Monopsony with Recruiting*

Justin Bloesch
Cornell Economics and ILR

Birthe Larsen
Copenhagen Business School

Anders Yding
UC Berkeley

November 13, 2024

Abstract

We develop a tractable model of monopsony where firms use wages and recruiting expenditures to attract workers. The model predicts that firms' labor supply curves are elastic in the long run, consistent with evidence that firms pay higher wages while growing, but the wage premium at large firms is small. We confirm these predictions using the effect of export demand shocks on the wage growth of job switchers in Denmark. Our results imply that monopsony rents are dissipated by recruiting costs, which can reconcile existing estimates of monopsony power with the profit share of national income in rich countries.

*Justin Bloesch: Cornell Department of Economics and School of Industrial and Labor Relations (email: jb2722@cornell.edu); Birthe Larsen: Copenhagen Business School (email: bl.eco@cbs.dk); Anders Yding: University of California, Berkeley (email: ayding@berkeley.edu). We thank Anastasia Burya, Gabriel Chodorow-Reich, Kyle Herkenhoff, Lawrence Katz, Philipp Kircher, Patrick Kline, Alan Manning, Guido Menzio, Simon Mongey, Chris Moser, Suresh Naidu, Kris Nimark, Tommaso Porzio, Seth Sanders, Anna Stansbury, Evan Soltas, Jacob Weber, and seminar participants at the Copenhagen Business School, the University of California, Berkeley, the IAB workshop Imperfect Competition in the Labor Market, the NBER SI, the EALE conference, Bergen, the EEA conference, Rotterdam, the local-MPC workhop, Copenhagen, the BSE forum, the macroeconomics of the labor market, Barcelona, the SED conference, Barcelona, and Columbia Junior Micro/Macro Labor conference and for helpful comments. JEL codes J42, J63, E25.

1 Introduction

There is a growing consensus that firms have significant power to set wages: firms that choose lower wages do not lose all of their employees (Card, 2022). In many standard models of labor market monopsony, firm wage setting power implies a finite labor supply elasticity, allowing researchers to infer the markdowns of wages from marginal product, and by extension, the profits that firms earn by exploiting their wage setting power. However, questions remain in estimating the labor supply elasticity and its implications for the distribution of income. First, common estimates of labor supply elasticities between 2-6 imply that firms must pay significantly higher wages to become large. However, in cross-sectional data, large firms pay only slightly higher wages, and other studies find labor supply elasticities that are at least an order of magnitude higher.[1] Additionally, numerous recent studies show that growing firms pay a wage premium.[2] Second, there is a puzzle regarding profits: commonly used low labor supply elasticities imply implausibly large profit shares of income in the aggregate.[3] Further, many monopsony models abstract from recruiting activity and assume that firms' only way to expand is by offering higher wages. In this article, we therefore ask, *if firms can spend on recruiting, what labor supply elasticity do firms face in the long run, and how large are profits from wage setting power?*

To answer these questions, we derive a novel and tractable model of dynamic monopsony where workers search on the job, and firms use both higher wages and a recruiting expenditure margin to attract workers. Workers have time-varying, horizontally differentiated preferences over firms, and the presence of both search frictions and these preferences gives firms wage setting power: firms that choose lower wages do not immediately lose all their workers. Firms can recruit new workers by spending on recruiting activity, and the number of searching workers that a firm matches with each period is a function of both the amount of recruiting expenditure incurred by the firm and the number of incumbent workers in the firm. In the short run the firm faces diminishing returns from its recruiting expenditure, and so in response to a positive demand shock for its output, firms offer higher wages to increase the number of hires from a limited pool of workers matched to the firm. Consequently, firms pay high wages when growing. However, in the long run, a greater number of incumbents in the firm eases recruiting, lessening the incentive to pay high wages to generate more hires. If the firm's recruiting function has constant returns to scale, then a firm's optimal wage in steady state is the same regardless of the firm's size, resulting in a infinite long-run labor

[1] Sokolova and Sorensen (2021) provide a meta-analysis of two broad approaches to estimating labor supply elasticities that provide dramatically different estimates. Bloom et al. (2018) show that the firm size-wage premium had declined dramatically in the United States. Matsudaira (2014) finds that nursing homes subject to employment level mandates did not increase wages.

[2] See Schmieder (2023), Tanaka et al. (2023), and Carrillo-Tudela et al. (2023). Engbom et al. (2022) find that conditional on capital and productivity, large firms do not pay higher wages.

[3] Quoting Manning (2021): "The low estimated wage elasticity of the labor supply curve implies that employers have a lot of monopsony power: If this power is exercised it is not clear how it can be reconciled with observed levels of the profit share."

supply elasticity. We then show that in steady state, the *recruiting cost-adjusted* wage markdown is tightly related to this long-run labor supply elasticity: if firms' long-run labor supply curve is elastic, then the recruiting cost-adjusted wage markdown is equal to one in steady state, and the gap between wages and marginal product is entirely consumed by recruiting expenditure. In total, if firms have a constant returns to scale recruiting function, then there are still monopsony rents, but these rents are dissipated by the costs incurred to acquire workers in the first place.[4]

Next, we estimate the parameter governing the returns to scale of the recruiting cost function, and by extension the long-run labor supply elasticity and recruiting cost-adjusted markdown, using the effect of idiosyncratic demand shocks on the path of firm size, employment growth, and wages. In particular, we look at the wage growth of job switchers around the time of firm export demand shocks using Danish administrative merged employer-employee and trade data. We focus on the wage growth of switchers, rather than wage response of stayers, as the response of stayer wages to firm demand shocks may reflect rent sharing motivations that are unrelated to the firm's long-run labor supply elasticity (Kline et al. (2019); Garin and Silverio (2023); Carvalho et al. (2023)). We find that a firm's employment growth and the wage growth of the firm's newly hired workers respond positively on impact to export demand shocks. This is consistent with diminishing returns to recruiting effort, as firms resort to offering high wages to increase hiring among a limited pool of potential workers. We also show that the effect of trade shocks on a firm's employment level is persistent: firms are larger for years following the export demand shock. Crucially, however, workers who switch into the shocked firm in the years following the trade shock, when the firm is larger but is no longer growing, do not see any wage premium relative to workers who were hired prior to the shock. This is consistent with a labor supply curve that is elastic in the long run: the firm has increased in size, but the firm pays similar wages as when the firm was smaller. Together, these results are consistent with diminishing returns to recruiting expenditure in the short run, but an elastic labor supply curve in the long run.

We then embed our monopsonistic firms with recruiting into an equilibrium setting with capital, imperfect product market competition, and labor market concentration. We solve for the labor share, profit share, and capital shares of aggregate income in closed form. We then demonstrate the profit puzzle in this equilibrium setting: in a model with standard labor supply elasticities but no recruiting margin, and otherwise standard values for price markups and the elasticity of output with respect to capital, the profit share of income is 10-15 percentage points too high, and the labor share of income is 10-15 percentage points too low, relative to national accounts data in rich countries. We show that the puzzle can be reconciled if instead firms have access to a recruiting margin: if firms have a constant returns to scale recruiting function and an elastic long-run labor supply curve, then our model can match aggregate profit and labor shares while maintaining realistic price markups over marginal cost.

As a part of this equilibrium analysis, we additionally consider nonatomistic firms and strategic

[4]This result that a constant returns to scale recruiting function make the firm's labor supply curve elastic is a familiar result from Manning (2003), Kuhn (2004), Manning (2006).

interactions in the presence of a recruiting margin. Firms compete according to an extended version of Bertrand competition, where firms solve their problems taking as given their competitors' wage and vacancy choices. We show that if labor market concentration rises such that the labor market Herfindahl-Hirschman Index (HHI) increases from 0 to realistic levels around 0.1 (Berger et al., 2022), then wages fall relative to marginal product by only 2.5 percent economy-wide, raising the aggregate profit share of income by 2 percentage points. Therefore our model of monopsony with recruiting can still match aggregate profit and labor shares, even when taking into account both price markups and strategic interactions from realistic levels of labor market concentration.

On the whole, our results point to a labor market with a substantial amount of rents: our evidence is consistent with wages that are around 8% below marginal product and a marginal hiring cost of a worker between 4-6 months of wages. However, our results suggest that competition by firms to access this flow of rents, in the form of spending on recruiting, limits the extent that these labor market rents result in profits for firms in the aggregate.

Our novel framework unifies key aspects of dynamic monopsony and on-the-job search models. In the dynamic monopsony literature, researchers use credible identification strategies to estimate the elasticity of recruits and the elasticity of separations with respect to a firm's wage policy. In our setting, the combination of search frictions and workers' heterogeneous preferences creates finite elasticities of recruiting and separations with respect to wages, giving firms wage setting power. To our knowledge, our paper is the first to explicitly integrate these recruiting and separation elasticities into a microfounded, equilibrium search framework. The result is a highly tractable model of on-the-job search that delivers many insights in closed form and can be easily disciplined by existing empirical evidence.

Lastly, we make an additional theoretical contribution: our environment has random on-the-job search and wage posting, but there is a point mass of wages in equilibrium when firms and workers are ex-ante homogenous. This stands in contrast to the famous result in Burdett and Mortensen (1998) that with on-the-job search and ex-ante homogenous workers and firms, the equilibrium must contain wage dispersion. Like Albrecht et al. (2018), we achieve this by workers having horizontally differentiated preferences over workplaces, but unlike much of the existing monopsony literature, we make these differentiated preferences time-varying.[5] The result is that identical firms choose identical wage policies, and job-to-job mobility occurs even if there is a point mass of wages. This equilibrium without wage dispersion permits a highly tractable starting point that can easily be enriched and may serve as a useful benchmark for future research relating to matching markets where currently matched agents continue to search.

The paper is organized as follows. Section 2 lays out the model of dynamic monopsony in partial equilibrium examining the firm's problem alone, and then provides a microfoundation for firms' wage setting power and characterizes the labor market equilibrium with atomistic firms. Log-linearizing the firm's problem, we show that a firm's wage policy is a function of its employment

[5] Menzio (2024) derives general conditions for when point masses of prices arise in the context of search frictions and diminishing returns in production.

level and employment growth rate. Section 3 presents the empirical evidence on the wage growth of switchers around firm export demand shocks in Denmark. Section 4 derives the quantitative model with labor market concentration and addresses the question about the aggregate profit and labor shares. Section 5 concludes.

Related Literature This paper is most closely related to the canonical models of dynamic monopsony, nesting and providing a microfoundation for the model of dynamic monopsony in Manning (2003) and providing a microfounded case of Manning (2006)'s generalized model of monopsony that explicitly integrates worker turnover and on-the-job search. Numerous authors have estimated recruiting and separation elasticities including Azar et al. (2019), Bassier et al. (2022), and Datta (2023). Our study also contributes to a growing literature that emphasizes the importance of concentration in the labor market, including Azar et al. (2020), Berger et al. (2022), Derenoncourt and Weil (2024), Jarosch et al. (2021), Naidu et al. (2018), and Schubert et al. (2023), among others.

The importance of non-wage amenities, which is essential in our model, is highlighted in Sorkin (2018) and Hall and Mueller (2018). Heise and Porzio (2022) also have related time-varying horizontal preferences over workplaces. Along with Berger et al. (2024), we are among the first to jointly consider labor market concentration, search frictions, and idiosyncratic workplace preferences, which are three of the most studied mechanisms for giving firms wage setting power, and we are the first to consider these jointly in a setting of wage posting.

Our study relates to the literature of firm size wage premia, such as Brown and Medoff (1989) and Oi and Idson (1999), and more recently Cobb and Lin (2017), Bloom et al. (2018), and Song et al. (2019). Numerous authors estimate labor supply elasticities using the response of stayer wages and firm size to idiosyncratic shocks to firms, such as Lamadon et al. (2022), Trottner (2022), Seegmiller (2021), and Chan et al. (2020).

2 Monopsony with Recruiting: Atomistic Firms

In this section, we derive our core model of dynamic monopsony. In Section 2.1, we begin with the problem of an atomistic firm, where a firm chooses its wage policy and how much to spend on recruiting, taking as given that a higher wage accelerates hiring and decreases worker turnover. We derive the recruiting cost-adjusted wage markdown, and we show that this markdown is closely related to the firm's long-run labor supply elasticity, which in turn is related to the returns to scale in the firm's recruiting function. In Section 2.2, we microfound the workers' problem, specifying the frictional labor market and workers' preferences. This allows us to fully characterize the equilibrium labor market and draw out additional insights analytically. Lastly, we log linearize the firm's problem and show that in response to an idiosyncratic firm demand shock, the firm's optimal wage is a function of the firm's employment growth rate and the employment level, motivating the empirical analysis later in Section 3.

2.1 Firm's Problem with Recruiting

In this subsection, we lay out the firm's problem, where firms choose the path of wages and recruiting expenditure (and implicitly their employment level and output), taking as given that a higher wage increases the rate of hiring and decreases worker turnover. We assume that firms produce with diminishing returns to labor, that firms can match with more workers by spending more on recruiting, and that higher wages can lower recruiting costs while attaining the same level of hires. We decompose the share of marginal product that is spent on wages, spent on recruiting, and retained by the firm as profits. We additionally derive the recruiting-cost adjusted wage markdown and the firm's inverse labor supply elasticity. We solve the firm's problem first without microfounding the worker's problem to demonstrate the generality of our results, as we show that from the firm's problem alone we can nest numerous monopsony models as special cases where the recruiting margin is shut down.

Production Firms produce using labor N with a diminishing returns to scale technology $F(N)$, where $F'(N) > 0$ and $F''(N) < 0$. For Section 2, we will parameterize the firm's value added as $F(N_t) = AN_t^{\alpha}$, where A is a demand shifter and $\alpha \in (0, 1)$.

Recruiting Cost Function Every period, firms begin the period with N_{t-1} incumbent workers and can choose a level of recruiting expenditures E_t. A firm's amount of effective recruiting activity V_t is

$$V_t = c_0 E_t^{\frac{1}{1+\chi}} N_{t-1}^{\frac{\chi(1-\sigma)}{1+\chi}}, \tag{1}$$

with $\chi \geq 0$ and $\sigma \in [0, 1]$. χ determines the amount of diminishing returns to recruiting expenditure given the number of incumbents N_{t-1}. σ governs both how important incumbents are in recruiting[6] and whether or not the firm's recruiting function is constant returns to scale: if $\sigma = 0$, then the recruiting function has constant returns to scale. Another notable case is $\sigma = 1$, where the number of incumbents has no effect on recruiting.[7]

This recruiting function is reduced form in the sense that effective recruiting activity V_t captures all possible activities that a firm can engage in when trying to match with more workers. These activities could include instructing current employees to share with their social network that the firm is hiring, recalling former employees, putting help-wanted signs in stores, paying for online job postings, attending career fairs, etc. The assumption behind our recruiting function is that firms rationally choose the least cost per unit of effective recruiting first, and firms choose more expensive forms of recruiting when the cheaper forms of recruiting are exhausted.

[6]There is a large literature that documents that many hires are done through referrals: see Burks et al. (2015), Caldwell and Harmon (2019), Jahn and Neugart (2020) and Dustmann et al. (2016).

[7]σ in theory could be negative, making the recruiting function an increasing returns to scale function. For atomistic firms, this will mean that optimal wages are decreasing in firm size. For certain parameter values later on, this could result in firms having no interior solution for optimal firm size. To avoid this, we restrict our attention to $\sigma \in [0, 1]$. For details, see equations (34) and (35) in Appendix A.6

We intentionally denote effective recruiting activity V_t, since for the rest of this paper, we will refer to V_t as "vacancies" to conform to notational convention in the search literature.[8] With this relabeling, the recruiting function (1) can be reframed as a convex vacancy posting cost such that total recruiting expenditure is

$$E_t = c\left(\frac{V_t}{N_{t-1}^{1-\sigma}}\right)^{\chi} V_t, \tag{2}$$

where $c = c_0^{-(1+\chi)}$. In this expression, χ determines the convexity of recruiting costs with respect to the number of vacancies, and σ determines the extent that additional incumbents N_{t-1} offsets these convex recruiting costs. In the case of a constant returns to scale recruiting function (i.e., $\sigma = 0$), then per-vacancy costs are increasing in the ratio of vacancies to incumbent workers V_t/N_{t-1} if $\chi > 0$.

Worker Flows Firms choose wages period by period, and firms are constrained to pay the same wage to all workers, both incumbents and new hires. Firms take as given that paying a higher wage increases recruiting and decreases worker turnover. Let $R(w_t) \in [0,1]$ be the share of vacancies that produce a hire, with $R'(w_t) > 0$, and let $S(w_t) \in [0,1]$ be the share of incumbent workers each period who separate from the firm, with $S'(w_t) < 0$.[9] The firm's law of motion for employment is

$$N_t = (1 - S(w_t))N_{t-1} + V_t R(w_t), \tag{3}$$

where the measure of workers N_t the firm employs in period t is the measure of incumbents retained from last period $(1 - S(w_t))N_{t-1}$ plus the measure of new hires $V_t R(w_t)$.

Firm's Problem Firms maximize the present discounted value of revenue less wage costs and recruiting costs. The firm solves

$$\max_{\{N_t\},\{V_t\},\{w_t\}} \sum_{t=0}^{\infty} \left(\frac{1}{1+\rho}\right)^t \left(AN_t^{\alpha} - w_t N_t - c\left(\frac{V_t}{N_{t-1}^{1-\sigma}}\right)^{\chi} V_t\right), \tag{4}$$

subject to (3). In steady state, the firm's optimal wage satisfies

$$w = c(1+\chi)\left(\frac{S(w)}{R(w)}\right)^{1+\chi}\left(\varepsilon_{R,w} - \varepsilon_{S,w}\right)N^{\sigma\chi}, \tag{5}$$

[8]A now extensive literature has documented the importance of accounting for recruiting intensity, beginning with Davis et al. (2013) and more recently by Gavazza et al. (2018) and Mongey and Violante (2023). These papers directly map vacancies in a model to measured job openings in the data, finding that an additional recruiting margin is necessary to rationalize increasing vacancies yields at firms that are growing. We sidestep this distinction, as our measure of "vacancies" is a reduced form composite of all possible recruiting activities.

[9]The lack of a t subscript on $R()$ and $S()$ denotes our assumption that firms are operating in a stationary environment. Additionally we assume that the wage policy in period t affects recruiting and separation probabilities only in period t, which will be justified by our microfoundation in the next section.

where $\varepsilon_{R,w}$ is the elasticity of the recruiting rate with respect to the wage, and $\varepsilon_{S,w}$ is the elasticity of the separation rate with respect to the wage. Note that the separation elasticity $\varepsilon_{S,w}$ is negative: $\varepsilon_{S,w} \equiv \frac{wS'(w)}{S(w)}$, so with $S'(w) < 0$, the separation elasticity $\varepsilon_{S,w}$ is negative as well.[10] Define $\mathcal{E}(w) \equiv (\varepsilon_{R,w} - \varepsilon_{S,w})$ as the recruiting elasticity minus the separation elasticity. The steady state ratio of vacancies to employment takes the following form:

$$\frac{V}{N} = \frac{S(w)}{R(w)}. \tag{6}$$

The two above equations demonstrate that the firm trades off between wage costs and turnover costs: a higher wage means that the firm will post fewer vacancies in steady state, decreasing recruiting costs. Conversely, a low wage would mean high steady state turnover costs. Additionally, if $\sigma = 0$, then the optimal wage w is independent of firm size N. In this case, the trade-off between turnover costs and wages is the same in steady state regardless of firm size. If the trade-off between wage and turnover costs is the same regardless of size, then the cost-minimizing choice of wages and vacancies per worker will be the same for firms of different sizes.

In the limit where firms do not discount[11] ($\rho \to 0$), the optimal steady-state level of employment is

$$N = \left(\frac{\alpha A}{w} \frac{\mathcal{E}(w)(1+\chi)}{1+\mathcal{E}(w)(1+\chi)+\sigma\chi} \right)^{\frac{1}{1-\alpha}}. \tag{7}$$

Recall the definition of the sum of recruiting and (negative) separation elasticities as:

$$\mathcal{E}(w) \equiv \varepsilon_{R,w} - \varepsilon_{S,w}.$$

For presentation purposes, we will suppress the (w), and simply write this sum of elasticities as $\mathcal{E}$.

We now decompose the share of marginal product that goes to wages, recruiting costs, and profits, as well as derive expressions for the recruiting cost-adjusted wage markdown and the inverse labor supply elasticity in Table 1. Each of these expressions now depend on the sum of recruiting and separation elasticities $\mathcal{E}$ and the recruiting cost parameters χ and σ. The first three terms simply represent a decomposition: what share of marginal product is spent on wages, on recruiting per worker, and what is left over. We term the share of marginal product that is left over after accounting for wage and recruiting costs "labor market profits." The next term is the recruiting cost-adjusted wage markdown, which expresses what share of marginal product is spent by the

[10] We show in Appendix B that if firms post fixed wage contracts for forward-looking workers, the optimal wage expression is the same as when firms set wages period-by-period for both incumbents and new hires. This is because, when firms are setting wages for the future, current new hires are future incumbents, so the optimal wage similarly takes into account both recruiting and retention.

[11] We show in the Appendix A.1 that when this model is calibrated to standard monthly parameters, the discount rate is quantitatively unimportant. Therefore, for the rest of the paper, we will assume $\rho = 0$.

firm on either wages or recruiting.[12] The last term in Table 1 is the inverse labor supply elasticity, which shows the elasticity of optimal wages as the firm's optimal employment level changes, under the assumption that the firm's hiring cost parameters c, χ, and σ are unchanged. This reflects how much wages would change relative to employment in response to a demand shock.

Table 1: Decomposing Marginal Product into Wages, Recruiting Costs, and Profits

Outcome	Formula
$\frac{w^*}{MRPL^*}$	$\frac{(1+\chi)\mathcal{E}}{1+(1+\chi)\mathcal{E}+\sigma\chi}$
$\frac{\text{Recruiting Costs/Worker}^*}{MRPL^*}$	$\frac{1}{1+(1+\chi)\mathcal{E}+\sigma\chi}$
Labor Market Profit Share of MRPL $\equiv \frac{1-w^*-\text{Recruiting Costs/Worker}^*}{MRPL^*}$	$\frac{\sigma\chi}{1+(1+\chi)\mathcal{E}+\sigma\chi}$
Recruiting Cost-Adjusted Markdown $\equiv \frac{w^*+\text{Recruiting Costs/Worker}^*}{MRPL^*}$	$\frac{1+(1+\chi)\mathcal{E}}{1+(1+\chi)\mathcal{E}+\sigma\chi}$
Inverse Labor Supply Elasticity $\equiv \varepsilon_{w^*,N^*}$	$\frac{\sigma\chi}{1+(1+\chi)\mathcal{E}-\varepsilon_{\mathcal{E},w}}$

This table shows the wages, recruiting cost, and labor market profit shares of marginal product, the recruiting cost-adjusted markdown, and the inverse labor supply elasticity. These results hold in partial equilibrium and are only a function of the firm's problem. Analogous results for the general ρ case can be found in Appendix A.1. Recruiting costs as a share of marginal product is calculated as $\big(c\big((V/N^{1-\sigma})^{\chi}V/N\big)/F'(N)$.

Two parameterizations are of particular interest. The first is the limit case where the recruiting

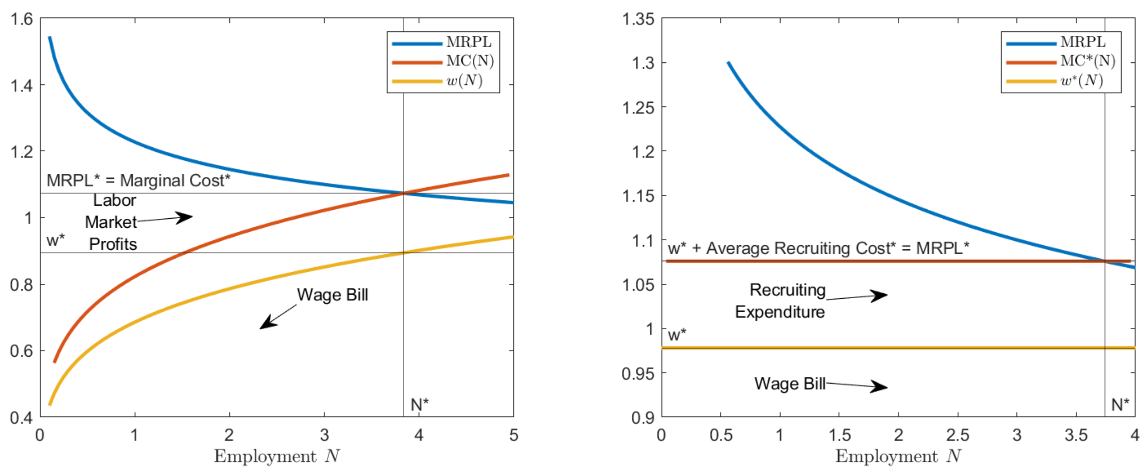
[12] There is a question of how to account for recruiting costs: they could be interpreted either an expenditure or an output cost in our setting. As our base case, we will assume that recruiting costs are an expenditure that accrues to labor, which is why the recruiting cost adjusted markdown is $(w^*+\text{recruiting costs/worker}^*)/MRPL^*$. An alternative is that recruiting costs are subtracted from output, in which case the recruiting cost adjusted wage markdown would be $w^*/(MRPL-\text{recruiting costs/worker}^*)$. A third alternative is that recruiting effort may divert labor away from productive activity, such that labor inputs are $N-cV^{1+\chi}N^{-\chi(1-\sigma)}$. In Section 4, it will be more convenient to count recruiting cost as an expenditure that accrues to labor, but assuming that recruiting costs reflect output losses or fewer labor inputs, or some combination thereof, has almost no effect on our results.

margin is shut down as a choice, where $\sigma = 1$ and $\chi \to \infty$. Plugging in these parameter values for χ and σ into Table 1, we can see that we generate the standard dynamic monopsony model without a recruiting margin, where the labor supply elasticity $\varepsilon_{N,w} = \mathcal{E} = \varepsilon_{Rw} - \varepsilon_{Sw}$, the firm spends nothing on recruiting, and the wage markdown is $\frac{\mathcal{E}}{1+\mathcal{E}}$. Thus, our model with a recruiting margin nests standard models of dynamic monopsony without a recruiting margin as a special case. Though our model is dynamic, in steady state it is isomorphic to static models of monopsony without a recruiting margin.

