when policies decrease employment, these dynamics delay their costs. In either case, comparing the present value of labor income under such policies taking these dynamics into account can lead to vastly different conclusions about their impact than either a short-run empirical analysis, which extrapolates from a few years of observations, or a long-run theoretical analysis, which focuses on steady-state changes, would suggest.

Given our interest in redistributive policies, we allow for rich heterogeneity in workers' productivity so as to reproduce the observed distribution of wages in the data. In particular, we allow for differences in worker productivity both across broad education groups—namely, between college and non-college educated workers—and across workers within the same education group. Since empirically the dispersion of wages within an education group is an order of magnitude larger than across them and the estimates of the substitutability of labor within education groups are over four times as large as those across groups, the key margin of substitution for the labor market policies we study is across non-college workers with different levels of productivity.

Our putty-clay technology provides a rich model of how firms substitute across workers with different productivity over time. Formally, in each period when making investment decisions, firms choose between types of capital that require different mixes of worker types to be operated (putty-clay capital). The resulting capital built in that period has vintage-specific productivity and permanent idiosyncratic productivity as in Gilchrist and Williams (2000). This setup gives rise to both an endogenous rigidity in the use of multiple inputs to production, which accounts for their different degrees of substitutability over time, and a key margin of variable capital utilization, which critically affects the speed of an economy's transition to the long run after any large policy change. Indeed, in the long run, firms can flexibly substitute across different worker mixes by investing in different types of capital that are differentially intensive in the use of the labor services of different workers. Once a unit of capital is in place, however, production is Leontief in its required mix of workers. Hence, in the short run, firms can only adjust their labor demand by choosing which of the units of already installed capital—which differ in their input mixes, vintage productivity, and idiosyncratic productivity—to utilize in production.

Our second contribution is to embed this production structure into a labor market environment with matching frictions and firm monopsony power. We account for monopsony power in order to capture the idea, dating back to Robinson (1933), that labor market policies may not only redistribute income but also help correct inefficiencies. We incorporate this insight into a modern dynamic search-theoretic framework of the labor market with long-term employment relationships and empirically realistic labor market flows into and out of unemployment. Importantly, such a framework allows us to avoid specifying ad-hoc rationing rules when policy-mandated constraints, such as the minimum wage, become binding.[1]. We differ from this literature in that we augment the standard search framework with monopsonistic competition among firms to provide a potential welfare enhancing role of policy. Additionally, we focus on the transition dynamics induced by the various labor market policies and this literature focuses on steady state changes.

Since the speed of an economy's transition triggered by a policy change is fundamentally a quantitative question, we first show how the mechanism of our model can be disciplined by readily available data and then explore its implications for the dynamic impact of policies. We specify a degree of firms' monopsony power in the labor market that matches recent estimates of wage markdowns documented by Seegmiller (2021), Lamadon, Mogstad and Setzler (2022), and Berger, Herkenhoff and Mongey (2022), which find that, on average, workers are paid between 65% and 85% of their marginal products. We then parametrize firms' production technology by targeting key moments that are informative about input substitution possibilities in the short and long run. For the short run, we ensure that the distribution of capital productivity reproduces average capital utilization rates in the data. The resulting model implies small employment responses to the minimum wage in the short run, thus matching our first motivating fact. For the long run, we specify a set of possible capital types that traces out a nested constant elasticity of substitution frontier in capital and labor of different levels of education and productive ability. We choose the parameters of this function to match the long-run effects of changes in the relative supply of workers with different levels of education on their relative wages as estimated by Card and Lemieux (2001), consistent with our second motivating fact.

Using the minimum wage as our leading example, we illustrate now important it is to account for

[1] Our paper is related to a long line of research that has evaluated the impact of labor market policies through the lens of frictional models of the labor market. See, in particular, Eckstein and Wolpin (1990), Flinn (2006), Ahn, Arcidiacono and Wessels (2011), Engbom and Moser (2022), and Drechsel-Grau (2022)

transition dynamics when evaluating labor market policies. We do so because this context provides a particularly clean mapping between the policy under consideration and the changes in the relative cost of different types of workers to firms. Our main result is that when technology is of the putty-clay form, employment adjusts only slowly to changes in the minimum wage. Crucially, the direction of this adjustment depends on the size of the change in the minimum wage being contemplated. Specifically, when firms have monopsony power in the labor market, a small increase in the minimum wage ultimately *increases* the long-run employment of workers for whom the minimum wage binds—typically workers of low productivity—by reducing the monopsony distortions they face. In our model, though, increasing the employment of these workers requires firms to build new types of capital that are more intensive in the use of these workers' services, which take time to accumulate.

By contrast, a large increase in the minimum wage ultimately *decreases* the long-run employment of low-productivity workers by making them too expensive to employ relative to their marginal products. But again it takes time for firms to substitute away from these workers. In the short run, firms do not have an incentive to dismiss employed workers because doing so would require forfeiting the monopsony profits they still earn on high-productivity workers, given that installed (Leontief) capital implies a high degree of complementarity across workers. Instead, firms progressively substitute away from less productive workers only by replacing them at a lower rate than the rate at which they naturally separate—that is, by *attrition* that leaves some installed capital idle—while simultaneously building new capital that is less intensive in the use of such workers' services. Note that if capital utilization were full, as simpler putty-clay models imply, then attrition would occur at most at the speed at which capital depreciates. Our endogenous utilization margin instead accelerates an economy's transition relative to these simpler models because firms choose to idle more of their existing capital and thus to phase out workers more quickly than capital depreciates.

The slow overall adjustment in employment just described shapes how a minimum wage policy impacts the labor income of affected workers. On impact, the minimum wage immediately raises the wages of all workers bound by the new minimum. For a small increase in the minimum wage, these labor income gains increase over time as firms gradually hire more of these workers. In this sense, our putty-clay technology delays the ultimate long-run *benefits* of the policy. On the contrary, for a large increase in the minimum wage, labor income gains for the lowest-productivity workers are eroded over time as firms slowly substitute away from these workers. The employment loss is large enough that the labor income of the lowest-productivity workers eventually falls. In this case, our putty-clay technology delays the long-run *costs* of the policy.

We use our model to illustrate the potential limitations of two strategies of policy evaluation that are common in the literature, which abstract from the dynamic effects we focus on, by way of two stylized examples. First, a *static short-run* analysis would measure in the data the effects of a minimum wage change over the first couple of subsequent years and then presume that what occurs during this period of time captures what occurs from then on. Second, a *steady-state* analysis would instead assess the impact of labor market policies by comparing the new steady state of an economy to the old one. We show that both the static short-run and the long-run steady-state approaches severely misstate the labor income gains from labor market policies over time that our model implies. For example, accounting for the full transition path following an increase in the minimum wage to $15, we find that a worker originally earning $7.50 experiences a present value increase in labor income of about 40%. A static calculation would predict instead that labor income increases by 80%—which is double the true gain. By contrast, a steady state calculation would predict that labor income *decreases* by 60%. Although these examples are admittedly simplified representations of standard approaches, they illustrate our key message: when an economy's response to a sizable change in policy is slow, as labor market data indicate, any comprehensive assessment of the associated benefits or costs must take the full dynamics of adjustment into account.

Despite these results being primarily illustrated in the context of the minimum wage, similar lessons apply to the EITC. The key difference between these two policies is that under the EITC, the government, rather than firms, pays the cost of increasing the wages of low-productivity workers. As a result, the EITC reduces monopsony distortions without creating an incentive for firms to substitute away from lower-productivity workers, which results in an increase in their employment in the long run. As is the case for a small increase in the minimum wage, it takes time for firms to increase the employment of these workers, which delays the ultimate long-run benefits of the EITC. Hence, here too both a naive static and a purely steady-state analysis provide a poor approximation to evaluating changes in the EITC over the full transition that an economy would experience.

Some of the insights that our paper offers are related to those in Sorkin (2015) and Aaronson et al. (2018), who also argue that the effect of the minimum wage on employment is smaller in the short run than in the long run. These papers propose partial equilibrium models of a small sector—such as restaurants—with a variant of the standard putty-clay capital to study small changes in the minimum wage that affect only that sector. Since both papers capture consumers' behavior through a reduced-form industry demand curve, they are silent about labor supply and consumption and thus refrain from examining the effects of the minimum wage on workers' behavior and welfare.

By contrast, we study the dynamic distributional effects of labor market policies within a general equilibrium framework that connects firms' labor demand to individual workers' labor supply. On the firm side, the models in these papers feature full utilization of inputs and limited heterogeneity among workers. We find that allowing for intensive margins of capital and labor utilization and rich patterns of labor-labor substitution—especially within education groups—are crucial for assessing the impact of labor market policies. Finally, we incorporate firm monopsony power so that policies can potentially improve on allocations to shed light on any static and dynamic trade-off between efficiency and redistribution.[2] We also show that firm monopsony power has key implications for the transition of an economy in response to changes in policies, since it affects firms' incentives to adjust their stock of labor and capital as the relative prices of inputs change.

## 1 Model

We propose a framework that allows for rich margins of an economy's response to large-scale labor market policies. First, we combine directed search with monopsony power to provide a novel notion of dynamic monopsony power in an environment with long-term employment relationships. Second, we allow for variable degrees of substitutability in production among different workers and among workers and capital at different time horizons, connecting firms' choices of employment of different types of workers to firms' decisions of how much to invest and utilize different vintages of capital. Third, by doing so, the framework we develop gives rise to dynamic patterns of responses to policy that are richer and more consistent with data than existing models.

Formally, our economy features a single output good used for consumption and investment over an infinite horizon.[3] Households maximize utility by choosing their consumption, labor supply, and intensity of job search. Firms maximize profits by investing in capital, deciding how much of their capital to use in production, and hiring workers. In this section, we describe the model without any labor market policies. Omitted proofs and derivations are in the Online Appendix.

[2] Berger, Herkenhoff and Mongey (2025) study the role of the minimum wage in alleviating monopsony distortions in a model with rich firm heterogeneity. They find that long-run efficiency gains are small because a single minimum wage is too blunt an instrument to correct monopsony distortions across heterogeneous firms. We find a similar result in the long run given the heterogeneity in monopsony distortions across heterogeneous workers. However, our primary focus is on the dynamics generated by endogenous changes in the elasticity of substitution across workers. Also see Mousavi (2022) and Berger et al. (2024) for work exploring the interaction of monopsony and taxation.

[3] See Aaronson and French (2007) and MaCurdy (2015) for multi-good analyses of the minimum wage that account for differential effects of the minimum wage on the relative prices of the goods produced by different sectors. To this end, these papers assume that different sectors employ different mixes of workers. Although such an extension of our framework is straightforward in principle, it is challenging in practice. Any quantitative analysis would require disciplining each sector's production structure, including the elasticity of substitution among skill groups within each sector, the elasticity of demand of each sector for the different skill groups, and input-output links among sectors. Embedding this structure in a dynamic general equilibrium model is beyond the scope of this paper.

## 1.1 Households

Households or families differ in their type $i \in \{1, 2, \ldots, I\}$, which describes attributes that are imperfectly substitutable in production. In our quantitative analysis, a type consists of a household's education level and labor efficiency, which allows us to match the distribution of wages within and across education groups in U.S. data. Families of type $i$ have measure $\mu_i$, and each is composed of a large number of members.[4] Each family has preferences represented by the utility function

$$\sum\nolimits_{t=0}^{\infty} \beta^t U_t(c_{it}, n_{it}, s_{it}) \quad \text{with} \quad n_{it} = \left( \sum\nolimits_j n_{ijt}^{\frac{\omega+1}{\omega}} \right)^{\frac{\omega}{\omega+1}} \quad \text{and} \quad s_{it} = \sum\nolimits_j s_{ijt},$$

where $c_{it}$ is the family's consumption of the output good, $n_{it}$ is an index of the disutility of work, $s_{it}$ is an index of the disutility of labor market search, and $j$ denotes a firm. The index $n_{it}$ describes how a family views jobs at different firms $j$ as imperfect substitutes for each other, which gives rise to firm $j$'s monopsony power in hiring workers as described below.[5] The extent of firm monopsony power is governed by the degree of substitutability of jobs in workers' preferences, captured by the parameter $\omega$. In particular, as $\omega \to \infty$ jobs at different firms become perfect substitutes and firm monopsony power vanishes. We make this point formally below.

In our directed search setting, at the beginning of period $t$, families observe the current employment offers from each firm $j$. A job offer for members of a family of type $i$ can be summarized by the tightness $\theta_{ijt}$ of the sub-market of firm $j$'s jobs, defined as the ratio of the firm's posted vacancies $\mu_i a_{ijt}$ to the workers searching for them $\mu_i s_{ijt}$, and the present value of wages over the life of the match, $W_{ijt+1}$, as detailed below. Given these offers, each type-$i$ family chooses the measure of family members $s_{ijt}$ searching for jobs at each firm $j$. A family member searching in $t$ for a job at firm $j$ finds that job with probability $\lambda_w(\theta_{ijt})$, begins working in $t+1$, is paid the wage $w_{ijt+\tau} = (1+g)^\tau w_{ijt+1}$ in each period $\tau \geq 1$ of employment, and exogenously separates with probability $\sigma$ at the end of each period, where $g$ is the growth rate of the economy and the stream of wage payments $\{w_{ijt+\tau}\}$ has present value $W_{ijt+1}$. A type $i$-family's problem is to choose sequences of consumption, $\{c_{it}\}$, measures of family members searching for jobs at each firm $j$, $\{s_{ijt}\}$, and

[4] This formulation implies perfect consumption risk sharing within each type-$i$ family, as in Merz (1995) and Andolfatto (1996), but still allows for imperfect risk sharing *across* families of different types.

[5] See Berger, Herkenhoff and Mongey (2022) and Deb et al. (2024) for related preferences and discussions of their microfoundation. This specification can be thought as arising from a family's idiosyncratic valuation of work at a particular firm due to, say, locations or other non-wage amenities.

measures of family members employed at each such firm, $\{n_{ijt}\}$, to solve

$$\max_{c_{it},s_{ijt},n_{ijt+1}} \sum\nolimits_{t=0}^{\infty} \beta^t U_t(c_{it}, n_{it}, s_{it}) \tag{1}$$

$$\text{s.t. } n_{ijt+1} = (1-\sigma)n_{ijt} + \lambda_w(\theta_{ijt})s_{ijt}$$

$$\sum\nolimits_{t=0}^{\infty} Q_{0,t}c_{it} = \zeta_i \mathbb{P} + \mathbb{I}_i + \sum\nolimits_{t=0}^{\infty} Q_{0,t} \sum\nolimits_j \lambda_w(\theta_{ijt})s_{ijt}W_{ijt+1},$$

where the first constraint is the employment transition law in which $\sigma$ is the exogenous separation rate and $\lambda_w(\theta_{ijt})s_{ijt}$ is the measure of searching members that find jobs. The second constraint is the period 0 budget constraint where $Q_{0,t}$ denotes the price of a claim to output in period $t$ in units of output in period 0, $\zeta_i\mathbb{P}$ is the family's share of the present value of firm profits $\mathbb{P}$, and $\mathbb{I}_i$ is the present value of wages promised to family members initially employed in period 0.

The family's labor supply decisions are summarized by the first-order condition for the measure of a family's members searching for jobs at firm $j$ in period $t$, given by

$$-\frac{U_{sit}}{U_{cit}} = \lambda_w(\theta_{ijt})Q_{t,t+1}(W_{ijt+1} + V_{ijt+1}) \text{ if } s_{ijt} > 0,$$

where $U_{cit}$ and $U_{sit}$ denotes the derivative of $U_t(\cdot)$ with respect to $c_{it}$ and $s_{it}$. The left side of this equation is the marginal disutility of searching for jobs in period-$t$ consumption units, which is equated across all firms $j$ for which workers search. The right side, which is the marginal benefit of searching for jobs, reflects that a worker finds a job at firm $j$ with probability $\lambda_w(\theta_{ijt})$, begins working in the next period, and receives the present value of wage payments $W_{ijt+1}$ net of the present value of the disutility of working at firm $j$, $V_{ijt+1}$, both discounted by $Q_{t,t+1}$. Here and throughout we use the notation $Q_{t,s} = Q_{0,s}/Q_{0,t}$ for the price of the output good in period $s$ in units of the output good in period $t$. The present value of the disutility of work at firm $j$ is recursively defined as

$$V_{ijt+1} = \frac{U_{nit+1}}{U_{cit+1}} \left(\frac{n_{ijt+1}}{n_{it+1}}\right)^{\frac{1}{\omega}} + Q_{t+1,t+2}(1-\sigma)V_{ijt+2}. \tag{2}$$

If workers search for jobs at firm $j$ in period $t$, then the value of doing so must be at least as large as the value of searching for jobs at any other firm $j'$ in that $\lambda_w Q_{t,t+1}(\theta_{ijt})(W_{ijt+1} + V_{ijt+1}) \geq \mathcal{W}_{it} \equiv \max_{j'}\{\lambda_w(\theta_{ij't})Q_{t,t+1}(W_{ij't+1} + V_{ij't+1})\}$. In a symmetric equilibrium in which all firms other than firm $j$ offer workers the common value $\lambda_w(\theta_{it})Q_{t,t+1}(W_{it+1} + V_{it+1})$, this inequality reduces to

$$\lambda_w(\theta_{ijt})Q_{t,t+1}(W_{ijt+1} + V_{ijt+1}) \geq \mathcal{W}_{it} = \lambda_w(\theta_{it})Q_{t,t+1}(W_{it+1} + V_{it+1}), \tag{3}$$

where $\mathcal{W}_{it}$ is the market value of an offer. We refer to (3) as the *participation constraint* because firms understand that any job offer they make must satisfy this condition for them to be able to attract workers. As we formalize below, this constraint is our model's version of the firm-specific labor supply curve that arises in static models of firm monopsony power—see Robinson (1933) for an early reference—since imposing this constraint implies that each firm views itself as facing an upward-sloping supply curve of workers for its jobs. The supply curve is dynamic in our framework because both workers' job search decisions and firms' hiring decisions are intertemporal due to search frictions. In particular, forming an employment relationship entails incurring costs now—time for workers and resources for firms—for the prospect of future benefits—wages for workers and profits for firms. As we will show, firms' monopsony power affects not only the wages that firms pay to workers but also the number of job vacancies firms post or, equivalently, these workers' probabilities of finding a job. It turns out that the impact of firms' monopsony power on both of these margins leads to distinct sources of inefficiencies in the labor market: both wages and vacancies are lower than in the corresponding competitive equilibrium.

## 1.2 Firms

We develop a production structure in which short-run and long-run elasticities of substitution differ not only between capital and labor but also across different types of labor. To this end, we extend the putty-clay setup of Gilchrist and Williams (2000) with endogenous capital utilization to an environment with heterogeneous workers, search frictions, and firm monopsony power.[6]

**Production Technology.** A large but finite number of identical firms indexed by $j = 1, \ldots, J$ operate installed capital whose productivity differs along two dimensions. First, all units of capital produced in $t$ have the same permanent *vintage productivity* $A_t$. Vintage productivity grows at a constant rate $g_A$ chosen so that its evolution, $A_{t+1} = (1 + g_A)A_t$, generates an aggregate economic growth rate of $g$. This productivity growth leads to natural obsolescence of older vintages of capital, since the older the vintage the less productive is capital relative to the newest vintage. This force leads firms to progressively shut down older, less productive capital before shutting down newer, more productive capital. In period $t$, we index a firm's history of past installed capital vintages by their date of installation $t - \tau$; along a balanced growth path, we equivalently index them by the number of periods since they have been installed, $\tau \geq 0$.

Second, within each vintage of capital, any new unit of capital is subject to a permanent *idiosyncratic productivity* shock $\varepsilon$. Specifically, if a firm installs $K$ units of capital, referred to as *machines*, then a measure $\pi(\varepsilon)K$ of these machines has idiosyncratic productivity $\varepsilon$. Here $\pi(\varepsilon)$ is the

[6] A major point of Gilchrist and Williams's (2000) paper, whose technology we adopt, is that their version of the putty-clay model captures well business-cycle dynamics. Indeed, the authors argue that in many dimensions, their putty-clay setup actually fits the data better than the standard putty-putty setup.

p.d.f. of $\varepsilon$ with mean 1 and $\Pi(\varepsilon)$ is the corresponding c.d.f. One interpretation of these idiosyncratic productivity differences among new units of capital is that machines are often standardized and, as a result, can be more or less adequate for their particular use at a firm after they are installed. We show that these differences generate an active margin of variable capital utilization *within* each capital vintage, because firms have an incentive to shut down machines with lower $\varepsilon$ within a vintage. The total productivity of a type $(A_{t-\tau}, \varepsilon)$ unit of capital is $A_{t-\tau}\varepsilon$. We first describe the familiar putty-putty version of this production structure and then move to the putty-clay version.

*Putty-Putty Production.* Here, if firm $j$ combines $K_{jt}(A_{t-\tau}, \varepsilon)$ units of type $(A_{t-\tau}, \varepsilon)$ capital with $N_{1jt}(A_{t-\tau}, \varepsilon), \ldots, N_{Ijt}(A_{t-\tau}, \varepsilon)$ units of each type of labor, then the firm produces

$$A_{t-\tau}\varepsilon \times F\left(K_{jt}(A_{t-\tau}, \varepsilon), N_{1jt}(A_{t-\tau}, \varepsilon), \ldots, N_{Ijt}(A_{t-\tau}, \varepsilon)\right) \tag{4}$$

units of output in period $t$, where $F(K, N_1, \ldots, N_I)$ is a constant returns-to-scale production function. We assume that $F(K, N_1, \ldots, N_I) = K^{\alpha} G(N_1, \ldots, N_I)^{1-\alpha}$, where $G(N_1, \ldots, N_I)$ is also a constant returns-to-scale function.[7] In our quantitative work, we specify $G(N_1, \ldots, N_I)$ to be a constant elasticity of substitution (CES) function for consistency with the empirical literature on the elasticity of substitution among different groups of workers in production. Since the production function $F$ features constant returns to scale, we can write it in intensive form as $F(K, N_1, \ldots, N_I) = K \times F(1, N_1/K, \ldots, N_I/K) \equiv K \times f(v)$, where $v = (v_1, \ldots, v_I)$ is the vector of labor-to-capital ratios or *labor intensities* $v_i = N_i/K$ and $f(v) = F(1, v)$. We can then express the output produced using $K_{jt}(A_{t-\tau}, \varepsilon)$ units of type $(A_{t-\tau}, \varepsilon)$ capital in (4) as

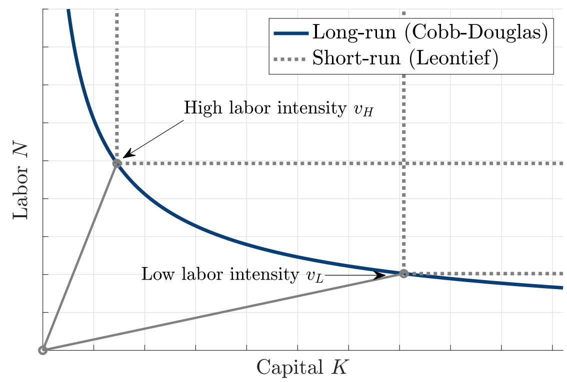
$$A_{t-\tau}\varepsilon \times K_{jt}(A_{t-\tau}, \varepsilon) \times f\left(v_{1jt}(A_{t-\tau}, \varepsilon), \ldots, v_{Njt}(A_{t-\tau}, \varepsilon)\right).$$

With putty-clay production, considered next, production at the firm level has a similar expression but with a Leontief production function.

*Putty-Clay Production.* In our putty-clay model, firms choose the labor intensity of capital of any vintage—when capital is *putty*—but cannot adjust it once capital is installed—when capital hardens to *clay*. This setup provides a parsimonious way to allow short-run elasticities of substitution across *all* inputs to differ from long-run ones. For example, a firm may invest in either a machine that needs to be operated by many low-skilled workers and few high-skilled ones or a machine that needs few low-skilled workers and many high-skilled ones. After this decision is made, the skill mix of

[7] This Cobb-Douglas form between capital and the labor aggregate is required to achieve balanced growth with vintage capital productivity growth. It is also broadly consistent with estimates of the elasticity of substitution between capital and labor in the literature. For instance, Oberfield and Raval (2021) suggests values for $\alpha$ between 0.5 and 0.7, whereas Karabarbounis and Neiman (2014) estimates a value of 1.25.

Figure 1: Illustration of Short-Run vs. Long-Run Isoquants




Notes: Short-run and long-run isoquants of putty-clay model with no capital heteorgeneity $A_{t-\tau} = \varepsilon = 1$ and one type of labor $G(N) = N$. Long-run isoquant corresponds to $F(K, N) = 1$. Short-run isoquants correspond to Leontief production functions for different labor intensities $v = N/K$.

workers needed to operate the machine is fixed over the life of the machine. Following Gilchrist and Williams (2000), we assume that the labor intensity of capital is chosen before capital is installed and the idiosyncratic productivity $\varepsilon$ is realized. This timing assumption leads to an active shutdown margin for firms. From now on, we index a unit of capital by $(v, A_{t-\tau}, \varepsilon)$, where $v$ is the fixed vector of labor intensities $v_i = N_i/K$. Specifically, if a firm combines $K_{jt}(v, A_{t-\tau}, \varepsilon)$ units of capital with $N_1(v, A_{t-\tau}, \varepsilon), \dots, N_I(v, A_{\tau-\tau}, \varepsilon)$ units of each type of labor, it produces output equal to

$$Y_{jt}(v, A_{t-\tau}, \varepsilon) = A_{t-\tau}\varepsilon \times \min\left\{K_{jt}(v, A_{t-\tau}, \varepsilon), \frac{N_{1jt}(v, A_{t-\tau}, \varepsilon)}{v_1}, \dots, \frac{N_{Ijt}(v, A_{t-\tau}, \varepsilon)}{v_I}\right\} \times f(v). \quad (5)$$

The min operator captures the Leontief nature of production once capital is installed. The maximum output that a firm can obtain from this set of machines is when it assigns to them $N_{ijt}(v, A_{t-\tau}, \varepsilon) = v_i K_{jt}(v, A_{t-\tau}, \varepsilon)$ units of each type of labor for all $i$—namely, in exactly the required proportion. We refer to $v_i K$ as the *labor requirement* of type-$i$ workers for a machine of type $v$. If the firm allocates too much labor of type $i$ to the machine in that $N_{ijt}(v, A_{t-\tau}, \varepsilon) > v_i K_{jt}(v, A_{t-\tau}, \varepsilon)$ , then the excess labor of type $i$ assigned to the machine, $N_{ijt}(v, A_{t-\tau}, \varepsilon) - v_i K_{jt}(v, A_{t-\tau}, \varepsilon)$, remains idle. Likewise, if the firm allocates too little labor of type $i$ to the machine in that $N_{ijt}(v, A_{t-\tau}, \varepsilon) < v_i K_{jt}(v, A_{t-\tau}, \varepsilon)$, then some of the capital $K(v, A_{t-\tau}, \varepsilon)$, namely $v_i K_{jt}(v, A_{t-\tau}, \varepsilon) - N_{ijt}(v, A_{t-\tau}, \varepsilon)$, remains idle, that is, the firm shuts down some of that type of capital. If a firm has a total amount $N_{ijt}$ of workers of type $i$, then the allocation of type $i$-workers across machines, namely $N_{ijt}(v, A_{t-\tau}, \varepsilon)$ for a machine of type $(v, A_{t-\tau}, \varepsilon)$, must satisfy the *adding-up constraint*

$$\sum\nolimits_{\tau} \int_{v,\varepsilon} N_{ijt}(v, A_{t-\tau}, \varepsilon)\pi(\varepsilon) dv d\varepsilon \leq N_{ijt}. \quad (6)$$

The long-run elasticity of substitution across inputs is governed by the intensive-form production function $f(v)$. To see how, Figure 1 plots isoquants of the short- and long-run production functions with $A_{t-\tau} = \varepsilon = 1$ and one type of labor, $G(N) = N$. The long-run isoquant is generated by the putty-putty production function $F(K, N) = K^{\alpha} N^{1-\alpha}$, with intensive form $f(v) = (N/K)^{1-\alpha}$. A firm with a putty-clay production function can choose ex ante among different types of capital with labor intensities $v = N/K$ along a long-run isoquant. However, once capital is installed, its labor intensity $v$ is fixed and the firm's only choice is how much of it to utilize, that is, where on the corresponding ray to the origin the firm chooses to locate itself.

Over time, though, firms can choose different points along the long-run isoquant by accumulating different types of capital. In this sense, the long-run substitution possibilities are determined by the putty-putty technology $f(v)$. With multiple types of workers, so that, say, $F(K, N_1, N_2) = K^{\alpha} G(N_1, N_2)$, likewise a firm over time can gradually substitute across workers by installing new machines that embody new labor-to-capital ratios. For example, when the minimum wage increases, firms can substitute away from low-productivity workers towards higher-productivity ones and/or capital. The speed of the economy's adjustment to labor market policies depends on how much of the existing capital is utilized and the rate at which new capital requiring different input mixes is installed.

To be able to characterize the firm's problem using standard first-order conditions, we express output in terms of capital utilization rates and replace the min-representation with inequality constraints on these rates. Specifically, we write the output of each machine of type $(v, A_{t-\tau}, \varepsilon)$ in terms of the *worker-to-capital ratios* $N_{ijt}(v, A_{t-\tau}, \varepsilon)/K_{jt}(v, A_{t-\tau}, \varepsilon)$ so that

$$Y_{jt}(v, A_{t-\tau}, \varepsilon) = A_{t-\tau} \varepsilon K_{jt}(v, A_{t-\tau}, \varepsilon) \min \left\{ 1, \frac{1}{v_1} \cdot \frac{N_{1jt}(v, A_{t-\tau}, \varepsilon)}{K_{jt}(v, A_{t-\tau}, \varepsilon)}, \ldots, \frac{1}{v_I} \cdot \frac{N_{Ijt}(v, A_{t-\tau}, \varepsilon)}{K_{jt}(v, A_{t-\tau}, \varepsilon)} \right\} f(v),$$

and then define the *utilization rate* $u_{jt}(v, A_{t-\tau}, \varepsilon)$ for each machine as

$$u_{jt}(v, A_{t-\tau}, \varepsilon) = \min \left\{ 1, \frac{1}{v_1} \cdot \frac{N_{1jt}(v, A_{t-\tau}, \varepsilon)}{K_{jt}(v, A_{t-\tau}, \varepsilon)}, \ldots, \frac{1}{v_I} \cdot \frac{N_{Ijt}(v, A_{t-\tau}, \varepsilon)}{K_{jt}(v, A_{t-\tau}, \varepsilon)} \right\}. \tag{7}$$

This utilization rate is less than 1 if a firm assigns fewer workers than required of *any* type to operate the machine at full capacity, namely, $N_{ijt}(v, A_{t-\tau}, \varepsilon) < v_i K_{jt}(v, A_{t-\tau}, \varepsilon)$ for any type $i$.

A firm chooses the utilization rate of each machine subject to three constraints:

$$u_{jt}(v, A_{t-\tau}, \varepsilon) \geq 0, u_{jt}(v, A_{t-\tau}, \varepsilon) \leq 1, \text{ and } u_{jt}(v, A_{t-\tau}, \varepsilon) \leq N_{ijt}(v, A_{t-\tau}, \varepsilon)/(v_i K_{jt}(v, A_{t-\tau}, \varepsilon)). \tag{8}$$

The first constraint, $u_{jt}(v, A_{t-\tau}, \varepsilon) \geq 0$, captures the non-negativity constraint on $N_{ijt}(v, A_{t-\tau}, \varepsilon)$, as inputs cannot be negative. We capture the constraints imposed by the min operator in (5) with the two constraints $u_{jt}(v, A_{t-\tau}, \varepsilon) \leq 1$, since capital cannot be utilized beyond its capacity, and

$u_{jt}(v, A_{t-\tau}, \varepsilon) \leq N_{ijt}(v, A_{t-\tau}, \varepsilon)/(v_i K_{jt}(v, A_{t-\tau}, \varepsilon))$. We can then rewrite (7) as

$$Y_{jt}(v, A_{t-\tau}, \varepsilon) = A_{t-\tau}\varepsilon \times u_{jt}(v, A_{t-\tau}, \varepsilon) \times K_{jt}(v, A_{t-\tau}) f(v), \tag{9}$$

and append the constraints on capital utilization just described to the firm problem.

**Investment.** Each period $t$, firms choose how much new capital $X_{jt}(v)$ of vintage $t$ to install for each labor intensity $v$. Since capital depreciates at rate $\delta$, the amount of type $(v, A_t, \varepsilon)$-type capital left in period $t + \tau$ is $(1 - \delta)^\tau \pi(\varepsilon) X_{jt}(v, A_t)$. A key assumption is that investment is irreversible in that $X_{jt}(v) \geq 0$. Otherwise, firms could replicate the putty-putty model by converting the existing capital back into output and then investing this output into the optimal type of putty-clay capital.

**Hiring.** Firms hire workers in a monopsonistically competitive labor market with directed search with a constant returns to scale matching function $m(\mu_i a, \mu_i s)$. In period $t$, firm $j$ creates a measure of vacancies $\mu_i a_{ijt}$ at cost $\kappa_{it}$ each and posts offers of the form $(\theta_{ijt}, W_{ijt+1})$ to attract each type-$i$ workers, where market tightness $\theta_{ijt} = (\frac{\mu_i a_{ijt}}{\mu_i s_{ijt}})$ determines the probability $\lambda_w(\theta_{ijt})$ that a worker searching in this market finds firm $j$ and the probability $\lambda_f(\theta_{ijt})$ that firm $j$ meets a worker. Firms are able to attract workers whenever their offers satisfy the participation constraint in (3)—we maintain that firms treat symmetrically individual families of the same type. The measure $\{N_{ijt}\}$ of workers of type-$i$ employed by firm $j$ evolves according to the transition law

$$N_{ijt+1} = (1 - \sigma) N_{ijt} + \lambda_f(\theta_{ijt}) \mu_i a_{ijt}. \tag{10}$$

We take the number of firms $J$ to be finite, so that each firm hires a measure of workers, but large enough that they act as monopsonistic competitors; see Burdett and Judd (1983).

**Firm Problem.** Let $Y_{jt} = \sum_\tau \int_{v,\varepsilon} Y_{jt}(v, A_{t-\tau}, \varepsilon) \pi(\varepsilon) dv d\varepsilon$ with $Y_{jt}(v, A_{t-\tau}, \varepsilon)$ as in (9). Taking as given the sequence of intertemporal prices, $\{Q_{0t}\}$, and the market value of offers, $\{\mathcal{W}_{it}\}$, each firm $j$ chooses sequences of investments in each capital type $v$, $\{X_{jt}(v)\}$, allocations of type-$i$ workers across installed capital, $\{N_{ijt}(v, A_{t-\tau}, \varepsilon)\}$ for each machine of type $(v, A_{t-\tau}, \varepsilon)$, measures of vacancies to post for each type-$i$ worker, $\{\mu_i a_{ijt}\}$, job offers for each such type, $\{\theta_{ijt}, W_{ijt+1}\}$, and total workers of type $i$, $\{N_{ijt}\}$, understanding that $N_{ijt} = \mu_i n_{ijt}$, to maximize

$$\mathbb{P} = \sum_{t=0}^{\infty} Q_{0,t} \left[ Y_{jt} - \int_v X_{jt}(v) dv - \sum_{i=1}^{I} \kappa_{it} \mu_i a_{ijt} - \sum_{i=1}^{I} \lambda_f(\theta_{ijt}) \mu_i a_{ijt} Q_{t,t+1} W_{ijt+1} \right], \tag{11}$$

the present value of profits. Constraints to this problem are the participation constraints (3) for all $i$ and $t$, the adding up constraints (6) for all $i$ and $t$, the utilization constraints (8) for all $t$, the transition laws for labor (10) for all $i$ and $t$, the investment irreversibility constraints $X_{jt}(v) \geq 0$ for all $t$, and the nonnegative vacancy constraints $\mu_i a_{ijt} \geq 0$ for all $i$ and $t$.

