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

October 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. Since these policies can greatly differ between the short run and the long run, an accurate assessment of their overall impact must take into account the entire time path of the responses they induce.

*This is a revised version of our paper titled "The Distributional Impact of the Minimum Wage in the Short Run and the Long Run". First Draft: May 2020. 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, Berkeley, Chicago, Cornell, Harvard, MIT, Michigan, Minnesota, North American Summer Meeting of the Econometric Society 2021, NBER Summer Institute, NBER Fall 2022 EFG Meeting, Northwestern, Princeton, NYU, Stanford, UCLA, UCSD, UC Santa Cruz, USC, Wharton, Wisconsin and the Federal Reserve Banks of Atlanta, Philadelphia, Richmond and San Francisco. 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.

1 Introduction

The high degree of wage inequality in the United States, coupled with growing evidence on firm monopsony power in labor markets, has recently renewed an 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 relative to other inputs of production, including other workers. Hence, the ultimate effects of these policies on employment, income, and welfare must jointly account for any potential input substitution that may induce over time. Such dynamic considerations are likely to be quantitatively important, since firms are known to substitute away in the long run from workers whose relative wages increase.[1]

Often, the short-run impacts of labor market policies are easy to infer using well-identified empirical strategies.[2] The standard approach in this vein estimates how wages and employment change in a group treated by the policy in the one or two years after its implementation using the corresponding changes in some other untreated group as a control. The vast majority of this literature finds that increases in the minimum wage of various sizes lead to small short-run changes in employment, which suggests that the short-run elasticity of substitution across inputs is low. Longer-run effects, however, are harder to identify through such a research design as many other forces can cause wages and employment to diverge between treatment and control groups over an extended period of time, thus confounding the effects of the policy of interest.

Given the lack of viable empirical strategies to directly measure long-run outcomes, a natural question is how the full dynamic effects of labor market policies can be analyzed. One approach is to specify a framework that accounts for the potential substitutability across inputs over time in response to a policy. But for the quantitative implications of such a framework to be taken seriously, they need to rely on credible estimates of (i) the long-run substitutability among inputs, (ii) how the substitutability among inputs evolves over time, and (iii) the structure of the labor market, in particular the degree of firms' labor market power. As we formalize, it is indeed the interplay between the long-run substitutability of inputs and monopsony distortions that crucially determines the long-run effects of an increase in the minimum wage or an expansion of the EITC.

In this paper, we fill a gap in the literature by developing a framework in which the elasticity of substitution across inputs of production endogenously changes over time so as to analyze the

[1] A large literature has shown the empirical importance of labor-labor substitution in production. See, for example, Katz and Murphy (1992), Card and Lemieux (2001), Borjas and Katz (2007), and Bils, Kaymak and Wu (2024).

[2] For work on the short-run effects of labor market policies such as the minimum wage and the EITC, see, for example, Neumark and Shirley (2022) for a recent survey on the minimum wage literature and Whitmore Schanzenbach and Strain (2021) for a recent paper on the EITC.

dynamic effects of labor market policies. To do so parsimoniously, we combine a model of putty-clay capital with rich worker heterogeneity, firm monopsony power, and frictional labor markets. We note that important parameters of this model can be disciplined with standard moments from the labor and macro literatures, which identify key features of production technologies through plausibly exogenous sources of variation unrelated to labor market policies. For example, the *long-run* substitutability across workers of different productivity has been frequently estimated by exploiting exogenous changes in the size of different productivity groups; see, for example Katz and Murphy (1992), Card and Lemieux (2001), Borjas and Katz (2007), and Bils, Kaymak and Wu (2024). In contrast to the small short-run elasticities of substitution across workers implied by changes in the minimum wage, this literature finds much larger longer-run elasticities of substitution across workers. Accordingly, we maintain that the long-run substitutability across workers of different productivities in response to changes in labor market policies is similar to the long-run substitutability across workers of different productivities in response to exogenous changes in the size of different demographic groups. Through the lens of our model, such a mapping is appropriate.

At its core, our framework is a new model of dynamic labor adjustment, which extends the putty-clay model of Gilchrist and Williams (2000) to incorporate worker heterogeneity.[3] The idea is that when making investment decisions, firms choose among units of capital that require different mixes of workers to be operated. In the long run, firms can flexibly substitute across workers by investing in different types of capital—a margin of adjustment that we discipline through the estimates of long-run labor-labor substitutability discussed above. Once a unit of capital is in place, however, production is Leontief in its required mix of workers. Thus, in the short run, firms can only adjust their labor demand by choosing how much to utilize each unit of installed capital. We show that information on capital utilization and depreciation rates can be used to discipline this short-run margin of adjustment as well as the speed at which the economy transitions to the long run. Through this dynamic structure, our model is jointly consistent with existing estimates of low labor-labor substitutability in the short run and high labor-labor substitutability in the long run.

Since firm monopsony power is often a justification for labor market policies, we incorporate it by assuming that workers have idiosyncratic preferences over firms.[4] This feature drives a wedge between firms' average and marginal costs of hiring a worker, resulting in them paying lower wages

[3]Putty-clay models have been shown to explain well the dynamic adjustment of the capital stock to a variety of macroeconomic shocks. See, for example, Gilchrist and Williams (2000), who show that unlike models with standard capital, models with putty-clay capital can generate a prolonged hump-shaped response in hours, persistent changes in output growth, and steeper recessions relative to expansions in response to business cycle shocks.

[4]That labor market policies such as the minimum wage may correct monopsony distortions dates back to Robinson (1933). For a recent discussion, see Manning (2021a), Manning (2021b), and Berger, Herkenhoff and Mongey (2025).

compared to the efficient level and hiring fewer workers. Ideal labor market policies that correct both distortions are a minimum wage for each worker set equal to a worker's competitive wage and a subsidy to vacancy creation for each worker. Since these policies are complex, we study two common policies used in practice: a uniform minimum wage and transfers to low-income workers. Based on the intuition gleaned from our ideal policies, we note that for any given worker, these policies can generate employment and welfare gains if they move wages closer to their efficient level. However, if they raise wages beyond this level, firms will substitute away from such workers at a speed governed by the putty-clay production structure outlined above. In this sense, the long-run effects of labor market policies are determined by both long-run input substitution elasticities and the extent of firm monopsony power in the economy.

We close the model in general equilibrium allowing for search frictions in the labor market. In practice, search frictions help us to pin down labor market allocations when the minimum wage binds. But, more importantly, combining search frictions and monopsony power allows us to interpret existing empirical estimates of wage markdowns. In particular, wage markdowns in our model consist of both an inefficient component, due to monopsony distortions, but also an efficient component, due to firms needing to be compensated for the costs incurred in hiring workers.[5] We discipline the total wage markdown, which reflects both components, using recent estimates 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 pin down the efficient share of the wage markdown using moments on labor market flows, which reveal the degree of search frictions. Quantitatively, we find the efficient component of the wage markdown to be small compared to the inefficient monopsony component.

Our main result is that taking into account the dynamic adjustment of labor is necessary to correctly assess the full impact of labor market policies. Our model is ideally suited for such a policy evaluation exercise because of its unique ability to simultaneously generate low labor-labor substitutability in the short run, consistent with the minimum wage literature, and high labor-labor substitutability in the long run, consistent with the classic work of Card and Lemieux (2001) and Borjas and Katz (2007). We illustrate this result using the minimum wage as our leading example because it provides an instance of a particularly clean mapping between the policy under consideration and the implied changes in the relative wages of different types of workers. Crucially, we show that the long-run direction of adjustment depends on the size of the change in the minimum wage being contemplated. Due to firm monopsony power, a small increase in the minimum wage ultimately *increases* the long-run employment of workers for whom the minimum wage binds by

[5] See, for example, Kline (2025) for a recent discussion.

reducing the monopsony distortions they face. However, increasing the relative employment of these workers requires firms to build new types of capital that are more intensive in the use of these workers' services, which takes time to accumulate. The fact that employment changes little in the short run can then be consistent with employment for low-productivity workers actually increasing in the long run for small changes in the minimum wage.

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 due to the putty-clay nature of technology. Namely, 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 capital is Leontief across workers. Instead, firms progressively substitute away from less productive workers by not fully replacing those who naturally attrit, thereby leaving idle some of their existing capital. At the same time, firms start installing 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 the decline in employment in response to a minimum wage increase would occur at most at the speed at which capital depreciates. Instead, our model with an endogenous capital utilization margin can accommodate richer patterns for the speed of an economy's transition. It does so by allowing firms to flexibly choose to idle more of their existing capital and thus to phase out workers more quickly than capital depreciates. In this sense, our model breaks the mechanical link between worker attrition and capital depreciation that is essentially hard-wired into first-generation putty-clay models.

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, the initial labor income gains for the lowest-productivity workers are eroded over time as firms slowly substitute away from these workers. The employment loss is eventually large enough that the labor income of the lowest-productivity workers ultimately falls. In this case, our putty-clay technology delays the long-run *costs* of the policy.

Quantitatively, we use our model to illustrate the potential limitations of two common strategies for the evaluation of policies, which fall short of accounting for their full dynamic effects. First, a typical *static or short-run* analysis measures in the data the effects of a minimum wage change

over the first couple of subsequent years and then presumes that what occurs during this period of time captures what occurs from then on. Second, a typical *long-run or steady-state* analysis instead assesses the impact of labor market policies by comparing the new and old steady states of an economy. We show that both the short-run and long-run approaches severely misstate the overall labor income gains from labor market policies. For example, by 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%—twice as much. 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. Namely, when an economy's response to a sizable change in policy is slow, as existing evidence suggests, 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 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, thus leading to 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. 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 path that an economy experiences. These results also suggest that a large increase in the EITC—possibly funded by a tax on corporate profits—may be a preferred way to support lower-income workers relative to a large increase in the minimum wage, because the EITC reduces firms' incentive to substitute away from the affected workers over time.

We show that our parameterized framework matches many of the broad empirical patterns documented in the literature. For example, Clemens and Strain (2021) examine the employment effects of minimum wage increases of various size in the short and medium run. Consistent with the predictions of our model with putty-clay capital and firm monopsony power, they find empirical evidence that small minimum wage increases generate no employment declines in either the short or the medium run. Larger minimum wage increases, however, lead to no short-run employment declines but result in much larger employment declines in the medium run. In line with our key labor-labor substitutability assumption, Horton (2025) relies on a large-scale experiment from an online job platform to show that after the imposition of a randomized minimum wage on the platform, treated

employers systematically substitute away from lower-skilled workers towards higher-skilled ones. Likewise, Clemens, Kahn and Meer (2021) provide evidence that firms substitute toward older and higher-educated workers after a minimum wage increase. Finally, Harasztosi and Lindner (2019) document that in response to a large and persistent increase in the minimum wage in Hungary, exposed firms adjusted their capital stock in the long run.

A core contribution of this paper is the creation of a quantitative model that researchers and policymakers can use to assess the full dynamic impact of labor market policies on worker welfare. Some of the key elements of our framework have antecedents in the existing literature. As noted above, we augment the model of putty-clay model developed in Gilchrist and Williams (2000) by including heterogeneous workers. Additionally, Sorkin (2015) uses a simple, partial equilibrium putty-clay model with one type of workers to illustrate that the short-run effects of the minimum wage measured by the empirical literature may be different from the long-run ones.[6] However, his model abstracts from worker heterogeneity and firm monopsony power, which we find are critical in shaping the dynamic effects of the minimum wage over time.[7] Our model of monopsony builds on that in Berger, Herkenhoff and Mongey (2022) by setting it in a frictional labor market, adding rich worker heterogeneity and putty-clay input adjustment.[8] Finally, the search literature has also studied the effects of labor market policies in various types of frictional labor markets without inefficient monopsony or putty-clay input adjustment.[9] By combining and extending these various strands of the literature, we have developed a framework that yields distinct theoretical and quantitative predictions about the dynamic effects of labor market policies.

Ultimately, we hope that our framework will help focus the discussion of the benefits or costs of the minimum wage and other income support policies on the magnitudes of two key parameters: the degree of monopsony distortion in labor markets and the degree of labor-labor substitutability across workers of different productivity. We believe that focusing on refinements in the measurement of these key parameters is a promising way forward in the decades-long debate over the pros and cons of labor market policies.

[6] Sorkin (2015) adapts the putty-clay setup of Atkeson and Kehoe (1999)—used to model the slow decline of an input of production, energy use, in response to an exogenous increase in its price—to model the slow decline of another input, the change in employment, in response to an exogenous increase in the minimum wage.

[7] Aaronson et al. (2018) develop a model with two types of workers in which firm entry and exit is the only intensive margin of labor adjustment to the minimum wage. Like Sorkin (2015), they abstract from firm monopsony power.

[8] Berger, Herkenhoff and Mongey (2025) uses the model in Berger, Herkenhoff and Mongey (2022) to study the minimum wage in an environment with rich firm heterogeneity. They focus on the long-run effects of the minimum wage in a static environment. See Mousavi (2022) and Berger et al. (2024) for work exploring the interaction between monopsony distortions and taxation more broadly.

[9] See, for example, Eckstein and Wolpin (1990), Flinn (2006), Ahn, Arcidiacono and Wessels (2011), Engbom and Moser (2022), and Drechsel-Grau (2022). Given the observed frequency of labor market flows in the U.S. data, transition dynamics solely due to search frictions are much faster than those generated by our putty-clay model.

2 Model

We now develop our model with putty-clay capital, firm monopsony power, and search frictions in the labor market. Our economy features a single output good used for consumption and investment over an infinite horizon in discrete time.[10] 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.

2.1 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 μ_i, and each is composed of a large number of members.[11] Each family has preferences represented by the utility function

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

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.[12] The extent of firm monopsony power is governed by the degree of substitutability of jobs in workers' preferences captured by the parameter ω. 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 each 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 θ_{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 W_{ijt+1} over the life of the match as detailed below. Given these offers, each type-i family chooses the

[10] See Aaronson and French (2007) and MaCurdy (2015) for analyses of the minimum wage that account for differential effects of the minimum wage on the relative prices of the goods of different sectors, which employ different mixes of workers. Although extending our framework along these dimensions 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 worker skill groups within each sector, the elasticity of demand of each sector for each skill group, and input-output links among sectors, which is beyond the scope of this paper.

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

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

measure of family members s_{ijt} searching for jobs at each firm j. This measure of family members finds a measure of jobs $\lambda_w(\theta_{ijt})s_{ijt}$ at firm j in period t, begins working in $t+1$, is paid the wage $w_{ijt+\tau} = (1+g)^\tau w_{ijt+1}$ per member in each period $\tau \geq 1$ of employment, and exogenously separates with probability σ 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 measures of family members employed at each such firm $\{n_{ijt}\}$ to solve

$$\begin{aligned} &\max_{c_{it}, s_{ijt}, n_{ijt+1}} \sum\nolimits_{t=0}^{\infty} \beta^t U_t(c_{it}, n_{it}, s_{it}) \qquad (1) \\ &\text{s.t. } n_{ijt+1} = (1-\sigma)n_{ijt} + \lambda_w(\theta_{ijt})s_{ijt} \\ &\quad \sum\nolimits_{t=0}^{\infty} Q_{0,t}c_{it} = \zeta_i \mathbb{P} + \mathbb{I}_i + \sum\nolimits_{t=0}^{\infty} Q_{0,t} \sum\nolimits_j \lambda_w(\theta_{ijt})s_{ijt}W_{ijt+1}, \end{aligned}$$

where the first constraint is the law of motion of employment in which σ is the exogenous separation rate and $\lambda_w(\theta_{ijt})s_{ijt}$ is the measure of searching members who 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

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

where U_{cit} and U_{sit} denote the derivative of U_t 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 whose jobs workers search for. The right side is the marginal benefit of searching for jobs, which reflects that a worker finds a job at firm j at rate $\lambda_w(\theta_{ijt})$, begins working in the next period, and receives the present value W_{ijt+1} of wage payments net of the present value V_{ijt+1} of the disutility of work at firm j, both discounted by $Q_{t,t+1}$. Throughout, we denote by $Q_{t,s} = Q_{0,s}/Q_{0,t}$ 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

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

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 $Q_{t,t+1}\lambda_w(\theta_{ijt})(W_{ijt+1} + V_{ijt+1}) \geq \mathcal{W}_{it} \equiv \max_{j'}\{Q_{t,t+1}\lambda_w(\theta_{ij't})(W_{ij't+1} + V_{ij't+1})\}$. In a symmetric equilibrium in which all firms other

than firm j offer workers the common value $Q_{t,t+1}\lambda_w(\theta_{it})(W_{it+1}+V_{it+1})$, this inequality reduces to

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

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. This 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 show below, firms' monopsony power affects not only the wages that firms pay to workers but also the number of job vacancies firms post or, equivalently, workers' probabilities of finding a job. It turns out that the impact of firms' monopsony power on these two margins leads to distinct sources of inefficiencies in the labor market: wages are lower and vacancies are fewer than in the corresponding competitive equilibrium.

2.2 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.[13]

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 is, the less productive capital is 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

[13] We adopt the technology of Gilchrist and Williams (2000), who demonstrate 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.

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 ε. 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 ε, where $\pi(\varepsilon)$ is the p.d.f. of ε, which has mean 1, and $\Pi(\varepsilon)$ is its 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 ε 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. If firm j combines $K_{jt}(A_{t-\tau}, \varepsilon)$ units of type-$(A_{t-\tau}, \varepsilon)$ capital with $N_{ijt}(A_{t-\tau}$,-$\varepsilon)$ units of each type of labor, it produces $A_{t-\tau}\varepsilon \times F\left(K_{jt}(A_{t-\tau}, \varepsilon), N_{1jt}(A_{t-\tau}, \varepsilon), \ldots, N_{Ijt}(A_{t-\tau}, \varepsilon)\right)$ 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.[14] 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 in production among different groups of workers. 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 as

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

In the putty-putty model, firms freely choose these labor intensities $v_{ijt}(A_{t-\tau}, \varepsilon)$ for each type of capital in any period. The optimal choice is to equate labor per efficiency unit of capital across machines, resulting in the aggregate production function $Y_{jt} = F(K_{jt}, N_{1jt}, \ldots, N_{Ijt})$, where K_{jt} is the aggregate efficiency units of capital; see Appendix C. In contrast, in the putty-clay model, firms cannot reassign workers ex post to equate labor per efficiency unit of capital so no analogous aggregate production function arises. We consider this version of the model next.

Putty-Clay Production. In the putty-clay model, firms choose the labor intensity of capital of any

[14] A Cobb-Douglas technology in capital and the labor aggregate is required to achieve balanced growth with vintage capital productivity growth. This specification is also broadly consistent with estimates of the elasticity of substitution between capital and labor in the literature. For instance, Oberfield and Raval (2021) suggest values for α between 0.5 and 0.7, whereas Karabarbounis and Neiman (2014) estimate a value of 1.25.

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 workers needed to operate the chosen machine is fixed over the life of the machine. Following Gilchrist and Williams (2000), we assume that the labor intensity of capital is set before capital is installed and the idiosyncratic productivity ε of each unit is realized. This timing assumption leads to an active shutdown margin for each unit of capital. 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 $N_1(v, A_{t-\tau}, \varepsilon), \ldots, N_I(v, A_{\tau-\tau}, \varepsilon)$ units of each type of labor, it produces output

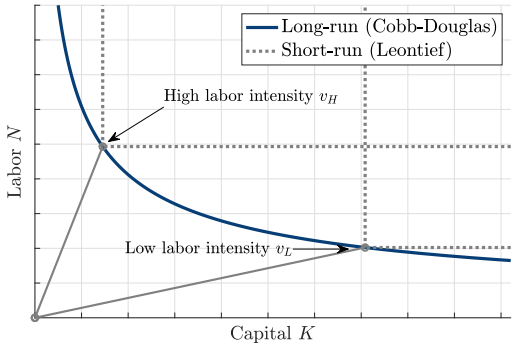
$$Y_{jt}(v, A_{t-\tau}, \varepsilon) = A_{t-\tau}\varepsilon \times \min\left\{K_{jt}(v, A_{t-\tau}, \varepsilon), \frac{N_{1jt}(v, A_{t-\tau}, \varepsilon)}{v_1}, \ldots, \frac{N_{Ijt}(v, A_{t-\tau}, \varepsilon)}{v_I}\right\} \times f(v). \quad (4)$$

The minimum operator captures the Leontief nature of production once capital is installed. The maximum output that a firm can obtain from these units of capital is when it assigns to them $N_{ijt}(v, A_{t-\tau}, \varepsilon) = v_i K_{jt}(v, A_{t-\tau}, \varepsilon)$ units of labor of each type 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, $v_i K_{jt}(v, A_{t-\tau}, \varepsilon) - N_{ijt}(v, A_{t-\tau}, \varepsilon)$, 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*

$$\sum\nolimits_{\tau} \int_{v,\varepsilon} N_{ijt}(v, A_{t-\tau}, \varepsilon)\pi(\varepsilon) dv d\varepsilon \leq N_{ijt}. \quad (5)$$

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 so that $G(N) = N$. 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 any labor intensity $v = N/K$ along a long-run isoquant. However, once capital is installed, its labor intensity v is fixed and a firm's only choice is how much of it to utilize, that is, where on the corresponding ray from the origin a firm chooses to locate itself. Over time, though, firms

FIGURE 1: 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, $G(N) = N$. The long-run isoquant corresponds to $F(K, N) = 1$. Short-run isoquants correspond to Leontief production functions for different labor intensities $v = N/K$.

can choose different points along the long-run isoquant by accumulating different types of capital. Likewise, with multiple types of workers so that, say, $F(K, N_1, N_2) = K^{\alpha}G(N_1, N_2)$, a firm over time can gradually substitute across workers by installing new machines that embody different labor-to-capital ratios. For example, when the minimum wage increases, firms can substitute away from low-productivity workers towards higher-productivity ones or capital. Below we show that the speed of the economy's adjustment to labor market policies critically depends on how much of the existing capital is utilized and the rate at which new capital with 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 Leontief representation of production 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

$$Y_{jt}(v, A_{t-\tau}, \varepsilon) = A_{t-\tau}\varepsilon K_{jt}(v, A_{t-\tau}, \varepsilon) \min\left\{1, \frac{1}{v_1} \cdot \frac{N_{1jt}(v, A_{t-\tau}, \varepsilon)}{K_{jt}(v, A_{t-\tau}, \varepsilon)}, \ldots, \frac{1}{v_I} \cdot \frac{N_{Ijt}(v, A_{t-\tau}, \varepsilon)}{K_{jt}(v, A_{t-\tau}, \varepsilon)}\right\} f(v), \tag{6}$$

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 the 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 Leontief structure in (4)

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 (6) using (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 δ, the amount of (v, A_t, ε)-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 subject to search frictions, in which matching is governed by the 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 κ_{it} each and posts offers of the form $(\theta_{ijt}, W_{ijt+1})$ to attract workers of each type i, where market tightness $\theta_{ijt} = \mu_i a_{ijt}/(\mu_i s_{ijt})$ determines the measure $\lambda_w(\theta_{ijt})\mu_i s_{ijt}$ of searching workers from the μ_i type-i families that find a job with firm j and the measure $\lambda_f(\theta_{ijt})\mu_i a_{ijt}$ of type-i workers hired by firm j. 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 in t evolves according to labor transition law $N_{ijt+1} = (1-\sigma)N_{ijt} + \lambda_f(\theta_{ijt})\mu_i a_{ijt}$. 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 $\{X_{jt}(v)\}$ in each capital type v, allocations of type-i workers $\{N_{ijt}(v, A_{t-\tau}, \varepsilon)\}$ across installed capital for each machine of type $(v, A_{t-\tau}, \varepsilon)$, measures of vacancies to post $\{\mu_i a_{ijt}\}$ for workers of each type i, job offers $\{\theta_{ijt}, W_{ijt+1}\}$ for each such type, and total workers $\{N_{ijt}\}$ of type i understanding that $N_{ijt} = \mu_i n_{ijt}$ to maximize the present value of profits

$$\mathbb{P} = \sum\nolimits_{t=0}^{\infty} Q_{0,t} \left[Y_{jt} - \int_v X_{jt}(v) dv - \sum\nolimits_{i=1}^{I} \kappa_{it} \mu_i a_{ijt} - \sum\nolimits_{i=1}^{I} \lambda_f(\theta_{ijt}) \mu_i a_{ijt} Q_{t,t+1} W_{ijt+1} \right]. \tag{10}$$

Constraints to this problem are the participation constraints in (3) for all i and t, the adding-up constraints in (5) for all i and t, the utilization constraints in (8) for all i and t, the labor transition laws for all i and t, the investment irreversibility constraints $X_{jt}(v) \geq 0$ for all t, and the nonnegative vacancy posting 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 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 workers of each 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 maximizes (10); *c*) at each date t, job-finding and job-filling rates are consistent with the matching function; *d*) total employment of workers of each type at each firm j, $\{N_{ijt}\}$, satisfies the adding-up constraint in (5) and the labor transition laws; *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, $\sum_i \mu_i c_{it} + \sum_{i,j} \kappa_{it} \mu_i a_{ijt} + \sum_j \int_v X_{jt}(v) dv = \sum_{j,\tau} Y_{jt}$. We build intuition along the balanced growth path or *BGP* of the economy. Along this path, consumption, investment, output, wages, and the disutility of work and search grow with the economy at rate $1+g$. Job search, employment, vacancies, and intertemporal prices are constant.

3 Equilibrium Characterization

We now characterize equilibrium. We mainly focus on the firm problem, since it determines the key margins affecting the economy's response to labor market policies. First, endogenous capital utilization governs the short-run labor demand for labor given the distribution of installed capital. Second, investment in new types of capital lead to an increasing degree of substitutability across different types of workers over time.

3.1 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 component.