The second parameterization of interest is when firms have a constant returns to scale recruiting function, i.e., $\sigma = 0$. In this case, firms have an elastic labor supply curve, as $\sigma = 0$ implies an inverse labor supply elasticity of zero. Additionally, a constant returns to scale recruiting function with $\sigma = 0$ implies that the recruiting cost adjusted wage markdown is equal to 1: even though wages are below marginal product, the gap between wages and marginal product is entirely spent on recruiting costs.

Figure 1: Wage Bill, Recruiting Cost, and Marginal Product

(a) Standard Monopsony: $\sigma = 1$, $\chi \to \infty$ (b) Constant Returns to Scale Recruiting: $\sigma = 0$



This figure shows the marginal product, marginal cost, and labor supply curves for firms with different recruiting functions. Panel (a) shows a firm in a standard monopsony model with no recruiting function, equivalent to $\sigma = 1$ and $\chi \to \infty$ in our model. The firm's marginal cost line lies of above the firm's labor supply curve, and the wage lies below marginal product when the marginal product curve and marginal cost curve intersect. Panel (b) shows a firm with a constant returns to scale recruiting function. For this firm, the optimal wage is the same regardles of firm size, and the share of recruiting costs in total costs is also constant across different firm sizes. The wage lies below marginal product at the optimum, but the gap between wages and marginal product is entirely consumed by recruiting costs.

Figure 1 shows visually shows the results in Table 1 for these two special cases: the recruiting function shut down ($\sigma = 1$ and $\chi \to \infty$) and a constant returns to scale recruiting function $\sigma = 0$.

Figure 1a plots the labor supply curve, the marginal cost curve, and the marginal product curve for the case where the recruiting margin is shut down. This figure is the same figure as in standard monopsony models where the labor supply curve is upward sloping, the firm's marginal cost curve lies above the labor supply curve, and the marginal product and marginal cost curves intersect at a point where wages lie below marginal product.

Figure 1b shows the same graph but for a firm with a constant returns to scale recruiting function. For any steady state level of employment, the trade-off between wage costs and turnover costs is the same. Therefore, the marginal cost of employment (across steady states) is constant. At the point where the marginal cost of employment crosses the marginal product curve, wages are below marginal product, but the entire gap between wages and marginal product is spent on recruiting. The fact that wages are below marginal product demonstrates that, once firms and workers are matched, the firm gets a flow of rents from employing the worker. However, just as in a free entry condition in labor search models, firms pay recruiting costs up front in order to get access to the flow of rents that comes from employing the worker.

Evidence on Recruiting Costs Later in the paper, we will present indirect evidence that firms have constant returns to scale recruiting functions, with the implication that the gap between wages and marginal product is consumed by recruiting costs, and also that firms face diminishing returns to recruiting expenditure. This exercise leads to natural questions. Are recruiting costs large enough to account for the difference between wages and marginal product, and is there evidence for these diminishing returns to recruiting expenditure?

The first part of the answer is that in a model with recruiting, the ratio of wages to marginal product can be much closer to 1 than would be implied by a model without a recruiting margin. As a consequence, the level of recruiting costs does not need to be large in order for the recruiting cost-adjusted markdown to be equal to 1. To see this, consider the evidence that the sum of recruiting and separation elasticities $\mathcal{E} \equiv \varepsilon_{R,w} - \varepsilon_{S,w}$ that is typically estimated to be around 4 (Bassier et al. (2022), Datta (2023), Sokolova and Sorensen (2021)). In a model without a recruiting margin ($\sigma = 1$, $\chi \to \infty$), this would imply a wage share of marginal product equal to $\frac{\mathcal{E}}{1+\mathcal{E}} = .8$. Consider instead a different benchmark of a constant returns to scale recruiting function ($\sigma = 0$) but firms face convexity in recruiting costs as a function of vacancies per incumbent, with $\chi = 2$. In that case, holding fixed the recruiting minus separation elasticities $\mathcal{E} \equiv \varepsilon_{R,w} - \varepsilon_{S,w} = 4$, the ratio of the wage to marginal product becomes $\frac{\mathcal{E}(1+\chi)}{1+\mathcal{E}(1+\chi)+0} = .92$. Thus, a higher value of χ pushes wages closer to marginal product, holding $\varepsilon_{R,w} - \varepsilon_{S,w}$ fixed.

Given the importance of χ in determining how close wages are to marginal product, do we have any evidence on χ? Looking first at models calibrated to match firm-level evidence, Kaas and Kircher (2015) find that in a directed search model, a value of 8 for the elasticity of recruiting costs with respect to vacancies is needed to match the relationship between firm vacancy rates and employment growth rates in Davis et al. (2013). In a model with two recruiting margins (vacancies and effort), Mongey and Violante (2023), find that the convexity on the vacancy cost is about three times higher than the convexity on effort. Gavazza et al. (2018) estimate analogous parameters to be 4.6 and 1.1, respectively. Since our model has only one recruiting margin, this would imply a value of χ that is at least 1.

When interpreting evidence on turnover or hiring costs, it is important to distinguish between average and marginal hiring costs. In our setting, marginal hiring costs = $(1 + \chi) \times$ average hiring costs. Evidence on average hiring costs tends to show that average hiring costs are moderate in size: Manning (2011) surveys evidence and Dube et al. (2010) point to average hiring costs around 2-5% of the wage bill. Blatter et al. (2012) finds the average cost of replacing skilled workers to be 10-17 weeks of salary. Michaillat and Saez (2024) argue that servicing a vacancy requires approximately one full time worker. In terms of marginal hiring costs, Muehlemann and Pfeifer (2016) find that marginal hiring costs are increasing, consistent with $\chi = 1.3$. Bloesch and Weber (2022) find evidence of congestion of onboarding software developers, consistent with convex costs in the rate of training new workers. Evidence on separation costs, and by implication the marginal cost of a hire, tend to be larger: Kline et al. (2019), Jäger and Heining (2022), Isen (2013), Bertheau et al. (2021), and Bloesch et al. (2022) all point to separation costs to be between 1-3 years of workers' wages. However, the latter four papers estimate these costs from worker deaths, which may be significantly more costly than replacing a worker after a typical quit.

Our model has strict implications for the present value of a match to a firm and the level of marginal hiring costs in steady state. Given the expressions for wages and recruiting costs in Table 1, the recruiting expenditure per incumbent each period in steady state is $\frac{w}{\mathcal{E}(1+\chi)}$. Since in steady state the number of hires is equal to the number of separations SN, the recruiting expenditure per hire is $\frac{w}{S\mathcal{E}(1+\chi)}$. Given the convexity parameter χ, the marginal hiring cost is $1 + \chi$ times average cost per hire, so the marginal hiring cost is $\frac{w}{S\mathcal{E}}$. Therefore, with $\mathcal{E} = 4$ and a monthly separation rate between .04 and .06,[13] this would imply a marginal hiring cost per worker of around 4-6 months of wages. To be consistent with evidence on the level of hiring costs, this would imply a value χ roughly between 1 and 3.

All in all, the existing evidence points to convexity in recruiting costs with respect to the recruiting rate or hiring rate, consistent with $\chi > 0$. For the rest of this paper, we will assume of a value of $\chi = 2$, and we will not use our empirical evidence in Section 3 to further make inference on χ. However, we will estimate σ in Section 3 using the wage growth of job switchers before and after idiosyncratic demand shocks. Importantly, only the value of σ, but not the value of χ, will matter for our quantitative analysis in Section 4 of the equilibrium distribution of income across labor, capital, and profits.

[13] The rate of worker flows across firms is similar Denmark and the US Caldwell and Harmon (2019). Bagger et al. (2022) find that the monthly separation rate in Denmark is around 6.6 percent. The total monthly job separation rate from the Job Opening and Labor Turnover Survey (JOLTS) in the US averages between 3 and 4 percent.

2.2 Microfounding Worker Mobility Decisions and Labor Market Equilibrium

In this section, we specify a structure for worker preferences and frictional labor markets. This will be used to derive results for the firm's problem as a function of primitives, including analytical expressions for the optimal wage, and it will allow us to characterize the labor market in equilibrium in closed form.We will also derive an explicit formula for the firm size-wage premium in terms of primitives and analytically link the inverse labor supply elasticity to the share of marginal product that firms retain as profits.

Population and Frictional Markets Workers $i \in [0, 1]$ and firms $j \in [0, 1]$ both have a unit mass. Firms can post vacancies to recruit workers, and workers can search on the job. The measure of aggregate vacancies is $\bar{V} = \int_{j\in J} V_j dj$, and the measure of searchers $\mathcal{S}$ will be all the workers who are enabled to search in period t. A worker can be matched with at most one vacancy per period, and vice-versa. Matching occurs according to a constant returns to scale matching function $M(\bar{V}, \mathcal{S})$. Define the labor market tightness as $\theta = \bar{V}/\mathcal{S}$. Conditional on searching, the probability that a worker encounters a job is $f(\theta) = M(\bar{V}, \mathcal{S})/\mathcal{S}$, and the probability that a firm's vacancy encounters a searching worker is $g(\theta) = M(\bar{V}, \mathcal{S})/\bar{V}$. Workers are exogenously separated with probability s, and exogenously separated workers cannot search this period. The remaining $1 - s$ share of workers can search on the job with probability λ_{EE}, and workers are able to consider quitting into unemployment each period with probability λ_{EU}, with $\lambda_{EE} + \lambda_{EU} \leq 1$. The unemployment rate is denoted U, and unemployed workers can search every period. The mass of searchers each period is $\mathcal{S} = U_{t-1} + (1 - s)\lambda_{EE}(1 - U_{t-1})$. The unemployment benefit is b.

The timing is as follows: each period firms inherit the number of incumbents N_{t-1} from last period. Firms decide their wage policy for this period and the number of vacancies they wish to post, and the firms incur recruiting costs. Workers are exogenously separated with probability s. Search and matching then occurs, and workers make their mobility decisions, both across firms and across employment and unemployment states. At the end of the period, firms produce and pay wages.

Workers Workers get indirect utility from wage income w and have time varying, idiosyncratic preferences over workplaces ι

$$U_{ijt} = \log(w_{jt}) + \iota_{ijt}.$$

Workers care only about the current period and have no ability to save.[14] The worker's non-wage utility from working at firm j in period t is ι_{ijt}. This preference ι_{ijt} is drawn i.i.d. each period and is distributed type-1 extreme value with scale parameter $1/\gamma$, with $\gamma \in (0, \infty)$. Workers can search for other jobs with probability λ_{EE}. Workers have the same i.i.d. preference draws over the state of being unemployed, and workers can consider quitting into unemployment probability λ_{EU}.

If a worker searches on the job and encounters a vacancy, the worker's choice of whether to stay in the firm or to switch becomes a simple discrete choice problem. For a worker employed at firm j that meets a vacancy of firm k, the probability of leaving firm j for firm k is

$$s_{jk}(w_j, w_k) = \frac{w_k^\gamma}{w_j^\gamma + w_k^\gamma}. \tag{8}$$

Workers who can consider quitting into unemployment separate with probability $b^\gamma/(w_j^\gamma + b^\gamma)$.

Analogously the probability that firm j poaches worker i who currently works at firm k, conditional on the worker matching with firm j's vacancy, is

$$r_{kj}(w_j, w_k) = \frac{w_j^\gamma}{w_j^\gamma + w_k^\gamma}, \tag{9}$$

and the probability that an unemployed worker accepts a job at firm j, conditional on the wage w_j, is $w_j^\gamma/(w_j^\gamma + b^\gamma)$.

Let the cumulative distribution of wage policies wages posted in vacancies be $\Upsilon(w)$, with corresponding density $\upsilon(w)$. Let the cumulative distribution of wages that workers are currently employed at be denoted $\Phi(w)$, with density $\phi(w)$. Let the share of searchers who are employed be Φ_E, and the share of searchers who are unemployed is $\Phi_U = 1 - \Phi_E$. Then the separation and recruiting functions are

$$S(w_j) = s + (1-s)\left(\lambda_{EE} f(\theta) \int_{w_k} \upsilon(w_k) \frac{w_k^\gamma}{w_j^\gamma + w_k^\gamma} dw_k + \lambda_{EU} \frac{b^\gamma}{w_j^\gamma + b^\gamma}\right) \tag{10}$$

$$R(w_j) = g(\theta)\left(\Phi_E \int_{w_k} \phi(w_k) \frac{w_j^\gamma}{w_j^\gamma + w_k^\gamma} dw_k + \Phi_U \frac{w_j^\gamma}{w_j^\gamma + b^\gamma}\right). \tag{11}$$

Now that we specified the workers problem and the frictional labor market, we have provided a microfoundation for the firm's recruiting and separation functions $R(w)$ and $S(w)$. Firms solve their

[14] In Appendix B, we generalize the model where workers are forward looking and firms post fixed-wage contracts for different cohorts of workers. We show that the expression for firms' optimal steady-state wage policy in equation (5), when firms are constrained to pay new hires and incumbents the same wage, is identical to the expression for optimal wages in steady state in the forward-looking case with fixed wage contracts. This is because in the forward looking case, new hires eventually become incumbents, and the optimal wage in both setting takes into account recruiting of new hires and retaining incumbents. To save on notation, we focus on the myopic case in the main body of the paper. Additionally, we abstract from optimal contracts, such as in Stevens (2004) and Balke and Lamadon (2022), to maintain our closed form results.

maximization problem, maximizing objective function (4) while plugging in the above recruiting and separation functions into the law of motion for employment (3).

Firms in theory may differ in their recruiting cost parameters χ_j, σ_j, and c_j as well as their demand shifter A_j and elasticity of revenue with respect to labor α_j. Next we will define an equilibrium and characterize key results about this equilibrium.

Symmetric Equilibrium A steady-state equilibrium is an employment policy $N^*(A_j, \alpha_j, c_j, \chi_j, \sigma_j, \tilde{w}, \theta)$, wage policy $w^*(A_j, \alpha_j, c_j, \chi_j, \sigma_j, \tilde{w}, \theta)$, and vacancy policy $V^*(A_j, \alpha_j, c_j, \chi_j, \sigma_j, \tilde{w}, \theta)$, a distribution of wages over employed workers $\Phi(w)$ and vacancies $\Upsilon(w)$, a mass of searchers $\mathcal{S}$, and labor market tightness θ such that (i) firms maximize profits, (ii) workers maximize utility, and (iii) flows of workers into and out of each firm balance each period.

Proposition 1 *If a steady-state equilibrium exists, and if there are no exogenous separations ($s = 0$), then at such an equilibrium*

1. *the recruiting elasticity minus the separation elasticity $\mathcal{E}(w_j) \equiv \varepsilon_{R(w_j),w_j} - \varepsilon_{S(w_j),w_j} = \gamma$ for any choice of wage level w_j, and*

2. *identical firms choose identical wage, employment, and vacancy policies.*

Proof: see Appendix A.4 for $\varepsilon_{R(w_j),w_j} - \varepsilon_{S(w_j),w_j} = \gamma$. We characterize the equilibrium in Appendix A.5, showing that we can solve for firms' optimal wage and employment level in closed form. Additionally, we prove in Appendix A.6 that if an equilibrium exists, and if all firms are have identical values for α, χ, and σ (but may differ in A_j and c_j), then the equilibrium is unique.

Corollary 2 *If firms are identical ex-ante identical, there are no exogenous separations ($s = 0$), and if steady-state equilibrium exists, then the equilibrium is unique and the distribution of wages is a point mass.*

This result stands in contrast to the famous result in Burdett and Mortensen (1998), that with on-the-job search and ex-ante homogeneous workers and firms, the equilibrium must contain wage dispersion. Like Albrecht et al. (2018), we achieve this by workers having horizontally differentiated preferences over workplaces. However, in our setting worker preferences over firms are time-varying rather than permanent. This both means that firms are not limited in their size in the long run by worker preferences, as well as that workers have non-degenerate outside utility distributions even if the outside wage distribution is a point mass. This means that the probability that a worker leaves a given job is a smooth function of their current wage and a competing outside wage, creating a smooth trade-off for firms between wage costs and turnover probabilities. If identical firms are facing the same smooth trade-off, then all firms will choose the same wage, generating the point mass of wages in equilibrium.[15] However, there may be wage dispersion if firms are ex-ante heterogeneous.

As we saw in Table 1, the recruiting cost adjusted markdown and the inverse labor supply elasticity are closely related. Now under this microfoundation, we can show that they are explicitly linked. The inverse labor supply elasticity is now

$$\varepsilon^{j}_{w,N} = \frac{\sigma_j \chi_j}{1+\gamma(1+\chi_j)}. \tag{12}$$

This is the ratio of the comparative static of wages with respect to firm demand shifter A_j over employment with respect to A_j: $\varepsilon_{w,N} = \frac{\partial \log w_j / \partial \log A_j}{\partial \log N_j / \partial \log A_j}$. As before, if $\sigma = 0$, then optimal wages will be unchanged in response to shifts in demand, but firms will increase their employment when demand for their product grows. This expression is the same as in Table 1, except that the superelasticity term drops out because the sum of the recruiting and negative separation elasticities $\mathcal{E}$ is constant and equal to γ. The recruiting cost adjusted markdown is

$$\frac{w_j + \text{recruiting cost per worker}_j}{MRPL_j} = \frac{1}{1+\varepsilon^{j}_{w,N}}. \tag{13}$$

If firms have an elastic labor supply curve, then the recruiting cost-adjusted wage markdown will be equal to 1.

2.3 Wages and Firm Employment Growth

In Section 2.2, we focused on the role of σ in determining whether firms have a constant returns to scale recruiting function, and by extension the inverse labor supply elasticity and the recruiting cost-adjusted markdown in steady state. In this section, we will focus on the theoretical response of wages to idiosyncratic firm demand shocks, which will be informative for whether firms have a constant returns to scale recruiting function or not, i.e., whether $\sigma = 0$ or $\sigma > 0$.

We consider how a firm's optimal wage responds to an idiosyncratic shock to A, which we interpret as an idiosyncratic demand shock. We log linearize the firm's problem around the firm's steady state, assuming an equilibrium with no wage dispersion ($s = 0$, and ex-ante identical firms). Define g_t as the firm's employment growth rate: $g_t \equiv \frac{N_t}{N_{t-1}}$. The firm's optimal wage is then

$$\hat{w}_t = \frac{1}{1+\gamma(1+\chi)} \left(\frac{\chi + \frac{\varepsilon_R}{\varepsilon_R - \varepsilon_S}}{S} - 1 \right) \hat{g}_t + \frac{\sigma \chi}{1+\gamma(1+\chi)} \hat{N}_{t-1}, \tag{14}$$

[15] This single-wage equilibrium can break down with exogenous separations. This is because, when there are exogenous separations, a greater number of workers accumulate in unemployment. With a larger pool of unemployed workers, firms may find it profitable to pay a low wage that on-the-job searchers are unlikely to take, but unemployed workers are likely to take, generating "stair-step" recruiting and separation functions. Higher values of γ makes targeting unemployed workers with a low wage more desirable, making it more likely for the single-wage equilibrium to break down. Numerical results suggest that for calibrations with reasonable unemployment rates, the single-wage equilibrium breaks down for values of γ between 15 and 20, far above the range of 4-6 needed to match empirical recruiting and separation elasticities.

where ε_R and ε_S are the recruiting and separation elasticities for this firm, evaluated at the steady state. A few points jump out from this expression. First, the coefficient on the number of incumbent workers N_{t-1} is $\frac{\sigma\chi}{1+\gamma(1+\chi)}$, which is the inverse labor supply elasticity and is consistent with our previous results. Given that we can calibrate γ with external evidence, and under the assumption that $\chi > 0$, empirically estimating the coefficient on $\hat{N}_{t-1}$ will test a null hypothesis of $\sigma = 0$, as we will test in the next section.

The coefficient on the employment growth rate is increasing in χ: when firms face diminishing returns to recruiting expenditure, it is increasingly costly for a firm to match with a marginal worker. To accelerate hiring in a limited pool of matched workers, firms want to pay higher wages, but only if matching with more workers is increasing expensive. Thus, a higher χ makes the firm's optimal wage response larger if the firm is trying to hire the same number of workers.

3 Empirical Evidence

In this section, we show two main results. First, we study the effect of export demand shocks on firm's employment level, employment growth rate, and the wage growth of switchers, as the export demand shocks provide plausibly exogenous, idiosyncratic, and persistent demand shocks across manufacturing firms. This allows us to estimate our parameter of interest σ, which determines the returns to scale in firms' recruiting function. Second, we document that the wage premium at large firms is small in Denmark across industries. This provides suggestive evidence that firms can become large without paying high wages in sectors outside of manufacturing.

3.1 Model to Data

Our model in Section 2.2 yields wages that may differ across firms, but there is no worker heterogeneity. When we turn to the data, workers will be heterogeneous, which is outside the model presented thus far. In Appendix A.8, we derive an extension where workers are heterogeneous and firms pay a wage per efficiency unit of labor. This provides a microfoundation for a wage equation where workers' wages are log additive in firm and worker effects, motivating empirical analysis that leverages an AKM (Abowd et al., 1999) framework. As we will see shortly in our empirical analysis, we leverage the log-additive nature of wages in order to difference out worker effects when observing the wage growth of job switchers. Let worker fixed effects be ζ_i and firm fixed effects ψ_j, with a time varying error term $e_{i,j,t}$. Then log wages are described by

$$\begin{aligned}\log(w_{i,j,t}) &= \zeta_i + \psi_{j,t} + e_{i,j,t} \\ &= \zeta_i + \tilde{\psi}_{j,t} + \psi_{j,0} + e_{i,j,t}\end{aligned}$$

where $\tilde{\psi}_{j,t}$ is the time-varying wage premium at firm j and $\psi_{j,0}$ is the time-invariant firm wage effect. First differencing this expression for workers who switch from firm k to firm j yields

$$\log(w_{i,j,t}) - \log(w_{i,k,t-2}) = \tilde{\psi}_{j,t} + \psi_{j,0} - \psi_k + e_{i,j,t} - e_{i,k,t-2},$$

assuming that firm j may be subject to shocks that affect is wage policy, while the wage effect of firm k is stable. Substituting the time varying firm effect $\tilde{\psi}_{j,t}$ with our model-implied expression for log deviations of wages $\hat{w}_t$ from firm j's initial wage from equation (14) yields

$$\log(w_{i,j,t}) - \log(w_{i,k,t-2}) = \beta_0 \hat{g}_{j,t} + \beta_1 \left(\log(N_{j,t-1}) - \log(N_{j,0})\right) + \psi_{j,0} - \psi_k + e_{i,j,t} - e_{i,k,t-2} \quad (15)$$

where $\beta_0 = \frac{1}{1+\gamma(1+\chi)} \frac{\chi + \frac{\varepsilon_R}{\varepsilon_R - \varepsilon_S}}{S} - 1$ and $\beta_1 = \frac{\sigma\chi}{1+\gamma(1+\chi)}$. Thus, the deviation of the wage growth of switchers in period t, relative to switcher wage growth in some pre-period when the firm was at steady state, will be a function of its employment growth rate in period t and the deviation of the firm's employment level prior to the period t shock $\hat{N}_{j,t-1}$ relative to its prior steady-state size. Given this equation, this estimation strategy requires estimating how switchers' wage growth changes in response to both contemporary firm demand shocks and past firm demand shocks.