**Equilibrium.** We focus on symmetric equilibria in which all firms make the same capital, labor, utilization, vacancy creation, and employment decisions. A *symmetric monopsonistically competitive search equilibrium* consists of *i*) allocations for each type $i$ of households, namely, sequences of consumption, $\{c_{it}\}$, measures of family members searching for jobs at each firm $j$, $\{s_{ijt}\}$, and measures of family members employed at each such firm $\{n_{ijt}\}$; *ii*) allocations for firms, namely, sequences of investments in each capital type $v$, $\{X_{jt}(v)\}$, associated capital stocks $\{K_{jt+\tau+1}(v, A_t, \varepsilon)\}$, allocations of each worker type $i$ to each machine type $(v, A_{t-\tau}, \varepsilon)$, $\{N_{ijt}(v, A_{t-\tau}, \varepsilon)\}$, utilization rates for each machine type $(v, A_{t-\tau}, \varepsilon)$, $\{u_{jt}(v, A_{t-\tau}, \varepsilon)\}$, measures of vacancies to post for each worker type $i$, $\{\mu_i a_{ijt}\}$, employment offers, $\{\theta_{ijt}, W_{ijt}\}$, and the total measure of employed workers of each type $\{N_{ijt}\}$; *iii*) intertemporal prices $Q_{0,t}$ for consumption goods such that at these allocations, *a*) each household $i$'s allocation solves (1); *b*) each firm $j$'s allocation solves (11); *c*) at each date $t$, the job-finding and job-filling rates are consistent with the matching function; *d*) total employment at each firm of employed workers of each type, $\{N_{ijt}\}$, satisfies the adding-up constraint (6) and the transition law (10); *e*) labor demand equals labor supply for each type $i$ of worker, $N_{ijt} = \mu_i n_{ijt}$, at each date $t$; and *f*) the output market clears,

$$\sum\nolimits_i \mu_i c_{it} + \sum\nolimits_{i,j} \kappa_{it} \mu_i a_{ijt} + \sum\nolimits_j \int_v X_{jt}(v) dv = \sum\nolimits_{j,\tau} Y_{jt}.$$

We use the balanced growth path, *BGP*, of this economy to build intuition. On that path, consumption, investment, output, wages, and the disutility of both working and searching grow along with the economy at rate $1 + g$ and searching, employment, vacancies, and intertemporal prices, $Q_{t,t+1}$ stay constant. Finally, since the capital stock grows but labor stays constant the labor to capital ratios shrink over them.

## 2 Equilibrium Characterization

We now characterize equilibrium. We mainly focus on the firm problem as it determines the key margins affecting the economy's response to market-wide labor policies. First, endogenous capital utilization governs short-run labor demand, given the distribution of installed capital. Second, investment in new types of capital generates an increasing degree of substitutability across workers over time. Together with firms' dynamic monopsony power, these features of our framework give rise to dynamic patterns of responses to policy that are richer and more consistent with the data than those implied by existing models.

## 2.1 The Allocation of Labor to Capital: Capital Utilization

We can decompose a firm's dynamic profit maximization problem into a static component, governing the allocations of employed workers to existing capital, and a dynamic component, governing the hiring of workers and the accumulation of new capital. We now turn to the first one.

**Utilization Problem.** A firm's key margin of adjustment in the short run is how much to utilize each type of installed capital by choosing how many workers to assign to each unit. The utilization decision is part of the solution to the dynamic problem in (11). But taking as given the multiplier $\hat{\chi}_{ijt}$, from this dynamic problem on the constraint (6) on the uses of labor, we show that the utilization rate for each machine solves a static problem given the firm's existing capital stock of previous vintages $\{K_{jt}(v, A_{t-\tau})\}_{\tau=1}^{\infty}$. The static utilization problem is

$$\max_{\{u_{jt}(v,A_{t-\tau},\varepsilon)\}_{\tau=1}^{\infty}} \sum\nolimits_{\tau} \int_{v,\varepsilon} A_{t-\tau}\varepsilon \times u_{jt}(v, A_{t-\tau}, \varepsilon) \times K_{jt}(v, A_{t-\tau})\pi(\varepsilon) dv d\varepsilon$$
$$+ \sum\nolimits_{i} \hat{\chi}_{ijt} \left[ N_{ijt} - \sum\nolimits_{\tau} \int_{v,\varepsilon} N_{ijt}(v, A_{t-\tau}, \varepsilon)\pi(\varepsilon) dv d\varepsilon \right] \tag{12}$$
$$\text{s.t. } 0 \leq u_{jt}(v, A_{t-\tau}, \varepsilon) \leq 1 \ \text{ and } \ u_{jt}(v, A_{t-\tau}, \varepsilon) v_i K_{jt}(v, A_{t-\tau}, \varepsilon) \leq N_{ijt}(v, A_{t-\tau}, \varepsilon).$$

We refer to this problem as a *pseudo-Lagrangian* problem because the multipliers $\hat{\chi}_{ijt}$ are taken from a different problem, namely the dynamic problem in (11). These multipliers capture the shadow value of an additional marginal measure of type-$i$ workers available for production only in period $t$. Note that this shadow value is also the opportunity cost of not assigning workers to their best alternative use in operating machines. At an optimal allocation, this opportunity cost equals the marginal product of labor assigned to the marginal operating machine. Therefore, we refer to it either as the *marginal product* or the *shadow cost* of labor. In the absence of search frictions, this shadow cost would simply equal the flow wage $\widehat{\chi}_{ijt} = w_{ijt}$.

**Proposition 1.** *The optimal utilization rates for the static pseudo-Lagrangian problem in* (12) *are also the optimal utilization rates for the dynamic problem in* (11)*. Both are given by a cutoff rule such that firms fully utilize capital of type* $(v, A_{t-\tau}, \varepsilon)$ *with idiosyncratic productivity* $\varepsilon$ *above the threshold* $\underline{\varepsilon}(v, A_{t-\tau}, \hat{\chi}_{jt}) = \sum_i \hat{\chi}_{ijt} v_i / [A_{t-\tau} f(v)]$ *and do not utilize capital with* $\varepsilon$ *below it.*

To prove this proposition, we first show that optimal capital utilization follows a cutoff rule. To see why, note that the first-order condition for utilization $u_{jt}(v, A_{t-\tau}, \varepsilon)$ in (11) is given by

$$A_{t-\tau}\varepsilon f(v) K_{jt}(v, A_{t-\tau})\pi(\varepsilon) - \sum\nolimits_{i} \lambda_{ijt}(v, A_{t-\tau}, \varepsilon) v_i K_{jt}(v, A_{t-\tau}) = \lambda_{ijt}^{U}(v, A_{t-\tau}, \varepsilon) - \lambda_{ijt}^{L}(v, A_{t-\tau}, \varepsilon),$$

where $\lambda_{ijt}^{L}(v, A_{t-\tau}, \varepsilon), \lambda_{ijt}^{U}(v, A_{t-\tau}, \varepsilon)$, and $\lambda_{ijt}(v, A_{t-\tau}, \varepsilon)$ are the multipliers on three utilization constraints in (8), respectively. Substituting the first-order condition for labor assignment $N_{ijt}$ in problem (11), namely, $\lambda_{ijt}(v, A_{t-\tau}, \varepsilon) = \hat{\chi}_{ijt}\pi(\varepsilon)$ where $\hat{\chi}_{ijt}$ is the multiplier on the adding-up constraint (6), into the first-order condition for utilization and dividing by $K_{jt}(v, A_{t-\tau})\pi(\varepsilon)$ yields

$$A_{t-\tau}\varepsilon f(v) - \sum_i \hat{\chi}_{ijt}v_i = \left[\lambda^U_{ijt}(v, A_{t-\tau}, \varepsilon) - \lambda^L_{ijt}(v, A_{t-\tau}, \varepsilon)\right] / \left[K_{jt}(v, A_{t-\tau})\pi(\varepsilon)\right]. \quad (13)$$

Now, if $A_{t-\tau}\varepsilon f(v) - \sum_i \hat{\chi}_{ijt}v_i > 0$ or, equivalently, $\varepsilon > \underline{\varepsilon} \equiv \sum_i \hat{\chi}_{ijt}v_i/[A_{t-\tau}f(v)]$, then (13) implies that $\lambda^U_{ijt}(v, A_{t-\tau}, \varepsilon) - \lambda^L_{ijt}(v, A_{t-\tau}, \varepsilon) > 0$ and so $u_{jt}(v, A_{t-\tau}, \varepsilon) = 1$ by complementary slackness. If $A_{t-\tau}\varepsilon f(v) - \sum_i \hat{\chi}_{ijt}v_i < 0$ or, equivalently, $\varepsilon < \underline{\varepsilon} \equiv \sum_i \hat{\chi}_{ijt}v_i[A_{t-\tau}f(v)]$, then $\lambda^U_{ijt}(v, A_{t-\tau}, \varepsilon) - \lambda^L_{ijt}(v, A_{t-\tau}, \varepsilon) < 0$ by (13), which yields that $u_{jt}(v, A_{t-\tau}, \varepsilon) = 0$ by complementary slackness. So, the utilization decision has the form: fully utilize if $\varepsilon > \underline{\varepsilon}$ and do not utilize at all if $\varepsilon < \underline{\varepsilon}$.[8] Note that this solution depends on time only through the multipliers $\hat{\chi}_{ijt}$ and the productivity $A_{t-\tau}$.

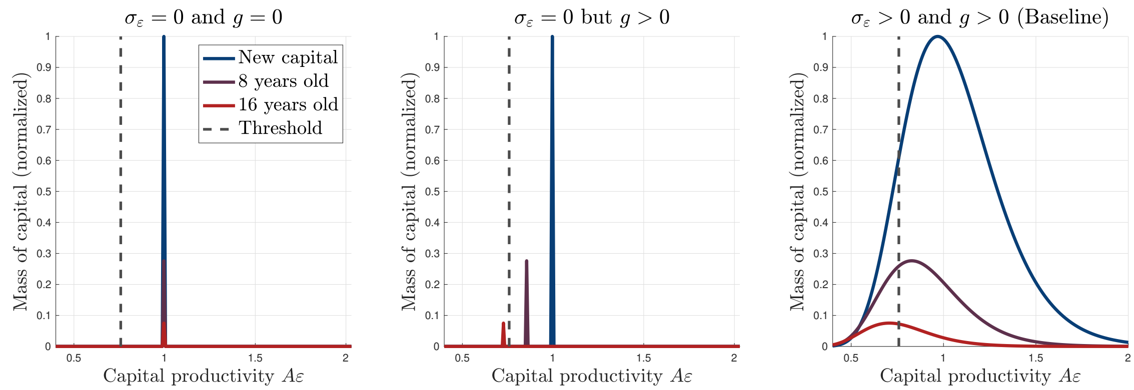
Next, we show that the solution to the static problem (12) for utilization coincides with that of the dynamic problem. To do so, we note that if we follow the same steps for this static problem as we did with the dynamic problem, we find that that the cutoff rule of Proposition 1 is the solution to this static problem—with the given $\hat{\chi}_{ijt}$ as well. This concludes the proof.

To understand this cutoff rule, suppose that the $\pi(\varepsilon)$ machines using capital $K_{jt}(v, A_{t-\tau})$ with productivity $A\varepsilon$ are not being fully utilized, so that $u_{jt}(v, A_{t-\tau}, \varepsilon) < 1$. Increasing their utilization by $\Delta$ would require an increase in each of the labor types assigned to them by $\pi(\varepsilon)\Delta N_{ijt} = \pi(\varepsilon)v_iK\Delta$. Doing so would tighten the adding-up constraint on all $I$ types of labor at a *shadow cost* of $\sum_i \hat{\chi}_{ijt}v_i\pi(\varepsilon)\Delta$. This shadow cost captures that using more of each type of labor on an existing non-fully utilized machine prevents other machines from being utilized at a higher rate. The cutoff rule for utilization prescribes that such machines should be fully utilized only if the increase in output from doing so, namely, $A_{t-\tau}\varepsilon f(v)K_{jt}(v, A_{t-\tau})\pi(\varepsilon)\Delta$, is at least large as the shadow cost of the resources required to so so, namely $\sum_i \hat{\chi}_{ijt}v_iK_{jt}(v, A_{t-\tau})\Delta$. In Proposition 1, we emphasized that the multiplier in the static problem came from the dynamic problem, hence we wrote it as $\hat{\chi}_{ijt}$. From now on, though, we let $\chi_{ijt}$ denote the multiplier from the dynamic problem.

The quantitative importance of this utilization rate margin is determined by the mass of existing capital near its operating threshold. Following Gilchrist and Williams (2000), we assume that $\log\varepsilon \sim N(-\sigma^2_\varepsilon/2, \sigma^2_\varepsilon)$ so that the distribution of idiosyncratic capital productivity is parameterized by its dispersion $\sigma_\varepsilon$. In our quantitative exercise we choose $\sigma_\varepsilon$ to match the average capital utilization rate in the data, implicitly relying on the log-normal form of the distribution of $\varepsilon$ to determine changes in capital utilization in response to the policies we examine.

[8] In the knife-edge case where $A_{t-\tau}\varepsilon f(v) - \sum_i \chi_{ijt}v_i = 0$, the firm is indifferent over any $u_{jt}(v, A_{t-\tau}, \varepsilon) \in [0, 1]$.

Figure 2: Illustrating the Capital Utilization Margin Along the BGP




Notes: Distribution of capital productivity $A_{t-\tau}\varepsilon$ along the BGP. Colored lines plot mass of capital of different vintages. Dashed vertical line is the operating threshold $A_{t-\tau}\underline{\varepsilon}_{t-\tau}$; all capital to the right of it is utilized.

**Intuition from the Balanced Growth Path.** To understand the determinants of firms' utilization decisions, consider our model's balanced growth path. As we show, on this path firms invest in a unique type of capital each period, with labor intensity $v_t = N_{ijt}(v, A_{t-\tau}, \varepsilon)/K_{jt}(v, A_{t-\tau}, \varepsilon)$ that shrinks at rate $g$, since $K_{jt}(v, A_{t-\tau}, \varepsilon)$ grows at rate $g$ and $N_{ijt}(v, A_{t-\tau}, \varepsilon)$ stays constant. By Proposition 1, a unit of capital of vintage $t-\tau$ is operated in period $t$ if and only if its overall productivity $A_{t-\tau}\varepsilon_{t-\tau}$ is greater than the cutoff level $A_{t-\tau}\underline{\varepsilon}_{t-\tau} = \sum_{i=1}^{I} \widetilde{\chi}_i \widetilde{v}_i / f(\widetilde{v})$, where $\widetilde{v}_i = \widetilde{v}_i(1+g)^t$ is the detrended type-$i$ labor requirement of the capital and $\widetilde{\chi}_i = \widetilde{\chi}_{it}/(1+g)^t$ is the detrended shadow cost of type-$i$ labor. Since $A_{t-\tau} = (1+g_A)^{-\tau} A_t$ and $A_0$ is normalized to 1, the operating threshold for a unit of capital from vintage $t-\tau$ in $t$ is $\underline{\varepsilon}_{t-\tau} = (1+g)^{\tau} \sum_{i=1}^{I} \widetilde{\chi}_i \widetilde{v}_i / f(\widetilde{v})$. This idiosyncratic productivity cutoff grows as the vintage ages because the shadow cost of labor grows at rate $g$ but the vintage productivity $A_{t-\tau}$ remains fixed. Therefore, a shrinking set of these machines—those with idiosyncratic productivity $\varepsilon \geq \underline{\varepsilon}_{t-\tau}$—are profitable to operate.

Figure 2 shows that both vintage productivity growth $g$ and the dispersion of idiosyncratic capital productivity $\sigma_\varepsilon$ play an important role in shaping capital utilization. Each panel plots the distribution of capital according to its overall productivity $A_{t-\tau}\varepsilon$. In this space, the threshold $A_{t-\tau}\underline{\varepsilon}_{t-\tau}$ is constant across vintages, but average overall productivity is lower for older vintages. All capital with productivity to the right of this operating threshold, denoted with a vertical dashed line, is utilized. As a baseline, the left panel of Figure 2 plots the distribution of capital productivity in a version of the model without vintage growth ($g = 0$) and without dispersion in capital productivity ($\sigma_\varepsilon = 0$). In this case, the distribution of capital productivity is degenerate. Moreover, the distribution lies strictly above the utilization threshold, and thus all capital is fully utilized as in Atkeson and Kehoe (1999). This result occurs because the variable shadow profit from operating capital, $A_{t-\tau}\varepsilon f(\widetilde{v}) - \sum_i \widetilde{\chi}_i \widetilde{v}_i$, is strictly positive due to firms having to recover the cost of producing capital and having monopsony power in the labor market. Hence, a small enough increase in the shadow cost of labor does not induce firms to shut down capital.

The middle panel of Figure 2 introduces capital heterogeneity across vintages (with $g = 2\%$ as in our calibrated model) but abstracts from heterogeneity within vintages ($\sigma_\varepsilon = 0$). The productivity of a given vintage of capital now declines relative to the frontier so that old enough vintages cross the operating threshold and are completely shut down. In this case, an increase in the shadow cost of labor induces firms to shut down some of the older vintages of capital near the threshold. However, the mass of such older vintages is fairly small because most of the old capital has already depreciated away at rate $\delta$ by the time this threshold is reached.

The right panel of Figure 2 allows for heterogeneity both across vintages ($g = 2\%$) and within vintages ($\sigma_\varepsilon > 0$ as in our calibrated model). As before, the average productivity of newer vintages is higher than that of older vintages. However, due to the significant dispersion in productivity within vintages, each vintage has a positive mass of capital that is shut down. Indeed, every vintage has marginal machines that are shut down for even a small increase in the shadow cost of labor.

## 2.2 Dynamic Labor and Capital Allocation

We now turn to characterizing the dynamic decisions about how capital and labor evolve over time.

**Firm Labor Market Decisions.** We start with a description of workers' search and firms' hiring decisions. Our key contribution in this dimension is to show that our framework with monopsonistically competitive search yields a tractable model of firm dynamic monopsony power with long-term employment relationships. These features lead to novel interactions between a firm's hiring decisions across periods that are absent from existing models, which consider either one-period employment relationships and firm monopsony power (Berger, Herkenhoff and Mongey (2022)) or long-term employment relationships in a competitive setting (Kehoe et al. (2023)). The resulting monopsony distortions open the door for labor market policies to improve the efficiency of the equilibrium.

To understand our model's dynamic monopsony power, consider the sequence of participation constraints in (3) that a firm in period 0 confronts when making labor market decisions given by (3) for any $t$. These constraints are different from the analogous ones that arise in competitive search models due to the imperfect substitutability of jobs in workers' preferences. Specifically, the key term in the present value of the disutility of work at firm $j$, $V_{ijt}$, defined in (2) that is specific to our imperfectly substitutable case is $(n_{ijt+1}/n_{it+1})^{\frac{1}{\omega}}$, which arises from $dn_{it+1}/dn_{ijt+1}$. When workers view jobs as imperfectly substitutable, firms take account that hiring workers of a family of type

$i$ in any period $s$ affects the utility of the future family members hired in any subsequent period. Intuitively, if firm $j$ hires additional $\lambda_w(\theta_{ijt})s_{ijt}\Delta$ workers in period $s \leq t$, then $(1-\sigma)^{t-s}\Delta$ of them are still working at the firm in period $t$. Through this term, the presence of these members of a family of type $i$ hired in $s$ affects the utility of all future members of this family hired in any period $t \geq s$.[9] By contrast, in the perfectly competitive search case ($\omega = \infty$), the term $dn_{it+1}/dn_{ijt+1}$ equals 1 and the participation constraints of workers hired in any period are disconnected from those hired in any other period (so $\partial V_{ijt+\tau+1}/\partial n_{ijt+1} \neq 0$ for all $\tau \geq 0$).

*Source of Dynamic Monopsony.* The intertemporal linkage across the participation constraints in (3) for workers hired in different periods makes the analysis of firm monopsony power much richer than in the typical static monopsony model. To keep track of these dynamic interactions, we collect the terms of the participation constraints in each period $t$ that are common across periods (as in Marcet and Marimon (2019)) to isolate the impact of additional hires of a type-$i$ family by firm $j$ in $t$ on the disutility of work of all family's members hired by the firm in future periods. Let then $Q_{0,t+1}\mu_i\gamma_{ijt+1}$ be the (scaled) multiplier on the time-$t$ participation constraint of the firm problem in (11). After grouping terms, the contribution of these constraints to the Lagrangian is

$$\sum\nolimits_{t=0}^{\infty} Q_{0,t+1}\mu_i M_{ijt+1}\frac{U_{nit+1}}{U_{cit+1}}\left(\frac{n_{ijt+1}}{n_{it+1}}\right)^{\frac{1}{\omega}} + \sum\nolimits_{t=0}^{\infty} Q_{0,t+1}\mu_i\gamma_{ijt+1}\left[W_{ijt+1} - \frac{\mathcal{W}_{it}}{Q_{t,t+1}\lambda_w(\theta_{ijt})}\right], \quad (14)$$

where $M_{ijt}$ is an auxiliary variable with transition law $M_{ijt+1} = (1-\sigma)M_{ijt} + \gamma_{ijt+1}$ summarizing a firm's *dynamic promises* to type-$i$ workers by cumulating past multipliers on (3) where we have divided the participation constraint by $\lambda_w(\theta_{ijt})$.

The firm's first-order condition for labor, $N_{ijt} = \mu_i n_{ijt}$, is key for our analysis

$$\nu_{ijt} = \underbrace{\chi_{ijt}}_{\text{marginal product of labor}} + \underbrace{M_{ijt}\frac{U_{nit}}{U_{cit}}\frac{1}{\omega}\left(\frac{n_{ijt}}{n_{it}}\right)^{\frac{1}{\omega}-1}\frac{1}{n_{it}}}_{\text{source of monopsony distortion}} + Q_{t,t+1}(1-\sigma)\nu_{ijt+1}. \quad (15)$$

Here $\nu_{ijt}$, referred to as the *value of hiring a worker of type $i$*, is the multiplier on the constraint $N_{ijt+1} \leq (1-\sigma)N_{ijt} + \lambda_f(\theta_{ijt})\mu_i a_{ijt}$ and hence gives the value of a marginal increase in the measure of type $i$ workers to firm $j$ in $t$. The first term in (15), which is positive, is the flow value in production of assigning these workers to a marginal machine as captured by $\chi_{ijt}$, which is the multiplier on the adding-up constraint on labor. The second term, which is negative, summarizes how the greater cost of hiring more workers of the same type tightens the participation constraints in (3). This term, which is equal to the derivative of the first term on the left side of (14) with

[9] We have suppressed notation for individual families within the type $i$, so these calculations can be thought of as performed at the individual family level after imposing symmetry across individual families of type $i$.

respect to $n_{ijt}$, describes the impact of an additional measure of employed workers of type $i$ on these constraints. Formally, it captures the effect of an increase in $n_{ijt}$ on the flow disutility of work for workers of type $i$, $\frac{\partial}{\partial n_{ijt}}\left(\frac{n_{ijt}}{n_{it}}\right)^{\frac{1}{\omega}} = \frac{1}{\omega}\left(\frac{n_{ijt}}{n_{it}}\right)^{\frac{1}{\omega}-1}\frac{1}{n_{it}}$, converted to units of consumption and, hence, output through the term $U_{nit}/U_{cit}$, and scaled by the shadow price $M_{ijt}$. The third term captures that only $1-\sigma$ of these wokers will remain employed at $t+1$.

*Monopsony Distortions.* Here monopsony power, $\omega < \infty$, reduces the present value of a worker to the firm in (15) and distorts firms' job creation and wage-setting relative to the competitive equilibrium. To see how, note that the first-order condition for vacancy posting $a_{ijt}$ is

$$\frac{\kappa_i}{\lambda_f(\theta_{ijt})} = Q_{t,t+1}\left(\nu_{ijt+1} - W_{ijt+1}\right). \tag{16}$$

That is, firms post job vacancies until the marginal cost of hiring a worker, $\kappa_{it}/\lambda_f(\theta_{ijt})$ on the left side of (16), equals the marginal benefit, the right side. Here the marginal benefit is the present value having a worker that starts working $t+1$ minus the wages paid to that worker. Because the monopsony distortion lowers a firm's present value of employing workers, $\nu_{ijt}$, firms post fewer vacancies and hire fewer workers than in a competitive search equilibrium.

Combining the first-order conditions for market tightness $\theta_{ijt}$ and wages $W_{ijt+1}$ gives

$$W_{ijt+1} = \eta\nu_{ijt+1} - (1-\eta)V_{it+1}.$$

Hence, the present value of wages $W_{ijt+1}$ is the marginal value of workers of type $i$ to firm $j$ minus a negative term, namely the disutility of work of workers of type $i$, weighted by the elasticity of the matching function through $\eta$ and $1-\eta$. Thus, since a firm's monopsony power lowers the value of a worker to a firm, $\nu_{ijt+1}$, it also lowers wages.

**Firm Investment Decisions.** We now examine firms' investment decisions, which determine the evolution of the capital stock and the different types of capital it consists of. Together, these decisions determine how the elasticity of substitution across workers evolves over time. The investment problem is potentially complex because a firm must choose $X_{jt}(v)$, namely, the amount to invest in capital of *each* possible labor intensity $v \in \mathbb{R}^I$ in each period $t$. As we show in the next proposition, this problem is tractable since in each $t$, a firm invests in only one type of labor intensity of capital, $v_t$. We characterize this problem under a *single-peakedness* assumption that

$$\sum\nolimits_{\tau=1}^{\infty} Q_{t,t+\tau}(1-\delta)^{\tau-1}\int_{\underline{\varepsilon}_{t,t+\tau}}^{\infty}\left[A_t\varepsilon f(v) - \sum\nolimits_i \chi_{it+\tau}v_i\right]\pi(\varepsilon)d\varepsilon \tag{17}$$

is single-peaked in $v$. We drop the firm subscript $j$ for simplicity.

**Proposition 2.** *Under (17), $X_t(v) > 0$ for at most one type of capital denoted $v_t$ that solves*

$$v_t = \arg\max_v \sum_{\tau=1}^{\infty} Q_{t,t+\tau}(1-\delta)^{\tau-1} \int_{\underline{\varepsilon}_{t,t+\tau}}^{\infty} \Big[A_t \varepsilon f(v) - \sum_i \chi_{it+\tau} v_i\Big] \pi(\varepsilon) d\varepsilon, \tag{18}$$

*where $\underline{\varepsilon}_{t,t+\tau} = \underline{\varepsilon}(v_t, A_t; \chi_{t+\tau})$ is the idiosyncratic productivity threshold for capital made in period $t$ to be utilized in period $t+\tau$ for $\tau \geq 1$. Also, if $X_t(v_t)$ is strictly positive, then*

$$1 = \sum_{\tau=1}^{\infty} Q_{t,t+\tau}(1-\delta)^{\tau-1} \int_{\underline{\varepsilon}_{t,t+\tau}}^{\infty} \Big[A_t \varepsilon f(v_t) - \sum_i \chi_{it+\tau} v_{it}\Big] \pi(\varepsilon) d\varepsilon. \tag{19}$$

To prove this proposition, we first consider the first-order condition for $K_{jt+\tau}(v, A_t)$, which is the capital installed in period $t$ with productivity $A_t$ in use in $t+\tau$, given by

$$q_{t,t+\tau}(v) = \int_{\varepsilon} u_{t+\tau}(v, A_t, \varepsilon) \Big[A_t \varepsilon f(v) \pi(\varepsilon) - \sum_i \lambda_{it+\tau}(v, A_t, \varepsilon) v_i\Big] d\varepsilon.$$

Using the cutoff form of the utilization rule and the first-order condition for $N_{it}(v, A_{t-\tau}, \varepsilon)$, namely, $\lambda_{it}(v, A_{t-\tau}, \varepsilon) = \chi_{it}\pi(\varepsilon)$, gives $q_{t,t+\tau}(v) = \int_{\underline{\varepsilon}_{t,t+\tau}}^{\infty} \Big[A_t \varepsilon f(v) - \sum_i \chi_{it+\tau} v_i\Big] \pi(\varepsilon) d\varepsilon$ which when substituted in the investment first-order condition $\mu_t(v) = 1 - \sum_{\tau=0}^{\infty} Q_{t,t+\tau}(1-\delta)^{\tau-1} q_{t,t+\tau}(v)$ gives

$$\mu_t(v) = 1 - \sum_{\tau=1}^{\infty} Q_{t,t+\tau}(1-\delta)^{\tau-1} \int_{\underline{\varepsilon}_{t,t+\tau}}^{\infty} \Big[A_t \varepsilon f(v) - \sum_i \chi_{it+\tau} v_i\Big] \pi(\varepsilon) d\varepsilon. \tag{20}$$

We use (20) to show that firms invest in only one type of capital in $t$. To do so, note that since $\mu_t(v)$ is a Lagrange multiplier, it has a minimum value of zero. Also, if the second term on the right side of (20) is single-peaked in $v$, then there exists a unique value of $v$—denoted by $v_t$—that achieves that minimum.[10] Then, $\mu_t(v) > 0$ for all $v \neq v_t$, which by complementary slackness implies that $X_t(v) = 0$ for all such $v$. For the optimal type $v_t$, (20) holds with $\mu_t(v_t) = 0$, which establishes (19). Since this optimal type $v_t$ minimizes the right side of (20), it equivalently solves the problem (18). So we have established the proposition.

The first-order condition for the choice of the optimal labor intensity $\{v_i\}$ in (18) is

$$\sum_{\tau=1}^{\infty} Q_{t,t+\tau}(1-\delta)^{\tau-1} \int_{\underline{\varepsilon}_{t,t+\tau}}^{\infty} A_t \varepsilon \frac{\partial f(v_t)}{\partial v_i} \pi(\varepsilon) d\varepsilon = \sum_{\tau=1}^{\infty} Q_{t,t+\tau}(1-\delta)^{\tau-1} \int_{\underline{\varepsilon}_{t,t+\tau}}^{\infty} \chi_{it+\tau} \pi(\varepsilon) d\varepsilon, \tag{21}$$

so firms choose this labor intensity to equate the present value of the marginal product of a type-$i$ worker to the present value of the shadow cost of employing that worker over the life of a unit of capital of type $v_t$. The discounting of these present values is by the price of output $Q_{t,t+\tau}$ in period $t+\tau$ times the share of capital remaining, $(1-\delta)^{\tau-1}$. The marginal product of type-$i$

[10]Note that the expression (17) being single-peaked in $v$ is weaker than it being concave in $v$. In our environment (17) with partial utilization of capital, it is not concave. In this sense, this proposition generalizes that of Proposition 3 in Atkeson and Kehoe (1999) which assumed full utilization and in which the analogous term to (17) was concave. We have found that the term in (17) is single-peaked in all our quantitative work.

worker in period $t+\tau$ is the product of the average idiosyncratic productivity of the capital that is utilized, $\int_{\underline{\varepsilon}_{t,t+\tau}}^{\infty} \varepsilon\pi(\varepsilon)d\varepsilon$, the vintage productivity of the capital, $A_t$, and the marginal product of the intensive-form putty-putty production function $f_i(v_t)$. The shadow cost of the worker in period $t+\tau$ is the share of the capital that is utilized, $\int_{\underline{\varepsilon}_{t,t+\tau}}^{\infty} \pi(\varepsilon)d\varepsilon$, times the shadow cost of the worker using that capital, $\chi_{it+\tau}$. Finally, equation (19) equates marginal revenues, namely the present value of a firm's shadow profits over the life of the marginal unit of capital of type $v_t$, to the marginal cost in consumption goods of that unit of capital, which is 1.

### 2.3 Equilibrium Along the Balanced Growth Path

To build intuition for how our model works we start with the analysis of its BGP. To be consistent with balanced growth, we assume that vacancy-posting costs grow at the same rate as the economy, $\kappa_{it} = (1+g)^t\kappa_i$, and that the representative type-$i$ family has preferences

$$U_t(c_{it}, s_{it}, n_{it}) = \log\left[c_{it} - (1+g)^t v(n_{it}) - (1+g)^t h(s_{it})\right].$$

As in Greenwood, Hercowitz and Huffman (1988), these preferences feature no wealth effects on labor supply so that equilibrium in the labor market is independent of the consumption allocation along the BGP. We scale the disutility of work and labor market search by $(1+g)^t$ so that as consumption and wages along the balanced growth path grow at rate $g$, and the optimal amount of work and search stay constant. Consumption, investment, output, wages, and the disutility of work and job search all grow with the economy. With "tilde", we denote detrended versions of these variable, say detrended wages $\widetilde{w}_{it} = w_{it}(1+g)^{-t}$, which are constant along the BGP. Letting $\Pi^u(\underline{\varepsilon}) = \int_{\underline{\varepsilon}}^{\infty} \pi(\varepsilon)d\varepsilon$ denote the fraction of capital above a cutoff level $\varepsilon$, $\Pi^p(\underline{\varepsilon}) = \int_{\underline{\varepsilon}}^{\infty} \varepsilon\pi(\varepsilon)d\varepsilon$ denote the average idiosyncratic productivity of utilized capital, and $\widetilde{\beta} = \beta/(1+g)$, we have

**Lemma 3.** *Along the balanced growth path, the labor allocations and wages are determined by*

*a) optimal cut-off for idiosyncratic productivity of capital* $\underline{\varepsilon}_1 = (1+g)(1-\alpha)m(\underline{\varepsilon}_1)$ *where*

$$m(\underline{\varepsilon}_1) = \frac{\sum_{\tau=0}^{\infty} \beta^{\tau+1}(1-\delta)^\tau(1+g)^{-\tau-1}\Pi^p((1+g)^t\underline{\varepsilon}_1)}{\sum_{\tau=0}^{\infty} \beta^{\tau+1}(1-\delta)^\tau\Pi^u((1+g)^t\underline{\varepsilon}_1)};$$

*b) zero profits on investment* $1 = \alpha\left[\sum_{\tau=0}^{\infty} \widetilde{\beta}^{\tau+1}(1-\delta)^\tau\Pi^u((1+g)^t\underline{\varepsilon}_1)\right] f(\widetilde{v})$;

*c) optimal labor intensities satisfy* $\widetilde{\chi}_i = f_i(\widetilde{v})m(\underline{\varepsilon}_1)$;

*d) free entry in vacancy posting and wages satisfy*

$$\kappa_i = \beta\lambda_f(\theta_i)\frac{f_i(\widetilde{v})m(\underline{\varepsilon}_1) - \widetilde{w}_i - v'(n_i)/\omega}{1-\beta(1-\sigma)} \text{ and } \widetilde{w}_i = \eta[f_i(\widetilde{v})m(\underline{\varepsilon}_1) - v'(n_i)/\omega] + (1-\eta)v'(n_i);$$

*e) household optimal search* $h'(s_i) = \beta\lambda_w(\theta_i)[\widetilde{w}_i - v'(n_i)]/[1-\beta(1-\sigma)]$;

*together with labor market clearing $\mu_i n_i/\widetilde{v}_i = \mu_1 n_1/\widetilde{v}_1$, $i = 2, \ldots, N$, market tightness $\theta_i = a_i/s_i$, and the steady state law of motion for employment, $\sigma n_i = \lambda_w(\theta_i)s_i$.*

Notice that these equilibrium equations define a system of $1+6N$ equations in $1+6N$ variables $\underline{\varepsilon}_1$, $\widetilde{v}_i$, $n_i$, $\theta_i$, $s_i$, $a_i$, and $\widetilde{w}_i$. So in this precise sense, with our preferences, the labor block of the model can be solved independently of the rest of the model, which determines consumption, investment, output, and so on. It will prove useful to define the corresponding labor allocations for the *competitive search equilibrium*, $\underline{\varepsilon}_1^c$, $\widetilde{v}_i^c$, $n_i^c$, $\theta_i^c$, $s_i^c$, $a_i^c$, and $\widetilde{w}_i^c$, as those the solve the same set of conditions except for those under $d$), which are replaced by

$$\kappa_i = \lambda_f(\theta_i)[f_i(\widetilde{v})m(\underline{\varepsilon}_1) - \widetilde{w}_i]/(\rho+\sigma) \text{ and } \widetilde{w}_i = \eta[f_i(\widetilde{v})m(\underline{\varepsilon}_1)] + (1-\eta)v'(n_i). \tag{22}$$

It is also useful to note for later that the cutoff level of idiosyncratic productivity is the same under the monopsonistic and the competitive allocations. The reason is that $a$) and $b$) determine the cutoff independently of the other equations. We summarize this feature with the following lemma.

**Lemma 4.** *Along any BGP, the utilization threshold $\underline{\varepsilon}_1$ is independent of the degree of firm monopsony power $\omega$, namely, $\underline{\varepsilon}_1 = \underline{\varepsilon}_1^c$.*

Intuitively, the utilization schedule is tied to capital's share of income, which is constant given our Cobb-Douglas formulation for $f(v)$. In Section 3, this lemma will be useful because it also implies that the utilization schedule is independent of labor market policies, which helps shed light on why the long-run effects of the those policies are similar to those in the putty-putty model.