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 of maximizing (10). But by taking as given from this dynamic problem the multiplier $\hat{\chi}_{ijt}$ on the adding-up constraint for the uses of

labor in (5) together with a firm's state in period t, namely, the existing stock of capital vintages $\{K_{jt}(v, A_{t-\tau})\}_{\tau=1}^{\infty}$ and the stock of employed workers of each type $\{N_{ijt}\}$ in t, the utilization rate of each machine solves a static problem. The static utilization problem is

$$\max_{\{u_{jt}(v,A_{t-\tau},\varepsilon),N_{ijt}(v,A_{t-\tau},\varepsilon)\}_{\tau=1}^{\infty}} \sum\nolimits_{\tau} \int_{v,\varepsilon} A_{t-\tau}\varepsilon u_{jt}(v, A_{t-\tau}, \varepsilon) K_{jt}(v, A_{t-\tau}) f(v)\pi(\varepsilon) dv d\varepsilon$$
$$+ \sum\nolimits_{i} \hat{\chi}_{ijt} \left[N_{ijt} - \sum\nolimits_{\tau} \int_{v,\varepsilon} N_{ijt}(v, A_{t-\tau}, \varepsilon)\pi(\varepsilon) dv d\varepsilon \right] \tag{11}$$
$$\text{s.t. } 0 \leq u_{jt}(v, A_{t-\tau}, \varepsilon) \leq 1 \text{ and } u_{jt}(v, A_{t-\tau}, \varepsilon) v_i K_{jt}(v, A_{t-\tau}, \varepsilon) \leq N_{ijt}(v, A_{t-\tau}, \varepsilon).$$

By (9), the problem is to maximize the expected value of output in period t, $\sum_{\tau} \int_{v,\varepsilon} Y_{jt}(v, A_{t-\tau}, \varepsilon)$, by efficiently allocating employed workers $\{N_{ijt}\}$ to the existing stock of machines $\{K_{jt}(v, A_{t-\tau}, \varepsilon)\}$ and choosing their utilization rates. We refer to this problem as a *pseudo-Lagrangian* problem because the multipliers $\hat{\chi}_{ijt}$ are taken from a different problem—the dynamic problem of maximizing (10). These multipliers capture the shadow value of an additional marginal measure of type-i workers available for production only in t or, equivalently, the *marginal product* of type-i workers in t.

Proposition 1. *The optimal utilization rates for the static pseudo-Lagrangian problem in* (11) *are also the optimal utilization rates for the dynamic problem in* (10)*. Both are given by a cutoff rule such that firms fully utilize capital of type* $(v, A_{t-\tau}, \varepsilon)$ *with idiosyncratic productivity* ε *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* ε *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 $u_{jt}(v, A_{t-\tau}, \varepsilon)$ that maximizes (10) is

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

where $\lambda^{L}_{ijt}(v, A_{t-\tau}, \varepsilon), \lambda^{U}_{ijt}(v, A_{t-\tau}, \varepsilon)$, and $\lambda_{ijt}(v, A_{t-\tau}, \varepsilon)$ are the multipliers on the utilization constraints in (8). Substituting the first-order condition for labor assignment $N_{ijt}(v, A_{t-\tau}, \varepsilon)$ that maximizes (10), namely, $\lambda_{ijt}(v, A_{t-\tau}, \varepsilon) = \hat{\chi}_{ijt}\pi(\varepsilon)$, into the first-order condition for utilization and dividing by $K_{jt}(v, A_{t-\tau})\pi(\varepsilon)$ yields

$$A_{t-\tau}\varepsilon f(v) - \sum\nolimits_{i} \hat{\chi}_{ijt} v_i = \left[\lambda^{U}_{ijt}(v, A_{t-\tau}, \varepsilon) - \lambda^{L}_{ijt}(v, A_{t-\tau}, \varepsilon)\right] / \left[K_{jt}(v, A_{t-\tau})\pi(\varepsilon)\right]. \tag{12}$$

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 (12) 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 (12), 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 $\varepsilon < \underline{\varepsilon}$.[15] Note

[15] 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 $u_{jt}(v, A_{t-\tau}, \varepsilon) \in [0, 1]$.

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 (11) 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 those just described for the dynamic problem, we find that the cutoff rule of Proposition 1 is the solution to this static problem with the same $\hat{\chi}_{ijt}$. This concludes the proof.

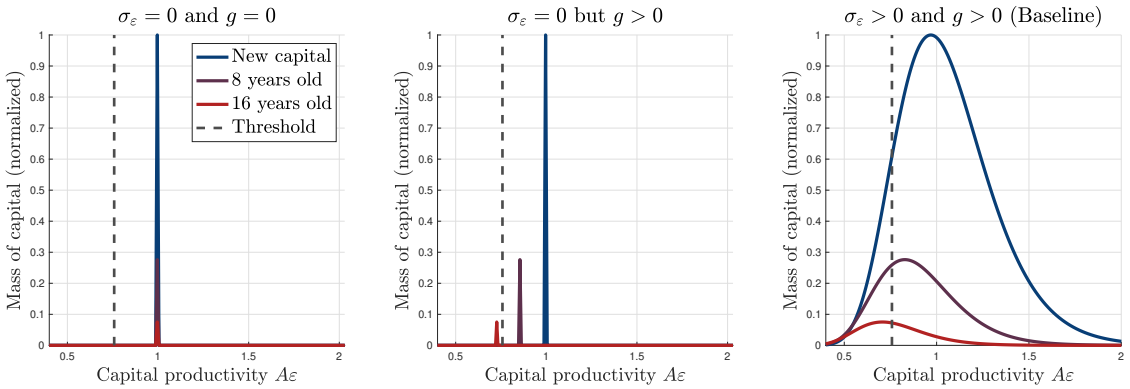
To understand this cutoff rule, suppose the $\pi(\varepsilon)$ machines using capital $K_{jt}(v, A_{t-\tau})$ with productivity $A\varepsilon$ are not being fully utilized so that $u_{jt}(v, A_{t-\tau}, \varepsilon) < 1$. Increasing their utilization by Δ would require an increase in each of the labor types assigned to them by $v_i K_{jt}(v, A_{t-\tau})\pi(\varepsilon)\Delta$. Doing so would tighten the adding-up constraint on all I types of labor by $\sum_i \hat{\chi}_{ijt} v_i K_{jt}(v, A_{t-\tau})\pi(\varepsilon)\Delta$. 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 value of the resources required to do so, namely, $\sum_i \hat{\chi}_{ijt} v_i K_{jt}(v, A_{t-\tau})\pi(\varepsilon)\Delta$. So far, to emphasize that the multiplier in the static problem came from the dynamic problem, we wrote it as $\hat{\chi}_{ijt}$. Hereafter, we let χ_{ijt} denote the multiplier from the dynamic problem.

The importance of the utilization rate margin is determined by the mass of existing capital near its operating threshold. Like Gilchrist and Williams (2000), we assume that $\log \varepsilon \sim N(-\sigma_\varepsilon^2/2, \sigma_\varepsilon^2)$ and, in our quantitative work, choose σ_ε to match the average capital utilization rate in the data.

Intuition from the Balanced Growth Path. To understand firms' utilization policy, 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 $v_{it} = N_{ijt}(v, A_{t-\tau}, \varepsilon)/K_{jt}(v, A_{t-\tau}, \varepsilon)$ that decreases 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 $\widetilde{v}_i = v_{it}(1+g)^t$ is the detrended type-i labor requirement of the capital and $\widetilde{\chi}_i = \chi_{it}/(1+g)^t$ is the detrended shadow value of type-i labor. Since $A_{t-\tau} = (1+g_A)^{-\tau} A_t$ and $A_0 = 1$, the idiosyncratic productivity cutoff 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 value of the labor required to operate the capital grows at rate g but its vintage productivity $A_{t-\tau}$ remains fixed. Thus, 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 σ_ε play an important role in shaping capital utilization. Each panel plots the distribution of overall capital productivity $A_{t-\tau}\varepsilon$. Note that the threshold $A_{t-\tau}\underline{\varepsilon}_{t-\tau}$ is constant across vintages τ, but since average overall productivity $A_{t-\tau}$ declines at rate $1+g$, the cutoff $\underline{\varepsilon}_{t-\tau}$

FIGURE 2: Capital Utilization Margin Along the BGP



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

must grow at rate $1+g$. All capital with productivity to the right of this threshold, denoted with a vertical dashed line, is utilized. The left panel of Figure 2 plots the distribution of capital productivity for a version of the model without vintage capital growth ($g = 0$) and without dispersion in capital productivity ($\sigma_\varepsilon = 0$) so $\varepsilon = 1$. In this *full-utilization* case, the distribution of capital productivity is degenerate and lies strictly above this utilization threshold. Thus, all capital is fully utilized as in Atkeson and Kehoe (1999). This result occurs because the shadow variable profit from operating one unit of capital, $A_{t-\tau}f(\widetilde{v}) - \sum_i \widetilde{\chi}_i \widetilde{v}_i$, is strictly positive due to firms needing to recover the cost of producing this capital and having monopsony power in the labor market. Hence, a small enough increase in the cost of labor does not induce firms to shut down capital—this intuition will prove key to understanding the short-run impact of the labor market policies we consider.

The middle panel of Figure 2 introduces capital heterogeneity across vintages (with $g = 2\%$ as in our quantitative model) but abstracts from heterogeneity within vintages ($\sigma_\varepsilon = 0$). The productivity of a given capital vintage $A_{t-\tau}$ now declines relative to that of the frontier vintage A_t so that old enough vintages cross the utilization threshold and are shut down. Intuitively, since $A_{t-\tau}$ decreases with τ, the shadow variable profit from operating one unit of capital, $A_{t-\tau}f(\widetilde{v}) - \sum_i \widetilde{\chi}_i \widetilde{v}_i$, becomes negative for old enough vintages. But the mass of such older vintages is fairly small because most of the old capital has depreciated away by the time this utilization 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 quantitative model). As before, the average productivity of newer vintages is higher than that of older ones. However, due to the significant dispersion in productivity within vintages, each vintage features a positive mass of capital that is shut down. Indeed, every vintage contains some marginal machines that are shut down for even a small increase in the cost of labor.

3.2 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 by describing 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 distortions provide a justification for labor market policies that can improve the efficiency of equilibrium.

To understand firms' dynamic monopsony power, consider the sequence of participation constraints in (3) that a firm in period 0 faces when making labor market decisions for any period 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 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 into account that hiring workers of a family of type i in any period t affects the utility of the family members hired in any subsequent period. Intuitively, if firm j hires additional $\lambda_w(\theta_{ijt})s_{ijt}\Delta$ workers in period $t \leq s$, then $(1-\sigma)^{s-t}\Delta$ of them are still working in period s. Through this term, the presence of these members of a family of type i hired in t affects the utility of all future members of this family hired in any future period s.[16] By contrast, in the perfectly competitive search case ($\omega \to \infty$), 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 the participation constraints of different 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

[16] 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.

time-t participation constraint of firm j's problem in (10). After grouping terms, the contribution of these constraints to the Lagrangian is

$$\sum_{t=0}^{\infty} Q_{0,t+1}\mu_i M_{ijt+1}\frac{U_{nit+1}}{U_{cit+1}}\left(\frac{n_{ijt+1}}{n_{it+1}}\right)^{\frac{1}{\omega}} + \sum_{t=0}^{\infty} Q_{0,t+1}\mu_i\gamma_{ijt+1}\left[W_{ijt+1} - \frac{\mathcal{W}_{it}}{Q_{t,t+1}\lambda_w(\theta_{ijt})}\right], \quad (13)$$

where M_{ijt} is an auxiliary variable with transition law $M_{ijt+1} = (1-\sigma)M_{ijt}+\gamma_{ijt+1}$ that summarizes a firm's *dynamic promises* to type-i workers by cumulating past multipliers on (3), where in (13) we have divided the participation constraint by $\lambda_w(\theta_{ijt})$.

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

$$\nu_{ijt} = \underbrace{\chi_{ijt}}_{\text{marginal product of labor}} + \underbrace{M_{ijt}\frac{U_{nit}}{U_{cit}}\frac{1}{\omega}\left(\frac{n_{ijt}}{n_{it}}\right)^{\frac{1}{\omega}-1}\frac{1}{n_{it}}}_{\text{source of monopsony distortion}} + Q_{t,t+1}(1-\sigma)\nu_{ijt+1}. \quad (14)$$

Here ν_{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 so gives the value of a marginal increase in the measure of type-i workers to firm j in t. The first term on the right side of (14), which is positive, is the flow value in production of assigning these workers to a marginal machine as captured by χ_{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 type-i workers from each of the μ_i families of type i tightens the participation constraints in (3). This term, which is equal to the derivative of the first term on the left side of (13) with 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 a family 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}—this disutility arises from the love-of-variety nature of preferences over different jobs. The third term captures that only $1-\sigma$ of these workers remain employed in $t+1$.

Monopsony Distortions. Here monopsony power reduces the present value of a worker to a firm in (14) and distorts firms' job creation and wage setting relative to a competitive equilibrium. To see how job creation is distorted, note that the first-order condition for vacancy posting a_{ijt} is

$$\frac{\kappa_i}{\lambda_f(\theta_{ijt})} = Q_{t,t+1}\left(\nu_{ijt+1} - W_{ijt+1}\right). \quad (15)$$

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 (15) equals the marginal benefit on the right side. This marginal benefit is the present value of having a worker who starts working in $t+1$ net of the present value of the wages paid to the worker. Because monopsony distortions lower firms' present value of employing workers ν_{ijt}, firms post fewer vacancies and hire fewer workers than in a competitive search equilibrium. To see how

wages are distorted, combine the first-order conditions for market tightness and wages to obtain

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

The present value of wages W_{ijt+1} is the marginal value of workers of type i to firm j minus the disutility of work of workers of type i at firm j, where each term is weighted by the elasticity of the matching function through η and $1-\eta$. Thus, since a firm's monopsony power lowers the value ν_{ijt+1} of a worker to a firm, 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. These decisions together determine how the elasticity of substitution across workers evolves over time. The investment problem is in principle 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 the *single-peakedness* assumption that

$$\sum\nolimits_{\tau=1}^{\infty} Q_{t,t+\tau}(1-\delta)^{\tau-1} \int_{\underline{\varepsilon}_{t,t+\tau}}^{\infty} \Big[A_t \varepsilon f(v) - \sum\nolimits_i \chi_{it+\tau} v_i\Big]\pi(\varepsilon)d\varepsilon \tag{16}$$

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

Proposition 2. *Under* (16)*, $X_t(v) > 0$ for at most one type of capital denoted v_t that solves*

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

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

$$1 = \sum\nolimits_{\tau=1}^{\infty} Q_{t,t+\tau}(1-\delta)^{\tau-1} \int_{\underline{\varepsilon}_{t,t+\tau}}^{\infty} \Big[A_t \varepsilon f(v_t) - \sum\nolimits_i \chi_{it+\tau} v_{it}\Big]\pi(\varepsilon)d\varepsilon. \tag{18}$$

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

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

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 $q_{t,t+\tau}(v) = \int_{\underline{\varepsilon}_{t,t+\tau}}^{\infty} [A_t \varepsilon f(v) - \sum_i \chi_{it+\tau} v_i]\pi(\varepsilon)d\varepsilon$, which when substituted in the investment first-order condition $\mu_t(v) = 1 - \sum_{\tau=1}^{\infty} Q_{t,t+\tau}(1-\delta)^{\tau-1} q_{t,t+\tau}(v)$ yields

$$\mu_t(v) = 1 - \sum\nolimits_{\tau=1}^{\infty} Q_{t,t+\tau}(1-\delta)^{\tau-1} \int_{\underline{\varepsilon}_{t,t+\tau}}^{\infty} \Big[A_t \varepsilon f(v) - \sum\nolimits_i \chi_{it+\tau} v_i\Big]\pi(\varepsilon)d\varepsilon. \tag{19}$$

To derive (17), we use (19) to show that firms invest in at most 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 (19) is single-peaked in v, then there exists a unique value of v, denoted by v_t, that achieves this minimum. So, firms invest in at most one type of capital. Note that minimizing the right side of (19) is equivalent to solving (17). Finally, if $X_t(v_t) > 0$, then $\mu_t(v_t) = 0$, which yields (18) establishing the proposition.

To interpret this proposition, note that the first-order condition for the problem in (17) is

$$\sum\nolimits_{\tau=1}^{\infty} Q_{t,t+\tau}(1-\delta)^{\tau-1} \int_{\underline{\varepsilon}_{t,t+\tau}}^{\infty} A_t \varepsilon \frac{\partial f(v_t)}{\partial v_i} \pi(\varepsilon) d\varepsilon = \sum\nolimits_{\tau=1}^{\infty} Q_{t,t+\tau}(1-\delta)^{\tau-1} \int_{\underline{\varepsilon}_{t,t+\tau}}^{\infty} \chi_{it+\tau} \pi(\varepsilon) d\varepsilon. \quad (20)$$

The left side of (20) is the present value of the output from making a machine marginally more intensive in labor of type i. Firms equate this present value to the shadow value of type-i workers assigned to the machine—the right side of (20). These two present values are discounted by the price of output $Q_{t,t+\tau}$ in $t+\tau$ multiplied by the share of capital remaining $(1-\delta)^{\tau-1}$ and accounting for the utilization rate of the machine changing over time according to the cutoff rule $\underline{\varepsilon}_{t,t+\tau}$.

3.3 Equilibrium Along the Balanced Growth Path

We start with the analysis of the model's 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 consumption along the BGP. We scale the disutility of work and labor market search by $(1+g)^t$ so that as consumption and wages grow at rate g along the BGP, 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 $\widetilde{x}$, we denote the detrended version of a variable x, say, detrended wages are $\widetilde{w}_{it} = w_{it}(1+g)^{-t}$, 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 the cutoff level ε, $\Pi^p(\underline{\varepsilon}) = \int_{\underline{\varepsilon}}^{\infty} \varepsilon \pi(\varepsilon) d\varepsilon$ denote the average idiosyncratic productivity of utilized capital, and $\widetilde{\beta} = \beta/(1+g)$, we can establish the following result.

Lemma 3. *Along the BGP, labor allocations and wages are determined by*

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

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

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 that satisfy $\widetilde{\chi}_i = \frac{\partial f(\widetilde{v})}{\partial \widetilde{v}_i} m(\underline{\varepsilon}_1)$*;*

d) free entry for job vacancies and wages that 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 that satisfies $h'(s_i) = \beta\lambda_w(\theta_i)[\widetilde{w}_i - v'(n_i)]/[1-\beta(1-\sigma)]$*;*
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* $\sigma n_i = \lambda_w(\theta_i)s_i$*.*[17]

Condition *c*) for firms' choice of the labor intensities of new capital equates the shadow value of type-i labor, $\widetilde{\chi}_i$, to the output from making a machine marginally more intensive in type-i labor, $\partial f(\widetilde{v})/\partial\widetilde{v}_i$, adjusted for changes in utilization and relative vintage productivity over the life of the machine, as captured by $m(\underline{\varepsilon}_1)$. Since this utilization adjustment is a feature of the capital stock, it is common to all worker types i so $m(\underline{\varepsilon}_1)$ does not depend on i.

Taking ratios of condition *c*) across workers of types i and i' yields $\widetilde{\chi}_i/\widetilde{\chi}_{i'} = f_i(\widetilde{v})/f_{i'}(\widetilde{v})$. In our quantitative work, we find that the shadow value of type-i workers $\widetilde{\chi}_i$ approximately satisfies $\widetilde{\chi}_i \approx (1+1/\omega)\widetilde{w}_i$ along the BGP. Combining these last two equations gives that the ratio of the wages of workers of types i and i' approximately satisfies $\widetilde{w}_i/\widetilde{w}_{i'} \approx f_i(\widetilde{v})/f_{i'}(\widetilde{v})$. Hence, the long-run elasticities of substitution across different types of workers are disciplined only by the putty-putty production function $f(v)$: the putty-clay aspect of production is irrelevant for the long-run substitutability across workers in production. Intuitively, 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. In turn, 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 labor literature to estimate the elasticity of substitution across different worker groups. We exploit this fact to rely on existing estimates of labor-labor substitutability to pin down the parameters of $G(N)$.

Notice that the equilibrium conditions of Lemma 3 define a system of $1+6N$ equations in $1+6N$ variables $\underline{\varepsilon}_1$ and $\{\widetilde{v}_i,\ n_i,\ \theta_i,\ s_i,\ a_i,\ \widetilde{w}_i\}$. In this precise sense, 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*, $\underline{\varepsilon}_1^c$ and $\{\widetilde{v}_i^c,\ n_i^c,\ \theta_i^c,\ s_i^c,\ a_i^c,\ \widetilde{w}_i^c\}$, as those that solve the same set of conditions except for those under *d*), which are replaced by

$$\kappa_i = \beta\lambda_f(\theta_i)\frac{f_i(\widetilde{v})m(\underline{\varepsilon}_1) - \widetilde{w}_i}{1-\beta(1-\sigma)} \text{ and } \widetilde{w}_i = \eta f_i(\widetilde{v})m(\underline{\varepsilon}_1) + (1-\eta)v'(n_i). \tag{21}$$

Finally, it is also useful to note for later that the cutoff level of idiosyncratic productivity $\underline{\varepsilon}_1$ for

[17] This labor market clearing condition is derived by dividing the (level) version of this equation for i, $\mu_i n_i = \sum_{\tau=1}^{\infty}\Pi^u(\underline{\varepsilon}_\tau)\widetilde{v}_i(1-\delta)^{\tau-1}\widetilde{X}$, by the corresponding one for type 1.

optimal capital utilization is the same under the monopsonistic and the competitive allocations, since conditions *a*) and *b*) of Lemma 3 determine this cutoff independently of the other conditions.

Lemma 4. *On the BGP, the utilization threshold $\underline{\varepsilon}_1$ is independent of monopsony power ($\underline{\varepsilon}_1 = \underline{\varepsilon}_1^c$).*

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 4, this lemma will be helpful because it also implies that the utilization schedule of capital is independent of labor market policies, which helps shed light on why their long-run effects are similar to those predicted by the putty-putty model.

4 Understanding Labor Market Policies

We now turn to the analysis of canonical labor market policies to illustrate the use of our framework. To motivate this discussion, we first show that a combination of worker type-specific minimum wages and vacancy-posting subsidies can eliminate monopsony distortions. Although these policies provide a useful benchmark, they are difficult to implement in practice because they require targeting each worker type. Therefore, we analyze two simple policies that are often used 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 qualitatively explore how the policies affect outcomes in the long run.

4.1 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 markets 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 such type'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 ω that describes the imperfect substitutability of jobs in workers' preferences enters the firm's first-order conditions associated with the choices of wages W_{ijt+1} and market tightness θ_{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 θ_{ijt} is novel and reflects the idea that as more vacancies are posted—holding wages fixed—market tightness increases. Hence, in this sense, firms' monopsony power distorts both the *supply* and the *demand* for labor. To build further intuition, consider the participation constraint along the BGP,

$$\frac{\lambda_w(\theta_{ij})}{\rho+\sigma}\left[\widetilde{w}_{ij} - v'(n_i)\left(\frac{n_{ij}}{n_i}\right)^{\frac{1}{\omega}}\right] \geq \widetilde{\mathcal{W}}_i, \tag{22}$$

where $\rho = \frac{1}{\beta} - 1$ is the rate of time preference, $\widetilde{w}_{ij}$ is the detrended flow wage of type-i workers at firm j, and $v'(n_i)$ is their 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 can attract such workers, and thus satisfy the participation constraint in (22) for them, in two ways. First, it can raise the flow wage it pays. At a symmetric equilibrium, the required wage increase—obtained by differentiating (22) holding as an equality with respect to $\widetilde{w}_{ij}$ and n_{ij} holding $\lambda_w(\theta_{ij})$ fixed—gives

$$\frac{d\widetilde{w}_{ij}}{dn_{ij}} = \frac{v'(n_i)}{\omega n_i} > 0. \tag{23}$$

Online Appendix A converts (23) into an elasticity showing that $d\log n_{ij}/d\log \widetilde{w}_{ij} \approx \omega$. Thus, the parameter controlling the degree of monopsony power can, loosely speaking, be interpreted as a firm-specific labor supply elasticity in response to permanent changes in wage offers along the BGP.

Second, firm j can attract additional type-i workers by posting more vacancies, thus raising the job-finding probability $\lambda_w(\theta_{ij})$. At a symmetric equilibrium, the required increase in $\lambda_w(\theta_{ij})$—obtained by differentiating (22) with respect to $\lambda_w(\theta_{ij})$ and n_{ij} holding $\widetilde{w}_{ij}$ fixed—gives

$$\frac{d\lambda_w(\theta_{ij})}{dn_{ij}} = \frac{(\lambda_w(\theta_{ij})/\omega)(v'(n_i)/n_i)}{\widetilde{w}_i - v'(n_i)} > 0. \tag{24}$$

Taken together, (23) and (24) 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 along 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 $\widetilde{w}$ to their marginal product. As shown in Online Appendix A, along the BGP this markdown is

$$\frac{\widetilde{w}_i}{f_i(\widetilde{v})m(\underline{\varepsilon}_1)} = \left[1 + \underbrace{\frac{1}{\omega} \times \frac{v'(n_i)}{v'(n_i) + \frac{\eta}{1-\eta}(\rho+\sigma)\frac{\kappa_i}{\lambda_f(\theta_i)}}}_{\text{monopsony component}} + \underbrace{\frac{(\rho+\sigma)\frac{\kappa_i}{\lambda_f(\theta_i)}}{v'(n_i) + \frac{\eta}{1-\eta}(\rho+\sigma)\frac{\kappa_i}{\lambda_f(\theta_i)}}}_{\text{efficient component}}\right]^{-1}. \tag{25}$$

The *monopsony component* reflects firms' monopsony power, which distorts allocations relative to the 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 cost $\kappa_i/\lambda_f(\theta_{ij})$ 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 cost, $(\rho+\sigma)\kappa_i/\lambda_f(\theta_i)$, which we find to be quantitatively small.

4.2 Scope for Policy

In our economy, monopsony power leads wages and vacancy posting to be too low for every worker type. In this context, as we now show, an ideal policy that achieves efficient outcomes consists of both type-specific minimum wages and type-specific vacancy-posting subsidies.

Proposition 5. *In the BGP of the symmetric monopsonistically competitive search equilibrium, the efficient allocations can be implemented through the 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 τ_i to vacancy posting for each worker type i such that $1-\tau_i = [\widetilde{w}_i^c - v'(n_i^c)]/[\widetilde{w}_i^c - v'(n_i^c)(1-1/\omega)]$.*

Intuitively, the competitive equilibrium differs from the monopsonistically competitive equilibrium only in the free-entry condition and the wage equation. So, if the government sets the minimum wage to its competitive level for each type, since firms wish to lower wages below their competitive levels, the minimum wage constraints would all bind. In this case, the monopsony wage equation would be replaced by $\underline{w}_i = \widetilde{w}_i^c$, which would fix the distortions to wages. But this policy alone would not eliminate the distortions to the vacancy-posting condition from the $v'(n_i^c)/\omega$ term. To do so, we can choose type-specific τ_i entry subsidies so that the free-entry condition

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

under these policies becomes $\kappa_i = \lambda_f(\theta_i^c)\frac{f_i(\widetilde{v}^c)m(\underline{\varepsilon}_1)-\widetilde{w}_i^c}{\rho+\sigma}$ as in the efficient equilibrium. This simply requires setting τ_i as specified in the proposition, which completes its proof. By allowing for time-varying minimum wages and subsidies to vacancy creation, the same result holds outside the BGP.