We emphasize looking at the wage growth of switchers. It is common in the rent sharing literature to look at idiosyncratic shocks to firms, as in Kline et al. (2019), Garin and Silverio (2023), Lamadon et al. (2019). Using the response of employment and stayer wages to firm demand shocks, one could infer a labor supply elasticity, under the assumption that the only reason that firms are raising the wages for incumbents is that firms face a finite labor supply elasticity. However, there may be many reasons why firms raise the wages of incumbents in response to a shock that is not related to the long-run availability of workers at a given wage. For instance, Carvalho et al. (2023) shows that firms that receive contracts that are more financially remunerative, but do not necessarily require more labor, generate larger wage responses for incumbent workers. This is consistent with either profit sharing or incumbent workers extracting greater rents from the firm. A temporary period of high match rents could occur if new hires and incumbents are not substitutable in the short run, as Kline et al. (2019) argue and Mercan et al. (2021) provide evidence for. Furthermore, firms may offer incumbent higher future wage *paths* or contracts that are downwardly rigid, so if incumbent wages are persistently elevated years after a firm-specific shock, this may mean nothing for the availability of new hires at a given wage in the long run. Therefore, comparing the wage growth of switchers before, during, and after idiosyncratic firm demand shocks provides a cleaner measure of how much the firm needs to pay in order to acquire labor as the firm's size changes.[16]

3.2 Data

We construct an annual merged employer-employee panel dataset using Danish administrative data from 2001–2019. The administrative registries we use are maintained by Statistics Denmark. For

[16]We explore two different extensions that allow firms to pay workers of different tenure different wages, both of which demonstrate that the wage of new hires is the relevant object to study to estimate σ. In Appendix A.7 we derive a two period model where the firm can pay workers hired in the later period a different wage than incumbent workers or workers hired in the earlier period. In Appendix B, we assume that workers are forward looking, and firms post fixed-wage contracts separately for different cohorts of arriving workers. In both settings, if $\sigma = 0$, then the wage of new hires depends only on the current hiring rate, while incumbent wages of incumbents and past hires are affected by past hiring rates. Thus in both extensions, estimating labor supply elasticities using the wages of incumbents or past hires would yield finite long-run labor supply elasticities, even if $\sigma = 0$.

workers' employment history and wages, we use the IDAN (Danish Statistics, 2019) registry, which reports workers' earnings, hours worked, occupation, and employer. We combine this dataset with additional information about firms characteristics from the Danish business register (FIRM, Statistics Denmark (2019a)) and the accounting statistics register (FIRE, Statistics Denmark (2019d)). These data include firm sales, value added, employment (measured in full-time equivalents), gross salaries, and gross profits. We include only firms that report value added, which excludes the public sector and financial firms. Demographic information comes from the IDAP (Statistics Denmark, 2019b) and UDDA (Statistics Denmark, 2019e) registries.

3.3 Export Demand Shocks and Switcher Wage Growth

To estimate σ, which will be informative for the long-run inverse labor supply elasticity to the firm, we need variation in firms' demand for labor. To do so, we will construct shift-share export demand shocks to generate plausibly exogenous, idiosyncratic shocks to demand.

Sample For this exercise, we restrict our sample to firms that export goods. To construct our demand shocks, we use data on sales and exports of Danish manufacturing firms at the level of 6-digit product categories (HS codes). We have information on firm sales by product codes from a large-scale administrative survey (the VARS dataset, Statistics Denmark (2019c)). This survey is carried out by Statistics Denmark to construct the Danish component of the Eurostat PRODCOM dataset and is mandatory for all Danish manufacturing firms with more than 10 employees. We combine the sales data with administrative survey and customs data (the UHDM registry, Statistics Denmark (2019f)) to get information about the goods exports of Danish firms. Finally, we use publicly available data from UN Comtrade (United Nations, 2019) on international trade flows at the country-product level.

Our sample of shocked firms has 1,594 unique firms, with a mean size of 117.5 full time equivalent employees, with a standard deviation of 426.2. The median firm size is 41.0 full-time equivalent employees.[17] These statistics are additionally reported in Table 2.

Table 2: Firm Summary Statistics

	Mean	Std. Dev.	Median
Firm Size	117.5	426.2	41.0

Trade Shocks We use shocks to firm export demand using a global export demand shift-share IV similar to that used in Dhyne et al. (2022) and Hummels et al. (2014). For the shift, for each country-product pair, we construct the world import demand $WID_{c,p,t}$, which is country c's imports of product p from all countries excluding Denmark. We then compute the change in log

[17] Due to reporting requirements, the reported median is the mean firm size for firms between the 45th and 55th percentiles of the size distribution.

world import demand $d \log WID_{c,p,t}$ for each country-product-year cell. In our main specification, we use 4-digit product codes.

For the shares, we construct the export share of sales for each country-product type. A Danish firm j is exposed to the shock $WID_{c,p,t}$ if they export product p to country c. We use the country-product share of firm total exports in the prior period:

$$\omega_{j,c,p,t-1} = \frac{exports_{j,c,p,t-1}}{\sum_{c,p} exports_{j,c,p,t-1}}$$

Using the product of the shifts and shares, we can construct the predicted percent change in exports for firm j:

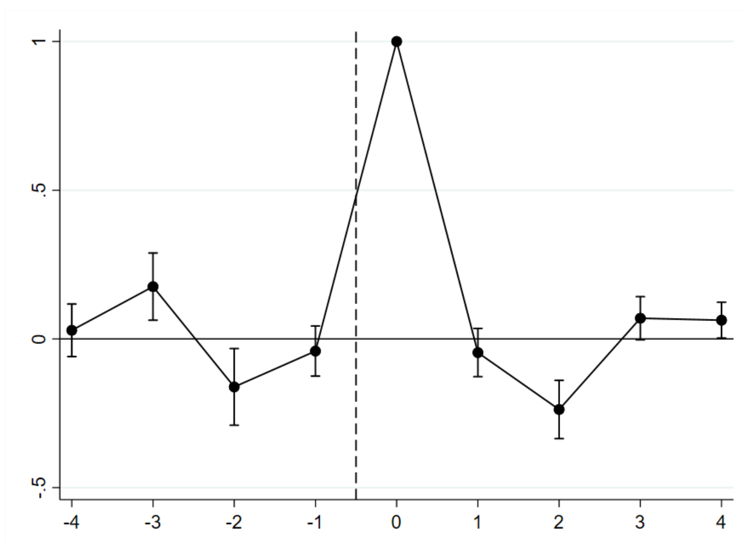
$$\tilde{z}_{j,t} = \sum_{c,p} \omega_{j,c,p,t-1} d \log WID_{c,p,t}.$$

Normalizing by the firm's prior export share of sales, we get a predicted percent change in sales for firm j

$$z_{j,t} = \Omega_{j,t-1} \tilde{z}_{j,t},$$

where $\Omega_{j,t-1}$ is the export share of sales for firm j in year $t-1$. We then winsorize this shock series at the 1st and 99th percentiles.

Figure 2: Trade Shock Autocorrelation



This figure show the coefficients of a regression a trade shock in period $t+\nu$ on a trade shock in period t, controlling for industry and year fixed effects. The coefficient of z_t on the current shock z_t is 1 by construction. Standard errors are clustered at the industry by year level.

These export demand shocks predict increases in actual sales: a 100 log point predicted increase in sales on average results in a 52 log point increase in actual sales.[18] On average, export demand shocks are relatively small: the standard deviation of shocks is 0.078. This means that in a typical

[18] Figure 5 in Appendix A.9 shows that on impact, a 100 log point increase in predicted sales increase actual sales

year, an exporting firm experiencing a 1-standard deviation shock in export demand should expect an increase of actual sales by around 4 percent. These shocks are also quite persistent and have little autocorrelation. In Figure 2 we show the results of regressing the leads and lags of the shock on the time t shock, including year and industry fixed effects. While there is some slight autocorrelation, these export demand shocks follow a near-unit root process (Dhyne et al., 2022).

Firm-Level Switcher Wage Growth In our main specification, we compute a firm level measure of switcher wage growth. When workers switch jobs, we want to control for the wage effect of the firm that they leave. So in the first step, we estimate an AKM regression to recover the pay policy of the firm that workers leave. We use a worker's wage growth over two years $\Delta \log(w)_{i,k,j,t-2,t} \equiv \log(w_{i,j,t}) - \log(w_{i,k,t-2})$. We use a two year horizon due to the annual nature of the data we use: the IDAN data reports a worker's "main job" in November and how many hours the worker had worked in that job in the past year. We define a worker as full time if they work at least 1400 hours within a year. A worker who switchers in year $t-1$ may not work full time in either job, and so we compare the wage in the last year of full time work in year $t-2$ relative to the wage in the first year of full time work in year t.[19] Using a sample of only switchers (worker-year observations when workers switch), we estimate

$$\Delta \log(w)_{i,k,j,t-2,t} = \alpha + \mathbf{X}_{i,t-2}B + \beta_\psi \hat{\psi}_k + e_{ijkt},$$

where $\mathbf{X}_{i,t-2}$ are mincer controls including years of education, potential experience and its interaction with education, potential experience squared and its interaction with education, and gender. $\hat{\psi}_k$ is the estimated AKM effect for the firm k that the worker leaves. We recover a predicted value of wage growth $\widehat{\Delta \log(w)}_{i,k,j,t-2,t}$. We then compute the residualized individual switcher wage $\widetilde{\Delta \log(w)}_{i,k,j,t-2,t} = \Delta \log(w)_{i,k,j,t-2,t} - \widehat{\Delta \log(w)}_{i,k,j,t-2,t}$. Next, we compute an average of the residualized individual switcher wage growth for the firm j that workers arrive at in year t

$$\Delta \log \tilde{w}_{j,t} = \frac{1}{n} \sum_{i=1}^{n} \widetilde{\Delta \log(w)}_{i,k,j,t-2,t}.$$

Thus, $\Delta \log \tilde{w}_{j,t}$ is firm-year average switcher log wage growth, residualized for worker demographics and the wage effect of the firm that the worker left. We winsorize our firm-level average switcher log wage growth at the 1st and 99th percentiles.

3.4 Empirical Design

Firm-Level Event Study Next, we will perform event studies to examine the effect of an export demand shock on the firm-level employment level and the wage growth of arriving job switchers.

by around 52 log points, with an F statistic of 28.78. The effect on actual sales falls somewhat in the following years, but is fairly persistent.

[19]The monthly BFL registry begins in 2008, and we study job switchers beginning in 2005. We do not have access to the SPELLS dataset that reports monthly observations prior to 2008.

We will estimate the effect of an export demand shock that occurs in year t. We want to know the path of log employment at firm j from $t-4$ to $t+3$. Similarly, we want to know the path of switcher wage growth for each periods $t-4$ through $t+3$.

$$\Delta \log N_{j,t-1,t+\nu} = \sum_{s=-4}^{0} \beta_s z_{t+s} + \epsilon_{j,t} \tag{16}$$

$$\Delta \log \tilde{w}_{j,t+\nu} = \sum_{s=-4}^{0} \zeta_s z_{t+s} + u_{j,t} \tag{17}$$

We estimate these regressions for $\nu \in [-4, 3]$. We additionally include firm fixed effects, industry-by-year fixed effects, as well as industry-by-year fixed effects interacted with the lagged export share (Borusyak et al., 2022).

IV Specification For the full IV specification, the ideal experiment would be to implement equation (15) by using the following IV specification,

$$\Delta \log \tilde{w}_{j,t+\nu} = \beta_0 \Delta \log N_{j,t-m,t-1} + \beta_1 \Delta \log N_{j,t-1,t} + \epsilon_{j,t},$$

instrumenting for current employment growth $\Delta \log N_{t-1,1}$ and past employment growth $\Delta \log N_{t-m,t-1}$ with current and past export demand shocks, where $t-m$ indicates some period in the past when the firm was at steady state. However, unbiased estimation of β_1 would require controlling for incomplete shares by including industry by year fixed effects interacted with export shares in year $t-1$ (Borusyak et al., 2022). These export shares will be correlated with export demand shocks in periods $t-1$ and earlier, and so coefficients on lagged employment β_0 in the same regression will be biased. To avoid this, we will instead fix the timing of the shock, and vary the timing of the outcome variables in our instrumental variables estimation. Specifically, we will estimate the wage growth of switchers who arrive in period t on the firm employment growth between $t-1$ and t, using as an instrument the export demand shock in period t. Second, we will regress wage growth of workers who arrive in future periods, specifically in $t+2$, on the employment growth of the firm from $t-1$ to $t+1$, again instrumenting employment growth with the export demand shock at time t. This yields the specifications

$$\Delta \log \tilde{w}_{t-2,t} = \alpha + \beta_0 \Delta \log N_{j,t-1,t} + \epsilon_{j,t} \tag{18}$$

$$\Delta \log \tilde{w}_{t,t+2} = \alpha + \beta_1 \Delta \log N_{j,t-1,t+1} + u_{j,t}, \tag{19}$$

where again we include firm fixed effects, industry by year fixed effects, as well industry by year fixed effects interacted with the firms' export share in year $t-1$. The coefficient β_1 will identify the combination of parameters $\frac{\sigma\chi}{1+\gamma(1+\chi)}$ from equation (15).

Robustness: Worker-Level Event Study As a robustness check, we use an individual-level local projection regression to estimate the dynamic path of an individual worker's wage as the worker switchers across firms, estimating the effect of current and past firm shocks on the path of an individual workers' wage. Equation (20) estimates the coefficient on the period t shock on the wage growth of a worker over a horizon from $t-2$ to $t+\nu$. This regression includes industry-by-year fixed effects, as wage as industry by year fixed effects interacted with a firm's export share of sales in $t-1$. This path of coefficients of β_s for each horizon ν will trace out the effect of an export demand shock hitting in year t on the path of wages for workers who switch into the firm in year t.

$$\Delta \log w_{i,t-2,t+\nu} = \alpha + \sum_{\nu=-4}^{0} \beta_s z_{t+s} + \mathbf{X}_{i,t-2} B + \beta_\psi \hat{\psi}_k + e_{ijkt}, \tag{20}$$

$$\Delta \log w_{i,t-2,t+\nu} = \alpha + \sum_{\nu=-4}^{-1} \beta_s z_{t+s} + \mathbf{X}_{i,t-2} B + \beta_\psi \hat{\psi}_k + e_{ijkt}, \tag{21}$$

As an additional test, we also estimate (21), which is the same regression except that we exclude the time t shock, and we include in our controls both industry by year fixed effects, as well as industry by year effects interacted with the export share of sales in year $t-2$. That way we can get an unbiased estimate for the coefficient β_{-1}, which estimates the effect of a export demand shock in period $t-1$ on the path of wages for workers who arrive in year t.

3.5 Results

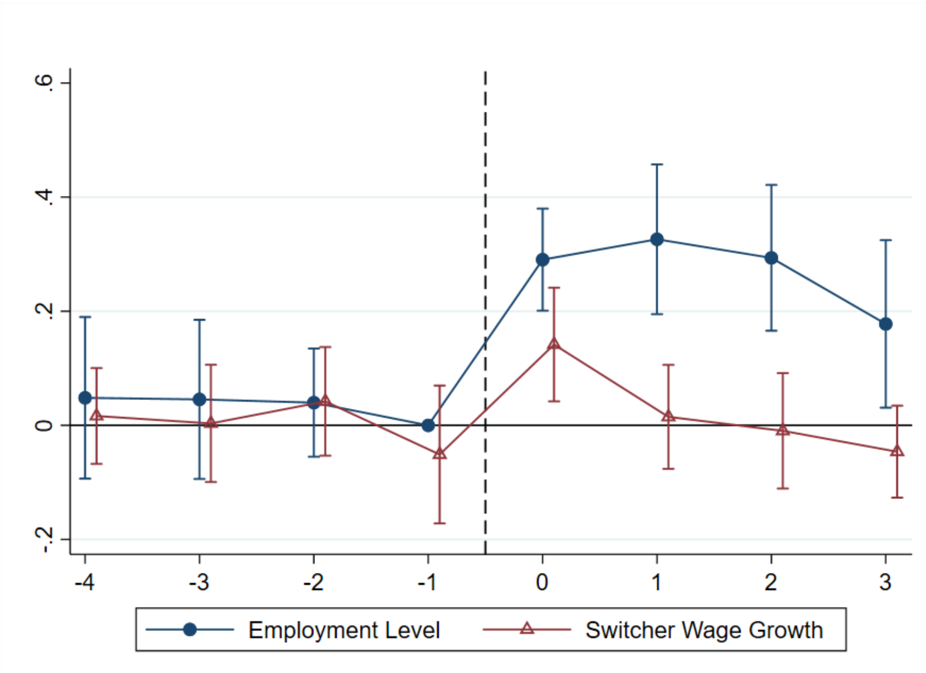
We begin by plotting the response of the firm employment level and wage growth of switchers (16) from equations (17) in Figure 3.[20] The solid circles show the effect of this shock on employment, and the hollow triangles report the effect on the wage growth of switchers, where each point represents a different cohort of new hires each year. Conditional on the sequence of shocks from $t-4$ to $t-1$, firms with different values of shocks in period t have similar pre-trends in both the level of employment and in the wage growth of workers who switched into the firm. On impact, an export demand shock that predicts a 100 log point increase in predicted sales increases a firm's employment level by approximately 30 log points.[21] The effect on employment is persistent: employment remains elevated at around 30 log points above the pre-shock level of employment two years after the shock occurs. For switcher wages, in the year of the shock, job switchers who move into firms that have a 100 log point export demand shock receive 14 log point higher wage growth than workers who

[20] We align the samples so that the coefficient on firm employment level and switcher wage growth within each time horizon reflects the same sample. However, this means that not every horizon has exactly the same sample, as not every firm makes a full time hire of a switcher each year. The sample for each horizon has between 1450 and 1580 unique firms. The sample of the period 0 horizon in Figure 3 , as well as the sample for the IV regression in Table 3 have 1492 unique firms.

[21] When using log sales as the outcome variable, a 100 log point increase in predicted sales increases actual sales by 50 log points. This is likely because the instrument is measured with error. We report the figure showing the response of log sales in Appendix A.9.

arrive at a firm with a 0 percent export demand shock. However, in the years after the shock, when the firm is larger but is no longer growing, new hire wage growth is similar to what is was before the shock occurred.

Figure 3: Response of Firm Employment Level and Switcher Wage Growth to Export Demand Shocks



The figure plots the path of the employment level and wage growth of switchers in response to a 100 log point predicted increase in sales due to an export demand shock in period t. The blue circles show that firms increase employment by approximately 30 log points, and the effect is persistent. The hollow red triangles show the effect of the period t trade shock on the wage growth of different cohorts of arriving switchers. Workers who switch into the firm in period t experience wage growth that is 14 log points greater than workers who switch into a firm that is not shocked. However, workers who switch into the shocked firm in subsequent years receive no greater wage increases that workers who arrived prior to the shock.

One concern is that the increase in labor demand for these firms may affect the demand for labor at the market level, and so that the stable unit treatment value assumption (SUTVA) would be violated. We believe that this is not a concern for two reasons. First, the median firm in our sample has 41 workers, and the standard deviation of an export demand shock is 0.078. This 7.8% increase in predicted sales would correspond to a roughly 2.6% increase in employment for treated firms. A 2.6% increase of employment for a 41 worker firm is growing by one worker, which is the smallest change of extensive margin employment possible. Given that Denmark is a relatively densely populated country, we assume that expanding by a single worker does not lead to a violation of SUTVA. Second, in Appendix A.9 we report robustness using a set of firms whose log employment average is below 4 during the years in the sample (which is average employment of 54 full-time equivalents or fewer). We find similar results, though the effect of switcher wages on impact appears larger in the subsample of smaller firms.

The results in Figure 3 provide the first stage and reduced form estimates for an IV estimation. The first three columns of Table 3 report the results for equation (19). In the first column, we include lagged shocks as instruments, while in column (2) we include lagged export demand shocks as controls. In third column, we control for employment growth between $t+1$ and $t+2$. In all of these specifications, the point estimate for wage growth for switching workers who arrive in period 2 is very close to zero. These coefficients estimate $\frac{\sigma\chi}{1+\gamma(1+\chi)}$, and since previous evidence suggests that $\chi > 0$, our empirical point estimates of zero for these regressions imply a point estimate of zero for σ.

Table 3: Instrumental Variable Estimation

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	$\Delta \log w_{t,t+2}$	$\Delta \log w_{t,t+2}$	$\Delta \log w_{t,t+2}$	$\Delta \log w_{t-2,t}$	$\Delta \log w_{t-2,t}$	$\Delta \log w_{t-2,t}$
$\Delta \log N_{t-1,t+1}$	0.0155	0.0178	0.0293			
	(0.154)	(0.173)	(0.184)			
$\Delta \log N_{t-1,t}$				0.673**	0.644**	0.488**
				(0.262)	(0.291)	(0.212)
Lagged Export Shocks	Instrument	Control	Control	Instrument	Control	Control
First Stage F Statistic	29.47	29.47	25.53	25.0	25.0	40.47
Control for $\Delta \log N_{t+1,t+2}$	N	N	Y	-	-	-
Observations	5,678	5,678	5,678	5,678	5,678	6,912

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

This table reports regression coefficients for equations (18) and (19). Columns 1-3 report instrumental variable regression of wage growth for workers who arrive in period $t+2$ on firm employment growth between period $t-1$ and $t+1$, where firm employment growth is instrumented by the firm export demand shock in period t. Columns 4-6 report instrumental variable regression of wage growth for workers who arrive in period t on firm employment growth between period $t-1$ and t, where firm employment growth is again instrumented by the firm export demand shock in period t. Standard errors are clustered at the industry-by-year level.

Columns 4, 5, and 6 of Table 3 report the regression for report results for equation (18). In the fourth column, we include lag export demand shocks as instruments, while in columns 5 and 6, we include lagged export demand shocks only as controls. Columns 5 and 6 differ only in their sample. The sample of firms in column 5 is the same sample used in columns 1-5, which requires that firms have non-missing values for all the variables needed to run each regression. Column 6 includes a larger sample that corresponds to the sample reported for the $t = 0$ coefficient in Figure 3. The range of point estimates is 0.49-0.67. This coefficient can be interpreted as the short run inverse labor supply elasticity. With short run inverse labor supply elasticities between 0.49 and

0.67, this would imply short run labor supply elasticities approximately between 1.5 and 2.

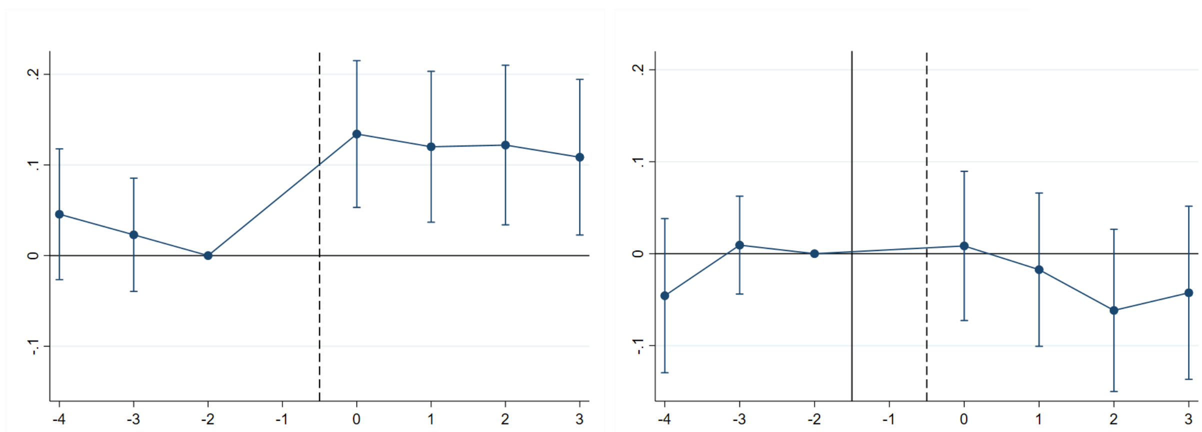
While we argue that σ is identified from the coefficients estimated in columns 1-3 in Table 3, we do not make an inference on the value of the parameter χ from the regression coefficients in columns 4-6 of Table 3. To do so would not be robust to a range of reasonable modifications to the model, such as wage adjustment costs or long-term contracting.[22] However, the fact that the wages of switchers are elevated in years of firm employment growth does suggest that firms are constrained in their ability to elastically increase recruiting activity. Therefore, we interpret our evidence in columns 4-6 that firms do face diminishing returns to recruiting expenditure, so $\chi > 0$, but we will not take a specific stance on the value of χ from this evidence.[23]

Finally, we report the dynamic path of individual worker wages for workers who switch jobs, estimating coefficients on the trade shocks that occur the year during and the year before workers switches with local projections. Figure 4(a) shows the coefficient of the time t shock on the wage growth of workers from $t-2$ to $t-\nu$. Visually inspecting the pre-trends, workers who switch into differently shocked firms appear to have similar paths of wages prior to switching. Consistent with evidence in Figure 3, wages of new hires are approximately 13% higher at firms facing a 100 log point shock relative to workers who arrive at firms with a shock of 0. The effect of switching on wages for workers who switch into the firm during the shock year is persistent.

