Monopsony and Nominal Rigidities *

Marc de la Barrera †

MIT

Job Market Paper

November 15, 2023

Click here for most recent version.

Abstract

This paper introduces monopsony in the New Keynesian model with on-the-job search and nominal rigidities. Firms set wages to poach and retain workers and, when setting prices, they consider the wage costs and hiring costs of a new worker. I show that the importance of hiring costs is directly linked to the degree of firm monopsony and is sizeable. This contrasts with standard search models where these costs turn out to be negligible. I develop a price Philips curve and a new wage Philips curve to show how both product and labor market power affect the transmission of shocks to price and wage inflation. Monopsony power steepens the price Philips curve and flattens the wage Philips curve. In the calibrated model the main inflation driver is hiring difficulties, rather than increases in real wages. Both positive demand shocks and negative labor supply shocks reduce the real wage.

*I'm greatly indebted to Iván Werning, Martin Beraja and Lawrence Schmidt for their invaluable guidance and support throughout this project. I'm also grateful to Daron Acemoglu, George-Marios Angeletos, David Autor, Isaac Baley, Ari Bronsoler, Ricardo Caballero, Tomás Caravello, Eduardo Davila, Masao Fukui, Jordi Galí, Basil Halperin, Chen Lian, Pedro Ignacio Martínez, Andreu Mas-Colell, Rafael Veiel, Jaume Ventura, Christian Wolf and participants at MIT seminars for helpful comments.

†MIT Department of Economics. Email: mbarrera@mit.edu

1 Introduction

There is little debate that most firms set prices and wages, and that those decisions are subject to nominal rigidities. Yet the leading macroeconomic models of wage rigidity abstract from these facts, either by assuming that firms take wages as given or they are bargained with workers. In this paper, we take a monopsonistic view of the labor market, where firms post wages in an environment with search frictions, and derive the macroeconomic implications of this natural assumption.

Monopsony, introduced by Robinson (1933) and popularized by Manning (2003), is a central topic in labor economics but it has not been developed in the wage rigidity literature. Given that most employment contracts are not bargained at all (Hall and Krueger (2012)), and wages are set infrequently (Grigsby, Hurst, and Yildirmaz, 2021), this is an important omission. Instead, the benchmark model for wage rigidity is the one popularized by Erceg, Henderson, and Levin (2000) where unions set wages and firms can hire at will. Wage dynamics are determined by the household's willingness to supply labor, as is standard in models with a neoclassical labor supply. Despite the model's success in fitting the dynamics of marginal costs, it is unlikely that this is the right model to think about wage setting in a country like the US, where union density is at 10%.

An alternative and more realistic approach is to include search frictions in the New Keynesian model, which is the avenue that we take. The standard way of defining wage formation under this view is by some sort of bargaining, with different options to introduce wage rigidity in this framework (Blanchard and Galí, 2010, Gertler, Sala, and Trigari, 2008). In all of them, what determines the dynamics of wages is the outside option of workers and their bargaining power. A problem with bargaining models is that we do not have evidence of such protocols nor guidance on how to estimate them. In contrast, in this paper wages are determined by competition across firms, and we can get estimates of the intensity of such competition. A major difference between bargaining models and the one presented lies on the calibration. In bargaining models, the relevance of hiring costs is negligible compared to the wage that firms have to pay to workers. In the monopsonistic model, it coincides with the wage markdown, which we have estimates for.

Section 2 introduces the elements of the monopsonistic model. It is a dynamic wage posting model with on-the-job search in the spirit of Burdett and Mortensen (1998) with endogenous vacancy creation, price setting power, and nominal rigidities.[1] Firms post wages to reduce turnover costs. By setting a higher wage, they have to pay their entire workforce

[1]There are several models that share some of those characteristics, but this is the first that combines all of them. The closest ones are Coles and Mortensen (2016) for a wage posting model with flexible prices and wages, or Krause and Lubik (2007) for a model with search with price rigidities and bargained wages.

more but the vacancy yield increases and the quit rate decreases. When setting prices, firms take into account the fact that having an extra employee entails two costs: the wage that has to be paid and their hiring cost.

The consumption side of the model is standard, and worker behaviour is simplified. Unemployed workers accept any job regardless of the wage, and job-to-job decisions are made by comparing current wages plus an idiosyncratic taste shock. Taste shocks are introduced for two reasons. First, they rationalize that many job-to-job transitions are to lower-paying jobs (Sorkin, 2018) and can account for heterogeneous moving costs. Standard models assume exogenous reallocation shocks to account for these transitions. Second, they reduce the labor supply elasticity at the individual firm level. If the variance of such shocks is sufficiently big, then Albrecht, Carrillo-Tudela, and Vroman (2018) show that a symmetric wage equilibrium exists, which will simplify the exposition. The model is closed by deriving the general equilibrium acceptance and turnover rates. While a single firm can raise wages to attract workers, if all firms do the same, the effect is offset and the only way to increase employment is by raising vacancies. This leads to a second round of effects, as more vacancies imply more quits, which further increases the incentive to increase wages and post more vacancies, and so on.

Having presented the main elements of the model, in Section 3 we derive the implications of monopsony in the price and wage Philips curve. We start by showing that firms with more monopsony power devote relatively more resources into hiring costs rather than paying wages. The importance of marginal hiring costs coincides with the wage markdown. If idiosyncratic taste shocks are dispersed, firms find it optimal to lower wages and wait for the workers with high preference to work with them.

Next we derive the price and wage Philips curve and express them as a function of labor market variables. In search models, inflation dynamics are determined by the evolution of the real wage and the marginal cost of hiring. Given that hiring costs are more volatile than the real wage, and monopsony increases the importance of hiring costs, more monopsonistic economies have a *steeper* price Philips curve. That is, given a sequence of aggregate output, higher monopsony implies a higher inflation response. The result is the opposite for the wage Philips curve. With a lower elasticity of labor supply, the cost of being missoptimized on wages is reduced, and therefore the incentive to move them is also lower. As in Lorenzoni and Werning (2023), with nominal rigiditites in prices and wages the sign of the real wage response to a demand shocks is not determined. For positive shocks, more monopsony makes it more likely that the real wage goes down, as we will get from the calibrated model.

We end the section by considering the case where firms pay per hire instead of paying per vacancy. This is the extreme case where finding workers is free but training them is not, in spirit of Salop (1979) turnover model. This assumption has two implications. First, since training costs are independent of the business cycle, both the price and wage Philips curve become flatter.[2] Second, this assumption increases monopsony power and the importance of hiring costs by reducing the elasticity of labor supply by a factor of two. The literature of monopsony has widely adopted the argument made by Manning (2003), which shows that in general equilibrium, the quit elasticity coincides with the acceptance elasticity. But when posting vacancies is free, they do not care about the latter. The measure of monopsony power should be reevaluated and this is left for future work.