4.3 Minimum Wage

We turn to studying 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 wage paid to 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 and are denoted by $w_{ijt}^0 = (1+g)^t w_{ij}^0$.

Intuition from the BGP. As Proposition 5 suggests, increasing the minimum wage up to a worker's competitive wage, which would bind for all workers, encourages firms to hire efficiently

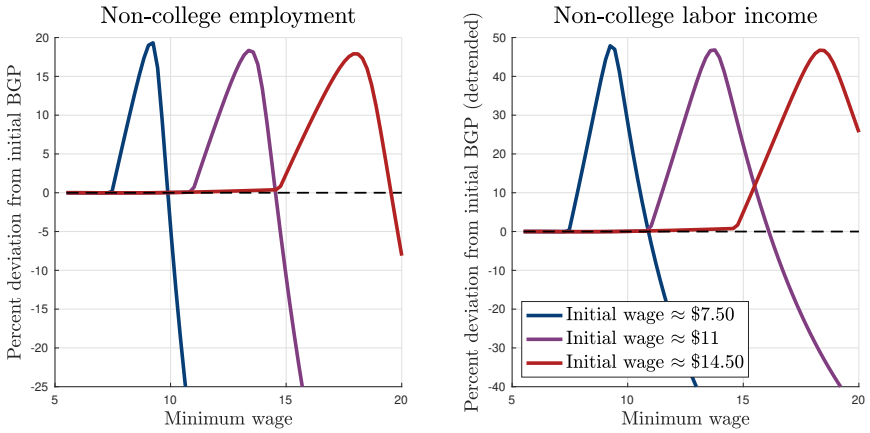
because it eliminates the monopsony force whereby a firm takes into account that hiring a marginal worker raises the wages of all inframarginal workers and so sets too low a wage. For increases in the minimum wage above a worker's competitive wage, however, the minimum wage discourages firms from hiring such a worker because it requires a firm to pay the worker more than the worker's marginal product. Now consider a given minimum wage. Intuitively, for worker types for whom the minimum wage is only slightly larger than their current wages—namely, the minimum wage is set no higher than what these workers would be paid in the competitive search equilibrium—monopsony distortions are reduced, so hiring increases. However, for worker types for whom the minimum wage is much larger than their current wages—namely, above their wages in the competitive search equilibrium—the minimum wage discourages firms from hiring them.

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 the employment of three types of non-college workers in the new BGP.[18] These lines, generated by our quantitative 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 w_i^c in the associated competitive search equilibrium. 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, roughly at 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 in Figure 3 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 higher-productivity workers but reduce employment and income of the lowest-productivity ones.

[18] Throughout, we index each of these minimum wage levels by the equivalent hourly wage that would bind for the same proportion of workers in the 2017-2019 ACS data to which we benchmark our model.

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.

4.4 Targeted Transfers

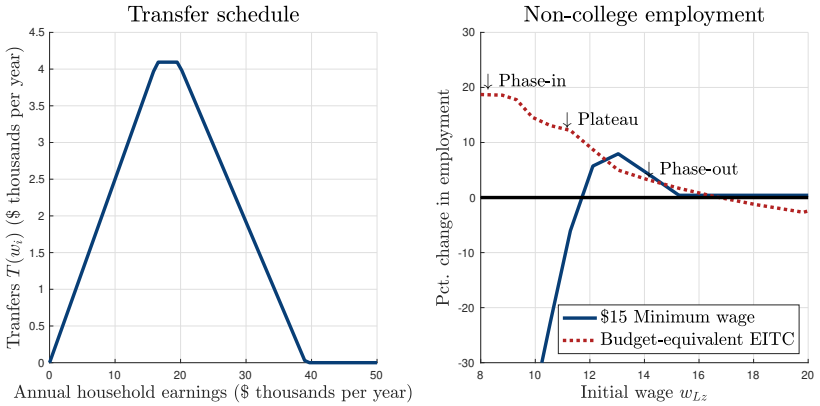
We turn to the long-run effects of transfer payments to households conditional on working. These transfer payments can be more targeted to specific worker types than a minimum wage and, therefore, be more effective at redistributing resources to their intended beneficiaries. In the long run, transfer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

Intuition from the BGP. We consider a general transfer policy such that a worker earning the flow wage w_{ijt} in t receives an additional payment $T_t(w_{ijt})$ so the worker's after-transfer earnings are $A_t(w_{ijt}) = w_{ijt} + T_t(w_{ijt})$. 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, which implies that the profit tax does not affect any of the firms' decisions. The first-order condition for households' search intensity s_{ijt} becomes

$$-\frac{U_{sit}}{U_{cit}} = Q_{t,t+1}\lambda_w(\theta_{ijt})(W^H_{ijt+1} + V_{ijt+1}) \text{ if } s_{ijt} > 0, \tag{26}$$

where $W^H_{ijt+1} = A_{t+1}(w_{ijt+1}) + Q_{t+1,t+2}(1-\sigma)A_{t+2}(w_{ijt+2}) + Q_{t+1,t+3}(1-\sigma)^2 A_{t+3}(w_{ijt+3}) + \ldots$ 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 (26), 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' labor market 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

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 worker 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.

effect lowers the shadow price of workers and reduces monopsony distortions. We next summarize the effect of the policy along the new BGP.

Lemma 6. *Under the transfer policy, labor allocations in the monopsonistically competitive search equilibrium solve the equations in Lemma 3 with the optimal search condition replaced by*

$$h'(s_i) = \frac{\lambda_w(\theta_i)}{\rho + \sigma} \left[A(w_i) - v'(n_i) \right] \tag{27}$$

and the vacancy posting condition replaced by

$$\frac{\kappa_i}{\lambda_f(\theta_i)} = \frac{1}{\rho + \sigma} \left[\chi_i - w_i - \frac{v'(n_i)}{\omega A'(w_i)} \right]. \tag{28}$$

A positive *level* of transfers $A(w_i) > w_i$ stimulates workers' job search in (27) by increasing the marginal gain to households of searching for jobs. Also, a positive *marginal* transfer $A'(w_i)$ in (28), 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, over 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 the (detrended) transfer schedule plotted in the left panel of Figure 4. We make this schedule budget-equivalent to a $15 minimum wage by choosing the rate of the linear profit tax that funds the EITC to raise tax revenues equal to firms' lost profits due to the minimum wage. The resulting EITC is such that 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 as income increases with a negative marginal transfer rate of 22%.

Note that 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. In the phase-out region, however, households face a positive average subsidy rate but a negative marginal subsidy rate, since the credit is being phased out.

The right panel of Figure 4 compares the effects of the EITC and a $15 minimum wage 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 effects contribute to higher long-run employment rates for this group. Workers in the plateau region still face a positive average transfer rate but their marginal transfer rate is zero, so their monopsony distortion is not impacted, leading to a smaller increase in employment. Yet, workers in the phase-out region face a negative marginal transfer rate, which exacerbates the monopsony distortion, leading to a decline in their employment.

Overall, although the EITC compounds monopsony distortions for high-wage workers, Figure 4 shows that the EITC is much more effective at stimulating the employment of non-college workers earning low wages, as it is more effective at correcting their monopsony distortions.

5 Quantification

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

5.1 Parametrization

We assume that a 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. We assume that z is drawn from an education-specific log-normal distribution with mean zero and variance σ_z^e. A type-i family has preferences within the Greenwood, Hercowitz and Huffman (1988) class of the form

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

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

$$G(N) = \left[\xi N_L(N)^{\frac{\varphi-1}{\varphi}} + (1-\xi) N_H(N)^{\frac{\varphi-1}{\varphi}}\right]^{\frac{\varphi}{\varphi-1}} \text{ and } N_e(N) = \left[\sum\nolimits_{i \in I_e} z_i (N_i)^{\frac{\phi-1}{\phi}}\right]^{\frac{\phi}{\phi-1}}. \tag{29}$$

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)).

5.2 Disciplining Key Features

We begin by providing intuition for how we parameterize the features of our model related to the three main forces that govern 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 control how much firms eventually substitute across different types of workers; and *iii*) the degree of firm monopsony power, which determines the size of the distortions that labor market policies can potentially correct.

Capital Utilization. The dynamics of capital utilization are shaped by the growth rate of the economy g and the depreciation rate of capital δ, which jointly determine the average utilization rate of different vintages of capital, together with the dispersion of idiosyncratic capital productivity σ_ε, which determines the share of capital utilized within a vintage of capital. We set the growth rate of the productivity of new capital to induce an average aggregate growth rate g of 2% per year and the capital depreciation rate to 15% annually, which matches the aggregate capital depreciation rate excluding structures. We choose the within-vintage dispersion parameter σ_ε to reproduce 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 the output actually produced divided by the output that would be produced if all capital was fully utilized, namely,

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

where $\Pi^p(\underline{\varepsilon})$ is the *average productivity of utilized capital.* The higher σ_ε is, the more capital is below the utilization threshold, and so the lower capacity utilization is; see Figure 2. We choose σ_ε so that the average capacity utilization rate is 75%, which is its approximate value in the post-2000 data.

Monopsony Power. We choose the degree of firm monopsony power ω to match existing estimates of wage markdowns. A growing literature has measured wage markdowns in the United States documenting 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 (25), wage markdowns in our model reflect both the degree of firm monopsony power ω 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 ω to reproduce the average wage markdown. This strategy implies that average hiring costs $\kappa_i/\lambda_f(\theta_i)$ are approximately 64% of one month's wage.[19] Given that we assume a job-separation rate of 2.8% per month and a 2% annual rate of time preference, the annuitized value of hiring costs that a firm must recoup each period is only $(\rho + \sigma)\kappa_i/\lambda_f(\theta_i) \approx 0.03 \times 0.64 = 1.9\%$ of one month's wage. Hence, the efficient component of the markdown is relatively small. Through the lens of our model, this result implies a large inefficient component of the markdown, yielding $\omega = 3.07$.

Long-Run Elasticities of Substitution. The long-run elasticities of substitution between workers of different abilities in $G(N)$ are governed by ϕ, the elasticity of substitution within education groups, and φ, the elasticity of substitution across education groups. We set the cross-group elasticity φ to 1.4 and the within-group elasticity ϕ to 4, which are values in the range of the estimates in the labor literature as in Card and Lemieux (2001) and Borjas and Katz (2007). This work assumes that different observable groups of workers—for instance, by age or immigration status—correspond to different ability groups z and exploits variation in the labor supply of these groups to estimate these elasticities of substitution based on the implied variation in their wages. In particular, our value of the between-group elasticity φ is consistent with the benchmark estimate in Katz and Murphy (1992). A concern with borrowing estimates of ϕ and φ 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 5.4, we address this concern by showing that the estimators of these parameters common in the literature recover nearly exactly the true long-run elasticity based on data generated from our model.

5.3 Details

As for the remaining parameters, we externally fix a subset of them and then endogenously choose 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 worker productivity z—pinned down as

[19] 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. Our estimate belongs to the middle of this range.

TABLE 1: Fixed Parameters

Parameter	Description	Value
Households		
$N_H/(N_H+N_L)$	Share of college workers	0.31
β	Discount factor	$(\frac{1.04}{1+g})^{-1/12}$
γ_n	Labor supply elasticity	1.00
γ_s	Search supply elasticity	5.00
χ_s	Scale of search disutility	3.8×10^6
Production function		
φ	Long-run elasticity b/t N_{ht} and $N_{\ell t}$	1.40
ϕ	Long-run elasticity within education group	4.00
δ	Capital depreciation (annualized)	15%
g	Long-run growth rate (annualized)	2.0%
Labor market frictions		
σ	Job destruction rate	2.8%
η	Elasticity of matching function w.r.t. unemployed	0.50

explained below—jointly determine the weights μ_i. We choose a value for the discount factor β of $[1.04/(1+g)]^{1/12}$ so that the annualized real interest rate r equals 4%. We set the parameter γ_n of the utility function, which primarily controls the elasticity of labor supply, to 1. There exists a locus of values for the parameters γ_s and ψ_s governing the disutility of labor market search that imply approximately identical labor market outcomes along the BGP but different responses of search effort to changes in labor market policy. 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 lead to a relatively muted response of search effort to the minimum wage as in the data (see Adams, Meer and Sloan (2022)).[20] We set the exogenous separation rate σ to 2.8% per month, which is the Abowd-Zellner corrected estimate of the separation rate in 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 η to 0.5 as in Ljungqvist and Sargent (2017).

Fitted Parameters. We set the remaining parameters in Table 2 panel A to match the statistics in panel B. The parameters of the worker productivity distribution and the production function govern the distribution of income in the economy. We choose the dispersion of productivity σ_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 ξ placed on non-college labor in $G(N)$ in the long-run production function determines the share of total labor income accruing to non-college workers, whereas the Cobb-Douglas coefficient $1-\alpha$ on labor in production controls labor's total share of income in aggregate output. We pin down these parameters accordingly.

Hiring costs $\kappa_{it}/\lambda_f(\theta_{it})$ are pinned down by the size of vacancy-posting costs (κ_0 in $\kappa_{it} = \kappa_0(1+$

[20] 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 and Targeted Moments

Panel A

Parameter	Description	Value
Worker Productivity Distribution $\log\mathcal{N}(0,\sigma_z^e)$		
σ_z^L	SD of non-college z	0.69
σ_z^H	SD of college z	0.84
Production Function		
α	Exponent on capital in production	0.24
ξ	Weight on non-college labor in production	0.42
Monopsony		
ω	Monopsony power	3.07
Search Frictions		
B	Matching function productivity	0.50
κ_0	Vacancy-posting costs (normalization)	0.32
Labor Disutility		
ψ_n^L	Weight of non-college labor disutility	3.84
ψ_n^H	Weight of college labor disutility	4.79
Capital Utilization		
σ_ε	Dispersion in capital productivity	0.25

Panel B

Statistic	Description	Data	Model
Average Wage Markdown			
$\mathbb{E}[w_i]/\mathbb{E}[\chi_i]$	Average wage markdown	0.75	0.75
Wage Distribution			
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			
$\mathbb{E}[w_iN_i]/Y$	Aggregate labor share	0.57	0.57
$\mathbb{E}[w_z^HN_z^H]/\mathbb{E}[w_iN_i]$	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: Parameters in panel A are chosen to match the statistics reported in panel B. The 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 employment rates are calculated using ACS data (2017–2019). The aggregate labor share is from Karabarbounis and Neiman (2014). The average unemployment rate is from the BLS. The capacity utilization rate is drawn from the U.S. Census Quarterly Survey of Plant Capacity Utilization (QPC).

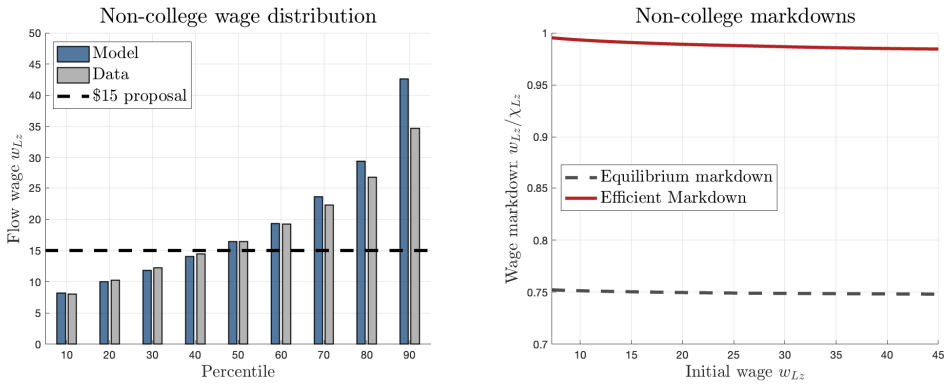
$g)^t z_i$) and the efficiency of the matching function (B in $m(a_{it}, s_{it}) = Bs_{it}^{\eta}a_{it}^{1-\eta}$) since $\lambda_f(\theta_{it}) = B\theta_{it}^{-\eta}$. As is typical in search models, the size of the vacancy-posting costs κ_0 is not separately identified from the efficiency B of the 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 κ_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, which is very close 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.

5.4 Model Fit and Validation

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

Wage Distribution and Markdowns. Table 2 shows that 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 largely matches the 50-10 wage ratio within each education group despite our parsimonious parametrization of the productivity distribution of z—specified as log-normally distributed for each education group with mean normalized to 1. The left panel of Figure 5 compares the wage distribution of non-college workers in the model with that in the ACS data. By targeting the left tail of this distribution, we ensure that a $15 minimum wage binds for the same fraction of non-

FIGURE 5: Non-College Wages and Wage Markdowns



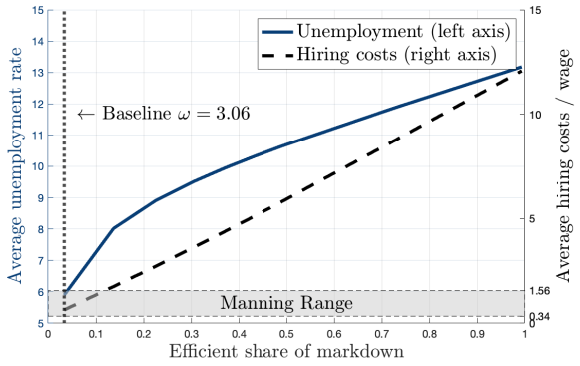
Notes: The left panel plots deciles of the non-college wage distribution in our model and the ACS data (2017-2019). The right panel plots the steady-state wage markdown w_{Lz}/χ_{Lz} for select non-college worker types. "Equilibrium markdown" corresponds to our baseline model. "Efficient markdown" corresponds 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.

college-educated workers in the model as in the data, 45%. The model slightly overpredicts wages at the top of the wage distribution for non-college workers, but they are barely affected by our experiments. The right panel of Figure 5 shows that the wage markdown for non-college workers is nearly constant at the value of 0.75 across workers with different wages. As explained below, this constancy of markdowns across workers is an important reason why our model implies values for ϕ in line with existing estimates based on static perfectly competitive models of the labor market.

Hiring Costs and Wage Markdown. In our benchmark model, we target an average markdown of 0.75 and a steady state unemployment rate of 5.9% to infer how much of the markdown is due to the inefficient component. In this subsection, we explore the identification of the inefficient component by examining how the average unemployment rate and average hiring costs change as we vary the efficient share of the markdown. Specifically, in Figure 6, we vary the degree of firm monopsony power ω and adjust the efficiency of the matching function B accordingly so as to match an average markdown of 0.75. Different combinations of ω and B imply different levels of the efficient component of the markdown—this is the variation shown along the x-axis of the figure. At our baseline parameterization (the dashed vertical line), the efficient component of the markdown is about 4%, the implied steady-state unemployment rate (solid line, left axis) is 5.9%, and the average hiring cost (dashed line, right axis) is 0.64 of one month's wage.

The range of estimates for hiring costs from Manning (2011)'s survey is shown in the shaded area of the figure. At the upper end of this range, the efficient share of the markdown increases to 10% but the implied monopsony power decreases only slightly as ω increases from 3.06 to 3.3. As the figure shows, this level of hiring costs seems implausible as it leads to an increase in the

FIGURE 6: Implications of Alternative Decompositions of the Wage Markdown



Notes: Steady-state moments for different shares of the total wage markdown accruing to the efficient component in (25). We calculate these moments by varying the degree of monopsony power ω and choosing the efficiency of the matching function B to match the total average markdown of 0.75. The x-axis plots the implied share of the markdown due to the efficient component. The left y-axis (blue) plots the steady-state unemployment rate and the right y-axis (black) plots average hiring costs relative to one month's wage. The shaded region includes the range of estimates of hiring costs from Manning (2011).

unemployment rate from its baseline value of 5.9% to 8% (solid blue line), which is well outside of its historical average. Put differently, a low steady state unemployment rate implies that hiring costs cannot be too large, all else equal. For example, increasing the share of the efficient component of the markdown to 50% increases the average hiring cost to six times one month's wage, which is far outside Manning (2011)'s range, and simultaneously raises the unemployment rate to over 10.5%. In summary, to match an unemployment rate of about 6%, hiring costs can only be about two-third's of a worker's monthly wage. Thus, it is hard to argue that plausible hiring costs would lead us to significantly revise downward the degree of monopsony power that we infer.

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, φ and ϕ. 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 putty-clay model. Nonetheless, we show that the estimators used in the literature recover the true values of ϕ of our long-run production function when applied to data generated from our model in response to the same exogenous variation and time aggregated as in Card and Lemieux (2001). This result obtains because these estimators 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.

In terms of the within-education group elasticity of substitution across workers, ϕ, which is

a key parameter, 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 $G(N)$.[21] For each education group, these authors assume that workers with different experience (age) differ in their productivity, but all workers with a given level of experience share the same productivity. 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 across workers with different productivity within an education group, $\widehat{\phi}$, exploits the variation in their relative wages induced by changes in the relative supply of workers with different experience over time, which they interpret as exogenous.

To validate our choice of ϕ, we perform the following exercise. 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 $\widehat{\phi}$ of ϕ. 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

$$\frac{w_{it}}{w_{i't}} = \frac{z_i}{z_{i'}} \left(\frac{N_{it}}{N_{i't}} \right)^{-\frac{1}{\phi}} \quad \text{or} \quad \Delta \log \frac{w_{it}}{w_{i't}} = -\frac{1}{\widehat{\phi}} \Delta \log \frac{N_{it}}{N_{i't}}, \tag{30}$$

where Δ denotes the difference operator across two time periods. By (30), we can construct a 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 E for details. 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 ϕ in our model.

One reason that this estimator may not perform well on data from our model is that workers are not paid their marginal products even in the long run when putty-clay frictions are irrelevant. Quantitatively, this concern is unwarranted because wage markdowns are nearly constant across workers with different productivity along the BGP. Hence, the ratios of wages of workers with different productivity approximately equal the ratios of their marginal products. Another concern is that using just five-year differences in employment as in Card and Lemieux (2001) may severely bias downward our estimate of the degree of substitutability across workers, since our technology

[21] Our results are not sensitive to the across-group elasticity φ, 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.

is Leontief in the short run. This concern is also unwarranted in practice because by varying the measure of families of each type, we modify the *supply* of workers to replicate Card and Lemieux (2001)'s measured changes in the employment rate of workers with different productivity within an education group. Since the incoming workers features a different skill mix than that of the existing supply of workers—which is embedded in the labor intensity of installed capital—as more workers are college-educated, firms must adjust the ratios v on new machines to accommodate them. This adjustment of labor-to-capital ratios is governed by the long-run function $f(v)$, as illustrated in Figure 1, which only reflects the *long-run* substitutability across workers captured by ϕ.

6 Dynamic Effects of Labor Market Policies

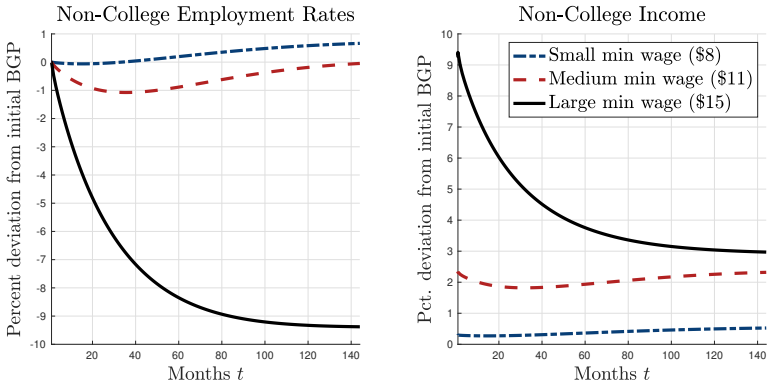
We use our quantitative model to study the dynamic effects of labor market policies. In particular, we highlight the importance of our model's key features in explaining the resulting transition dynamics. Our main result is that the transition to a new long-run equilibrium is slow, which implies that the short-run impacts of policies that the empirical literature mostly focuses on are not very informative about their *overall* effects. We begin by studying the minimum wage to illustrate the workings of our model. We then show that similar insights apply to the EITC.

6.1 Overview of the Dynamic Effects of the Minimum Wage

We assume that the economy starts on the 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 does not bind for any worker. 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, firms and workers perfectly anticipate the resulting transition path to the new BGP. In period 0, a firm may fire any worker. For retained workers, a firm must increase the flow wages of workers whose flow wages are initially below the new minimum to at least the new minimum and must honor its wage commitments from the initial BGP to workers for whom the minimum wage does not bind, 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 effects underlying these aggregate ones.

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.

FIGURE 7: 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 a minimum wage of various sizes, expressed relative to the initial BGP.

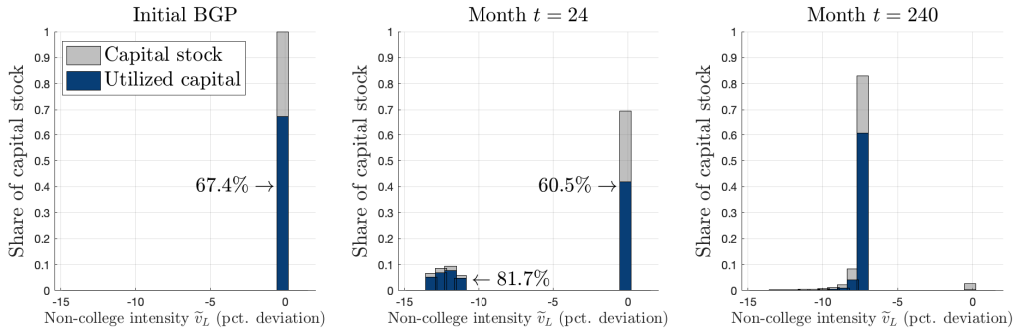
The left panel of Figure 7 illustrates our main result that aggregate non-college employment adjusts slowly to any such policy change. The small minimum wage induces firms to increase employment by reducing monopsony distortions for low-productivity workers, but it takes time to do so. By contrast, the 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 change in employment is only a small fraction of its long-run change.[22] The right panel of Figure 7 illustrates that this 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 flow labor 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}, which occurs 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, this slow adjustment of employment delays the total long-run *benefits* of the minimum wage on labor 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, these workers' labor income $w_{it}n_{it}$ slowly declines. We show below that this dynamics of labor income is 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

[22]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 8: Distribution of Capital Types Along the Transition After $15 Minimum Wage



Notes: The gray portion of each histogram represents the share of capital for the corresponding range of detrended non-college labor intensities $\widetilde{v}_L = \sum_{i \in L} \widetilde{v}_i$. 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.

workers in the short run, over time, this gain is eroded as their employment progressively falls. As we elaborate on later, the short-run effects of a large increase in the minimum wage greatly overstate its total benefits. At the same time, the long-run effects of such a policy overstate the total costs of it, because they altogether miss the short-run benefits.