Figure 4: Path of Switcher Wage Level by Trade Shock Timing Relative to Job Switch Timing

(a) Workers Arrive in Trade Shock Year

(b) Workers Arrive in Year After Trade Shock



These figure reports the coefficients from the equations (20) and (21). Panel (a) plots the path of an individual worker's wage level for a worker who switches in year t into a firm that experiences a 100 log point positive shock to sales in period t. Panel (b) plots the path of an individual worker's wage level for a worker who switchers in year t into a firm that experiences the same size of shock, but in period $t-1$.

[22]For example, Bloesch et al. (2024) embed the labor market outlined in this paper in a DSGE setting where firms are subject to quadratic nominal wage adjustment costs. For a given value of χ, larger wage adjustment costs would lead to a smaller slope of wages with respect to firm employment growth.

[23]As we discussed in Section 2.1, a value of χ between 1 and 4 would be consistent with evidence from the literatures on the relationship between firm growth rates and job filling rates, as well as evidence of the level of recruiting costs.

Figure 4(b) reports the coefficient on z_{t-1} on the wage changes of switchers who arrive in period t from a local projection like equation (20), except that we drop the trade shock in period t.[24] The point of this exercise is to study the wage path of workers who arrive in the firm a year after a trade shock occurs. This figure shows that workers who arrive in firms that had larger trade shocks in period $t-1$ receive no wage premium when they arrive in period t, and there is no subsequent catch-up wage growth for these late-arriving workers. This provides further evidence that firms do not need to pay new hires higher wages once firms become large.

It is worth noting that our model does not predict a persistent wage effect for new hires. Indeed, the worker side of the model is intentionally simple in order to focus on the effect of the firm's recruiting margin on the firm's wage setting behavior. One way to rationalize the persistent effect on wages for new hires is that firms offer workers contracts that specify a wage level or wage path in the future, either explicitly or implicitly. In Appendix B, we show that if firms offer fixed wage contracts to forward-looking workers, then the optimal wage that a firm posts is a function of the firm's current hiring rate and future hiring rates. Workers who receive a higher wage when hired receive persistently higher wages due to the long-term contract.[25]

3.6 Wages and Firm Size

So far in this empirical section, we have shown that when manufacturing firms are subject to export demand shocks, these firms do not pay any higher wages to new hires than the firms did prior to before the shock. In this subsection, we estimate the relationship between firm wage effects and firm size for Danish firms economy-wide. We estimate firm wage effects for private sector firms in the Danish economy from 2008-2019, using the results from estimating firm AKM effects in Section 3.3. We then regress these estimated firm wage effects on the log of a firm's average employment level in full time equivalent workers, averaged over 2008-2019. We do this for two separate samples of firms: first restricting on firms with 5 or more full time equivalent workers, and second, restricting on firms with 20 or more full time equivalent workers.

Table 4 reports the results. For firms with 5 or more workers, the elasticity of firm wage effects with respect to firm size is 0.019, or around 2%. Focusing only on firms with on average 20 workers or more, the elasticity drops to .009, or around 1%. These estimates are similar to the elasticities of firm wage effects on log firm size estimated in the United States (Bloom et al., 2018).[26] Together with our other results, this supports the view that firms pay high wages while growing, but firms do not need to pay high wages to be large.[27]

[24] We correspondingly include the industry-by-year fixed effects interacted with export shares for period $t-2$ instead of for year $t-1$.

[25] We also report the response of stayer wage growth to export demand shocks in Appendix A.9.

[26] Bloom et al. (2018) show that the elasticity of firm wage effects with respect to firm size is similarly small in the US. Other recent evidence in the United States has shown that the firm size wage premium has declined over time (Cobb and Lin, 2017; Song et al., 2019).

[27] This is consistent with evidence from Friedrich et al. (2023), who show that when firms suddenly expand (defined by firms increasing employment by at least 20% in one year), the residualized wages of newly hired workers is no

Table 4: AKM Firm Wage Effects and Log Firm Size

	$\hat{\psi}_j$	$\hat{\psi}_j$
$\log \bar{N}_j$	0.019***	0.0094***
	(0.0006)	(0.00098)
Observations	55,951	13,809
Sample	$\bar{N}_j \geq 5$	$\bar{N}_j \geq 20$

Robust standard errors in parentheses
*** p < 0.01, ** p<0.05, * p<0.1

Summary of Empirical Findings In total, we find that when firms are exposed to export demand shocks, firms increase employment on impact and offer higher wages to workers who arrive in the year of the shock. This is consistent with diminishing returns to recruiting expenditure ($\chi > 0$), so firms find it worth it to offer high wages in periods of growth. However, we find that once firms become large, they no longer offer new hires a premium relative to the wage that firms paid prior to the demand shock. We therefore cannot reject a constant returns to scale recruiting function, with $\sigma = 0$. As a result, our estimates show that firms have elastic supply curves in the long run, implying that the recruiting cost-adjusted markdowns are close to 1.

4 Equilibrium with Oligopsonistic Competition

In this section, we extend the equilibrium described in Section 2.2, taking into account an oligopsonistic labor market with strategic behavior. Additionally we introduce capital and imperfect product market competition, allowing us to fully characterize the capital, labor, and profit shares of aggregate income. We show that if firms have wage setting power but also have constant returns to scale recruiting functions, then we can match aggregate shares of income going to labor, capital, profits in national accounts data in rich countries, while also taking into account the effects of empirically realistic levels of labor market concentration.

Setting The unit of analysis is a local labor market. We assume there is no mobility of workers across local labor markets. In each local labor market, there is a mass 1 of workers. Firms are now non-atomistic, firms produce tradable goods only, and workers consume only traded goods. Each firm faces a constant elasticity of demand for its product, so output for firm j is $Y_j = CP_j^{-\epsilon}$, where C is a demand shifter. Firms produce with a Cobb-Douglas production technology $Y_j = A_j K_j^{\beta} N_J^{1-\beta}$. Capital depreciates at rate δ, and capital is rented elastically at rate r_K.

Firm's Problem and Equilibrium Firms behave strategically, knowing that their choices will affect equilibrium outcomes. The equilibrium concept that we choose is Nash in wages and vacan-

higher after the shock than prior to the shock.

cies: that is, when solving their problem, firm j takes as given the wage and vacancy policies of all other firms $-j$. We will focus on a labor market with unemployment, so $\lambda_{EU} = 0$ and $s = 0$. Firms have common recruiting cost parameters χ, σ, and c. From the mobility decision of workers, it must hold in steady state that

$$\frac{N_j}{N_k} = \frac{V_j}{V_k} \left(\frac{w_j}{w_k} \right)^{\gamma}.$$

Using the labor market clearing condition that $\sum_{j \in J} N_j = 1$, we have that

$$N_j = \frac{V_j w_j^{\gamma}}{V_j w_j^{\gamma} + \sum_{k \neq j} V_k w_k^{\gamma}}. \tag{22}$$

In the limit where the discount rate $\rho \to 0$, firms maximize the steady state flow of profits. Therefore, firms solve

$$\max_{N_j, V_j, w_j} P_j Y_j - r^K K_j - w_j N_j - \left(\frac{V_j}{N_j} \right)^{\chi} V_j N_j^{\sigma\chi}$$

subject to (22), $Y_j = C P_j^{-\epsilon}$, and $Y_j = A_j K_j^{\beta} N_J^{1-\beta}$, taking as given the wage and vacancy policies of the other firms. Since the mass of workers is 1, firm j's share of employment is just equal to their employment level $s_j = N_j$. When firms are optimizing, the recruiting cost-adjusted markdown is now equal to

$$\mu_j \equiv \frac{w_j + c \left(\frac{V_j}{N_j} \right)^{1+\chi} N_j^{\sigma\chi}}{MRPL_j} = \frac{1 + \gamma(1+\chi)}{1 + \gamma(1+\chi) + \sigma\chi + (1+\chi)\frac{s_j}{1-s_j}}, \tag{23}$$

where a firm with a larger share of employment s_j marks down wages more relative to marginal product.

A steady-state equilibrium then is a set of price levels, capital levels, wage policies, vacancies, and employment levels such that workers maximize utility, firms maximize profits given the wage and vacancy policies of other firms, and flows of workers into and out of firms balance.

Labor and Profit Shares in Equilibrium We will assume that turnover costs accrue to labor income.[28] Then in equilibrium, the share of income that goes to wage income and turnover costs (which are allocated to wages) as a share of gross output is

[28] Our results are not sensitive to this choice. Consider instead that recruiting activity takes labor away from production, so labor inputs are equal to $N_t - c(V_t/N_{t-1}^{1-\sigma})^{\chi} V_t$. For atomistic firms, if $\sigma = 0$, then $\mu = w/MRPL = 1$ and the gross labor share is simply $(1-\beta)\frac{\epsilon-1}{\epsilon}$. Intuitively, $\sigma = 0$ means that the tradeoff between wages and turnover costs is invariant to firm size, and so there is a constant marginal cost of labor. Firms then choose employment that equalizes the marginal product and marginal cost of labor.

$$\text{Gross Labor Share} \equiv \sum_j \frac{w_j N_j + c\left(\frac{V_j}{N_j}\right)^{1+\chi} N_j^{\sigma\chi}}{P_j Y_j}$$
$$= \frac{\epsilon - 1}{\epsilon}(1-\beta)\frac{\sum_j \mu_j P_j Y_j}{\sum_j P_j Y_j}.$$

When computing the net labor and profit share, we will just focus on J identical firms, so $s_j = 1/J$ for all j, in which case we also get a Herfindahl-Hirschman index of $HHI = 1/J$. Then the economy-wide wage markdown will be $\mu = \frac{1+\gamma(1+\chi)}{1+\gamma(1+\chi)+\sigma\chi+1/(J-1)}$. As we will discuss momentarily, we will be interested in the labor share of net value added, which subtracts depreciation of capital from GDP. Subtracting out depreciation from the denominator, the net labor share is

$$\text{Net Labor Share} = \frac{\frac{\epsilon-1}{\epsilon}(1-\beta)\mu}{1 - \frac{\epsilon-1}{\epsilon}\beta\frac{\delta}{\delta+r^K}}.$$

The net profit share is then

$$\text{Net Profit Share} = \frac{\frac{1}{\epsilon} + \frac{\epsilon-1}{\epsilon}(1-\beta)(1-\mu)}{1 - \frac{\epsilon-1}{\epsilon}\beta\frac{\delta}{\delta+r^K}},$$

where the share of gross value added from product market power is $\frac{1}{\epsilon}$, and the profit share of gross value added from wage markdowns is $\frac{\epsilon-1}{\epsilon}(1-\beta)(1-\mu)$.

Evidence on Profit and Labor Shares There are multiple different empirical measures of aggregate factor shares. The measures we will focus on are net shares of domestic corporate value added, which excludes housing and proprietor's income. Our preferred measure of the labor share exclude housing because housing is not traditionally a factor of production, and focusing on the corporate share allows us to avoid taking a stance on how to allocate proprietor's income, which can be challenging (Elsby et al., 2013). The "net" part of net domestic corporate value added means that depreciation is subtracted from value added, since payments to capital that pay off depreciation is not available for consumption (Koh et al., 2020). Rognlie (2016) shows for advanced economies, the labor share of net value added for the domestic corporate sector is between 70-80%.[29] The allocation of non-labor income between capital income and pure profit is usually not measured directly, so we will use additional evidence and the model to infer a capital and profit share of income.

Calibration Table 5 presents our choice of parameter values. We set the inverse scale parameter of workers' preferences over non-wage amenities $\gamma = 4$ to match microeconomic evidence on recruiting and separation elasticities. We set $\sigma = 0$ based on our evidence that atomistic firms have

[29] For the United States, data on value added of nonfinancial domestic corporate business can be found in Table 1.15 of the National Income and Product Accounts.

an elastic long run labor supply curve. We set the convexity of recruiting costs with respect to the number of vacancies $\chi = 2$ to match evidence on the level of turnover costs and marginal hiring costs. We set the product demand elasticity $\epsilon = 7$ to generate a price markup over marginal cost of 14%, which is relatively small but consistent with empirical evidence (Kline et al., 2019). We assume an ad-hoc risk premium on capital to be $r = 0.05$ (Jordà et al., 2019), a risk free real rate of zero, and a depreciation rate of $\delta = .075$. This yields a rental rate of capital $r^K = .125$. Together with a Cobb-Douglas capital share $\beta = .3$, the parameters jointly deliver a ratio of capital to output $K/Y \approx 2$, consistent with Rognlie (2018) and capital depreciation share of GDP of 15%,[30] and a gross capital share of income of 25%, consistent with Barkai (2020).

Table 5: Calibration

Parameter		Value	Reason
ϵ	Demand elasticity	7	Price Markup
r^K	Rental rate of capital	.125	Return on capital=.05
δ	Depreciation rate	.075	Match depreciation/GDP
β	Cobb-Douglas capital elasticity	.3	Match K/Y$\approx$ 2
γ	Inverse scale parameter of non-wage preferences	4	Match $\varepsilon_R - \varepsilon_S = 4$
s	Exogenous separation rate	0	Tractability
c	Hiring cost constant	1024	Match vacancy rate
λ^{EE}	On-the-job search probability	.16	Match J-J mobility

For the labor-related parameters, we set the rate of exogenous separations at $s = 0$, so we get the result that $\varepsilon_{R,w} - \varepsilon_{S,w} = \gamma$ at any wage level. We set the probability that workers can consider unemployment $\lambda_{EU} = 0$ to shut down the unemployment state. As can be seen in adjusted markdown expression in equation (23), it turns out that all other parameters that govern the search aspect of the model drop out. For example, the on-the-job search probability λ_{EE} and the cost of posting a vacancy c do not matter for markdowns and consequently labor and profit shares. Regardless, we set $\lambda_{EE} = .16$ to get a monthly job-to-job separation rate just above 2% (2.3% in (Bagger et al., 2022)) and $c = 1024$ to generate a job openings rate of 4.3% (4.1% rate in Bagger et al. (2022)).

Results Table 6 reports the share of income that goes to labor, capital, and profits under three scenarios. In the first column, we return to the assumption of atomistic firms, so $s_j = 0$ for all firms, but the recruiting margin is shut down ($\sigma = 1$, $\chi \to \infty$). In this case, wages are 20% below marginal product. Profits from price markups over marginal cost account for 17% of net value added, and net capital income (excluding depreciation) accounts for 12% of net value added. In total, this leaves a profit share of net value added equal to 31%, while the labor share of net value added is only 57%, significantly below the empirical range of 70-80%.

[30]See National Income and Product Accounts (NIPA) Table 1.11.

In the second column, we assume a constant returns to scale recruiting function ($\sigma = 0$) and only atomistic firms ($J \to \infty$). Product market profits and net capital as a share of value added are the same as in column 1, so the only difference is that the labor share of net value added is 71%, 14 percentage points higher than in the no-recruiting case. As consequence, the profit share of net value added is 14 percentage points lower. This case matches national accounts data on labor and non-labor shares of national income in rich countries.

Table 6: Distribution of Income: Standard Monopsony, CRS Recruiting, and Concentration: $\gamma = 4$

	Standard Monopsony $\sigma = 1$, $\chi \to \infty$	CRS Recruiting $\sigma = 0$	CRS + Oligopsony HHI = 0.1	National Accounts Data
Recruiting Cost-Adjusted Markdown	.80	1	.975	
Product Profits/Net Output	.17	.17	.17	
Total Profit/Net Output	.31	.17	.19	
Net Capital Income/Net Output	.12	.12	.12	
Non-Labor Income/Net Output	.43	.29	.31	.20-.30
Labor Income/Net Output	.57	.71	.69	.70-.80

This table reports the recruiting cost-adjusted markdown, and the following shares of aggregate net value added: product market profits, total profit, net capital income, non-labor income, labor income. For each model, we assume that the labor supply elasticity conditional on the number of vacancies $\mathcal{E} = \varepsilon_{R,w} - \varepsilon_{S,w} = 4$. The three separate economies being compared are (i) a standard monospony model where firms have no recruiting margin; (ii) our dynamic monopsony model where firms have a constant returns to scale recruiting function; and (iii) a finite number (10) of equally sized nonatomistic firms with constant returns to scale recruiting functions, leading to a labor market HHI of 0.1.

In third column, we allow for 10 equally sized non-atomistic firms and a constant returns to scale recruiting function ($\sigma = 0$). Compared to the case with no concentration, the recruiting cost-adjusted wage markdown is 2.5% percent lower, lowering the aggregate labor share of net income by 2 percentage points, putting the labor share just outside the empirical range. If local labor market as less concentrated as found by Schubert et al. (2023) (median HHI of .018 on a 0-1 scale), then the labor share of net value added would rise back above 70%.

Could the puzzle be resolved by eliminating price markups? As can be seen in in Column 1 of Table 6, profits from pricing power in the product market as a share of net value added (17%) are roughly equal to profits from wage setting power in the labor market (31% -17% = 14%). What would the profit share be if we shut down product market rents? To calculate this, we would want to re-calibrate the capital elasticity to $\beta = .25$ to get capital to output ratio of 2. In that case, with all other parameters the same, and $\gamma = 4$, $\chi \to \infty$, $\sigma = 1$, and $\epsilon \to \infty$, we would end up with a labor share of net output of 71%. That is, the labor share of net output can be matched, under a labor supply elasticity of 4 and no recruiting margin, only if price markups are completely eliminated.

Relation to Existing Models A few existing articles address labor market concentration in equilibrium with additional imperfect labor market concentration. Berger et al. (2022) study an environment where nonatomistic firms face upward sloping labor supply curves both within and across labor markets. The authors' model differs from ours in that theirs assumes Cournot competition with a nested labor market structure, while our equilibrium concept is closer to Bertrand, and we assume that there is no mobility across labor markets. Indeed, assuming Cournot competition in our setting with no cross-market labor mobility would lead to structural labor supply elasticities of zero. The main differences are that we find that (i) labor supply elasticities for atomistic firms (i.e., within market elasticities) are infinitely elastic, rather than an elasticity of around 11, (ii) for a given level of concentration, we estimate wages that are closer to marginal product, and (iii), to the extent that wages are below marginal product, a most of the gap is spent on recruiting and is not retained as profits. As such, we match aggregate capital, labor, and profit shares allowing for price markups over marginal cost, while Berger et al. (2022) match aggregate shares with profits being almost entirely driven by labor market markdowns.

Berger et al. (2024) write down a theoretical model with three party bargaining, labor market concentration, on-the-job search, and convex vacancy posting costs. Our models differ in that ours is a model of wage posting rather than bargaining, and we allow for incumbent workers to affect vacancy costs. Jarosch et al. (2021) write down a model of bargaining where firms can punish workers such that if negotiations between the firm and worker break down, the firm can remove its own job opening from that worker's search, weakening the worker's outside option if the firm is granular in the labor market. We abstract from this mechanism, allowing workers to switch to other jobs within the same firm.

5 Conclusion

In this article, we derive a model of dynamic monopsony where workers have heterogeneous preferences over firms and can search on the job, and firms can attract workers with higher wages and recruiting expenditures. We use this model to analytically decompose the share of marginal product into wages, recruiting costs, and profits. We provide empirical evidence consistent with constant returns to scale recruiting functions, implying that firm's long-run labor supply curves are very elastic, and that the rents firms get from exploiting their wage setting power are spent on recruiting. We show in equilibrium that elastic labor supply curves help match aggregate profit, capital, and labor shares of income if firms also have pricing power in the product market, even taking into account labor market concentration and strategic interactions. Our setting tractably combines numerous sources of monopsony, including preference heterogeneity, search frictions, and labor market concentration, in a way that is easily disciplined by new and existing empirical evidence and nests numerous existing monopsony models.

The tractability of our model may be useful for researchers who wish to model a labor market with on-the-job search, but do not want the added complexity of an equilibrium distribution of wages. For example, Bloesch et al. (2024) use a similar setting as the labor block of a New Keynesian model to study how wages evolve in response to cost-of-living shocks if firms post wages and workers search on the job. These authors find that the quit rate is a good predictor of nominal wage growth, as found empirically by Faberman and Justiniano (2015) and recently confirmed by Heise et al. (2024). Heterogeneity in firm wage policies can be easily included by including heterogeneous firms.

Our setting of time-varying idiosyncratic preferences and on-the-job search may be useful for tractably modeling the product market and firm choices of expenditure on advertising. Recasting spending on recruiting as spending on advertising, and idiosyncratic preferences over products rather than employers, there is a natural extension to characterize firms' pricing as a trade-off between advertising activity and retaining a base of customers, while extracting a flow of rents from existing customers. More broadly, this framework could be a tractable benchmark for future research relating to matching markets where currently matched agents continue to search.

References

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999). High Wage Workers and High Wage Firms. *Econometrica 67*(2), 251–333.

Albrecht, J., C. Carrillo-Tudela, and S. Vroman (2018). On-the-Job Search with Match-Specific Amenities. *Economics Letters 162*, 15–17.

Azar, J., S. Berry, and I. E. Marinescu (2019). Estimating Labor Market Power. *Available at SSRN 3456277*.

Azar, J., I. Marinescu, and M. Steinbaum (2020). Labor Market Concentration. *Journal of Human Resources*, 1218–9914R1.

Bagger, J., F. Fontaine, M. Galenianos, and I. Trapeznikova (2022). Vacancies, Employment Outcomes and Firm Growth: Evidence from Denmark. *Labour Economics 75*, 102103.

Balke, N. and T. Lamadon (2022). Productivity shocks, long-term contracts, and earnings dynamics. *American Economic Review 112*(7), 2139–2177.

Barkai, S. (2020). Declining Cabor and Capital Shares. *The Journal of Finance 75*(5), 2421–2463.

Bassier, I., A. Dube, and S. Naidu (2022). Monopsony in Movers The Elasticity of Labor Supply to Firm Wage Policies. *Journal of Human Resources 57*(S), S50–s86.

Berger, D., K. Herkenhoff, A. R. Kostøl, and S. Mongey (2024). An Anatomy of Monopsony: Search Frictions, Amenities, and Bargaining in Concentrated Markets. *NBER Macroeconomics Annual 38*(1), 1–47.

Berger, D., K. Herkenhoff, and S. Mongey (2022). Labor Market Power. *American Economic Review 112*(4), 1147–1193.

Bertheau, A., P. Cahuc, S. Jager, and R. Vejlin (2021). Turnover Costs: Evidence from Unexpected Worker Separations. Technical report, Working Paper.

Blatter, M., S. Muehlemann, and S. Schenker (2012). The Costs of Hiring Skilled Workers. *European Economic Review 56*(1), 20–35.

Bloesch, J., B. Larsen, and B. Taska (2022). Which Workers Earn More at Productive Firms? Position Specific Skills and Individual Worker Hold-up Power. Working Paper.

Bloesch, J., S. J. Lee, and J. Weber (2024). Do Cost-of-Living Shocks Pass Through to Wages? Working Paper.

Bloesch, J. and J. P. Weber (2022). Congestion in onboarding workers and sticky r&d. Working Paper.

Bloom, N., F. Guvenen, B. S. Smith, J. Song, and T. von Wachter (2018). The Disappearing Large-Firm Wage Premium. In *AEA Papers and Proceedings*, Volume 108, pp. 317–22.

Borusyak, K., P. Hull, and X. Jaravel (2022). Quasi-Experimental Shift-Share Research Designs. *The Review of Economic Studies 89*(1), 181–213.

Brown, C. and J. Medoff (1989). The Employer Size-Wage Effect. *Journal of Political Economy 97*(5), 1027–1059.

Burdett, K. and D. T. Mortensen (1998). Wage Differentials, Employer Size, and Unemployment. *International Economic Review*, 257–273.

Burks, S. V., B. Cowgill, M. Hoffman, and M. Housman (2015). The Value of Hiring through Employee Referrals. *The Quarterly Journal of Economics 130*(2), 805–839.

Caldwell, S. and N. Harmon (2019). Outside Options, Bargaining, and Wages: Evidence from Coworker Networks. *Unpublished manuscript, Univ. Copenhagen*, 203–207.

Card, D. (2022). Who Set Your Wage? *American Economic Review 112*(4), 1075–1090.

Carrillo-Tudela, C., H. Gartner, and L. Kaas (2023). Recruitment Policies, Job-Filling Rates, and Matching Efficiency. *Journal of the European Economic Association 21*(6), 2413–2459.

Carvalho, M., J. Galindo da Fonseca, and R. Santarrosa (2023). How are Wages Determined?: a Quasi-Experimental Test of Wage Determination Theories. *Cahier de recherche* (2023-01).