How important are those marginal hiring costs, relative to the wage that firms have to pay? In Section 4 we answer this question using the sufficient statistics derived in Proposition 1. By knowing the wage markdown we can infer marginal hiring costs. All we need is the quit elasticity at the firm level, which has been estimated by many studies, and the turnover rate which is observable. Sokolova and Sorensen (2021) propose using 3.5 as a best practice estimate, and the US turnover rate is 12% quarterly. Those values imply a wage markdown of 13%, which is equivalent to say that marginal hiring costs represent 13% of the total cost of employing a new worker. From here, we can obtain the value of a worker, which we estimate it to be 15 weeks of its wage. This value is in line with direct estimates of hiring costs provided by Muehlemann and Strupler (2018) or the actual pricing of staffing firms, who usually charge from 8 to 16 weekly wages. In contrast, traditional calibrations of search models like Gertler, Sala, and Trigari (2008), Blanchard and Galí (2010) or Christiano, Eichenbaum, and Trabandt (2016) imply that these costs are below one week of wages, one order of magnitude smaller. Hiring costs in bargaining models are usually calibrated to satisfy the free entry condition. In those models, the real wage is the main driver of marginal costs.

Finally, Section 5 calibrates the model and performs several exercises. An advantage of wage posting is that the calibration does not depend on unobserved and contested parameters like the bargaining power of workers or the value of unemployment.[3] Instead of calibrating these two parameters we target the quit elasticity, which can be estimated. When the model is subject to a demand shock, individual firms post more vacancies, increase nominal wages, and prices. All other firms do the same, which offset the benefits of raising wages. More vacancies mean (i) more workers quit and (ii) the quit elasticity increases. The resulting

[2] C. A. Pissarides (2009) and Christiano, Eichenbaum, and Trabandt (2016) underscore the importance in fixed costs of hiring, independent of the state of the labor market. However, they crucially assume these costs are paid *after* a match has been made but *before* the bargaining takes place, which makes them sunk.

[3] For example, Hagedorn and Manovskii (2008) argues that worker's bargaining power is to zero, Gertler, Sala, and Trigari (2008) calibrates it to be close to one, and Gagliardone, Gertler, et al. (2023) exogenously set to 0.5 to satisfy the Hosios condition.

response is high price and wage inflation, and a drop in tehr real wage. Since the real wage falls, the sole driver of inflation is the increase in hiring costs, as in Krause and Lubik (2007). This is not the only paper that emphasises that demand shocks can reduce the real wage. Lorenzoni and Werning (2023) get the same result by imposing strong diminishing returns to labor, which increase marginal costs despite nominal wages not rising. We compare the model to an alternative model where wages are bargained with real wage rigidities as in Blanchard and Galí (2010). The bargained model features higher employment response and lower inflation response, with an increase in the real wage.

We extend the model to account for labor supply shocks, supported by the evidence that not all unemployed accept all job offers Faberman et al. (2022). Instead they receive a flow value of being unemployed and are subject to the same idiosyncratic taste shocks than workers. We can then represent labor supply shocks by an increase in these unemployment benefits, or an increase in the disutility of work, as happened during the Covid recovery. A unique feature of the monopsonistic model is that labor supply shocks can *lower* the real wage. Other search models would predict the opposite, even those with nominal rigidities like Gertler, Sala, and Trigari (2008). When workers are unwilling to work, any individual firm has the incentive to increase wages to attract them, but also they pass the extra cost into prices. Whether they do most of the adjustment on prices or wages depends on the relative product and labor market elasticities and the rigidities they face. Firms do not internalize that by raising its own price, they are lowering the real wage of everyone. This mechanism is a distinctive feature of this model which requires the notion of 'pay workers more' in order to ease labor shortages.

The recent post-Covid recovery period has been characterized by a sharp increase in labor shortages, price inflation, and nominal wage inflation, as shown in Figure 1. Business owners claimed they could not find workers and the general response, as exemplified by President Biden in June 2021, was to "pay them more". This narrative is consistent with a monopsonistic view of the labor market. Firms' narrative implies that there are frictions in the labor market that prevent them from hiring at will, and the response implies that by setting higher wages, labor shortages will ease. However, this argument misses two points in general equilibrium. Any firm can raise its wage to attract more workers, but this effect is offset by other firms doing the same. And firms that face difficulties hiring workers can opt to raise prices instead in order to reduce their individual demand, so the effect on the real wage is undetermined. Autor, Dube, and Mcgrew (2023) find that during the Covid recovery, regions with a tighter labor market saw a larger nominal wage increase but also an equally large price increase that offset any effect on the real wage.

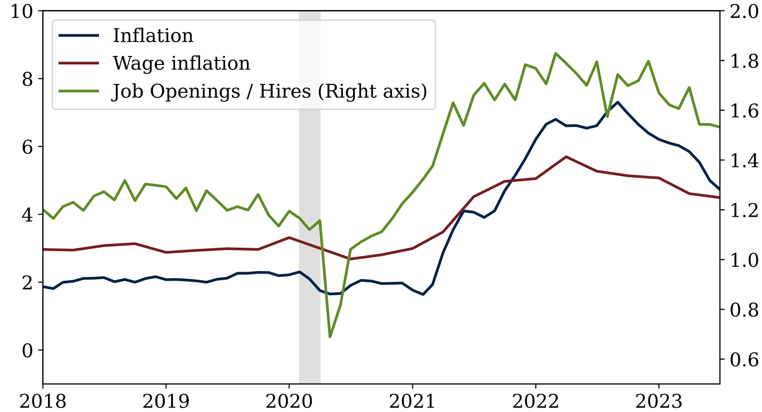

Figure 1: Dynamics of inflation (All-items less energy), wage inflation (Employment Cost Index) and job openings prior and after the Covid pandemic. Left axis correspond to annual inflation rates, right axis the ratio of job openings divided by hires in a given month.

Related Literature. This paper is mainly related to four strands of literature. It brings monopsony to a dynamic, general equilibrium models in a tractable way. It compares the model mechanism to standard New Keynesian models, with special focus on those with search frictions. It emphasises the hiring costs, and finally it relates to papers discussing the recent inflation surge.

The monopsony literature distinguishes several sources of firm wage-setting power: search frictions, and preference heterogeneity Manning (2021). By having a wage posting model with idiosyncratic preference shocks, this paper contains elements of both. The first view is pioneered by the Burdett and Mortensen (1998) wage posting model with on-the-job search, while the second got traction with Berger, Herkenhoff, and Mongey (2022a) and bought into general equilibrium by Berger, Herkenhoff, and Mongey (2022b). Albrecht, Carrillo-Tudela, and Vroman (2018) shows how idiosincratic shocks in an otherwise standard wage posting model allows for the existence of a unique wage equilibrium, an assumption that we make here.

The notion that the labor supply at the firm level is not completely elastic dates back to Robinson (1933) and has gained popularity over the recent decades (Manning (2003)). Whether firms post wages or those are bargained is an empirical question, and the true answer is that it is a little bit of both. Hall and Krueger (2012) document that two-thirds of workers do not bargain at all their wage. The SCE Job Search Supplement (Faberman et al. (2022)) corroborate this finding and add that only 12% of job offers receive a counter-offer, giving empirical support for wage posting models over bargaining models. A critical measure of monopsony power is the elasticity of labor supply at the firm level, which many papers atempt to estimate. Sokolova and Sorensen (2021) provides a meta-study from which we will take the main estimate of the quit elasticity, and Bassier, Dube, and Naidu (2022) provide estimates by broad sectors.