6.2 Understanding the Slow Employment Dynamics

To understand the model's slow employment dynamics, it is useful to start by considering how the capital stock gradually shifts from its original non-college labor intensity on the initial BGP to a significantly lower one along the transition. Figure 8 plots the distribution of capital types in the initial BGP as well as after 2 and 20 years 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 utilized—the blue bars. As the left panel shows, in the initial BGP, all capital has the same detrended labor intensity $\widetilde{v}_L^*$, and firms utilize roughly 68% of the different 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.

As the middle panel illustrates, after 2 years, over 30% of the original capital has depreciated and the utilization rate for the remaining capital has also declined. The new capital has lower non-college labor intensity and a very high utilization rate. As the right panel shows, after 20 years, almost all of the original capital has depreciated, and the economy has nearly finished its transition to a new BGP in which capital features a much lower non-college labor intensity.

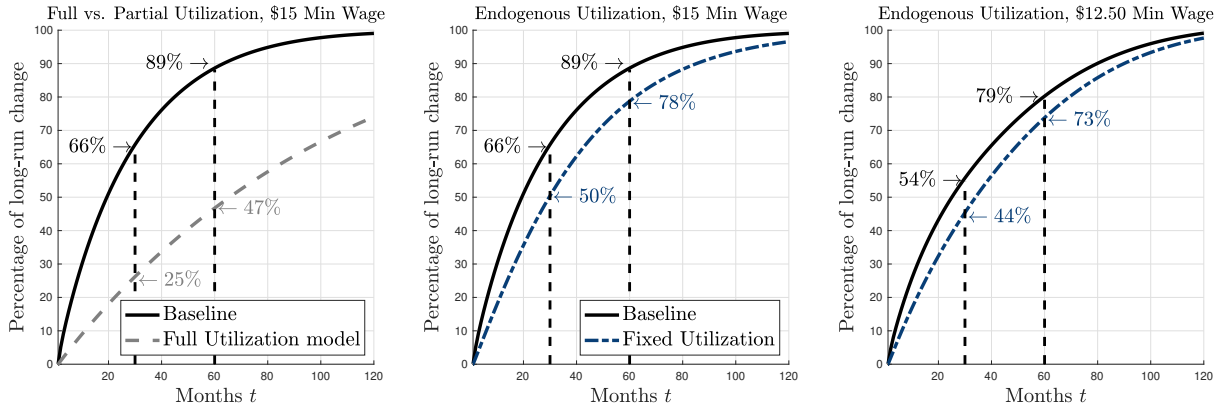
We now turn to analyzing in more detail how our model generates slow employment dynamics and explain the role of the vintage capital structure. Recall that in our baseline model, the average productivity of capital of vintage t is $(1+g)^t$ times larger than that of capital built in period 0 and

each unit of capital of any vintage has permanent idiosyncratic productivity ε with a density with mean 1. The *full-utilization* model is defined as the baseline model with one change: our baseline settings for g and $\pi(\varepsilon)$ are replaced by $g = 0$ and a density concentrated on $\varepsilon = 1$; see Figure 2. Thus, in this model, it is optimal for firms to have a 100% utilization rate of capital, both along the BGP and for all the minimum wage experiments that we conduct—for high enough minimum wages, the utilization rate falls below 100%. In our baseline model, capital utilization declines with age along the BGP with no minimum wage, and a higher minimum wage leads the rate of capital utilization to fall endogenously along the transition. For both models, the *speed of adjustment* is defined as the percentage change in non-college employment in the t-th month following the introduction of the minimum wage, relative to its long-run change between the new and the old BGP.

The left panel of Figure 9 shows that the speed of adjustment to a \$15 minimum wage in our baseline model (solid line) is significantly faster than in the model with full utilization (dashed line). In particular, in the full-utilization model, over the first 30 months after the introduction of a \$15 minimum wage, only 25% of the eventual change in non-college employment has occurred, and by the 60th month, only 47% has occurred. With full utilization, a firm has two margins of adjustment. For the old capital from the original BGP, a firm's only margin of adjustment is to wait for this capital to depreciate before replacing workers. Indeed, since workers separate at the rate of 30% per year and capital depreciates at the rate of only 15% per year, each year a firm actively hires 15% of each type of workers assigned to the old capital—no matter how low their skill level—to keep this capital running at full capacity. Hence, if a firm wishes to fully utilize all of its old capital, then it needs to continue actively hiring low-skill workers in the short run in the same proportion as it did in the original BGP. The second margin of adjustment is that a firm can invest in new types of capital that require a much lower fraction of low-skill workers. Thus, a firm can change the skill mix of its workers only through this second margin, which takes time.

The left panel also shows that our baseline model with vintage capital generates a much faster transition in response to a \$15 minimum wage: over the first 30 months, 66% of the eventual change in non-college employment has occurred and by 60th month, 89% has occurred. One reason for this faster transition is that, even in the initial BGP, this model features *endogenous obsolescence* of capital in addition to standard capital depreciation. To elaborate, in the baseline model with endogenous utilization, capital with sufficiently low productivity $A\varepsilon$ ceases to be utilized, namely, it becomes endogenously obsolete. So it is as if in this model, the rate of reduction of the useful capital stock is the sum of the depreciation rate and the obsolescence rate, whereas in the full-utilization model, the rate of reduction of the useful capital stock is just the depreciation rate.

FIGURE 9: 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. The left panel compares effects of a $15 minimum wage in our baseline model and a full-utilization model parameterized to match the same targets. The middle panel compares the effects of a $15 minimum wage in our baseline to $\widehat{N}_{Lt} = \sum_{\tau=1}^{\infty} \sum_{i \in L} (1-\delta)^{\tau-1} X_{t-\tau} \Pi^u(\underline{\varepsilon}_\tau) v_{Lt-\tau}$, which holds the capital utilization schedule fixed at its BGP value. The right panel does the same for a $12.50 minimum wage.

Therefore, even if the obsolescence rate remains constant after the introduction of the minimum wage, capital turns over much faster in the vintage model than in the full-utilization model due to the steady-state obsolescence effect in the original BGP. The second, more subtle, reason why the transition is faster in the model with vintage capital is that utilization rates endogenously respond to the minimum wage. Indeed, as we saw in the middle panel of Figure 8 above, the utilization rate of existing capital declines after the imposition of the minimum wage.

We quantify these steady-state obsolescence and endogenous utilization effects as follows. Since $\Pi^u(\underline{\varepsilon}) = \int_{\underline{\varepsilon}}^{\infty} \pi(\varepsilon) d\varepsilon$ is the utilization rate for cutoff $\underline{\varepsilon}$ and $\underline{\varepsilon}_{t-\tau,t}$ is the cutoff ε for capital installed in period $t-\tau$ to be utilized in t, we can express the *employment* of type-i labor in t as

$$N_{it} = \sum_{\tau=1}^{\infty} (1-\delta)^{\tau-1} X_{t-\tau} \Pi^u(\underline{\varepsilon}_{t-\tau,t}) v_{it-\tau}, \tag{31}$$

which is the sum of the employed workers of type i assigned to current and past vintages of capital. The type-i labor allocated to capital made in $t-\tau$ is the product of the amount of capital of that vintage remaining in t, $(1-\delta)^{\tau-1} X_{t-\tau}$, the share of that vintage that is still utilized, $\Pi^u(\underline{\varepsilon}_{t-\tau,t})$, and the type-$i$ labor intensity of that vintage, $v_{it-\tau}$. Employment changes as any of these terms changes. Let $\widehat{N}_{it} = \sum_{\tau=1}^{\infty} (1-\delta)^{\tau-1} X_{t-\tau} \Pi^u(\underline{\varepsilon}_\tau) v_{it-\tau}$ be the *fixed partial-utilization* component of type-i labor, which is the analog of (31) but with utilization rates held fixed at their initial BGP levels, where $\underline{\varepsilon}_\tau$ is the utilization threshold along the BGP for capital installed τ periods ago. Thus, $\widehat{N}_{it}$ changes across periods only because of capital accumulation dynamics.

We can decompose the labor of each type i as $N_{it} = \widehat{N}_{it} + (N_{it} - \widehat{N}_{it})$ and define the associ-

ated aggregate non-college labor as $N_{Lt} = \sum_{i \in L} N_{it}$ and its fixed partial-utilization component as $\hat{N}_{Lt} = \sum_{i \in L} \hat{N}_{it}$. Then, we can express non-college labor as the sum of the fixed partial-utilization component and the residual, the *endogenous utilization* component. Hence,

$$\underbrace{N_{Lt}}_{\text{Labor}} = \underbrace{\widehat{N}_{Lt}}_{\text{Fixed Partial-Utilization Component}} + \underbrace{N_{Lt} - \widehat{N}_{Lt}}_{\text{Endogenous Utilization Component}} . \tag{32}$$

The middle panel of Figure 9 compares the transition dynamics of non-college employment (solid line) and its fixed partial-utilization component (dashed line) in response to a \$15 minimum wage. Clearly, the fixed partial-utilization component accounts for most of the increased speed of transition from the full-utilization model to the baseline model (in the left panel). For example, at 30 months, the fixed partial-utilization component accounts for over 76% of the 66% share (50%/66%) of the transition in the baseline model—the endogenous utilization component accounts for the rest. At 60 months, this component accounts for 88% (78%/89%). Hence, the obsolescence effect on past vintages of capital already present in the initial BGP in the baseline model accounts for the bulk of the faster speed of transition in the baseline model relative to the full-utilization model.

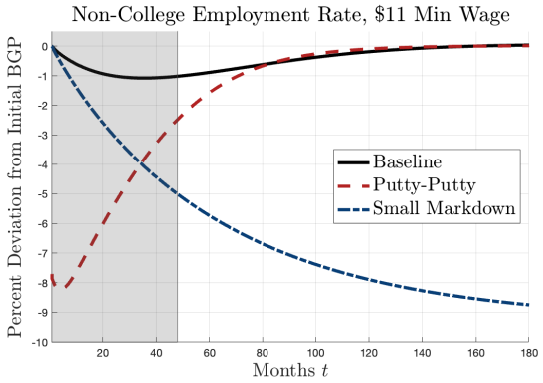
Next, from the middle and right panels, we see that the endogenous utilization component (dashed-dotted line) in the baseline model converge to zero over time because capital utilization eventually reverts to its initial value (by an extension of Lemma 3). Comparing the middle and right panels of Figure 9 further reveals that the endogenous utilization component is bigger for larger minimum wages: after 60 months, the difference between the transition paths of the baseline model and fixed-utilization model is 6 percentage points under a \$12.50 minimum wage (right panel) and 11 percentage points under a \$15 minimum wage (middle panel). Intuitively, the larger is the minimum wage, the greater is firms' incentive to idle more of the marginal capital stock.

Role of Monopsony and Putty-Clay Technology. The combination of putty-clay technology and firm monopsony power is essential for our model to generate slow employment responses to the minimum wage as in the data. Figure 10 illustrates this result for an \$11 minimum wage, which is within the set of minimum wage changes studied in the empirical literature. We shade the first four years of the figure because most of this literature focuses on employment effects over this horizon. Consistent with estimates in the literature, our baseline model implies small employment declines in response to modest increases in the minimum wage, as firms slowly substitute away from the lowest-productivity workers.

Consider now the response of non-college employment in the putty-putty version of our model with standard capital adjustment costs.[23] Here, even after modest minimum wage increases, firms

[23]Here X_{jt} units of investment generate output costs of $\zeta_t[A_t^{1/\alpha} X_{jt}/K_{jt} - (\delta + g_K)]^2 K_{jt}/2$, where g_K is the growth rate of the efficiency units of capital along the BGP and the parameter $\zeta_t = (1 + g_\zeta)^t \zeta$ grows at rate

FIGURE 10: Role of Putty-Clay and Firm Monopsony Power



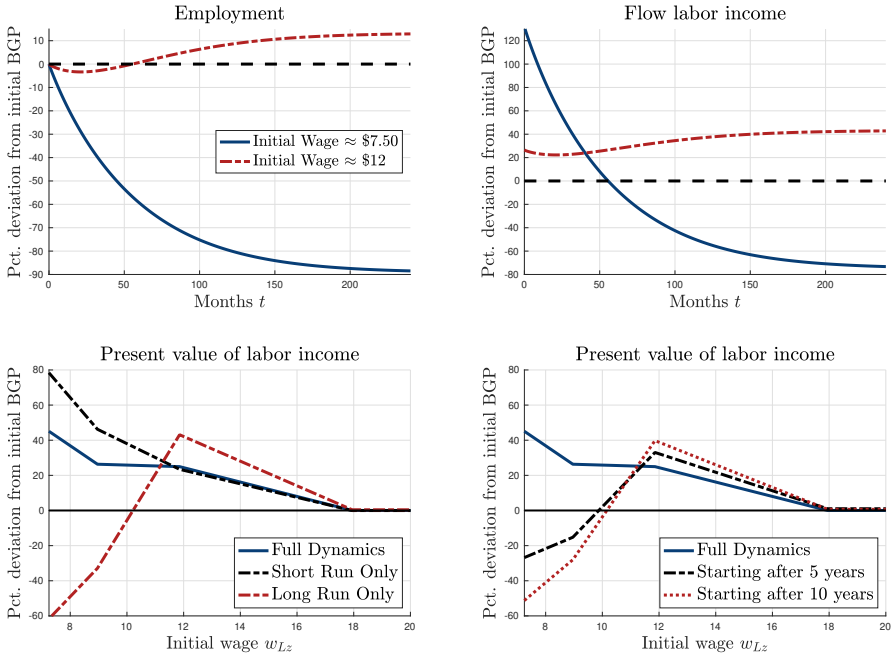
Notes: Transition paths of non-college employment after the introduction of an $11 minimum wage expressed relative to the initial BGP. "Baseline" corresponds to our baseline model (the same as in Figure 7). "Putty-putty" corresponds to the putty-putty version of our model from Appendix C with the adjustment cost parameter $\zeta = 3$. "Small markdown" corresponds to our putty-clay model with ω reparameterized to target an average wage markdown of 0.95 (instead of 0.75 as in our baseline).

substitute away from low-productivity workers by immediately firing them, which leads to an 8% decline in aggregate non-college employment upon impact—well outside the range of estimates in the data. In both the baseline and putty-putty models, employment slowly recovers over time as firms hire more of the workers for whom the minimum wage reduces monopsony distortions. It turns out that in the final steady state, this reallocation leaves aggregate non-college employment rate virtually unchanged in both versions of our model.

To highlight the importance of firm monopsony power for the slow response of employment to the minimum wage, consider next reducing firm monopsony power by choosing ω to target an average wage markdown of only 5% so that most of it is due to the efficient component. In this *small markdown* case, as the dash-dotted line in Figure 10 shows, an $11 minimum wage lowers steady-state non-college employment rate by nearly 9%. Thus, even with putty-clay capital, non-college employment declines sharply in the first few years after the introduction of the new minimum wage—again, outside the range of existing estimates. In other words, putty-clay models that abstract from monopsony power also imply counterfactually fast employment dynamics—it is the combination of putty-clay capital and monopsony power that allows our model to match the slow and small response of employment to the minimum wage.

$1 + g_\zeta = (1 + g_A)^{-1/\alpha}$. We set $\zeta = 3$. Note that capital adjustment costs slow down the adjustment of the efficiency units of capital K_t, but they do not prevent firms from adjusting the employment ratios of different worker groups because the production function is weakly separable between capital and labor, $F(K, \mathbf{N}) = K^\alpha G(\mathbf{N})$. As a result, labor-labor substitution within $G(\mathbf{N})$ would still occur even if adjustment costs $\zeta \to \infty$ so that K was constant. However, the slow adjustment of capital affects the overall level of labor demand, generating the small decline in non-college employment from the impact period to the first year after in Figure 10. See Appendix C for details.

FIGURE 11: Individual Employment and Labor Income Dynamics After $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 workers in our baseline model. *Full Dynamics* refers to the present value of labor income along the transition path in our baseline. *Short Run Only* calculates this present value assuming that labor income gains during the first two years persist over time; *Long Run Only* calculates this present value assuming that labor income gains in the new steady state persist over time (bottom left). *Starting after 5 years* and *Starting after 10 years* calculate this present value starting 5 or 10 years after the minimum wage is introduced (bottom right).

6.3 Heterogeneous Employment and Income Responses

To understand the implications of the slow dynamics of aggregate employment and labor income, the top left panel of Figure 11 illustrates how the employment rates n_{it} of a low-productivity non-college worker earning $7.50 and a medium-productivity one earning $12 change over time in response to a $15 minimum wage. On impact, the employment of both workers remains constant. Over time, though, firms substitute away from low-productivity workers leading to a gradual decrease in their employment. By contrast, firms substitute towards medium-productivity workers for whom monopsony distortions are reduced, but this adjustment process also occurs slowly.[24]

The top right panel of Figure 11 plots these workers' labor income. Since their employment rate does not change on impact, with a $15 minimum wage the labor income of workers initially earning $12 increases by 25%. As firms increase the demand for these workers, their labor income increases

[24]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 and attriting many of the existing lower-productivity workers before they start purchasing new capital, which uses higher ratios of higher-productivity workers. This pattern of firm response also explains why employment declines for lower-productivity workers are faster than employment increases for higher-productivity workers, thus generating a nonmonotone response in aggregate non-college employment to a $12 minimum wage, as shown in the dashed line of the left panel of Figure 11.

by about 40% in the long run. Hence, the slow employment dynamics delays the long-run *benefits* of the minimum wage for these workers. This slow dynamics also delays the long-term *costs* of the minimum wage for workers initially earning $7.50. Specifically, on impact, a $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 labor income of these workers significantly falls relative to its level in the initial BGP.

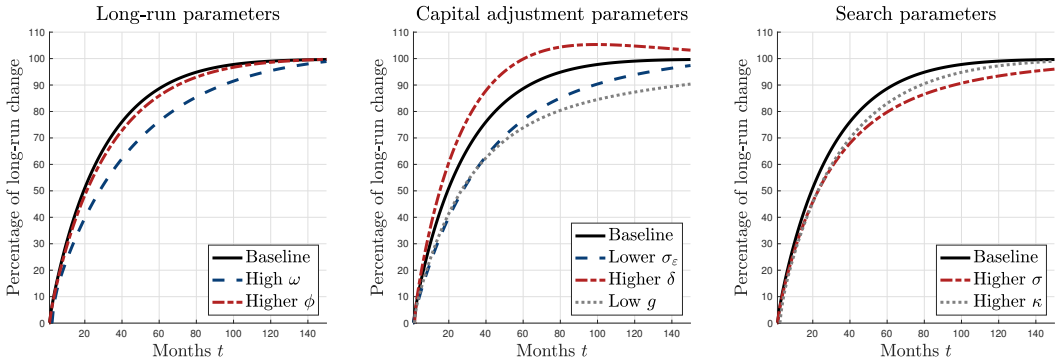
A full evaluation of the policy must then integrate both its short-run and long-run effects. The blue line in the bottom left panel of Figure 11 plots the percentage increase in the present value of labor income for each non-college worker type along the transition, relative to its present value in the initial BGP (as measured by their initial wage). The present value of labor income (weakly) increases for all non-college workers and proportionately more for the lowest-wage ones.

We also plot the change in this present value if we assume that the average change in labor income over the first two years after the introduction of the new minimum is permament. We refer to this change as the *Short Run Only* counterfactual. Clearly, for low-wage workers, this *naive static analysis* greatly overstates the gains from the minimum wage. For example, for workers initially earning $7.50, it would imply an increase in the present value of labor income by 80%, which is twice as large as the true gains after accounting for the transition.

The bottom left panel of Figure 11 also shows how this present value would change in a *Long Run Only* counterfactual that assumes that on impact the economy immediately reaches the new BGP. For workers initially earning $7.50, this *naive long-run analysis* would predict that they experience a decrease in the present value of labor income by over 60%, whereas the true gains calculated accounting for the transition are positive and over 40%. The calculations under the Short-Run Only and Long-Run Only counterfactuals differ significantly from those under our baseline, especially for lower-productivity workers, as they experience both larger short-run gains and larger long-run losses in response to a $15 minimum wage.

These observations show the limitations of two common approaches to evaluating the impact of the minimum wage. The first approach is to extrapolate empirical measures of the impact of a change in the minimum wage, estimated over the first few years following a change, to much longer time horizons, presuming that what happens during such a short-run period is informative about what happens thereafter. The second approach is to either rely on a long-run framework that ignores the transition altogether or one with an overly fast transition to assess the dynamic impact of a change in the minimum wage. Neither of these common approaches would correctly measure the change in the present value of labor income along the full transition for lower-wage workers.

FIGURE 12: 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 7. *Higher* ω corresponds to a degree of monopsony power of $\omega^{-1} = 1/6$ that produces an 85% markdown. *Higher* ϕ corresponds to a long-run elasticity of substitution within education groups of $\phi = 4.5$. *Lower* σ_ε 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* δ sets the depreciation rate to $\delta = 20\%$ annually. *Lower* g corresponds to a trend growth rate g of 0.01% annually. *Higher* σ sets the job-destruction rate to $\sigma = 3.5\%$ monthly. *Higher* κ increases the baseline vacancy-posting cost κ_0 by 2.5 times, which roughly doubles average hiring costs $\kappa_i/\lambda_f(\theta_i)$ to 125% of average monthly wage.

Finally, throughout this discussion, we have focused on the effects of minimum wage changes on workers who are already in the market. The bottom right panel of Figure 11 considers the present value of income of workers who enter the market either 5 or 10 years after the imposition of a $15 minimum wage. The present value of labor income for low productivity workers who enter in the future is now negative. The reason is that such workers miss out on the early part of the transition during which wages increase but firms have not yet substituted away from them. Interestingly, for workers with higher skill levels, entering later increases their gains. The reason is that for such workers, firms substitute towards them rather than away from them over time.

6.4 Sensitivity Analysis and Relationship to Empirical Literature

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

Sensitivity Analysis. We first examine how various features of our model affect how quickly aggregate non-college employment responds to the minimum wage. The left panel of Figure 12 focuses on two parameters that are key for the long-run effect of the minimum wage: the degree of firm monopsony power, governed by ω, and the long-run elasticity of labor-labor substitution, governed by ϕ. Reducing the degree of monopsony power from $\omega = 3.07$ to $\omega = 6$—which changes the average markdown from 0.75 to 0.85—also leads to a slower transition.[25] Increasing the long-run elasticity of labor-labor substitution from $\phi = 4$ to $\phi = 4.5$ leads to a slightly slower transition. As the online appendix shows, both of these changes lead to a larger long-run decline in employment.

[25]The present value of benefits from a $15 minimum are smaller for lower monopsony distortions; see Appendix F.

The middle panel of Figure 12 illustrates the sensitivity of the non-college employment dynamics to the parameters that govern the capital adjustment process. As for the capital accumulation dynamics, consider an increase in the capital depreciation rate from $\delta = 15\%$ to $\delta = 20\%$ per year. Since a firm's installed capital depreciates more quickly, employment adjusts more rapidly, as firms more quickly purchase new capital with different labor intensity. As for the capital utilization dynamics, consider a decrease in the dispersion of idiosyncratic capital productivity from $\sigma_\varepsilon = 0.25$, with an average utilization rate of 75% to $\sigma_\varepsilon = 0.01$, with an average utilization rate of 97%. Since, as Figure 2 shows, fewer units of capital are near the utilization threshold, changes in utilization are small in response to changes in the minimum wage, slowing down the transition of employment. Similarly, if we decrease the growth rate of the economy g from 2% to 1%, the transition is again slower as fewer units of capital are near the utilization threshold.

The right panel of Figure 12 shows the robustness of our main results to two parameters governing search frictions: the vacancy-posting cost κ_i and the job separation rate σ. Increasing the baseline vacancy-posting cost κ_0 to approximately double average hiring costs $\kappa_i/\lambda_f(\theta_i)$ and increasing the job separation rate σ from 2.8% to 3.5% have a relatively minor effect on the speed of transition of employment. The reason is that most of the slow adjustment implied by our model is due to our putty-clay technology rather than 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.[26] 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 7 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 dif-

[26] We focus 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 either across neighboring cities or from cities to suburbs where the minimum wage remained unchanged. Although city-level studies are informative about relative spatial reallocation effects, say, city vs. suburb employment changes, they are not informative about the aggregate effects of policies that we focus on.

ferent 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 7. However, Clemens and Strain (2021) also finds 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 minimum wage changes shown in Figure 7. We take it as a strength of our framework that we can match these dynamic employment responses to increases in the minimum wage.[27]

Finally, a separate literature provides evidence that firms slowly adjust their mix of labor in response to increases in the minimum wage, in line with the predictions of our model. For example, Horton (2025) and Clemens, Kahn and Meer (2021) provide evidence that U.S. firms substitute away from low-productivity workers towards higher-productivity ones when the minimum wage increases. Harasztosi and Lindner (2019) document that firms increase their capital stock in response to a large minimum wage change in Hungary. This adjustment is consistent with our model.

Indeed, it is important to note that the speed and amount of capital adjustment in response to a minimum wage change does not provide direct evidence for or against putty-clay models. The reason, as emphasized throughout our paper, is that putty-clay imposes a friction to adjusting the *mix of labor inputs on existing capital*, not a friction to accumulating new capital. The ideal data for validating the key mechanism of putty-clay adjustment in our model would explore the change in the labor intensity of existing units of capital in response to a large increase in the minimum wage. For example, suppose a firm produced the same product with different plants that varied in the time they were built and in their original low-skill labor shares. Our model implies that, in response to a large minimum wage, the utilization rates of *existing* capital in older or more low-skill intensive plants would fall more than the utilization rates in other plants. Furthermore, plants built after the minimum wage change would have higher utilization rates and high-skill labor shares. We think such types of analysis would be a fruitful area of future research.