Condition $c$) for firms choice of optimal labor intensities on new capital equates the the shadow cost of type-$i$ labor $\widetilde{\chi}_i$ to the marginal product of that labor, $f_i(\widetilde{v})$, adjusted for changes in utilization and the relative vintage productivity over the life of the machine. Since this utilization adjustment is a feature of the capital stock, it is common for all worker types $i$, in that the second term in this expression does not depend on $i$. Thus, taking ratios of this expression across two workers of different types $i$ and $i'$ gives $\widetilde{\chi}_i/\widetilde{\chi}_{i'} = f_i(\widetilde{v})/f_{i'}(\widetilde{v})$. In our quantitative work in Section 4, we will find that the markdown of wages below workers' marginal products is approximately $\widetilde{\chi}_i \approx (1+1/\omega)\widetilde{w}_i$ along the BGP. This relationship implies that the ratio of these two workers' wages satisfies $\widetilde{w}_i/\widetilde{w}_{i'} \approx f_i(\widetilde{v})/f_{i'}(\widetilde{v})$. Hence, the long-run elasticities of substitution between workers are inherited from the putty-putty production function $f(v)$ and the putty-clay aspect of production is irrelevant because by being able to choose different capital intensities, a firm in the long run can freely adjust the use of any inputs, as shown in Figure 1.

Since $f(v) = G(v)^{1-\alpha}$ is the intensive form of $F(K,N) = K^\alpha G(N)^{1-\alpha}$, the wage ratio $\widetilde{w}_i/\widetilde{w}_{i'}$

approximately equals $G_i(\widetilde{v})/G_{i'}(\widetilde{v})$, which is the key condition used in the literature to estimate the elasticity of substitution in production between different worker groups. We exploit this fact later to rely on existing estimates of labor-labor substitutability to pin down the parameters of $G(n)$.

# 3 Understanding Labor Market Policies

We now turn to the labor market policies that we examine to illustrate our framework. To motivate this discussion, we first show that a combination of worker type-specific minimum wages and vacancy-posting subsidies can eliminate the monopsony distortions in our model. Although these policies provide a useful benchmark, they are difficult to implement because they require targeting each worker type. Therefore, we analyze two simple policies that are often used in practice to help improve the labor market outcomes of low-earning workers: a uniform minimum wage or a targeted transfer conditional on working, such as the Earned Income Tax Credit (EITC). In this section, we use the BGP to qualitatively explain how the policies affect outcomes in the long run.

## 3.1 Distortions to Wages and Job-Finding Probabilities

We start by examining the distortions that monopsony power introduces to wages and job creation.

*Monopsony Distortions.* Firms in our model compete for workers along two distinct margins: the value of wages they offer and, through the number of vacancies they post, the tightness of the market for their jobs. Both margins are distorted. Wage distortions arise because hiring a marginal worker of a given type increases the marginal disutility of work of all inframarginal hires of a given type in the same family. Hence, a firm needs to compensate those inframarginal workers with a higher wage when it hires additional workers of the same type. In this precise sense, firms face an upward-sloping supply curve of workers in offered wages. Likewise, as a firm increases its vacancies for a given type of worker, it increases the worker's job-finding probability. So a firm also faces an upward-sloping curve for workers in offered vacancies and, hence, in offered market tightness.

Formally, the parameter $\omega$ enters the firm's first-order conditions associated with the choices of wages $W_{ijt+1}$ and market tightness $\theta_{ijt}$, leading to a downward distortion of both relative to their levels in a competitive search equilibrium. The markdown for wages is an extension to a dynamic setting of the standard markdown that arises in imperfectly competitive models of the labor market with static labor supply. The markdown for market tightness $\theta_{ijt}$ is novel and reflects the idea that as more vacancies are posted—holding wages fixed—market tightness increases. Hence, in this precise sense, firms' monopsony power distorts both the *supply* and the *demand* for labor.

To build further intuition, consider the participation constraint along the BGP:

$$\frac{\lambda_w(\theta_{ij})}{\rho+\sigma}\left[\widetilde{w}_{ij} - v'(n_i)\left(\frac{n_{ij}}{n_i}\right)^{\frac{1}{\omega}}\right] \geq \widetilde{\mathcal{W}}_i, \tag{23}$$

where $\rho = \frac{1}{\beta} - 1$ is the rate of time preference, $\widetilde{w}_{ij}$ is the detrended flow wage, and $v'(n_i)$ is the flow disutility of labor supply. Now, suppose that firm $j$ contemplates hiring $dN_{ij} = \mu_i dn_{ij}$ additional workers of type $i$ along the BGP. A firm attract such workers, and thus satisfy the participation constraint (23) for them, in two ways. First, it can raise the flow wage it pays. The required wage increase—obtained by differentiating (23) holding as an equality with respect to $\widetilde{w}_{ij}$ and $n_{ij}$ holding $\lambda_w(\theta_{ij})$ fixed—at a symmetric equilibrium gives

$$\frac{d\widetilde{w}_{ij}}{dn_{ij}} = \frac{v'(n_i)}{\omega n_i} > 0. \tag{24}$$

Online Appendix A converts (24) into an elasticity and shows that $d\log n_{ij}/d\log \widetilde{w}_{ij} \approx \omega$. Hence, the parameter controlling the degree of monopsony power can, loosely speaking, be interpreted as firm-specific labor supply elasticity in response to permanent changes in wage offers along the BGP.

The second way in which firm $j$ can attract such workers is by posting more vacancies, thereby raising the job-finding probability $\lambda_w(\theta_{ij})$. The required increase in $\lambda_w(\theta_{ij})$—obtained by differentiating (23) with respect to $\lambda_w(\theta_{ij})$ and and $n_{ij}$ holding $\widetilde{w}_{ij}$ fixed—at a symmetric equilibrium gives

$$\frac{d\lambda_w(\theta_{ij})}{dn_{ij}} = \frac{(\lambda_w(\theta_{ij})/\omega)(v'(n_i)/n_i)}{\widetilde{w}_i - v'(n_i)} > 0. \tag{25}$$

Taken together, (24) and (25) illustrate how a worker's participation constraint encodes firm-specific (inverse) labor supply curves with respect to both wages and the job-finding rate, which are the two dimensions in which firm monopsony power manifests itself.

The empirical literature often summarizes the degree of firm monopsony power in wage setting using the *wage markdown*, namely, the ratio of workers' wages relative to their marginal product, $\widetilde{\chi}_i = f_i(\widetilde{v})m(\underline{\varepsilon}_1)$. Along the BGP this markdown is

$$\frac{\widetilde{w}_i}{f_i(\widetilde{v})m(\underline{\varepsilon}_1)} = \left[1 + \underbrace{\frac{1}{\omega} \times \frac{v'(n_i)}{v'(n_i) + \frac{\eta}{1-\eta}(\rho+\sigma)\frac{\kappa_i}{\lambda_f(\theta_i)}}}_{\text{monopsony component}} + \underbrace{\frac{(\rho+\sigma)\frac{\kappa_i}{\lambda_f(\theta_i)}}{v'(n_i) + \frac{\eta}{1-\eta}(\rho+\sigma)\frac{\kappa_i}{\lambda_f(\theta_i)}}}_{\text{efficient component}}\right]^{-1}, \tag{26}$$

The *monopsony component* reflects firms' monopsony power, which distorts allocations relative to an efficient equilibrium. The *efficient component* exists even in the absence of firm monopsony power ($\omega \to \infty$) and captures that hiring a worker requires firms to incur the costs $\kappa_i/\lambda_f(\theta_{ij})$ for which they must be compensated over the life of a match. Indeed, in each period a firm needs only recoup the annuitized value of such costs, $(\rho+\sigma)\kappa_i/\lambda_f(\theta_i)$, which we find to be quantitatively small.

### 3.2 Scope for Policy

In our economy, monopsony power, as captured by the degree $\omega$ of substitutability of jobs in workers' preferences, is the only source of distortions: when $\omega$ becomes large, monopsony power disappears and the equilibrium becomes efficient.

**Lemma 5.** *As $\omega \to \infty$, the allocations of the monopsonistically competitive search equilibrium converge to those of the competitive search equilibrium, which are efficient.*

Hence, implementing the efficient allocation requires a set of instruments that can support the corresponding competitive search equilibrium. We show that it is feasible to do so with type-specific minimum wages combined with type-specific vacancy-posting subsidies.

**Proposition 6.** *In the BGP of the symmetric monopsonistically competitive search equilibrium the efficient allocations can be implemented through a combination of a minimum wage for each worker type $i$ equal to that type's competitive search equilibrium wage $w_i^c$ and a subsidy to vacancy posting for each worker type $i$ equal to $1-\tau_i = [\widetilde{w}_i^c - v'(n_i^c)]/[\widetilde{w}_i^c - v'(n_i^c)(1-1/\omega)]$.*

To understand this result, recall that the conditions defining a competitive equilibrium differ from those defining a monopsonistically competitive equilibrium only in the free entry condition and the wage equation. If our model had static relationships between firms and workers, the government could implement the competitive allocations by setting the minimum wage policy $\underline{w}_i = \widetilde{w}_i^c$ for each $i$. To see what would happen in our model with dynamic relationships, note first that under such policies, since firms want to lower wages below their competitive levels, the minimum wage constraints would all bind. Hence, under these policies the monopsony wage equation would be replaced by $\underline{w}_i = \widetilde{w}_i^c$ and that would fix the direct distortion to the wage and the indirect distortions to optimal search and vacancy posting that emanate from them.

But that policy alone would not eliminate the *direct* distortions to the vacancy posting condition from the $v'(n_i^c)/\omega$ term. To build intuition, suppose by way of contradiction that the minimum wage policy did implement the competitive allocations. Online Appendix B shows that evaluating the vacancy-posting condition at the competitive allocations with this minimum wage policy gives

$$\kappa_i = \lambda_f(\theta_i^c)\frac{f_i(\widetilde{v}^c)m(\underline{\varepsilon}_1) - \widetilde{w}_i^c}{\rho+\sigma}\left[\frac{\widetilde{w}_i^c - v'(n_i^c)}{\widetilde{w}_i^c - v'(n_i^c)(1-1/\omega)}\right]. \tag{27}$$

Comparing (27) to its competitive analog, (22), we see that in the competitive allocations the term in square brackets in (27) equals 1. Since in (27) the term in square brackets in less than one, at such policies a monopsonist would choose to deviate down to lower the job-finding rate below the competitive level. Intuitively, this policy fails because it only eliminates one of the two distortions.

To eliminate the distortion to vacancy posting, we use a second set of instruments, namely, type specific entry subsidies $\tau_i$. We choose these subsidies so that when evaluating the monopsony free entry condition under the joint policies of entry subsidies and minimum wages the condition

$$\kappa_i(1-\tau_i) = \lambda_f(\theta_i^c)\frac{f_i(\widetilde{v}^c) - \widetilde{w}_i^c}{\rho+\sigma}\left[\frac{\widetilde{w}_i^c - v'(n_i^c)}{\widetilde{w}_i^c - v'(n_i^c)(1-1/\omega)}\right]$$

becomes

$$\kappa_i = \lambda_f(\theta_i^c)\frac{f_i(\widetilde{v}^c)m(\underline{\varepsilon}_1) - \widetilde{w}_i^c}{\rho+\sigma}.$$

But it does, since dividing these two equations gives us $1-\tau_i$ from the proposition, completing the proof. Also note that as monopsony power diminishes, $\omega$ gets larger,the optimal subsidy falls to zero. Finally, if we expand the set of available instruments to include time-varying minimum wages and subsidies for vacancy creation, then the same result holds outside the BGP.

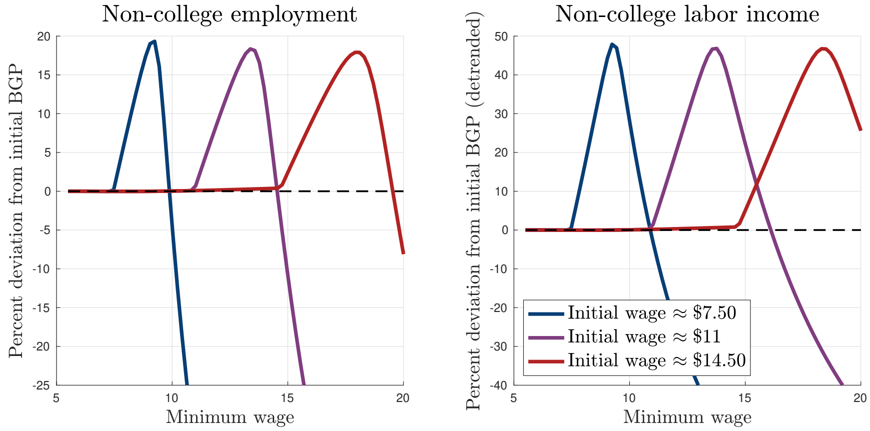
### 3.3 Minimum Wage

We study the long-run implications of a uniform minimum wage policy defined by a floor $\underline{w}_t = (1+g)^t\underline{w}$ on the flow paid wage for all workers that grows along with the economy. Along the BGP, this policy can be interpreted as introducing a fixed wage floor $\underline{w}$ in detrended terms. We compare the outcomes of this policy to those under an initial BGP in which wages are determined by the monopsonistically competitive equilibrium that follow $w_{ijt}^0 = (1+g)^t w_{ij}^0$.

For a given worker, an increase in the minimum wage up to that worker's competitive wage encourages firms to hire that worker because it eliminates the monopsony force a firm takes into account that hiring such a worker raises the wages of all inframarginal workers. For increases in the minimum wage above a worker's competitive wage, the minimum wage discourages firms from hiring that worker because it requires a firm to pay that worker more than the worker's marginal product. In our model, a small minimum wage can reduce the marginal cost because it increases the return from searching, making it easier to satisfy the participation constraint. Intuitively, for worker types for whom the minimum wage is only slightly larger than their current wage—namely, the minimum wage is set no higher than what that worker would be paid in the competitive equilibrium—the second effect dominates, and hiring increases. But for worker types whose current wage is much lower than the minimum wage, the first effect dominates, and hiring decreases.

**Intuition from the BGP.** To illustrate the forces that shape the long-run effects of the minimum wage in our model, Figure 3 shows how different levels of the detrended minimum wage $\underline{w}$ affect

Figure 3: Long-Run Effects of the Minimum Wage on Employment




Notes: Percentage change in BGP employment rates (left panel) and detrended labor income (right panel) of select non-college worker types as a function of the minimum wage.

the employment of three types of non-college workers in the new BGP.[11] These lines, generated by our quantified model described later show that long run employment has an inverted U shape with respect to the minimum wage. To understand why, consider the employment curve for the worker who earns the equivalent of a $7.50 wage in the initial BGP. Because of monopsony power, this initial wage $w_i$ is marked down relative to its efficient level in the associated competitive search equilibrium, $w_i^c$. Then, small increases in the minimum wage above the initial wage $w_i$ increase employment by bringing wages closer to their efficient level. This employment curve eventually reaches a peak, which is roughly the efficient wage level $w_i^c$. Increasing the minimum wage above this level leads employment to fall. Intuitively, for any type of worker, as the minimum wage is raised from the monopsony level to the efficient level, monopsony distortions are mitigated. For wages higher than that there is no monopsony distortion to eliminate, so employment falls.

**Distributional Implications.** Comparing the employment response across the three different worker types reveals a long-run distributional trade-off: a single minimum wage cannot simultaneously correct the monopsony distortion for all workers. Since workers have different levels of productivity, the levels of efficient wages $w_i^c$ and therefore the peaks of the employment curves are heterogeneous across workers. This feature implies that the minimum wage is too blunt an instrument to benefit all workers in the economy. For example, a small minimum wage would increase the labor income of workers initially earning low wages but would not affect the labor income of higher-productivity workers. By contrast, a high minimum wage would increase the income of

[11]Throughout the paper, we index each of these levels by the equivalent hourly wage which would bind for the same amount of workers in the 2017-2019 ACS data to which we calibrated the model.

higher-productivity workers but reduce employment and income of the lowest-productivity ones.

### 3.4 Targeted Transfers

We turn to the long-run effects of transfer payments to households conditional on labor market earnings. These transfer payments can be more targeted to specific worker types than a minimum wage and, therefore, be more effective at redistributing resources to their intended beneficiaries. In the long run, transfer payments can either alleviate or exacerbate monopsony distortions, depending on how they are designed. We then discuss a particular program, the Earned Income Tax Credit.

**Transfers.** We consider a general transfer policy such that a worker earning a flow wage $w_{ijt}$ in period $t$ receives an additional payment $T_t(w_{ijt})$, so that a worker's after-transfer earnings are $A_t(w_{ijt}) = w_{ijt}+T_t(w_{ijt})$. We let $A_t(w_{ijt}) = (1+g)^t A(w_{ijt}/(1+g)^t)$ for some time-invariant function $A$ so that the policy is consistent with balanced growth. These transfers are financed through a linear tax on firms' profits. We assume that wage payments, vacancy-posting costs, and investment costs are fully deducted from this profit tax, implying that the profit tax does not affect any of the firms' decisions. The first-order condition for households' search intensity $s_{ijt}$ becomes

$$-\frac{U_{sit}}{U_{cit}} = \lambda_w(\theta_{ijt})Q_{t,t+1}(W^H_{ijt+1} + V_{ijt+1}) \text{ if } s_{ijt} > 0, \tag{28}$$

where $W^H_{ijt+1} = A_{t+1}(w_{ijt+1})+Q_{t+1,t+2}(1-\sigma)A_{t+2}(w_{ijt+2})+Q_{t+1,t+3}(1-\sigma)^2A_{t+3}(w_{ijt+3})+\ldots$ is the present value of after-transfer wage payments to a type-$i$ family. Hence, transfer payments raise the return from searching on the right side of (28), incentivizing families to increase their search effort and the participation constraint becomes $\lambda_w(\theta_{ijt})(W^H_{ijt+1}+V_{ijt+1}) \geq \mathcal{W}_{it}$. This policy affects firms' decisions by changing the shadow price on this constraint from $\lambda_f(\theta_{ijt})a_{ijt}$ to $\lambda_f(\theta_{ijt})a_{ijt}/A'_t(w_{ijt+1})$. If the marginal transfer rate is positive, so that $A'_t(w_{ijt+1}) > 1$, then a marginal increase in a firm's wage offer is accompanied by an increase in the transfer payment. This effect lowers the shadow price of workers and reduces the monopsony distortion.

**Intuition from the BGP.** We next summarize the effect of the policy along the new BGP.

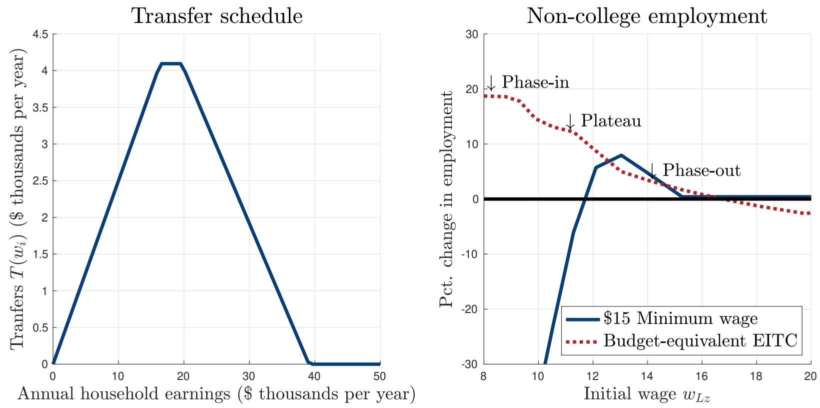
**Lemma 7.** *Under the transfer policy the labor allocations for the monopsonistic equilibrium solve the same equations as those in Lemma 3 with the optimal search condition replaced by*

$$h'(s_i) = \frac{\lambda_w(\theta_i)}{\rho+\sigma}\left[A(w_i) - v'(n_i)\right], \tag{29}$$

*and the vacancy posting condition replaced by*

$$\frac{\kappa_i}{\lambda_f(\theta_i)} = \frac{1}{\rho+\sigma}\left[\chi_i - w_i - \frac{v'(n_i)}{\omega A'(w_i)}\right]. \tag{30}$$

Figure 4: The EITC and its Long-Run Effects on Non-College Employment




Notes: EITC transfer schedule (left panel) and its impact on employment rates for individual non-college types in the new BGP, as a function of workers' wage in the initial BGP (right panel). For reference, right panel also plots effect of the $15 minimum wage on employment rates as well.

A positive *level* of transfers $A(w_i) > w_i$ stimulates workers' search for jobs in (29) by increasing the marginal gain to households of increasing search. In contrast, a positive *marginal* transfer $A'(w_i)$ in (30), stimulates firms to post vacancies because a firm internalizes that as it raises the wage to attract an additional worker, the government transfer rises along with it. So, on the upward-sloping portion of the EITC the government effectively subsidizes the hiring of additional workers.

**Earned Income Tax Credit.** We model the EITC using (detrended) the transfer schedule plotted in the left panel of Figure 4. We make this EITC schedule budget-equivalent to a $15 minimum wage by choosing the linear profit tax rate which funds the EITC to raise tax revenues equal to the firms' lost profits due to the minimum wage. In the *phase-in* region the transfer is proportional to household income with a positive marginal transfer rate of 25%; in the *plateau* region, the transfer is constant at its maximum benefit; and in the *phase-out* region, the transfer is reduced proportionally to any additional income with a negative marginal transfer rate of 22%. In the phase-in region, households face both a positive average subsidy rate, since the total tax credit is positive, and a positive marginal subsidy rate, since the credit is being phased in.

The right panel of Figure 4 plots the effects of the EITC on the employment of non-college workers in the new BGP. The lowest-wage workers in the phase-in region face a positive average transfer rate, which increases their search effort, and a positive marginal transfer rate, which decreases their monopsony distortion. Both of these effects contribute to higher long-run employment rates for this group. Workers in the plateau region still face a positive average transfer but their marginal transfer is zero, so their monopsony distortion is not impacted, leading to a smaller increase in employment. Finally, workers in the phase-out region face a negative marginal transfer rate, which exacerbates the monopsony distortion leading to a decline in their employment.

# 4 Quantification

We now parameterize our model in order to quantitatively study the speed of transition induced by these labor market policies. We set a model period to one month.

## 4.1 Parametrization

We assume that household type $i = (e, z)$ is an education-productivity pair with $e \in \{L, H\}$, where $L$ denotes workers with less than a bachelor's degree, $H$ denotes workers with at least a bachelor's degree, and $z$ denotes their permanent productivity level $z$. The productivity level $z$ is drawn from an education-specific log-normal distribution with mean zero and variance $\sigma_z^e$. A type-$i$ family has preferences within the Greenwood, Hercowitz and Huffman (1988) class:

$$U_t(c_{it}, s_{it}, n_{it}) = \log\left[c_{it} - (1+g)^t\left(\psi_n^e \frac{n_{it}^{1+1/\gamma_n}}{1+\gamma_n}\right) - (1+g)^t\left(\psi_s^e \frac{s_{it}^{1+1/\gamma_s}}{1+\gamma_s}\right)\right].$$

A family' disutility of work and labor market search are governed by the scale parameters $\psi_n^e$ and $\psi_s^e$, which differ between education groups, and the elasticity parameters $\gamma_n$ and $\gamma_s$. The long-run production function is $F(K, N) = K^\alpha G(N)^{1-\alpha}$, where $N$ is the vector of labor types. The labor aggregator $G(N)$ with $e \in \{L, H\}$, is given by

$$G(N) = \left[\xi N_L(N)^{\frac{\varphi-1}{\varphi}} + (1-\xi) N_H(N)^{\frac{\varphi-1}{\varphi}}\right]^{\frac{\varphi}{\varphi-1}} \text{ and } N_e(N) = \left[\sum\nolimits_{i\in I_e} z_i (N_i)^{\frac{\phi-1}{\phi}}\right]^{\frac{\phi}{\phi-1}}. \tag{31}$$

This CES production function features an outer nest between labor of each education group $e$ and an inner nest among labor of different productivity levels within an education group. Versions of this functional form has been extensively used and estimated in the labor economics literature (see Katz and Murphy (1992), Card and Lemieux (2001), and Borjas and Katz (2007)).

## 4.2 Disciplining Key Features

We begin by providing intuition for how we parameterize key features of our model that relate to the three main forces that determine the transition dynamics of the economy in response to labor market policies. These are *i*) capital utilization, which is connected to how quickly firms choose to let existing workers attrit in the short run when wages change; *ii*) the long-run elasticities of substitution across workers, which controls how much firms eventually substitute across different types of workers; and *iii*) the degree of firm monopsony power, which determines the magnitude of the labor market distortions that labor market policies can potentially correct.

**Capital Utilization.** The dynamics of capital utilization is shaped by three parameters: the growth rate of the economy $g$ and the depreciation rate of capital $\delta$, which together determine the average utilization rate of different vintages of capital, and the dispersion of idiosyncratic capital productivity $\sigma_\varepsilon$, which determines the share of capital utilized within a vintage of capital. We set the growth rate of the productivity of new capital to imply an average aggregate growth rate $g$ of 2% per year and the depreciation rate to 15% annually, which matches the aggregate capital depreciation rate excluding structures. We choose the within-vintage dispersion parameter $\sigma_\varepsilon$ to match the average utilization rate across all vintages of capital from the U.S. Census Quarterly Survey of Plant Capacity Utilization (QPC). This survey measures the market value of plants' production relative to what they could have produced if all capital in a firm was fully utilized and labor was freely available. We define this *capacity utilization* rate as

$$\frac{\sum_\tau (1-\delta)^\tau X_{t-\tau} A_{t-\tau-1} \Pi^p(\underline{\varepsilon}_{t,t-\tau-1})}{\sum_\tau (1-\delta)^\tau X_{t-\tau} A_{t-\tau-1}},$$

namely, output actually produced divided by the output that would be produced if all capital was fully utilized, where $\Pi^p(\underline{\varepsilon})$ is the *total productivity of utilized capital.* The higher is $\sigma_\varepsilon$, the more capital is below the utilization threshold, and so the lower is capacity utilization; see Figure 2. We choose $\sigma_\varepsilon$ so that the average capacity utilization rate is 75% as in post-2000 data.

**Monopsony Power.** We choose the degree of firm monopsony power, $\omega$, to match existing estimates of wage markdowns. A growing literature has measured wage markdowns in the United States and estimated that on average workers are paid between 0.65 and 0.85 of their marginal products (see Seegmiller (2021), Lamadon, Mogstad and Setzler (2022), Berger, Herkenhoff and Mongey (2022), and Yeh, Macaluso and Hershbein (2022)) For our baseline, we target the midpoint of this range, 0.75, but we also explore the sensitivity of our results to smaller wage markdowns. By the decomposition in (26), wage markdowns in our model reflect both the degree of firm monopsony power $\omega$ and the annuitized value of hiring costs $(\rho+\sigma)\kappa_i/\lambda_f(\theta_i)$. We use data on labor market flows to pin down hiring costs and residually infer $\omega$ to match the average wage markdown.

According to Manning (2011), who surveys the empirical literature on hiring costs, a plausible range for (average) hiring costs is around 34%-156% of one month's wages. However, since the average job-separation rate is around 2.5%-4% per month, only a small fraction of these costs needs to be recouped per month by a firm. In our calibration, average hiring costs are approximately 64% of one month's wage, the job-separation rate is 2.8% per month, the interest rate is 4% annually, and the growth rate is 2% annually. Therefore, the annuitized value of hiring costs that firms must recoup each period is only $(\rho + \sigma)\frac{\kappa_i}{\lambda_f(\theta_i)} \approx 0.03 \times 0.64 = 1.9\%$ of one month's wage. Through the lens of our model, this annuitized value of hiring costs along with our estimates of $v'(n_i)$ yields that $\omega = 3.07$. Hence, the efficient component of the markdown is modest and so much so that reasonable variations in it have a very small effect on the estimate of $\omega$. For example, if we assumed hiring costs were zero, $\omega$ would be equal to 3 and if we assumed they were twice as large as in our baseline, $\omega$ would be equal to 3.18.

**Long-Run Elasticities of Substitution.** The long-run elasticities of substitution between workers of different abilities in $G(N)$ are governed by $\phi$, the elasticity of substitution within education groups, and $\varphi$, the elasticity of substitution across education groups. We set the cross-group elasticity $\varphi$ to 1.4 and the within-group elasticity $\phi$ to 4, which are values in the range of the estimates for them in the labor literature (Card and Lemieux (2001) and Borjas and Katz (2007)). Work in this literature assumes that different observable groups of workers—for instance, by age or immigration status—correspond to different productivity groups $z$ and use residual variation in the labor supply of these groups to estimate these elasticities of substitution based on the implied variation in their wages. Although the specific sources of variation that identify these elasticities differ across studies, the range of these estimates is broadly consistent with our values (see, for example, Card and Lemieux (2001)). Our value of the between-group elasticity $\varphi$ is also consistent with the benchmark value in Katz and Murphy (1992). A concern with borrowing estimates of $\phi$ and $\varphi$ from the literature is that the empirical variation under which these values have been estimated is different from the time variation on which our analysis mostly relies. In Section 4.4, we address this concern by showing that the estimators of these parameters common in the literature recover the true long-run elasticity nearly exactly based on data generated from our model.

## 4.3 Details

As for the remaining parameters, we externally fix a subset of them and then endogenously pin down the remaining ones to reproduce important features of U.S. labor markets.

**Fixed Parameters.** As shown in Table 1, we set the share of college-educated households to 31% in order to match their proportion in the American Community Survey (ACS) over the period between 2017 and 2019. This share and the distribution of productivity $z$—pinned down as explained below—jointly determine the weights $\mu_i$. We choose a value for the discount factor $\beta$ of $[1.04/(1+g)]^{1/12}$ so that the annualized real interest rate $r$ equals 4%. We set the parameter $\gamma_n$ of the utility function, which primarily controls the elasticity of labor supply, to 1. Now, there exists a locus of values for the parameters $\gamma_s$ and $\psi_s$ governing the disutility of labor market search that imply

Table 1: Fixed Parameters

| Parameter | Description | Value |
|---|---|---|
| *Households* | | |
| $N_H/(N_H+N_L)$ | Share of college workers | 0.31 |
| $\beta$ | Discount factor | $(\frac{1.04}{1+g})^{-1/12}$ |
| $\gamma_n$ | Labor supply elasticity | 1.00 |
| $\gamma_s$ | Search supply elasticity | 5.00 |
| $\chi_s$ | Scale of search disutility | $3.8 \times 10^6$ |
| *Production function* | | |
| $\varphi$ | Long-run elasticity b/t $N_{ht}$ and $N_{\ell t}$ | 1.40 |
| $\phi$ | Long-run elasticity within education group | 4.00 |
| $\delta$ | Capital depreciation (annualized) | 15% |
| $g$ | Long-run growth rate (annualized) | 2.0% |
| *Labor market frictions* | | |
| $\sigma$ | Job destruction rate | 2.8% |
| $\eta$ | Elasticity of matching function w.r.t. unemployed | 0.50 |

approximately identical labor market outcomes along the BGP but different responses of search effort to policy changes. We choose a pair of values on this locus, $\gamma_s = 5$ and $\psi_s^H = \psi_s^L = 3.8 \times 10^6$, that imply a relatively muted response of search effort to the minimum wage as in data (see Adams, Meer and Sloan (2022)).[12] We set the exogenous rate of match separation $\sigma$ to 2.8% per month, which is the Abowd-Zellner corrected estimate of the separation rate estimated by Krusell et al. (2017) from the Current Population Survey (CPS). We fix the elasticity of the matching function with respect to the measure of searchers $\eta$ to 0.5 as in Ljungqvist and Sargent (2017).

**Fitted Parameters.** We set the remaining parameters in Table 2 to match the statistics in Table 3. The parameters of the worker productivity distribution and the production function govern the distribution of income in the economy. We then choose the dispersion of productivity $\sigma_z^e$ to match the ratio of the 50-th to the 10-th percentiles of the wage distribution within each education group $e \in \{L, H\}$. This target ensures that we reproduce the left tail of each distribution, which is most directly affected by the policies we examine. Note that the weight placed on non-college labor in $G(N)$ in the long-run production function, $\xi$, determines the share of total labor income accruing to non-college workers, whereas the Cobb-Douglas coefficient on labor in production, $1-\alpha$, controls labor's total share of income in aggregate output. We pin down these parameters accordingly.

We choose the parameters governing hiring costs to match labor market flows so that the degree of monopsony power $\omega$ is determined residually so as to reproduce average wage markdowns. Hiring costs $\kappa_{it}/\lambda_f(\theta_{it})$ are pinned down by both the size of vacancy-posting costs ($\kappa_0$ in $\kappa_{it} = \kappa_0(1+g)^t z_i$) and the efficiency of the matching function (the parameter $B$ in $m(a_{it}, s_{it}) = B s_{it}^{\eta} a_{it}^{1-\eta}$) since $\lambda_f(\theta_{it}) = B\theta_{it}^{-\eta}$. As is typical in search models, the size of the vacancy-posting costs $\kappa_0$ is

[12] These parameters primarily govern the response of unemployment and labor force participation to the policies we analyze. Our results focus on changes in the employment rate which are unaffected by the choice of these parameters.

Table 2: Fitted Parameters

| Parameter | Description | Value |
|---|---|---|
| *Worker productivity distribution* $\log \mathcal{N}(0, \sigma_z^e)$ | | |
| $\sigma_z^L$ | SD of non-college $z$ | 0.69 |
| $\sigma_z^H$ | SD of college $z$ | 0.84 |
| *Production function* | | |
| $\alpha$ | Exponent on capital in production | 0.24 |
| $\xi$ | Weight on non-college labor in production | 0.42 |
| *Monopsony* | | |
| $\omega$ | Monopsony power | 3.07 |
| *Search frictions* | | |
| $B$ | Matching function productivity | 0.50 |
| $\kappa_0$ | Vacancy-posting costs (normalization) | 0.32 |
| *Labor Disutility* | | |
| $\psi_n^L$ | Weight of non-college labor disutility | 3.84 |
| $\psi_n^H$ | Weight of college labor disutility | 4.79 |
| *Capital Utilization* | | |
| $\sigma_\varepsilon$ | Dispersion in capital productivity | 0.25 |

Notes: Parameters that are endogenously chosen to match the statistics in Table 3.

not separately identified from the efficiency $B$ of the matching function because, by the free-entry condition, only the ratio $\kappa_{it}/(B\theta_{it}^{-\eta})$ is relevant for labor market outcomes. Following Shimer (2005), we normalize $\kappa_0$ so that average market tightness is 1 and choose $B$ to match an average unemployment rate of 5.9% in the United States before the Great Recession. This value implies an average job-finding rate of 0.44 in our model, compared to 0.45 in the data (see Shimer (2005) and Kehoe et al. (2023)). The remaining parameters relevant for labor market outcomes are those for the disutility of labor supply, $\{\psi_n^e\}$, which are pinned down by the average employment rates of each education group. Finally, the dispersion of the distribution of idiosyncratic capital productivity $\sigma_\varepsilon$ is primarily determined by the average capacity utilization rate, as described above.

## 4.4 Model Fit and Validation

We now discuss how the model fits well the targeted moments, how it also matches key untargeted features of the data, and how we validate our key parameter values.

**Wage Distribution and Markdowns.** As apparent from Table 3, the average wage markdown, the aggregate and college income shares, the employment rates by education group, the unemployment rate, and average capacity utilization are nearly identical in the model and in the data. The model also matches the 50-10 wage ratio within each education group, though not perfectly given our parsimonious parametrization of the productivity distribution—we specify worker productivity $z$ as log-normally distributed for each education group with the mean normalized to 1.