Chan, M., S. Salgado, and M. Xu (2020). Heterogeneous Passthrough from TFP to Wages. *Available at SSRN 3538503*.

Cobb, J. A. and K.-H. Lin (2017). Growing Apart: The Changing Firm-Size Wage Premium and its Inequality Consequences. *Organization Science 28*(3), 429–446.

Danish Statistics (2001-2019). IDA ansættelser (IDAN.

Datta, N. (2023). The Measure of Monopsony: the Labour Supply Elasticity to the Firm and its Constituents. Working Paper.

Davis, S. J., R. J. Faberman, and J. C. Haltiwanger (2013). The Establishment-Level Behavior of Vacancies and Hiring. *The Quarterly Journal of Economics 128*(2), 581–622.

Derenoncourt, E. and D. Weil (2024). Voluntary Minimum Wages. NBER Working Paper No. 32546.

Dhyne, E., A. K. Kikkawa, T. Komatsu, M. Mogstad, and F. Tintelnot (2022). Foreign Demand Shocks to Production Networks: Firm Responses and Worker Wmpacts. No. w 30447. National Bureau of Economic Research.

Dube, A., E. Freeman, and M. Reich (2010). Employee Replacement Costs.

Dustmann, C., A. Glitz, U. Schönberg, and H. Brücker (2016). Referral-Based Job Search Networks. *The Review of Economic Studies 83*(2), 514–546.

Elsby, M. W., B. Hobijn, and A. Şahin (2013). The Decline of the US Labor Share. *Brookings papers on economic activity 2013*(2), 1–63.

Engbom, N., C. Moser, and J. Sauermann (2022). Firm Pay Dynamics. *Journal of Econometrics*.

Faberman, R. J. and A. Justiniano (2015). Job Switching and Wage Growth. *Chicago Fed Letter*.

Friedrich, M., I. Helm, and U. Schoönberg (2023). Firm Expansion in Imperfect Labor Markets. Presented at the Imperfect Competition in the Labor Market Conference, IZA Bonn, June 2023.

Garin, A. and F. Silverio (2023). How Responsive are Wages to Firm-Specific Changes in Labor Demand? Evidence from Idiosyncratic Export Demand Shocks. *Forthcoming, Review of Economic Studies*.

Gavazza, A., S. Mongey, and G. L. Violante (2018). Aggregate Recruiting Intensity. *American Economic Review 108*(8), 2088–2127.

Hall, R. E. and A. I. Mueller (2018). Wage Dispersion and Search Behavior: The Importance of Nonwage Job Values. *Journal of Political Economy 126*(4), 1594–1637.

Heise, S., J. Pearce, and J. Weber (2024). Wage Growth and Labor Market Tightness. *FRB of New York Staff Report* (1128).

Heise, S. and T. Porzio (2022). Labor Misallocation Across Firms and Regions. Working Paper.

Hummels, D., R. Jørgensen, J. Munch, and C. Xiang (2014). The Wage Effects of Offshoring: Evidence from Danish Matched Worker-Firm Data. *American Economic Review 104*(6), 1597–1629.

Isen, A. (2013). Dying to Know: Are Workers Paid Their Marginal Product? Working Paper.

Jäger, S. and J. Heining (2022). How Substitutable are Workers? Evidence from Worker Deaths. National Bureau of Economic Research.

Jahn, E. and M. Neugart (2020). Do Neighbors Help Finding a Job? Social Networks and Labor Market Outcomes After Plant Closures. *Labour Economics 65*, 101825.

Jarosch, G., J. S. Nimczik, and I. Sorkin (2021). Granular Search, Market Structure, and Wages. *Revise and Resubmit, Review of Economic Studies*.

Jordà, Ò., K. Knoll, D. Kuvshinov, M. Schularick, and A. M. Taylor (2019). The Rate of Return on Everything, 1870–2015. *The Quarterly Journal of Economics 134*(3), 1225–1298.

Kaas, L. and P. Kircher (2015). Efficient Firm Dynamics in a Frictional Labor Market. *American Economic Review 105*(10), 3030–3060.

Kline, P., N. Petcova, H. Williams, and O. Zidar (2019). Who Profits from Patents? Rent-Sharing at Innovative Firms. *The Quarterly Journal of Economics 1343*, 1404.

Koh, D., R. Santaeulàlia-Llopis, and Y. Zheng (2020). Labor Share Decline and Intellectual Property Products Capital. *Econometrica 88*(6), 2609–2628.

Kuhn, P. (2004). Is Monopsony the Right Way to Model Labor Markets? A Review of Alan Manning's Monopsony in Motion. *International Journal of the Economics of Business 11*(3), 369–378.

Lamadon, T., M. Mogstad, and B. Setzler (2019). Imperfect Competition, Compensating Differentials and Rent Sharing in the US Labor Market. Technical report, National Bureau of Economic Research.

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.

Manning, A. (2003). *Monopsony in Motion*. Princeton University Press.

Manning, A. (2006). A Generalised Model of Monopsony. *The Economic Journal 116*(508), 84–100.

Manning, A. (2011). Imperfect Competition in the Labor Market. In *Handbook of Labor Economics*, Volume 4, pp. 973–1041. Elsevier.

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

Matsudaira (2014). Monopsony in the Low-Wage Labor Market? Evidence from Minimum Nurse Staffing Regulations. *The Review of Economics and Statistics 96*(1), 92–102.

Menzio, G. (2024). Search Theory of Imperfect Competition with Decreasing Returns to Scale. *Journal of Economic Theory 218*, 105827.

Mercan, Y., B. Schoefer, and P. Sedláček (2021). A Congestion Theory of Unemployment Fluctuations. Working Paper.

Michaillat, P. and E. Saez (2024). u= uv: The Full-Employment Rate of Unemployment in the United States. *Brookings Papers on Economic Activity 55*(2).

Mongey, S. and G. L. Violante (2023). Macro Recruiting Intensity from Micro Data. NBER Working Paper No. w26231.

Muehlemann, S. and H. Pfeifer (2016). The Structure of Hiring Costs in Germany: Evidence from Firm-Level Data. *Industrial Relations: A Journal of Economy and Society 55*(2), 193–218.

Naidu, S., E. A. Posner, and G. Weyl (2018). Antitrust Remedies for Labor Market Power. *Harvard Law Review 132*(2), 536–601.

Oi, W. Y. and T. L. Idson (1999). Firm Size and Wages. *Handbook of Labor Economics 3*, 2165–2214.

Rognlie, M. (2016). Deciphering the Fall and Rise in the Net Capital Share: Accumulation or Scarcity? *Brookings Papers on Economic Activity 2015*(1), 1–69.

Rognlie, M. (2018). Comment on "Accounting for Factorless Income". In *NBER Macroeconomics Annual 2018, volume 33*, pp. 235–248. University of Chicago Press.

Schmieder, J. F. (2023). Establishment Age and Wages. *Journal of Econometrics 233*(2), 424–442.

Schubert, G., A. Stansbury, and B. Taska (2023). Employer Concentration and Outside Options. Working Paper.

Seegmiller, B. (2021). Valuing Labor Market Power: The Role of Productivity Advantages. Working Paper.

Sokolova, A. and T. Sorensen (2021). Monopsony in Labor Markets: A Meta-Analysis. *ILR Review 74*(1), 27–55.

Song, J., D. J. Price, F. Guvenen, N. Bloom, and T. Von Wachter (2019). Firming up Inequality. *The Quarterly Journal of Economics 134*(1), 1–50.

Sorkin, I. (2018). Ranking Firms using Revealed Preference. *The Quarterly Journal of Economics 133*(3), 1331–1393.

Statistics Denmark (2001-2019a). Generel firmastatistik (FIRM).

Statistics Denmark (2001-2019b). IDA persondata (IDAP).

Statistics Denmark (2001-2019c). Industriens salg af varer (VARE).

Statistics Denmark (2001-2019d). Regnskabsstatistikken (FIRE).

Statistics Denmark (2001-2019e). Uddannelse (UDDA, Education).

Statistics Denmark (2001-2019f). Udenrigshandel diskretioneret (UHDM).

Stevens, M. (2004). Wage-tenure Contracts in a Frictional LabourMmarket: Firms' Strategies for Recruitment and Retention. *The Review of Economic Studies 71*(2), 535–551.

Tanaka, S., L. Warren, and D. Wiczer (2023). Earnings Growth, Job Flows and Churn. *Journal of Monetary Economics 135*, 86–98.

Trottner, F. (2022). Who Gains from Scale? Working Paper.

United Nations (2001-2019). COMTRADE Database. https://comtradeplus.un.org/.

A Appendix

A.1 Firm's Dynamic Problem

This appendix section derives the decomposition of marginal product in Section 2.1, quantifies the effect of non-zero discounting, and derives the inverse labor supply elasticity in Section 2.1.

If the firm is operating within a stationary environment, (i.e., $S(w)$ and $R(w)$ are not changing), the firm maximizes the present discounted value of profits. Ignoring a firm subscript, the firm's problem is:

$$\max_{\{N_t\},\{w_t\},\{V_t\}} \sum_{t=0}^{\infty} \Big(\frac{1}{1+\rho}\Big)^t (AN_t^{\alpha} - w_t N_t - c\Big(\frac{V_t}{N_{t-1}}\Big)^{\chi} N_{t-1}^{\chi\sigma} V_t)$$
$$\text{s.t. } N_t = (1 - S(w_t))N_{t-1} + R(w_t)V_t.$$

The lagrangian is:

$$\mathcal{L} : \sum_{t=0}^{\infty} \Big(\frac{1}{1+\rho}\Big)^t \big(AN_t^{\alpha} - w_t N_t - cV_t^{1+\chi} N_{t-1}^{\chi(\sigma-1)} + \lambda_t[(1 - S(w_t))N_{t-1} + R(w_t)V_t - N_t]\big).$$

The first order conditions are:

$$\begin{aligned}
\mathcal{L}_{N_t} :& \alpha A N_t^{\alpha-1} - w_t - \frac{1}{1+\rho} c\chi(\sigma-1)V_{t+1}^{1+\chi} N_t^{\chi(\sigma-1)-1} - \lambda_t + \frac{1}{1+\rho}\lambda_{t+1}(1 - S(w_{t+1})) = 0 \\
\mathcal{L}_{w_t} :& - N_t - \lambda_t S'(w_t)N_{t-1} + \lambda_t R'(w_t)V_t = 0 \\
\mathcal{L}_{V_t} :& - c(1+\chi)V_t^{\chi} N_{t-1}^{\chi(\sigma-1)} + \lambda_t R(w_t) = 0.
\end{aligned}$$

Rearranging first order condition on V_t yields, solving in steady state, and that $\frac{V}{N} = \frac{S(w)}{R(w)}$:

$$\lambda = \frac{c(1+\chi)(\frac{S(w)}{R(w)})^{\chi}}{R(w)} N^{\sigma\chi}.$$

Using this expression for λ and the first order condition on wages yields

$$w = c(1+\chi)\Big(\frac{S(w)}{R(w)}\Big)^{1+\chi}\big(\varepsilon_{R,w} - \varepsilon_{S,w}\big)N^{\sigma\chi}. \tag{24}$$

The steady state value of employment N can be solved for using the previous two expresssions and the first order condition on employment. The choice of firm size N will depend on the discount factor, as growing requires an upfront costs, so firms that discount more steeply will choose to be smaller. Optimal employment is:

$$\alpha A N^{\alpha-1} = \frac{w}{\mathcal{E}(1+\chi)}\Big(1 + \mathcal{E}(w)(1+\chi) + \sigma\chi + \frac{\rho}{1+\rho}\big(\chi(1-\sigma) + (1+\chi)\frac{1-S(w)}{S(w)})\Big). \tag{25}$$

In steady state, the level of vacancies is given by

$$V = \frac{S(w)}{R(w)}N. \tag{26}$$

Collectively, equations (24), (25), and (26) characterize the firm's optimal choice of wages, employment and vacancies in steady state.

With these expressions, we can solve for what shares of marginal product go to wages, turnover costs, and profits in steady state. Marginal revenue product is $MRPL = \alpha A N^{\alpha-1}$.

$$\frac{w}{MPRL} = \frac{\mathcal{E}(1+\chi)}{1+\mathcal{E}(1+\chi)+\sigma\chi+\frac{\rho}{1+\rho}\big(x(1-\sigma)+(1+\chi)\frac{S(w)}{1-S(w)}\big).}$$

Recruiting costs per worker, in steady state, is $c(V/N)^{1+\chi}N^{1+\sigma\chi}/N$. As a share of marginal product, these costs are:

$$\frac{\text{Recruiting Costs per Worker}}{MPRL} = \frac{1}{1+\mathcal{E}(1+\chi)+\sigma\chi+\frac{\rho}{1+\rho}\big(x(1-\sigma)+(1+\chi)\frac{S(w)}{1-S(w)}\big).}$$

The labor market profits per worker is the gap between marginal product and the sum of wages and per-incumbent recruiting costs.

$$\frac{\text{Labor Market Profits per Worker}}{MPRL} = \frac{\sigma\chi+\frac{\rho}{1+\rho}\big(\frac{1-S(w)}{S(w)}\big)}{1+\mathcal{E}(1+\chi)+\sigma\chi+\frac{\rho}{1+\rho}\big(x(1-\sigma)+(1+\chi)\frac{S(w)}{1-S(w)}\big).}$$

How large is this additional term due to discounting? At a monthly frequency, total monthly separation rates are approximately 0.04, and given an annual discount rate, ρ can be approximated to being equal to .004. Setting $\sigma = 0$ and $\chi = 1$, we have:

$$\frac{\rho}{1+\rho}\big(\chi(1-\sigma)+(1+\chi)\frac{1-S(w)}{S(w)}\big) = \frac{.004}{.996}\times(1+2\times\frac{.96}{.04}) \approx .2.$$

Picking $\mathcal{E} = 4$, the profit per worker as a share of marginal product collected by firms due to discounting is then (suppose $\sigma = 0$):

$$\frac{.2}{1+4(1+2)+.2} \approx .015.$$

This calculation implies that under standard parameters, the additional profits collected on the margin are 1.5% of marginal product. Economically, standard time preferences is quantitatively unimportant because the discounting from time preference rate is an order of magnitude smaller than the separation rate. Because matching with a worker requires an upfront investment with a future flow of payoffs, but the life of that flow of payoffs is affected by the separation rate, the relevant discount rate to the firm is the sum of the separation rate and time preference parameter.

Deriving the Inverse Labor Supply Elasticity in General Form To derive the inverse labor supply elasticity, we start with the optimal wage equation:

$$w = c(1+\chi)\left(\frac{S(w)}{R(w)}\right)^{1+\chi}\big(\varepsilon_{R,w} - \varepsilon_{S,w}\big)N^{\sigma\chi}.$$

Define $\mathcal{E}(w) = \big(\varepsilon_{R,w} - \varepsilon_{S,w}\big)$, which is the same as before. In logs, this is:

$$\log(w) = \log(c(1+\chi)) + (1+\chi)\log\left(\frac{S(w)}{R(w)}\right) + \log(\mathcal{E}(w)) + \sigma\chi\log(N).$$

Taking the total derivative with respect to $\log(w)$ yields:

$$\begin{aligned} 1 &= (1+\chi)\big(\varepsilon_{S,w} - \varepsilon_{R,w}\big) + \varepsilon_{\mathcal{E},w} + \log(\sigma\chi)\varepsilon_{N,w} \\ &= -(1+\chi)\mathcal{E}(w) + \varepsilon_{\mathcal{E},w} + \log(\sigma\chi)\varepsilon_{N,w}. \end{aligned}$$

Noting that $\varepsilon_{w,N} = \varepsilon_{N,w}^{-1}$ and rearranging yields:

$$\varepsilon_{w,N} = \frac{\sigma\chi}{1 + (1+\chi)\mathcal{E}(w) - \varepsilon_{\mathcal{E}(w),w}}.$$

A.2 Firm's Problem with Nonparametric Recruiting Costs

In this section, we derive a general firm problem of dynamic monopsony with a non-parametric hiring cost function in partial equilibrium, and we show that marginal product can be analytically decomposed into wages, recruiting costs, and profit as a function of the separation elasticity and elasticities of the hiring cost function.

In this general firm problem, the firm maximizes an infinite sum of discounted profits. Firms produce with a diminishing returns to scale production function using only labor $F(N)$, with $F'(N) > 0$ and $F''(N) < 0$. Firms set wages w_t each period and pays cost $\mathcal{C}(H_t, N_{t-1}, w_t)$ to make H_t hires, with a cost that depends on the number of hires H_t, the wage w_t, and the number of incumbent workers from the prior period N_{t-1}. Firms solve:

$$\max_{\{N_t\},\{H_t\},\{w_t\}} \sum_{t=0}^{\infty}\left(\frac{1}{1+\rho}\right)^t\Big(F(N_t) - w_t N_t - \mathcal{C}(H_t, N_{t-1}, w_t)\Big),$$

subject to the law of motion for employment:

$$N_t = (1 - S(w_t))N_{t-1} + H_t,$$

where the rate at which workers separate $S(w_t) \in [0,1]$ is a decreasing function of the wages this period: $S'(w_t) < 0$.

In steady state, the firm's optimal wage trades off wage costs and recruiting costs. Supressing the time subscripts and solving for the optimal steady state wage yields:

$$w^* = \frac{\mathcal{C}(H,N,w)}{N}\big(- \varepsilon_{\mathcal{C},w} - \varepsilon_{S,w} \times \varepsilon_{\mathcal{C},H}\big)$$

There are natural restrictions that we apply to the cost function. First, conditional on the number of hires and incumbents, the recruiting cost function should be at least weakly decreasing in the wage: $\mathcal{C}_w(H,N,w) \leq 0$, and so $\varepsilon_{C,w} \leq 0$ as well. Hiring costs should be strictly increasing in the number of hires so $\varepsilon_{C,h} > 0$, and by assumption the separation rate is decreasing in the wage, so $\varepsilon_{S,w} < 0$. The optimal wage finds the point at which lowering wages further yields an equal increase in hiring costs. Letting $\rho \to 0$, we have the ratio of wages to marginal product

$$\frac{w}{F'(N)} = \frac{-\varepsilon_{\mathcal{C},w} - \varepsilon_{S,w} \times \varepsilon_{\mathcal{C},H}}{\varepsilon_{\mathcal{C},N} - \varepsilon_{\mathcal{C},w} + \varepsilon_{\mathcal{C},H}(1 - \varepsilon_{S,w})}$$

Hiring costs per worker as a share of marginal product are

$$\frac{\mathcal{C}(H,N,w)/N}{F'(N)} = \frac{1}{\varepsilon_{\mathcal{C},N} - \varepsilon_{\mathcal{C},w} + \varepsilon_{\mathcal{C},H}(1 - \varepsilon_{S,w})}$$

What is not paid in wages or turnover costs is retained as profit

$$\frac{F'(N) - w - \mathcal{C}(H,N,w)/N}{F'(N)} = \frac{\varepsilon_{\mathcal{C},N} + \varepsilon_{\mathcal{C},H} - 1}{\varepsilon_{\mathcal{C},N} - \varepsilon_{\mathcal{C},w} + \varepsilon_{\mathcal{C},H}(1 - \varepsilon_{S,w})}.$$

This final expression captures what Manning (2006) calls "diseconomies of scale in recruiting." If the sum of the elasticity of hiring costs with respect to hires $\varepsilon_{\mathcal{C},H}$ and the elasticity of hiring costs with respect to incumbents $\varepsilon_{\mathcal{C},N}$ is greater than 1. Another way of looking at this is, if hiring costs are linear in the number of hires, then $\varepsilon_{\mathcal{C},H} = 1$. To the extent that hiring costs are more convex than linear with respect to the number of hires, the hiring cost of function exhibits diseconomies of scale if negative values of $\varepsilon_{\mathcal{C},N}$ does not offset values of $\varepsilon_{\mathcal{C},H}$ above 1; that is, being large does not offset the higher cost of hiring due to convexity in hiring cost.

A.3 Recruiting Costs as Reallocating Labor

Suppose that rather than an expenditure, recruiting activity by firms means taking labor away from production. The a firm's labor input is

$$N_t - cV_t^{1+\chi} N_{t-1}^{-\chi(1-\sigma)}.$$

If product demand is downward sloping and $Y_t = P_t^{-\epsilon}$, then a firm that doesn't discount ($\rho \to 0$) maximizes the flow rate of profits:

$$\left(K_t^{\beta} \left(N_t - cV_t^{1+\chi} N_{t-1}^{-\chi} \right)^{1-\beta} \right)^{\frac{\epsilon-1}{\epsilon}} - w_t N_t - r^K K_t.$$

Given that $V_t = \frac{S(w_t)}{R(w_t)} N_t$ in steady state, and dropping time subscripts, we have that the firm solves

$$\max_{K,N,w} \left(K^{\beta} \left(N - c \left(\frac{S(w)}{R(w)} \right)^{1+\chi} N \right)^{1-\beta} \right)^{\frac{\epsilon-1}{\epsilon}} - wN - r^K K$$

$$K^{\beta\frac{\epsilon-1}{\epsilon}} N^{(1-\beta)\frac{\epsilon-1}{\epsilon}} \left(1 - c \left(\frac{S(w)}{R(w)} \right)^{1+\chi} \right)^{(1-\beta)\frac{\epsilon-1}{\epsilon}} - wN - r^K K.$$

Taking first order conditions, we have

$$FOC_N : (1-\beta)\frac{\epsilon-1}{\epsilon} K^{\beta\frac{\epsilon-1}{\epsilon}} N^{(1-\beta)\frac{\epsilon-1}{\epsilon}-1} \left(1 - c \left(\frac{S(w)}{R(w)} \right)^{1+\chi} \right)^{(1-\beta)\frac{\epsilon-1}{\epsilon}} = w.$$

$$FOC_w : (1-\beta)\frac{\epsilon-1}{\epsilon} K^{\beta\frac{\epsilon-1}{\epsilon}} N^{(1-\beta)\frac{\epsilon-1}{\epsilon}} \left(1 - c \left(\frac{S(w)}{R(w)} \right)^{1+\chi} \right)^{(1-\beta)\frac{\epsilon-1}{\epsilon}-1} \times$$
$$(-c(1+\chi)) \left(\frac{S(w)}{R(w)} \right)^{1+\chi} \left(\frac{S'(w)}{S(w)} - \frac{R'(w)}{R(w)} \right) = N.$$

Rearranging the first order condition on N and plugging it into the first order condition for w, we have

$$(1-\beta)\frac{\epsilon-1}{\epsilon} K^{\beta\frac{\epsilon-1}{\epsilon}} N^{(1-\beta)\frac{\epsilon-1}{\epsilon}} \left(1 - c \left(\frac{S(w)}{R(w)} \right)^{1+\chi} \right)^{(1-\beta)\frac{\epsilon-1}{\epsilon}} = wN \tag{27}$$

Rearranging the first order condition on wages, we have

$$\frac{wN}{1 - c\left(\frac{S(w)}{R(w)}\right)^{1+\chi}} (-c(1+\chi)) \left(\frac{S(w)}{R(w)} \right)^{1+\chi} \left(\frac{S'(w)}{S(w)} - \frac{R'(w)}{R(w)} \right) = N. \tag{28}$$

$$c(1+\chi) \left(\frac{S(w)}{R(w)} \right)^{1+\chi} (\varepsilon_{R,w} - \varepsilon_{S,w}) = 1 - c \left(\frac{S(w)}{R(w)} \right)^{1+\chi}. \tag{29}$$

So we once again have the result that if $\sigma = 0$, then optimal wages are not a function of firm size N. Rearranging the first order condition on N once again, we come away with the simple formula

$$(1-\beta)\frac{\epsilon-1}{\epsilon} PY = wN,$$

and consequently

$$\frac{wN}{PY} = (1-\beta)\frac{\epsilon-1}{\epsilon}.$$

A.4 Constant Sum of Recruiting and Separation Elasticities

This section proves the result from Section 2.2 that the sum of recruiting and separation elasticities $\varepsilon_{Rw} - \varepsilon_{Sw} = \gamma$ for any choice of w in steady state if there are no exogenous separations and if a steady state equilibrium exists.

Suppose that an steady state equilibrium exists, characterized by tightness θ, the aggregate wage index $\tilde{w}$, distribution of posted wages $\Upsilon(w)$, distribution of employed wages $\Phi(w)$, and wage, employment, and vacancy policies w_j^*, N_j^*, and V_j^*. Suppose there are K wage levels. Given $w_1, ...w_k$ and $V_1, ..., V_K$, we want to solve for steady state employment shares $\phi_1, ..., \phi_K$. We also want to show that $\varepsilon_{Rw} - \varepsilon_{Sw} = \gamma$ in this steady state. Let us assume that:

$$\frac{\phi_i}{\phi_k} = \frac{\upsilon_i}{\upsilon_k}\left(\frac{w_i}{w_k}\right)^\gamma$$

for any two wage levels i and k. Then we need (1) to show that the employment share in any wage level is in a steady state, (2) solve for the level of the ϕ_k's, and (3) show that $\varepsilon_{Rw} - \varepsilon_{Sw} = \gamma$ for any firm j's wage policity w_j.