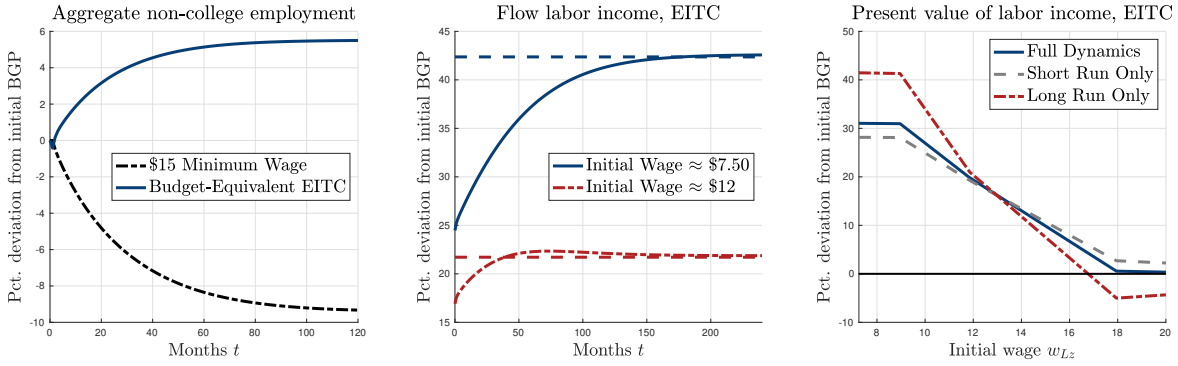
6.5 Dynamics Effects of the EITC

We now show that the mechanisms governing the dynamic effects of the EITC and the minimum wage are similar. To this end, we introduce the EITC in period 0 starting from the economy's BGP

[27]Similarly, Cengiz et al. (2019) estimates the short- and the long-run effects of small minimum wage increases, which, averaging across all the increases they study, were about 10% or $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 has similar predictions for small minimum wage changes. However, as our model shows, these findings should not be extrapolated to larger minimum wage changes.

without any policies assuming that agents have perfect foresight thereafter. We maintain the EITC is budget-equivalent to a $15 minimum wage in the long run—namely, financed though a linear tax on profits such that tax revenues equal the profit loss associated with a $15 minimum.

FIGURE 13: Employment and Labor Income Dynamics in Response to EITC



Notes: Transition paths following the introduction of the EITC. The left panel plots the path of aggregate non-college employment expressed relative to the initial BGP. The middle panel plots detrended flow labor income for two non-college worker types. The 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 as in Figure 11.

The left panel of Figure 13 compares the transition path of aggregate non-college employment following the introduction of the EITC with that following a $15 minimum wage. Under the 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 pay to hire them. Lower wages and reduced monopsony distortions 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 13 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 labor 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 labor income. These income gains are proportionally larger for the low-productivity worker.

The right panel of Figure 13 shows 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. Here also the Short Run Only and Long Run Only counterfactuals are unsatisfactory. In particular, 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 the present value of labor income, 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 its true effects on the present value of labor income.

7 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 their full dynamic impact into account. In this paper, we provide a general equilibrium model of dynamic labor and capital adjustment that features rich worker heterogeneity, complementarity in production among workers, putty-clay capital, and monopsony and search frictions to capture the key forces required for such an assessment. Our novel framework accounts for many of the features that have informed the long-standing debate on canonical policies such as the minimum wage and the EITC—namely, the substitutability among workers of different skills in production, the labor displacing or labor augmenting effect of capital, and firms' monopsony power in labor markets. Used as a policy evaluation tool, our framework leads to substantially different conclusions about the aggregate and distributional effects of these common policies than standard static-only or long-run-only approaches would imply. Our model is ideally suited for such policy evaluation because of its unique ability to simultaneously generate low labor-labor substitutability in the short run consistent with the minimum wage literature and high labor-labor substitutability in the long run consistent with the classic work of Card and Lemieux (2001) and Borjas and Katz (2007).

An original contribution of our exercise is to microfound the notion of firm monopsony power within a frictional equilibrium model of the labor market. Search frictions lead to labor market allocations that imply realistic labor market flows into and out of employment, allow exploring the impact of policies on both the participation and employment margin, which tend to move at different frequencies, and help to pin down allocations when policy-mandated constraints such as the minimum wage become binding. We hope that the framework we propose will stimulate more work on the dynamic evaluation of policies that have first-order effects on the price and allocation of different types of labor in imperfectly competitive labor markets.

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.

Bils, M., B. Kaymak, and K.-J. Wu (2024): "Labor substitutability among schooling groups," *American Economic Journal: Macroeconomics*, 16(4), 1–34.

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.

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

Horton, J. J. (2025): "Price Floors and Employer Preferences: Evidence from a Minimum Wage Experiment," *American Economic Review*, 115(1), 117–46.

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.

Kline, P. M. (2025): "Labor Market Monopsony: Fundamentals and Frontiers," Working Paper 33467, National Bureau of Economic Research.

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.

Ljungqvist, L., and T. J. Sargent (2017): "The fundamental surplus," *American Economic*

Review, 107(9), 2630–2665.

MaCurdy, 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.

——— (2021a): "The Elusive Employment Effect of the Minimum Wage," *Journal of Economic Perspectives*, 35(1), 3–26.

——— (2021b): "Monopsony in Labor Markets: A Review," *ILR Review*, 74(1), 3–26.

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.

Whitmore Schanzenbach, D., and M. R. Strain (2021): "Employment Effects of the Earned Income Tax Credit: Taking the Long View," *Tax Policy and the Economy*, 35, 87–129.

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

$$\begin{aligned}
&\max_{c_{it},s_{ijt},n_{ijt+1}} \sum\nolimits_{t=0}^{\infty} \beta^t U_t(c_{it}, n_{it}, s_{it}) \\
&\text{s.t. } n_{ijt+1} = (1-\sigma)n_{ijt} + \lambda_w(\theta_{ijt})s_{ijt} \quad (\times \beta^{t+1}\widehat{V}_{ijt+1}) \\
&\quad \sum\nolimits_{t=0}^{\infty} Q_{0,t}c_{it} = \zeta_i \mathbb{P} + \mathbb{I}_i + \sum\nolimits_{t=1}^{\infty} Q_{0,t} \sum\nolimits_j \lambda_w(\theta_{ijt-1})s_{ijt-1}W_{ijt}. \quad (\times \Gamma).
\end{aligned} \tag{A1}$$

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

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

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

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

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 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 θ_{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:

$$\begin{aligned}
&\sum_t Q_{0,t}\Bigg(\sum_\tau \int_{v,\varepsilon} u_{jt}(v,A_{t-\tau},\varepsilon)A_{t-\tau}\varepsilon f(v)K_{jt}(v,A_{t-\tau})\pi(\varepsilon)d\varepsilon dv - \sum_i \mu_i(\lambda_f(\theta_{ijt-1})a_{ijt-1}W_{ijt} + \kappa_{it}a_{ijt}) \\
&- \int X_{jt}(v)dv\Bigg) + \sum_{t=0}^{\infty} Q_{0,t+1}\mu_i M_{ijt+1}\frac{U_{nit+1}}{U_{cit+1}}\left(\frac{n_{ijt+1}}{n_{it+1}}\right)^{\frac{1}{\omega}} + \sum_{t=0}^{\infty} Q_{0,t+1}\mu_i\gamma_{ijt+1}\left[W_{ijt+1} - \frac{\mathcal{W}_{it}}{Q_{t,t+1}\lambda_w(\theta_{ijt})}\right] \\
&\text{such that} \quad u_{jt}(v,A_{t-\tau},\varepsilon) \geq 0 \quad (\times Q_{0,t}\lambda^L_{jt}(v,A_{t-\tau},\varepsilon)) \\
&u_{jt}(v,A_{t-\tau},\varepsilon) \leq 1 \quad (\times Q_{0,t}\lambda^U_{jt}(v,\varepsilon,A_{t-\tau})) \\
&u_{jt}(v,A_{t-\tau},\varepsilon)v_i K_{jt}(v,A_{t-\tau})\pi(\varepsilon) \leq N_{ijt}(v,A_{t-\tau},\varepsilon) \text{ for all } i \quad (\times Q_{0,t}\lambda_{ijt}(v,A_{t-\tau},\varepsilon)) \\
&\sum_\tau \int_{v,\varepsilon} N_{ijt}(v,A_{t-\tau},\varepsilon)d\varepsilon dv \leq \mu_i n_{ijt} \text{ for all } i \quad (\times Q_{0,t}\chi_{ijt}) \\
&\mu_i n_{ijt+1} \leq (1-\sigma)\mu_i n_{ijt} + \lambda_f(\theta_{ijt})\mu_i a_{ijt} \text{ for all } i \quad (\times Q_{0,t+1}\nu_{ijt+1}) \\
&K_{jt+\tau+1}(v,A_t) = (1-\delta)^\tau X_{jt}(v) \quad (\times Q_{0,t+\tau+1}q_{jt,t+\tau+1}(v)) \\
&X_{jt}(v) \geq 0 \quad (\times Q_{0,t}\mu_{jt}(v)),
\end{aligned}$$

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 $\lambda_{ijt}(v,A_{t-\tau},\varepsilon) = \chi_{ijt}$ for all machines $(v,A_{t-\tau},\varepsilon)$ that are operated. The first-order condition for utilization $u_{jt}(v,A_{t-\tau},\varepsilon)$ is given by

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

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

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

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 $\varepsilon < \underline{\varepsilon}(v, A_{t-\tau}; \chi_{jt})$. In the knife-edge case in which $A_{t-\tau}\varepsilon f(v) - \sum_i \chi_{ijt} v_i = 0$, the firm is indifferent over any $u_{jt}(v, A_{t-\tau}, \varepsilon) \in [0, 1]$. Note that the solution to the static utilization problem (11) 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 ν_{ijt+1} as the present value of a marginal worker to the firm, taking into account both their marginal product χ_{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

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

The first-order condition for market tightness θ_{ijt} is

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

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

$$\gamma_{ijt+1} = \lambda_f(\theta_{ijt})a_{ijt}. \tag{A9}$$

Plugging in this expression for γ_{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

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

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

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

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

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

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

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

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.[28] However, each term in the summand in this second line is zero at the optimum. To see this, plug in $\underline{\varepsilon}(v, A_t; \chi_{jt+\tau}) = \frac{\sum_i \chi_{ijt+\tau} v_i}{A_t f(v)}$ to see that the term becomes 0 for each future period τ. 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,

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

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. $c_{it} = \widetilde{c}_{it}(1+g)^t$.

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

(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 Households

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

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

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

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

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

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

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

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

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

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

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,

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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 (25) 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 ε in the putty-clay model, but after the realization of ε in the model with standard capital. Therefore, firms in the standard model will implicitly assign more workers to high-ε 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 ε in the putty-clay model, then we would not have an active utilization margin, which is a key feature of our analysis.

Reduced-Form 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 7 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 7. *In the BGP, the labor allocations and wages are determined by the following equations:*

(i) optimal cut-off for the 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 profits

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

(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, θ_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, α, and the dispersion of idiosyncratic shocks σ_ε.

(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 can be rearranged to obtain 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) can be rearranged to obtain

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

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 (25) 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 obtain

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

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 θ_{ij} and $\widetilde{\mathcal{W}}_i$ fixed, we obtain

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

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

$$W_{ijt+1} \geq \underline{W}_{t+1} \quad \text{for all } t \geq 0 \ (\times Q_{0,t}\rho_{ijt+1}) \tag{B1}$$

and a nonnegativity condition on vacancies

$$a_{ijt} \geq 0 \text{ for all } t \geq 0 \quad (\times Q_{0,t}\xi^a_{ijt}), \tag{B2}$$

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 (B1) 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 why we specify this constraint in terms of flow wages is that in practice this is how minimum wage legislation works. Our formulation restricts wages in minimal ways consistent with the constraint that in each period the flow wage is at least as high as the legislated minimum. Specifically, it prevents firms from offering present values of wages such that in some periods the associated flow wage falls below the 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 θ_{ijt}, (A8), are not affected.

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

$$\begin{aligned} &-Q_{0,t}\kappa_{it} - Q_{0,t+1}\lambda_f(\theta_{ijt})W_{ijt+1} + Q_{0,t+1}\lambda_f(\theta_{ijt})\nu_{ijt+1} + Q_{0,t}\xi^a_{ijt} = 0 \\ &\quad \Longrightarrow \frac{\kappa_i}{\lambda_f(\theta_{ijt})} + Q_{t,t+1}W_{ijt+1} \geq Q_{t,t+1}\nu_{ijt+1}, \text{ with equality if } a_{ijt} > 0. \end{aligned} \tag{B3}$$

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

$$-\lambda_f(\theta_{ijt})a_{ijt} + \gamma_{ijt+1} + \rho_{ijt+1} = 0, \tag{B4}$$

where ρ_{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 γ_{ijt+1} given that $W_{ijt+1} = \underline{W}_t$. That is, we solve for γ_{ijt+1} using the following algebra

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

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

$$\gamma_{ijt+1} = \left\{ \begin{array}{c} \lambda_f(\theta_{ijt})a_{ijt} \text{ if slack} \\ \frac{\eta}{1-\eta} \frac{\nu_{ijt+1} - \underline{W}_{t+1}}{\underline{W}_{t+1} + V_{ijt+1}} \lambda_f(\theta_{ijt})a_{ijt} \text{ if binds} \end{array} \right\} \tag{B5}$$

$$W_{ijt+1} = \left\{ \begin{array}{c} \eta\nu_{ijt+1} - (1-\eta)V_{ijt+1} \text{ if slack} \\ \underline{W}_{t+1} \text{ if binds} \end{array} \right\} \tag{B6}$$

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

$$\nu_{ijt+1} = \chi_{ijt+1} + M_{ijt+1} \frac{u_{nit+1}}{u_{cit+1}} \frac{1}{\omega} \left(\frac{n_{ijt+1}}{n_{it+1}} \right)^{\frac{1}{\omega}-1} \frac{1}{n_{it+1}} + Q_{t+1,t+2}(1-\sigma)\nu_{ijt+2}, \tag{B7}$$

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 (B5) becomes

$$\gamma_{it+1} = \left\{ \begin{array}{c} \lambda_f(\theta_{it})a_{it} \text{ if slack} \\ \frac{\eta}{1-\eta}\frac{\widetilde{\nu}_{it+1} - \underline{\widetilde{W}}_{t+1}}{\underline{\widetilde{W}}_{t+1} + \widetilde{V}_{it+1}}\lambda_f(\theta_{it})a_{it} \text{ if binds} \end{array} \right\} \tag{B8}$$

and in detrended form wages are

$$\widetilde{W}_{it+1} = \left\{ \begin{array}{c} \eta\widetilde{\nu}_{it+1} - (1-\eta)\widetilde{V}_{it+1} \text{ if slack} \\ \underline{\widetilde{W}}_{t+1} \text{ if binds} \end{array} \right\}. \tag{B9}$$

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

$$\widetilde{\nu}_{it+1} = \widetilde{\chi}_{it+1} - M_{it+1}v'(n_{it+1})\frac{1}{\omega n_{it+1}} + Q_{t+1,t+2}(1+g)(1-\sigma)\widetilde{\nu}_{it+2}. \tag{B10}$$

B.1.2 Balanced Growth Path and Proof of Proposition 6

We summarize how the minimum wage impacts the BGP In the following result, which is analogous to Lemma 7 in Appendix A.

Lemma 8. *In the BGP, the labor allocations and wages are determined by the following equations:*

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

$$\underline{\varepsilon}_1 = (1+g)(1-\alpha)m(\underline{\varepsilon}_1), \tag{B11}$$

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

$$m(\underline{\varepsilon}_1) = \frac{\sum_{\tau=1}^{\infty}\widetilde{\beta}^{\tau}(1-\delta)^{\tau-1}\Pi^p((1+g)^{\tau-1}\underline{\varepsilon}_1)}{\sum_{\tau=1}^{\infty}\widetilde{\beta}^{\tau}(1-\delta)^{\tau-1}(1+g)^{\tau}\Pi^u((1+g)^{\tau-1}\underline{\varepsilon}_1)}; \tag{B12}$$

(ii) the marginal unit of capital earns zero profits

$$1 = \alpha\left[\sum_{\tau=1}^{\infty}\widetilde{\beta}^{\tau}(1-\delta)^{\tau-1}\Pi^p((1+g)^{\tau-1}\underline{\varepsilon}_1)\right]f(\widetilde{v}); \tag{B13}$$

(iii) flow wages

$$\widetilde{w}_i = \left\{ \ \eta[f_i(\widetilde{v})m(\underline{\varepsilon}_1) - v'(n_i)/\omega] + (1-\eta)v'(n_i) \ \textit{if slack}, \underline{w} \ \textit{if bind} \ \right\}; \tag{B14}$$

(iv) optimal vacancy posting

$$\kappa_i = \left\{ \begin{array}{c} \beta\lambda_f(\theta_i)\frac{f_i(\widetilde{v})m(\underline{\varepsilon}_1)-\widetilde{w}_i-v'(n_i)/\omega}{1-\beta(1-\sigma)} \ \textit{if slack} \\ \beta\lambda_f(\theta_i)\frac{f_i(v)m(\underline{\varepsilon}_1)-\underline{w}}{1-\beta(1-\sigma)}\left[\frac{\underline{w}-v'(n_i)}{\underline{w}-v'(n_i)(1-1/\omega)}\right] \ \textit{if bind} \end{array} \right\}; \tag{B15}$$

(v) optimal household search

$$h'(s_i) = \beta\lambda_w(\theta_i)\frac{\widetilde{w}_i - v'(n_i)}{1-\beta(1-\sigma)}; \tag{B16}$$

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

$$\sigma n_i = \lambda_w(\theta_i)s_i; \tag{B17}$$

(vii) labor market clearing

$$\frac{\mu_i n_i}{\widetilde{v}_i} = \frac{\mu_1 n_1}{\widetilde{v}_1}. \tag{B18}$$

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 (B8) in the detrended system) and (ii) the wage equation (equation (B9) 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 7 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 γ_{it} from (B8) in the expression for the present value of a worker $\widetilde{\nu}_{it+1}$ from (B14) and simplified to arrive at the non-binding version of condition (iv) from Lemma 7. When the minimum wage is binding, we must follow a different strategy because the multiplier on the participation constraint γ_{it} in (B8) 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 (B10) implicitly depends on the value of γ_{it} through the auxiliary variable $M_{it+1} = (1-\sigma)M_{it} + \gamma_{it+1}$.

Note that imposing balanced growth on (B10) gives

$$\widetilde{\nu}_i = \frac{\widetilde{\chi}_i - M_i v'(n_i)/(\omega n_i)}{1-\beta(1-\sigma)}. \tag{B19}$$

Next, note that the BGP version of the multiplier on the participation constraint (B8) becomes

$$\gamma_i = \frac{\eta}{1-\eta}\frac{\widetilde{\nu}_i - \underline{W}}{\underline{W} - \widetilde{V}_i}\lambda_f(\theta_i)a_i \implies \frac{\gamma_i}{\sigma}\frac{1}{n_i} = \frac{\widetilde{\nu}_i - \underline{W}}{\underline{W} + \widetilde{V}_i}, \tag{B20}$$

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 (B19) 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 (B20) by dividing both the numerator and denominator by $1-\beta(1-\sigma)$ to get

$$\frac{\gamma_i}{\sigma}\frac{1}{n_i} = \frac{\widehat{\nu}_i - \underline{w}}{\underline{w} - v'(n_i)}. \tag{B21}$$

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 (B21) into this equation to get

$$\widehat{\nu}_i = \widetilde{\chi}_i - \frac{v'(n_i)}{\omega}\frac{\widehat{\nu}_i - \underline{w}}{\underline{w} - v'(n_i)}. \tag{B22}$$

We will use this implicit expression for $\widehat{\nu}_i$ to obtain the expression for optimal vacancy posting (B15). Now subtract the minimum wage $\underline{w}$ and solve for $\widehat{\nu}_i - \underline{w}$ to get

$$\begin{aligned}
\widehat{\nu}_i - \underline{w} = \widetilde{\chi}_i - \underline{w} - \frac{\widehat{\nu}_i - \underline{w}}{\underline{w} - v'(n_i)}\frac{v'(n_i)}{\omega} &\implies (\widehat{\nu}_i - \underline{w})\left[1 + \frac{1}{\underline{w} - v'(n_i)}\frac{v'(n_i)}{\omega}\right] = \widetilde{\chi}_i - \underline{w} \\
&\implies (\widehat{\nu}_i - \underline{w})\left[\frac{\underline{w} - v'(n_i) + v'(n_i)/\omega}{\underline{w} - v'(n_i)}\right] = \widetilde{\chi}_i - \underline{w} \\
&\implies \widehat{\nu}_i - \underline{w} = \left[\frac{\underline{w} - v'(n_i)}{\underline{w} - v'(n_i)(1-1/\omega)}\right](\widetilde{\chi}_i - \underline{w}).
\end{aligned} \tag{B23}$$

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

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

to get

$$\kappa_i = \frac{\beta}{1-\beta(1-\sigma)}\lambda_f(\theta_i)\left[\frac{\underline{w} - v'(n_i)}{\underline{w} - v'(n_i)(1-1/\omega)}\right](\widetilde{\chi}_i - \underline{w}). \tag{B24}$$

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

$$\kappa_i = \frac{\beta}{1-\beta(1-\sigma)}\lambda_f(\theta_i)\left[\frac{\underline{w}-v'(n_i)}{\underline{w}-v'(n_i)(1-1/\omega)}\right][f_i(v)m(\underline{\varepsilon}_1)-\underline{w}], \tag{B25}$$

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 κ_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

$$\underline{\varepsilon}_1 = (1+g)(1-\alpha)m(\underline{\varepsilon}_1) \tag{B26}$$

$$1 = \alpha\left[\sum_{\tau=1}^{\infty}\widetilde{\beta}^{\tau}(1-\delta)^{\tau-1}\Pi^p((1+g)^{\tau-1}\underline{\varepsilon}_1)\right]f(\widetilde{v}) \tag{B27}$$

$$\kappa_i(1-\tau_i) = \lambda_f(\theta_i)\frac{f_i(v)m(\underline{\varepsilon}_1)-\underline{w}_i}{\rho+\sigma}\left[\frac{\underline{w}_i-v'(n_i)}{\underline{w}_i-v'(n_i)(1-1/\omega)}\right] \tag{B28}$$

$$h'(s_i) = \beta\lambda_w(\theta_i)\frac{\underline{w}_i-v'(n_i)}{1-\beta(1-\sigma)} \tag{B29}$$

$$\sigma n_i = \lambda_w(\theta_i)s_i \tag{B30}$$

$$\frac{\mu_i n_i}{\widetilde{v}_i} = \frac{\mu_1 n_1}{\widetilde{v}_1}. \tag{B31}$$

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

$$\underline{\varepsilon}_1 = (1+g)(1-\alpha)m(\underline{\varepsilon}_1) \tag{B32}$$

$$1 = \alpha\left[\sum_{\tau=1}^{\infty}\widetilde{\beta}^{\tau}(1-\delta)^{\tau-1}\Pi^p((1+g)^{\tau-1}\underline{\varepsilon}_1)\right]f(\widetilde{v}^c) \tag{B33}$$

$$\kappa_i = \beta\lambda_f(\theta_i^c)\frac{f_i(\widetilde{v}^c)m(\underline{\varepsilon}_1)-\widetilde{w}_i^c}{1-\beta(1-\sigma)} \tag{B34}$$

$$h'(s_i^c) = \beta\lambda_w(\theta_i^c)\frac{\widetilde{w}_i^c-v'(n_i^c)}{1-\beta(1-\sigma)} \tag{B35}$$

$$\sigma n_i^c = \lambda_w(\theta_i^c)s_i^c \tag{B36}$$

$$\frac{\mu_i n_i^c}{\widetilde{v}_i^c} = \frac{\mu_1 n_1^c}{\widetilde{v}_1^c} \tag{B37}$$

$$\widetilde{w}_i^c = \eta f_i(\widetilde{v}^c)m(\underline{\varepsilon}_1) + (1-\eta)v'(n_i^c). \tag{B38}$$

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, (B26)—(B31), 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, (B32)—(B38).

Given this guessed allocation, the following equations are clearly the same because the wage does not enter them directly: (i) the productivity cutoffs (B26) and (B32), (ii) the zero profit condition for the marginal unit of capital (B27) and (B34), (iii) the law of of motion for employment (B30), and (iv) the labor ratios (B31) and (B37). 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 (B29) and (B35) also coincide. Finally, under the choice for τ_i from the main text, the optimal vacancy-posting conditions (B28) and (B34) also coincide.

Note that dividing the monopsony vacancy-posting condition (B28) 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 χ_{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 ν_{ij1} is the scaled multiplier on this constraint. Finally, firms must satisfy the non-negativity constraint $F_{ij0} \geq 0$ $(\times \xi^f_{ij0})$.