While Burdett and Mortensen (1998) pioneered the idea that firms post higher wages to increase in firm size, the idea that firms increase wages of incumbents to save in turnover costs dates back to Salop (1979). Manning (2006) also acknowledge that firms can increase in size by posting higher wages or spending more in recruiting costs. The closest paper in terms of modeling the firm wage setting problem is Bloesch and Larsen (2023). The paper notes that when firms pay vacancy costs in terms of vacancy *rates*, then in steady state there is not a size-wage relationship at the firm level. Empirically, the importance of hiring costs has been documented in a series of papers by Blatter, Muehlemann, and Schenker (2012) and Muehlemann and Strupler (2018) using a rich dataset from Switzerland. In bargaining model, Silva and Toledo (2009) and C. A. Pissarides (2009) both emphasize the importance of post-matching hiring costs.

The goal of this paper is to bring monopsony and nominal wage rigidities to an otherwise standard New Keynesian model. The standard approach is the one pioneered by Erceg, Henderson, and Levin (2000) in which unions set wages and firms determine quantities. While it might be a good model for some European countries, unions penetration is minimal in the United States and the assumption that firms hire a bundle of labor services is hard to map into reality. Yet, the simplicity of this model has made it very popular, used in most papers with wage rigidity Christiano, Eichenbaum, and Evans (2005), Huo and Ríos-Rull (2020), Lorenzoni and Werning (2023). In those models, nominal wages are set by unions which target some marginal rate of substitution. By definition there are not search frictions, but Galí (2011) reinterprets this model to be able to talk about unemployment.

Closer to this paper are New Keynesian models that feature search frictions. The general practice (Blanchard and Galí (2010),Gertler, Sala, and Trigari (2008),Gagliardone and Gertler (2023),Moscarini and Postel-Vinay (2023)) is to separate the product and the labor market trough a perfectly competitive intermediate layer of 'labor services' produced by firms subject to search frictions. This assumption is very convenient, as it allows one to disentangle forward-looking vacancy-posting and pricing decisions and thus simplify the analysis. Final good firms buy these labor services at the equilibrium price, produce goods and are subject to price rigidities. This separation makes the problem of the labor packer a real one, where productivity is the price of the intermediate good. Then several papers consider different wage determination protocols. The current paper does not use this two-layer economy and instead it is the same firm that posts prices and vacancies. Another that takes this approach is Thomas (2011). However, in that paper wages are negotiated among firms and workers, and it is the value that induces the right amount of hours required to supply the labor that is needed to satisfy demand.

Among those papers, Blanchard and Galí (2010) and Gertler, Sala, and Trigari (2008) explicitly model wage rigidities in a search environment. In a model with matching frictions, the bargaining set for wage determination is relatively wide, because the difficulty in locating matches creates match capital the moment a tentative match is made. Any wage within the bargaining set could be an outcome of the bargain. Wage rigidity helps pin down which of these points is selected as in R. Hall (2005). The main implication of this form of wage rigidity is that wages affect hires trough the vacancy creation incentives. Lower wages make matches more profitable from the firm's prespective which induces higher vacancy creation. In this paper, lower wages lowers the vacancy yield, which makes the hiring process costlier.

Finally, the recent inflation period has spurred several papers that try to explain it. Lorenzoni and Werning (2023) and Gagliardone, Gertler, et al. (2023) highlight the importance of the low substitutability of labor with other inputs like oil. Autor, Dube, and Mcgrew (2023) study the wage compression over this period, and in one of their analysis find that at the state level, labor market tightness spurred nominal wage growth *and* price growth, with a resulting null effect on the real wage. Cerrato and Gitti (2022) document a sharp steepening of the price Philips curve during the Covid recovery.

2 A Monopolistic Model with Nominal Rigidities

This section presents the model. The starting point is a New Keynesian model with on-the-job search. The household block is standard. It supplies one unit of labor inelastically and makes consumption-saving decisions. Firms set both prices and wages subject to Rotemberg costs and post vacancies to hire workers. Unemployed workers always accept jobs and job-to-job transitions are subject to an idiosyncratic taste shock that allows for the existence of a unique equilibrium.

Households. The household block is standard. There is a representative household with a continuum of members of measure unity that make consumption-saving decisions. They supply one unit of labor exogenously and take the total labor income $\int w_{it} n_{it} di$ as given. Employment is determined through a search and matching process that we describe below. The family provides perfect consumption insurance for its members, implying that consumption is the same for each person, regardless of whether he or she is currently employed. The preferences of the representative household are the equally weighted average of the preferences of its workers,

$$U_t = E_t \sum_{k=0}^{\infty} \beta^k u(C_{t+k}),$$

where C_t is a CES aggregator of individual varieties c_{it} with elasticity ϵ_p across goods i, priced at p_{it}. Households can save on risk-free bonds B_t sold at a price Q_t set by the central bank that pays one nominal unit at $t+1$ and get profits from firms rebated Π_t. The budget constraint is

$$\int_0^1 p_{it} c_{it} di + Q_t B_t = \int_0^1 w_{it} n_{it} + B_{t-1} + \Pi_t.$$

Demand for each variety is $c_{it} = \left(\frac{p_{it}}{P_t}\right)^{-\epsilon_p} C_t$ and the price level is $P_t^{1-\epsilon_p} \equiv \int p_{it}^{1-\epsilon_p} di$. The Euler equation is

$$Q_t = \beta E_t \frac{u'(C_{t+1})}{u'(C_t)} \frac{1}{\Pi_{t+1}},$$

where $\Lambda_{t,t+1} = \beta \frac{u'(C_{t+1})}{u'(C_t)}$ is the discount factor.

Monopsonistic firms with nominal rigidities. Each variety i is sold by an infinitely lived firm that hires workers in a market subject to search frictions by posting wages and vacancies. Firms post prices p_{it} and face a downward sloping demand curve for their goods $\mathcal{D}_t\left(\frac{p_{it}}{P_t}\right)$ with elasticity ϵ_p, the same as the households. They use labor as the sole input of production, and output is $y_{it} = n_{it}$. Production is linear and productivity is normalized to one to simplify the exposition, and we relax this assumption in the quantitative section.

Firms can post vacancies v_{it} and nominal wages w_{it}. Each vacancy costs κ_v to post and it becomes a hire given an acceptance rate $a_t\left(\frac{w_{it}}{W_t}\right)$. The acceptance rate depends on the relative wage of the firm versus the market wage. In principle, firms should care about the entire wage distribution, but this formulation already assumes the existence of a symmetric equilibrium, which will be discussed later. The firm ends the period $t-1$ with n_{it-1} employees but before period t starts, an exogenous fraction $\bar{\delta}$ quit. Workers become unemployed and ready to search for a job at t. These exogenous quits represent layoffs, retirements, or reallocations, which in the data represent around two thirds of job separations and are not related to the wage paid. Once wages have been posted, a fraction $\delta_t\left(\frac{w_t}{W_t}\right)$ of the remaining workforce is poached by other firms. The employment law of motion is given by

$$n_{it} = \left(1 - (\bar{\delta} + (1-\bar{\delta})\delta_t\left(\frac{w_{it}}{W_t}\right)\right) n_{it-1} + a_t\left(\frac{w_{it}}{W_t}\right) v_{it}. \quad (1)$$