First let's show that inflows are equal to outflows for any sector i. Inflows to sector i are:

$$\begin{aligned}
&V_i g(\theta) \sum_{k \neq i} \phi_k \frac{w_i^\gamma}{w_i^\gamma + w_k^\gamma} \\
=&V_i \frac{g(\theta)}{f(\theta)} f(\theta) \sum_{k \neq i} \phi_i \frac{\upsilon_k}{\upsilon_i} w_k^\gamma \frac{1}{w_i^\gamma + w_k^\gamma} \\
=&V_i \frac{1}{\theta} f(\theta) \phi_i \frac{1}{\upsilon_i} \sum_{k \neq i} \upsilon_k \frac{w_k^\gamma}{w_i^\gamma + w_k^\gamma} \\
=&V_i \frac{S}{V} f(\theta) \phi_i \frac{V}{V_i} \sum_{k \neq i} \upsilon_k \frac{w_k^\gamma}{w_i^\gamma + w_k^\gamma} \\
=&\lambda N f(\theta) \phi_i \sum_{k \neq i} \upsilon_k \frac{w_k^\gamma}{w_i^\gamma + w_k^\gamma} \\
=&N_i \lambda f(\theta) \sum_{k \neq i} \upsilon_k \frac{w_k^\gamma}{w_i^\gamma + w_k^\gamma},
\end{aligned}$$

which is the formula for outflows from sector i (using $f(\theta)/g(\theta) = \theta$, $\theta = V/S$, $S = \lambda N$, $\phi_i = N_i/N$). Thus, when $\frac{\phi_i}{\phi_k} = \frac{\upsilon_i}{\upsilon_k}\left(\frac{w_i}{w_k}\right)^\gamma$, this labor market is in steady state.

Next we solve for the values of ϕ_i. First, define some constant C such that

$$\frac{\upsilon_i w_i^\gamma}{\phi_i} = \frac{\upsilon_k w_k^\gamma}{\phi_k} = C,$$

thus

$$\phi_k = \frac{\upsilon_k w_k^\gamma}{C}, \ \ \forall k.$$

We also know

$$\sum_{k=1}^{K} \phi_k = 1,$$

thus

$$\sum_{k=1}^{K} \frac{w_k^\gamma \upsilon_k}{C} = 1 \longrightarrow \sum_{k=1}^{K} w_k^\gamma \upsilon_k = C.$$

Thus

$$\phi_i = \frac{\upsilon_i w_i^\gamma}{C} = \frac{\upsilon_i w_i^\gamma}{\sum_{k=1}^{K} \upsilon_k w_k^\gamma}.$$

Lastly, we show that for any firm j, the firm faces a constant $\varepsilon_{Rw} - \varepsilon_{Sw} = \gamma$ at any value of w_j. The firm's ratio of sepration rate to recruiting rate is:

$$\begin{aligned}
\frac{S(w_j)}{R(w_j)} &= \frac{f(\theta)\lambda\Big(\upsilon_1 \frac{w_1^\gamma}{w_1^\gamma + w_j^\gamma} + ... + \upsilon_K \frac{w_K^\gamma}{w_K^\gamma + w_j^\gamma}\Big)}{g(\theta)\Big(\phi_1 \frac{w_j^\gamma}{w_1^\gamma + w_j^\gamma} + ... + \phi_K \frac{w_j^\gamma}{w_K^\gamma + w_j^\gamma}\Big)} \\
&= \lambda\theta \frac{\Big(\upsilon_1 \frac{w_1^\gamma}{w_1^\gamma + w_j^\gamma} + ... + \upsilon_K \frac{w_K^\gamma}{w_K^\gamma + w_j^\gamma}\Big)}{\Big(\phi_1 \frac{w_j^\gamma}{w_1^\gamma + w_j^\gamma} + ... + \phi_K \frac{w_j^\gamma}{w_K^\gamma + w_j^\gamma}\Big)}
\end{aligned}$$

Using that $\upsilon_i w_i^\gamma = \phi_i \sum_{k=1}^{K} w_k^\gamma \upsilon_k$, we can rewrite the above equation as:

$$\begin{aligned}
\frac{S(w_j)}{R(w_j)} &= \lambda\theta \frac{\sum_{k=1}^{K} \upsilon_k w_k^\gamma}{w_j^\gamma} \frac{\Big(\phi_1 \frac{1}{w_1^\gamma + w_j^\gamma} + ... + \phi_K \frac{1}{w_K^\gamma + w_j^\gamma}\Big)}{\Big(\phi_1 \frac{1}{w_1^\gamma + w_j^\gamma} + ... + \phi_K \frac{1}{w_K^\gamma + w_j^\gamma}\Big)} \\
&= \lambda\theta w_j^{-\gamma} \sum_{k=1}^{K} \upsilon_k w_k^\gamma = \lambda\theta \Big(\frac{\tilde{w}}{w_j}\Big)^\gamma,
\end{aligned}$$

with $\tilde{w} = (\sum_{k=1}^{K} \upsilon_k w_k^\gamma)^{\frac{1}{\gamma}}$, where the elasticity of this function with respect to w_j is $-\gamma$.

Therefore, when firms solve (4) subject to (3), (10) and (11), the firm's optimal choices follow (5), (7), and (6). Plugging in $\varepsilon_{Rw} - \varepsilon_{Sw} = \gamma$ and $S(w_j)/R(w_j) = \lambda\theta(\tilde{w}/w_j)^\gamma$ into (5) and (7) yield (30) and (31) from Appendix A.5.

Adding Unemployment Here we show that the previous result about a constant value of $\mathcal{E}$ holds even with unemployment.

Inflows into and outflows from unemployment must balance:

$$(1-U)\lambda_{EU}\sum_{k=1}^{K}\phi_k\frac{b^\gamma}{b^\gamma+w_k^\gamma}=U\lambda_{UE}f(\theta)\sum_{k=1}^{K}\upsilon_k\frac{w_k^\gamma}{w_k^\gamma+b^\gamma}.$$

Using that

$$\frac{\upsilon_i w_i^\gamma}{\phi_i}=\frac{\upsilon_k w_k^\gamma}{\phi_k}=C,$$

the prior equation becomes

$$(1-U)\lambda_{EU}b^\gamma\sum_{k=1}^{K}\phi_k\frac{1}{b^\gamma+w_k^\gamma}=U\lambda_{UE}f(\theta)\frac{\upsilon_i w_i^\gamma}{\phi_i}\sum_{k=1}^{K}\phi_k\frac{1}{w_k^\gamma+b^\gamma}.$$

Reusing that $\phi_i=\frac{\upsilon_i w_i^\gamma}{\sum_{k=1}^K \upsilon_k w_k^\gamma}$, the previous equation becomes

$$(1-U)\lambda_{EU}b^\gamma=U\lambda_{UE}f(\theta)\sum_{k=1}^{K}\upsilon_k w_k^\gamma$$

Solving for U yields:

$$U=\frac{\lambda_{EU}b^\gamma}{\lambda_{EU}b^\gamma+\lambda_{UE}f(\theta)\sum_{k=1}^K\upsilon_k w_k^\gamma}.$$

The ratio of separations to recruits is:

$$\frac{S(w_j)}{R(w_j)}=\frac{\lambda_{EU}\frac{b^\gamma}{b^\gamma+w_j^\gamma}+f(\theta)\lambda_{EE}\Big(\upsilon_1\frac{w_1^\gamma}{w_1^\gamma+w_j^\gamma}+...+\upsilon_K\frac{w_K^\gamma}{w_K^\gamma+w_j^\gamma}\Big)}{g(\theta)\Big(\Phi_U\frac{w_j^\gamma}{w_j^\gamma+b^\gamma}+\Phi_E\big(\phi_1\frac{w_j^\gamma}{w_1^\gamma+w_j^\gamma}+...+\phi_K\frac{w_j^\gamma}{w_K^\gamma+w_j^\gamma}\big)\Big)}$$

Define the share of searchers that are unemployed Φ_U as:

$$\Phi_U=\frac{\lambda_{UE}U}{\lambda_{UE}U+\lambda_{EE}(1-U)},$$

Plugging in our value for U, this expression becomes

$$\Phi_U=\frac{\lambda_{EU}b^\gamma}{\lambda_{EU}b^\gamma+\lambda_{EE}f(\theta)\tilde{w}^\gamma},$$

where $\tilde{w}^\gamma=\sum_{k=1}^K\upsilon_k w_k^\gamma$, and $\Phi_E=1-\Phi_U$. Plugging in these terms and following familiar algebra yields:

$$\frac{S(w_j)}{R(w_j)} = \frac{\lambda_{EU}\frac{b^\gamma}{b^\gamma+w_j^\gamma} + f(\theta)\lambda_{EE}\tilde{w}^\gamma\left(\phi_1\frac{1}{w_1^\gamma+w_j^\gamma} + ... + \phi_K\frac{1}{w_K^\gamma+w_j^\gamma}\right)}{g(\theta)\frac{w_j^\gamma}{\lambda_{EU}b^\gamma+\lambda_{EE}f(\theta)\tilde{w}^\gamma}\left(\lambda_{EU}b^\gamma\frac{1}{w_j^\gamma+b^\gamma} + f(\theta)\lambda_{EE}\tilde{w}^\gamma\left(\phi_1\frac{1}{w_1^\gamma+w_j^\gamma} + ... + \phi_K\frac{1}{w_K^\gamma+w_j^\gamma}\right)\right)}.$$

As before, large terms on the top and bottom cancel, yielding:

$$\frac{S(w_j)}{R(w_j)} = \frac{\lambda_{EU}b^\gamma + \lambda_{EE}f(\theta)\tilde{w}^\gamma}{g(\theta)}w_j^{-\gamma}.$$

Nesting the case above when $\lambda_{EU} = 0$ yields

$$\frac{S(w_j)}{R(w_j)} = \lambda_{EE}\theta\tilde{w}^\gamma w_j^{-\gamma}.$$

A.5 Characterization of Constant Elasticity of Equilibrium

We now characterize the equilibrium, focusing on the case of $\lambda_{EU} = 0$, $s = 0$, and $\rho \to \infty$. Given parameters α_j, A_j, χ_j, σ_j, and c_j, aggregate labor market tightness θ, and aggregate wage index $\tilde{w}$, a firm j's optimal wage is

$$w_j^* = \left(c_j\gamma(1+\chi_j)(\lambda\theta)^{1+\chi_j}\tilde{w}^{\gamma(1+\chi_j)}\left(\frac{\alpha_j A_j\gamma(1+\chi_j)}{1+\gamma(1+\chi_j)+\sigma_j\chi_j}\right)^{\frac{\sigma_j\chi_j}{1-\alpha_j}}\right)^{\frac{1-\alpha_j}{(1-\alpha_j)(1+\gamma(1+\chi_j))+\sigma_j\chi_j}}, \tag{30}$$

where $\tilde{w}$ is a vacancy-weighted index of the distribution of wages: $\tilde{w} = \left(\int_{w_k}\nu(w_k)w_k^\gamma dw_k\right)^{\frac{1}{\gamma}}$. Taking $\rho \to 0$ for simplicity, optimal employment at firm j is

$$N_j^* = \left(\frac{\alpha_j A_j}{w_j^*} \times \frac{\gamma(1+\chi_j)}{1+\gamma(1+\chi_j)+\sigma_j\chi_j}\right)^{\frac{1}{1-\alpha_j}}. \tag{31}$$

The optimal vacancy policy is simply $V_j^* = \bar{V}(\tilde{w}/w_j)^\gamma N_j^*$, and labor market tightness is $\theta = (\int_{k\in J} V_k dk)/\lambda$.

As we saw in Table 1, the share of marginal product retained by the firm as profits and the inverse labor supply elasticity have very similar formulas. Now under this microfoundation, we can show that they are explicitly linked. The formula for the inverse labor supply elasticity is now

$$\varepsilon_{w,N}^j = \frac{\sigma_j\chi_j}{1+\gamma(1+\chi_j)}. \tag{32}$$

This is the ratio of the comparative static of wages with respect to firm TFP A_j over employment with respect to A_j: $\varepsilon_{w,N} = \frac{\partial\log w_j/\partial\log A_j}{\partial\log N_j/\partial\log A_j}$. As before, if $\sigma = 0$, then firm size will have no effect on wages for firms at their steady state size. This expression is the same as in Table 1, except that the superelasticity term drops out because the sum of the recruiting and negative separation elasticities $\mathcal{E}$ is constant and equal to γ. The profit share of marginal product is then

$$\frac{\text{Profit}_j}{MRPL_j} = \frac{\varepsilon^j_{w,N}}{1+\varepsilon^j_{w,N}}. \tag{33}$$

The profit share of marginal product that the firm earns will be very close to, and increasing in, the inverse labor supply elasticity. We will next estimate the elasticity of wages to firm size in Section 3.

A.6 Proving Uniqueness for Limited Firm Heterogeneity

In this appendix section, we show that all the endogenous outcomes θ, $\tilde{w}$, N_j, w_j, and V_j can be solved for as functions of parameters. Using the result from Appendix A.4 that in any equilibrium firms with have a constant value of $\mathcal{E}$, and as such optimal wages and employment will take the form of (30) and (31), we then confirm that if an steady-state equilibrium exists, it is then unique. We will again restrict our attention to the $\lambda_{EU} = 0$ and $\rho \to 0$ case.

Proposition 3 *If there are no exogenous separations $s = 0$, γ is positive and finite, firms have identical α, χ, and σ, and if a steady-state equilibrium exists, then the equilibrium is unique.*

Given labor market tightness θ and an aggregate wage index $\tilde{w}$, a firm's optimal wage in steady state is:

$$w^*_j = \left(c_j\gamma(1+\chi)(\lambda\theta)^{1+\chi}\tilde{w}^{\gamma(1+\chi)}\right)^{\frac{1-\alpha}{(1-\alpha)(1+\gamma(1+\chi))+\sigma\chi}} \left(\frac{\alpha A_j\gamma(1+\chi)}{1+\gamma(1+\chi)+\sigma\chi}\right)^{\frac{\sigma\chi}{(1-\alpha)(1+\gamma(1+\chi))+\sigma\chi}} \tag{34}$$

Taking $\rho \to 0$ for simplicity, optimal employment at firm j is:

$$N^*_j = \left(c_j\gamma(1+\chi)(\lambda\theta)^{1+\chi}\tilde{w}^{\gamma(1+\chi)}\right)^{\frac{-1}{(1-\alpha)(1+\gamma(1+\chi))+\sigma\chi}} \left(\alpha A_j\frac{\gamma(1+\chi)}{(1+\gamma(1+\chi))+\sigma\chi}\right)^{\frac{1+\gamma(1+\chi)}{(1-\alpha)(1+\gamma(1+\chi))+\sigma\chi}} \tag{35}$$

To have an interior solution, we must have that $\sigma\chi < (1-\alpha)(1+\gamma(1+\chi))$, i.e., σ cannot be too negative. This is ensured by restricting σ to $\sigma \in [0, 1]$. In steady state, optimal vacancies are:

$$V_j = \lambda_{EE}\theta\left(\frac{\tilde{w}}{w_j}\right)^{\gamma}N_j.$$

Lemma 4 *Vacancy shares υ_j are a function of only parameters and do not depend on aggregate wages $\tilde{w}$ or tightness θ.*

First we want to show that relative wages are not a function of $\tilde{w}$ or θ. For two firms j and k, relative wages are:

$$\frac{w_j}{w_k} = \left(\frac{c_j}{c_k}\right)^{\frac{1-\alpha}{(1-\alpha)(1+\gamma(1+\chi))+\sigma\chi}} \left(\frac{A_j}{A_k}\right)^{\frac{\sigma\chi}{(1-\alpha)(1+\gamma(1+\chi))+\sigma\chi}}$$

Next we show that relative employment is not a function of $\tilde{w}$ or θ. For two firms j and k, relative employment is:

$$\frac{N_j}{N_k} = \left(\frac{c_j}{c_k}\right)^{\frac{-1}{(1-\alpha)(1+\gamma(1+\chi))+\sigma\chi}} \left(\frac{A_j}{A_k}\right)^{\frac{1+\gamma(1+\chi)}{(1-\alpha)(1+\gamma(1+\chi))+\sigma\chi}}$$

Thus relative vacancies are a function of parameters only:

$$\frac{V_j}{V_k} = \frac{N_j}{N_k}\left(\frac{w_j}{w_k}\right)^{-\gamma} = \left(\frac{c_j}{c_k}\right)^{\frac{-1-\gamma(1+\alpha)}{(1-\alpha)(1+\gamma(1+\chi))+\sigma\chi}} \left(\frac{A_j}{A_k}\right)^{\frac{1+\gamma(1+\chi)-\gamma\sigma\chi}{(1-\alpha)(1+\gamma(1+\chi))+\sigma\chi}}.$$

Let ω_j be the mass of firms with parameters c_j and A_j, with $\int_A \int_c \omega_j dc dA = 1$. The share of vacancies of firm type j are

$$\upsilon_j = \frac{\omega_j V_j}{\int_{k\in J} \omega_k V_k} = \frac{1}{\int_{k\in J} \frac{\omega_k}{\omega_j}\frac{V_k}{V_j}}.$$

Since V_j/V_k is a function of only parameters, and the share of each time of firm ω_j is a parameter, then the vacancy share υ_j is a function of only parameters.

Lemma 5 *The ratio of a firm's optimal wage w_j to the index of aggregate wages $\tilde{w}$ is a function of parameters.*

$$\frac{\tilde{w}}{w_j} = \frac{\int_k \upsilon_k w_k dk}{w_j} = \frac{\int_k \upsilon_k \left(c_k\gamma(1+\chi)(\lambda\theta)^{1+\chi}\tilde{w}^{\gamma(1+\chi)}\right)^{\frac{1-\alpha}{(1-\alpha)(1+\gamma(1+\chi))+\sigma\chi}} \left(\frac{\alpha A_k\gamma(1+\chi)}{1+\gamma(1+\chi)+\sigma\chi}\right)^{\frac{\sigma\chi}{(1-\alpha)(1+\gamma(1+\chi))+\sigma\chi}} dk}{\left(c_j\gamma(1+\chi)(\lambda\theta)^{1+\chi}\tilde{w}^{\gamma(1+\chi)}\right)^{\frac{1-\alpha}{(1-\alpha)(1+\gamma(1+\chi))+\sigma\chi}} \left(\frac{\alpha A_j\gamma(1+\chi)}{1+\gamma(1+\chi)+\sigma\chi}\right)^{\frac{\sigma\chi}{(1-\alpha)(1+\gamma(1+\chi))+\sigma\chi}}}.$$

All γ, χ, λ, θ, and σ terms can be factored outside the intregral and cancel on the top and bottom. Therefore, a firm's wage w_j relative to the market index $\tilde{w}$ is:

$$\frac{\tilde{w}}{w_j} = \frac{\int_k \upsilon_k w_k dk}{w_j} = \frac{\int_k \upsilon_k c_k^{\frac{1-\alpha}{(1-\alpha)(1+\gamma(1+\chi))+\sigma\chi}} A_k^{\frac{\sigma\chi}{(1-\alpha)(1+\gamma(1+\chi))+\sigma\chi}} dk}{c_j^{\frac{1-\alpha}{(1-\alpha)(1+\gamma(1+\chi))+\sigma\chi}} A_j^{\frac{\sigma\chi}{(1-\alpha)(1+\gamma(1+\chi))+\sigma\chi}}}.$$

Because the vacancy shares υ_k are also a function of parameters, a firms wage w_j relative to the market $\tilde{w}$ is a function of only parameters and not functions of labor market tightness θ.

Thus, we can write relative wages as a function of parameters only:

$$\frac{\tilde{w}}{w_j} = \frac{\tilde{w}}{w_j}\Big(\mathbb{c}, \mathbb{A}\Big).$$

Lemma 6 *Given distributions of parameters $\mathbb{A}$ and $\mathbb{c}$, the wage index $\tilde{w}$ is an increasing function of tightness θ.*

Based on equation (34), individual firm wages w_j are increasing in aggregate wages $\tilde{w}$ and tightness θ. Solving out individual wages, and solving $\tilde{w}$ in terms of only θ and parameters yields:

$$\tilde{w} = \Big(\int_j \upsilon_j w_j^{\gamma} dj\Big)^{\frac{1}{\gamma}}$$

$$\tilde{w} = \Big(\int_j \upsilon_j \Big(\theta^{1+\chi}\tilde{w}^{\gamma(1+\chi)}\Big)^{\frac{\gamma(1-\alpha)}{(1-\alpha)(1+\gamma(1+\chi))+\sigma\chi}} g(c_j, A_j) dj\Big)^{\frac{1}{\gamma}}$$

$$\tilde{w} = \Big(\theta^{1+\chi}\tilde{w}^{\gamma(1+\chi)}\Big)^{\frac{(1-\alpha)}{(1-\alpha)(1+\gamma(1+\chi))+\sigma\chi}} \Big(\int_j \upsilon_j g(c_j, A_j) dj\Big)^{\frac{1}{\gamma}}, \tag{36}$$

with

$$g(c_j, A_j) = \Big(c_j\gamma(1+\chi)\lambda^{1+\chi}\Big)^{\frac{1-\alpha}{(1-\alpha)(1+\gamma(1+\chi))+\sigma\chi}} \Big(\frac{\alpha A_j \gamma(1+\chi)}{1+\gamma(1+\chi)+\sigma\chi}\Big)^{\frac{\sigma\chi}{(1-\alpha)(1+\gamma(1+\chi))+\sigma\chi}}.$$

Grouping all $\tilde{w}$ terms on the left hand side yields:

$$\tilde{w}^{\frac{(1-\alpha)+\sigma\chi}{(1-\alpha)(1+\gamma(1+\chi))+\sigma\chi}} = \theta^{\frac{(1+\chi)(1-\alpha)}{(1-\alpha)(1+\gamma(1+\chi))+\sigma\chi}} \Big(\int_j \upsilon_j g(c_j, A_j) dj\Big)^{\frac{1}{\gamma}}. \tag{37}$$

Simplifying yields:

$$\tilde{w} = \theta^{\frac{(1+\chi)(1-\alpha)}{(1-\alpha)+\sigma\chi}} \Big(\int_j \upsilon_j g(c_j, A_j) dj\Big)^{\frac{(1-\alpha)(1+\gamma(1+\chi))+\sigma\chi}{\gamma((1-\alpha)+\sigma\chi)}}.$$

Proof of Proposition Let $\omega(A,c)$ be the joint density function of the mass of firm populations over parameters c and A. Then the mass aggregate vacancies V is

$$V = \int_A \int_c V_j \omega(A,c) dc dA$$

$$V = \int_A \int_c \lambda_{EE}\theta\Big(\frac{\tilde{w}}{w_j}\Big)^{\gamma} N_j \omega(A,c) dc dA$$

We showed already that $\tilde{w}/w_j$ is a function of parameters:

$$V = \int_A \int_c \lambda_{EE}\theta\Big(\frac{\tilde{w}}{w_j}(\mathbb{c}, \mathbb{A})\Big)^{\gamma} N_j \omega(A,c) dc dA \tag{38}$$

We can abbreviate the expression for optimal employment as

$$N_j = \Bigg(\theta^{1+\chi}\tilde{w}^{\gamma(1+\chi)}\Bigg)^{\frac{-1}{(1-\alpha)(1+\gamma(1+\chi))+\sigma\chi}} f(c_j, A_j), \tag{39}$$

with

$$f(c_j, A_j) = \Big(c_j\gamma(1+\chi)\lambda^{1+\chi}\Big)^{\frac{-1}{(1-\alpha)(1+\gamma(1+\chi))+\sigma\chi}} \Big(\frac{\alpha A_j \gamma(1+\chi)}{1+\gamma(1+\chi)+\sigma\chi}\Big)^{\frac{1+\gamma(1+\chi)}{(1-\alpha)(1+\gamma(1+\chi))+\sigma\chi}}.$$

Plugging in (37) and (39) into equation (38) yields:

$$V = \int\int \theta \times \theta^{\frac{-1-\chi}{(1-\alpha)(1+\gamma(1+\chi))+\sigma\chi}} \times \theta^{\frac{(1+\chi)(1-\alpha)}{(1-\alpha)+\sigma\chi}\frac{-\gamma(1+\chi)(1-\alpha)}{(1-\alpha)(1+\gamma(1+\chi))+\sigma\chi}} H(A_j, c_j, \mathbb{A}, \mathbb{c}) dcdA,$$

with

$$H(A_j, c_j, \mathbb{A}, \mathbb{c}) = \Big(\frac{\tilde{w}}{w_j}(\mathbb{c}, \mathbb{A})\Big)^{\gamma} f(c_j, A_j)\Big(\int_j \upsilon_j g(c_j, A_j) dj\Big)^{\frac{-1-\chi}{(1-\alpha)+\sigma\chi}}$$

In equilibrium where all workers are employed, and employed workers search with probability λ, then (normalize the population of workers to 1) tightness θ is:

$$\theta = \frac{V}{\lambda}.$$

The prior equation then becomes:

$$\lambda\theta = \int\int \theta \times \theta^{\frac{-1-\chi}{(1-\alpha)(1+\gamma(1+\chi))+\sigma\chi}} \times \theta^{\frac{(1+\chi)(1-\alpha)}{(1-\alpha)+\sigma\chi}\frac{-\gamma(1+\chi)(1-\alpha)}{(1-\alpha)(1+\gamma(1+\chi))+\sigma\chi}} H(A_j, c_j, \mathbb{A}, \mathbb{c}) dcdA.$$

The θ terms on the right hand side can be pulled out of the integral, and a θ term on both sides cancels. With some algebra, we have

$$\theta = \Big(\frac{1}{\lambda}\int\int H(A_j, c_j, \mathbb{A}, \mathbb{c}) dcdA\Big)^{\frac{(1-\alpha)+\sigma\chi}{1+\chi}}.$$

Thus equilibrium labor market tightness can be calculated strictly as a function of parameters. Plugging in our value of θ in equation (37), we can solve for the aggregate wage $\tilde{w}$ as a function of parameters. Given $\tilde{w}$, we can compute individual firms' N_j, w_j, and V_j.