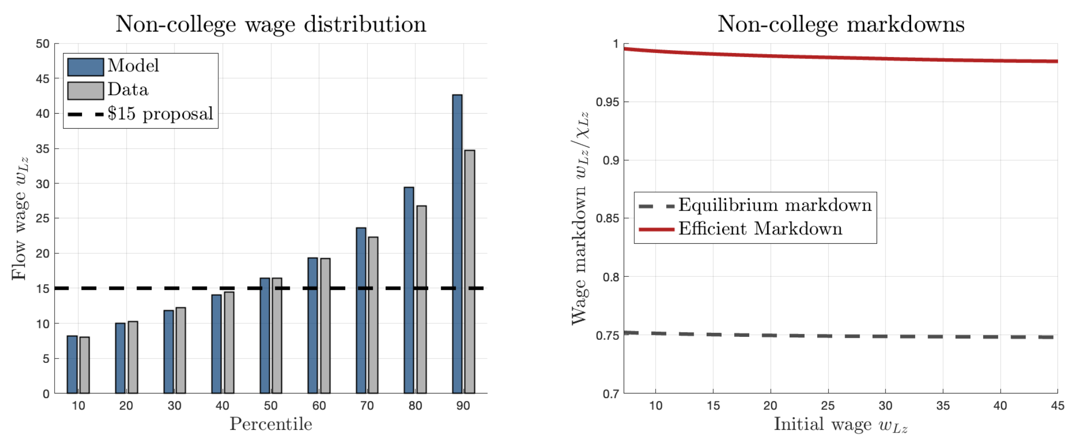
The model replicates key features of the distribution of wages of non-college workers. Figure 5 left panel compares this distribution in the model and in ACS data. In the model, we discipline it

Table 3: Targeted Moments

| Statistic | Description | Data | Model |
|---|---|---|---|
| *Average wage markdown* | | | |
| $\mathbb{E}[w_i]/\mathbb{E}[\chi_i]$ | Average wage markdown | 0.75 | 0.75 |
| *Wage Distribution, ACS 2017-2019* | | | |
| $w_{50}^L/w_{10}^L$ | Non-college 50-10 ratio | 2.04 | 2.00 |
| $w_{50}^H/w_{10}^H$ | College 50-10 ratio | 2.30 | 2.17 |
| *Income shares* | | | |
| $\mathbb{E}[w_i N_i]/Y$ | Aggregate labor share | 0.57 | 0.57 |
| $\mathbb{E}[w_z^H N_z^H]/\mathbb{E}[w_i N_i]$ | College income share | 0.55 | 0.55 |
| *Unemployment rate* | | | |
| $\mathbb{E}[s_i]/(\mathbb{E}[s_i]+\mathbb{E}[N_i])$ | Average unemployment rate | 5.9% | 5.9% |
| *Employment Rates* | | | |
| $\mathbb{E}^L[n_i]$ | Non-college employment rate | 0.47 | 0.47 |
| $\mathbb{E}^H[n_i]$ | College employment rate | 0.62 | 0.62 |
| *Capacity Utilization Rate* | | | |
| $\mathbb{E}[\Phi^p(\varepsilon_\tau)]$ | Average capacity utilization | 0.75 | 0.75 |

Notes: Average wage markdown is the midpoint of the range of estimates in the literature. The statistics on the wage distribution by education group, the college income share, and the employment rates are calculated using ACS data (2017-2019). The aggregate labor share is from Karabarbounis and Neiman (2014). The average unemployment rate is the average unemployment rate from the BLS. The average capacity utilization rate is drawn from U.S. Census Quarterly Survey of Plant Capacity Utilization (QPC).

Figure 5: Non-College Wages and Markdowns




Notes: The left panel plots deciles of the non-college wage distribution in our model and the ACS data (2017-2019). The right panel plots steady-state markdowns $w_{Lz}/\chi_{Lz}$ for select non-college worker types. "Equilibrium markdown" corresponds to our baseline model. "Efficient markdown" corresponds to the equilibrium of our model without monopsony power ($\omega \to \infty$). The $x$-axis is the wage $w_{Lz}$ of a type-$z$ non-college worker.

by targeting the 50-10 wage ratio to match the left tail of this distribution. Hence, we ensure that a $15 minimum wage binds for the same fraction of non-college-educated workers in the model as it does in the data—both 45%. Although the model slightly overpredicts wages at the top of the wage distributions for non-college workers, but these groups are barely affected by our experiments.

The right panel of Figure 5 plots the wage markdown for non-college workers of different productivity as a function of their wage, which is nearly constant at the value of 0.75 we target. As explained below, this is an important reason why our model implies values for $\phi$ in line with existing estimates, which are predicated on static, perfectly competitive models of the labor market.

**Long-Run Elasticities of Substitution Across Workers.** We rely on estimates from the labor economics literature to discipline the long-run elasticities of substitution across workers, $\varphi$ and $\phi$. These parameters are commonly estimated using the variation over time in the supply of workers with different levels of education, age or immigration status and the associated changes in their wage rates. This approach is based on static models of competitive labor markets with CES production—assumptions that do not hold in our model. Nonetheless, we now show that the implied estimators perform well on data generated from our model because they are derived from a reduced-form relationship between wage ratios and labor supplies implied by those models that approximately holds in our model as well.

To elaborate, we focus on the within-education group elasticity of substitution across workers, $\phi$, which is a key parameter.[13] We follow Card and Lemieux (2001)'s classic estimation strategy based on a static, competitive labor market in which firms have the same production function as our long-run function $G(N)$. For each education group, these authors assume that workers with different experience (age) differ in their productivity but that all workers with a given level of experience share the same productivity level. Hence, the ratio of wages of workers with different experience within an education group is informative about their relative productivity. Their estimator of the elasticity of substitution of workers with different productivity within an education group, $\widehat{\phi}$, exploits the variation in their relative wages induced by changes over time in the relative supply of workers with different levels of experience—which they interpret as exogenous.

We perform the following exercise to validate our model. Suppose that Card and Lemieux (2001) observe data generated by our model when the elasticity of substitution across workers is $\phi = 4$, a value in the low range of their estimates. We reproduce the variation in employment used by Card and Lemieux (2001) by allowing the measures of families $\mu_{it}$ of each type to vary over time and then construct a version of Card and Lemieux (2001)'s estimator $\widehat{\phi}$. In their model, the ratios of wages of workers with productivity $i$ and $i'$ from a given education group equals the ratio of their marginal products, which yields

$$\frac{w_{it}}{w_{i't}} = \frac{z_i}{z_{i'}}\left(\frac{N_{it}}{N_{i't}}\right)^{-\frac{1}{\widehat{\phi}}} \quad \text{or} \quad \Delta \log \frac{w_{it}}{w_{i't}} = -\frac{1}{\widehat{\phi}}\Delta \log \frac{N_{it}}{N_{i't}}, \tag{32}$$

where $\Delta$ denotes the difference operator across two time periods. By (32), we can construct a

[13] Our results are not sensitive to the across-group elasticity $\varphi$ since the dispersion in wages *within* education groups is much larger than the dispersion *across* education groups and the substitutability across groups is much lower.

Table 4: Literature's Estimation Strategy for Long-Run Elasticity $\phi$

| True Value | Card and Lemieux (2001) Variation |
|---|---|
| $\phi = 4$ | $\widehat{\phi} = 3.94$ |

Notes: Results from simulating the model for the time path of measures of families $\mu_{it}$ as described in text. *True value* reports the value of the long-run elasticity of substitution across workers $\phi = 4$ used to simulate the model. *Card and Lemieux (2001) Variation* reports the estimate $\widehat{\phi}$ from the regression in (32) using 5-year time differences, with the exception of the first observation which is a 10-year difference. See Online Appendix D for details.

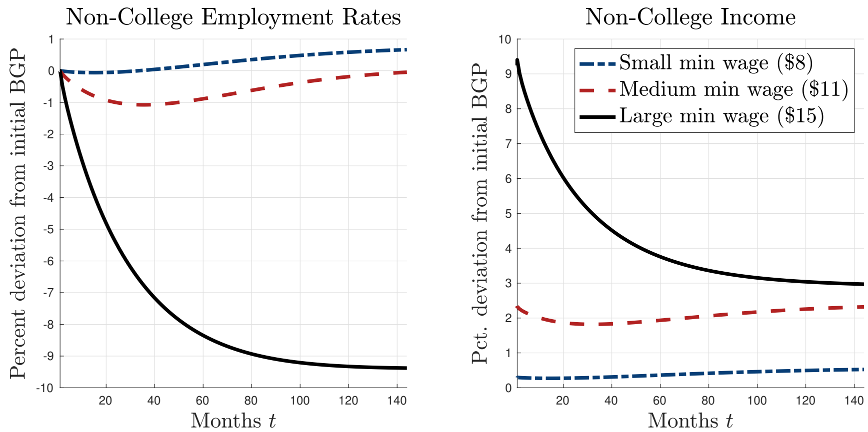
standard estimator $\widehat{\phi}$ by linearly projecting changes in wage ratios on changes in employment ratios. Following Card and Lemieux (2001), we assume that the first observation is a ten-year time difference and the rest are five-year ones; see Online Appendix D for details of this exercise. The second column of Table 4 shows that when we simulate data using $\phi = 4$, the estimator described recovers $\widehat{\phi} = 3.94$, which is very close to the true value. Hence, we conclude that the estimates of the long-run elasticity of substitution in the literature well discipline the value of $\phi$ in our model.

One reason that our estimator may not perform well is that workers are not paid their marginal products even in the long run (when putty-clay frictions are irrelevant). Quantitatively, the concern is unwarranted because wage markdowns are nearly constant across workers of different abilities along the BGP, the ratios of wages of workers with different productivity approximately equal the ratios of their marginal products along the BGP. Another concern is that using five-year differences in employment as in Card and Lemieux (2001) may severely bias our estimates of workers' substitutability in production downward toward non-substitutability, since our technology is Leontief in the short run. This concern is unwarranted in practice because in the experiment described, we modify the *supply* of workers (by varying the measure of families of each type) to replicate Card and Lemieux (2001)'s measured changes in the employment rate of workers with different productivity within an education group. Since the job-finding rate is high in the data and the model, firms rapidly absorb incoming workers by *adjusting* the labor requirements of new capital. Critically, this adjustment of labor-to-capital ratios is governed by the long-run function $f(v)$, as illustrated in Figure 1, which reflects the *long-run* substitutability across workers captured $\phi$.

# 5 Dynamic Effects of Labor Market Policies

We use our quantitative model to study the dynamic effects of various labor market policies. In particular, we highlight the quantitative importance of our key model features in explaining transition dynamics to these policies. Our main result is that the transition to a new long-run equilibrium is slow due to sluggish capital adjustment. This slow-transition result implies that

FIGURE 6: Dynamic Effects of the Minimum Wage




Notes: Transition paths of non-college employment (left panel) and detrended labor income (right panel) after the introduction of the minimum wage of various sizes, expressed relative to the initial BGP.

the short-run impacts of policies, which the empirical literature mostly focuses on, are not very informative about their *overall* effects. We begin by focusing on the minimum wage to illustrate the framework's key mechanisms because it allows us to clearly parse out how firms substitute among different types of workers. We then show that similar insights also apply to the EITC.

## 5.1 Overview of the Dynamic Effects of the Minimum Wage

We assume that the economy starts in a BGP without any labor market policies and workers are paid the flow wages $w^0_{ijt} = (1+g)^t w^0_{ij}$—note that this is equivalent to starting with a minimum wage of $7.25 since in our quantitative model this minimum wage binds on no one. In period 0, a minimum wage policy—represented by a floor $\underline{w}_t = (1+g)^t \underline{w}$ on the flow wage a firm can pay that grows with the economy—is unexpectedly introduced. Afterwards, agents perfectly anticipate the resulting transition path to the new BGP. In the initial BGP, the wages of some of the lower-productivity workers are below those dictated by the new minimum wage. In period 0, a firm may fire any measure of these workers it chooses. For the remaining workers, a firm must increase its flow wages to satisfy the minimum-wage constraint. Finally, for workers for whom the minimum wage does not bind, a firm must honor its wage commitments from the initial BGP, that is, $w_{ijt} \geq \max\{\underline{w}_{it}, w^0_{ijt}\}$. We begin by studying the aggregate effects of the policy on non-college workers and then examine the micro-level adjustments underlying these aggregate effects.

The aggregate effects of the minimum wage in the long run depend on its level. We consider three illustrative levels that differ in how they qualitatively affect aggregate non-college employment in the long run: a small minimum wage (equivalent to $8 per hour) that increases long-run employment, a medium minimum wage (equivalent to $11 per hour) that leaves long-run employment unchanged, and a large minimum wage (equivalent to $15 per hour) that reduces long-run employment.

The left panel of Figure 6 illustrates our main result that aggregate employment adjusts slowly to any such policy change. The small minimum wage induces firms to increase employment by reducing monopsony distortions for low-productivity workers, but it takes time to do so. Conversely, the large minimum wage induces firms to substitute away from low-productivity workers, but it takes more than ten years for this substitution to fully play out. In both cases, the short-run adjustment of employment is only a small fraction of its eventual long-run change.[14] The right panel of Figure 6 illustrates that these slow employment dynamics delays the long-run effect of the minimum wage on labor income. To see why, note that the minimum wage has two effects on the flow income $w_{it}n_{it}$: a direct effect through the increase in the wage $w_{it}$ of workers bound by the minimum wage, which occurs immediately, and an indirect effect through the implied change in employment $n_{it}$ that plays out over time. For the small minimum wage, these two effects reinforce each other, namely, the minimum wage directly increases wages upon impact, and indirectly increases employment gradually over time. Thus, here, the slow adjustment of employment to the minimum wage delays the total long-run *benefits* of the minimum wage on income.

For the large minimum wage, firms do not fire any of their initially employed workers on impact so the minimum wage leads to an immediate increase in the labor income of affected workers. Over time, however, firms reduce their hiring of the lowest-productivity workers and, hence, their labor income $w_{it}n_{it}$ slowly declines. We show below that these dynamics in labor income are critical for understanding the total effect of the minimum wage on individual workers.

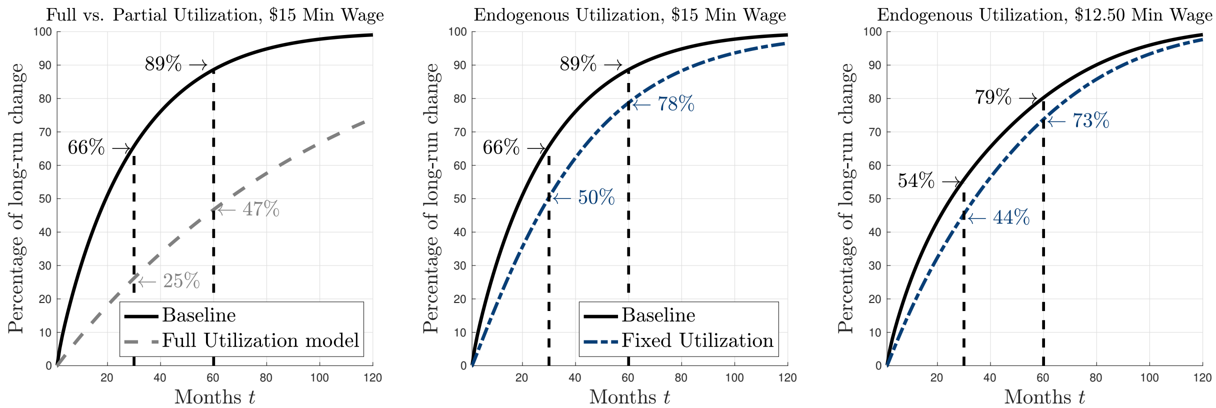
An important lesson from this analysis is that although a large increase in the minimum wage might seem desirable, since it leads to a large increase in the labor income of low-productivity workers in the short run, over time this gain is eroded as employment progressively falls. Hence, the short-run effects of a large minimum wage increase greatly overstate its total benefits. The counterpart of this result is that the long-run effects of such a policy by themselves overstate the total costs of it, because they altogether miss the short-run benefits.

## 5.2 Understanding the Slow Employment Dynamics

Here we analyze why our model generates slow employment dynamics. We also explain the role of our vintage structure defined by capital of vintage $t$ having average productivity that is $(1+g)^t$ times as large as capital built in period 0 and in every vintage, each unit of capital having having

[14]The medium minimum wage induces non-monotonic transition dynamics due to two conflicting forces: it induces firms to substitute *away* from some low-productivity workers, who have become relatively more expensive to employ, and to substitute *towards* higher-productivity workers, who are now relatively more attractive.

Figure 7: Channels of Adjustment of Non-College Employment




Notes: Transition paths of aggregate non-college employment, expressed as percentage of the total change in the new BGP. Left panel compares effects of a \$15 minimum wage in our baseline model and a full utilization model calibrated to match the same targets. Middle panel compares the effects of a \$15 minimum wage in our baseline to $\widehat{N}_{Lt} = \sum_{\tau=1}^{\infty} \sum_{i \in L} (1-\delta)^{\tau-1} X_{t-\tau} \Pi^u(\varepsilon_\tau) v_{Lt-\tau}$, which holds the utilization schedule fixed at its BGP value. Right panel does the same for \$12.50 minimum wage.

a permanent idiosyncratic productivity $\varepsilon$. The *full utilization* model has neither of these features, $g = 0$ and $\varepsilon = 1$. As shown in the left panel of Figure 2, the full utilization model has 100% utilization of capital along a BGP whereas utilization in our baseline vintage model declines with age. For both models the *speed of adjustment* is defined as the change in non-college employment in the $t^{th}$ month following the introduction of the minimum wage as a percentage of the long-run change in non-college employment achieved in the new BGP.

Figure 7's left panel shows that the transition in the *full utilization* model is slow: in the first 30 months after the introduction of a \$15 minimum wage, employment of non-college workers has adjusted only 25% of the way to the new BGP level and by 60 months employment adjusts by only 47%. Here, the only operative margin of adjustment is to wait for existing capital to depreciate before replacing workers. Indeed, since workers separate at 30% per year and capital depreciates at only 15% per year, the firm actively hires 15% each type of worker that was working on the machines in the original BGP—no matter how low the skill level—to keep the old machines running at full capacity. Hence, the desire of the firm to keep all the old capital active is saving the jobs of the low-skill workers from being eliminated in the short run. At the same time, the firm is building new types of capital with a much lower fraction of low-skill types. In short, the full utilization model has only these type of *capital dynamics* to adjust the speed of transition in the short run.

The left panel also shows that our baseline model with a vintage structure generates a much faster transition in response to a large minimum wage change: in the first 30 months employment adjusts by 66% and by 60 months it adjusts by 89%. One reason for this faster transition is that, even in the initial BGP, the vintage model, has *endogenous obsolescence* of capital in addition to standard capital depreciation. To see how, compare the left panel of Figure 2 with full utilization to the right panel with vintage capital in which all the capital with productivity $A\varepsilon$ to the left of the cutoff level is never utilized again. So it is as if capital disappears from use at rate equal to the sum of the depreciation rate and the obsolescence rate whereas in the full utilization model capital disappears from use only at the rate of depreciation. So, even if the obsolescence rate stays fixed after the introduction of the minimum wage, capital turns over much faster in the vintage model than the full utilization model due to this *steady state obsolescence effect* in the original BGP.

The second, more subtle, reason why transition is faster in the model with vintage capital is that utilization rates endogenously respond to the minimum wage change. Specifically, they cause the cutoff $\underline{\varepsilon}$ in the right panel of Figure 2 to shift to the right because the binding minimum wage on some low-skilled workers makes marginal capital less attractive and, hence, more capital is idled. Since this force is bigger the larger is the minimum wage, it leads large minimum wages to have faster transitions than small minimum wages.

We can quantify the sizes of the steady state obsolescence effect and the endogenous utilization effect as follows. To do so, recall that $\Pi^u(\underline{\varepsilon}) = \int_{\underline{\varepsilon}}^{\infty} \pi(\varepsilon)d\varepsilon$ and $\underline{\varepsilon}_{t-\tau,t}$ is the cutoff $\varepsilon$ for capital made in period $t-\tau$ to be utilized in $t$, and write the *total movements* in type $i$ labor at $t$ as

$$N_{it} = \sum_{\tau=1}^{\infty} (1-\delta)^{\tau-1} X_{t-\tau} \Pi^u(\underline{\varepsilon}_{t-\tau,t}) v_{it-\tau}, \tag{33}$$

Hence, this labor is the sum of its use in current and all past vintages of capital. The type-$i$ labor allocated to capital made in $t-\tau$ is the product of the capital of that vintage remaining in $t$, $(1-\delta)^{\tau-1}X_{t-\tau}$, the share of that vintage that is utilized, $\Pi^u(\underline{\varepsilon}_{t-\tau,t})$, and the type-$i$ labor intensity of that vintage, $v_{it-\tau}$. Hence, employment changes as any of these pieces change.

Next, let the *fixed partial utilization* employment of type $i$, $\widehat{N}_{it}$, be the analog of (33) but with utilization rates held fixed at their BGP levels, so $\widehat{N}_{it} = \sum_{\tau=1}^{\infty}(1-\delta)^{\tau-1}X_{t-\tau}\Pi^u(\underline{\varepsilon}_{\tau})v_{it-\tau}$, where $\underline{\varepsilon}_{\tau}$ is the utilization threshold along the BGP for capital installed $\tau$ periods ago. Thus, $\widehat{N}_{it}$ moves only because of capital accumulation dynamics, but in a world where there is fixed partial utilization. For each labor type $i$ we can decompose the total movement in its labor as $N_{it} = \widehat{N}_{it} + [N_{it} - \widehat{N}_{it}]$ and define the associated aggregate of non-college labor by summing over the types $i$ without college by letting $N_{Lt} = \sum_{i \in L} N_{it}$ and $\hat{N}_{Lt} = \sum_{i \in L} \hat{N}_{it}$. Then we can decompose the total movements in non-college labor into the sum of those movements with (counterfactually) fixed partial utilization and the rest—which is due to endogenous utilization as

$$\underbrace{N_{Lt}}_{\text{Total Movement}} = \underbrace{\widehat{N}_{Lt}}_{\text{Fixed Partial Utilization}} + \underbrace{N_{Lt} - \widehat{N}_{Lt}}_{\text{Endogenous Utilization}} \tag{34}$$

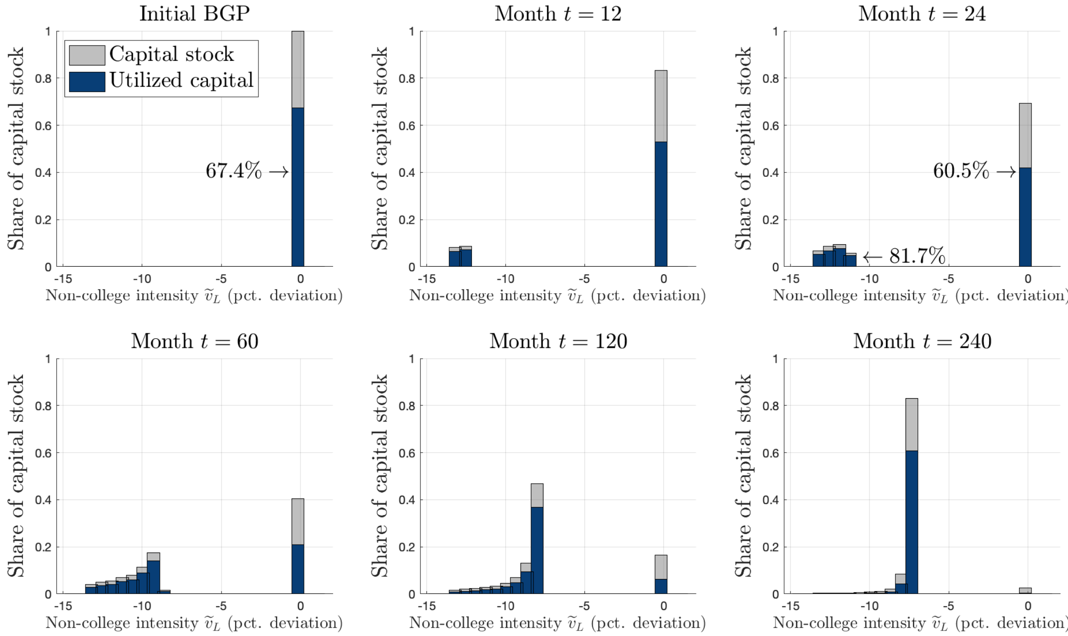
Figure 7's middle panel adds the path of non-college employment under fixed partial utilization to the left panel. Clearly, the fixed partial utilization model generates the bulk of the speeding up in the transition. For example, of the 41 percentage point increase in the speed of transition at 30 months shown in the left panel between the baseline model and the full utilization model (66% vs. 25%), the fixed partial utilization component generates about 60% of it ((50%-25%)/41%) and the endogenous utilization component generates the rest. Moreover, the difference between the fixed partial utilization counterfactual and the full utilization model, which is due to the capital utilization dynamics, converges to zero along the transition path because capital utilization eventually reverts to its initial value—see Lemma 3. Hence, the obsolescence effect on past vintages of capital present in the initial BGP of the benchmark model accounts for the bulk of the increased transition speed.

Comparing Figure 7's middle and right panels shows that the endogenous utilization component discussed above is bigger for larger minimum wage changes: after 60 months, the difference between the two transition paths is 6 percentage points with a \$12.50 minimum wage (right panel) but 11 percentage points with a \$15 minimum wage (middle panel). This result is intuitive: the larger is the minimum wage, the greater is the endogenous force to idle more of the marginal capital stock.

Figure 8 illustrates these margins of adjustment along the transition path by plotting the distribution of capital types at various points in time following the introduction of a \$15 minimum wage. Indexing a unit of capital by its detrended aggregate non-college labor intensity $\widetilde{v}_L = \sum_{i \in I_L} \widetilde{v}_i$, we calculate the total amount of capital of each type, the gray bars, and the total amount of capital of that type that is actually utilized, the blue bars. In the initial BGP, all capital has the same *detrended* labor intensity $\widetilde{v}_L^*$, and firms utilize roughly 68% of the various vintages that make up that capital. Once a large minimum wage is introduced, firms begin to invest in capital with lower detrended non-college labor intensity than in the initial BGP. Now remember that the initial BGP has the fixed partial utilization pattern present in the right-most panel of Figure 2. As Figure 7's middle panel shows, by itself this partial utilization pattern, even when held fixed at the initial BGP levels, leads to faster transitions than the full utilization model.

We now turn to how the endogenous response of utilization rates to the minimum wage speeds up the transition even more than would occur with the fixed partial utilization present in the initial BGP of the vintage capital model. To see how suppose that the capital utilization schedule was

Figure 8: Distribution of Capital Types Along the Transition, \$15 Minimum Wage



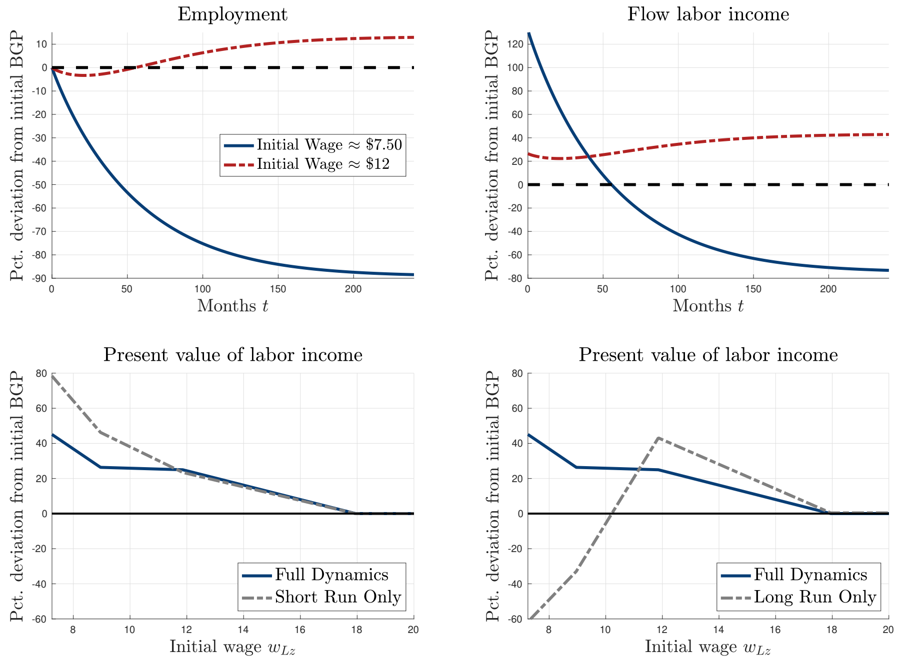

Notes: The gray portion of each histogram represents the share of capital in the corresponding range of detrended non-college labor intensities $\tilde{v}_L = \sum_{i \in L} \tilde{v}_i$. The blue portion of each histogram represents the share of capital that is utilized in production. The different panels represent the response at different time horizons after the introduction of the minimum wage.

held constant at the schedule in the initial BGP so that the blue bars were a constant share of the gray bars for all capital types. Given our depreciation rate of 15%, it takes more than ten years for this old capital to fully depreciate. The employment dynamics induced by this capital investment channel is the capital accumulation dynamics we were discussing earlier. To see the size of the endogenous utilization dynamics compare the 67.4% utilization rate in the initial BGP with utilization rates in month 24, namely 60.5% for old capital and 81.7% for new capital. These differences are the reason that the endogenous utilization margin speeds up the transition.

### 5.3 Heterogeneous Employment and Income Responses

To understand the implications of the slow dynamics of aggregate employment and labor income, Figure 9's top left panel plots the employment rates $n_{it}$ for a low productivity non-college worker earning \$7.50 and a medium productivity one earning \$12 in response to a \$15 minimum wage. On impact, the employment of both worker types remains constant. Over time, though, firms substitute away from low-productivity workers, leading to a prolonged decrease in their employment, which takes many years to complete. By contrast, firms substitute towards medium-productivity workers,

Figure 9: Individual Employment and Labor Income Dynamics to $15 Minimum Wage




Notes: The top panels show the transition dynamics for employment and labor income for two non-college worker types in response to a $15 minimum wage. The bottom panels show the change in the present value of labor income for all non-college worker types under our baseline model. *Full Dynamics* refers to the present value of income computed under the transition path in our baseline model. *Short Run Only* computes the present value assuming that labor income gains during the first two years persists over time (bottom left). *Long Run Only* computes the present value assuming that labor income gains in the new steady state persist over time (bottom right).

for whom the monopsony distortion is reduced, but again this adjustment process is slow.[15]

Figure 9's top right panel plots these workers' labor incomes. Since the employment rate of workers does not change on impact, with a $15 minimum wage the labor income of workers initially earning $12 increases by 25%. Over time, firms slowly increase the demand for these workers, so their labor income eventually increases by about 40% in the long run. Hence, the slow employment dynamics delay the long-run *benefits* of the minimum wage for these workers. By contrast, the slow employment dynamics delay the long-term *costs* of the minimum wage for workers initially earning $7.50. On impact, the $15 minimum wage doubles their labor income because it doubles their wages, and firms have not yet reduced their employment. Over time, though, firms slowly substitute away from these workers which reduces both their employment and their labor income. Eventually, the income of these workers falls significantly relative to that in the initial BGP.

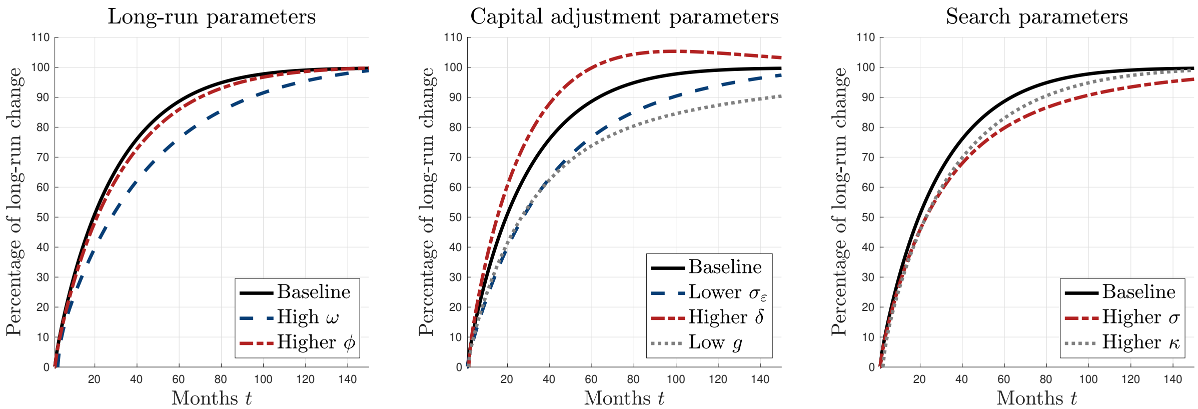
[15]In the early stages of the transition, the employment of medium-productivity workers falls slightly before rising. This decline results from firms lowering the utilization rate of their existing capital before they start purchasing new capital. Additionally, the employment declines for lower-productivity workers are faster than the employment increases for higher-productivity workers which explains why aggregate non-college employment responds nonmonotonically to a $12 minimum wage, as shown in the red dashed line of the left panel of Figure 6.

A full evaluation of the policy must then integrate both short-run and long-run effects. The blue line in the bottom panels of Figure 9 plots the percentage increase in the present value of labor income along the transition relative to its present value in the initial BGP for each non-college worker type (as measured by their initial wage). The present value of income (weakly) increases for all non-college workers, and proportionately more for the lowest-wage ones. The bottom left panel shows how this present value would change if we assumed that the average change in labor income over the first two years after the introduction of the new minimum lasted forever. We refer to this as the *Short Run Only* counterfactual. Clearly, for low-wage workers, this *naive static analysis* greatly overstates the gain from the minimum wage. For example, for workers initially earning \$7.50 an hour, it would imply an increase in the present value of labor income by 80%, which is double the true gain accounting for the transition.

The bottom right panel shows how this present value would change if we assumed that on impact the economy immediately reached the new BGP. We refer to this as the *Long Run Only* counterfactual. For workers initially earning \$7.50 an hour, this *naive long-run analysis* would predict that they experience a decrease in the present value of earnings of over 60%, whereas the true gains calculated accounting for the transition are positive and over 40%. The extent to which the *Short Run Only* and *Long Run Only* counterfactuals differ from a model accounting for the actual transition path is much larger for lower productivity workers. The reason for this is that workers experience both larger short run gains and larger long run losses in response to a \$15 minimum wage.

These calculations highlight the limitations of two common approaches in the literature. The first one is to extrapolate empirical measures of the impact of a change in minimum wage estimated over the first couple of years following this change to much longer time horizons, presuming that what happens during such a short-run period is informative about what happens from then on. The second one is to use a long-run framework that ignores an economy's transition to assess the dynamic impact of a minimum wage change. For example, imagine that we eliminated putty-clay and search frictions and all other sources of dynamics in our model so that the economy immediately reached the new BGP. The first approach would give the same predictions as the naive static analysis as the *Short Run Only* counterfactual while the second approach would give the same predictions as the *Long Run Only* counterfactual. Neither of these common approaches would be close to measuring the correct present value of the change as calculated along the transition for lower wage workers.

FIGURE 10: Sensitivity Analysis for \$15 Minimum Wage




Notes: Transition paths of aggregate non-college employment, expressed as percentage of the total change in the new BGP. *Baseline* corresponds to the model shown in Figure 6. *Higher* $\omega$ corresponds to a degree of monopsony power of $\omega^{-1} = 1/6$ that produces an 85% markdown. *Higher* $\phi$ corresponds to a long-run elasticity of substitution within education groups of $\phi = 4.5$. *Lower* $\sigma_\varepsilon$ corresponds to a standard deviation of idiosyncratic capital productivity of $\sigma_\varepsilon = 0.01$ that generates a steady-state capacity utilization rate of 97%. *Higher* $\delta$ sets the depreciation rate to $\delta = 20\%$ annually. *Lower* $g$ corresponds to a trend growth rate $g$ of 0.01% annually. *Higher* $\sigma$ sets the job-destruction rate to $\sigma = 3.5\%$ monthly. *Higher* $\kappa$ increases the baseline vacancy-posting cost $\kappa_0$ by 2.5 times, which roughly doubles the average hiring costs $\kappa_i/\lambda_f(\theta_i)$ to 125% of average monthly wage.

## 5.4 Sensitivity Analysis and Relationship to Empirical Literature

We now examine the sensitivity of our results to alternative parameterizations and discuss their relationship to the empirical literature.

**Sensitivity Analysis.** Figure 10 illustrates how various features of our model impact the speed at which aggregate non-college employment responds to the minimum wage. The left panel focuses on the two parameters that are key in determining the long-run effect of the minimum wage: the degree of firm monopsony power, governed by $\omega$, and the long-run elasticity of labor-labor substitution, governed by $\phi$. When the degree of monopsony power is reduced to $\omega = 6$, which implies an average markdown of 0.85, the minimum wage leads to a larger decline in employment in the long run and a slower transition to it. Similarly, increasing the long-run elasticity of labor-labor substitution to $\phi = 4.5$ leads to a larger decline in employment and a slightly slower transition.[16]

The middle panel of Figure 10 illustrates the sensitivity of the employment response with respect to the parameters that govern the capital adjustment process. For the capital accumulation dynamics, consider an increase in the capital depreciation rate to $\delta = 20\%$ per year. Since the firm's installed capital depreciates more quickly, employment also adjusts more quickly to the minimum wage as firms more quickly purchase new capital with different labor intensities. For the capital

[16]The sensitivity analysis for the *level* of non-college employment in the new BGP can be found in the Online Appendix.

utilization dynamics, consider a decrease in the dispersion of idiosyncratic capital productivity to $\sigma_\varepsilon = 0.01$, which implies that the average utilization rate of capital is 97% as in the middle panel of Figure 2. As very few units of capital are near the utilization threshold, changes in utilization are small in response to changes in the minimum wage, which slows down the transition of employment. Similarly, if we instead set $g = 0.01\%$ annually, then fewer units of capital are near the utilization threshold and the transition is again slower.