Let $\tilde{\delta}\left(\frac{w_{it}}{W_t}\right) \equiv \bar{\delta} + (1-\bar{\delta})\delta_t\left(\frac{w_{it}}{W_t}\right)$ be the total turnover at the firm level. Workers hired at time t are ready to produce. The firm faces nominal rigidities a la Rotemberg in price and wage setting. Let $x_{it} = (n_{it}, p_{it}, w_{it})$ be the state of the firm and $J_t(x_{t-1})$ be the value of a firm at time t that ended $t-1$ with $n_{it-1}, p_{it-1}, w_{it-1}$. Firms take as given the aggregate sequences $\{N_{t+k}, V_{t+k}, P_{t+k}, W_{t+k}\}_{k=0}^{\infty}$, which defines the function $\{\delta_{t+k}, a_{t+k}, \mathcal{D}_{t+k}\}_{k=0}^{\infty}$. To simplify notation, we drop the subindex i from the Bellman equation, which is

$$
\begin{aligned}
J_t(x_{t-1}) = \max_{p_t, w_t, v_t, n_t} \; & \frac{p_t}{P_t}\mathcal{D}_t\left(\frac{p_t}{P_t}\right) - \frac{w_t}{P_t}n_t - \kappa_v v_t \\
& - \frac{\kappa_p}{2}\left(\frac{p_t}{p_{t-1}} - 1\right)^2 Y_t - \frac{\kappa_w}{2}\left(\frac{w_t}{w_{t-1}} - 1\right)^2 N_t + E_t[\Lambda_{t,t+1}J_{t+1}(x_t)]
\end{aligned}
$$

subject to (1) and $\mathcal{D}_t\left(\frac{p_t}{P_t}\right) = n_t$. κ_p and κ_w are price and wage rigidity parameters, respectively. The assumption of Rotemberg pricing simplifies the formulation by making the equilibrium symmetric. Vacancy costs are assumed to be linear for simplicity. Traditional wage posting models like Burdett and Mortensen (1998) assume linear production and then they require convexity in vacancy costs to determine firm size. Here, even with linear production and linear vacancy costs, firm size is determined by the shape of the demand function.

The problem formulation differs from standard New Keynesian models with search like Gertler, Sala, and Trigari (2008), Blanchard and Galí (2010) or Christiano, Eichenbaum, and Trabandt (2016) where the identity of the price setter and the wage setter is differentiated to keep both problems tractable. Krause and Lubik (2007) has a similar model where firms hire by posting vacancies workers and have price-setting power, but they take wages as given. In Bloesch and Larsen (2023), firms solve the same model in steady state and with flexible prices and wages.

To pay for vacancy costs and the Rotemberg adjustment costs, firms buy a bundle of goods from all other firms as it is the case in a roundabout economy. They aggregate this bundle using the same elasticity as households do, so demand at the firm level is $y_{it} = \mathcal{D}_t\left(\frac{p_{it}}{P_t}\right) = \left(\frac{p_{it}}{P_t}\right)^{-\epsilon_p} Y_t$

Labor Markets. Labor markets are subject to search frictions and feature on-the-job search. There is a unit of workers willing to supply labor. Aggregate unemployment is measured at the end of the period and is given by $U_t = 1 - N_t$. As it has been noted, at the beginning of the period a fraction $\bar{\delta}$ of workers separate from their firms and search for another job. Thus, the pool of unemployed workers that search for a job at period t is $U_{t-1} + \bar{\delta}N_{t-1}$. On top of that, employed workers search with efficiency s relative to the unemployed ones. Then market tightness is

$$\theta_t = \frac{V_t}{U_{t-1} + \bar{\delta} N_{t-1} + s(1-\bar{\delta})N_{t-1}},$$

where $S_t \equiv U_{t-1} + \bar{\delta}N_{t-1} + s(1-\bar{\delta})N_{t-1}$ is total search effort. A constant returns matching function generates $m(V_t, S_t)$ matches at period t. The probability that a vacancy meets a potential worker is $q(\theta_t) = m(\theta^{-1}, 1) = \frac{m(S_t,V_t)}{V_t}$ with elasticity $\frac{d\log q}{d\log\theta} = -\eta$. Conditional on a successful match, the applicant is an employed worker with probability

$$p_t^E = \frac{s(1-\bar{\delta})N_{t-1}}{U_{t-1} + \bar{\delta}N_{t-1} + s(1-\bar{\delta})N_{t-1}}$$

and unemployed with $p_t^U = \frac{U_{t-1}+\bar{\delta}N_{t-1}}{U_{t-1}+\bar{\delta}N_{t-1}+s(1-\bar{\delta})N_{t-1}}$. This description of the labor market is standard in any model with on the job search like Faberman et al. (2022) or Moscarini and Postel-Vinay (2023). For a given level of vacancies, a higher employment level makes it costlier hire a worker because (i) it is less likely to match one of them and (ii) it is more likely that the matched worker is already employed, who is likely to reject the job offer, as we describe next.

For this section, an unemployed worker that receives an offer accepts it regardless of the wage, which will be relaxed in Section 5.4. This is a simplifying assumption that makes aggregate labor supply fixed and allows firms to set wages only considering the behavior of the employed workers. In particular, a worker earning w_{it} accepts an offer that pays w_{jt} if $\varepsilon_{jt}w_{jt} \geq w_{it}$, where $\varepsilon_{jt} \sim F_\varepsilon$ is a multiplicative taste shock. Therefore, the probability that a worker accepts the offer from firm j is $1 - F_\varepsilon\left(\frac{w_{it}}{w_{jt}}\right)$. We can think of $F_\varepsilon(1) > 0.5$ as moving costs, in the sense that the probability of accepting a job offer that pays the same as the current job is less than $1/2$. This behavior corresponds to a myopic worker that fully discounts the future and makes job decisions according to the time t best option. At a steady state equilibrium, workers expect $w_{it} = w_{it+1}$ so the assumption is innocuous. But it greatly simplifies the problem of the firm out of the steady state, that otherwise would be untractable. The implications of the effect of monopsony would be the same at the expense of losing analytical tractability.

It is well known that in a wage posting model a la Burdett and Mortensen (1998), a symmetric equilibrium where all firms post the same wage does not exist. If the wage distribution had any mass point, it would be profitable for firms in that point to deviate and offer a slightly higher wage. This argument misses the fact that jobs are heterogeneous and people do not move from job to job solely based on the wage paid. Preference heterogeneity is also regarded as one of the main sources of monopsony power by firms. Albrecht, Carrillo-

Tudela, and Vroman (2018) show that if workers have idiosyncratic taste shocks for different firms with $f_\varepsilon(1)$ high enough, then a symmetric equilibrium where all firms post the same wage can be sustained. The reason is that preference heterogeneity reduces the elasticity of the turnover and acceptance rates, which makes deviating unprofitable because the wage increase required to the entire workforce does not compensate the reduced turnover and increased acceptance rate. $f_\varepsilon(1)$ is the mass of workers indifferent between two jobs that pay the same. The traditional search model is the case where $f_\varepsilon(1) \to \infty$, and workers always move to the better paying job. Trought the paper, we assume that the taste dispersion is big enough such that the symmetric equilibrium exists and let W_t be the symmetric market wage.

With the labor markets defined, the acceptance rate of vacancies is given by

$$a_t\left(\frac{w_t}{W_t}\right) = q(\theta_t)\left(1 - p_t^E F_\varepsilon\left(\frac{W_t}{w_t}\right)\right).$$

With probability $q(\theta_t)$, the vacancy meets a worker. Conditional on the match, with probability p_t^E this worker is already employed and rejects the job offer with probability $F_\varepsilon\left(\frac{W_t}{w_t}\right)$.