First-Order Conditions. 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, $\widehat{W}_{i0} - \chi_{ij0} - Q_{0,1}(1-\sigma)\nu_{ij1} + \xi^f_{ij0} = 0$, which implies

$$\chi_{ij0} + Q_{0,1}(1-\sigma)\nu_{ij1} \geq \widehat{W}_{i0}, \text{ with equality if } F_{ij0} > 0. \tag{B39}$$

That is, firms do not fire workers if the present value of the workers' benefits to the firm — their marginal product in period $t = 0$, χ_{ij0} plus their present value going forward, $Q_{0,1}(1-\sigma)\nu_{ij1}$ on the LHS of (B39)—is strictly greater than the present value of wage payments to those workers—$\widehat{W}_{i0}$

on the RHS of (B39). 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 $A'_t(w) = A'(\widetilde{w})$.[29] We use the discounter operator $d_{t+1} = 1 + Q_{t+1,t+2}(1+g)(1-\sigma) + Q_{t+1,t+3}(1+g)^2(1-\sigma)^2 + \ldots$ 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 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

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

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

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

[29] 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 τ_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 that changes is the one for search effort to take into account that wage payments are now W^H_{ijt+1},

$$-\frac{u_{sit}}{u_{cit}} = \lambda_w(\theta_{ijt})Q_{t,t+1}\left(V_{ijt+1} + W^H_{ijt+1}\right). \tag{B40}$$

The optimal search condition (B40) also changes the participation constraint to

$$\lambda_w(\theta_{ijt})(W^H_{ijt+1} + V_{ijt+1}) \geq \mathcal{W}_{it}. \tag{B41}$$

B.2.2 Firms

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

Firm Problem. The presence of the transfer system changes the firm's problem in three ways. First, the participation constraint (B41) 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

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

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

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

$u_{jt}(v, A_{t-\tau}, \varepsilon)v_i K_{jt}(v, A_{t-\tau})\pi(\varepsilon) \leq N_{ijt}(v, A_{t-\tau}, \varepsilon)$ 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 $i \quad (\times Q_{0,t+1}\nu_{ijt+1})$

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

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. At the hiring stage, the first-order conditions for employment n_{ijt+1}, vacancies a_{ijt}, and market tightness θ_{ijt} are unaffected. However, the transfer system changes how the first-order condition for market tightness (A8), so we reproduce it here:

$$W_{ijt+1} = \nu_{ijt+1} + \frac{\gamma_{ijt+1}}{a_{ijt}} \frac{\mathcal{W}_{it}}{\lambda_w(\theta_{ijt})^2} \frac{\lambda_w'(\theta_{ijt})}{\lambda_f'(\theta_{ijt})}. \tag{B42}$$

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

$$-Q_{0,t+1}\mu_i\lambda_f(\theta_{ijt})a_{ijt}d_{t+1} + Q_{0,t+1}\mu_i\gamma_{ijt+1}d_{t+1}A_t'(w_{ijt+1}) = 0 \implies \gamma_{ijt+1} = \frac{\lambda_f(\theta_{ijt})a_{ijt}}{A_t'(w_{ijt+1})}. \tag{B43}$$

The transfer system changes the multiplier on the participation constraint, which has two effects. First, this multiplier enters the law of motion for the auxiliary variable $M_{ijt+1} = (1-\sigma)M_{ijt} + \gamma_{ijt+1}$. Second, this multiplier affects how we simplify the FOC for market tightness (B42). Plugging the expression for the multiplier (B43) into the FOC for market tightness (B42) gives

$$\begin{aligned} W_{ijt+1} &= \nu_{ijt+1} + \frac{\lambda_f(\theta_{ijt})}{\lambda_w(\theta_{ijt})} \frac{\mathcal{W}_{it}}{\lambda_w(\theta_{ijt})} \frac{\lambda_w'(\theta_{ijt})}{\lambda_f'(\theta_{ijt})} \frac{1}{A_t'(w_{ijt+1})} \\ &= \nu_{ijt+1} + \frac{\lambda_f(\theta_{ijt})}{\lambda_w(\theta_{ijt})} \left(W_{ijt+1}^H + V_{ijt+1}\right) \frac{\lambda_w'(\theta_{ijt})}{\lambda_f'(\theta_{ijt})} \frac{1}{A_t'(w_{ijt+1})} = \nu_{ijt+1} - \frac{1-\eta}{\eta} \frac{W_{ijt+1}^H + V_{ijt+1}}{A_t'(w_{ijt+1})}, \end{aligned} \tag{B44}$$

where in the second line we used $\frac{\mathcal{W}_{it}}{\lambda_w(\theta_{ijt})} = W_{ijt+1}^H + V_{ijt+1}$ 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 (B44) 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 Balanced Growth Path

The only equilibrium conditions which have changed relative to the baseline are the search FOC (B40) and the wage equation (B44). We therefore focus our discussion of detrending on those conditions. For the search FOC (B40), first note that

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

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

$$\begin{aligned} (1+g)^t h'(s_{it}) &= Q_{t,t+1}\lambda_w(\theta_{it})(1+g)^{t+1}\left(\widetilde{W}_{it+1}^H + \widetilde{V}_{it+1}\right) \\ \implies h'(s_{it}) &= Q_{t,t+1}(1+g)\lambda_w(\theta_{it})\left(\widetilde{W}_{it+1}^H + \widetilde{V}_{it+1}\right). \end{aligned} \tag{B45}$$

To detrend the wage equation (B44), recall the property that $A'_t(w_{ijt+1}) = A'(\widetilde{w}_{ijt+1})$. Therefore,

$$\widetilde{W}_{ijt+1} = \widetilde{\nu}_{ijt+1} - \frac{1-\eta}{\eta}\frac{\widetilde{W}^H_{ijt+1} + \widetilde{V}_{ijt+1}}{A'(\widetilde{w}_{ijt+1})} \text{ and } \gamma_{ijt+1} = \frac{\lambda_f(\theta_{ijt})a_{ijt}}{A'(\widetilde{w}_{ijt+1})}. \tag{B46}$$

C Putty-Putty Model

This appendix derives the putty-putty version of our model, referenced in Section 2 of the main text. Section C.1 shows that the firm's problem in the putty-putty model can be represented using an aggregate production function in terms of efficiency units of capital. Section C.2 solves that problem and compares the solution to the putty-clay model.

C.1 Aggregate Production Function

In each period t, firm j can freely allocate $N_{ijt}(A_{t-\tau}, \varepsilon)$ units of type-i labor to capital with vintage productivity $A_{t-\tau}$ and idiosyncratic productivity ε. The resulting output from combining $N_{ijt}(A_{t-\tau}, \varepsilon)$ units of labor with $K_{jt}(A_{t-\tau}, \varepsilon)$ units of capital is

$$\begin{aligned}
Y_{jt}(A_{t-\tau}, \varepsilon) &= \Big[(A_{t-\tau}\varepsilon)^{\frac{1}{\alpha}} K_{jt}(A_{t-\tau}, \varepsilon)\Big]^{\alpha} G(\mathbf{N}_{jt}(A_{t-\tau}, \varepsilon))^{1-\alpha} \\
G(\mathbf{N}_{jt}(A_{t-\tau}, \varepsilon)) &= \Big[\xi N_{Ljt}(A_{t-\tau}, \varepsilon)^{\frac{\varphi-1}{\varphi}} + (1-\xi) N_{Hjt}(A_{t-\tau}, \varepsilon)^{\frac{\varphi-1}{\varphi}}\Big]^{\frac{\varphi}{\varphi-1}} \\
N_{jet}(A_{t-\tau}, \varepsilon) &= \Big[\sum_{i \in I_e} z_i (N_{ijt}(A_{t-\tau}, \varepsilon)^{\frac{\phi-1}{\phi}})\Big]^{\frac{\phi}{\phi-1}},
\end{aligned}$$

where $\mathbf{N}_{jt}(A_{t-\tau})$ is the vector of labor across types $i = 1, 2, \ldots, I$, $e \in \{L, H\}$ and I_e is the set of types i in group e. These are the functional forms used in our quantitative analysis in Section 5.

Labor Allocation Problem. As in the putty-clay model, a key component of the firm's profit maximization problem is to allocate workers across different units of capital, subject to the constraints $\sum_\tau \int_\varepsilon N_{ijt}(A_{t-\tau}, \varepsilon) d\varepsilon \leq N_{ijt}$, where N_{ijt} is the total measure of workers of type i employed by firm j at the beginning of period t. We isolate this labor allocation problem by solving

$$\max_{N_{ijt}(A_{t-\tau},\varepsilon)} \sum_\tau \int_\varepsilon [(A_{t-\tau}\varepsilon)^{\frac{1}{\alpha}} K_{jt}(A_{t-\tau}, \varepsilon)]^{\alpha} G(\mathbf{N}_{jt}(A_{t-\tau}, \varepsilon))^{1-\alpha} d\varepsilon + \sum_i \widehat{\chi}_{ijt}[N_{ijt} - \sum_\tau \int_\varepsilon N_{itj}(A_{t-\tau}, \varepsilon) d\varepsilon,$$

where, as in the main text, we use the notation $\widehat{\chi}_{ijt}$ to emphasize the fact that the multiplier $\widehat{\chi}_{ijt}$ comes from the firm's full dynamic profit maximization problem. Solving this pseudo-Lagrangian yields the aggregate production function in the putty-putty model.

The first-order condition with respect to $N_{ijt}(A_{t-\tau}, \varepsilon)$ is

$$(1-\alpha)\left[\frac{(A_{t-\tau}\varepsilon)^{\frac{1}{\alpha}} K_{jt}(A_{t-\tau}, \varepsilon)}{G(\mathbf{N}_{jt}(A_{t-\tau}, \varepsilon))}\right]^{\alpha} G_i(\mathbf{N}_{jt}(A_{t-\tau}, \varepsilon)) = \widehat{\chi}_{ijt}. \tag{C1}$$

With our nested CES functional form, the partial derivative $G_i(\mathbf{N}_{jt}(A_{t-\tau},\varepsilon))$ can be written as

$$G_i(\mathbf{N}_{jt}(A_{t-\tau},\varepsilon)) = \xi_e \left[\frac{N_{ejt}(A_{t-\tau},\varepsilon)}{G(\mathbf{N}_{jt}(A_{t-\tau},\varepsilon))}\right]^{-\frac{1}{\varphi}} z_i \left[\frac{N_{ijt}(A_{t-\tau},\varepsilon)}{N_{ejt}(A_{t-\tau},\varepsilon)}\right]^{-\frac{1}{\phi}},$$

where $\xi_e = \xi$ if $i \in I_L$ and $\xi_e = 1-\xi$ if $i \in I_H$. Plugging this expression into (C1) gives

$$(1-\alpha)\left[\frac{(A_{t-\tau}\varepsilon)^{\frac{1}{\alpha}}K_{jt}(A_{t-\tau},\varepsilon)}{G(\mathbf{N}_{jt}(A_{t-\tau},\varepsilon))}\right]^{\alpha} \xi_e \left[\frac{N_{ejt}(A_{t-\tau},\varepsilon)}{G(\mathbf{N}_{jt}(A_{t-\tau},\varepsilon))}\right]^{-\frac{1}{\varphi}} z_i \left[\frac{N_{ijt}(A_{t-\tau},\varepsilon)}{N_{ejt}(A_{t-\tau},\varepsilon)}\right]^{-\frac{1}{\phi}} = \widehat{\chi}_{ijt}. \quad \text{(C2)}$$

We next proceed in two steps. First, we manipulate the first-order conditions (C2) across types i to show that relative labor demands are independent of the capital type $(A_{t-\tau},\varepsilon)$. Second, leveraging this result, we show that the level of labor demand depends linearly on the capital type $(A_{t-\tau},\varepsilon)$. This linearity allows us to aggregate across capital types to derive the aggregate production function.

Relative Labor Demands. To derive relative labor demands, divide the first-order condition (C2) for two types i and i_e which are both in the same broad group I_e:

$$\frac{z_i}{z_{i_e}}\left(\frac{N_{ijt}(A_{t-\tau},\varepsilon)}{N_{i_ejt}(A_{t-\tau},\varepsilon)}\right)^{-\frac{1}{\phi}} = \frac{\widehat{\chi}_{ijt}}{\widehat{\chi}_{i_ejt}}.$$

Invert this mapping to get

$$\frac{N_{ijt}(A_{t-\tau},\varepsilon)}{N_{i_ejt}(A_{t-\tau},\varepsilon)} = \left(\frac{\widehat{\chi}_{ijt}}{\widehat{\chi}_{i_ejt}}\frac{z_{i_e}}{z_i}\right)^{-\phi} \equiv n^e_{ijt}.$$

Hence, the ratio of employment between any two worker types i within the same education group is independent of capital type $(A_{t-\tau},\varepsilon)$. We denote the ratio of type i to the "base type" i_e n^e_{it} for all $i \in I_e$. The base type i_e is arbitrary, but it may be helpful to think of it as the first index i within the education group I_e. Note that, by definition, $n^e_{i_ejt} = 1$.

We can use this result to re-write the labor aggregate for the education group $N_{ejt}(A_{t-\tau},\varepsilon)$ as

$$N_{ejt}(A_{t-\tau},\varepsilon) = \left[\sum_{i\in I_e} z_i (N_{ijt}(A_{t-\tau},\varepsilon)^{\frac{\phi-1}{\phi}})\right]^{\frac{\phi}{\phi-1}} = N_{i_ejt}(A_{t-\tau},\varepsilon)[\sum_{i\in I_e} z_i (n^e_{ijt})^{\frac{\phi-1}{\phi}}]^{\frac{\phi}{\phi-1}}. \quad \text{(C3)}$$

Taking the ratio of this expression for $e=H$ divided by $e=L$ gives

$$\frac{N_{Hjt}(A_{t-\tau},\varepsilon)}{N_{Ljt}(A_{t-\tau},\varepsilon)} = \frac{N_{i_Hjt}(A_{t-\tau},\varepsilon)}{N_{i_Ljt}(A_{t-\tau},\varepsilon)} \frac{\left[\sum_{i\in I_H} z_i (n^H_{ijt})^{\frac{\phi-1}{\phi}}\right]^{\frac{\phi}{\phi-1}}}{\left[\sum_{i\in I_L} z_i (n^L_{ijt})^{\frac{\phi-1}{\phi}}\right]^{\frac{\phi}{\phi-1}}} \equiv \frac{N_{i_Hjt}(A_{t-\tau},\varepsilon)}{N_{i_Ljt}(A_{t-\tau},\varepsilon)} n_{ijt} \quad \text{(C4)}$$

for $n_{ijt} = \frac{\left[\sum_{i\in I_H} z_i (n^H_{ijt})^{\frac{\phi-1}{\phi}}\right]^{\frac{\phi}{\phi-1}}}{\left[\sum_{i\in I_L} z_i (n^L_{ijt})^{\frac{\phi-1}{\phi}}\right]^{\frac{\phi}{\phi-1}}}$, which is independent of capital type $(A_{t-\tau},\varepsilon)$.

Next, take the ratio of the FOC (C2) to the two base types i_H and i_L in each education group:

$$\frac{1-\xi}{\xi}\left(\frac{N_{Hjt}(A_{t-\tau},\varepsilon)}{N_{Ljt}(A_{t-\tau},\varepsilon)}\right)^{\frac{1}{\phi}-\frac{1}{\varphi}} \frac{z_{i_H}}{z_{i_L}}\left(\frac{N_{i_Hjt}(A_{t-\tau},\varepsilon)}{N_{i_Ljt}(A_{t-\tau},\varepsilon)}\right)^{-\frac{1}{\phi}} = \frac{\widehat{\chi}_{i_Hjt}}{\widehat{\chi}_{i_Ljt}}.$$

Plug in the ratio of group-level aggregates from (C4) to get

$$\frac{1-\xi}{\xi}\left(\frac{N_{i_Hjt}(A_{t-\tau},\varepsilon)}{N_{i_Ljt}(A_{t-\tau},\varepsilon)}n_{ijt}\right)^{\frac{1}{\phi}-\frac{1}{\varphi}}\frac{z_{i_H}}{z_{i_L}}\left(\frac{N_{i_Hjt}(A_{t-\tau},\varepsilon)}{N_{i_Ljt}(A_{t-\tau},\varepsilon)}\right)^{-\frac{1}{\phi}}=\frac{\widehat{\chi}_{i_Hjt}}{\widehat{\chi}_{i_Ljt}}$$
$$\implies \frac{1-\xi}{\xi}\frac{z_{i_H}}{z_{i_L}}n_{ijt}^{\frac{1}{\phi}-\frac{1}{\varphi}}\left(\frac{N_{i_Hjt}(A_{t-\tau},\varepsilon)}{N_{i_Ljt}(A_{t-\tau},\varepsilon)}\right)^{-\frac{1}{\varphi}}=\frac{\widehat{\chi}_{i_Hjt}}{\widehat{\chi}_{i_Ljt}}.$$

Invert this mapping to get

$$\frac{N_{i_Hjt}(A_{t-\tau},\varepsilon)}{N_{i_Ljt}(A_{t-\tau},\varepsilon)}=\left(\frac{\widehat{\chi}_{i_Hjt}}{\widehat{\chi}_{i_Ljt}}\frac{\xi}{1-\xi}\frac{z_{i_L}}{z_{i_H}}n_{ijt}^{\frac{1}{\varphi}-\frac{1}{\phi}}\right)^{-\phi}.$$

Finally, plug this back into (C4) to get

$$\frac{N_{Hjt}(A_{t-\tau},\varepsilon)}{N_{Ljt}(A_{t-\tau},\varepsilon)}=n_{ijt}\left[\frac{\widehat{\chi}_{i_Hjt}}{\widehat{\chi}_{i_Ljt}}\frac{\xi}{1-\xi}\frac{z_{i_L}}{z_{i_H}}n_{ijt}^{\frac{1}{\varphi}-\frac{1}{\phi}}\right]^{-\phi}\equiv N_{Ejt}, \tag{C5}$$

which is also independent of capital type $(A_{t-\tau},\varepsilon)$.

Level of Labor Demand. We have now shown that all ratios of individual and group level labor demands are independent of capital type $(A_{t-\tau},\varepsilon)$. We now turn to deriving the level of labor demand and showing that it depends linearly on the efficiency units of capital. To do so, we return once more to the first-order condition (C2), this time for the base type $i=i_L$:

$$(1-\alpha)\left[\frac{(A_{t-\tau}\varepsilon)^{\frac{1}{\alpha}}K_{jt}(A_{t-\tau},\varepsilon)}{G(\mathbf{N}_{jt}(A_{t-\tau},\varepsilon))}\right]^{\alpha}\xi\left[\frac{N_{Ljt}(A_{t-\tau},\varepsilon)}{G(\mathbf{N}_{jt}(A_{t-\tau},\varepsilon))}\right]^{-\frac{1}{\varphi}}z_{i_L}\left[\frac{N_{i_Ljt}(A_{t-\tau},\varepsilon)}{N_{Ljt}(A_{t-\tau},\varepsilon)}\right]^{-\frac{1}{\phi}}=\widehat{\chi}_{i_Ljt}. \tag{C6}$$

Note that we can write the ratio

$$\frac{N_{Ljt}(A_{t-\tau},\varepsilon)}{G(\mathbf{N}_{jt}(A_{t-\tau},\varepsilon))}=\frac{N_{Ljt}(A_{t-\tau},\varepsilon)}{N_{Ljt}(A_{t-\tau},\varepsilon)[\xi+(1-\xi)N_{Ejt}^{\frac{\varphi-1}{\varphi}}]^{\frac{\varphi}{\varphi-1}}}=\left[\xi+(1-\xi)N_{Ejt}^{\frac{\varphi-1}{\varphi}}\right]^{-\frac{\varphi}{\varphi-1}},$$

which is independent of $(A_{t-\tau},\varepsilon)$. Similarly, we can write the ratio

$$\frac{N_{i_Ljt}(A_{t-\tau},\varepsilon)}{N_{Ljt}(A_{t-\tau},\varepsilon)}=\left[\sum_{i\in I_L}z_i(n_{ijt}^L)^{\frac{\phi-1}{\phi}}\right]^{-\frac{\phi}{\phi-1}}.$$

Plugging both of these back (C6) gives

$$(1-\alpha)\left[\frac{(A_{t-\tau}\varepsilon)^{\frac{1}{\alpha}}K_{jt}(A_{t-\tau},\varepsilon)}{G(\mathbf{N}_{jt}(A_{t-\tau},\varepsilon))}\right]^{\alpha}\xi\left[\xi+(1-\xi)N_{Ejt}^{\frac{\varphi-1}{\varphi}}\right]^{\frac{1}{\varphi-1}}z_{i_L}\left[\sum_{i\in I_L}z_i(n_{ijt}^e)^{\frac{\phi-1}{\phi}}\right]^{\frac{1}{\phi-1}}=\widehat{\chi}_{i_Ljt}.$$

The only terms that involve $(A_{t-\tau},\varepsilon)$ are in $\frac{(A_{t-\tau}\varepsilon)^{\frac{1}{\alpha}}K_{jt}(A_{t-\tau},\varepsilon)}{G(\mathbf{N}_{jt}(A_{t-\tau},\varepsilon))}$, which we can write as

$$\frac{(A_{t-\tau}\varepsilon)^{\frac{1}{\alpha}}K_{jt}(A_{t-\tau},\varepsilon)}{G(\mathbf{N}_{jt}(A_{t-\tau},\varepsilon))}=\frac{(A_{t-\tau}\varepsilon)^{\frac{1}{\alpha}}K_{jt}(A_{t-\tau},\varepsilon)}{N_{Ljt}(A_{t-\tau},\varepsilon)[\xi+(1-\xi)N_{Ejt}^{\frac{\varphi-1}{\varphi}}]^{\frac{\varphi}{\varphi-1}}}.$$

Plugging this, we have

$$(1-\alpha)\left[\frac{(A_{t-\tau}\varepsilon)^{\frac{1}{\alpha}}K_{jt}(A_{t-\tau},\varepsilon)}{N_{Ljt}(A_{t-\tau},\varepsilon)[\xi+(1-\xi)N_{Ejt}^{\frac{\varphi-1}{\varphi}}]^{\frac{\varphi}{\varphi-1}}}\right]^{\alpha}\xi\left[\xi+(1-\xi)N_{Ejt}^{\frac{\varphi-1}{\varphi}}\right]^{\frac{1}{\varphi-1}}z_{i_L}\left[\sum_{i\in I_L}z_i(n_{ijt}^e)^{\frac{\phi-1}{\phi}}\right]^{\frac{1}{\phi-1}}=\widehat{\chi}_{i_Ljt}.$$

Solve this equation for $N_{Ljt}(A_{t-\tau},\varepsilon)$ to get

$$\begin{aligned}N_{Ljt}(A_{t-\tau},\varepsilon)&=(A_{t-\tau}\varepsilon)^{\frac{1}{\alpha}}K_{jt}(A_{t-\tau},\varepsilon)\times\left[\frac{(1-\alpha)\xi[\xi+(1-\xi)N_{Ejt}^{\frac{\varphi-1}{\varphi}}]^{\frac{1}{\varphi-1}}z_{i_L}[\sum_{i\in I_L}z_i(n_{ijt}^e)^{\frac{\phi-1}{\phi}}]^{\frac{1}{\phi-1}}}{\widehat{\chi}_{i_Ljt}[\xi+(1-\xi)N_{Ejt}^{\frac{\varphi-1}{\varphi}}]^{\frac{\varphi}{\varphi-1}}}\right]^{\frac{1}{\alpha}}\\&\equiv(A_{t-\tau}\varepsilon)^{\frac{1}{\alpha}}K_{jt}(A_{t-\tau},\varepsilon)\times L_{jt}.\end{aligned}\tag{C7}$$

Hence, we've achieved our goal of writing the labor allocation in terms of relative labor demands, which are independent of capital type $(A_{t-\tau},\varepsilon)$, and the level of labor demand, which is linear in the efficiency units of capital $(A_{t-\tau}\varepsilon)^{\frac{1}{\alpha}}$.

Aggregation. We now show how to aggregate different types of capital to derive an aggregate production function. We start from the aggregate labor input $G(\mathbf{N}_{jt}(A_{t-\tau},\varepsilon))$,

$$G(\mathbf{N}_{jt}(A_{t-\tau},\varepsilon))=N_{Ljt}(A_{t-\tau},\varepsilon)[\xi+(1-\xi)N_{Ejt}^{\frac{\varphi-1}{\varphi}}]^{\frac{\varphi}{\varphi-1}}=(A_{t-\tau}\varepsilon)^{\frac{1}{\alpha}}K_{jt}(A_{t-\tau},\varepsilon)L_{jt}[\xi+(1-\xi)N_{Ejt}^{\frac{\varphi-1}{\varphi}}]^{\frac{\varphi}{\varphi-1}}.$$

Note that $K_{jt}(A_{t-\tau},\varepsilon)=K_{jt}(A_{t-\tau})\pi(\varepsilon)$ where $K_{jt}(A_{t-\tau})=(1-\delta)^{\tau}X_{jt-\tau}$ is the total amount of capital with vintage productivity $A_{t-\tau}$. Integrate over ε to get

$$G(\mathbf{N}_{jt}(A_{t-\tau}))=A_{t-\tau}^{\frac{1}{\alpha}}(1-\delta)^{\tau}X_{jt-\tau}L_{jt}[\xi+(1-\xi)N_{Ejt}^{\frac{\varphi-1}{\varphi}}]^{\frac{\varphi}{\varphi-1}}\int\varepsilon^{\frac{1}{\alpha}}\pi(\varepsilon)d\varepsilon,$$

where, abusing notation somewhat, we define $G(\mathbf{N}_{jt}(A_{t-\tau}))=\int G(\mathbf{N}_{jt}(A_{t-\tau},\varepsilon))d\varepsilon$. We normalize the mean of ε such that $\int\varepsilon^{\frac{1}{\alpha}}\pi(\varepsilon)d\varepsilon=1$, in which case we have

$$G(\mathbf{N}_{jt}(A_{t-\tau}))=A_{t-\tau}^{\frac{1}{\alpha}}(1-\delta)^{\tau}X_{jt-\tau}L_{jt}[\xi+(1-\xi)N_{Ejt}^{\frac{\varphi-1}{\varphi}}]^{\frac{\varphi}{\varphi-1}}.$$

Now sum over vintages τ get

$$G(\mathbf{N}_{jt})=L_{jt}[\xi+(1-\xi)N_{Ejt}^{\frac{\varphi-1}{\varphi}}]^{\frac{\varphi}{\varphi-1}}\sum_{\tau=1}^{\infty}A_{t-\tau}^{\frac{1}{\alpha}}(1-\delta)^{\tau}X_{jt-\tau}=L_{jt}[\xi+(1-\xi)N_{Ejt}^{\frac{\varphi-1}{\varphi}}]^{\frac{\varphi}{\varphi-1}}K_{jt},\tag{C8}$$

where $K_{jt}=\sum_{\tau=1}^{\infty}A_{t-\tau}^{\frac{1}{\alpha}}(1-\delta)^{\tau}X_{jt-\tau}$ is the aggregate efficiency units of capital.