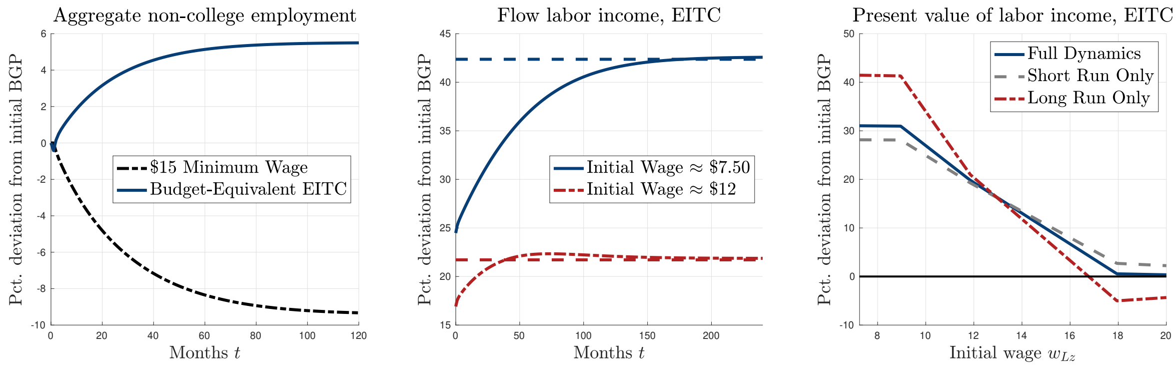
The right panel of Figure 10 shows the sensitivity of our main results to two parameters governing the degree of search frictions: the vacancy-posting cost $\kappa_i$ and the job destruction rate $\sigma$. We increase the baseline vacancy-posting cost $\kappa_0$ to approximately double the average hiring costs $\kappa_i/\lambda_f(\theta_i)$ and increase the job separation rate to $\sigma = 3.5\%$ monthly. As the figure shows, these parameters have a relatively minor effect on the speed of transition of employment, because most of the slow adjustment implied by our model is due to our putty-clay technology not search frictions.

**Comparison with the Empirical Literature.** Our model's slow employment dynamics is consistent with the large empirical literature documenting the effects of the minimum wage.[17] Neumark and Shirley (2022) recently reviewed this vast literature by conducting a meta-analysis of 109 published studies based on cross-state variation in the minimum wage in the United States and calculating the implied short-run elasticity of the employment response. All of the papers reviewed focus on the employment effects *i*) stemming from small minimum wage changes (increases of \$3 or less); *ii*) analyzed over short time horizons (12 to 24 months after the policy takes place); and *iii)* for initially lower-earning workers, such as teenagers and young adults. Neumark and Shirley (2022) find that roughly 80% of the studies imply zero to small short-run employment declines over the first two years following a minimum wage increase. Our model's response to small- and medium-sized minimum wage increases in Figure 6 are consistent with these findings.

Directly related to our focus on the transition dynamics in response to policies, Clemens and Strain (2021) provides evidence on how employment responds to minimum wage increases of different sizes at different horizons. These authors estimate the employment effects of small changes (less than \$2.50) and larger changes (more than \$2.50) in the minimum wage in both the short run (1 to 3 years) and the medium run (4 to 6 years). They find that small minimum wage changes have insignificant effects on employment in the short and medium run, consistent with our small minimum wage experiment in the left panel of Figure 6. However, Clemens and Strain (2021) also finds

[17] We focus our discussion on work examining state or national changes in the minimum wage. Studies of changes in the minimum wage at the city level tend to find larger short-run declines in employment as households and firms can easily substitute their consumption and production across neighboring cities where the minimum wage did not change. Such margins of adjustment are outside the scope of our model.

Figure 11: Employment and Labor Income Dynamics in Response to EITC




Notes: Transition paths following the introduction of the EITC. Left panel plots the path of aggregate non-college employment, expressed relative to initial BGP. Middle panel plots detrended flow labor income for two non-college worker types. Right panel plots the present value of labor income for select non-college worker types. *Full Dynamics*, *Short Run Only*, and *Long Run Only* are defined similarly as in Figure 9.

that larger changes in the minimum wage have statistically significant negative employment effects after 4 to 6 years. These results are in line with the implications of our model for the medium- and large-sized minimum wage changes shown in Figure 6. We take it as a strength of our framework that we can match these dynamic employment responses to increases in the minimum wage.[18]

Finally, a separate literature also provides evidence that firms slowly adjust their input mix in response to minimum wage increases, in line with the predictions of our model. For example, in response to a large and persistent minimum wage increase in Hungary, Lindner and Harasztosi (2019) documents that firms responded by substituting away from labor towards capital. Relatedly, Clemens, Kahn and Meer (2021) provide evidence that U.S. firms substitute away from low-productivity workers towards higher-productivity ones in response to minimum wage increases.

## 5.5 Dynamics Effects of the EITC

We now show that the mechanisms governing the dynamic effects of the EITC are similar to those of the minimum wage. As with the minimum wage, we assume that the economy starts along a BGP without any policies and that the EITC is introduced starting in period 0, after which agents have perfect foresight of the transition path to the new BGP. We study an EITC that is budget-equivalent to a $15 minimum wage in the long run—namely, financed though a linear tax on profits such that the revenue from the tax equals the loss in profits associated with a $15 minimum.

[18] Similarly, Cengiz et al. (2019) estimate both the short- and the long-run effects of small minimum wage increases—averaging across all the increases they study, the minimum wage was raised by about 10% or about $0.75 in current dollars. These authors find that employment effects are small and positive for lower wage workers in the few years after a small minimum wage increase, and persist over a seven-year horizon. Again, our model also predicts these findings for small minimum wage changes on low wage workers. However, as our model shows, one should not extrapolate these findings to larger minimum wage changes.

The left panel of Figure 11 compares the transition path of aggregate non-college employment following the introduction of the EITC with that following a $15 minimum wage. Under a minimum wage, firms pay the marginal cost of the policy and hence substitute away from the affected workers. Under the EITC, the government pays the marginal cost of it, thereby subsidizing the wages of affected workers, stimulating their labor supply, and lowering the pre-transfer wage that firms need to pay to hire them. These lower wages and reduced monopsony distortion induce firms to substitute toward the subsidized workers. Thus, a budget-equivalent EITC expansion leads to better medium- and long-run employment outcomes for non-college workers than does a $15 minimum wage.

The middle panel of Figure 11 shows the implications of these employment dynamics for the labor income of a low-productivity worker earning $7.50 and a medium-productivity worker earning $12 before the EITC. On impact, the EITC raises the income of both workers. Over time, firms substitute toward them and their income continues to grow. Hence, the gradual employment dynamics slows down the long-run benefits of the policy on workers' labor income. Proportionally, income gains are larger for the low-productivity worker. The right panel of Figure 11 illustrates how the present value of labor income increases for all workers affected by the EITC, especially for the lower-productivity ones primarily targeted by the policy. For workers initially earning less than $13, comparing the new BGP to the old one overstates the true gain in labor income, because such a comparison ignores the slow dynamics of employment. For higher-productivity workers, a BGP comparison implies a decline in income altogether, because these workers in the phase-out region of the EITC experience an increase in their monopsony distortion in the new BGP, which lowers their employment in the long run. Hence, for this policy as well, taking into account the entire transition path of the economy is critical to accurately gauging the true effects of an EITC policy on the present value of income.

## 6 Conclusion

When an economy's response to a sizable policy change is slow, any comprehensive assessment of the policy's benefits or costs must take the full dynamics of adjustment into account. In this paper, we provide a tractable dynamic general equilibrium model that captures the key forces required for such an assessment. We show that using our novel framework to evaluate policies leads to substantially different conclusions about their aggregate and distributional effects than standard static-only or long-run-only approaches would imply.

## References


Aaronson, D., and E. French (2007): "Product Market Evidence on the Employment Effects of the Minimum Wage," *Journal of Labor Economics*, 25(1), 167–200.

Aaronson, D., E. French, I. Sorkin, and T. To (2018): "Industry Dynamics and the Minimum Wage: A Putty-Clay Approach," *International Economic Review*, 59(1), 51–84.

Adams, C., J. Meer, and C. Sloan (2022): "The Minimum Wage and Search Effort," *Economics Letters*, 212(110288), 1–8.

Ahn, T., P. Arcidiacono, and W. Wessels (2011): "The Distributional Impacts of Minimum Wage Increases When Both Labor Supply and Labor Demand Are Endogenous," *Journal of Business and Economic Statistics*, 29(1), 12–23.

Andolfatto, D. (1996): "Business cycles and labor-market search," *The American Economic Review*, 112–132.

Atkeson, A., and P. J. Kehoe (1999): "Models of energy use: Putty-putty versus putty-clay," *American Economic Review*, 89(4), 1028–1043.

Berger, D., K. Herkenhoff, and S. Mongey (2022): "Labor Market Power," *American Economic Review*, 112(4), 1147–93.

——— (2025): "Minimum Wages, Efficiency, and Welfare," *Econometrica*, 93(1), 265–301.

Berger, D., K. Herkenhoff, S. Mongey, and N. Mousavi (2024): "Monopsony amplifies distortions from progressive taxes," in *AEA Papers and Proceedings*, vol. 114, 555–560. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Borjas, G. J., and L. F. Katz (2007): "The evolution of the Mexican-born workforce in the United States," in *Mexican immigration to the United States*, 13–56. University of Chicago Press.

Burdett, K., and K. L. Judd (1983): "Equilibrium Price Dispersion," *Econometrica*, 51(4), 955–969.

Card, D., and T. Lemieux (2001): "Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis*," *The Quarterly Journal of Economics*, 116(2), 705–746.

Cengiz, D., A. Dube, A. Lindner, and B. Zipperer (2019): "The Effect of Minimum Wages on Low-Wage Jobs," *The Quarterly Journal of Economics*, 134(3), 1405–1454.

Clemens, J., L. B. Kahn, and J. Meer (2021): "Dropouts Need Not Apply? The Minimum Wage and Skill Upgrading," *Journal of Labor Economics*, 39(S1), S107–S149.

Clemens, J., and M. R. Strain (2021): "The Heterogeneous Effects of Large and Small Minimum

Wage Changes: Evidence over the Short and Medium Run Using a Pre-Analysis Plan," Discussion paper, National Bureau of Economic Research.

Deb, S., J. Eeckhout, A. Patel, and L. Warren (2024): "Walras–Bowley Lecture: Market power and wage inequality," *Econometrica*, 92(3), 603–636.

Drechsel-Grau, M. (2022): "Employment and Reallocation Effects of Higher Minimum Wages," *Unpublished Working Paper*.

Eckstein, Z., and K. I. Wolpin (1990): "Estimating a Market Equilibrium Search Model from Panel Data on Individuals," *Econometrica*, 58(4), 783–808.

Engbom, N., and C. Moser (2022): "Earnings inequality and the minimum wage: Evidence from Brazil," *American Economic Review*, 112(12), 3803–3847.

Flinn, C. J. (2006): "Minimum Wage Effects on Labor Market Outcomes under Search, Matching, and Endogenous Contact Rates," *Econometrica*, 74(4), 1013–1062.

Gilchrist, S., and J. C. Williams (2000): "Putty-clay and investment: a business cycle analysis," *Journal of political Economy*, 108(5), 928–960.

Greenwood, J., Z. Hercowitz, and G. W. Huffman (1988): "Investment, Capacity Utilization, and the Real Business Cycle," *American Economic Review*, 78(3), 402–417.

Karabarbounis, L., and B. Neiman (2014): "The global decline of the labor share," *The Quarterly Journal of Economics*, 129(1), 61–103.

Katz, L. F., and K. M. Murphy (1992): "Changes in relative wages, 1963–1987: supply and demand factors," *The Quarterly Journal of Economics*, 107(1), 35–78.

Kehoe, P. J., P. Lopez, V. Midrigan, and E. Pastorino (2023): "Asset prices and unemployment fluctuations: A resolution of the unemployment volatility puzzle," *The Review of Economic Studies*, 90(3), 1304–1357.

Krusell, P., T. Mukoyama, R. Rogerson, and A. Şahin (2017): "Gross worker flows over the business cycle," *American Economic Review*, 107(11), 3447–3476.

Lamadon, T., M. Mogstad, and B. Setzler (2022): "Imperfect competition, compensating differentials, and rent sharing in the US labor market," *American Economic Review*, 112(1), 169–212.

Lindner, A., and P. Harasztosi (2019): "Who Pays for the Minimum Wage?," *American Economic Review*, 109(8), 2693–2727.

Ljungqvist, L., and T. J. Sargent (2017): "The fundamental surplus," *American Economic Review*, 107(9), 2630–2665.

MacCurdy, T. (2015): "How Effective Is the Minimum Wage at Supporting the Poor?," *Journal of Political Economy*, 123(2), 497 – 545.

Manning, A. (2011): "Imperfect Competition in the Labor Market," In *Handbook of Labor Economics*, Volume 4b, Orley Ashenfelter and David Card, Eds., Amsterdam, North–Holland, 973–1041.

Marcet, A., and R. Marimon (2019): "Recursive contracts," *Econometrica*, 87(5), 1589–1631.

Merz, M. (1995): "Search in the labor market and the real business cycle," *Journal of Monetary Economics*, 36(2), 269–300.

Mousavi, N. S. (2022): "Optimal Labor Income Taxes, Incomplete Markets, and Labor Market Power," Ph.D. thesis, The University of Chicago.

Neumark, D., and P. Shirley (2022): "Myth or Measurement: What Does the New Minimum Wage Research Say about Minimum Wages and Job Loss in the United States?," Discussion paper, National Bureau of Economic Research.

Oberfield, E., and D. Raval (2021): "Micro Data and Macro Technology," *Econometrica*, 89(2), 703–732.

Robinson, J. (1933): "The Economics of Imperfect Competition," St. Martin's Press.

Seegmiller, B. (2021): "Valuing Labor Market Power: The Role of Productivity Advantages," Discussion paper, MIT Sloan Working Paper.

Shimer, R. (2005): "The cyclical behavior of equilibrium unemployment and vacancies," *American Economic Review*, 95(1), 25–49.

Sorkin, I. (2015): "Are There Long-Run Effects of the Minimum Wage?," *Review of Economic Dynamics*, 18(2), 306–333.

Yeh, C., C. Macaluso, and B. Hershbein (2022): "Monopsony in the US labor market," *American Economic Review*, 112(7), 2099–2138.

# Online Appendix

## A Model Without Labor Market Policies

We start by studying the model without any labor market policies. This analysis underlies the results in Sections 1 and 2 of the main text. In Appendix B, we introduce labor market policies and use this BGP as the initial condition.

### A.1 Households

From the main text, the household's utility maximization problem is

$$\max_{c_{it}, s_{ijt}, n_{ijt+1}} \sum_{t=0}^{\infty} \beta^t U_t(c_{it}, n_{it}, s_{it})$$

$$\text{s.t. } n_{ijt+1} = (1-\sigma)n_{ijt} + \lambda_w(\theta_{ijt})s_{ijt} \quad (\times \beta^{t+1}\widehat{V}_{ijt+1})$$

$$\sum_{t=0}^{\infty} Q_{0,t}c_{it} = \zeta_i \mathbb{P} + \mathbb{I}_i + \sum_{t=1}^{\infty} Q_{0,t} \sum_j \lambda_w(\theta_{ijt-1})s_{ijt-1}W_{ijt}. \quad (\times \Gamma) \tag{A1}$$

Here, the variables in parentheses denote the (often rescaled) Lagrange multiplier associated with the constraint. The first-order condition for consumption $c_{it}$ is $\beta^t U_{cit} = \Gamma Q_{0,t}$. Taking ratios of this equation across adjacent time periods gives $Q_{t,t+1} = \beta \frac{U_{cit+1}}{U_{cit}}$. The first-order condition for employment $n_{ijt+1}$ is

$$\beta^{t+1} U_{nit+1} \left( \frac{n_{ijt+1}}{n_{it+1}} \right)^{\frac{1}{\omega}} - \beta^{t+1}\widehat{V}_{ijt+1} + \beta^{t+2}(1-\sigma)\widehat{V}_{ijt+2} = 0,$$

which implies

$$\widehat{V}_{ijt+1} = U_{nit+1} \left( \frac{n_{ijt+1}}{n_{it+1}} \right)^{\frac{1}{\omega}} + \beta(1-\sigma)\widehat{V}_{ijt+2}.$$

which uses the fact that $\frac{\partial n_{it+1}}{\partial n_{ijt+1}} = \left( \frac{n_{ijt+1}}{n_{it+1}} \right)^{\frac{1}{\omega}}$. Note that $\widehat{V}_{ijt+1}$ is in units of utility at $t+1$. Going forward, it will be useful to put this object in consumption units by dividing by the marginal utility of consumption in period $t+1$:

$$\begin{aligned} V_{ijt+1} \equiv \frac{\widehat{V}_{ijt+1}}{U_{cit+1}} &= \frac{U_{nit+1}}{U_{cit+1}} \left( \frac{n_{ijt+1}}{n_{it+1}} \right)^{\frac{1}{\omega}} + \beta \frac{U_{cit+2}}{U_{cit+2}}(1-\sigma)\frac{\widehat{V}_{ijt+2}}{U_{cit+1}} \\ &= \frac{U_{nit+1}}{U_{cit+1}} \left( \frac{n_{ijt+1}}{n_{it+1}} \right)^{\frac{1}{\omega}} + Q_{t+1,t+2}(1-\sigma)V_{ijt+2}, \end{aligned} \tag{A2}$$

where the second line uses the fact that $Q_{t+1,t+2} = \beta \frac{U_{cit+2}}{U_{cit+1}}$. This equation defines $V_{ijt+1}$ as the present value of marginal disutilities of work for workers that are hired in period $t$ and begin working in period $t+1$, in terms of their consumption units in period $t+1$.

The first-order condition for search effort $s_{it}$ is

$$\beta^t U_{sit} + \beta^{t+1}\widehat{V}_{ijt+1}\lambda_w(\theta_{ijt}) + \Gamma Q_{0,t+1}\lambda_w(\theta_{ijt})W_{ijt+1} = 0$$
$$\implies -U_{sit} = \lambda_w(\theta_{ijt})\beta U_{cit+1}\left(\frac{\widehat{V}_{ijt+1}}{U_{cit+1}} + W_{ijt+1}\right)$$
$$\implies -\frac{U_{sit}}{U_{cit}} = \lambda_w(\theta_{ijt})Q_{t,t+1}\left[V_{it+1} + W_{ijt+1}\right], \tag{A3}$$

where the second line uses the fact that $\Gamma Q_{0,t+1} = U_{cit+1}$ and the third line uses the fact that $Q_{t,t+1} = \beta\frac{U_{cit+1}}{U_{cit}}$. Recall from the main text that we will write the participation constraint as

$$\lambda_w(\theta_{ijt})Q_{t,t+1}(W_{ijt+1} + V_{ijt+1}) \geq \mathcal{W}_{it} \equiv \lambda_w(\theta_{it})Q_{t,t+1}(W_{it+1} + V_{it+1}). \tag{A4}$$

## A.2 Firms

We now turn to the firm's profit maximization problem, which is the main challenge of solving the model. We abstract from initial conditions because we use these results to derive the limiting BGP. We also ignore the non-negativity constraint on vacancy posting $a_{ijt} \geq 0$ because that constraint is not binding along the BGP. We begin by showing how to group collect the Lagrange multipliers on the participation constraints in terms of the auxiliary variable $M_{ijt}$ defined in the main text. This allows us to specify the full profit maximization problem of the firm. We then derive the first-order conditions of the profit maximization problem. Finally, we summarize the resulting conditions which characterize the solution to the firm's problem.

**Grouping Multipliers on the Participation Constraint.** We collect across time the corresponding terms of the participation constraints of each period, in a way that has become standard in the dynamic contracting literature (Marcet and Marimon 2019), so as to isolate the impact of additional hires of a type-$i$ family by firm $j$ in $t$ on the disutility of work of all members of the family hired by the firm in future periods. To do this, we attach the (scaled) Lagrange multiplier $Q_{0,t+1}\mu_i\gamma_{ijt+1}$ to the time-$t$ participation constraint (5) from the main text. It is instructive to write out how the first few period's participation constraints enter firm $j$'s expected profit maximization problem:

$$\begin{aligned}
&Q_{0,1}\gamma_{ij1}\Big[\frac{U_{ni1}}{U_{ci1}}\left(\frac{n_{ij1}}{n_{i1}}\right)^{\frac{1}{\omega}} + Q_{1,2}(1-\sigma)\frac{U_{ni2}}{U_{ci2}}\left(\frac{n_{ij2}}{n_{i2}}\right)^{\frac{1}{\omega}} + Q_{1,3}(1-\sigma)^2\frac{U_{ni3}}{U_{ci3}}\left(\frac{n_{ij3}}{n_{i3}}\right)^{\frac{1}{\omega}} + \ldots + W_{ij1} - \frac{\mathcal{W}_{i0}}{Q_{0,1}\lambda_w(\theta_{ij0})}\Big]\\
&Q_{0,2}\gamma_{ij2}\Big[\frac{U_{ni2}}{U_{ci2}}\left(\frac{n_{ij2}}{n_{i2}}\right)^{\frac{1}{\omega}} + Q_{2,3}(1-\sigma)\frac{U_{ni3}}{U_{ci3}}\left(\frac{n_{ij3}}{n_{i3}}\right)^{\frac{1}{\omega}} + \ldots + W_{ij2} - \frac{\mathcal{W}_{i1}}{Q_{1,2}\lambda_w(\theta_{ij1})}\Big]\\
&Q_{0,3}\gamma_{ij3}\Big[\frac{U_{ni3}}{U_{ci3}}\left(\frac{n_{ij3}}{n_{i3}}\right)^{\frac{1}{\omega}} + \ldots + W_{ij3} - \frac{\mathcal{W}_{i2}}{Q_{2,3}\lambda_w(\theta_{ij2})}\Big].
\end{aligned}$$

By collecting the multipliers associated with the terms $\frac{U_{nit}}{U_{cit}}\left(\frac{n_{ijt}}{n_{it}}\right)^{\frac{1}{\omega}}$ for some $t$ and noting that $Q_{0,\tau}Q_{\tau,t} = Q_{0,t}$ for $\tau < t$, it is easy to see that all such terms are summarized by the auxiliary variable $M_{ijt+1} = (1-\sigma)M_{ijt} + \gamma_{ijt+1}$ as in Marcet and Marimon (2019). Hence, the contributions of the participation constraint to the Lagrangian can be reduced to

$$\sum_{t=0}^{\infty} Q_{0,t+1}\mu_i M_{ijt+1}\frac{U_{nit+1}}{U_{cit+1}}\left(\frac{n_{ijt+1}}{n_{it+1}}\right)^{\frac{1}{\omega}} + \sum_{t=0}^{\infty} Q_{0,t+1}\mu_i\gamma_{ijt+1}\left[W_{ijt+1} - \frac{\mathcal{W}_{it}}{Q_{t,t+1}\lambda_w(\theta_{ijt})}\right].$$

**Profit-Maximization Problem.** Using these results, we can write the firm's problem as choosing utilization $u_{jt}(v,\varepsilon,A_{t-\tau})$, the labor allocation $N_{ijt}(v,A_{t-\tau},\varepsilon)$, total employment $N_{ijt}$, vacancy posting $a_{ijt}$, market tightness $\theta_{ijt}$, present value of wage offers $W_{ijt+1}$, investment $X_{jt}(v)$, and capital $K_{jt+\tau+1}(v,A_t)$, in order to maximize the expected present value of profits:

$$\sum_t Q_{0,t}\left(\sum_\tau \int_{v,\varepsilon} u_{jt}(v,A_{t-\tau},\varepsilon)A_{t-\tau}\varepsilon f(v)K_{jt}(v,A_{t-\tau})\pi(\varepsilon)d\varepsilon dv - \sum_i \mu_i(\lambda_f(\theta_{ijt-1})a_{ijt-1}W_{ijt} + \kappa_{it}a_{ijt})\right.$$
$$\left. - \int X_{jt}(v)dv\right) + \sum_{t=0}^{\infty} Q_{0,t+1}\mu_i M_{ijt+1}\frac{U_{nit+1}}{U_{cit+1}}\left(\frac{n_{ijt+1}}{n_{it+1}}\right)^{\frac{1}{\omega}} + \sum_{t=0}^{\infty} Q_{0,t+1}\mu_i\gamma_{ijt+1}\left[W_{ijt+1} - \frac{\mathcal{W}_{it}}{Q_{t,t+1}\lambda_w(\theta_{ijt})}\right]$$

such that $u_{jt}(v,A_{t-\tau},\varepsilon) \geq 0 \quad (\times Q_{0,t}\lambda^L_{jt}(v,A_{t-\tau},\varepsilon))$

$u_{jt}(v,A_{t-\tau},\varepsilon) \leq 1 \quad (\times Q_{0,t}\lambda^U_{jt}(v,\varepsilon,A_{t-\tau}))$

$u_{jt}(v,A_{t-\tau},\varepsilon)v_i K_{jt}(v,A_{t-\tau})\pi(\varepsilon) \leq N_{ijt}(v,A_{t-\tau},\varepsilon)$ for all $i \quad (\times Q_{0,t}\lambda_{ijt}(v,A_{t-\tau},\varepsilon))$

$\sum_\tau \int_{v,\varepsilon} N_{ijt}(v,A_{t-\tau},\varepsilon)d\varepsilon dv \leq \mu_i n_{ijt}$ for all $i \quad (\times Q_{0,t}\chi_{ijt})$

$\mu_i n_{ijt+1} \leq (1-\sigma)\mu_i n_{ijt} + \lambda_f(\theta_{ijt})\mu_i a_{ijt}$ for all $i \quad (\times Q_{0,t+1}\nu_{ijt+1})$

$K_{jt+\tau+1}(v,A_t) = (1-\delta)^\tau X_{jt}(v) \quad (\times Q_{0,t+\tau+1}q_{jt,t+\tau+1}(v))$

$X_{jt}(v) \geq 0 \quad (\times Q_{0,t}\mu_{jt}(v)),$

with the side conditions that $M_{ijt+1} = (1-\sigma)M_{ijt} + \gamma_{ijt+1}$ and $\frac{\mathcal{W}_{it}}{Q_{t,t+1}\lambda_w(\theta_{it})} = W_{it+1} + V_{it+1}$. As before, variables in parenthesis denote scaled Lagrange multipliers on the associated constraint. In this problem, we have explicitly written the measure of workers as $N_{ijt} = \mu_i n_{ijt}$, where $n_{ijt}$ is the share of family $i$ working at firm $j$. We make this substitution because the participation constraint naturally depends on per-capital $n_{ijt}$ rather than the total measure $N_{ijt}$.

We now proceed to take the first-order conditions of this problem. We group these conditions into three blocks: the utilization block, the hiring block, and the investment block.

**Utilization Block.** The first-order condition for labor assignment $N_{ijt}(v, A_{t-\tau}, \varepsilon)$ is simply $\lambda_{ijt}(v, A_{t-\tau}, \varepsilon) = \chi_{ijt}$. The first-order condition for utilization $u_{jt}(v, A_{t-\tau}, \varepsilon)$ is is given by

$$A_{t-\tau}\varepsilon f(v)K_{jt}(v, A_{t-\tau})\pi(\varepsilon) - \sum_i \lambda_{ijt}(v, A_{t-\tau}, \varepsilon)v_i K_{jt}(v, A_{t-\tau})\pi(\varepsilon) = \lambda^U_{ijt}(v, A_{t-\tau}, \varepsilon) - \lambda^L_{ijt}(v, A_{t-\tau}, \varepsilon).$$

Substituting the first-order condition for labor assignment from above, namely $\lambda_{ijt}(v, A_{t-\tau}, \varepsilon) = \chi_{ijt}$, we get

$$A_{t-\tau}\varepsilon f(v) - \sum_i \chi_{ijt}v_i = \frac{\lambda^U_{ijt}(v, A_{t-\tau}, \varepsilon) - \lambda^L_{ijt}(v, A_{t-\tau}, \varepsilon)}{K_{jt}(v, A_{t-\tau})\pi(\varepsilon)}. \tag{A5}$$

If $A_{t-\tau}\varepsilon f(v) - \sum_i \chi_{ijt}v_i > 0$ or, equivalently, $\varepsilon > \frac{\sum_i \chi_{ijt}v_i}{A_{t-\tau}f(v)} \equiv \underline{\varepsilon}(v, A_{t-\tau}; \chi_{jt})$ for $\chi_{jt} = (\chi_{1jt}, \ldots, \chi_{Ijt})$, then (A5) implies that $\lambda^U_{ijt}(v, A_{t-\tau}, \varepsilon) - \lambda^L_{ijt}(v, A_{t-\tau}, \varepsilon) > 0$ and so $u_{jt}(v, A_{t-\tau}, \varepsilon) = 1$ by complementary slackness. If $A_{t-\tau}\varepsilon f(v) - \sum_i \chi_{ijt}v_i < 0$ or, equivalently, $\varepsilon < \frac{\sum_i \chi_{ijt}v_i}{A_{t-\tau}f(v)} = \underline{\varepsilon}(v, A_{t-\tau}; \chi_{jt})$, then $\lambda^U_{ijt}(v, A_{t-\tau}, \varepsilon) - \lambda^L_{ijt}(v, A_{t-\tau}, \varepsilon) < 0$ by (A5), which implies that $u_{jt}(v, A_{t-\tau}, \varepsilon) = 0$ by complementary slackness. So, the utilization decision has the form: fully utilize if $\varepsilon > \underline{\varepsilon}(v, A_{t-\tau}; \chi_{jt})$ and do not utilize at all if $\varepsilon < \underline{\varepsilon}(v, A_{t-\tau}; \chi_{jt})$.[19]

Note that the solution to the static utilization problem (12) from the main text coincides with this solution from the dynamic problem if we set the static multipliers $\widehat{\chi}_{ijt} = \chi_{ijt}$.

**Hiring Block.** The first-order condition for employment $n_{ijt+1}$ is

$$\nu_{ijt+1} = \chi_{ijt+1} + M_{ijt+1}\frac{U_{nit+1}}{U_{cit+1}}\frac{1}{\omega}\left(\frac{n_{ijt+1}}{n_{it+1}}\right)^{\frac{1}{\omega}-1}\frac{1}{n_{it+1}} + Q_{t+1,t+2}(1-\sigma)\nu_{ijt+2}, \tag{A6}$$

which uses the fact that $\frac{\partial}{\partial n_{ijt+1}}\left(\frac{n_{ijt+1}}{n_{it+1}}\right)^{\frac{1}{\omega}} = \frac{1}{\omega}\left(\frac{n_{ijt+1}}{n_{it+1}}\right)^{\frac{1}{\omega}-1}\frac{1}{n_{it+1}}$. This equation identifies the multiplier $\nu_{ijt+1}$ as the present value of a marginal worker to the firm, taking into account both their marginal product $\chi_{ijt+1}$ and the monopsony distortion $M_{ijt+1}\frac{u_{nit+1}}{u_{cit+1}}\frac{1}{\omega}\left(\frac{n_{ijt+1}}{n_{it+1}}\right)^{\frac{1}{\omega}-1}\frac{1}{n_{it+1}}$.

The first-order condition for vacancy posting $a_{ijt}$ is

$$\begin{aligned} &-Q_{0,t}\kappa_{it} - Q_{0,t+1}\lambda_f(\theta_{ijt})W_{ijt+1} + Q_{0,t+1}\lambda_f(\theta_{ijt})\nu_{ijt+1} = 0 \\ \Longrightarrow\ &\frac{\kappa_i}{\lambda_f(\theta_{ijt})} = Q_{t,t+1}\left(\nu_{ijt+1} - W_{ijt+1}\right). \end{aligned} \tag{A7}$$

[19] In the knife-edge case where $A_{t-\tau}\varepsilon f(v) - \sum_i \chi_{ijt}v_i = 0$, the firm is indifferent over any $u_{jt}(v, A_{t-\tau}, \varepsilon) \in [0, 1]$.

The first-order condition for market tightness $\theta_{ijt}$ is

$$
\begin{aligned}
&- Q_{0,t+1}\lambda_f'(\theta_{ijt})a_{ijt}W_{ijt+1} + Q_{0,t+1}\nu_{ijt+1}\lambda_f'(\theta_{ijt})a_{ijt} + Q_{0,t+1}\gamma_{ijt+1}\frac{\mathcal{W}_{it}}{Q_{t,t+1}\lambda_w(\theta_{ijt})^2}\lambda_w'(\theta_{ijt}) = 0 \\
&\implies W_{ijt+1} = \nu_{ijt+1} + \frac{\gamma_{ijt+1}}{a_{ijt}}\frac{\mathcal{W}_{it}}{Q_{t,t+1}\lambda_w(\theta_{ijt})^2}\frac{\lambda_w'(\theta_{ijt})}{\lambda_f'(\theta_{ijt})}. \qquad \text{(A8)}
\end{aligned}
$$

The first-order condition for wages $W_{ijt+1}$ is

$$
\gamma_{ijt+1} = \lambda_f(\theta_{ijt})a_{ijt}, \qquad \text{(A9)}
$$

Plugging in this expression for $\gamma_{ijt+1}$ into the first-order condition for market tightness (A8) gives

$$
\begin{aligned}
W_{ijt+1} &= \nu_{ijt+1} + \frac{\lambda_f(\theta_{ijt})}{\lambda_w(\theta_{ijt})}\frac{\mathcal{W}_{it}}{Q_{t,t+1}\lambda_w(\theta_{ijt})}\frac{\lambda_w'(\theta_{ijt})}{\lambda_f'(\theta_{ijt})} \\
&= \nu_{ijt+1} + \frac{\lambda_f(\theta_{ijt})}{\lambda_w(\theta_{ijt})}\left(W_{ijt+1} + V_{ijt+1}\right)\frac{\lambda_w'(\theta_{ijt})}{\lambda_f'(\theta_{ijt})} \\
&= \nu_{ijt+1} - \frac{1-\eta}{\eta}(W_{ijt+1} + V_{ijt+1}),
\end{aligned}
$$

where in the second line we used $\frac{\mathcal{W}_{it}}{Q_{t,t+1}\lambda_w(\theta_{ijt})} = W_{ijt+1} + V_{ijt+1}$ and in the third line we used $\frac{\lambda_f(\theta_{ijt})}{\lambda_w(\theta_{ijt})}\frac{\lambda_w'(\theta_{ijt})}{\lambda_f'(\theta_{ijt})} = -\frac{1-\eta}{\eta}$. Solving for the present value of wages $W_{ijt+1}$ yields

$$
W_{ijt+1} = \eta\nu_{ijt+1} - (1-\eta)V_{it+1}, \qquad \text{(A10)}
$$

which is the expression in the main text.

**Investment Block.** We now characterize the investment stage and, in the process, prove Proposition 2 from the main text. First, consider capital installed in period $t$ — and therefore with vintage productivity $A_t$ — in use in period $t+\tau$, $K_{jt+\tau}(v, A_t)$. The first-order condition for this variable is

$$
\begin{aligned}
&Q_{0,t+\tau}\int u_{jt+\tau}(v, A_t, \varepsilon)A_t\varepsilon f(v)\pi(\varepsilon)d\varepsilon \\
&\quad - Q_{0,t+\tau}\sum\nolimits_i v_i \int \lambda_{ijt+\tau}(v, A_t, \varepsilon)u_{jt+\tau}(v, \varepsilon, A)\pi(\varepsilon)d\varepsilon - Q_{0,t+\tau}q_{jt,t+\tau}(v) = 0 \\
&\implies q_{jt,t+\tau}(v) = \int_{\underline{\varepsilon}(v, A_t;\chi_{jt+\tau})}\left(A_t\varepsilon f(v) - \sum_i \chi_{ijt+\tau}v_i\right)\pi(\varepsilon)d\varepsilon, \qquad \text{(A11)}
\end{aligned}
$$

where the second line uses the facts that $\lambda_{ijt+\tau}(v, A_t, \varepsilon) = \chi_{ijt+\tau}$ and that $u_{jt}(v, A, \varepsilon) = 1$ for $\varepsilon \geq \underline{\varepsilon}(v, A_t; \chi_{jt+\tau})$ and 0 otherwise. The first-order condition for investment $X_{jt}(v)$ is

$$\mu_{jt}(v) = 1 - \sum_{\tau=1}^{\infty} Q_{t,t+\tau}(1-\delta)^{\tau-1} q_{jt,t+\tau}(v)$$

$$\implies \mu_{jt}(v) = 1 - \sum_{\tau=1}^{\infty} Q_{t,t+\tau}(1-\delta)^{\tau-1} \int_{\underline{\varepsilon}(v,A_t;\chi_{jt+\tau})} \left( A_t \varepsilon f(v) - \sum_i \chi_{ijt+\tau} v_i \right) \pi(\varepsilon) d\varepsilon, \qquad \text{(A12)}$$

where the second line uses the expression for $q_{jt,t+\tau}(v)$ from (A11).