Similarly, the endogenous job-to-job turnover rate is

$$\delta_t\left(\frac{w_t}{W_t}\right) = \frac{V_t q(\theta_t) p_t^E \left(1 - F_\varepsilon\left(\frac{w_t}{W_t}\right)\right)}{(1-\bar{\delta})N_{t-1}}.$$

Out of $V_t q(\theta_t)$ matches, a fraction p_t^E are to employed workers, who accept the market wage offer with probability $1 - F_\varepsilon\left(\frac{w_t}{W_t}\right)$. All these matches are divided by the current mass of workers $(1-\bar{\delta})N_{t-1}$ under the assumption that no worker receives more than one offer at any given period. Finally, since there is on the job search, the aggregate law of motion of employment differs from the firm law of motion because job-to-job transitions do not add new workers into the workforce. Only vacancies that match unemployed workers add to employment. The law of motion for aggregate employment is

$$N_t = (1-\bar{\delta})N_{t-1} + V_t q(\theta_t) p_t^U. \tag{2}$$

Given that the unemployed workers accept all job offers, this law of motion is independent on the wage.

Market clearing and monetary policy To close the model, a central bank sets interest rates according to a Taylor rule that targets current inflation. It sets the bond prices to

$$Q_t = e^{m_t} \beta \left(\Pi_t^p\right)^{\phi_\pi},$$

where Π_t^p is the inflation rate and ϕ_π is its Taylor coefficient. m_t is a monetary policy shock that is interpreted as a demand shock.

Total output Y_t is devoted to consumption and to pay for the vacancy and Rotemberg costs. The market clearing condition is:

$$Y_t = C_t + \kappa_v V_t + \frac{\kappa_p}{2} \left(\Pi_t^p - 1\right)^2 Y_t - \frac{\kappa_w}{2} \left(\Pi_t^w - 1\right)^2 N_t$$

Equilibrium The equilibrium definition is standard. Firms take as given aggregate variables as given and maximize profits by choosing p_{it}, w_{it}, v_{it} and n_{it} given their initial state $(n_{it-1}, p_{it-1}, w_{it-1})$. In the symmetric equilibrium, $P_t = p_{it}, W_t = w_{it}, V_t = v_{it}$ and $N_t = n_{it}$ for all i and no firm finds it profitable to deviate.

3 Monopsony and the Philips Curves

In this section, we solve the model and derive the price Philips curve and wage Philips curve. We show how monopsony is indicative of the marginal hiring costs and how those costs affect both curves.

Optimal wage setting. First, we solve the partial equilibrium problem of a firm, which takes the acceptance $a_t(\cdot)$ and turnover $\delta_t(\cdot)$ functions as given. Let μ_t be the value of an incumbent employee. The first order condition for v_t yields

$$\mu_t = \frac{\kappa_v}{a_t\left(\frac{w_t}{W_t}\right)}.$$

The value of a worker is the cost it requires to hire it. Posting an extra vacancy costs κ_v, but it is only converted into a hire with probability $a_t\left(\frac{w_t}{W_t}\right)$. This condition is similar to the standard free entry condition that many search models have like C. Pissarides (2017), with two differences. In this model, there is no free entry but rather firms that make hiring decisions based on the demand for goods they face, and crucially, the wage posted affects the probability of a vacancy being filled.

Start by assuming that wages are flexible, $\kappa_w = 0$. Then the first order condition of the firm problem with respect to the wage is

$$\frac{1}{P_t} n_t = \frac{\kappa_v}{a_t\left(\frac{w_t}{W_t}\right)} \left(-\tilde{\delta}_t' \left(\frac{w_t}{W_t}\right) \frac{1}{W_t} n_{t-1} + a_t' \left(\frac{w_t}{W_t}\right) \frac{1}{W_t} v_t \right). \tag{3}$$

The left hand side of equation (3) represents the cost of raising wages. If the wage is increased by dw_t, the firm has to pay the entire workforce this extra amount. By assumption, it cannot discriminate between incumbent and new workers. The right hand side represents the benefits of raising it. It reduces turnover by $\tilde{\delta}'\left(\frac{w_t}{W_t}\right)\frac{1}{W_t}$ and increases the vacancy yield by $a_t'\left(\frac{w_t}{W_t}\right)\frac{1}{W_t}$. The term inside the parenthesis is how many extra workers the firm is able to attract (or keep) given v_t, n_{t-1}, when wages increase by dw_t. Each one of these workers is valued at $\mu_t = \frac{\kappa_v}{a_t\left(\frac{w_t}{W_t}\right)}$. The term $\frac{1}{W_t}$ because workers consider relative, not absolute, wage increases when they make job-to-job acceptance decisions. When the market wage is high, the relative effect of increasing wages by dw_t diminishes.

The appendix shows by using the definitions of $a_t(\cdot)$ and $\delta_t(\cdot)$ and the property that the equilibrium is symmetric, we can get the flexible real wage of the economy, which is given by

$$\omega_t^{flex} = \kappa 2 \epsilon_{a,t} \frac{V_t}{N_t}, \tag{4}$$

where $\epsilon_{a,t} \equiv \frac{d \log a_t}{d \log w} = \frac{p_t^E f_\varepsilon(1)}{1 - p_t^E F_\varepsilon(1)}$ is the individual firm acceptance elasticity. It coincides with the quit elasticity $\epsilon_{\tilde{\delta},t} \equiv -\frac{d \log \tilde{\delta}_t}{d \log w}$ in steady state.[4] Manning (2003) shows that, under certain conditions satisfied here, both elasticities coincide and hence the 2 multiplying in the right-hand side. This elasticity varies over the business cycle because p_t^E is increasing in the employment level as documented by Autor, Dube, and Mcgrew (2023). An increase in worker search effort, which is assumed to be constant here, would also increase the quit elasticity.

Equation (4) substitutes the wage setting equation of other models. In those with bargaining, the real wage is a weighted average between the unemployment benefits and the labor productivity. In models with classical labor supply, the real wage is equal to the marginal rate of substitution of households. Here, real wages raise when the quit elasticity and vacancies are high, which imply that more firms poach workers from each other. It decreases

[4]Out of the steady state, the relationship is given by

$$\epsilon_{\tilde{\delta}} = \frac{1}{1 + \frac{1}{\tilde{\delta}_t(1)\frac{N_t - N_{t-1}}{N_t}}} \epsilon_{a,t} \tag{5}$$

with employment, because the cost of raising wages is proportional to the workforce.