Now let's aggregate output $Y_{jt}(A_{t-\tau},\varepsilon)$ across vintages. We have

$$\begin{aligned}Y_{jt}(A_{t-\tau},\varepsilon)&=[(A_{t-\tau}\varepsilon)^{\frac{1}{\alpha}}K_t(A_{t-\tau},\varepsilon)]^{\alpha}[(A_{t-\tau}\varepsilon)^{\frac{1}{\alpha}}K_{jt}(A_{t-\tau},\varepsilon)L_{jt}[\xi+(1-\xi)N_{Ejt}^{\frac{\varphi-1}{\varphi}}]^{\frac{\varphi}{\varphi-1}}]^{1-\alpha}\\&=(A_{t-\tau}\varepsilon)^{\frac{1}{\alpha}}K_{jt}(A_{t-\tau},\varepsilon)(L_{jt}[\xi+(1-\xi)N_{Ejt}^{\frac{\varphi-1}{\varphi}}]^{\frac{\varphi}{\varphi-1}})^{1-\alpha}.\end{aligned}$$

Only the first term depends on $(A_{t-\tau}, \varepsilon)$. Integrating over ε as above gives

$$Y_{jt}(A_{t-\tau}) = A_{t-\tau}^{\frac{1}{\alpha}}(1-\delta)^{\tau} X_{jt-\tau}(L_t[\xi + (1-\xi)N_{Ejt}^{\frac{\varphi-1}{\varphi}}]^{\frac{\varphi}{\varphi-1}})^{1-\alpha}.$$

Summing over vintages gives

$$Y_{jt} = K_{jt}(L_{jt}[\xi + (1-\xi)N_{Ejt}^{\frac{\varphi-1}{\varphi}}]^{\frac{\varphi}{\varphi-1}})^{1-\alpha}.$$

Now, write this as

$$Y_{jt} = K_{jt}^{\alpha} K_{jt}^{1-\alpha}(L_{jt}[\xi + (1-\xi)N_{Ejt}^{\frac{\varphi-1}{\varphi}}]^{\frac{\varphi}{\varphi-1}})^{1-\alpha} = K_{jt}^{\alpha} G(\mathbf{N}_{jt})^{1-\alpha},$$

where the last line uses that $G(\mathbf{N}_{jt}) = L_{jt}[\xi + (1-\xi)N_{Ejt}^{\frac{\varphi-1}{\varphi}}]^{\frac{\varphi}{\varphi-1}} K_{jt}$ from above.

Equivalence of Labor Allocations. Hence, we've shown that an aggregate production function exists in terms of the aggregate efficiency units of capital K_{jt} and the labor aggregate $G(\mathbf{N}_t)$. The last step that we need is to show that the labor aggregate $G(\mathbf{N}_{jt})$ is equal to

$$G(\mathbf{N}_{jt}) = \left[\xi N_{Ljt}^{\frac{\varphi-1}{\varphi}} + (1-\xi)N_{Hjt}^{\frac{\varphi-1}{\varphi}}\right]^{\frac{\varphi}{\varphi-1}} \text{ where } N_{ejt} = \left[\sum_{i\in I_e} z_i (N_{ijt}^{\frac{\phi-1}{\phi}})\right]^{\frac{\phi}{\phi-1}}.$$

This result follows from the derivation of the results above. To see why, start the bottom nest defining N_{ejt}. The steps above show that $N_{ijt}(A_{t-\tau}, \varepsilon) = n_{ijt} \times (A_{t-\tau}\varepsilon)^{\frac{1}{\alpha}} K_{jt}(A_{t-\tau}, \varepsilon)$ for some constant n_{it} which is independent of $A_{t-\tau}\varepsilon$. Plugging this into group-level $N_{ejt}(A_{t-\tau}, \varepsilon)$ gives

$$N_{ejt}(A_{t-\tau}, \varepsilon) = \left[\sum_{i\in I_e} z_i (N_{ijt}(A_{t-\tau}, \varepsilon)^{\frac{\phi-1}{\phi}})\right]^{\frac{\phi}{\phi-1}} = (A_{t-\tau}\varepsilon)^{\frac{1}{\alpha}} K_{jt}(A_{t-\tau}, \varepsilon) \left[\sum_{i\in I_e} z_i (n_{ijt}^{\frac{\phi-1}{\phi}})\right]^{\frac{\phi}{\phi-1}}.$$

Aggregating this across types ε and vintages τ gives

$$\begin{aligned} N_{ejt} &= \sum_{\tau} \int_{\varepsilon} N_{ejt}(A_{t-\tau}, \varepsilon) d\varepsilon \\ &= \left[\sum_{i\in I_e} z_i (n_{ijt}^{\frac{\phi-1}{\phi}})\right]^{\frac{\phi}{\phi-1}} \sum_{\tau} \int_{\varepsilon} (A_{t-\tau}\varepsilon)^{\frac{1}{\alpha}} K_{jt}(A_{t-\tau}, \varepsilon) = \left[\sum_{i\in I_e} z_i (n_{ijt}^{\frac{\phi-1}{\phi}})\right]^{\frac{\phi}{\phi-1}} K_{jt}. \end{aligned} \tag{C9}$$

From the adding up constraint, we have

$$\begin{aligned} N_{ijt} &= \sum_{\tau} \int_{\varepsilon} N_{ijt}(A_{t-\tau}, \varepsilon) d\varepsilon \\ &= \sum_{\tau} \int_{\varepsilon} n_{ijt} \times (A_{t-\tau}\varepsilon)^{\frac{1}{\alpha}} K_{jt}(A_{t-\tau}, \varepsilon) d\varepsilon = n_{ijt} \sum_{\tau} \int_{\varepsilon} (A_{t-\tau}\varepsilon)^{\frac{1}{\alpha}} K_{jt}(A_{t-\tau}, \varepsilon) d\varepsilon = n_{ijt} K_{ijt}. \end{aligned}$$

Hence, $N_{ijt} = n_{ijt} K_{ijt}$. Plugging this in to (C9) gives

$$N_{ejt} = \left[\sum_{i\in I_e} z_i (n_{ijt}^{\frac{\phi-1}{\phi}})\right]^{\frac{\phi}{\phi-1}} K_{jt} = \left[\sum_{i\in I_e} z_i (n_{ijt} K_{ijt})^{\frac{\phi-1}{\phi}}\right]^{\frac{\phi}{\phi-1}} = \left[\sum_{i\in I_e} z_i (N_{ijt}^{\frac{\phi-1}{\phi}})\right]^{\frac{\phi}{\phi-1}},$$

yielding the bottom nest of the CES. To derive the top nest defining $G(\mathbf{N}_{jt})$, aggregate (C7) to get

$$N_{Ljt}(A_{t-\tau},\varepsilon) = (A_{t-\tau}\varepsilon)^{\frac{1}{\alpha}} K_{jt}(A_{t-\tau},\varepsilon) L_{jt} \implies \sum_{\tau}\int_{\varepsilon} N_{Ljt}(A_{t-\tau},\varepsilon) = L_{jt}\sum_{\tau}\int_{\varepsilon}(A_{t-\tau}\varepsilon)^{\frac{1}{\alpha}} K_{jt}(A_{t-\tau},\varepsilon)$$

$$\implies N_{Ljt} = L_{jt}K_{jt}.$$

Similarly, aggregate (C5) to get

$$N_{Hjt}(A_{t-\tau},\varepsilon) = N_{Ejt}N_{Ljt}(A_{t-\tau},\varepsilon) \implies \sum_{\tau}\int_{\varepsilon} N_{Hjt}(A_{t-\tau},\varepsilon) = N_{Ejt}\sum_{\tau}\int_{\varepsilon} N_{Ljt}(A_{t-\tau},\varepsilon)$$

$$\implies N_{Hjt} = N_{Ejt}N_{Ljt}.$$

Plugging these two results into our aggregated labor input $G(\mathbf{N}_{jt})$ from (C8) gives

$$G(\mathbf{N}_{jt}) = Ljt\left[\xi + (1-\xi)N_{Ejt}^{\frac{\varphi-1}{\varphi}}\right]^{\frac{\varphi}{\varphi-1}} K_{jt} \implies G(\mathbf{N}_{jt}) = \frac{N_{Ljt}}{K_{jt}}\left[\xi + (1-\xi)\left(\frac{N_{Hjt}}{N_{Ljt}}\right)^{\frac{\varphi-1}{\varphi}}\right]^{\frac{\varphi}{\varphi-1}} K_{jt}$$

$$\implies G(\mathbf{N}_{jt}) = \left[\xi N_{Ljt}^{\frac{\varphi-1}{\varphi}} + (1-\xi)N_{Hjt}^{\frac{\varphi-1}{\varphi}}\right]^{\frac{\varphi}{\varphi-1}},$$

which is what we wanted to show.

C.2 Solving the Firm Problem

We now embed the putty-putty production structure derived above into the dynamic profit maximization problem of the firm. Recall that the derivation above resulted in the following production function and law of motion for efficiency units of capital

$$\begin{aligned} Y_{jt} &= K_{jt}^{\alpha} G(\mathbf{N}_{jt})^{1-\alpha} \equiv F(K_{jt}, \mathbf{N}_{jt}) \\ G(\mathbf{N}_{jt}) &= \left[\xi N_{Ljt}^{\frac{\varphi-1}{\varphi}} + (1-\xi)N_{Hjt}^{\frac{\varphi-1}{\varphi}}\right]^{\frac{\varphi}{\varphi-1}} \\ N_{ejt} &= \left[\sum_{i\in I_e} z_i N_{ijt}^{\frac{\phi-1}{\phi}}\right]^{\frac{\phi}{\phi-1}} \\ K_{jt+1} &= (1-\delta)K_{jt} + A_t^{\frac{1}{\alpha}} X_{jt}, \end{aligned}$$

where $\mathbf{N}_{jt}$ denotes the vector of all types of workers employed at firm j and recall that $A_t = (1+g_A)^t$. In addition, we assume that there are adjustment costs to changing the efficiency units of capital. In particular, X_{jt} units of investment incurs the following costs, in units of output:

$$-\frac{\zeta_t}{2}\left[\frac{A_t^{\frac{1}{\alpha}} X_{jt}}{K_{jt}} - (\delta + g_K)\right]^2 K_{jt},$$

where g_K is the growth rate of efficiency units of capital along the BGP (to be derived below) and the parameter $\zeta_t = (1+g_\zeta)^t\zeta$ grows at rate $1+g_\zeta = (1+g_A)^{-\frac{1}{\alpha}}$. This growth rate $g_\zeta < 0$ ensures

that the product $\zeta_t K_{jt}$ grows at the rate of physical output, rather than efficiency units of capital, which we show below is necessary for balanced growth.

C.2.1 Setting Up the Firm Problem

We solve the firm's problem using the same approach as in our putty-clay model. It will be convenient to have the firm choose per capita employment rates n_{ijt} rather than aggregate employment $N_{ijt} = \mu_i n_{ijt}$ because the per-capita employment rate enters the household's participation constraint. The firm's objective function is

$$F(K_{j0}, \widehat{\mathbf{N}}_{j0}) + \sum_i \mu_i f_{ij0} \widehat{W}_{i0} - \sum_i \mu_i \kappa_{i0} a_{ij0} - X_{j0} - \frac{\zeta}{2}\Big[\frac{X_{j0}}{K_{j0}} - (\delta + g_K)\Big]^2 K_{j0}$$

$$+\sum_{t=1}^{\infty} Q_{0,t}\Big[F(K_{jt}, \mathbf{N}_{jt}) - \sum_i \mu_i\big(\lambda_f(\theta_{ijt-1}) a_{ijt-1} W_{ijt} + \kappa_{it} a_{ijt}\big) - X_{jt} - \frac{\zeta_t}{2}\big(\frac{A_t^{\frac{1}{\alpha}} X_{jt}}{K_{jt}} - (\delta + g_K)\big)^2 K_{jt}\Big],$$

where we have normalized $A_0 = 1$. Here, we've separated out decisions in the initial period $t = 0$ from the remaining period because firms have the ability to fire workers only in that period. $\widehat{N}_{j0}$ denotes the measure of workers used in production after firing, i.e. $\widehat{N}_{ij0} = \mu_i(n_{ij0} - f_{ij0})$, where n_{ij0} is the per-capita number of workers at the beginning of the period. $\widehat{W}_{i0}$ denotes the present value of wage payments promised to those initially employed workers, taking into account both any policy changes and the changes in prices $Q_{0,t}$. Hence, firing f_{ij0} workers has the cost of giving up production today and in the future, because it affects the law of motion for employment as we'll see below. On the other hand, firing f_{ij0} workers has the benefit of saving on wage payments.

This problem is subject to the constraints (with associated Lagrange multipliers):

$$\begin{aligned}
n_{ij1} &\leq (1-\sigma)(n_{ij0} - f_{ij0}) + \lambda_f(\theta_{ij0}) a_{ij0} \quad (\times Q_{0,1}\mu_i \nu_{ij1}) \\
f_{ij0} &\geq 0 \quad (\times \xi^f_{ij0}) \\
n_{ijt+1} &\leq (1-\sigma) n_{ijt} + \lambda_f(\theta_{ijt}) a_{ijt} \quad (Q_{0,t+1}\mu_i \nu_{ijt+1}) \\
W_{ijt+1} + V_{ijt+1} &\geq \frac{\mathcal{W}_{it}}{Q_{t,t+1}\lambda_w(\theta_{ijt})} \quad (\times Q_{0,t+1}\mu_i \gamma_{ijt+1}) \\
W_{ijt+1} &\geq \overline{W}_{t+1} \quad (\times Q_{0,t+1}\mu_i \rho_{ijt+1}) \\
K_{jt+1} &\leq (1-\delta) K_{jt} + A_t^{\frac{1}{\alpha}} X_{jt} \quad (\times Q_{0,t} q_{jt}).
\end{aligned}$$

It is important to remember that, in the participation constraint $W_{ijt+1} + V_{ijt+1} \geq \frac{\mathcal{W}_{it}}{Q_{t,t+1}\lambda_w(\theta_{ijt})}$, the present value of disutilities V_{ijt+1} depends on the firm's choices of n_{ijt}. As in the main text, we will collect common terms following Marcet and Marimon (2019), which summarizes the multipliers

on those terms with the auxiliary variable $M_{ijt+1} = (1-\sigma)M_{ijt} + \gamma_{ijt+1}$. The Lagrangian becomes

$$
\begin{aligned}
&F(K_{j0}, \widehat{\mathbf{N}}_{j0}) + \sum_i \mu_i f_{ij0} \widehat{W}_{i0} - \sum_i \mu_i \kappa_{i0} a_{ij0} - X_{j0} - \frac{\zeta}{2}\big(\frac{X_{j0}}{K_{j0}} - (\delta + g_K)\big)^2 K_{j0} + \xi^f_{ij0} f_{ij0} \\
&+ q_{j0}\big[(1-\delta)K_{j0} + X_{j0} - K_{j1}\big] + Q_{0,1}\mu_i \nu_{ij1}\big[(1-\sigma)(n_{ij0} - f_{ij0}) + \lambda_f(\theta_{ij0})a_{ij0} - n_{ij1}\big] \\
&+ \sum_{t=1}^{\infty} Q_{0,t} \left\{ \begin{array}{l} F(K_{jt}, \mathbf{N}_{jt}) - \sum_i \mu_i\big[\lambda_f(\theta_{ijt-1})a_{ijt-1}W_{ijt} + \kappa_{it}a_{ijt}\big] - X_{jt} \\ - \frac{\zeta_t}{2}\Big[\frac{A_t^{\frac{1}{\alpha}} X_{jt}}{K_{jt}} - (\delta + g_K)\Big]^2 K_{jt} + q_{jt}\big[(1-\delta)K_{jt} + A_t^{\frac{1}{\alpha}} X_{jt} - K_{jt+1}\big] \\ + Q_{t,t+1}\mu_i \nu_{ijt+1}\big[(1-\sigma)n_{ijt} + \lambda_f(\theta_{ijt})a_{ijt} - n_{ijt+1}\big] + Q_{t,t+1}\mu_i \rho_{ijt+1}\big(W_{ijt+1} - \overline{W}_{t+1}\big) \end{array} \right\} \\
&+ \sum_{t=0}^{\infty} Q_{0,t+1}\Big[\mu_i M_{ijt+1}\frac{U_{nit+1}}{U_{cit+1}}\Big(\frac{n_{ijt+1}}{n_{it+1}}\Big)^{\frac{1}{\omega}} + \mu_i \gamma_{ijt+1}\Big(W_{ijt+1} - \frac{\mathcal{W}_{it}}{Q_{t,t+1}\lambda_w(\theta_{ijt})}\Big)\Big].
\end{aligned}
$$

C.2.2 First-Order Conditions

We now take first order conditions of this problem. Note that there is no "utilization block" in the putty-putty model, so we only have the hiring block and investment block.

Hiring Block. The first order condition for employment n_{jt+1} for $t \geq 0$ is

$$
\begin{aligned}
&-Q_{0,t+1}\mu_i \nu_{ijt+1} + Q_{0,t+1}\mu_i F_i(K_{jt+1}, \mathbf{N}_{jt+1}) + Q_{0,t+1}\mu_i M_{ijt+1}\frac{U_{nit+1}}{U_{cit+1}}\frac{1}{\omega}\big(\frac{n_{ijt+1}}{n_{it+1}}\big)^{\frac{1}{\omega}-1}\frac{1}{n_{it+1}} \\
&\qquad + Q_{0,t+2}\mu_i(1-\sigma)\nu_{ijt+2} = 0,
\end{aligned}
$$

where $F_i(K_{jt}, \mathbf{N}_{jt}) = \frac{\partial F(K_{jt}, \mathbf{N}_{jt})}{\partial N_{ijt}}$ so that $\frac{\partial F(K_{jt}, \mathbf{N}_{jt})}{\partial n_{ijt}} = \frac{\partial F(K_{jt}, \mathbf{N}_{jt})}{\partial N_{ijt}}\frac{\partial N_{ijt}}{\partial n_{ijt}} = \mu_i F_i(K_{jt}, \mathbf{N}_{jt})$. Simplifying the FOC above gives

$$
\nu_{ijt+1} = F_i(K_{jt+1}, \mathbf{N}_{jt+1}) + M_{ijt+1}\frac{U_{nit+1}}{U_{cit+1}}\frac{1}{\omega}\big(\frac{n_{ijt+1}}{n_{it+1}}\big)^{\frac{1}{\omega}-1}\frac{1}{n_{it+1}} + Q_{t+1,t+2}(1-\sigma)\nu_{ijt+2}. \tag{C10}
$$

This equation is similar to the corresponding FOC in the putty-clay model except that the μ_{ijt+1} in the putty-clay model is replaced with the marginal product $F_i(K_{jt+1}, \mathbf{N}_{jt+1})$. The first-order conditions for vacancy posting a_{ijt}, market tightness θ_{ijt}, and wages W_{ijt+1} are the same as in the putty-clay model since none of those variables interact with the production function.

The first-order condition for the first period firing f_{ij0} is

$$
-\mu_i F_i(K_{j0}, \widehat{\mathbf{N}}_{j0}) + \mu_i \widehat{W}_{ij0} - Q_{0,1}\mu_i(1-\sigma)\nu_{ij1} + \mu_i \xi^f_{ij0} = 0,
$$

which uses that $\frac{\partial F(K_{j0}, \widehat{\mathbf{N}}_{j0})}{\partial f_{ij0}} = \frac{\partial F(K_{j0}, \widehat{\mathbf{N}}_{j0})}{\partial \widehat{N}_{ij0}}\frac{\partial \widehat{N}_{ij0}}{\partial f_{ij0}} = -\mu_i F_i(K_{j0}, \widehat{N}_{j0})$, or, equivalently,

$$
F_i(K_{j0}, \widehat{\mathbf{N}}_{j0}) + Q_{0,1}(1-\sigma)\nu_{ij1} \geq \widehat{W}_{ij0}, \text{ with equality if } f_{ij0} > 0. \tag{C11}
$$

This expression is also similar to the corresponding FOC in the putty-clay model except that the initial χ_{ij0} is replaced with the marginal product $F_i(K_{j0}, \widehat{\mathbf{N}}_{j0})$.

Investment Block. The first-order condition for investment X_{jt} is

$$-Q_{0,t}\big[1+\zeta_t\big(\frac{A_t^{\frac{1}{\alpha}}X_{jt}}{K_{jt}}-(\delta+g_K)\big)A_t^{\frac{1}{\alpha}}\big]+Q_{0,t}A_t^{\frac{1}{\alpha}}q_{jt}=0 \implies A_t^{\frac{1}{\alpha}}q_{jt}=\big[1+\zeta_t\big(\frac{A_t^{\frac{1}{\alpha}}X_{jt}}{K_{jt}}-(\delta+g_K)\big)A_t^{\frac{1}{\alpha}}\big]. \tag{C12}$$

The first-order condition for capital K_{jt+1} is

$$\begin{aligned}Q_{0,t+1}F_k(K_{jt+1},\mathbf{N}_{jt+1})-Q_{0,t}q_{jt}+Q_{0,t+1}(1-\delta)q_{jt+1}-Q_{0,t+1}\frac{\zeta_{t+1}}{2}\big(\frac{A_{t+1}^{\frac{1}{\alpha}}X_{jt+1}}{K_{jt+1}}-(\delta+g_K)\big)^2\\ +Q_{0,t+1}\zeta_{t+1}\big(\frac{A_{t+1}^{\frac{1}{\alpha}}X_{jt+1}}{K_{jt+1}}-(\delta+g_K)\big)\frac{A_{t+1}^{\frac{1}{\alpha}}X_{jt+1}}{K_{jt+1}}=0.\end{aligned}$$

Rearranging this expression gives

$$q_{jt}=Q_{t,t+1}\left[\begin{aligned}F_k(K_{jt+1},\mathbf{N}_{jt+1}-\frac{\zeta_{t+1}}{2}\big(\frac{A_{t+1}^{\frac{1}{\alpha}}X_{jt+1}}{K_{jt+1}}-(\delta+g_K)\big)^2\\ +\zeta_{t+1}\big(\frac{A_{t+1}^{\frac{1}{\alpha}}X_{jt+1}}{K_{jt+1}}-(\delta+g_K)\big)\frac{A_{t+1}^{\frac{1}{\alpha}}X_{jt+1}}{K_{jt+1}}+(1-\delta)q_{jt+1}\end{aligned}\right]. \tag{C13}$$

In the special case with no adjustment costs $\zeta_t=0$, solving this equation forward gives

$$1=A_t^{\frac{1}{\alpha}}\sum_{\tau=1}^{\infty}Q_{t,t+\tau}(1-\delta)^{\tau-1}\big[F_k(K_{jt+\tau},\mathbf{N}_{jt+\tau})\big]$$

analogously to our "zero-profit" condition for the optimal capital type in the putty-clay model except that the RHS involves the marginal product of capital in F. By Euler's theorem, the latter is related to output net of labor costs yielding a similar expression to that in the putty-clay model.

Other Conditions. The only other condition to change in the putty-putty model is the market clearing condition for output. In a symmetric equilibrium, this condition is

$$Y_t=F(K_t,\mathbf{N}_t)=\sum_i\mu_i c_{it}+\sum_i\mu_i\kappa_{it}a_{it}+X_t+\frac{\zeta_t}{2}\big(\frac{A_t^{\frac{1}{\alpha}}X_t}{K_t}-(\delta+g_K)\big)^2K_t.$$

This condition differs from its analog in the putty-clay model in two ways. First, aggregate output is given by the production function $F(K_t,\mathbf{N}_t)$ rather than an aggregation of capital types. Second, adjustment costs are taken out of output. Note that adjustment costs are zero along the BGP.

Summary. We summarize the conditions that differ from the putty-clay model below. In this system, we impose a symmetric equilibrium, allowing us to drop the j subscripts:

$$\nu_{it+1} = F_i(K_{t+1}, \mathbf{N}_{t+1}) + M_{it+1}\frac{1}{\omega}\frac{U_{nit+1}}{U_{cit+1}}\frac{1}{n_{it+1}} + Q_{t+1,t+2}(1-\sigma)\nu_{it+2} \tag{C14}$$

$$\widehat{W}_{i0} \leq F_i(K_0, \widehat{\mathbf{N}}_0) + Q_{0,1}(1-\sigma)\nu_{i1}, \text{ with equality if } f_{i0} > 0 \tag{C15}$$

$$A_t^{\frac{1}{\alpha}} q_{jt} = \big[1 + \zeta_t\big(\frac{A_t^{\frac{1}{\alpha}} X_{jt}}{K_{jt}} - (\delta + g_K)\big) A_t^{\frac{1}{\alpha}}\big] \tag{C16}$$

$$q_t = Q_{t,t+1}\begin{bmatrix} F_k(K_{t+1}, \mathbf{N}_{t+1}) - \frac{\zeta_{t+1}}{2}\big(\frac{A_{t+1}^{\frac{1}{\alpha}} X_{t+1}}{K_{t+1}} - (\delta + g_K)\big)^2 \\ + \zeta_{t+1}\big(\frac{A_{t+1}^{\frac{1}{\alpha}} X_{t+1}}{K_{t+1}} - (\delta + g_K)\big)\frac{A_{t+1}^{\frac{1}{\alpha}} X_{t+1}}{K_{t+1}} + (1-\delta) q_{t+1} \end{bmatrix} \tag{C17}$$

$$K_{t+1} = (1-\delta)K_t + A_t^{\frac{1}{\alpha}} X_t \tag{C18}$$

$$Y_t = F(K_t, \mathbf{N}_t) \tag{C19}$$

$$Y_t = \sum_i \mu_i c_{it} + \sum_i \mu_i \kappa_{it} a_{it} + X_t + \frac{\zeta_t}{2}\big(\frac{A_t^{\frac{1}{\alpha}} X_t}{K_t} - (\delta + g_K)\big)^2 K_t. \tag{C20}$$

C.2.3 Detrending

We now detrend the model relative to its BGP; tildes denote detrended variables. We show that:

- The following variables grow at rate $1 + g = (1 + g_A)^{\frac{1}{1-\alpha}}$: ν_{it}, X_t, Y_t (in addition to all the other variables from the putty-clay model which we already showed grow at that rate). We will sometimes refer to this growth rate g as the growth rate of "physical units."
- The efficiency units of capital K_t grows at rate $1 + g_K = (1 + g)^{\frac{1}{\alpha}}$. We will sometimes refer to this as the growth rate of "efficiency units." Note that $g_K > g$.
- The adjustment cost parameter ζ_t shrinks at rate $1 + g_\zeta = (1 + g_A)^{-\frac{1}{\alpha}} = (1 + g)^{-\frac{1-\alpha}{\alpha}}$.
- The same variables are stationary as in the putty-clay model. In addition, $A_t^{\frac{1}{\alpha}} q_{jt}$ is stationary.