*Optimal Capital Type.* As in the main text, we use (A12) to show that there is unique type of capital in which firms invest in period $t$. To do so, first note that since $\mu_{jt}(v)$ is a Lagrange multiplier, it has a minimum value at zero. Furthermore, if the RHS of (A12) is single-peaked, then there is a unique value of $v$ — call it $v_{jt}$ — which achieves that minimum. Therefore, we have $\mu_{jt}(v) > 0$ for all $v \neq v_{jt}$, which by complementary slackness implies that $X_{jt}(v) = 0$ for all $v_{jt} \neq 0$. For the optimal type $v_{jt}$, we have that (A12) holds with $\mu_{jt}(v_{jt}) = 0$. Hence, under the optimal choice of capital type, the first-order condition for investment becomes

$$1 = \sum_{\tau=1}^{\infty} Q_{t,t+\tau}(1-\delta)^{\tau-1} \int_{\underline{\varepsilon}(v,A_t;\chi_{jt+\tau})} \left( A_t \varepsilon f(v_{jt}) - \sum_i \chi_{ijt+\tau} v_{ijt} \right) \pi(\varepsilon) d\varepsilon.$$

Since this optimal type $v_{jt}$ is the minimizer of the RHS of (A12), it equivalently solves the maximization problem

$$v_{jt} = \underset{v}{\operatorname{argmax}} \sum_{\tau=1}^{\infty} Q_{t,t+\tau}(1-\delta)^{\tau-1} \int_{\underline{\varepsilon}(v,A_t;\chi_{jt+\tau})} \left( A_t \varepsilon f(v_{jt}) - \sum_i \chi_{ijt+\tau} v_{ijt} \right) \pi(\varepsilon) d\varepsilon.$$

The first-order condition for $v_{ijt}$ in this problem is

$$0 = \sum_{\tau=1}^{\infty} Q_{t,t+\tau}(1-\delta)^{\tau-1} \left( A_t \frac{\partial f(v)}{\partial v_i} \int_{\underline{\varepsilon}(v,A_t;\chi_{jt+\tau})}^{\infty} \varepsilon \pi(\varepsilon) d\varepsilon - \chi_{ijt+\tau} \int_{\underline{\varepsilon}(v,A_t;\chi_{jt+\tau})}^{\infty} \pi(\varepsilon) d\varepsilon \right)$$
$$- \sum_{\tau=1}^{\infty} Q_{t,t+\tau}(1-\delta)^{\tau-1} \frac{\partial \underline{\varepsilon}(v, A_t; \chi_{jt+\tau})}{\partial v_i} \pi(\underline{\varepsilon}(v, A_t; \chi_{jt+\tau})) \left( A_t \underline{\varepsilon}(v, A_t; \chi_{jt+\tau}) f(v) - \sum_i \chi_{ijt+\tau} v_i \right).$$

The top line is the derivatives holding fixed $\underline{\varepsilon}(v, A_t; \chi_{jt+\tau})$ and the second line is the derivatives with respect to $\underline{\varepsilon}(v, A_t; \chi_{jt+\tau})$, using the fundamental theorem of calculus.[20] However, each term in the summand in this second line is zero at the optimum. To see this, plug in $\underline{\varepsilon}(v, A_t; \chi_{jt+\tau}) = \frac{\sum_i \chi_{ijt+\tau} v_i}{A_t f(v)}$

[20]That is, $\int_a^b F'(x)dx = F(b) - F(a)$ so $\frac{\partial}{\partial a} \int_a^b F''(a)$.

to see that the term becomes 0 for each future period $\tau$. In the language of Gilchrist and Williams (2000), the marginal unit of capital is earns zero quasi-rents in each period.

To summarize, the optimal investment policy is characterized by two conditions:

$$0 = \sum_{\tau=1}^{\infty} Q_{t,t+\tau}(1-\delta)^{\tau-1} \left( A_t f_i(v_t) \int_{\underline{\varepsilon}(v,A_t;\chi_{jt+\tau})}^{\infty} \varepsilon \pi(\varepsilon) d\varepsilon - \chi_{ijt+\tau} \int_{\underline{\varepsilon}(v,A_t;\chi_{jt+\tau})}^{\infty} \pi(\varepsilon) d\varepsilon \right) \ \forall i \tag{A13}$$

$$1 = \sum_{\tau=1}^{\infty} Q_{t,t+\tau}(1-\delta)^{\tau-1} \int_{\underline{\varepsilon}(v,A_t;\chi_{jt+\tau})} \left( A_t \varepsilon f(v_{jt}) - \sum_i \chi_{ijt+\tau} v_{ijt} \right) \pi(\varepsilon) d\varepsilon, \tag{A14}$$

where $f_i(v_t) = \frac{\partial f(v_{jt})}{\partial v_i}$. The first equation is the first-order condition for the optimal type of capital and the second equation is the first-order condition for investment in the optimal type.

Going forward, it will be useful to simplify notation. First, we let $\underline{\varepsilon}_{t,t+\tau} = \underline{\varepsilon}(v, A_t; \chi_{jt+\tau})$ denote the utilization cutoff for capital installed in period $t$ to be used in period $t+\tau$. Second, following the main text, we define $\Pi^u(\underline{\varepsilon}_{t,t+\tau}) = \int_{\underline{\varepsilon}_{t,t+\tau}} \pi(\varepsilon) d\varepsilon$ and $\Pi^p(\underline{\varepsilon}_{t,t+\tau}) = \int_{\underline{\varepsilon}_{t,t+\tau}} \varepsilon \pi(\varepsilon) d\varepsilon$. With this notation, we can write (A13) and (A14) more compactly as

$$0 = \sum_{\tau=1}^{\infty} Q_{t,t+\tau}(1-\delta)^{\tau-1} \left( \Pi^p(\underline{\varepsilon}_{t,t+\tau}) A_t f_i(v_t) - \Pi^u(\underline{\varepsilon}_{t,t+\tau}) \chi_{ijt+\tau} \right)$$

$$1 = \sum_{\tau=1}^{\infty} Q_{t,t+\tau}(1-\delta)^{\tau-1} \left( \Pi^p(\underline{\varepsilon}_{t,t+\tau}) A_t \varepsilon f(v_{jt}) - \Pi^u(\underline{\varepsilon}_{t,t+\tau}) \sum_i \chi_{ijt+\tau} v_{ijt} \right).$$

**Summary of Equilibrium Conditions.** We now collect all of the equilibrium conditions of the model, including the optimality conditions from the households and the firms. Since we study a symmetric equilibrium, we drop the $j$ notation for individual firms going forward.

$$Q_{t,t+1} = \beta \frac{U_{cit+1}}{U_{cit}} \tag{A15}$$

$$V_{it+1} = \frac{U_{nit+1}}{U_{cit+1}} + Q_{t+1,t+2}(1-\sigma) V_{it+2} \tag{A16}$$

$$-\frac{U_{sit}}{U_{cit}} = Q_{t,t+1} \lambda_w(\theta_{it})(W_{it+1} + V_{it+1}) \tag{A17}$$

$$\sum\nolimits_{t=0}^{\infty} Q_{0,t} c_{it} = \zeta_i \mathbb{P} + \mathbb{I}_i + \sum\nolimits_{t=1}^{\infty} Q_{0,t} \lambda_w(\theta_{it-1}) s_{it-1} W_{it} \tag{A18}$$

$$\underline{\varepsilon}_{t,t+\tau} = \frac{\sum_i \chi_{it+\tau} v_{it}}{A_t f(v_t)} \tag{A19}$$

$$\nu_{it+1} = \chi_{it+1} + M_{it+1} \frac{U_{nit+1}}{U_{cit+1}} \frac{1}{\omega n_{it+1}} + Q_{t+1,t+2}(1-\sigma) \nu_{it+2} \tag{A20}$$

$$\frac{\kappa_{it}}{\lambda_f(\theta_{it})} = Q_{t,t+1}(\nu_{it+1} - W_{it+1}) \tag{A21}$$

$$W_{it+1} = \eta \nu_{it+1} - (1-\eta) V_{it+1} \text{ and } \gamma_{it+1} = \lambda_f(\theta_{it}) a_{it} \tag{A22}$$

$$n_{it+1} = (1-\sigma)n_{it} + \lambda_f(\theta_{it})a_{it} \tag{A23}$$

$$M_{it+1} = \gamma_{it+1} + (1-\sigma)M_{it} \tag{A24}$$

$$\theta_{it} = a_{it}/s_{it} \tag{A25}$$

$$1 = \sum_{\tau=1}^{\infty} Q_{t,t+\tau}(1-\delta)^{\tau-1}\left(\Pi^p(\underline{\varepsilon}_{t,t+\tau})A_t\varepsilon f(v_t) - \Pi^u(\underline{\varepsilon}_{t,t+\tau})\sum_i \chi_{ijt+\tau}v_{it}\right) \tag{A26}$$

$$0 = \sum_{\tau=1}^{\infty} Q_{t,t+\tau}(1-\delta)^{\tau-1}\left(\Pi^p(\underline{\varepsilon}_{t,t+\tau})A_t f_i(v_t) - \Pi^u(\underline{\varepsilon}_{t,t+\tau})\chi_{ijt+\tau}\right) \tag{A27}$$

$$Y_t = \sum_{\tau=1}^{\infty} \Pi^p(\underline{\varepsilon}_{t-\tau,t})A_{t-\tau}f(v_{t-\tau})(1-\delta)^{\tau-1}X_{t-\tau} \tag{A28}$$

$$Y_t = \sum_i \mu_i c_{it} + X_t + \sum_i \mu_i \kappa_{it} a_{it}. \tag{A29}$$

$$\mu_i n_{it} = \sum_{\tau=1}^{\infty} \Pi^u(\underline{\varepsilon}_{t-\tau,t})v_{it-\tau}(1-\delta)^{\tau-1}X_{t-\tau}. \tag{A30}$$

Equation (A28) simplifies the expression for aggregate output from the main text using our results about optimal investment and the definition of $\Pi^p(\underline{\varepsilon}_{t-\tau,t})$. In particular, aggregate output equals output produced by each vintage of capital, $\Pi^p(\underline{\varepsilon}_{t-\tau,t})A_{t-\tau}f(v_{t-\tau})$, times the amount of capital of that vintage which is remaining, $(1-\delta)^{\tau-1}X_{t-\tau}$. Equation (A29) is market clearing for aggregate output. Finally, equation (A30) equates aggregate employment of type-$i$ worker in period $t$ to the amount of that type of labor assigned to each vintage.

### A.3 Detrending

Due to capital-embodied technological progress in vintage productivity $A_t$, the equilibrium allocation is not stationary over time. In this subsection, we describe how to detrend the model into stationary form, which will be useful for numerically solving the model. As in the main text, we assume that $\kappa_{it} = (1+g)^t\kappa_i$ so that vacancy-posting costs grow along with the economy.

A balanced growth path will have the following properties:

(i) The following variables grow along with the economy: $c_{it}, W_{it+1}, V_{it+1}, \chi_{it}, \nu_{it}, Y_t, X_t$. Let tildes denote detrended variables, e.g. $\widetilde{c}_{it} = c_{it}/(1+g)^t$.

(ii) The following variables shrink over time: $v_{it}$. Let $\widetilde{v}_{it} = v_{it}(1+g)^t$.

(iii) The following variables are stationary: $s_{it}, n_{it+1}, a_{it}, \theta_{it}, \gamma_{it+1}, M_{it+1}, Q_{t,t+1}$.

We now go through each of the equilibrium conditions and replace the original non-stationary variables with their stationary version.

### A.3.1 Household

Using our functional form for the utility function, namely

$$U_t(c_{it}, s_{it}, n_{it}) = \log\left(c_{it} - (1+g)^t v(n_{it}) - (1+g)^t h(s_{it})\right),$$

the ratio of date-0 output prices (A15) becomes

$$\begin{aligned}
Q_{t,t+1} &= \beta \frac{c_{it} - (1+g)^t v(n_{it}) - (1+g)^t h(s_{it})}{c_{it+1} - (1+g)^{t+1} v(n_{it+1}) - (1+g)^{t+1} h(s_{it+1})} \\
&= \beta \frac{(1+g)^t \widetilde{c}_{it} - (1+g)^t v(n_{it}) - (1+g)^t h(s_{it})}{(1+g)^{t+1}\widetilde{c}_{it+1} - (1+g)^{t+1} v(n_{it+1}) - (1+g)^{t+1} h(s_{it+1})} \\
&= \frac{\beta}{1+g} \frac{\widetilde{c}_{it} - v(n_{it}) - h(s_{it})}{\widetilde{c}_{it+1} - v(n_{it+1}) - h(s_{it+1})}.
\end{aligned}$$

The equation defining the disutility of labor (A16) becomes

$$\begin{aligned}
\widetilde{V}_{it+1}(1+g)^{t+1} &= -(1+g)^{t+1} v'(n_{it+1}) + Q_{t+1,t+2}(1-\sigma)\widetilde{V}_{it+2}(1+g)^{t+2} \\
\implies \widetilde{V}_{it+1} &= -v'(n_{it+1}) + Q_{t+1,t+2}(1+g)(1-\sigma)\widetilde{V}_{it+2}.
\end{aligned}$$

The first-order condition for optimal search effort (A17) becomes

$$\begin{aligned}
(1+g)^t h'(s_{it}) &= Q_{t,t+1}\lambda_w(\theta_{it})(1+g)^{t+1}\left(\widetilde{W}_{it+1} + \widetilde{V}_{it+1}\right) \\
\implies h'(s_{it}) &= Q_{t,t+1}(1+g)\lambda_w(\theta_{it})\left(\widetilde{W}_{it+1} + \widetilde{V}_{it+1}\right).
\end{aligned}$$

The budget constraint (A18) becomes

$$\sum_{t=0}^{\infty} Q_{0,t}(1+g)^t \widetilde{c}_{it} = \zeta_i \mathbb{P} + \mathbb{I}_i + \sum_{t=0}^{\infty} Q_{0,t+1}(1+g)^{t+1}\lambda_w(\theta_{it}) s_{it} \widetilde{W}_{it+1}.$$

### A.3.2 Firms

We go through the utilization block, the hiring block, and the investment block.

**Utilization Block.** The production stage is summarized by the expression for the productivity threshold $\underline{\varepsilon}_{t,t+\tau} = \sum_i \chi_{it+\tau} v_{it}/A_t f(v_t)$. Recalling that $v_{it}$ *shrinks* are rate $g$ so that $\widetilde{v}_{it} = (1+g)^t v_{it}$, in detrended terms, the numerator of the threshold is

$$\sum_i \chi_{it+\tau} v_{it} = \sum_i \left[ (1+g)^{t+\tau} \frac{\chi_{it+\tau}}{(1+g)^{t+\tau}} \right] \left[ \frac{1}{(1+g)^t} v_{it} (1+g)^t \right] = (1+g)^\tau \sum_i \widetilde{\chi}_{it+\tau} \widetilde{v}_{it}.$$

Recalling that $F(K, N_1, \ldots, N_I) = K^\alpha G(N_1, \ldots, N_I)^{1-\alpha}$ and both $F$ and $G$ are CRS gives

$$f(v_1, \ldots, v_I) = F(1, \frac{N_1}{K}, \ldots, \frac{N_I}{K}) = 1^\alpha G(\frac{N_1}{K}, \ldots, \frac{N_I}{K})^{1-\alpha} = G(v_1, \ldots, v_I)^{1-\alpha}.$$

Since $f(v_t) = G(v)^{1-\alpha}$, the denominator of the cutoff is

$$\begin{aligned} A_t f(v_t) &= ((1+g)^{1-\alpha})^t G(v_t)^{1-\alpha} = ((1+g)^{1-\alpha})^t G(\frac{\widetilde{v}_t}{(1+g)^t})^{1-\alpha} \\ &= ((1+g)^{1-\alpha})^t \left[ \frac{1}{(1+g)^t} G(\widetilde{v}_t) \right]^{1-\alpha} = f(\widetilde{v}_t), \end{aligned}$$

where the first equality uses that $A_t = ((1+g)^{1-\alpha})^t$ and $f(v_t) = G(v_t)^{1-\alpha}$, and the third equality uses the fact that $G(v_t)$ is constant returns to scale. Putting these results about the numerator and denominator together, we get

$$\underline{\varepsilon}_{t,t+\tau} = (1+g)^\tau \frac{\sum_i \widetilde{\chi}_{it+\tau} \widetilde{v}_{it}}{f(\widetilde{v}_t)}.$$

**Hiring Block.** The expression for the present value of a worker (A20) becomes

$$\begin{aligned} &\widetilde{\nu}_{it+1}(1+g)^{t+1} = \widetilde{\chi}_{it+1}(1+g)^{t+1} + M_{it+1} v'(n_{it+1})(1+g)^{t+1} \frac{1}{\omega n_{it+1}} + Q_{t+1,t+2}(1-\sigma)\widetilde{\nu}_{it+2}(1+g)^{t+2} \\ &\implies \widetilde{\nu}_{it+1} = \widetilde{\chi}_{it+1} + M_{it+1} v'(n_{it+1}) \frac{1}{\omega n_{it+1}} + Q_{t+1,t+2}(1+g)(1-\sigma)\widetilde{\nu}_{it+2}. \end{aligned}$$

The first-order condition for optimal vacancy-posting (A21) becomes

$$\frac{\kappa_i (1+g)^t}{\lambda_f(\theta_{it})} = Q_{t,t+1}(1+g)^{t+1} \left( \widetilde{\nu}_{it+1} - \widetilde{W}_{it+1} \right) \implies \frac{\kappa_i}{\lambda_f(\theta_{it})} = Q_{t,t+1}(1+g) \left( \widetilde{\nu}_{it+1} - \widetilde{W}_{it+1} \right).$$

The-condition for wages (A22) becomes

$$\widetilde{W}_{it+1}(1+g)^{t+1} = \eta \widetilde{\nu}_{it+1}(1+g)^{t+1} - (1-\eta)\widetilde{V}_{it+1}(1+g)^{t+1} \implies \widetilde{W}_{it+1} = \eta \widetilde{\nu}_{it+1} - (1-\eta)\widetilde{V}_{it+1}$$

and $\gamma_{it+1} = \lambda_f(\theta_{it}) a_{it}$ is already stationary. The evolution of employment (A23), the definition of the quasi-multipliers (A24), and the definition of market tightness (A25) are already stationary.

**Investment Block.** First, consider the condition that equates marginal cost with marginal benefit of new capital, (A26). As argued with the productivity cutoff above, the terms with $A_t f(v_t) = f(\widetilde{v}_t)$, and the terms with $\sum_i \chi_{it+\tau} v_{it} = (1+g)^\tau \sum_i \widetilde{\chi}_{it+\tau} \widetilde{v}_{it}$. Thus

$$1 = \sum_{\tau=1}^{\infty} Q_{t,t+\tau}(1-\delta)^{\tau-1} \left( \Pi^p(\underline{\varepsilon}_{t,t+\tau}) f(\widetilde{v}_t) - \Pi^u(\underline{\varepsilon}_{t,t+\tau}) \sum_i \widetilde{\chi}_{it+\tau} \widetilde{v}_{it} \right).$$

Next consider the first-order condition for the optimal type of capital, (A27). Note that

$$\begin{aligned} f_i(v_t) &= \frac{\partial}{\partial v_i} f(v_t) = \frac{\partial}{\partial v_i} \left[G(v_t)\right]^{1-\alpha} \\ &= (1-\alpha) G(v_t)^{-\alpha} G_i(v_t) \\ &= (1-\alpha) G(\frac{\widetilde{v}_t}{(1+g)^t})^{-\alpha} G_i(\frac{\widetilde{v}_t}{(1+g)^t}) \\ &= ((1+g)^\alpha)^t (1-\alpha) G(\widetilde{v}_t)^{-\alpha} G_i(\widetilde{v}_t) = ((1+g)^\alpha)^t f_i(\widetilde{v}_t), \end{aligned}$$

where the fourth line uses the fact that $G(v_t)$ is homogenous of degree one (and therefore its derivatives are homogenous of degree zero). Therefore, we have that the terms $A_t f_i(v_t) = ((1+g)^{1-\alpha})^t ((1+g)^\alpha)^t f_i(\widetilde{v}_t) = (1+g)^t f_i(\widetilde{v}_t)$. Plugging these into the first-order condition for the optimal type of capital (A27) gives

$$\begin{aligned} 0 &= \sum_{\tau=1}^{\infty} Q_{t,t+\tau}(1-\delta)^{\tau-1} \left( \Pi^p(\underline{\varepsilon}_{t,t+\tau})(1+g)^t f_i(\widetilde{v}_t) - \Pi^u(\underline{\varepsilon}_{t,t+\tau})(1+g)^{t+\tau} \widetilde{\chi}_{it+\tau} \right) \\ \implies 0 &= \sum_{\tau=1}^{\infty} Q_{t,t+\tau}(1-\delta)^{\tau-1} \left( \Pi^p(\underline{\varepsilon}_{t,t+\tau}) f_i(\widetilde{v}_t) - \Pi^u(\underline{\varepsilon}_{t,t+\tau})(1+g)^\tau \widetilde{\chi}_{it+\tau} \right). \end{aligned}$$

### A.3.3 Aggregate Conditions

We start with the definition of aggregate output (A28). As we argued above, the terms $A_{t-\tau} f(v_{t-\tau}) = f(\widetilde{v}_{t-\tau})$. Using this result, the equation becomes

$$\begin{aligned} \widetilde{Y}_t (1+g)^t &= \sum_{\tau=1}^{\infty} \Pi^p(\underline{\varepsilon}_{t-\tau,t}) f(\widetilde{v}_{t-\tau})(1-\delta)^{\tau-1}(1+g)^{t-\tau} \widetilde{X}_{t-\tau} \\ \implies \widetilde{Y}_t &= \sum_{\tau=1}^{\infty} \Pi^p(\underline{\varepsilon}_{t-\tau,t}) f(\widetilde{v}_{t-\tau})(1-\delta)^{\tau-1}(1+g)^{-\tau} \widetilde{X}_{t-\tau} \\ \implies \widetilde{Y}_t &= \sum_{\tau=1}^{\infty} \Pi^p(\underline{\varepsilon}_{t-\tau,t}) f(\widetilde{v}_{t-\tau}) \left(\frac{1-\delta}{1+g}\right)^{\tau-1} \frac{\widetilde{X}_{t-\tau}}{1+g}. \end{aligned}$$

For aggregate employment (A30) , note that $v_{it-\tau}X_{t-\tau} = \frac{\widetilde{v}_{it-\tau}}{(1+g)^{t-\tau}}\widetilde{X}_{t-\tau}(1+g)^{t-\tau} = \widetilde{v}_{it-\tau}\widetilde{x}_{t-\tau}$ is already stationary. So we have

$$\mu_i n_{it} = \sum_{\tau=1}^{\infty} \Pi^u(\underline{\varepsilon}_{t-\tau,t})\widetilde{v}_{it-\tau}(1-\delta)^{\tau-1}\widetilde{X}_{t-\tau}.$$

Finally, the output market clearing condition (A29) is

$$\widetilde{Y}_t(1+g)^t = \sum_i \mu_i \widetilde{c}_{it}(1+g)^t + \widetilde{X}_t(1+g)^t + \sum_i \mu_i\kappa_i(1+g)^t a_{it} \implies \widetilde{Y}_t = \sum_i \mu_i\widetilde{c}_{it} + \widetilde{X}_t + \sum_i \mu_i\kappa_i a_{it}.$$

### A.3.4 Summary of Detrended Equilibrium Conditions

We now collect all of these detrended equilibrium conditions. In our quantitative work, we compute the transition paths by solving this large nonlinear system,

$$Q_{t,t+1} = \frac{\beta}{1+g}\frac{\widetilde{c}_{it} - v(n_{it}) - h(s_{it})}{\widetilde{c}_{it+1} - v(n_{it+1}) - h(s_{it+1})} \tag{A31}$$

$$\widetilde{V}_{it+1} = -v'(n_{it+1}) + Q_{t+1,t+2}(1+g)(1-\sigma)\widetilde{V}_{it+2} \tag{A32}$$

$$h'(s_{it}) = Q_{t,t+1}(1+g)\lambda_w(\theta_{it})\left(\widetilde{W}_{it+1} + \widetilde{V}_{it+1}\right) \tag{A33}$$

$$\sum_{t=0}^{\infty} Q_{0,t}(1+g)^t\widetilde{c}_{it} = \zeta_i\mathbb{P} + \mathbb{I}_i + \sum_{t=0}^{\infty} Q_{0,t+1}(1+g)^{t+1}\lambda_w(\theta_{it})s_{it}\widetilde{W}_{it+1} \tag{A34}$$

$$\underline{\varepsilon}_{t,t+\tau} = (1+g)^\tau\frac{\sum_i \widetilde{\chi}_{it+\tau}\widetilde{v}_{it}}{f(\widetilde{v}_t)} \tag{A35}$$

$$\widetilde{\nu}_{it+1} = \widetilde{\chi}_{it+1} + M_{it+1}v'(n_{it+1})\frac{1}{\omega n_{it+1}} + Q_{t+1,t+2}(1+g)(1-\sigma)\widetilde{\nu}_{it+2} \tag{A36}$$

$$\frac{\kappa_i}{\lambda_f(\theta_{it})} = Q_{t,t+1}(1+g)\left(\widetilde{\nu}_{it+1} - \widetilde{W}_{it+1}\right) \tag{A37}$$

$$\widetilde{W}_{it+1} = \eta\widetilde{\nu}_{it+1} - (1-\eta)\widetilde{V}_{it+1} \text{ and } \gamma_{it+1} = \lambda_f(\theta_{it})a_{it} \tag{A38}$$

$$\mathbb{P} = \sum_{t=0}^{\infty} Q_{0,t}(1+g)^t\big[\widetilde{Y}_t - \widetilde{X}_t - \sum_i \mu_i\left(\kappa_i a_{it} + \lambda_f(\theta_{it})a_{it}Q_{t,t+1}(1+g)\widetilde{W}_{it+1}\right)\big] - \sum_i \mu_i\mathbb{I}_i \tag{A39}$$

$$M_{it+1} = (1-\sigma)M_{it} + \gamma_{it+1} \tag{A40}$$

$$1 = \sum_{\tau=1}^{\infty} Q_{t,t+\tau}(1-\delta)^{\tau-1}\left(\Pi^p(\underline{\varepsilon}_{t,t+\tau})f(\widetilde{v}_t) - \Pi^u(\underline{\varepsilon}_{t,t+\tau})\sum_i \widetilde{\chi}_{it+\tau}\widetilde{v}_{it}\right) \tag{A41}$$

$$0 = \sum_{\tau=1}^{\infty} Q_{t,t+\tau}(1-\delta)^{\tau-1}\left(\Pi^p(\underline{\varepsilon}_{t,t+\tau})f_i(\widetilde{v}_t) - \Pi^u(\underline{\varepsilon}_{t,t+\tau})(1+g)^\tau\widetilde{\chi}_{it+\tau}\right) \tag{A42}$$

$$n_{it+1} = (1-\sigma)n_{it} + \lambda_w(\theta_{it})s_{it} \tag{A43}$$

$$\theta_{it} = \frac{a_{it}}{s_{it}} \tag{A44}$$

$$\widetilde{Y}_t = \sum_{\tau=1}^{\infty} \Pi^p(\underline{\varepsilon}_{t-\tau,t}) f(\widetilde{v}_{t-\tau}) \left(\frac{1-\delta}{1+g}\right)^{\tau-1} \frac{\widetilde{X}_{t-\tau}}{1+g} \tag{A45}$$

$$\mu_i n_{it} = \sum_{\tau=1}^{\infty} \Pi^u(\underline{\varepsilon}_{t-\tau,t}) \widetilde{v}_{it-\tau}(1-\delta)^{\tau-1}\widetilde{X}_{t-\tau} \tag{A46}$$

$$\widetilde{Y}_t = \sum_i \mu_i \widetilde{c}_{it} + \widetilde{X}_t + \sum_i \mu_i \kappa_i a_{it}. \tag{A47}$$

### A.4 Balanced Growth Path

The balanced growth path is simply the steady state of the detrended system. In this subsection, we collect the conditions that define the BGP and then simplify them to prove Lemma 3 and Lemma 4 from the main text. We also derive the formula for the wage markdown (26) from the main text.

**Summary of BGP Conditions.** Collecting the summary of detrended equilibrium conditions from above and imposing a steady state, we get the system

$$Q_{t,t+1} \equiv \widetilde{\beta} = \frac{\beta}{1+g} \tag{A48}$$

$$h'(s_i) = \frac{\beta}{1-\beta(1-\sigma)}\lambda_w(\theta_i)\left[\widetilde{w}_i - v'(n_i)\right] \tag{A49}$$

$$\sum_{t=0}^{\infty} \beta^t \widetilde{c}_i = \zeta_i \mathbb{P} + \mathbb{I}_i + \sum_{t=0}^{\infty} \beta^{t+1}\lambda_w(\theta_i)s_i\widetilde{W}_i \tag{A50}$$

$$\widetilde{\nu}_i = \frac{1}{1-\beta(1-\sigma)}\left[\widetilde{\chi}_i - \frac{1}{\omega}v'(n_i)\right] \tag{A51}$$

$$\widetilde{W}_i = \eta\widetilde{\nu}_i + (1-\eta)\frac{v'(n_i)}{1-\beta(1-\sigma)} \tag{A52}$$

$$\frac{\kappa_i}{\lambda_f(\theta_i)} = \beta(\widetilde{\nu}_i - \widetilde{W}_i) \tag{A53}$$

$$\underline{\varepsilon}_\tau = (1+g)^\tau \frac{\sum_i \widetilde{\chi}_i \widetilde{v}_i}{f(\widetilde{v})} \tag{A54}$$

$$0 = f_i(v)\sum_{\tau=1}^{\infty} \widetilde{\beta}^\tau (1-\delta)^{\tau-1}\Pi^p(\underline{\varepsilon}_\tau) - \widetilde{\chi}_i \sum_{\tau=1}^{\infty} \widetilde{\beta}^\tau (1-\delta)^{\tau-1}(1+g)^\tau \Pi^u(\underline{\varepsilon}_\tau) \tag{A55}$$

$$1 = f(\widetilde{v})\sum_{\tau=1}^{\infty} \widetilde{\beta}^\tau (1-\delta)^{\tau-1}\Pi^p(\underline{\varepsilon}_\tau) - \sum_{\tau=1}^{\infty} \widetilde{\beta}^\tau (1-\delta)^{\tau-1}\Pi^u(\underline{\varepsilon}_\tau)\sum_i \widetilde{\chi}_i\widetilde{v}_i \tag{A56}$$

$$\widetilde{Y} = \sum_{\tau=1}^{\infty} \Pi^p(\underline{\varepsilon}_\tau) f(\widetilde{v}) \left(\frac{1-\delta}{1+g}\right)^{\tau-1} \frac{\widetilde{X}}{1+g} \tag{A57}$$

$$\mu_i n_i = \sum_{\tau=1}^{\infty} \Pi^u(\underline{\varepsilon}_\tau) \widetilde{v}_i (1-\delta)^{\tau-1} \widetilde{X} \tag{A58}$$

$$\widetilde{Y} = \sum_i \mu_i \widetilde{c}_i + \widetilde{X} + \sum_i \mu_i \kappa_i a_i \tag{A59}$$

$$\theta_i = a_i / s_i \tag{A60}$$

$$\mathbb{P} = \frac{\widetilde{Y} - \widetilde{X} - \sum_i \mu_i \left(\kappa_i a_i + \beta \sigma n_i W_i\right)}{1-\beta} - \sum_i \mu_i \mathbb{I}_i \tag{A61}$$

$$\sigma n_i = \lambda_w(\theta_i) s_i. \tag{A62}$$

Note that the BGP of the putty-clay model is not the same as the model with standard capital. The reason is that the labor intensities $v$ in the putty-clay model are chosen before the realization of the capital quality shock $\varepsilon$ in the putty-clay model, but after the realization of $\varepsilon$ in the model with standard capital. Therefore, firms in the standard model will implicitly assign more workers to high-$\varepsilon$ machines, which is not possible in the putty-clay model. If we allowed firms to choose the labor intensities $v$ after the realization of capital quality shocks $\varepsilon$ in the putty-clay model, then we would not have an active utilization margin, which is a key feature of our analysis.

**Reduced System Characterizing the BGP.** Under our preference specification, the labor market equilibrium and investment decisions are separable from the consumption allocation. This property allows us to significantly reduce the number of equations which characterize the BGP in Lemma 8 below (a version of Lemma 3 from the main text). In Appendix B, we use this Lemma to show that a combination of type-specific minimum wages and vacancy-posting subsidies can achieve the competitive allocation (Proposition 6 from the main text).

**Lemma 8.** *Along the balanced growth path, the labor allocations and wages are determined by the following equations:*

*(i) optimal cut-off for idiosyncratic productivity of capital*

$$\underline{\varepsilon}_1 = (1+g)(1-\alpha)m(\underline{\varepsilon}_1), \tag{A63}$$

*where $m(\underline{\varepsilon}_1)$ is defined by*

$$m(\underline{\varepsilon}_1) = \frac{\sum_{\tau=1}^{\infty} \widetilde{\beta}^{\tau}(1-\delta)^{\tau-1}\Pi^p((1+g)^{\tau-1}\underline{\varepsilon}_1)}{\sum_{\tau=1}^{\infty} \widetilde{\beta}^{\tau}(1-\delta)^{\tau-1}(1+g)^{\tau}\Pi^u((1+g)^{\tau-1}\underline{\varepsilon}_1)}; \tag{A64}$$

*(ii) the marginal unit of capital earns zero profit*

$$1 = \alpha \left[\sum_{\tau=1}^{\infty} \widetilde{\beta}^{\tau}(1-\delta)^{\tau-1}\Pi^p((1+g)^{\tau-1}\underline{\varepsilon}_1)\right] f(\widetilde{v}); \tag{A65}$$

*(iii) flow wages*

$$\widetilde{w}_i = \eta[f_i(\widetilde{v})m(\underline{\varepsilon}_1) - v'(n_i)/\omega] + (1-\eta)v'(n_i); \tag{A66}$$

*(iv) optimal vacancy posting*

$$\kappa_i = \beta\lambda_f(\theta_i)\frac{f_i(\widetilde{v})m(\underline{\varepsilon}_1) - \widetilde{w}_i - v'(n_i)/\omega}{1-\beta(1-\sigma)}; \tag{A67}$$

*(v) optimal household search*

$$h'(s_i) = \beta\lambda_w(\theta_i)\frac{\widetilde{w}_i - v'(n_i)}{1-\beta(1-\sigma)}; \tag{A68}$$

*(vi) the steady-state law of motion for employment*

$$\sigma n_i = \lambda_w(\theta_i)s_i; \tag{A69}$$

*(vii) labor market clearing*

$$\frac{\mu_i n_i}{\widetilde{v}_i} = \frac{\mu_1 n_1}{\widetilde{v}_1}. \tag{A70}$$

*Proof.* The system of equations consists of variables $1+5N$ variables $\underline{\varepsilon}_1$, $\widetilde{v}_i$, $n_i$, $\theta_i$, $s_i$, $\widetilde{w}_i$, with $1+5N$ equations (A63)–(A70).

*(i)* On a BGP, the optimal investment equation (A55) equates capital's marginal product to the marginal cost of operation. Hence, the shadow value of a worker $\widetilde{\chi}_i$ is simply that worker's marginal product, given by

$$\widetilde{\chi}_i = f_i(\widetilde{v})m(\underline{\varepsilon}_1), \tag{A71}$$

where $m$ is defined in equation (A64) as a weighted mean idiosyncratic productivity of capital that is utilized. Substituting $\widetilde{\chi}_i$ into the equation for $\underline{\varepsilon}_\tau$ in (A54) and evaluating at $\tau = 1$ gives

$$\underline{\varepsilon}_1 = (1+g)\frac{\sum_{i=1}^{N} f_i(\widetilde{v})\widetilde{v}_i}{f(\widetilde{v})} m(\underline{\varepsilon}_1).$$

From the definition of $f(v) = F(1, v) = v^{1-\alpha}$, we know that $f(v)$ is homogeneous of degree $1-\alpha$, which is the labor share of the production function. Applying Euler's theorem for homogeneous equations gives $\sum_{i=1}^{I} f_i(\widetilde{v})v_i = (1-\alpha)f(\widetilde{v})$, so we have

$$\underline{\varepsilon}_1 = (1+g)(1-\alpha)m(\underline{\varepsilon}_1),$$

which is equation (A63). This is independent of any other labor market condition and only depends on the parameters $g$, $\alpha$, and the dispersion of idiosyncratic shocks $\sigma_\varepsilon$.