Adding back the wage rigidity, we get the non-linear wage Philips curve

$$\left(\Pi_t^w - 1\right)\Pi_t^w = \frac{1}{\kappa_w}\left(\kappa 2\epsilon_{a,t}\frac{V_t}{N_t} - \omega_t\right) + E_t\left[\Lambda_{t,t+1}\left(\Pi_{t+1}^w - 1\right)\Pi_{t+1}^w\frac{N_{t+1}}{N_t}\right]. \tag{6}$$

Equation (6) is a novel Philips curve derived from intra-firm competition for workers subject to nominal rigidities. The standard treatment of wage Philips curves comes from the Erceg, Henderson, and Levin (2000) treatment of unions. It is used in many papers where wage rigidity is central like Schmitt-Grohé and Uribe (2005) or Lorenzoni and Werning (2023). Appendix **??** builds a model where unions set wages, whose wage Philips curve is given by

$$\left(\Pi_t^w - 1\right)\Pi_t^w = \frac{1}{\kappa_w}\left(\mu_w\frac{v'(N_t)}{u'(C_t)} - \omega_t\right) + E_t\left[\Lambda_{t,t+1}\left(\Pi_{t+1}^w - 1\right)\Pi_{t+1}^w\frac{N_{t+1}}{N_t}\right], \tag{7}$$

where $v(N_t)$ is the labor waste of households and μ_w is the wage increase charged by the unions. Under the union's view of the world, nominal wage inflation is determined by the gap between the current real wage and the optimal wage that households would like, which is a wage markup over the marginal rate of substitution. That model implicitly assumes that employment is fully determined by firms, there is no notion of hiring costs nor labor shortages. It is assumed that each union sells a differentiated labor variety, and firms require all of them to produce, which gives the union wage-setting power. Households dislike to work so that is why higher employment raises nominal wages, and also they want to work less hours if consumption increases. The motivation for the introduction of the wage Philips curve was to prevent wages from moving a lot and generate positive comovement between consumption and employment. While it satisfies the objective, the underlying assumptions required to get the desired result does not have a clear interpretation and are hard to map into reality.

Price setting. Next we turn on how firms set prices. With linear production, the (real) marginal cost λ_t is the cost of an extra employer, which includes the real wage that has to be paid plus the cost of hiring the worker. Hiring costs take into account that a worker hired at t most likely stays at the firm at $t+1$. Letting γ_t be this cost of hiring, marginal costs are

$$\lambda_t = \omega_t + \gamma_t. \tag{8}$$

The net cost of hiring is

$$\gamma_t = \frac{\kappa}{a_t\left(\frac{w_t}{W_t}\right)} - E_t\left[\Lambda_{t,t+1}\left(1-\tilde{\delta}_t\left(\frac{w_{t+1}}{W_{t+1}}\right)\right)\frac{\kappa}{a_t\left(\frac{w_{t+1}}{W_{t+1}}\right)}\right]. \tag{9}$$

The net cost of hiring a worker is its value net of the benefit of having the worker at $t+1$ as long as it remains on the firm, which happens with probability $1-\tilde{\delta}_t\left(\frac{w_{t+1}}{W_{t+1}}\right)$.Equation (9) shares similarities with other models with job-search like Krause and Lubik (2007) or Gertler, Sala, and Trigari (2008). However, in most bargaining models, the acceptance and turnover decisions are independent of the wage. Another difference is that a model that includes on-the-job search like this or Moscarini and Postel-Vinay (2023) has an endogenous job-destruction rate that evolves over the business cycle and adds variability to the net cost of hiring. When the labor market is tight, not only workers are harder to hire but they are also more likely to quit, an amplifying mechanism that is absent in models with exogenous job destruction.

Given marginal costs, the price Philips curve is standard for any model with Rotemberg costs,

$$\Pi_t^p(\Pi_t^p-1) = \frac{1}{\kappa_p}(1-\epsilon_p+\epsilon_p\lambda_t) + E_t\left[\Lambda_{t,t+1}\Pi_{t+1}^p(\Pi_{t+1}^p-1)\frac{Y_{t+1}}{Y_t}\right]. \tag{10}$$

Absent of nominal rigidities, firms would set a price equal to a markup $\mathcal{M}_p \equiv \frac{\epsilon_p}{\epsilon_p-1}$ equal to the marginal cost.

Wage markdown and hiring costs . Both hiring costs and the actual wage are important to determine marginal costs and therefore prices. Now we show that monopsony power is indicative of the relative importance of each cost. Linearizing (8) we obtain

$$\hat{\lambda}_t = \tau\hat{\omega}_t + (1-\tau)\hat{\gamma}_t, \tag{11}$$

where $\tau \equiv \frac{\omega}{\omega+(1-\tilde{\beta})\frac{\kappa}{a}}$ is the share of real wage costs relative to the total cost of having an extra employee in steady state. This is a general result for all models with search frictions and the next proposition shows its relationship with the degree of monopsony power.

Proposition 1 (Monopsony and marginal hiring costs) *In steady state, the share of the wage costs relative to the marginal cost of a new employee is equal to the monopsonistic wage markdown.*

$$\tau \equiv \frac{\omega}{\omega+(1-\tilde{\beta})\frac{\kappa}{a}} = \frac{2\tilde{\delta}\epsilon_{\tilde{\delta}}}{2\tilde{\delta}\epsilon_{\tilde{\delta}}+1-\tilde{\beta}} \equiv \mathcal{M}_w, \tag{12}$$

where $\tilde{\beta} \equiv \beta(1-\tilde{\delta})$ is the efective discount rate.

Proposition 1 shows that firms, or economies, where labor market power is high devote a relatively higher amount of resources into hiring workers rather than paying them in wages. This is natural since higher monopsony implies lower wages and higher turnover, which increases turnover costs. The expression for the wage markdown coincides with Manning (2003) which corresponds to a firm that faces a labor supply elasticity $2\epsilon_{\tilde{\delta}}$.[5] Throughout the paper, we will be referring to τ as the degree of monopsony power, with lower τ meaning higher monopsony. In order to get the result of the proposition, use the optimal wage setting condition in steady state $\omega = \frac{\kappa}{a}\tilde{\delta}2\varepsilon_{\tilde{\delta}}$ in the definition of τ to derive (12). The result holds for any production function, not only the linear one we assumed for simplicity. It also generalizes to vacancy cost functions that are not linear in v_t, although then the formula for the wage markdown might change slightly. But it would be still the case that the wage markdown coincides with the importance of hiring costs.

Most search models pin down employment by imposing a free entry condition on the firm side and calibrate the cost of hiring a worker. In bargaining models, this calibration depends on the surplus captured by the firm when a match is realized, which in turn depends on the bargaining weights of the wage-setting protocol ϑ. We do not have a clear understating on what is this bargaining weight and there is disagreement on its value. Hagedorn and Manovskii (2008) calibrate it to be 0.052, Gertler, Sala, and Trigari (2008) set it to 0.907. Another standard practice is to set it such the Hosios condition holds, as in Faberman et al. (2022) or Gagliardone and Gertler (2023). Krause and Lubik (2007) write that "in the absence of direct evidence on the worker's share parameter ϑ we follow the literature and set $\vartheta = 0.5$". Alternatively, others have tried to assess the importance of hiring and turnover costs by directly observing such costs, like Blatter, Muehlemann, and Schenker (2012) or Silva and Toledo (2009). With Proposition 1, we can get the relative importance of hiring costs relative to the wage that firms pay by knowing the turnover rate $\tilde{\delta}$ and quit elasticity $\varepsilon_{\tilde{\delta}}$. The former is observable and easy to compute, and for the later there are many papers estimating it, summarized in Sokolova and Sorensen (2021).

Monopsonistic Philips Curves. Equipped with that result, we can get the linearized price and wage Philips curves as a function of aggregate labor market variables. The next proposition shows how we can express the Philips curves using observable labor market variables and how monopsony power affects them.

[5]When $\beta = 0$, the wage markdown becomes the familiar $\frac{\epsilon_w}{\epsilon_w+1}$ that would come from a firm facing a static labor supply $n = \mathcal{LS}(w)$ with elasticity ϵ_w.