Capital Accumulation Equation. To see where how the growth rate of efficiency units is related to the growth rate of physical units, detrend the capital accumulation equation (C18):

$$\begin{aligned}
&(1 + g_K)^{t+1}\widetilde{K}_{t+1} = (1-\delta)(1 + g_K)^t \widetilde{K}_t + [(1 + g_A)^t]^{\frac{1}{\alpha}}(1 + g)^t \widetilde{X}_t \\
&\implies (1 + g_K)^{t+1}\widetilde{K}_{t+1} = (1-\delta)((1 + g_K)^t \widetilde{K}_t + [((1 + g)^{1-\alpha})^t]^{\frac{1}{\alpha}}(1 + g)^t \widetilde{X}_t \\
&\implies (1 + g_K)^{t+1}\widetilde{K}_{t+1} = (1-\delta)((1 + g_K)^t \widetilde{K}_t + ((1 + g)^{\frac{1}{\alpha}})^t \widetilde{X}_t = (1-\delta)((1 + g_K)^t \widetilde{K}_t + (1 + g_K)^t \widetilde{X}_t \\
&\implies (1 + g_K)\widetilde{K}_{t+1} = (1-\delta)\widetilde{K}_t + \widetilde{X}_t.
\end{aligned}$$

The growth rate of efficiency units is $g_K > g$ as it is driven by the growth in the physical amount of investment, which occurs at rate g, and in vintage productivity, which occurs at rate $(1+g_A)^{\frac{1}{\alpha}}$.

Production Function. Next, detrend the aggregate production function (C19):

$$Y_t = K_t^{\alpha} G(\mathbf{N}_t)^{1-\alpha} \implies (1+g)^t \widetilde{Y}_t = [(1+g_K)^{\alpha}]^t \widetilde{K}_t^{\alpha} G(\mathbf{N}_t)^{1-\alpha}$$
$$\implies (1+g)^t \widetilde{Y}_t = [1+g]^t \widetilde{K}_t^{\alpha} G(\mathbf{N}_t)^{1-\alpha} \implies \widetilde{Y}_t = \widetilde{K}_t^{\alpha} G(\mathbf{N}_t)^{1-\alpha}.$$

Even though efficiency units are growing at the rate $(1+g_K) = (1+g)^{\frac{1}{\alpha}} > 1+g$, output still grows at rate $1+g$ because capital is raised to the power α in the production function.

Hiring Block. To detrend (C14), note that the marginal product of labor grows at rate g so

$$F_i(K_t, \mathbf{N}_t) = ((1+g_K)^t \widetilde{K}_t)^{\alpha}(1-\alpha) G(\mathbf{N}_t)^{-\alpha} G_i(\mathbf{N}_t) = [(1+g_K)^{\alpha}]^t \widetilde{K}_t{}^{\alpha}(1-\alpha) G(\mathbf{N}_t)^{-\alpha} G_i(\mathbf{N}_t)$$
$$\implies F_i(K_t, \mathbf{N}_t) = (1+g)^t \widetilde{K}_t{}^{\alpha}(1-\alpha) G(\mathbf{N}_t)^{-\alpha} G_i(\mathbf{N}_t) = (1+g)^t F_i(\widetilde{K}_t, \mathbf{N}_t).$$

Using this result and the functional form of the utility function, we can detrend (C14):

$$(1+g)^{t+1}\widetilde{\nu}_{it+1} = (1+g)^{t+1} F_i(\widetilde{K}_{t+1}, \mathbf{N}_{t+1}) - \frac{M_{it+1}}{\omega}(1+g)^{t+1}\frac{v'(n_{it+1})}{n_{it+1}} + Q_{t+1,t+2}(1-\sigma)(1+g)^{t+2}\widetilde{\nu}_{it+2}$$
$$\implies \widetilde{\nu}_{it+1} = F_i(\widetilde{K}_{t+1}, \mathbf{N}_{t+1}) - \frac{M_{it+1}}{\omega}\frac{v'(n_{it+1})}{n_{it+1}} + Q_{t+1,t+2}(1-\sigma)(1+g)\widetilde{\nu}_{it+2}.$$

Since variables in period $t=0$ require no adjustments, the firing condition in detrended terms is

$$F_i(K_0, \widehat{\mathbf{N}}_0) + Q_{0,1}(1-\sigma)(1+g)\widetilde{\nu}_{i1} \geq \widehat{W}_{i0}, \text{ with equality if } f_{i0} > 0.$$

Investment Block. We now turn to the first order condition for investment (C16). Define $\widehat{q}_{jt} = A_t^{\frac{1}{\alpha}} q_{jt}$. As described in the beginning of this section, it turns out that this variable will be stationary. In this case, detrending the first order condition for investment (C16) gives

$$A_t^{\frac{1}{\alpha}} q_t = 1 + \zeta_t\Big(\frac{A_t^{\frac{1}{\alpha}} X_{jt}}{K_{jt}} - (\delta + g_K)\Big) A_t^{\frac{1}{\alpha}}$$
$$\implies \widehat{q}_t = 1 + (1+g_\zeta)^t \zeta\Big(\frac{[(1+g_A)^{\frac{1}{\alpha}}]^t (1+g)^t \widetilde{X}_t}{(1+g_K)^t \widetilde{K}_t} - (\delta + g_K)\Big)[(1+g_A)^{\frac{1}{\alpha}}]^t$$
$$\implies \widehat{q}_t = 1 + [(1+g_\zeta)(1+g_A)^{\frac{1}{\alpha}}]^t \zeta \left(\left[\frac{(1+g_A)^{\frac{1}{\alpha}}(1+g)}{(1+g_K)}\right]^t \frac{\widetilde{X}_t}{\widetilde{K}_t} - (\delta + g_K)\right) = 1 + \zeta\Big(\frac{\widetilde{X}_t}{\widetilde{K}_t} - (\delta + g_K)\Big).$$

The last equality follows for two reasons. First, $1+g_\zeta = (1+g_A)^{-\frac{1}{\alpha}}$ so that $(1+g_\zeta)(1+g_A)^{\frac{1}{\alpha}} = 1$, that is, the shrinking adjustment cost parameter ζ_t offsets the increasing efficiency units $A_t^{\frac{1}{\alpha}}$. Second, $1+g_K = (1+g)^{\frac{1}{\alpha}} = (1+g)(1+g)^{\frac{1-\alpha}{\alpha}} = (1+g)(1+g_A)^{\frac{1}{\alpha}}$, which implies that $\frac{(1+g_A)^{\frac{1}{\alpha}}(1+g)}{(1+g_K)} = 1$.

We now turn to the definition of q_{jt} in (C17) Multiply by $A_t^{\frac{1}{\alpha}}$ to get an expression for $\widehat{q}_{jt}$:

$$
\begin{aligned}
q_t &= Q_{t,t+1}\big[F_k(K_{t+1},\mathbf{N}_{t+1}) - \frac{\zeta_{t+1}}{2}\big(\frac{A_{t+1}^{\frac{1}{\alpha}}X_{t+1}}{K_{t+1}} - (\delta+g_K)\big)^2 \\
&\quad + \zeta_{t+1}\big(\frac{A_{t+1}^{\frac{1}{\alpha}}X_{t+1}}{K_{t+1}} - (\delta+g_K)\big)\frac{A_{t+1}^{\frac{1}{\alpha}}X_{t+1}}{K_{t+1}} + (1-\delta)q_{t+1}\big] \\
\implies A_t^{\frac{1}{\alpha}}q_t &= Q_{t,t+1}\big[A_t^{\frac{1}{\alpha}}F_k(K_{t+1},\mathbf{N}_{t+1}) - \frac{A_t^{\frac{1}{\alpha}}\zeta_{t+1}}{2}\big(\frac{A_{t+1}^{\frac{1}{\alpha}}X_{t+1}}{K_{t+1}} - (\delta+g_K)\big)^2 \\
&\quad + A_t^{\frac{1}{\alpha}}\zeta_{t+1}\big(\frac{A_{t+1}^{\frac{1}{\alpha}}X_{t+1}}{K_{t+1}} - (\delta+g_K)\big)\frac{A_{t+1}^{\frac{1}{\alpha}}X_{t+1}}{K_{t+1}} + (1-\delta)A_t^{\frac{1}{\alpha}}q_{t+1}\big] \\
\implies A_t^{\frac{1}{\alpha}}q_t &= Q_{t,t+1}\big(\frac{A_t}{A_{t+1}}\big)^{\frac{1}{\alpha}}\big[A_{t+1}^{\frac{1}{\alpha}}F_k(K_{t+1},\mathbf{N}_{t+1}) - \frac{A_{t+1}^{\frac{1}{\alpha}}\zeta_{t+1}}{2}\big(\frac{A_{t+1}^{\frac{1}{\alpha}}X_{t+1}}{K_{t+1}} - (\delta+g_K)\big)^2 \\
&\quad + A_{t+1}^{\frac{1}{\alpha}}\zeta_{t+1}\big(\frac{A_{t+1}^{\frac{1}{\alpha}}X_{t+1}}{K_{t+1}} - (\delta+g_K)\big)\frac{A_{t+1}^{\frac{1}{\alpha}}X_{t+1}}{K_{t+1}} + (1-\delta)A_{t+1}^{\frac{1}{\alpha}}q_{t+1}\big] \\
\implies \widehat{q}_t &= Q_{t,t+1}\big(1+g_A\big)^{-\frac{1}{\alpha}}\big[A_{t+1}^{\frac{1}{\alpha}}F_k(K_{t+1},\mathbf{N}_{t+1}) - \frac{A_{t+1}^{\frac{1}{\alpha}}\zeta_{t+1}}{2}\big(\frac{A_{t+1}^{\frac{1}{\alpha}}X_{t+1}}{K_{t+1}} - (\delta+g_K)\big)^2 \\
&\quad + A_{t+1}^{\frac{1}{\alpha}}\zeta_{t+1}\big(\frac{A_{t+1}^{\frac{1}{\alpha}}X_{t+1}}{K_{t+1}} - (\delta+g_K)\big)\frac{A_{t+1}^{\frac{1}{\alpha}}X_{t+1}}{K_{t+1}} + (1-\delta)\widehat{q}_{t+1}\big].
\end{aligned}
$$

Above, we showed that $A_{t+1}^{\frac{1}{\alpha}}\zeta_{t+1} = \zeta$ and that $\frac{A_{t+1}^{\frac{1}{\alpha}}X_{t+1}}{K_{t+1}} = \frac{\widetilde{X}_{t+1}}{\widetilde{K}_{t+1}}$. Plugging those results in gives

$$
\widehat{q}_t = Q_{t,t+1}\big(1+g_A\big)^{-\frac{1}{\alpha}}\big[A_{t+1}^{\frac{1}{\alpha}}F_k(K_{t+1},\mathbf{N}_{t+1}) - \frac{\zeta}{2}\big(\frac{\widetilde{X}_{t+1}}{\widetilde{K}_{t+1}} - (\delta+g_K)\big)^2 + \zeta\big(\frac{\widetilde{X}_{t+1}}{\widetilde{K}_{t+1}} - (\delta+g_K)\big)\frac{\widetilde{X}_{t+1}}{\widetilde{K}_{t+1}} + (1-\delta)\widehat{q}_{t+1}\big].
$$

We are left with detrending the marginal product term $A_{t+1}^{\frac{1}{\alpha}}F_k(K_{t+1},\mathbf{N}_{t+1})$, which becomes

$$
\begin{aligned}
& A_{t+1}^{\frac{1}{\alpha}}F_k(K_{t+1},\mathbf{N}_{t+1}) = \alpha A_{t+1}^{\frac{1}{\alpha}}K_{t+1}^{\alpha-1}G(\mathbf{N}_{t+1})^{1-\alpha} \\
\implies & A_{t+1}^{\frac{1}{\alpha}}F_k(K_{t+1},\mathbf{N}_{t+1}) = \alpha[(1+g_A)^{\frac{1}{\alpha}}]^{t+1}[(1+g_K)^{t+1}]^{\alpha-1}\widetilde{K}_{t+1}^{\alpha-1}G(\mathbf{N}_{t+1})^{1-\alpha} \\
\implies & A_{t+1}^{\frac{1}{\alpha}}F_k(K_{t+1},\mathbf{N}_{t+1}) = [(1+g_A)^{\frac{1}{\alpha}}(1+g_K)^{\alpha-1}]^{t+1}\alpha\widetilde{K}_{t+1}^{\alpha-1}G(\mathbf{N}_{t+1})^{1-\alpha} \\
\implies & A_{t+1}^{\frac{1}{\alpha}}F_k(K_{t+1},\mathbf{N}_{t+1}) = [(1+g)^{\frac{1-\alpha}{\alpha}}(1+g)^{\frac{\alpha-1}{\alpha}}]^{t+1}\alpha\widetilde{K}_{t+1}^{\alpha-1}G(\mathbf{N}_{t+1})^{1-\alpha} \\
\implies & A_{t+1}^{\frac{1}{\alpha}}F_k(K_{t+1},\mathbf{N}_{t+1}) = \alpha\widetilde{K}_{t+1}^{\alpha-1}G(\mathbf{N}_{t+1})^{1-\alpha} = F_k(\widetilde{K}_{t+1},\mathbf{N}_{t+1}).
\end{aligned}
$$

The second to last line uses that $1+g_A = (1+g)^{1-\alpha}$ and $1+g_K = (1+g)^{\frac{1}{\alpha}}$. We thus obtain

$$
\widehat{q}_t = Q_{t,t+1}\big(1+g_A\big)^{-\frac{1}{\alpha}}\big[F_k(\widetilde{K}_{t+1},\mathbf{N}_{t+1}) - \frac{\zeta}{2}\big(\frac{\widetilde{X}_{t+1}}{\widetilde{K}_{t+1}} - (\delta+g_K)\big)^2 + \zeta\big(\frac{\widetilde{X}_{t+1}}{\widetilde{K}_{t+1}} - (\delta+g_K)\big)\frac{\widetilde{X}_{t+1}}{\widetilde{K}_{t+1}} + (1-\delta)\widehat{q}_{t+1}\big].
$$

Production Function. Detrending the production function (C19) gives

$$Y_t = F(K_t, \mathbf{N}_t) = K_t^{\alpha} G(\mathbf{N}_t)^{1-\alpha} \implies (1+g)^t \widetilde{Y}_t = [(1+g_K)^t \widetilde{K}_t]^{\alpha} G(\mathbf{N}_t)^{1-\alpha}$$
$$\implies (1+g)^t \widetilde{Y}_t = (1+g)^t \widetilde{K}_t^{\alpha} G(\mathbf{N}_t)^{1-\alpha} \implies \widetilde{Y}_t = \widetilde{K}_t^{\alpha} G(\mathbf{N}_t)^{1-\alpha} = F(\widetilde{K}_t, \mathbf{N}_t).$$

The second to last line uses the fact that $(1+g_K) = (1+g)^{\alpha}$.

Output Market Clearing. Finally, we turn to the output market clearing condition (C20):

$$Y_t = \sum_i \mu_i c_{it} + \sum_i \mu_i \kappa_{it} a_{it} + X_t + \frac{\zeta_t}{2}\Big(\frac{A_t^{\frac{1}{\alpha}} X_t}{K_t} - (\delta + g_K)\Big)^2 K_t$$

$$\Rightarrow (1+g)^t \widetilde{Y}_t = (1+g)^t \sum_i \mu_i \widetilde{c}_{it} + (1+g)^t \sum_i \mu_i \kappa_i a_{it} + (1+g)^t \widetilde{X}_t + (1+g_\zeta)^t \frac{\zeta}{2}\Big(\frac{\widetilde{X}_t}{\widetilde{K}_t} - (\delta + g_K)\Big)^2 (1+g_K)^t \widetilde{K}_t$$

$$\Rightarrow (1+g)^t \widetilde{Y}_t = (1+g)^t \sum_i \mu_i \widetilde{c}_{it} + (1+g)^t \sum_i \mu_i \kappa_i a_{it} + (1+g)^t \widetilde{X}_t + [(1+g_\zeta)(1+g_K)]^t \frac{\zeta}{2}\Big(\frac{\widetilde{X}_t}{\widetilde{K}_t} - (\delta + g_K)\Big)^2 \widetilde{K}_t$$

$$\Rightarrow (1+g)^t \widetilde{Y}_t = (1+g)^t \sum_i \mu_i \widetilde{c}_{it} + (1+g)^t \sum_i \mu_i \kappa_i a_{it} + (1+g)^t \widetilde{X}_t + (1+g)^t \frac{\zeta}{2}\Big(\frac{\widetilde{X}_t}{\widetilde{K}_t} - (\delta + g_K)\Big)^2 \widetilde{K}_t$$

$$\Rightarrow \widetilde{Y}_t = \sum_i \mu_i \widetilde{c}_{it} + \sum_i \mu_i \kappa_i a_{it} + \widetilde{X}_t + \frac{\zeta}{2}\Big(\frac{\widetilde{X}_t}{\widetilde{K}_t} - (\delta + g_K)\Big)^2 \widetilde{K}_t.$$

The second to last line uses the fact that $(1+g_\zeta)(1+g_K) = (1+g_A)^{-\frac{1}{\alpha}}(1+g)(1+g_A)^{\frac{1}{\alpha}} = (1+g)$.

D 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).[30] Our sample includes all individuals 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.

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

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*
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 calculate hourly wages for our sample of full-time workers by dividing annual labor earning by annual hours worked. We determine 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.

E 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, μ_{it}. Section E.1 shows how we extend our model to incorporate time-varying measures of families. Section E.2 then explains how we choose the specific path of μ_{it} to mimic the empirical variation utilized by Card and Lemieux (2001).

Finally, Section E.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).

E.1 Model Extension

We model changes in the measure of families, μ_{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 μ_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 μ_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 μ_{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 μ_i from the baseline replaced by the time-varying measure μ_{it}. Results are available upon request.

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

Card and Lemieux (2001) estimate the elasticity of substitution ϕ 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, μ_i, changes over time as denoted by μ_{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 μ_{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).[31] 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 steady state of our model. Hence, we will choose the time path of the measures of families μ_{it} such that the percentage change in the model-generated employment series relative to the final BGP corresponds to Card and Lemieux (2001)'s variation relative to the end of their sample.[32] We construct the percentage changes in the employment

[31] 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.

[32] The last year of the model's 35-year sample does not exactly equal the new BGP as it takes time for the economy to converge to the new BGP once the measures of families stop changing. Yet, we find that the two points are close.

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 on the BGP of our model, determine an initial value a_{j0} from the 1960 Census data described in Appendix D, 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 μ_{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 i in t). In principle, this step requires solving a complicated fixed-point problem. Namely, for any candidate path of measures μ_{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 μ_{it}. Computationally, this procedure is prohibitively costly because it requires solving for the entire transition path for each candidate μ_{it}.

Instead, we compute the path of measures μ_{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 μ_{i0}, where we choose μ_{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 μ_{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 μ_{iT} correspond to the BGP of our quantitative model from Section 5. In order to construct the time path of measures in between these two points, we compute the path of μ_{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 μ_{it}.

E.3 Results

Given the variation described, we construct an estimator for the long-run elasticity of substitution ϕ 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 ϕ. 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

$$\Delta \log\left(\frac{w_{it}}{w_{jt}}\right) = -\frac{1}{\phi}\Delta \log\left(\frac{N_{it}}{N_{jt}}\right). \tag{E1}$$

Intuitively, we can use variation in labor supply to trace out firms' labor demand schedule, which depends on ϕ. We obtain an estimate of ϕ by estimating (E1) 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.

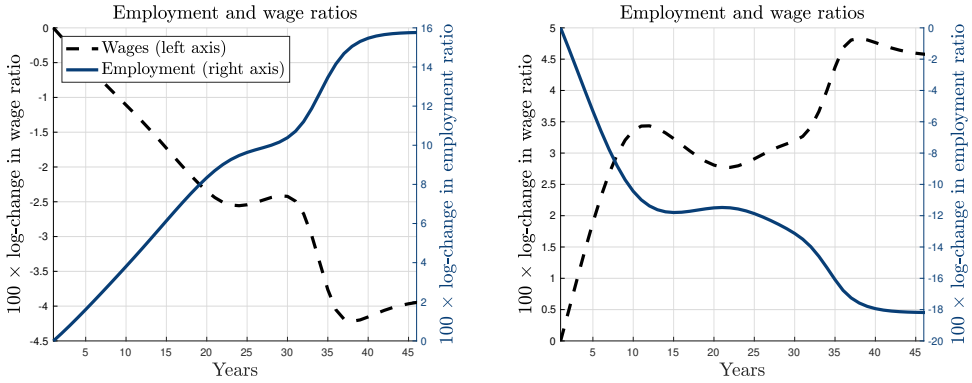
Results. Table E.1 illustrates the results of this exercise. The first column displays the true value of $\phi = 4$. The second column shows that our estimator leads to an estimate $\widehat{\phi} = 3.94$ of ϕ, very close to the true value. We conclude from this exercise that the estimates of the long-run elasticity of substitution in the literature provide appropriate discipline to pin down the value of ϕ in our model as well. The last column of the table shows that by estimating ϕ 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 ϕ.

TABLE E.1: Estimation Strategy for Long-Run Elasticity ϕ 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 μ_{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 (E1) 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 (E1) using a long difference between the new BGP and the initial BGP.

FIGURE E.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 μ_{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.

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 ϕ that we recover than in shaping the response of the economy to, say, changes in the minimum wage. Figure E.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 their employment ratios is approximately equal to $-\phi = -4$ times the log-change in their wage ratios by (E1). 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 $t = 0$.[33] In contrast, the policy experiments analyzed in the main text involve an immediate and unexpected jump in policy.

[33]Consistent with this predictability, the adjustment of employment ratios relative to wage ratios is farthest from ϕ in the early stages of the transition.

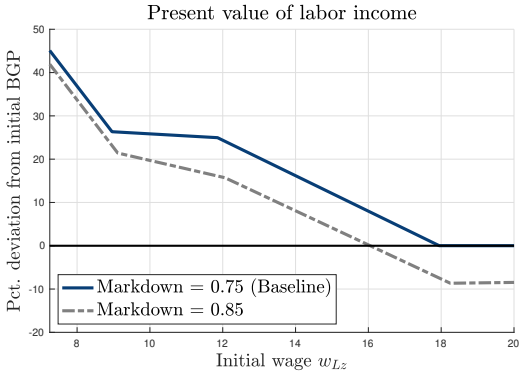
F Additional Quantitative Results

This Appendix contains additional quantitative results about the minimum wage.

F.1 Sensitivity of Labor Income Dynamics to Firm Monopsony Power

Figure F.1 shows how the present value of labor income varies with firm monopsony power. We decrease the degree of monopsony power by choosing ω to target a smaller markdown of 0.85 rather than 0.75 as in our baseline. This change barely affects the present value of the lowest-wage workers because both their short-run gains—their wages nearly double—and their long-run losses—substantial declines in their employment rates—are similar in both parameterizations. However, for workers who were initially earning close to the minimum wage policy, the benefits of the minimum wage are smaller because the monopsony distortion they face is smaller and fixing it is less valuable.

FIGURE F.1: Sensitivity of Labor Income Dynamics to Monopsony Power



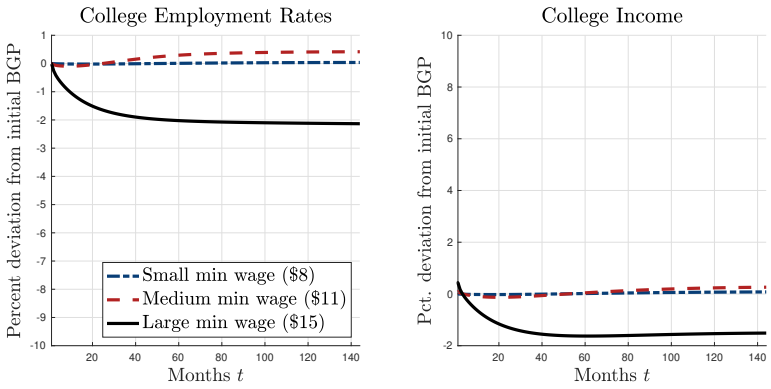
Notes: Change in the present value of labor income for all non-college workers in our baseline model and under an alternative parameterization in which we choose ω to match an average wage markdown of 0.85.

F.2 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-college workers. Appendix Figure F.2 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 F.2: Dynamic Effects of the Minimum Wage for College Workers

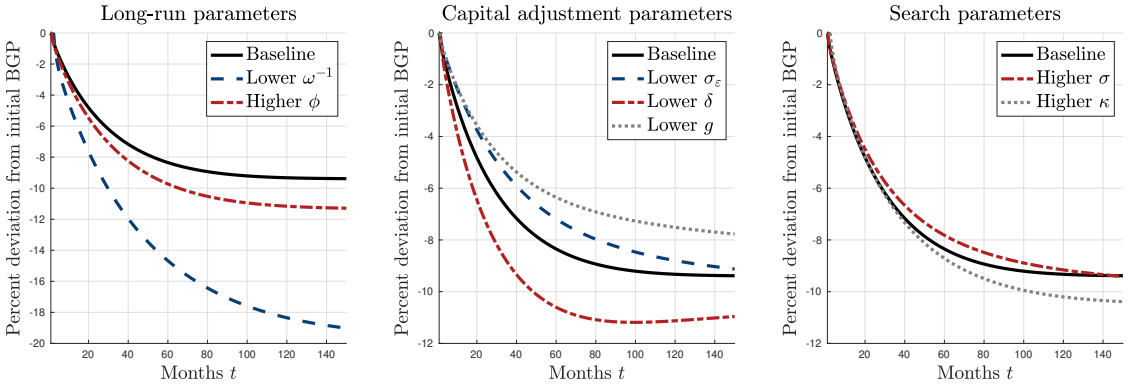


Notes: Transition paths of aggregate college employment (left panel) and labor income (right panel) following the introduction of the minimum wage. Employment is expressed in percentage deviation from the initial BGP. Labor income, net of trend growth $(1+g)^t$, is expressed relative to the initial BGP.

F.3 Additional Robustness Exercises

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 F.3 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 the various parameterizations. For example, the parameterization with less monopsony power (i.e. higher ω) leads to a larger long-run decline in non-college employment.

FIGURE F.3: Sensitivity Analysis to $15 Minimum Wage for Non-College Employment



Notes: Transition paths of non-college employment expressed as the percentage deviation from the original BGP. "Baseline" corresponds to the model shown in Figure 7. "Higher ω^{-1}" corresponds to a degree of monopsony power of $\omega^{-1} = 1/6$ that produces an 85% markdown. "Higher ϕ" corresponds to a long-run elasticity of substitution within education groups of $\phi = 4.5$. "Lower σ_ε" 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 δ" sets the depreciation rate to $\delta = 20\%$ annually. "Lower g" corresponds to a trend growth rate of $g = 1\%$. "Higher σ" sets the job-destruction rate to $\sigma = 3.5\%$ monthly. "Higher κ" increases the baseline vacancy-posting cost κ_0 by 2.5 times, which approximately doubles average hiring costs $\kappa_i/\lambda_f(\theta_i)$ to 125% of average monthly wage.