*(ii)* Substituting our expression for $\sum_{i=1}^{I} f_i(\widetilde{v})\widetilde{v}_i$ into the optimal investment condition (A56) obtains

$$\alpha f(\widetilde{v}) = \left[\sum_{\tau=1}^{\overline{\tau}} \widetilde{\beta}^\tau (1-\delta)^{\tau-1} \Pi^p((1+g)^{\tau-1}\underline{\varepsilon}_1)\right]^{-1}$$

which rearranges to the expression in equation (A65).

*(iii)* Substituting the value of $\widetilde{\chi}_i$ from (A71) into the definition of $\widetilde{\nu}_i$ (A51), the present value of a type $i$ worker to the firm $\widetilde{\nu}_i$ is

$$\widetilde{\nu}_i = \frac{f_i(\widetilde{v})m(\underline{\varepsilon}_1) - v'(n_i)/\omega}{1-\beta(1-\sigma)}.$$

Substituting $\widetilde{\nu}_i$ into the vacancy-posting condition (A53) gives the simplified vacancy-posting condition in equation (A67).

*(iv)* The household optimal search condition (A68) is a restatement of equation (A49).

*(v)* Substituting $\widetilde{\nu}_i$ into the BGP wage equation (A52) gives (A66).

*(vi)* The transition law for labor (A69) is a restatement of equation (A62).

*(vii)* The BGP labor market clearing condition (A58) rearranges to

$$\frac{\mu_i n_i}{\widetilde{v}_i} = \widetilde{X} \sum_{\tau=1}^{\overline{\tau}} \Pi^u(\underline{\varepsilon}_\tau)(1-\delta)^{\tau-1}$$

Observe that the right-hand side of this equation is independent of $i$, so the left-hand side must be the same for all $i$, which gives equation (A70). □

**Wage Markdowns.** We now describe how we arrive at the expression for the BGP wage markdown (26) from the main text. We will use equations (A51), (A53), and (A52) from the balanced growth path, reproduced here in rearranged form:

$$\widehat{\nu}_i = \widetilde{\chi}_i - \frac{1}{\omega} v'(n_i) \tag{A72}$$

$$\frac{1-\beta(1-\sigma)}{\beta} \frac{\kappa_i}{\lambda_f(\theta_{it})} = (\rho+\sigma) \frac{\kappa_i}{\lambda_f(\theta_{it})} = \widehat{\nu}_i - \widetilde{w}_i \tag{A73}$$

$$\widetilde{w}_i = \eta \widehat{\nu}_i + (1-\eta) v'(n_i) \tag{A74}$$

where $\widehat{\nu}_i = [1-\beta(1-\sigma)]\widetilde{\nu}_i$ is the flow value of the worker to the firm and $\rho = \frac{1}{\beta} - 1$ is the rate of time preference such that $\frac{1-\beta(1-\sigma)}{\beta} = \frac{1}{\beta} - (1-\sigma) = \rho + \sigma$ in (A73).

The expression for $\widehat{\nu}_i$, (A72), can be rewritten as $\widetilde{\chi}_i = \widehat{\nu}_i + \frac{1}{\omega} v'(n_i)$. The expression for the annuitized vacancy posting costs, (A73), can be written as $\widehat{\nu}_i = (\rho+\sigma)\frac{\kappa_i}{\lambda_f(\theta_i)} + \widetilde{w}_i$. Substituting this expression for $\widehat{\nu}_i$ into $\widetilde{\chi}_i$ implies that the ratio of $\widetilde{w}_i$ to $\widetilde{\chi}_i$ is given by

$$\frac{\widetilde{w}_i}{\widetilde{\chi}_i} = \frac{\widetilde{w}_i}{\widetilde{w}_i + (\rho+\sigma)\frac{\kappa_i}{\lambda_f(\theta_i)} + \frac{1}{\omega} v'(n_i)} = \frac{1}{1 + (\rho+\sigma)\frac{\kappa_i}{\lambda_f(\theta_i)}\frac{1}{\widetilde{w}_i} + \frac{1}{\omega}\frac{v'(n_i)}{\widetilde{w}_i}}, \tag{A75}$$

where the second equation divides the numerator and denominator by $\widetilde{w}_i$. We now eliminate the wage from the RHS of (A75). Equation (A73) can be written $\widetilde{w}_i = \widehat{\nu}_i - (\rho+\sigma)\frac{\kappa_i}{\lambda_f(\theta_{it})}$. Plug this into the wage equation (A74) and rearrange to get $\widehat{\nu}_i = v'(n_i) + (\rho+\sigma)\frac{\kappa_i}{\lambda_f(\theta_{it})}\frac{1}{1-\eta}$. Then plug this back into (A73) to get $\widetilde{w}_i = v'(n_i) + \frac{\eta}{1-\eta}(\rho+\sigma)\frac{\kappa_i}{\lambda_f(\theta_i)}$. Finally, plug this into (A75) to get

$$\frac{\widetilde{w}_i}{\widetilde{\chi}_i} = \left[1 + \frac{1}{\omega} \times \frac{v'(n_i)}{v'(n_i) + \frac{\eta}{1-\eta}(\rho+\sigma)\frac{\kappa_i}{\lambda_f(\theta_i)}} + \frac{(r+\sigma)\frac{\kappa_i}{\lambda_f(\theta_i)}}{v'(n_i) + \frac{\eta}{1-\eta}(\rho+\sigma)\frac{\kappa_i}{\lambda_f(\theta_i)}}\right]^{-1} \tag{A76}$$

as in the main text.

**Firm-Specific Labor Supply Elasticity.** We can also use this algebra to derive equation the firm-specific labor supply elasticity along the BGP from the main text. As in the main text, first consider the participation constraint for firm $j$ along the BGP:

$$\frac{\beta}{1-\beta(1-\sigma)} \lambda_w(\theta_{ij}) \left[\widetilde{w}_{ij} - v'(n_i)\left(\frac{n_{ij}}{n_i}\right)^{\frac{1}{\omega}}\right] \geq \widetilde{\mathcal{W}}_i.$$

Differentiating with respect to the wage $\widetilde{w}_{ij}$ and $n_{ij}$ holding $\theta_{ij}$ and $\widetilde{\mathcal{W}}_i$ fixed, we get

$$d\widetilde{w}_{ij} - v'(n_i)\frac{1}{\omega}\left(\frac{n_{ij}}{n_i}\right)^{\frac{1}{\omega}-1}\frac{dn_{ij}}{n_i} = 0 \implies d\log\widetilde{w}_{ij}\cdot\widetilde{w}_i = \frac{v'(n_i)}{\omega}d\log n_{ij},$$

where the second line uses the fact that $n_{ij} = n_i$ and $\widetilde{w}_{ij} = \widetilde{w}_i$ in a symmetric equilibrium. From the derivation of the markdown equation above, we know that $\widetilde{w}_i = v'(n_i) + \frac{\eta}{1-\eta}\frac{1-\beta(1-\sigma)}{\beta}\frac{\kappa_i}{\lambda_f(\theta_i)}$. Furthermore, from our calibration results, we also know that the annuitized portion of vacancy-posting costs are small, implying that $\widetilde{w}_i \approx v'(n_i)$. Plugging this in gives

$$d\log\widetilde{w}_{ij}\cdot v'(n_i) \approx \frac{v'(n_i)}{\omega}d\log n_{ij} \implies \frac{d\log n_{ij}}{d\log\widetilde{w}_{ij}} \approx \omega. \tag{A77}$$

# B Labor Market Policies

In this appendix, we show how to add to the model the two labor market policies that we study in the main text: the minimum wage and transfer programs (like the EITC).

## B.1 Minimum Wage

As in Appendix A, we first focus on the firm's problem ignoring initial conditions in order to see how the minimum wage changes the key decisions of the firm. We use this analysis to characterize the long-run effects of the minimum wage along the BGP and prove Proposition 6 from the main text. Finally, we add back in the initial conditions and discuss why firms are reluctant to fire workers in our quantitative work. Throughout, we focus only on the equations that change relative to the baseline model from Appendix A.

### B.1.1 Introducing the Minimum Wage

The firm's problem is the same as in Appendix A except that we add a minimum wage constraint

$$W_{ijt+1} \geq \underline{W}_{t+1} \quad \text{for all } t \geq 0 \ (\times Q_{0,t}\rho_{ijt+1}) \tag{A1}$$

and a nonnegativity condition on vacancies

$$a_{ijt} \geq 0 \text{ for all } t \geq 0 \quad (\times Q_{0,t}\xi^a_{ijt}), \tag{A2}$$

where $Q_{0,t}\rho_{ijt+1}$ and $Q_{0,t}\xi^a_{ijt}$ are the scaled multipliers. We assume that the firm fulfills the present value by a constant wage per period that grows with time and satisfies the legislated minimum wage constraint on the flow minimum wage. That is, if the wage offered to workers in period $t$ who begin working in period $t+1$ is $w_{ijt+1} \geq \bar{w}_{t+1}$, then in the net period we have $w_{ijt+2} = (1+g)w_{ijt+1} \geq (1+g)\bar{w}_{t+1}$ and so on. This leads to the constraint (A1) in terms of the present value $W_{ijt+1} = d_{t+1}w_{ijt+1}$ where $d_{t+1}$ is a discounter defined by

$$d_{t+1} = 1 + Q_{t+1,t+2}(1-\sigma)(1+g) + Q_{t+1,t+3}(1-\sigma)^2(1+g)^2 + \ldots,$$

which accounts for discounting, separations, and growth. The reason that we specify this constraint in terms of flow wages is that in practice that is how minimum wage legislation works. Our formulation restricts wages inwill minimal ways consistent with the constraint that in each period the flow wage is at least as high as its legislated minimum. Specifically, it prevents firms from offering present values of wages in which in some periods the associated flow wage falls below the legislated minimum.

**First-Order Conditions.** The only part of the firm's problem that is affected by the minimum wage are the equations in the hiring stage. Within the hiring stage, the first-order conditions for employment $n_{ijt+1}$, (A6), and market tightness $\theta_{ijt}$, (A8), are not affected.

The first-order condition for vacancy posting $a_{ijt}$ is now

$$\begin{aligned} & -Q_{0,t}\kappa_{it} - Q_{0,t+1}\lambda_f(\theta_{ijt})W_{ijt+1} + Q_{0,t+1}\lambda_f(\theta_{ijt})\nu_{ijt+1} + Q_{0,t}\xi^a_{ijt} = 0 \\ \Longrightarrow & \frac{\kappa_i}{\lambda_f(\theta_{ijt})} + Q_{t,t+1}W_{ijt+1} \geq Q_{t,t+1}\nu_{ijt+1}, \text{ with equality if } a_{ijt} > 0. \end{aligned} \tag{A3}$$

Here, we explicitly keep track of the multiplier on the nonnegativity constraint on vacancies, since it will never bind without a minimum wage policy but it could bind with one. The only condition that is directly affected is the first-order condition for wages $W_{ijt+1}$, which now is

$$-\lambda_f(\theta_{ijt})a_{ijt} + \gamma_{ijt+1} + \rho_{ijt+1} = 0, \tag{A4}$$

where $\rho_{ijt+1}$ is the multiplier on the minimum wage constraint. There are two cases. First, if the minimum wage is not binding, then this equation is reduces to $\gamma_{ijt+1} = \lambda_f(\theta_{ijt})a_{ijt}$. In this case, plugging this expression back into the first-order condition for market tightness (A8) we get the same equation as when there is no minimum wage, that is, (A10), which repeat here for convenience

$$W_{ijt+1} = \eta\nu_{ijt+1} - (1-\eta)V_{ijt+1}.$$

The interesting case is when the minimum wage is binding. Here we will simply use the first-order condition for market tightness (A8) as an equation that defines the multiplier $\gamma_{ijt+1}$ given that $W_{ijt+1} = \underline{W}_t$. That is, we solve for $\gamma_{ijt+1}$ using the following algebra

$$
\begin{aligned}
W_{ijt+1} &= \nu_{ijt+1} + \frac{\gamma_{ijt+1}}{a_{ijt}} \frac{\mathcal{W}_{it}}{Q_{t,t+1}\lambda_w(\theta_{ijt})^2} \frac{\lambda_w'(\theta_{ijt})}{\lambda_f'(\theta_{ijt})} \\
\underline{W}_{t+1} &= \nu_{ijt+1} + \frac{\gamma_{ijt+1}}{a_{ijt}} \frac{\mathcal{W}_{it}}{Q_{t,t+1}\lambda_w(\theta_{ijt})} \frac{\lambda_w'(\theta_{ijt})}{\lambda_f'(\theta_{ijt})} \frac{1}{\lambda_w(\theta_{ijt})} \\
\underline{W}_{t+1} &= \nu_{ijt+1} + \frac{\gamma_{ijt+1}}{a_{ijt+1}} (\underline{W}_{t+1} + V_{ijt+1}) \frac{\lambda_w'(\theta_{ijt})}{\lambda_f'(\theta_{ijt})} \frac{1}{\lambda_w(\theta_{ijt})} \\
\Longrightarrow \underline{W}_{t+1} &= \nu_{ijt+1} - \gamma_{ijt+1} \frac{1-\eta}{\eta} \frac{\underline{W}_{t+1} + V_{ijt+1}}{\lambda_f(\theta_{ijt}) a_{ijt}} \\
\Longrightarrow \gamma_{ijt+1} &= \frac{\eta}{1-\eta} \frac{\nu_{ijt+1} - \underline{W}_{t+1}}{\underline{W}_{t+1} + V_{ijt+1}} \lambda_f(\theta_{ijt}) a_{ijt}.
\end{aligned}
$$

In the second line, we plugged in $W_{ijt+1} = \underline{W}_{t+1}$. In the third line we used $\mathcal{W}_{it} / \left[Q_{t,t+1}\lambda_w(\theta_{ijt})\right] = \underline{W}_{t+1} + V_{ijt+1}$ and in the fourth line we used $\frac{\lambda_f(\theta_{ijt})}{\lambda_w(\theta_{ijt})} \frac{\lambda_w'(\theta_{ijt})}{\lambda_f'(\theta_{ijt})} = -\frac{1-\eta}{\eta}$. Summarizing

$$
\gamma_{ijt+1} = \left\{ \begin{array}{c} \lambda_f(\theta_{ijt}) a_{ijt} \text{ if slack} \\ \frac{\eta}{1-\eta} \frac{\nu_{ijt+1} - \underline{W}_{t+1}}{\underline{W}_{t+1} + V_{ijt+1}} \lambda_f(\theta_{ijt}) a_{ijt} \text{ if bind} \end{array} \right\} \tag{A5}
$$

$$
W_{ijt+1} = \left\{ \begin{array}{c} \eta \nu_{ijt+1} - (1-\eta) V_{ijt+1} \text{ if slack} \\ \underline{W}_{t+1} \text{ if bind} \end{array} \right\} \tag{A6}
$$

and the sequence of multipliers on the participation constraint, $\gamma_{ij1}, ..., \gamma_{it+1}$ show up in the value of a worker equation

$$
\nu_{ijt+1} = \chi_{ijt+1} + M_{ijt+1} \frac{u_{nit+1}}{u_{cit+1}} \frac{1}{\omega} \left( \frac{n_{ijt+1}}{n_{it+1}} \right)^{\frac{1}{\omega}-1} \frac{1}{n_{it+1}} + Q_{t+1,t+2}(1-\sigma)\nu_{ijt+2}, \tag{A7}
$$

since $M_{ijt+1} = \gamma_{ijt+1} + (1-\sigma)\gamma_{ijt} + ... + (1-\sigma)^t \gamma_{ij1}$. So, in general, the value of $M_{ijt+1}$ depends on the entire binding pattern of the minimum wage.

**Detrending.** Here we state the conditions of the problem in stationary form to anticipate the balanced growth path and we impose symmetry. The only conditions that change are those for when the minimum wage is binding. When it is slack then, as before, $\gamma_{it+1} = \lambda_f(\theta_{it}) a_{it}$ is already stationary. When it binds then in detrended form the multiplier is

$$\gamma_{it+1} = \frac{\eta}{1-\eta} \frac{(1+g)^{t+1}(\widetilde{\nu}_{it+1} - \underline{\widetilde{W}}_{t+1})}{(1+g)^{t+1}(\underline{\widetilde{W}}_{t+1} + \widetilde{V}_{it+1})} \lambda_f(\theta_{it}) a_{it} = \frac{\eta}{1-\eta} \frac{\widetilde{\nu}_{it+1} - \underline{\widetilde{W}}_{t+1}}{(\underline{\widetilde{W}}_{t+1} + \widetilde{V}_{it+1})} \lambda_f(\theta_{it}) a_{it}.$$

So in detrended form (A5) becomes

$$\gamma_{it+1} = \left\{ \begin{array}{c} \lambda_f(\theta_{it}) a_{it} \text{ if slack} \\ \frac{\eta}{1-\eta} \frac{\widetilde{\nu}_{it+1} - \underline{\widetilde{W}}_{t+1}}{\underline{\widetilde{W}}_{t+1} + \widetilde{V}_{it+1}} \lambda_f(\theta_{it}) a_{it} \text{ if bind} \end{array} \right\} \tag{A8}$$

and in detrended form wages are

$$\widetilde{W}_{it+1} = \left\{ \begin{array}{c} \eta \widetilde{\nu}_{it+1} - (1-\eta) \widetilde{V}_{it+1} \text{ if slack} \\ \underline{\widetilde{W}}_{t+1} \text{ if bind} \end{array} \right\}. \tag{A9}$$

Of course, here also since the value of $M_{ijt+1}$ depends on the entire binding pattern of the minimum wage, so does the value of a worker given by

$$\widetilde{\nu}_{it+1} = \widetilde{\chi}_{it+1} - M_{it+1} v'(n_{it+1}) \frac{1}{\omega n_{it+1}} + Q_{t+1,t+2}(1+g)(1-\sigma) \widetilde{\nu}_{it+2}. \tag{A10}$$

### B.1.2 BGP and Proof of Proposition 6

We summarize how the minimum wage impacts the BGP using the following Lemma, analogous to Lemma 8 from Appendix A:

**Lemma 9.** *Along the balanced growth path, the labor allocations and wages are determined by the following equations:*

*(i) optimal cut-off for idiosyncratic productivity of capital*

$$\underline{\varepsilon}_1 = (1+g)(1-\alpha) m(\underline{\varepsilon}_1), \tag{A11}$$

*where $m(\underline{\varepsilon}_1)$ is defined by*

$$m(\underline{\varepsilon}_1) = \frac{\sum_{\tau=1}^{\infty} \widetilde{\beta}^{\tau} (1-\delta)^{\tau-1} \Pi^p((1+g)^{\tau-1} \underline{\varepsilon}_1)}{\sum_{\tau=1}^{\infty} \widetilde{\beta}^{\tau} (1-\delta)^{\tau-1} (1+g)^{\tau} \Pi^u((1+g)^{\tau-1} \underline{\varepsilon}_1)}; \tag{A12}$$

*(ii) the marginal unit of capital earns zero profit*

$$1 = \alpha \left[ \sum_{\tau=1}^{\infty} \widetilde{\beta}^{\tau} (1-\delta)^{\tau-1} \Pi^{p}((1+g)^{\tau-1} \underline{\varepsilon}_1) \right] f(\widetilde{v}); \tag{A13}$$

*(iii) flow wages*

$$\widetilde{w}_i = \left\{ \; \eta[f_i(\widetilde{v}) m(\underline{\varepsilon}_1) - v'(n_i)/\omega] + (1-\eta) v'(n_i) \;\; \textit{if slack}, \underline{w} \;\; \textit{if bind} \; \right\}; \tag{A14}$$

*(iv) optimal vacancy posting*

$$\kappa_i = \left\{ \begin{array}{c} \beta \lambda_f(\theta_i) \frac{f_i(\widetilde{v}) m(\underline{\varepsilon}_1) - \widetilde{w}_i - v'(n_i)/\omega}{1-\beta(1-\sigma)} \;\; \textit{if slack} \\ \beta \lambda_f(\theta_i) \frac{f_i(v) m(\underline{\varepsilon}_1) - \underline{w}}{1-\beta(1-\sigma)} \left[ \frac{\underline{w} - v'(n_i)}{\underline{w} - v'(n_i)(1-1/\omega)} \right] \;\; \textit{if bind} \end{array} \right\}; \tag{A15}$$

*(v) optimal household search*

$$h'(s_i) = \beta \lambda_w(\theta_i) \frac{\widetilde{w}_i - v'(n_i)}{1-\beta(1-\sigma)}; \tag{A16}$$

*(vi) the steady state law of motion for employment*

$$\sigma n_i = \lambda_w(\theta_i) s_i; \tag{A17}$$

*(vii) labor market clearing*

$$\frac{\mu_i n_i}{\widetilde{v}_i} = \frac{\mu_1 n_1}{\widetilde{v}_1}. \tag{A18}$$

*Proof.* As described above, the only two equations change due to the presence of the minimum wage: (i) the expression for the multiplier on the participation constraint (equation (A8) in the detrended system) and (ii) the wage equation (equation (A9) in the detrended system). All the other equations characterizing the equilibrium from the summary of BGP conditions from the baseline model (A48)-(A62) continue to hold. Furthermore, the proof of conditions (i), (ii), (v), (vi), and (vii) from Lemma 8 relied only on those other conditions, so they apply equally here. Therefore, we only need to focus on the part of conditions (iii) and (iv) when the minimum wage binds. Clearly, the wage equation when the minimum wage binds is simply $\widetilde{w}_i = \underline{w}$, giving us the binding part of condition (iii).

The remaining challenge is to prove condition (iv). Recall, that in the case where the minimum wage is not binding, we substituted for the multiplier on the participation constraint $\gamma_{it}$ from (A8) in the expression for the present value of a worker $\widetilde{\nu}_{it+1}$ from (A14) and simplified to arrive at the non-binding version of condition (iv) from Lemma 8. When the minimum wage is binding, we must follow a different strategy because the multiplier on the participation constraint $\gamma_{it}$ in (A8) itself depends on the present value of a worker $\widetilde{\nu}_{it+1}$, and the present value of a worker $\widetilde{\nu}_{it+1}$ from (A10) implicitly depends on the value of $\gamma_{it}$ through the auxiliary variable $M_{it+1} = (1-\sigma)M_{it} + \gamma_{it+1}$. Note that imposing balanced growth on (A10) gives

$$\widetilde{\nu}_i = \frac{\widetilde{\chi}_i - M_i v'(n_i)/(\omega n_i)}{1-\beta(1-\sigma)}. \tag{A19}$$

Next, note that the BGP version of the multiplier on the participation constraint (A8) becomes

$$\gamma_i = \frac{\eta}{1-\eta}\frac{\widetilde{\nu}_i - \underline{W}}{\underline{W} - \widetilde{V}_i}\lambda_f(\theta_i)a_i \implies \frac{\gamma_i}{\sigma}\frac{1}{n_i} = \frac{\widetilde{\nu}_i - \underline{W}}{\underline{W} + \widetilde{V}_i}, \tag{A20}$$

where the second line uses the BGP law of motion for employment $\lambda_f(\theta_i)a_i = \sigma n_i$. Next, define the numerator of (A19) as the *flow value of a worker* to the firm along the BGP by letting $\widehat{\nu}_i = \widetilde{\chi}_i - M_i v'(n_i)\frac{1}{\omega n_i}$. We then convert all the terms on the right side of (A20) by dividing both the numerator and denominator by $1-\beta(1-\sigma)$ to get

$$\frac{\gamma_i}{\sigma}\frac{1}{n_i} = \frac{\widehat{\nu}_i - \underline{w}}{\underline{w} - v'(n_i)}. \tag{A21}$$

Next, we will plug this expression into the flow value of a worker $\widehat{\nu}_i = \widetilde{\chi}_i - M_i v'(n_i)\frac{1}{\omega n_i}$. Note that the BGP version of the evolution of $M_{it+1} = (1-\sigma)M_{it} + \gamma_{it+1}$ is $M_i = \frac{\gamma_i}{\sigma}$, so we can write this flow value as

$$\widehat{\nu}_i = \widetilde{\chi}_i - \frac{\gamma_i}{\sigma}v'(n_i)\frac{1}{\omega n_i} \implies \widehat{\nu}_i = \widetilde{\chi}_i - \frac{\gamma_i}{\sigma n_i}\frac{v'(n_i)}{\omega}.$$

Now plug in the expression for $\frac{\gamma_i}{\sigma n_i}$ from (A21) into this equation to get

$$\widehat{\nu}_i = \widetilde{\chi}_i - \frac{v'(n_i)}{\omega}\frac{\widehat{\nu}_i - \underline{w}}{\underline{w} - v'(n_i)}. \tag{A22}$$

We will use this implicit expression for $\widehat{\nu}_i$ to obtain the expression for optimal vacancy posting (A15).

Now subtract the minimum wage $\underline{w}$ and solve for $\widehat{\nu}_i - \underline{w}$ to get

$$\widehat{\nu}_i - \underline{w} = \widetilde{\chi}_i - \underline{w} - \frac{\widehat{\nu}_i - \underline{w}}{\underline{w} - v'(n_i)} \frac{v'(n_i)}{\omega} \implies (\widehat{\nu}_i - \underline{w}) \left[1 + \frac{1}{\underline{w} - v'(n_i)} \frac{v'(n_i)}{\omega}\right] = \widetilde{\chi}_i - \underline{w}$$
$$\implies (\widehat{\nu}_i - \underline{w}) \left[\frac{\underline{w} - v'(n_i) + v'(n_i)/\omega}{\underline{w} - v'(n_i)}\right] = \widetilde{\chi}_i - \underline{w}$$
$$\implies \widehat{\nu}_i - \underline{w} = \left[\frac{\underline{w} - v'(n_i)}{\underline{w} - v'(n_i)(1 - 1/\omega)}\right] (\widetilde{\chi}_i - \underline{w}) \,. \tag{A23}$$

Finally, plug this into the detrended vacancy-posting condition given by

$$\kappa_i = \frac{\beta}{1 - \beta(1 - \sigma)} \lambda_f(\theta_i) \left(\widehat{\nu}_i - \underline{w}\right).$$

to get

$$\kappa_i = \frac{\beta}{1 - \beta(1 - \sigma)} \lambda_f(\theta_i) \left[\frac{\underline{w} - v'(n_i)}{\underline{w} - v'(n_i)(1 - 1/\omega)}\right] (\widetilde{\chi}_i - \underline{w}) \,. \tag{A24}$$

Finally, note that the expression $\widetilde{\chi}_i = f_i(v) m(\underline{\varepsilon}_1)$ continues to be true from the proof of Lemma 3. Using this we obtain

$$\kappa_i = \frac{\beta}{1 - \beta(1 - \sigma)} \lambda_f(\theta_i) \left[\frac{\underline{w} - v'(n_i)}{\underline{w} - v'(n_i)(1 - 1/\omega)}\right] \left[f_i(v) m(\underline{\varepsilon}_1) - \underline{w}\right]. \tag{A25}$$

which completes the proof. □

**Proof of Proposition 6.** We will now build on this characterization of the BGP under the minimum wage to prove Proposition 6 from the main text. To do so, we must first extend the space of policies to incorporate the two policies from Proposition 6. First, a type-specific minimum wage can be represented as the detrended flow minimum wage, $\underline{w}_i$, specific for a type $i$ worker. Since the policy will set the minimum wage to its competitive level, which is strictly above the monopsonistically competitive equilibrium value, the minimum wage will be binding for each type of worker. Second, a subsidy to vacancy-posting can be represented by replacing the detrended vacancy-posting cost $\kappa_i$ with its after-subsidy version $\kappa_i(1 - \tau_i)$. With these two changes, the system of equations characterizing the allocation under the policies is

$$\underline{\varepsilon}_1 = (1+g)(1-\alpha)m(\underline{\varepsilon}_1) \tag{A26}$$

$$1 = \alpha\left[\sum_{\tau=1}^{\infty}\widetilde{\beta}^{\tau}(1-\delta)^{\tau-1}\Pi^{p}((1+g)^{\tau-1}\underline{\varepsilon}_1)\right]f(\widetilde{v}) \tag{A27}$$

$$\kappa_i(1-\tau_i) = \lambda_f(\theta_i)\frac{f_i(v)m(\underline{\varepsilon}_1) - \underline{w}_i}{\rho+\sigma}\left[\frac{\underline{w}_i - v'(n_i)}{\underline{w}_i - v'(n_i)(1-1/\omega)}\right] \tag{A28}$$

$$h'(s_i) = \beta\lambda_w(\theta_i)\frac{\underline{w}_i - v'(n_i)}{1-\beta(1-\sigma)} \tag{A29}$$

$$\sigma n_i = \lambda_w(\theta_i)s_i \tag{A30}$$

$$\frac{\mu_i n_i}{\widetilde{v}_i} = \frac{\mu_1 n_1}{\widetilde{v}_1}. \tag{A31}$$

The allocation in the competitive model can be obtained by evaluating Lemma 3 at $\omega = \infty$:

$$\underline{\varepsilon}_1 = (1+g)(1-\alpha)m(\underline{\varepsilon}_1) \tag{A32}$$

$$1 = \alpha\left[\sum_{\tau=1}^{\infty}\widetilde{\beta}^{\tau}(1-\delta)^{\tau-1}\Pi^{p}((1+g)^{\tau-1}\underline{\varepsilon}_1)\right]f(\widetilde{v}^c) \tag{A33}$$

$$\kappa_i = \beta\lambda_f(\theta_i^c)\frac{f_i(\widetilde{v}^c)m(\underline{\varepsilon}_1) - \widetilde{w}_i^c}{1-\beta(1-\sigma)} \tag{A34}$$

$$h'(s_i^c) = \beta\lambda_w(\theta_i^c)\frac{\widetilde{w}_i^c - v'(n_i^c)}{1-\beta(1-\sigma)} \tag{A35}$$

$$\sigma n_i^c = \lambda_w(\theta_i^c)s_i^c \tag{A36}$$

$$\frac{\mu_i n_i^c}{\widetilde{v}_i^c} = \frac{\mu_1 n_1^c}{\widetilde{v}_1^c} \tag{A37}$$

$$\widetilde{w}_i^c = \eta f_i(\widetilde{v}^c)m(\underline{\varepsilon}_1) + (1-\eta)v'(n_i^c). \tag{A38}$$

We will show that the competitive allocation from equations also solves the equilibrium allocation under the policy choices from Proposition 6. To do so, we evaluate the system of equations characterizing the equilibrium under the policies, (A26)—(A31), at the competitive allocation (i.e. with $\widetilde{v} = \widetilde{v}^c$, $\widetilde{n}_i = \widetilde{n}_i^c$, and so on). We will show that doing so gives the same system of equations as the competitive system, (A32)—(A38).

Given this guessed allocation, the following equations are clearly the same because the wage does not enter them directly: (i) the productivity cutoffs (A26) and (A32), (ii) the zero profit condition for the marginal unit of capital (A27) and (A34), (iii) the law of of motion for employment (A30), and (iv) the labor ratios (A31) and (A37). The household's optimal search condition depends on the wage; since the policy sets the type-specific minimum wage $\underline{w}_i = \widetilde{w}_i^c$, these two conditions (A29) and (A35) also coincide. Finally, under the choice for $\tau_i$ from the main text, the optimal vacancy-posting conditions (A28) and (A34) also coincide.

Note that dividing the monopsony vacancy-posting condition (A28) evaluated at the competitive allocation with the minimum wage policy $\underline{w}_i = \widetilde{w}_i^c$ and the $1 - \tau_i$ in the proposition gives that

$$1 - \tau_i = \left[ \frac{\widetilde{w}_i c - v'(n_i^c)}{\widetilde{w}_i^c - v'(n_i^c)(1 - 1/\omega)} \right],$$

which is exactly how these subsidies are set. This establishes the result.

### B.1.3 Initial Conditions

In all of our experiments, we assume that the economy is initially growing along the BGP without policies, characterized in Appendix A, and then the policy is unexpectedly introduced in the initial period $t = 0$. Up to this point, we have largely ignored the initial conditions faced by firms in this initial period $t = 0$ and focused on the behavior of the economy from period $t \geq 1$ onward. In this subsection, we specify the initial conditions and whether firms fire workers in the initial period $t = 0$. The firm takes as given four sets of initial conditions drawn from the initial BGP when solving its problem in period $t = 0$. First, the firm inherits a distribution of capital stocks $K_{j0}(v_{-\tau}, A_{-\tau})$, where $v_{-\tau}$ is the vector of labor intensities chosen along the initial BGP in periods $-1, -2, \ldots$ and $A_{-\tau}$ is the corresponding level of vintage productivity. Second, the firm inherits a measure of employed workers of each type $i$, $N_{ij0} = N_{i0}$, equal to the employment rate $N_i$ of each group from the BGP. Third, the firm inherits the flow wage schedule initially promised to each of these worker types along the BGP. Given that flow wages grow at a constant rate within a match, this flow wage schedule is summarized by $w_{i0}$, the flow wage promised to workers of group $i$ in period $t = 0$. Under the minimum wage, firms must now pay these workers $\widehat{w}_{i0} = \max\{w_{i0}, \underline{w}\}$ in period $t = 0$, $\widehat{w}_{i1} = \max\{(1 + g)w_{i0}, (1 + g)\underline{w}\}$ in the following period $t = 1$, and so on. Let $\widehat{W}_{i0} = \sum_{t=0}^{\infty} Q_{0,t}(1 - \sigma)^t \widehat{w}_{it}$ denote the present value of wage payments promised to these workers going forward. Finally, the firm inherits the Marcet-Marimon cumulation of multipliers $M_{i0}$ from the initial BGP to reflect promises made to workers hired before period $t = 0$.

We assume that when the minimum wage is unexpectedly introduced in period $t = 0$, a firm $j$ can choose to fire a measure $F_{ij0}$ of its initially employed workers. However, for all workers that it does not fire, the firm must pay them at least the flow minimum wage each period.

**Initial Period Decisions.** The majority of the firm's problem is identical to the what we have already studied except for decisions in the initial period $t = 0$. Furthermore, nearly all decisions about hiring and investment made in this period only impact the firm's objective starting in period $t \geq 1$ onward, so those are unchanged. The only exception is the option for the firm to fire $F_{ij0}$ workers in the initial period. The option to fire workers affects the profit maximization problem in four ways. First, for each fired worker, the firm saves itself the present value of flow wages it would have been obliged to fire that worker had they remained employed. Hence, the term $\sum_i \widehat{W}_{i0} F_{ij0}$ is added to the firm's objective function. Second, the adding up constraint in the assignment of workers to machines must reflect the fact that the firm may fire some of the existing workers:

$$\sum_{\tau=1}^{\infty} v_{i,-\tau} u_{j0}(v_{-\tau}, \varepsilon, A_{-\tau}) K_{j0}(v_{-\tau}, A_{-\tau}) \pi(\varepsilon) d\varepsilon dv \leq N_{ij0} - F_{ij0} \quad (\times \chi_{ij0}),$$

where $\chi_{ij0}$ is the multiplier on this constraint. Note that, since this constraint holds with equality along the initial BGP, positive firing $F_{ij0} > 0$ requires lowering the utilization rates of existing capital. Third, we must modify the law of motion for employment to account for firings as well:

$$N_{ij1} \leq (1 - \sigma)(N_{ij0} - F_{ij0}) + \lambda_f(\theta_{ij0}) \mu_i a_{ij0} \quad (\times Q_{0,1} \nu_{ij1}),$$

where $\nu_{ij1}$ is the scaled multiplier on this constraint. Finally, firms must satisfy the non-negativity constraint $F_{ij0} \geq 0$ for $(\times \xi^f_{ij0})$.