A.7 Two Period Model with Fixed Wage Contracts

In this section, we derive a two period version of our main model from Section 2, with periods 0 and 1. We demonstrate that if $\sigma = 0$, then the wages of new hires in period 1 are a function only of the vacancy rate in period 1. However, the wage in period 1 of incumbents (workers already employed in period 0) and the wage in period 1 of of workers hired in period 0 depend on the firm's hiring rate in period 0. Therefore when $\sigma = 0$, a sequence of shocks that would lead to growth in period 0, but no growth in period 1, would lead to elevated wages in period 1, but only for incumbents and workers hired in period 0, but not for workers hired in period 1. This would generate a relationship between incumbent wages and firm size even though $\sigma = 0$.

Let the mass of new hires in period t be J_t and the mass of incumbent workers be N_t. The firm inherits N_0 incumbents and is obligated to pay these incumbents w_0^N. The firm posts V_0 vacancies in period 0, and successfully recruits J_0 new hires. New hires are promised w_0 in both periods, and incumbents are offered w_0 in period 1. Incumbents and new hires from period 1 are retained into the next period with probability $\rho(w_0)$, so next period incumbent is $N_1 = \rho(w_0)(N_0 + J_0) =$

$\rho(w_0)(N_0 + R(w_0)V_0)$. Firms recruit in period 1, where the new hires in period 1 J_1 is the share $R(w_1^J)$ of vacancies V_1 that are successfully filled.

$$J_0 = R(w_0)V_0 \tag{40}$$
$$N_1 = \rho(w_0)(N_0 + J_0) \tag{41}$$
$$J_1 = R(w_1^J)V_1 \tag{42}$$

The firm produces a decreasing returns to scale function in the number of workers. The firm solves

$$\max_{J_0,J_1,N_1,w_0,w_1,V_0,V_1} \mathcal{L} : A_0 \left(N_0 + J_0\right)^{\zeta} - w_0^N N_0 - w_0 J_0 - c\left(\frac{V_0}{N_0^{1-\sigma}}\right)^{\chi} V_0 + A_1 \left(N_1 + J_1\right)^{\zeta} - w_0 N_1 - w_1^J J_1 - c\left(\frac{V_1}{(N_0+J_0)^{1-\sigma}}\right)^{\chi} V_1$$

subject to (40), (41), and (42). The optimal wage for new hires in period 1 w_1^J satisfies

$$w_1^J = c(1+\chi)\varepsilon_{R,w_1^J} \frac{1}{R(w_1^J)^{1+\chi}} \left(\frac{J_1}{N_0+J_0}\right)^{\chi} (N_0+J_0)^{\sigma\chi},$$

which is the 2-period analogue to equation (5) in the main text. The key part to notice is that the wage of new hires in period 1 w_1^J depends positively on the hiring rate $J_1/(N_0 + J_0)$ if $\chi > 0$ and prior firm size $N_0 + J_0$ if $\sigma > 0$. This is the two-period analogue to the result in equation (14) that wages a function of only the firm growth rate if $\sigma = 0$, but wages can depend on the level of employment if $\sigma > 0$.

For incumbents and workers who were hired in period 0, the optimal wage satisfies

$$w_0 = \frac{\frac{c(1+\chi)\left(\frac{J_0}{N_0}\right)^{\chi} N_0^{\sigma\chi}}{R(w_0^J)^{1+\chi}} \varepsilon_{R,w_0} \frac{J_0}{J_0+N_1} + \left(1+\varepsilon_{R,w_1^J}\right) \frac{c(1+\chi)\left(\frac{J_1}{J_0+N_0}\right)^{\chi} (J_0+N_0)^{\sigma\chi}}{R(w_1^J)^{1+\chi}} \varepsilon_{\rho,w_0} \frac{N_1}{J_0+N_1}}{\left(1+\varepsilon_{\rho,w_0}\frac{N_1}{J_0+N_1}\right)}.$$

While this expression is more complicated, it shows how the presence of fix-wage contracts can bias inference of σ when focusing on the wages of incumbents. The optimal wage w_0 is a function of hiring rates J_0/N_0 and $J_1/(J_0 + N_0)$ if $\chi > 0$, as well as the level of employment at the beginning of each period N_0 and $N_0 + J_0$ if $\sigma > 0$. Even if σ were truly zero, and if reseacher ran a regression of the form of (15) from the main text using the wages of incumbent workers or workers hired in period 0 (i.e. the year of the shock), the researcher would find a positive coefficient on β_1 and infer that $\sigma > 0$. Since σ is the determinant in our model of the firm's long-run labor supply elasticity, a researcher estimating labor supply elasticities using the long-run effect of shocks on the wages of incumbents or a continuous cohort of new hires would incorrectly find that the labor supply curve is not perfectly elastic.

A.8 Extension with Worker Heterogeneity

In this section, we show that we can extend the model in Section 2.2 where workers differ in ability and firms pay wages per efficiency unit of labor. We show that the law of motion for the firm is identical up to a constant as in Section 2.2. Wages are therefore log-additive in worker types and firm pay policies, motivating the AKM specification we use is Section 3.

Suppose that workers have immutable ability a_i. Let the number of workers at a firm be $N_j = \int_i m_j(a_i)di$, and the total about of labor be $L_j = \int_i m_j(a_i)a_i di$, where $m_j(a_i)$ is the density of ability of workers employed at firm j. Let $h(a_i)$ be the probability density of searchers of ability a_i, so $\int_i h(a_i)di = 1$.

Firms post vacancies as usual, and firms and workers engage in random search. Firms choose efficiency wage policies ω_j, so the amount paid to a worker of type a_i is $w_{ij} = a_i\omega_j$. Firms cannot direct search towards workers of different abilities. If a worker employed at firm j earning wage rate ω_j meets a firm k offering ω_k, the probability that the worker will leave is

$$s(\omega_j, \omega_k, a_i) = \frac{(a_i\omega_k)^\gamma}{(a_i\omega_j)^\gamma + (a_i\omega_k)^\gamma} = \frac{\omega_k^\gamma}{\omega_j^\gamma + \omega_k^\gamma}.$$

The separation rate $\mathbf{S}$ is the share of effective labor that leaves the firm each period.

$$\mathbf{S}(\omega_j) = \frac{\int_i \int_k m_j(a_i) f(\theta) \frac{(a_i\omega_k)^\gamma}{(a_i\omega_j)^\gamma + (a_i\omega_k)^\gamma} \upsilon(\omega_k) dk di}{\int_i m_j(a_i)di}$$

$$\mathbf{S}(\omega_j) = \frac{1}{N_j} \int_i \int_k m_j(a_i) f(\theta) \frac{\omega_k^\gamma}{\omega_j^\gamma + \omega_k^\gamma} \upsilon(\omega_k) dk di$$

$$\mathbf{S}(\omega_j) = \frac{1}{N_j} f(\theta) \int_k \frac{\omega_k^\gamma}{\omega_j^\gamma + \omega_k^\gamma} \upsilon(\omega_k) dk \int_i m_j(a_i) di$$

$$\mathbf{S}(\omega_j) = f(\theta) \int_k \frac{\omega_k^\gamma}{\omega_j^\gamma + \omega_k^\gamma} \upsilon(\omega_k) dk.$$

Thus the separation rate $\mathbf{S}$ is just a function of the firm's wage policy. The rate at which the firm acquires effective labor is

$$\mathbf{H}(V_j, \omega_j) = V_j g(\theta) \int_i \int_k \frac{(a_i\omega_j)^\gamma}{(a_i\omega_j)^\gamma + (a_i\omega_k)^\gamma} \phi(\omega_k, a_i) dk di$$

where $\phi(\omega_k, a_i)$ is the joint distribution of wage policies and abilities of searchers. Let the recruiting rate $\mathbf{R}(\omega_j)$ be the rate at which vacancies are converted into labor.

$$\mathbf{R}(\omega_j) = \frac{H(V_j, \omega_j)}{V_j} = g(\theta) \int_i \int_k \frac{(a_i\omega_j)^\gamma}{(a_i\omega_j)^\gamma + (a_i\omega_k)^\gamma} \phi(\omega_k, a_i) dk di$$

Suppose a candidate equilibrium where the distribution of wage policies is the same across worker types i. Then we can rewrite the recruiting rate as

$$\begin{aligned}
\mathbf{R}(\omega_j) &= g(\theta)\int_i\int_k \frac{(a_i\omega_j)^\gamma}{(a_i\omega_j)^\gamma+(a_i\omega_k)^\gamma}\phi(\omega_k)h(a_i)dkdi \\
\mathbf{R}(\omega_j) &= g(\theta)\int_k \frac{(a_i\omega_j)^\gamma}{(a_i\omega_j)^\gamma+(a_i\omega_k)^\gamma}\phi(\omega_k)dk\int_i h(a_i)di \\
\mathbf{R}(\omega_j) &= g(\theta)\int_k \frac{(a_i\omega_j)^\gamma}{(a_i\omega_j)^\gamma+(a_i\omega_k)^\gamma}\phi(\omega_k)dk.
\end{aligned}$$

The last thing to show is that in equilibrium, the distribution of employed wages $\phi(\omega_k)$ is constant for all worker types. Since the on-the-job search parameter for each worker type is identical, and since the flow rates across firm wage levels are the same, then the distribution of wage levels in equilibrium for each ability type a will be identical. As a result, the density of employment $m_j(a_i)/N_j = h(a_i)\ \forall i$ and j.

As a result, the law of motion of effective labor at the firm follows the same form as when there is no worker heterogeneity.

$$L_{j,t} = (1-\mathbf{S}(\omega_{j,t}))L_{j,t-1} + \mathbf{R}(\omega_{j,t})V_t.$$

Since all firms will have the same distribution of worker types a_i, so $m_j(a_i) = m(a_i)\forall j$, we can simply take the average $L_j = \int_i m_j(a_i)a_i di = \int_i m(a_i)a_i di = \frac{\int_i m(a_i)a_i di}{N_j}N_j = \bar{a}N_j$. Since $\mathbf{S}$ is the share of labor that is lost to turnover each period, and since worker quality will be the same at each firm in every period, we simply have that $\mathbf{S} = S$. As for recruiting, $\mathbf{R}$ is rate of converting vacancies into effective labor, while R is the rate of converting vacancies into the number of workers. Since the average quality of workers will be the same in each period, then $\mathbf{R} = R\bar{a}$. With some algebra, we have

$$\begin{aligned}
L_{j,t} &= (1-\mathbf{S}(\omega_{j,t}))L_{j,t-1} + \mathbf{R}(\omega_{j,t})V_t \\
N_{j,t}\bar{a} &= (1-S(\omega_{j,t}))N_{j,t-1}\bar{a} + R(\omega_{j,t})\bar{a}V_t \\
N_{j,t} &= (1-S(\omega_{j,t}))N_{j,t-1} + R(\omega_{j,t})V_t,
\end{aligned}$$

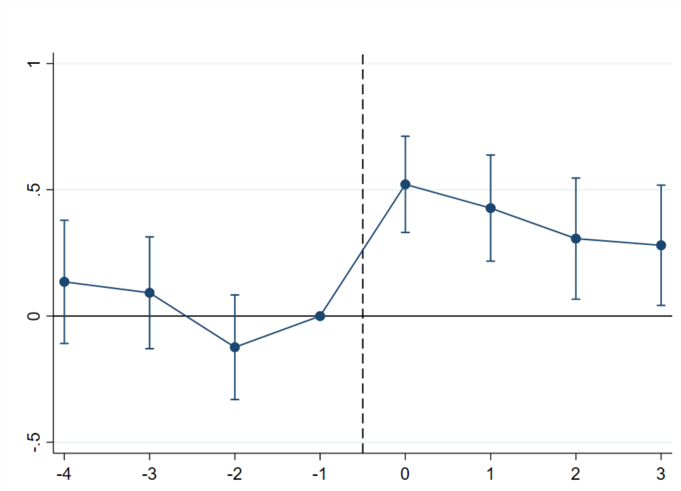
which is the same law of motion as in the original text except in wage policies ω_j rather than wage levels w_j.

Because the payments to workers are $w_{ij} = a_i\omega_j$, we can simply write log wages as $\log w_{ij} = \log a_i + \log\omega_j$, so the worker effect $\zeta_i = \log a_i$, and the firm effect $\psi_j = \log\omega_j$. For workers who switch, the change in wages $\log w_{ij} - \log_{ik} = \zeta_i + \omega_j - (\zeta_i + \omega_k) = \omega_j - \omega_k$.

A.9 Additional Empirical Results

Sales Figure 5 reports a local projection of the effect of an export demand shock on log sales, of a similar specification as equation (16), controlling for industry by year fixed effects, and industry by year fixed effects interacted with lagged export shares. A 100 log point increase in predicted sales increases actual sales by only 52 log points. The effect partially reverse, with sales peristently around 30 log points higher than prior to the shock.

Figure 5: Response of Log Sales to Export Demand Shock



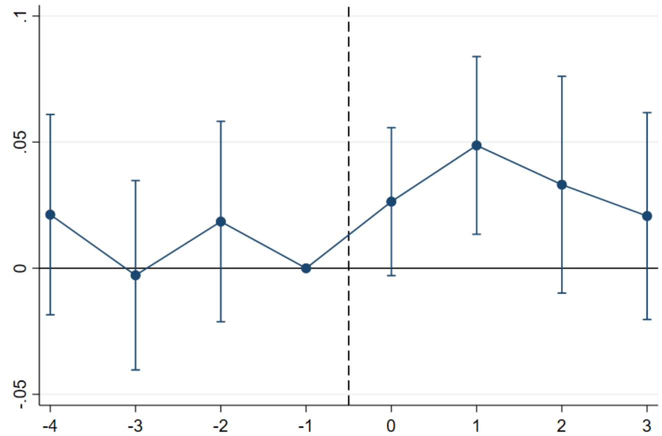
This figure plots the effect of 100 log point predicted increase in sales on a firm's log sales. A predicted 100 log point export demand shock increases sales by 50 log points on impact. The effect partially fades over time, and sales are thirty log points higher three years after the shock.

Stayers In addition to estimating the path of wages for switchers around export demand shocks, we also estimate the wage growth of job stayers. We define a stayer as a worker who was a full time worker ($\geq$ 1400 hours in a year) firm in years $t-2$, $t-1$, and t. We estimate a local projection of the change in workers' wages relative to year $t-1$, with the firm export demand shock in year t as the independent variable. We regress

$$\Delta \log w_{i,t-1,t+\nu} = \alpha + \sum_{\nu=-4}^{0} \beta_s z_{t+s} + \mathbf{X}_{i,t-2} B + e_{ijkt},$$

for $\nu \in [-4, 3]$ also controlling for industry by year fixed effects, as well as industry by year fixed effects interacted with $t-1$ export shares.

Figure 6: Wage Response of Stayers to Export Demand Shock

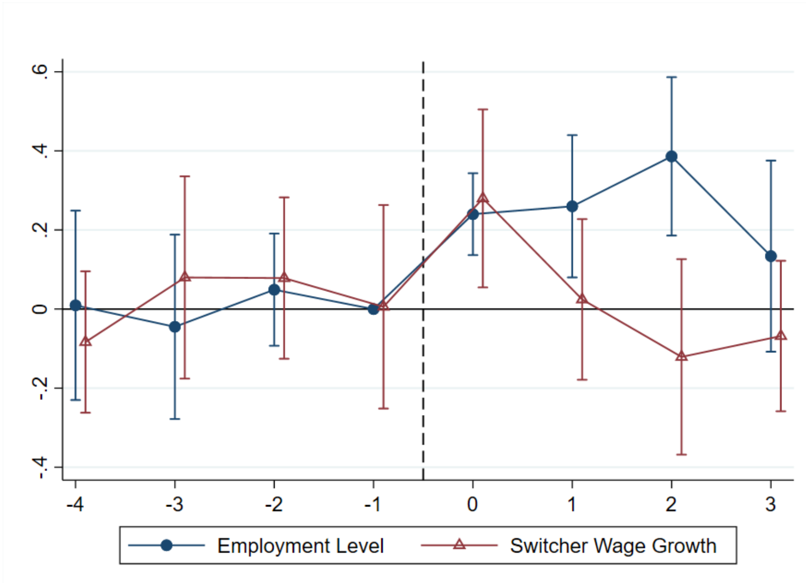


This figure plots the effect of a period t export demand shock, standardized to be a 100 log point predicted increase in sales, on the wage of workers who stay continuous at the firm through years $t-4$ to $t+3$.

Figure 6 shows the results. On impact, the point estimate shows that for a firm with a 100 log point predicted increase in sales, stayers wages rise by a statistically insignificant 2.5 log points in the first year, and a statistically significant 5 log points in the following year. The effect on wages appears to recede in the following years, but we cannot reject that the wage effect on stayers is persistent.

Switchers at Small Firms In the main text, we report the path of log employment and the growth of log wages of switchers in response to idiosyncratic firm export demand shocks. One concern is that the increasing in hiring by firms experiencing an export demand shock may have effects on the overall labor market, and so SUTVA would be violated. To avoid this concern, we focus on only small firms, defined as the log of the average employment in the sample window is below 4, which translates to an average firm size below 54 full time equivalent workers. This reduces the number of firms in the sample by about 43% We run the same regression as (16) and (17), but on this smaller sample of smaller firms. Figure 7 shows the results. On impact in response to a 100 log point predicted increase in sales, employment rises by a little over 20 log points, and switchers see 20 percent higher wages growth than workers who switch into firms with a zero shock. The effect on the employment level is persistent, which the extra wage growth received by job switchers who switch into the firm in $t=1$ and later into shock firms goes back to zero.

Figure 7: Response of Firm Employment Level and Switcher Wage Growth to Export Demand Shocks



The figure plots the path of the employment level and wage growth of switchers in response to a 100 log point predicted increase in sales due to an export demand shock in period t, but including only firms who have an average log firm size of less than 4 (i.e., full time equivalent workers below 54 FTE). This is a similar figure as is plotted in Figure 3 in the main text. The blue circles plot show that firms increase employment by approximately 23 log points, and the effect is persistent. The hollow red triangle show the effect of the period t trade shock on the wage growth of different cohorts of arriving switchers. Workers who switch into the firm in period t experience wage growth that is approximately 26 log points greater than workers who switch into a firm that is not shocked. However, workers who switch into the shocked firm in subsequent years receive no greater wage increases that workers who arrived prior to the shock.

B Supplementary Appendix

The purpose of this appendix is to show that if workers are forward looking, then the core of the main results in the paper still hold. The challenge if workers are forward looking is that their actions will depend on their beliefs about wages at a firm in the future. While there is an extensive literature on optimal contracting (and in particular, that firms will backload wages to encourage retention), we will focus on a simpler contracting environment where firms post fixed wage contracts, but the wages they offer to workers hired in different years may differ. We will also assume a stationary environment, so there are no aggregate shocks.

B.1 Forward Looking Workers

Workers get flow utility $\ln(w_j)$ and discount the future at rate β. Workers can search on the job with probability λ_{EE} and find an outside job with probabiliity $\lambda_{EE} f(\theta)$. If workers are matched with a job, they choose the job that provides the highest present discounted value of utility. A workers's value function is

$$V_j = \ln(w_j) + \beta \left(f(\theta)\lambda_{EE} V^{choose} + (1 - f(\theta)\lambda_{EE}) V_j' \right) \tag{43}$$

When workers match with an outside offer, they draw a preference shock for the new job, and they redraw a preference shock for their current job. Draws are i.i.d. across firms and across time and are Type-1 extreme value with scale parameter $\sigma = \frac{1}{\gamma(1-\beta)}$. So, given β, a high γ, which means small variance of taste shocks, means that a small σ also means a low variance of taste shocks.

The probability of choosing firm j over firm k is

$$\frac{\exp\left(\frac{V_j}{\sigma}\right)}{\exp\left(\frac{V_j}{\sigma}\right) + \exp\left(\frac{V_k}{\sigma}\right)}$$

The value of getting two draws and taking the best of them is

$$V^{choose}(w_j, w_k) = \sigma \log \left(\exp\left(\frac{V_j}{\sigma}\right) + \exp\left(\frac{V_k}{\sigma}\right) \right)$$

Value in a Single Wage Equilibrium Suppose we're in a single wage equilibrium, so $w_j = \bar{w}$ $\forall j$. Consider the value function of a worker at a firm paying $\bar{w}$.

$$\begin{aligned} V^{choose}(\bar{w}, \bar{w}) =& \sigma \log \left(2 \exp\left(\frac{\bar{V}}{\sigma}\right) \right) \\ =& \sigma \log(2) + \sigma \log(\exp(\bar{V}/\sigma)) \\ =& \sigma \log(2) + \bar{V} \end{aligned}$$

As $\sigma \to 0$ (the variance of taste shocks shrinks to zero), then $V^{choose}(\bar{w}, \bar{w}) \to \bar{V}$.

Plugging our expression for V^{choose} in to our value function in equation (43), we have

$$\bar{V} = \ln(\bar{w}) + \beta\left(f(\theta)\lambda_{EE}(\sigma\log(2) + \bar{V}') + (1 - f(\theta)\lambda_{EE})\bar{V}'\right) \tag{44}$$

$$\bar{V} = \ln(\bar{w}) + \beta\left(\bar{V} + f(\theta)\lambda_{EE}\sigma\log(2)\right) \tag{45}$$

$$(1-\beta)\bar{V} = \ln(\bar{w}) + \beta f(\theta)\lambda_{EE}\sigma\log(2). \tag{46}$$

$$\bar{V} = \frac{\ln(\bar{w}) + \beta f(\theta)\lambda_{EE}\sigma\log(2)}{1-\beta}. \tag{47}$$

$$\tag{48}$$

Rescaling our value function by multiplying both sides by $1-\beta$, we have

$$\tilde{V}(\bar{w}) = \log(\bar{w}) + \beta f(\theta)\lambda_{EE}\sigma\log(2)$$

Value of paying w_j, All other firms pay $\bar{w}$ Now we solve for the value to a worker employed at a firm paying w_j indefinitely. We assume this firm is atomistic, so if the worker leaves, the worker will never find this firm again. Since all other firms pay $\bar{w}$, a worker who leaves firm j will earn $\bar{w}$ for the rest of their life.

$$V^{choose}(w_j, \bar{w}_k) = \sigma\log\left(\exp\left(\frac{V_j}{\sigma}\right) + \exp\left(\frac{\bar{V}}{\sigma}\right)\right)$$

$$V_j = \log(w_j) + \beta\left(f(\theta)\lambda_{EE}\sigma\log\left(\exp\left(\frac{V_j}{\sigma}\right) + \exp\left(\frac{\bar{V}}{\sigma}\right)\right) + (1 - f(\theta)\lambda_{EE})V_j'\right)$$

This value is computed using value function interation.

Figure 8: Probability of Staying at firm Paying w_j by discount rate β

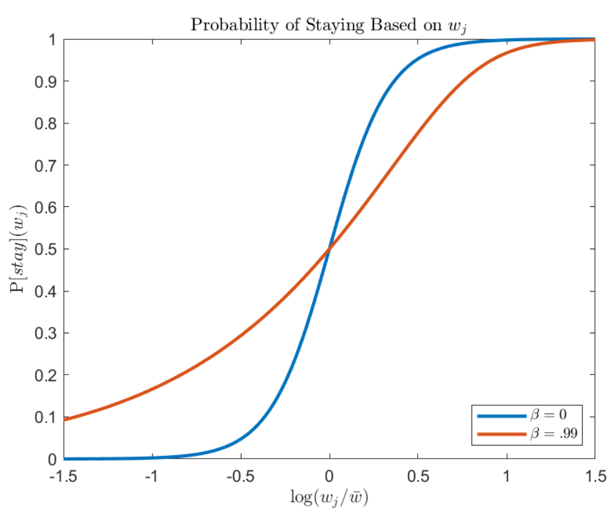


Figure 8 plots the probability that a workers stays if firm j given w_j, conditional on having matched with another firm paying $\bar{w}$. We can see that like in the myopic case, the probability of staying in a job increases with a firm's wage policy. Interestingly, the probability of staying in a job paying w_j, conditional on being offered a wage of $\bar{w}$, is increasing less in the wage. If a worker in a firm paying a lower wage than average, they can always quit at a later period, so having lower wages today has less effect on their utility. This can be seen in that the semielasticity of value functions with respect to wages is convex. At lower wages, a further lower wage doesn't affect utility that much, because the worker anticipates that they will leave soon anyway. At higher levels of wages, the worker knows they will stay longer, so an incrementally higher wage means that the worker will keep that higher income for a long time.

Figure 9: Worker Value V_j as a Function of Wages w_j, $\beta = .99$

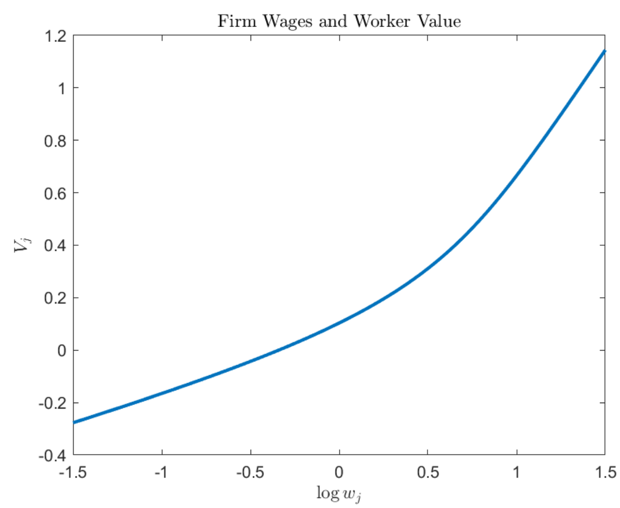
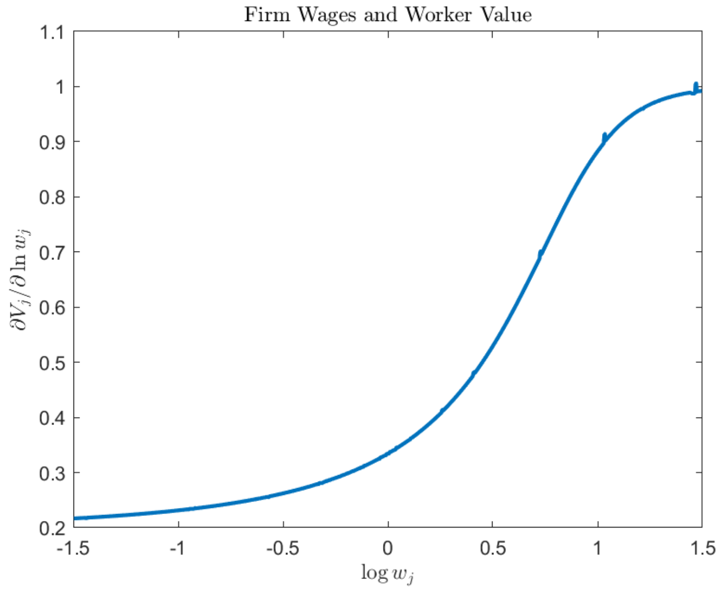


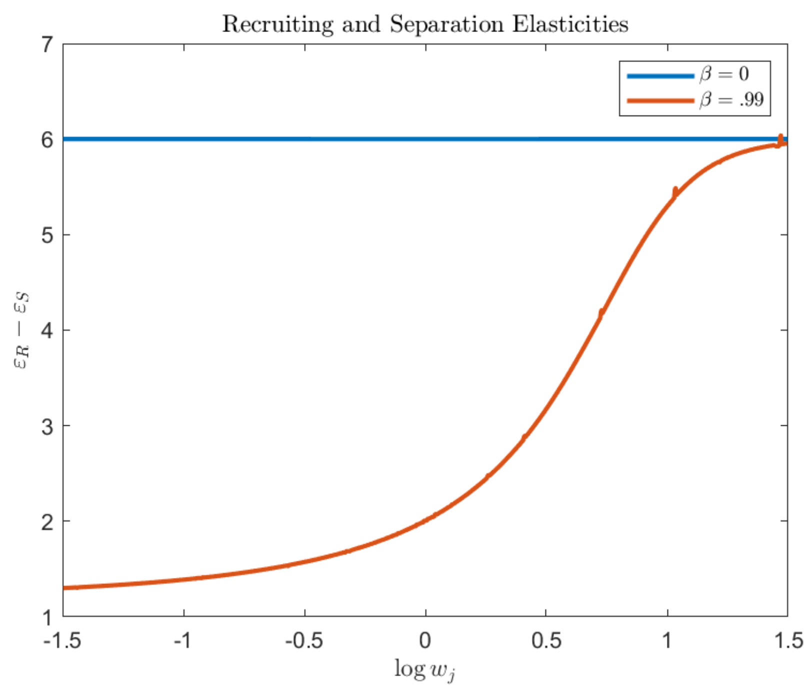
Figure 9 plots the relationship between the value function V_j and the log wage $\ln w_j$. You can that the semi-elasticity of the value function with respect to the wage approaches 1: for every percent increase in the wage, utility increases by one log points.

Figure 10: Semi-Elasticity of Worker Value to Wages



Lastly, we can plot the sum of the recruiting elasticity and (negative) separation elasticities for different rates of discounting in Figure 11. If workers are myopic, then $\varepsilon_{Rw} - \varepsilon_{Sw} = \gamma$. However, if workers are forward looking, then workers are as sensitive to relative wage differences as myopic workers only if wages are very high. At lower wage levels, workers become less sensitive to wage differences because of the option value of leaving low wage firms.

Figure 11: Recruiting and Separation Elasticities, $\beta = 0$ and $\beta = .99$



In total, having forward looking workers, and a contracting environment where firms post fix-wage contracts generates recruiting less separation elasticities, just as in the myopic case. However, the recruiting minus separation elasticities $\varepsilon_{Rw} - \varepsilon_{Sw}$ is no longer constant for different values of wages w_j, and so we would lose the tractability of many of the results in the paper.

B.2 Fixed Wage Contract Posting with Forward-Looking Workers

In the previous section, we derived the recruiting and separation elasticities for a firm choosing its wage policy, under the assumption that all its competitors are choosing the same wage. In this section, we will show that a firm that offers fix-wage contracts to forward-looking workers has the same optimal wage expression as firms that are posting wages period by period for myopic workers that are constrained to pay new hires and incumbents the same wage.

As for notation, the present period is time t. s denotes the number of periods prior to t, τ denotes periods into the future. Let $N_{s,t+\tau}$ be the number of workers hired in period s who are still in the firm at period $t+\tau$. For a cohort of workers who arrive in year s, firms promise to pay wage w_s as long as those workers are in the firm.

$$\max_{\{\{N_{t,t+\tau}\}\},\{w_{t+\tau}\},\{V_{t+\tau}\}} \sum_{\tau=0}^{\infty} \left(\frac{1}{1+\rho}\right)^{\tau} \left(A \left(\sum_{s=-\infty}^{t+\tau} N_{s,t+\tau} \right)^{\alpha} - \sum_{s=-\infty}^{t+\tau} w_s N_{s,t+\tau} - c \left(\frac{V_{t+\tau}}{N_{t+\tau-1}} \right)^{\chi} V_{t+\tau} N_{t+\tau-1}^{\sigma\chi} \right)$$

subject to

$$\begin{aligned} N_{s,t+\tau} &= (1 - S(w_s)) N_{s,t+\tau-1} \\ N_{t,t} &= R(w_t) V_t. \end{aligned}$$

The firm's lagrangian is

$$\begin{aligned} \mathcal{L} : \sum_{\tau=0}^{\infty} \left(\frac{1}{1+\rho}\right)^{\tau} & \left(A \left(\sum_{s=-\infty}^{t+\tau} N_{s,t+\tau} \right)^{\alpha} - \sum_{s=-\infty}^{t+\tau} w_s N_{s,t+\tau} - c \left(\frac{V_{t+\tau}}{N_{t-1+\tau}} \right)^{\chi} V_{t+\tau} N_{t-1+\tau}^{\sigma\chi} + \right. \\ & + \sum_{s=-\infty}^{t+\tau-1} \mu_{s,t+\tau} [-N_{s,t+\tau} + N_{s,t+\tau-1}(1 - S(w_s))] \\ & \left. + \lambda_{t+\tau} [-N_{t+\tau,t+\tau} + R(w_{t+\tau}) V_{t+\tau}] \right). \end{aligned}$$

The first order condition on w_t (i.e., $\tau = 0$)

$$
\begin{aligned}
\mathcal{L}_{w_t} :& -\sum_{\tau=0}^{\infty}\left(\frac{1}{1+\rho}\right)^{\tau} N_{t,t+\tau} - \sum_{\tau=1}^{\infty}\left(\frac{1}{1+\rho}\right)^{\tau} \mu_{t,t+\tau} N_{t,t+\tau-1} S'(w_t) + \lambda_t R'(w_t) V_t = 0 \\
\mathcal{L}_{V_t} :& \lambda_t R(w_t) - c(1+\chi) V_t^{\chi} N_{t-1}^{-\chi(1-\sigma)} = 0 \\
\mathcal{L}_{N_{t,t}} :& MRPL_t - w_t + \left(\frac{1}{1+\rho}\right) c\chi(1-\sigma) V_{t+1}^{1+\chi} N_t^{\chi(1-\sigma)-1} - \lambda_t + \left(\frac{1}{1+\rho}\right)\mu_{t,t+1}(1 - S(w_t)) = 0 \\
\mathcal{L}_{N_{t,t+\tau}}|_{\tau\geq 1} :& \left(\frac{1}{1+\rho}\right)^{\tau} MRPL_{t+\tau} - \left(\frac{1}{1+\rho}\right)^{\tau} w_t + \left(\frac{1}{1+\rho}\right)^{\tau+1} c\chi(1-\sigma) V_{t+\tau+1}^{1+\chi} N_{t+\tau}^{\chi(1-\sigma)-1} - \left(\frac{1}{1+\rho}\right)^{\tau} \mu_{t,t+\tau} +
\end{aligned}
$$

In steady state, these equations all future variables are equal to time t variables, except that $N_{t,t+\tau} = (1 - S(w_t))N_{t,t+\tau-1}$. Therefore we have

$$
\begin{aligned}
\lambda_t R'(w_t) V_t =& \sum_{\tau=0}^{\infty}\left(\frac{1}{1+\rho}\right)^{\tau} (1 - S(w))^{\tau} N_{t,t} + \mu_t S'(w_t) N_{t,t} \sum_{\tau=1}^{\infty}\left(\frac{1}{1+\rho}\right)^{\tau} (1 - S(w_t))^{\tau-1} \\
\lambda_t =& \frac{c(1+\chi) V_t^{\chi} N_{t,t}^{-\chi(1-\sigma)}}{R(w_t)} \\
\lambda_t =& MRPL_t - w_t + \left(\frac{1}{1+\rho}\right) c\chi(1-\sigma) V_t^{1+\chi} N_t^{\chi(1-\sigma)-1} + \left(\frac{1}{1+\rho}\right)\mu_t(1 - S(w_t)) \\
\mu_t =& MRPL_t - w_t + \left(\frac{1}{1+\rho}\right) c\chi(1-\sigma) V_t^{1+\chi} N_t^{\chi(1-\sigma)-1} + \left(\frac{1}{1+\rho}\right)\mu_t(1 - S(w_t))
\end{aligned}
$$

The first order condition on wages can be re-written, with a change of variables $\kappa = \tau - 1$, and $\tilde{\beta} = \frac{1}{1+\rho}$ as

$$
\begin{aligned}
\lambda_t R'(w_t) V_t =& \sum_{\tau=0}^{\infty}\left(\frac{1}{1+\rho}\right)^{\tau} (1 - S(w))^{\tau} N_{t,t} + \mu_t S'(w_t) N_{t,t} \tilde{\beta} \sum_{\tau=1}^{\infty} \tilde{\beta}^{\tau-1} (1 - S(w_t))^{\tau-1} \\
=& \sum_{\tau=0}^{\infty}\left(\frac{1}{1+\rho}\right)^{\tau} (1 - S(w))^{\tau} N_{t,t} + \mu_t S'(w_t) N_{t,t} \tilde{\beta} \sum_{\kappa=0}^{\infty} \tilde{\beta}^{\kappa} (1 - S(w_t))^{\kappa} \\
=& \sum_{\tau=0}^{\infty}\left(\frac{1}{1+\rho}\right)^{\tau} (1 - S(w))^{\tau} N_{t,t} + \mu_t S'(w_t) N_{t,t} \tilde{\beta} \frac{1}{1 - \tilde{\beta}(1 - S(w_t))}
\end{aligned}
$$

In the limit at $\rho \to 0$ (and $\tilde{\beta} \to 1$), the first first order condition in steady state becomes

$$
\lambda_t R'(w_t) V_t = \frac{N_{t,t}}{S(w_t)} + \frac{\mu_t S'(w_t) N_{t,t}}{S(w_t)}
$$

Note at this point that $N_{t,t}$ is the number of workers who arrive in each cohort, but N_t is the total number of workers. Since the wages for each cohort are constant at $w = w_t$. In steady state, the number of workers in the firm at period t is an infinte sum backwards of workers who were hired in the past:

$$N_t = \sum_{s=-\infty}^{t} (1 - S(w_s))^{t-s} N_{s,s}.$$

Again with a change of variables where $\xi = t - s$, and $w_s = w_t$ and $N_{s,s} = N_{t,t}$ $\forall s$, we have

$$N_t = \sum_{\xi=0}^{\infty} (1 - S(w_t))^{\xi} N_{t,t}. = \frac{N_{t,t}}{S(w_t)}$$

Plugging this in, as well as that $\lambda_t = \mu_t$, we get

$$\lambda_t R'(w_t) V_t = N_t + \lambda_t S'(w_t) N_t$$

With some algebra:

$$\begin{aligned}
\lambda_t \left(R'(w_t) V_t - S'(w_t) N_t \right) &= N_t \\
c(1+\chi) \left(\frac{V_t}{N_t} \right)^{\chi} N_t^{\sigma\chi} \frac{1}{R(w_t)} \left(R'(w_t) V_t - S'(w_t) N_t \right) &= N_t \\
c(1+\chi) \left(\frac{V_t}{N_t} \right)^{\chi} N_t^{\sigma\chi} \frac{1}{R(w_t)} \left(R'(w_t) \frac{V_t}{N_t} - \frac{S'(w_t)}{S(w_t)} S(w_t) \right) &= 1
\end{aligned}$$

If the wage is equal to w_t at all horizons, then the number of hires $V_t R(w_t) = S(w_t) N_t$. Therefore using that $V_t/N_t = S(w_t)/R(w_t)$, we have

$$\begin{aligned}
c(1+\chi) \left(\frac{V_t}{N_t} \right)^{\chi} N_t^{\sigma\chi} \frac{1}{R(w_t)} \left(R'(w_t) \frac{S(w_t)}{R(w_t)} - \frac{S'(w_t)}{S(w_t)} S(w_t) \right) &= 1 \\
c(1+\chi) \left(\frac{V_t}{N_t} \right)^{\chi} N_t^{\sigma\chi} \frac{S(w_t)}{R(w_t)} \left(\frac{R'(w_t)}{R(w_t)} - \frac{S'(w_t)}{S(w_t)} \right) &= 1 \\
c(1+\chi) \left(\frac{S(w_t)}{R(w_t)} \right)^{1+\chi} \left(\varepsilon_{R,w} - \varepsilon_{S,w} \right) N_t^{\sigma\chi} &= w_t
\end{aligned}$$

Thus, a firm in steady state that knows it will be in steady state forever has the same optimal wage expression as if workers were myopic and the firm posted wages period by period.

Out of Steady State Wages We now want to consider that w_t may not be equal to $w_{t+\tau}$ for all τ. Revisiting the first order conditions of the firm's lagrangian, and rolling the first order condition on $N_{t,t}$ forward to period $N_{t+\tau,t+\tau}$, we have

$$\begin{aligned}
\mathcal{L}_{N_{t+\tau,t+\tau}} &: MRPL_{t+\tau} - w_t + \beta c\chi(1-\sigma) V_{t+\tau+1}^{1+\chi} N_{t+\tau}^{\chi(1-\sigma)-1} - \lambda_{t+\tau} + \beta \mu_{t,t+\tau+1}(1 - S(w_t)) = 0 \\
\mathcal{L}_{N_{t,t+\tau}}|_{\tau \geq 1} &: MRPL_{t+\tau} - w_t + \beta c\chi(1-\sigma) V_{t+\tau+1}^{1+\chi} N_{t+\tau}^{\chi(1-\sigma)-1} - \mu_{t,t+\tau} + \beta \mu_{t,t+\tau+1}(1 - S(w_t)) = 0
\end{aligned}$$

It's evident from these expressions that $\lambda_{t+\tau} = \mu_{t,t+\tau}\ \forall \tau$.

$$\begin{aligned}
\lambda_t R'(w_t) V_t =& \sum_{\tau=0}^{\infty} \beta^\tau N_{t,t+\tau} + \sum_{\tau=1}^{\infty} \beta^\tau \mu_{t,t+\tau} N_{t,t+\tau-1} S'(w_t) \\
\lambda_t R(w_t) =& c(1+\chi) V_t^\chi N_{t-1}^{-\chi(1-\sigma)}
\end{aligned}$$

We once again have that $N_{t,t+\tau} = N_{t,t}(1-S(w_t))^\tau$. Substituting this in, as well as the expressions for the lagrange mutlipliers λ, we have

$$\begin{aligned}
\lambda_t R'(w_t) V_t =& \sum_{\tau=0}^{\infty} \beta^\tau N_{t,t}(1-S(w_t))^\tau + \sum_{\tau=1}^{\infty} \beta^\tau \lambda_{t,t+\tau} N_{t,t+\tau-1} S'(w_t) \\
c(1+\chi)\frac{R'(w_t)}{R(w_t)} V_t^{1+\chi} N_{t-1}^{-\chi(1-\sigma)} =& \frac{N_{t,t}}{S(w_t)} + S'(w_t) \sum_{\tau=1}^{\infty} \beta^\tau c(1+\chi) \frac{V_{t+\tau,t+\tau}^\chi N_{t+\tau-1,t+\tau-1}^{-\chi(1-\sigma)}}{R(w_{t+\tau})} N_{t,t+\tau-1}
\end{aligned}$$

Let $\beta \to 1$, so

$$\begin{aligned}
c(1+\chi)\frac{R'(w_t)}{R(w_t)} V_t^{1+\chi} N_{t-1}^{-\chi(1-\sigma)} =& \frac{N_{t,t}}{S(w_t)} + S'(w_t) \sum_{\tau=1}^{\infty} \beta^\tau \frac{c(1+\chi) V_{t+\tau,t+\tau}^\chi N_{t+\tau-1}^{-\chi(1-\sigma)}}{R(w_{t+\tau})} N_{t,t+\tau-1} \\
c(1+\chi)\frac{R'(w_t)}{R(w_t)} V_t^{1+\chi} N_{t-1}^{-\chi(1-\sigma)} =& \frac{N_{t,t}}{S(w_t)} + S'(w_t) \sum_{\tau=1}^{\infty} \beta^\tau \frac{c(1+\chi) V_{t+\tau,t+\tau}^\chi N_{t+\tau-1}^{-\chi(1-\sigma)}}{R(w_{t+\tau})} N_{t,t}(1-S(w_t))^{\tau-1} \\
c(1+\chi)\frac{R'(w_t)}{R(w_t)} V_t^{1+\chi} N_{t-1}^{-\chi(1-\sigma)} =& \frac{N_{t,t}}{S(w_t)} + S'(w_t) N_{t,t} \sum_{\tau=1}^{\infty} \beta^\tau \frac{c(1+\chi) V_{t+\tau,t+\tau}^\chi N_{t+\tau-1}^{-\chi(1-\sigma)}}{R(w_{t+\tau})} (1-S(w_t))^{\tau-1} \\
c(1+\chi)\frac{R'(w_t)}{R(w_t)} V_t^{1+\chi} N_{t-1}^{-\chi(1-\sigma)} =& \frac{N_{t,t}}{S(w_t)} \left(1 + \frac{S'(w_t)}{S(w_t)} S(w_t)^2 \sum_{\tau=1}^{\infty} \beta^\tau \frac{c(1+\chi) V_{t+\tau,t+\tau}^\chi N_{t+\tau-1}^{-\chi(1-\sigma)}}{R(w_{t+\tau})} (1-S(w_t))^{\tau-1}\right)
\end{aligned}$$

Rearranging and multiplying both sides by w_t yields

$$\begin{aligned}
w_t =& \frac{S(w_t)}{N_{t,t}} c(1+\chi)\varepsilon_{R,w_t} V_t^{1+\chi} N_{t-1}^{-\chi(1-\sigma)} - \varepsilon_{S,w_t} S(w_t)^2 \sum_{\tau=1}^{\infty} \beta^\tau \frac{c(1+\chi) V_{t+\tau,t+\tau}^\chi N_{t+\tau-1}^{-\chi(1-\sigma)}}{R(w_{t+\tau})} (1-S(w_t))^{\tau-1} \\
w_t =& N_{t-1} \frac{S(w_t)}{N_{t,t}} c(1+\chi)\varepsilon_{R,w_t} \left(\frac{V_t}{N_{t-1}}\right)^{1+\chi} N_{t-1}^{\sigma\chi} - \varepsilon_{S,w_t} S(w_t)^2 \sum_{\tau=1}^{\infty} \beta^\tau \frac{c(1+\chi)\left(\frac{V_{t+\tau,t+\tau}}{N_{t+\tau-1}}\right)^\chi N_{t+\tau-1}^{\chi\sigma}}{R(w_{t+\tau})} (1-S(w_t))^{\tau-1}
\end{aligned}$$

We can again see that in steady state, we would have $N_{t-1} = N_t = N_{t,t}/S(w_t)$ and $S(w_t)/R(w_t) = V_t/N_t$, which would yield

$$\begin{aligned} w_t =& c(1+\chi)\varepsilon_{R,w}\left(\frac{V}{N}\right)^{1+\chi} N_t^{\sigma\chi} - c(1+\chi)\varepsilon_{S,w}\left(\frac{V}{N}\right)^{1+\chi} N^{\sigma\chi} \\ =& c(1+\chi)\left(\frac{V}{N}\right)^{1+\chi}\left(\varepsilon_{R,w} - \varepsilon_{S,w}\right) N^{\sigma\chi}. \end{aligned}$$

Revisiting the out of steady state equation

$$w_t = N_{t-1}\frac{S(w_t)}{N_{t,t}}c(1+\chi)\varepsilon_{R,w_t}\left(\frac{V_t}{N_{t-1}}\right)^{1+\chi} N_{t-1}^{\sigma\chi} - \varepsilon_{S,w_t}S(w_t)^2\sum_{\tau=1}^{\infty}\beta^{\tau}\frac{c(1+\chi)\left(\frac{V_{t+\tau,t+\tau}}{N_{t+\tau-1}}\right)^{\chi} N_{t+\tau-1}^{\chi\sigma}}{R(w_{t+\tau})}(1-S(w_t))^{\tau-1}$$

we can see that this equation says that the optimal wage is a function of current and future vacancy rates V_t/N_{t-1}, where future vacancy rates are weighted by the number of employees hired in period t who are still remaining in period $t+\tau$. Intuitively, since new hires and incumbents are perfectly substitutable, and because incumbents and new hires provide the same value next period by decreasing marginal hiring costs, the vacancy rate of future cohorts reveals the value to the firm of getting new workers in the door. Given diminishing returns to recruiting, high vacancy rates indicate high marginal value of having more workers, and hence a higher marginal value of retaining workers hired in period t in future periods $t+\tau$. Thus, in total, the wage of workers hired today will be a function of the vacancy rate today and future vacancy rates. This can rationalize the results in Figure 4 in the main text, where workers who are hired in years that firms are growing rapidly receive higher wage increases that persistently raise the workers' wage level, while workers who are hired when the firm is no longer growing do not see higher wages from being at the now larger firm.