**First-Order Condition.** The first-order condition with respect to $F_{ij0}$ is $\widehat{W}_{i0} - \chi_{ij0} - Q_{0,1}(1 - \sigma)\nu_{ij1} + \xi^f_{ij0} = 0$ or, equivalently,

$$\widehat{W}_{i0} - \chi_{ij0} - Q_{0,1}(1 - \sigma)\nu_{ij1} + \xi^f_{ij0} = 0 \implies \chi_{ij0} + Q_{0,1}(1 - \sigma)\nu_{ij1} \geq \widehat{W}_{i0}, \text{ with equality if } F_{ij0} > 0. \tag{A39}$$

That is, firms do not fire workers if the present value of the workers' benefits to the firm — their marginal product in period $t = 0$, $\chi_{ij0}$ plus their present value going forward, $Q_{0,1}(1 - \sigma)\nu_{ij1}$ on the LHS of (A39) — is strictly greater than the present value of wage payments to those workers — $\widehat{W}_{i0}$ on the RHS of (A39). This is the only new condition for the initial period $t = 0$; all other conditions are the same as in the baseline model.

### B.2 Transfer Programs

We now turn to the transfer programs. Section B.2.1 shows how transfers impact the household's problem. Section B.2.2 shows how transfers impact the firms problem. Section B.2.3 shows how to detrend those conditions and arrives at the BGP.

#### B.2.1 Households

We first provide some additional notation related to the transfer system, and then show how it affects the solution to the household's problem.

**Notation.** As in the main text, we will represent the transfer system in terms of the after-transfers wages that households receive. In particular, if the firm pays the flow wage $w_{ijt}$, then households receive the flow payment $A_t(w_{ijt})$ which includes the transfers from the government. Also as in the main text, we assume that the transfer system satisfies the property $A_t(w_{ijt}) = (1+g)^t A(\widetilde{w}_{ijt})$ for some time-invariant function $A(\widetilde{w})$ where, as usual, tildes denote detrended variables. Note that this assumption also implies that $A'_t(w) = A'(\widetilde{w})$.[21] We use the discounter operator $d_{t+1} = 1 + Q_{t+1,t+2}(1+g)(1-\sigma) + Q_{t+1,t+3}(1+g)^2(1-\sigma)^2 + \dots$ to convert this stream of flow payments to the worker to the present value $W^H_{ijt+1} = d_{t+1}A_t(w_{ijt+1})$. From the firm's perspective, the present value of wage costs is the same as in the baseline model $W_{ijt+1} = d_{t+1}w_{ijt+1}$.

**Household's Problem.** The transfer program changes two parts of the household's problem relative to our baseline model. First, the wages received in the budget constraint are $W^H_{ijt}$ from above, to reflect the present value of transfer payments. Second, the present value of profits need to reflect the corporate taxes to fund the program.

More formally, the household's utility maximization problem is now

$$\max_{c_{it}, s_{ijt}, n_{ijt+1}} \sum_{t=0}^{\infty} \beta^t U_t(c_{it}, n_{it}, s_{it}) \quad \text{such that}$$

$$n_{ijt+1} = (1-\sigma)n_{ijt} + \lambda_w(\theta_{ijt})s_{ijt} \quad (\times \beta^t \widehat{V}_{ijt+1})$$

$$\sum_{t=0}^{\infty} Q_{0,t}c_{it} = \psi_i(1-\tau_c)\mathbb{P} + \mathbb{I}_i + \sum_{t=1}^{\infty} Q_{0,t} \sum_j \lambda_w(\theta_{ijt-1})s_{ijt-1}W^H_{ijt} \quad (\times\Gamma),$$

[21] To see this, note that

$$A'_t(w) = \frac{d}{dw}A_t(w) = \frac{d}{dw}(1+g)^t A(\widetilde{w}_{ijt}) = (1+g)^t \frac{d}{dw}A(\frac{w}{(1+g)^t}) = \frac{(1+g)^t}{(1+g)^t}A'(\widetilde{w}) = A'(\widetilde{w}).$$

where $\tau_c$ is the profits tax rate. As usual, the variables in parentheses denote the (often rescaled) Lagrange multiplier associated with the constraint. Compared to the household problem in the baseline model from Appendix A, the only first-order condition which changes is the one for search effort to take into account that wage payments are now $W^H_{ijt+1}$:

$$-\frac{u_{sit}}{u_{cit}} = \lambda_w(\theta_{ijt})Q_{t,t+1}\left(V_{ijt+1} + W^H_{ijt+1}\right). \tag{A40}$$

The fact that the optimal search condition (A40) also changes the participation constraint which firms will take as given:

$$\lambda_w(\theta_{ijt})(W^H_{ijt+1} + V_{ijt+1}) \geq \mathcal{W}_{it}. \tag{A41}$$

### B.2.2 Firms

We now turn to how the transfer program affects the solution to the firm's problem. We first restate the profit maximization problem and then derive the FOCs. As with our minimum wage analysis, we abstract from initial conditions in this section and focus on the conditions that change due to the presence of the transfer program.

**Firm's Problem.** The presence of the transfer system changes the firm's problem in three ways. First, the participation constraint (A41) now reflects the fact that households receive transfers, as derived above. Second, and related, we re-state the firm's wage choice in terms of the initial flow wage $w_{ijt+1}$ instead of the present value. Using $w_{ijt+1}$ as the choice variables allows us to capture how firms' wage-posting decisions affect both the present value of households' post-transfer income and firms' pre-transfer costs. Third, the profits tax multiplies flow profits each period by $1 - \tau_c$. However, since the tax rate is constant, this change amounts to multiplying the objective function by $1 - \tau_c$, which doesn't affect the profit-maximizing decisions. We therefore drop the $1 - \tau_c$ from the exposition to keep the equations as close to our baseline model as possible. We abstract from the possibility that firms will want to fire initial workers given that it will not be relevant for this policy. With these changes, the profit maximization problem is

$$\sum_t Q_{0,t}\Bigg(\sum_\tau \int_{v,\varepsilon} u_{jt}(v,A_{t-\tau},\varepsilon)A_{t-\tau}\varepsilon f(v)K_{jt}(v,A_{t-\tau})\pi(\varepsilon)d\varepsilon dv - \sum_i \mu_i(\lambda_f(\theta_{ijt-1})a_{ijt-1}d_t w_{ijt} + \kappa_{it}a_{ijt})$$

$$-\int X_{jt}(v)dv\Bigg) + \sum_{t=0}^{\infty} Q_{0,t+1}\mu_i M_{ijt+1}\frac{U_{nit+1}}{U_{cit+1}}\left(\frac{n_{ijt+1}}{n_{it+1}}\right)^{\frac{1}{\omega}} + \sum_{t=0}^{\infty} Q_{0,t+1}\mu_i\gamma_{ijt+1}\left[d_{t+1}A_{t+1}(w_{ijt+1}) - \frac{\mathcal{W}_{it}}{Q_{t,t+1}\lambda_w(\theta_{ijt})}\right]$$

such that $\quad u_{jt}(v,A_{t-\tau},\varepsilon)\geq 0 \quad (\times Q_{0,t}\lambda^L_{jt}(v,A_{t-\tau},\varepsilon))$

$u_{jt}(v,A_{t-\tau},\varepsilon)\leq 1 \quad (\times Q_{0,t}\lambda^U_{jt}(v,\varepsilon,A_{t-\tau}))$

$u_{jt}(v,A_{t-\tau},\varepsilon)v_iK_{jt}(v,A_{t-\tau})\pi(\varepsilon)\leq N_{ijt}(v,A_{t-\tau},\varepsilon)$ for all $i \quad (\times Q_{0,t}\lambda_{ijt}(v,A_{t-\tau},\varepsilon))$

$\sum_\tau\int_{v,\varepsilon} N_{ijt}(v,A_{t-\tau},\varepsilon)d\varepsilon dv \leq \mu_i n_{ijt}$ for all $i \quad (\times Q_{0,t}\chi_{ijt})$

$\mu_i n_{ijt+1}\leq (1-\sigma)\mu_i n_{ijt} + \lambda_f(\theta_{ijt})\mu_i a_{ijt}$ for all $i \quad (\times Q_{0,t+1}\nu_{ijt+1})$

$K_{jt+\tau+1}(v,A_t) = (1-\delta)^\tau X_{jt}(v) \quad (\times Q_{0,t+\tau+1}q_{jt,t+\tau+1}(v))$

$X_{jt}(v)\geq 0 \quad (\times Q_{0,t}\mu_{jt}(v)).$

**First-Order Conditions.** As with the minimum wage, the only part of the firm's problem that is affected are the equations in the hiring stage. Within the hiring stage, the first-order conditions for employment $n_{ijt+1}$, vacancies $a_{ijt}$, and market tightness $\theta_{ijt}$ are unaffected. However, the transfer system will change how we simplify the first-order condition for market tightness (A8), so we reproduce it here:

$$W_{ijt+1} = \nu_{ijt+1} + \frac{\gamma_{ijt+1}}{a_{ijt}}\frac{\mathcal{W}_{it}}{\lambda_w(\theta_{ijt})^2}\frac{\lambda_w'(\theta_{ijt})}{\lambda_f'(\theta_{ijt})}. \tag{A42}$$

The first-order condition which changes is the one for wages $w_{ijt+1}$:

$$-Q_{0,t+1}\mu_i\lambda_f(\theta_{ijt})a_{ijt}d_{t+1} + Q_{0,t+1}\mu_i\gamma_{ijt+1}d_{t+1}A_t'(w_{ijt+1}) = 0 \implies \quad \gamma_{ijt+1} = \frac{\lambda_f(\theta_{ijt})a_{ijt}}{A_t'(w_{ijt+1})}. \tag{A43}$$

Hence, as stated in the main text, the the transfer system changes the multiplier on the participation constraint. This multiplier then affects two things. First, it enters the law of motion for the auxiliary variable $M_{ijt+1} = (1-\sigma)M_{ijt} + \gamma_{ijt+1}$. Second, it affects how we simplify the FOC for market tightness (A42). Plugging the expression for the multiplier (A43) into the FOC for market tightness

(A42) gives

$$
\begin{aligned}
W_{ijt+1} &= \nu_{ijt+1} + \frac{\lambda_f(\theta_{ijt})}{\lambda_w(\theta_{ijt})} \frac{\mathcal{W}_{it}}{\lambda_w(\theta_{ijt})} \frac{\lambda'_w(\theta_{ijt})}{\lambda'_f(\theta_{ijt})} \frac{1}{A'_t(w_{ijt+1})} \\
&= \nu_{ijt+1} + \frac{\lambda_f(\theta_{ijt})}{\lambda_w(\theta_{ijt})} \left(W^H_{ijt+1} + V_{ijt+1}\right) \frac{\lambda'_w(\theta_{ijt})}{\lambda'_f(\theta_{ijt})} \frac{1}{A'_t(w_{ijt+1})} \\
&= \nu_{ijt+1} - \frac{1-\eta}{\eta} \frac{W^H_{ijt+1} + V_{ijt+1}}{A'_t(w_{ijt+1})}. \qquad \text{(A44)}
\end{aligned}
$$

where in the second line we used $\frac{\mathcal{W}_{it}}{\lambda_w(\theta_{ijt})} = W^H_{ijt+1} + V_{ijt+1}$ and in the third line we used $\frac{\lambda_f(\theta_{ijt})}{\lambda_w(\theta_{ijt})} \frac{\lambda'_w(\theta_{ijt})}{\lambda'_f(\theta_{ijt})} = -\frac{1-\eta}{\eta}$. In the baseline model, we were able to further simplify (A44) in order to get a closed-form expression for the present value of wage payments $W_{ijt+1}$. However, we're not able to do so in this case due to the transfer system.

### B.2.3 Detrending and BGP

The only equilibrium conditions which have changed relative to the baseline are the search FOC (A40) and the wage equation (A44). We therefore focus our discussion of detrending on those conditions. For the search FOC (A40), first note that

$$
W^H_{it+1} = d_{t+1} A_t(w_{ijt+1}) = d_{t+1}(1+g)^t A(\widetilde{w}_{ijt+1}) \implies \widetilde{W}^H_{ijt+1} = d_{t+1} A(\widetilde{w}_{ijt+1}),
$$

where the second equation uses our assumption that $A_t(w_{ijt+1}) = (1+g)^t A(\widetilde{w}_{ijt+1})$. Plugging this into the search FOC gives

$$
\begin{aligned}
(1+g)^t h'(s_{it}) &= Q_{t,t+1} \lambda_w(\theta_{it}) (1+g)^{t+1} \left(\widetilde{W}^H_{it+1} + \widetilde{V}_{it+1}\right) \\
\implies h'(s_{it}) &= Q_{t,t+1}(1+g) \lambda_w(\theta_{it}) \left(\widetilde{W}^H_{it+1} + \widetilde{V}_{it+1}\right). \qquad \text{(A45)}
\end{aligned}
$$

To detrend the wage equation (A44), recall the property that $A'_t(w_{ijt+1}) = A'(\widetilde{w}_{ijt+1})$. Therefore,

$$
\widetilde{W}_{ijt+1} = \widetilde{\nu}_{ijt+1} - \frac{1-\eta}{\eta} \frac{\widetilde{W}^H_{ijt+1} + \widetilde{V}_{ijt+1}}{A'(\widetilde{w}_{ijt+1})} \text{ and } \gamma_{ijt+1} = \frac{\lambda_f(\theta_{ijt}) a_{ijt}}{A'(\widetilde{w}_{ijt+1})}. \qquad \text{(A46)}
$$

# C Data Appendix

This appendix contains details about our data sources and targeted moments. We use data from the pooled 2017-2019 American Community Survey (ACS).[22] Our sample includes all individuals

[22] We downloaded the data directly from `https://usa.ipums.org/usa/`.

aged 16 and over. All observations are weighted using the weights provided by the ACS.

**Share of College Workers.** We define *college* individuals as those individuals who report having a bachelor's degree or higher. During the 2017-2019 period, 31.3% of our sample had at least a bachelor's degree.

**Employment Rates.** We focus on full-time employment and on workers strongly attached to the labor force. We define individuals as being *full-time* employed if 1) they are currently working at least 30 hours per week; 2) they reported working at least 29 weeks during the prior year; and 3) they reported positive labor earnings during the prior 12 month period. For our 2017-2019 sample, 46.8% of non-college individuals and 62.4% of college individuals worked full-time.

**Share of Income Earned by College Workers.** For the 2017-2019 period, 37.8% of individuals working full-time were college educated. Conditional on being full-time employed, mean annual earnings for college individuals total $91,706, whereas mean annual earnings for non-college individuals total $44,871. Given these statistics, we calculate that 55.5% of all earnings of full-time workers accrued to workers with at least a bachelor's degree.

Table 1: Average Wages by Education Group in ACS Data

| | Less than High School | High School | Some College | *College* |
|---|---|---|---|---|
| [0.5ex] Average wage | $16.6 | $19.6 | $21 | $37.4 |

*Notes:* Average wages of full-time workers by education group in ACS data.

**Wage Distributions.** We compute hourly wages for our sample of full-time workers by dividing annual labor earning by annual hours worked. We calculate annual hours worked as the product of weeks worked last year and reported usual hours worked. We impose two additional sample restrictions when measuring the wage distribution. First, we restrict the sample to only those workers who report at least $5,000 of labor earnings during the prior year. Second, we truncate the resulting distribution of hourly wages of each education group at the top and bottom 1%. All wages are converted to 2019 dollars using the June CPI-U. From these data, we compute the median wage and standard deviation of wages for each education group as well as the ratios of wages between the 10th percentile and the median for each of the education groups. These moments are used as part of our parameterization strategy. We also show that even though only those moments are targeted for each education group, our model matches the full distribution of wages for each education group quite closely. The heterogeneity of wages within education groups swamps the heterogeneity across education groups, motivating our choice to primarily focus on within-group heterogeneity. Related to this choice, Table 1 shows that the average wage of *each* education group is higher than \$15 per hour. Hence, modeling within-group heterogeneity is necessary for even a high minimum wage to be binding for any worker.

# D Validating Use of Long-Run Elasticity Estimates

We now provide the details about how we replicate Card and Lemieux (2001)'s estimation strategy in our model. As explained in the main text, Card and Lemieux (2001) exploit within-education-group variation in employment rates by age, which they identify as a skill level, $z_i$. We replicate this variation through exogenous changes in the measure of families, $\mu_{it}$. Section D.1 shows how we extend our model to incorporate time-varying measures of families. Section D.2 then explains how we choose the specific path of $\mu_{it}$ to mimic the empirical variation utilized by Card and Lemieux (2001). Finally, Section D.3 shows the results of this exercise and explains why this procedure recovers a value for the long-run elasticity of substitution among workers very close to that estimated by Card and Lemieux (2001).

## D.1 Model Extension

We model changes in the measure of families, $\mu_{it}$, as a one time unanticipated shock after which agents in the model have perfect foresight. Specifically, the economy starts at an initial BGP with measures $\mu_i$, and at time $t = 0$ agents learn about a new path of measures $\{\mu_{it}\}$ over time that converge to a new constant level $\mu_i^*$ at some point $T$ in the future.

To proceed, we must first specify the objective function of a type-$i$ family now that the measure of the family changes over time. We assume the family maximizes

$$\sum_{t=0}^{\infty} \beta^t \mu_{it} U_t(c_{it}, n_{it}, s_{it}),$$

where $c_{it}$, $n_{it}$, and $s_{it}$ are per-capita variables. This "utilitarian" utility function captures the average utility of each household member and weighs each member equally. In our numerical experiments, the path of $\mu_{it}$ is such that there are new members of the family available to work in each period. We assume that each of these new family members' initial labor market state is unemployment. Hence, they must search for a period before they can be hired by a firm. Therefore, the law of motion for employment of family $i$ at firm $j$ is

$$\mu_{it+1} n_{ijt+1} = (1-\sigma)\mu_{it} n_{ijt} + \lambda_w(\theta_{ijt})\mu_{it} s_{ijt},$$

where $n_{ijt}$ and $s_{ijt}$ correspond to per-capita variables as in the baseline model. In our baseline, $\mu_{it+1} = \mu_{it}$ so this equation reduces to $n_{ijt+1} = (1-\sigma)n_{ijt} + \lambda_w(\theta_{ijt})s_{ijt}$.

In per-capita terms, the budget constraint is now

$$\sum_{t=0}^{\infty} Q_{0,t}\mu_{it} c_{it} \leq \mu_{i0}\left(\zeta_i \mathbb{P}_0 + \mathbb{I}_i\right) + \sum_{t=1}^{\infty} Q_{0,t}\mu_{it-1}\sum_j \lambda_w(\theta_{ijt-1})s_{ijt-1}W_{ijt},$$

where, as before, $\zeta_i \mathbb{P}_0$ is the per-capita share of family $i$ in the present value of firm's profits and $\mathbb{I}_i$ is the per-capita present value of wages promised to initial workers. Putting all this together, the household's utility maximization problem is

$$\begin{aligned}
&\max_{c_{it}, s_{ijt}, n_{ijt+1}} \sum_{t=0}^{\infty} \beta^t \mu_{it} U_t(c_{it}, n_{it}, s_{it}) \\
&\text{s.t. } \mu_{it+1} n_{ijt+1} = (1-\sigma)\mu_{it} n_{ijt} + \lambda_w(\theta_{ijt})\mu_{it} s_{ijt} \\
&\quad \sum_{t=0}^{\infty} Q_{0,t}\mu_{it} c_{it} \leq \mu_{i0}\left(\zeta_i \mathbb{P}_0 + \mathbb{I}_i\right) + \sum_{t=1}^{\infty} Q_{0,t}\mu_{it-1}\sum_j \lambda_w(\theta_{ijt-1})s_{ijt-1}W_{ijt}
\end{aligned}$$

It turns out that this utility maximization problem leads to equilibrium conditions that naturally extend those of our baseline model, with the constant measure of each family $\mu_i$ from the baseline replaced by the time-varying measure $\mu_{it}$. Results are available upon request.

### D.2 Mimicking the Variation in Card and Lemieux (2001)

Card and Lemieux (2001) estimate the elasticity of substitution $\phi$ using residual variation in employment rates across different groups of workers within a given education group. A key assumption is that this residual variation reflects changes in labor supply. In this spirit, we mimic their variation by assuming that the measure of each family $i$, $\mu_i$, changes over time as denoted by $\mu_{it}$ in a way consistent with their data. Card and Lemieux (2001) identify $z_i$ by assuming that, within an education group, all workers within a 5-year age group share the same $z_i$. In their published paper, Card and Lemieux (2001) report the time series of the ratios of college to non-college employment rates $\frac{N_{jHt}}{N_{jLt}}$ for each age group $j$, but not the employment rates of each group $N_{jHt}$ and $N_{jLt}$ separately. We therefore proceed in two steps. First, we use additional assumptions to infer the time-series variation in $N_{jHt}$ and $N_{jLt}$ from what Card and Lemieux (2001) report about $\frac{N_{jHt}}{N_{jLt}}$

in the data. Second, we then back out the variation in the measures of families $\mu_{it}$ which replicates this variation in employment rates within our model.

**Mimicking the Employment Rate Variation in Card and Lemieux (2001).** Let $x_{jt} = \frac{N_{jHt}}{N_{jLt}}$ denote the ratio of college to non-college employment for age group $j$ reported by Card and Lemieux (2001).[23] Card and Lemieux (2001) classify workers into $N = 6$ different age groups, which we assume correspond to different within-education group skill levels $z_{e,j}$ in our model so that the youngest group corresponds to $z_{e,1}$, the next youngest group corresponds to to $z_{e,2}$, and so on. The sample used by Card and Lemieux (2001) covers the period between 1960 and 1995. The spirit of our modeling exercise will be that the measures of families change over this 35-year sample, at which point they remain constant and the economy settles into a new BGP. We will further assume that this new BGP corresponds to the calibrated steady state of our model. Hence, we will choose the time path of the measures of families $\mu_{it}$ such that the percentage change in the model's employment series relative to the final BGP corresponds to Card and Lemieux (2001)'s variation relative to the end of their sample.[24]

We construct the percentage changes in the employment series $N_{jLt}$ and $N_{jHt}$ in the following way. First, we construct a time series of our targeted ratios $x_{jt} = N_{jHt}/N_{jLt}$ relative to their 1995 endpoint in Card and Lemieux (2001)'s sample. Second, in order to separately construct the levels of $N_{jHt}$ and $N_{jLt}$, we will bring in additional information about the total employment of all workers in age group $j$, denoted $a_{jt} \equiv N_{jLt} + N_{jHt}$. Given a value of $a_{jt}$ and the ratio $x_{jt}$, we can solve for $N_{jLt} = a_{jt}/(1 + x_{jt})$ and $N_{jHt} = x_{jt}a_{jt}/(1 + x_{jt})$. To compute the time series of $a_{jt}$ for each age group $j$, we assume that the terminal value $a_{jT}$ equals its value in our calibrated BGP, compute an initial value $a_{j0}$ from the 1960 Census data described in Appendix C, and assume that $a_{jt}$ grows at a constant rate over the sample.

**Backing Out Variation in Measures of Families.** Equipped with the resulting series of employment rates $N_{it}$ for each worker type $i = (e, z)$, our second step consists of determining the measure of families $\mu_{it}$ in our model such that the model's equilibrium employment series $N_{it} = \mu_{it}n_{it}$ matches the data (recall that $n_{it}$ is the per-capita employment rate of family of type

[23]In particular, Figure IV in Card and Lemieux (2001) reports cohort fixed effects, which capture 98% of the time-series variation in the employment ratios $x_{jt}$—intuitively, because college attainment decisions are made before workers enter the labor force. We back out the time-series variation in $x_{jt}$ from these cohort effects. Since Card and Lemieux (2001) effectively report $x_{jt}$ for 5-year intervals, we log-linearly interpolate between these points.

[24]Of course, the last year of the model's 35-year sample does not exactly equal the new BGP because it takes time for the economy to exactly converge to the new BGP once the measures stop changing. However, we find that the two points are close.

$i$ in $t$). In principle, this step requires solving a complicated fixed-point problem. Namely, for any candidate path of measures $\mu_{it}$, we must solve for the equilibrium of the model, derive the implied path of aggregate employment series $N_{it}$, check if it matches the empirical path, and if not, update the guess of each $\mu_{it}$. Computationally, this procedure is prohibitively costly because it requires solving for the entire transition path for each candidate $\mu_{it}$.

Instead, we compute the path of measures $\mu_{it}$ from a simpler approach that approximately matches the path of $N_{it}$ from the data. In this exercise, we assume that the economy is initially along some BGP with measures $\mu_{i0}$, where we choose $\mu_{i0}$ to match the initial values of $N_{i0}$ from the data described above. At date $t = 0$, all agents unexpectedly learn about a new exogenous path of measures $\mu_{it}$ from $t = 0$ to a final point 35 years later, after which these measures remain constant. As described above, we assume that these final measures $\mu_{iT}$ correspond to the calibrated BGP from Section 4. In order to construct the time path of measures in between these two points, we compute the path of $\mu_{it}$ which solves the approximate relationship $N_{it}^{\text{data}} = \mu_{it} n_i^*$, where $n_i^*$ is the per-capita employment rate for each family $i$ in the final BGP. This relationship is an approximation because per-capital employment rates $n_{it}$ may endogenously change over time in equilibrium in response to the changes in measures $\mu_{it}$.

### D.3 Results

Given the variation described, we construct an estimator for the long-run elasticity of substitution $\phi$ following an established approach in the literature. In particular, suppose — like the majority of the literature which estimates this elasticity — that we interpret the data through the lens of a static CES production function $G(N)$ under perfect labor market competition, and accordingly construct the natural estimator $\widehat{\phi}$ of $\phi$. Our question now is: How biased would that estimator be if the true data-generating process were our model? We set to address this question next.

**Estimator.** To motivate the estimator we use, note first that under the assumption of a static and competitive labor market, the ratios of wages of workers with different skills equal the ratios of their marginal products,

$$\frac{w_{it}}{w_{jt}} = \frac{z_i}{z_j} \left( \frac{N_{it}}{N_{jt}} \right)^{-\frac{1}{\phi}} .$$

Under the CES form for $G(N)$, the ratio of marginal products depends on the ratio of employment rates as well as the ratio of skill levels $z_i/z_j$. Since these skill levels are constant over time, taking log-differences over time of the above condition gives

$$\Delta \log \left( \frac{w_{it}}{w_{jt}} \right) = -\frac{1}{\phi} \Delta \log \left( \frac{N_{it}}{N_{jt}} \right). \tag{D1}$$

Intuitively, we can use variation in labor supply to trace out firms' labor demand schedule, which depends on $\phi$. We obtain an estimate of $\phi$ by estimating (D1) through a simple linear regression using our model-simulated data constructed above. To construct the ratios $\frac{w_{it}}{w_{jt}}$ and $\frac{N_{it}}{N_{jt}}$ in our model, we choose the middle skill level $j_0 = N/2$ to be in the denominator, and then compute $\frac{w_{e,j't}}{w_{e,j_0t}}$ for all $j' \neq j_0$ and $e \in \{L, H\}$. Following Card and Lemieux (2001), for each of these ratios, we compute 5-year time differences except for the first observation, which is a 10-year difference. We then run the unweighted regression across each group and time period.

**Results.** Table D.1, which reproduces Table 4 from the main text, illustrates the results of this exercise. The first column displays the true value of $\phi = 4$. The second column shows that our estimator leads to an estimate $\widehat{\phi} = 3.94$ of $\phi$, very close to the true value. We conclude from this exercise that the estimates of the long-run elasticity of substitution in the literature provide appropriate discipline to pin down the value of $\phi$ in our model as well. The last column of the table shows that by estimating $\phi$ using a long difference between the final BGP and the initial BGP, we recover almost exactly the true value. The only slight difference stems from the presence of search frictions which, as noted in the main text, do not vary meaningfully across workers with different productivity. As noted in the main text, the estimator may not perform well in our model due to the existence of the putty-clay frictions. Because of these frictions, the 5-year time differences employed by Card and Lemieux (2001) reflects not only the true long-run substitution possibilities among workers assigned to new capital but also the employment ratios required to operate existing capital. Since these old employment ratios are rigid, taking them into account would bias the estimator $\widehat{\phi}$ downward relative to the true value of $\phi$. From this perspective, it may be surprising that our estimator does so well, despite the importance of putty-clay frictions in determining the slow response of the economy to changes in policy.

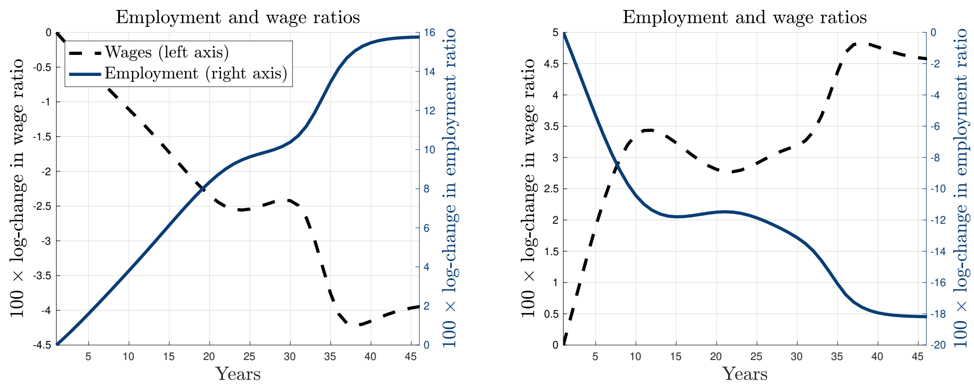
As argued in the main text, we interpret this finding as implying that putty-clay frictions are less important in shaping the value of $\phi$ that we recover than in shaping the response of the economy to, say, changes in the minimum wage. Figure D.1 further illustrates this logic. The two panels in the figure illustrate the wage and employment ratios for two non-college worker types $j'$ relative to the base type $j_0$ over the entire transition path. For both types of workers, the log-change in

Table D.1: Estimation Strategy for Long-Run Elasticity $\phi$ in Literature

| True value | Card and Lemieux (2001) Variation | Comparing BGPs Only |
|---|---|---|
| $\phi = 4$ | $\widehat{\phi} = 3.94$ | $\widehat{\phi} = 3.99$ |

Notes: Results from simulating the model for the time path of measures of families $\mu_{it}$ as described in text. *True value* reports the value of the long-run elasticity of substitution across workers $\phi = 4$ used to simulate the model. *Card and Lemieux (2001) Variation* reports the estimate $\widehat{\phi}$ from the regression in (D1) using 5-year time differences, with the exception of the first observation which is a 10-year difference. "Comparing BGPs Only" reports the estimator $\widehat{\phi}$ from (D1) using a long difference between the new BGP and the initial BGP.

Figure D.1: Variation in Employment and Wage Ratios for Estimating Elasticities




Notes: time path of employment ratios and wage ratios for individual non-college worker types for variation in measures of families $\mu_{it}$. Employment ratios are $N_{L,j't}/N_{L,j_0t}$ where $N_{L,j't} = \mu_{L,j't}n_{L,j't}$ and $N_{L,j_0t} = \mu_{L,j_0t}n_{L,j_0t}$ are the total measures of workers of that type and $j_0 = N/3$ is the base type, as described in the text. Wage ratios are similarly defined as $w_{L,j't}/w_{L,j_0t}$, where $w_{L,j't}$ and $w_{L,j_0t}$ are average wages among employed workers.

their employment ratios is approximately equal to $-\phi = -4$ times the log-change in their wage ratios by (D1). Firms' adjustment of these ratios is smooth because the underlying variation in the measures of families is itself smooth, and firms can predict it starting at $t = 0$.[25] In contrast, the policy experiments analyzed in the main text involve an immediate and unexpected jump in policy.

# E Additional Quantitative Results

This Appendix contains two additional quantitative results about the minimum wage.

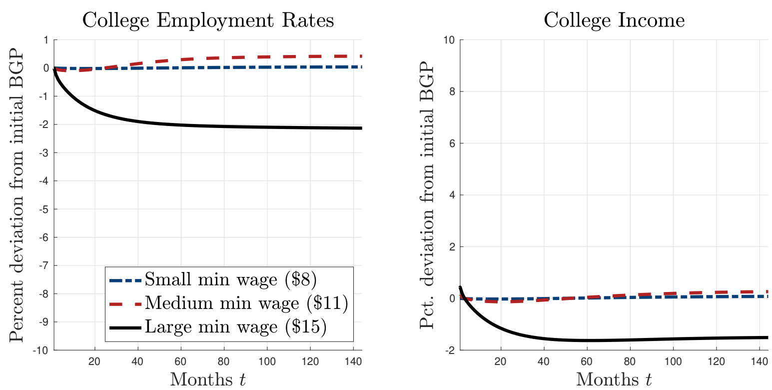
## E.1 Impact of Minimum Wage Changes on College Workers

In the main text, we focused on how labor market policies affected employment and labor earnings of non-college workers. These workers are most likely to be effected by the labor market policies we studied. In this subsection, we show how minimum wage changes of various sizes would affect non-

[25] Consistent with this predictability, the adjustment of employment ratios vs. wage ratios is farthest from $\phi$ in the early stages of the transition.

college workers. Appendix Figure E.1 plots the transition paths of *aggregate college employment and labor income* following the introduction of our three illustrative minimum wages. The small minimum wage has almost no effect on college employment and labor income because it is not binding for nearly all college workers. The medium minium wage has a very small positive effect on college employment and income because it reduces the monopsony distortion of a few low productivity college workers. Finally, the large minimum wage has a slightly negative effect on college employment for two reasons. First, the large minimum wage now binds on some lower productivity college workers. Second, the large reduction in non-college employment also reduces the marginal product of college workers because they are complementary in production.

FIGURE E.1: Dynamic Effects of the Minimum Wage for College Workers

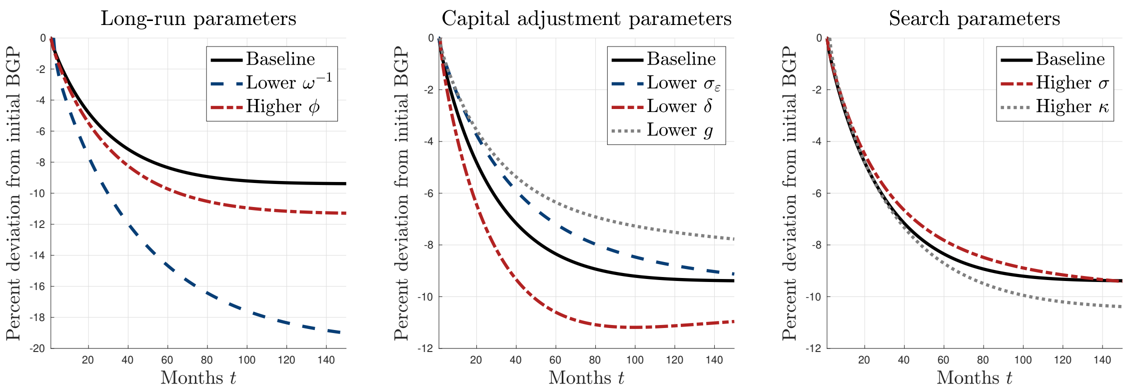



Notes: Transition paths of aggregate college employment (left panel) and labor income (right panel) following the introduction of the minimum wage. Employment is expressed in percentage deviation from the initial BGP. Labor income, net of trend growth $(1+g)^t$, is expressed relative to the initial BGP.

## E.2 Additional Robustness

Figure 10 in the main paper showed the robustness of the speed of transition to the new BGP when we change various parameters. In this subsection of the appendix we show how those parameters affect long rung changes in aggregate employment for non-college workers in response to a $15 minimum wage. Specifically, Appendix Figure E.2 plots our sensitivity analysis from Figure 10 of the main paper in terms of percentage deviations from the initial BGP rather than percentage of the total long run change. In this space, we can better assess the long-run effects of these various parameterizations. For example, the parameterization with less monopsony power (i.e. higher $\omega$) leads to a larger long-run decline in non-college employment.

Figure E.2: Sensitivity Analysis for \$15 Minimum Wage, Non-College Employment




Notes: Figure shows transition of non-college employment expresses as the percent deviation from the original BGP. "Baseline" corresponds to the model shown in Figure 6. "Higher $\omega^{-1}$" corresponds to a degree of monopsony power of $\omega^{-1} = 1/6$ that produces an 85% markdown. "Higher $\phi$" corresponds to a long-run elasticity of substitution within education groups of $\phi = 4.5$. "Lower $\sigma_\varepsilon$" corresponds to a standard deviation of idiosyncratic capital productivity of $\sigma_\varepsilon = 0.01$ that generates a steady-state capacity utilization rate of 97%. "Higher $\delta$" sets the depreciation rate to $\delta = 20\%$ annually. "Lower $g$" corresponds to a trend growth rate of $g = 1\%$. "Higher $\sigma$" sets the job-destruction rate to $\sigma = 3.5\%$ monthly. "Higher $\kappa$" increases the baseline vacancy-posting cost $\kappa_0$ by 2.5 times, which approximately doubles the average hiring costs $\kappa_i/\lambda_f(\theta_i)$ to 125% of average monthly wage.