Proposition 2 (Price and Wage Philips curves) *The dynamics of price and wage inflation of the monopsonistic economy satisfy the following system of equations:*

$$\pi_t^p = \frac{\epsilon_p}{\kappa_p}\left(\tau\hat{\omega}_t + \frac{1-\tau}{1-\tilde{\beta}}\left(\hat{V}_t - \hat{H}_t - \beta E_t\left[(1-\tilde{\delta})(\hat{V}_{t+1} - \hat{H}_{t+1}) - \tilde{\delta}\hat{\tilde{\delta}}_{t+1}\right]\right)\right) + \beta E_t[\pi_{t+1}^p] \quad (13)$$

$$\pi_t^w = \frac{1}{\kappa_w}\frac{\tau}{\mathcal{M}_p}\left(\hat{V}_t - \hat{H}_t + \frac{1-\bar{\delta}}{\bar{\delta}}\Delta\hat{N}_t + \hat{EE}_t - \hat{UE}_t - \hat{\omega}_t\right) + \beta E_t[\pi_{t+1}^w] \quad (14)$$

$$\Delta\hat{\omega}_t = \pi_t^w - \pi_t^p \quad (15)$$

a higher degree of monopsony (lower τ) raises the importance of hiring costs in the price Philips curve and flattens the wage Philips curve.

In models without labor market frictions and constant marginal product of labor, the only driver of inflation is the real wage. Search frictions acknowledge the cost of hiring workers and the fact that when workers are harder to find, firms instead raise wages. The term $\hat{V}_t - \hat{H}_t$ is a measure of how hard is to find workers, the inverse of the acceptance rate. When vacancies are high relative to the total hires in a period, it means that many of those vacancies are either unable to match a worker or they are turned down. If workers are more likely to leave the firm at $t+1$, then the net cost of hiring it at t also increases. This price Philips curve is similar to Krause, Lopez-Salido, and Lubik (2008) with two differences. On the job search makes the separation rate is endogenous and depends on market conditions, and (nominal) wages are posted by firms instead of having the real wage bargained.

Like prices, nominal wages also increase when workers are harder to find, when firms want to grow, and when the market is tight in the sense that more hires come from other employees, consistent with evidence provided by Autor, Dube, and Mcgrew (2023). Unemployed acceptance decision is independent of the wage offered so firms have low incentives to increase wages when they are mostly hiring from the unemployment pool. Moscarini and Postel-Vinay (2023) argue for the oposite sign in the EE/UE ratio which they call Acceptance Ratio AC. Theirs is a model of a job ladder, and a high AC ratio means that workers are missmatched, containing wage pressure.

A novel feature of this model is that both labor market power and product market power affect both Philips curves. Labor market power (lower ϵ_p, higher $\mathcal{M}_p$) flattens both curves. This result comes from the assumption that nominal rigidities have a menu cost component and firms set wages. When firms face Rotemberg costs, they trade off the explicit costs of increasing prices versus the benefits of having a price closer to the optimum. These benefits depend on the curvature of the profit function with respect to the price, which decreaess with product market power. If the elasticity of demand is low, firms can afford being misspriced so the incentive to change prices is reduced. If demand is very elastic, then Rotemberg pricing converges to flexible prices. A similar thing happens in the wage Philips curve. Higher monopsony power implies that having an optimal wage is less important, and wages move by less.

We have assumed that the marginal product of labor is constant for simplicity, which implies that the only driving force of inflation is related to the labor market. In the quantitative section, we allow for production to have decreasing returns to scale. Lorenzoni and Werning (2023) using a stylized unions model and Gagliardone and Gertler (2023) in a bigger DSGE model with Nash bargaining, present a model with two inputs, labor and oil, with very low elasticity of substitution. Then demand shocks or oil supply shocks are inflationary because they sharply reduce the marginal product of labor which raises marginal costs.

The results on Proposition 2 are not enough to conclude whether monopsonistic economies are more or less inflationary. Monopsony affects both the price and wage equation in different ways and they are both related by the real wage. However, we can study a simplified pricing problem where at $t = -1$ the economy is at steady state, wages are fully rigid at $t = 0$ and at $t \geq 1$ both prices and wages are fully flexible. The assumption that $\hat{\omega}_{-1} = 0$ and $\pi_t^w = 0$ imply that $\hat{\omega}_0 = -\pi_t^p$. These simplifying assumptions allow to write inflation at 0 as

$$\pi_0 = \frac{\frac{\epsilon_p}{\kappa_p}}{1 + \frac{\epsilon_p}{\kappa_p}\tau} \frac{1-\tau}{1-\tilde{\beta}} \left(\hat{V}_t - \hat{H}_t + \tilde{\beta}(\hat{V}_{t+1} - \hat{H}_{t+1}) + \beta(1-\bar{\delta})\delta\hat{\delta}_{t+1} \right).$$

In this case, it is clear that for demand shock that raises the net cost of hiring workers, monopsony is inflationary. Since nominal wages are fixed, any positive inflation level lowers the real wage. This lower real wage reduces the need to raise prices, but it is offset by increased hiring costs in the labor market. Monopsony dampens the first effect and amplifies the latter, and thus we are left with a more inflationary economy.

3.1 Alternative formulation: free vacancies, costly hire

The model previously presented assumes that vacancies are costly, as is standard in the search literature. C. A. Pissarides (2009) and Christiano, Eichenbaum, and Trabandt (2016) emphasize the importance of fixed costs of hiring that are independent of the labor market conditions to reduce wage volatility. Empirically, Muehlemann and Strupler (2018) find that pre-matching hiring costs (those related to search) account for just 21% of a firm's hiring costs. In this section we take the extreme assumption that firms face hiring costs and vacancies are virtually free.

The objective of considering this case is twofold. First, it makes it clear how the wage-setting problem differs from standard search models. Witouth search costs, bargaining models collapse to a competitive model, because there is no surplus to be shared. The dynamics of the wage are driven by the cost of matching workers. Here, while firms can costlessly post vacancies, hiring is costly and wages are set to reduce quits, more in line with a model where there are abundant applicants as in Salop (1979). Second, it challenges the monopsonistic theory that the relevant elasticity at the firm level is the sum of the quit elasticity and the acceptance elasticity.

The structure of the model is exactly the same with the exception of the firm Bellman equation, which now becomes

$$J_t(x_{t-1}) = \max_{p_t, w_t, v_t} \frac{p_t}{P_t}\mathcal{D}_t\left(\frac{p_t}{P_t}\right) - \frac{w_t}{P_t}n_t - \kappa_h h_t$$
$$- \frac{\kappa_p}{2}\left(\frac{p_t}{p_{t-1}} - 1\right)^2 Y_t - \frac{\kappa_w}{2}\left(\frac{w_t}{w_{t-1}} - 1\right)^2 N_t + \beta E_t[J_{t+1}(x_t)]$$

where $h_t = a(w_t)v_t$ are hires at period t. This formulation makes vacancies payoff-irrelevant, which means they are free to post. What allows for a positive wage in equilibrium are fairness constraints, the fact that firms have to pay the same to new and incumbent workers. Wages are not set to attract new workers since firms can post as many vacancies as needed to achieve this goal, but rather to keep incumbents from being poached by other firms. Having costless vacancies is not the same as not having search frictions. In this case, while firms can find a worker without incurring any cost, workers can only wait unemployed or at a firm until a job offer is handled to them. The the real wage is given by

$$\omega = \tilde{\delta}\kappa_h \epsilon_{\tilde{\delta}}.$$

The next proposition adapts Proposition 2 to the case when posting vacancies is free:

Proposition 3 (Costly hire) *If firms face costs to hire workers instead of posting vacancies, the price and wage Philips curves become:*

$$\pi_t^p = \frac{\epsilon_p}{\kappa_p}\left(\tau\hat{\omega}_t + \frac{1-\tau}{1-\tilde{\beta}}\beta(1-\bar{\delta})\delta\hat{\delta}_{t+1}\right) + \beta E_t[\pi_{t+1}^p]$$

$$\pi_t^w = \frac{1}{\kappa_w}\frac{\tau}{\mathcal{M}_p}\left(\frac{1-\bar{\tilde{\delta}}}{\bar{\tilde{\delta}}}\hat{\Delta}N_t + \hat{EE}_t - \hat{UE}_t - \hat{\omega}_t\right) + \beta E_t[\pi_{t+1}^w]$$

and (15), where $\tau = \frac{\tilde{\delta}\epsilon_{\tilde{\delta}}}{\tilde{\delta}\epsilon_{\tilde{\delta}}+1-\tilde{\beta}}$. Paying per hire instead of per vacancy reduces τ and flattens both the price and the wage Philips curves.

There are 2 differences between Proposition 2 and Proposition 3. First, the wage markdown increases as the effective elasticity of labor supply is reduced by a half. This observation is important for the literature that estimates wage markdowns and monopsony power. It is conventional to assume that the labor supply elasticity is $\epsilon_a + \epsilon_{\tilde{\delta}}$ and by lack of estimation of ϵ_a and using general equilibrium arguments, set $\epsilon_a = \epsilon_{\tilde{\delta}}$. This is the case in this model, but when hiring is costly ϵ_a does not matter in the wage determination. Firms do not set wages to increase the acceptance rate but rather to reduce the turnover rate. Most certainly the reality is in between, with some costs being related to finding workers (job ads, interviews...) and those can be reduced by increasing the wage, and some others are independent of the wage. This means that the effective labor supply that firms face lie on the interval $[\epsilon_{\tilde{\delta}}, 2\epsilon_{\tilde{\delta}}]$, depending on how important are costs related to vacancies versus costs related to hiring.

The second difference is that the terms related to the acceptance rate $V_t - H_t$ are not present in Proposition 3. This reduces the price and wage reaction to labor market conditions, because the cost of getting an extra employee no longer depends on market tightness. However, market tightness is still what drives the dynamics of wages because when the market is hot, it is more likely that workers recieve outside offers, which pushes firms to increase wages to prevent them from quitting. On the marginal cost side, an increased turnover rate reduces the net cost of hiring a worker since it is more likely that he will quit at $t + 1$.

4 The Importance of Hiring Costs

Hiring workers is costly, and firms internalize that when making price and wage-setting conditions. The previous section has emphasized the importance of such costs and how they relate to monopsony power, but it remains to assess whether they are indeed a significant driver of marginal costs. First, we note that previous search models already have the notion that hiring costs matter for inflation dynamics, but their calibration implies that their effect is negligible. Then using the result in Proposition 1, that in the monopsonistic model presented here, a reasonable calibrated model implies that search frictions are an important driver of the inflation.

The traditional approach. The general approach of modeling New Keynesian models with search frictions is to assume a two-layer economy. Christiano, Eichenbaum, and Trabandt (2016), Gertler, Sala, and Trigari (2008), Moscarini and Postel-Vinay (2023), Blanchard and Galí (2010), to put some examples, share this structure. This assumption allows to separate the identity of the price setter and the wage setter, simplifying both problems. Firms subject to search frictions hire workers, and sell 'labor services' at a perfectly competitive price ϑ_t, using Christiano, Eichenbaum, and Trabandt (2016) notation, and retailers buy these services, differentiate them, and sell them subject to nominal rigidities and a downward sloping demand curve.

Let's consider the labor market, where firms can post vacancies paying some cost to be defined. Letting J_t be the value of a single job, the bellman equation that governs it

$$J_t = \vartheta_t - w_t + \beta(1-\bar{\delta})J_{t+1}. \tag{16}$$

When there is a match, the job generates a flow surplus of $\vartheta_t - w_t$ at t and with probability $(1-\bar{\delta})$ it survives another period. The papers mentioned, to the exception of Moscarini and Postel-Vinay (2023), do not feature on-the-job search so separations are exogenous. We do not need to specify how the wage is defined, different papers have different wage-setting protocols. The model is closed by imposing a free entry condition, equating the expected benefit of a vacancy to its cost. In reduced form, it pins down the value of a job given some function $c(\cdot)$ that potentially depends on market tightness,

$$J_t = c(\theta_t).$$

The standard case, as in C. Pissarides (2017), is when vacancies have a cost to post κ and are matched to a worker with probability $q(\theta_t)$. Then the value of a job is $J_t = \frac{\kappa}{q(\theta_t)}$. Other models consider convex vacancy costs, or as in C. A. Pissarides (2009) or Christiano, Eichenbaum, and Trabandt (2016), fixed costs that are paid after the match has been realized. With the free-entry condition, we can rewrite (16) as

$$\vartheta_t = w_t + \frac{\kappa}{q(\theta_t)} - \beta(1-\bar{\delta})\frac{\kappa}{q(\theta_{t+1})}. \tag{17}$$

The second layer in the economy are retailers that buy the labor services at price ϑ_t, and sell their own variety subject to a downard sloping demand curve and nominal rigidities. If we assume that production is linear $y_t = n_t$, then the marginal cost is the prive of the labor service, so $\lambda_t = \vartheta_t$.[6] Linearized, we obtain equation (11), which we rewrite here to ease of

[6] If that was not the case, marginal costs would be the price of the labor services over their marginal product, but any conclusion with respect to the importance of hiring costs would remain unchanged

Paper	τ	Method/Target
Christiano, Eichenbaum, and Trabandt (2016)	0.994	Free entry
Gertler, Sala, and Trigari (2008)	0.996	Free entry
Blanchard and Galí (2010)	0.989	Hiring costs is 1% of GDP
Krause, Lopez-Salido, and Lubik (2008)	0.95	Labor share and output elasticity of labor
This paper	0.87	Quit elasticity and turnover

Table 1: Implied τ by popular models with search frictions.

exposition:

$$\hat{\lambda}_t = \tau \hat{w}_t + (1-\tau)\hat{\gamma}_t$$

with $\hat{\gamma}_t$ being the net cost of hiring a worker. The value of τ is rarely reported in the aforementioned papers. In most cases, the cost of posting a vacancy is calibrated to target the employment level, instead of being a parameter that can be observed in the data.

What are the costs implied by standard calibrations of models of the labor market? In Christiano, Eichenbaum, and Trabandt (2016) case, the total cost associated with hiring a new worker is roughly 7 percent of their quarterly wage rate. This means that in steady state, $c(\theta) = 0.07w$, but $\tau = \frac{w}{w+(1-\tilde{\beta})c(\theta)}$ takes into the account that a worker hired at t will most likely stay at the firm at $t+1$. Therefore, it would be wrong that hiring costs represent 7% of the cost of hiring a worker because those costs are only paid once. Taking that into account, we get $\tau = 0.994$. The real wage represents virtually all the cost of hiring a worker, the labor market tightness has little effect on marginal costs besides it's effect it has on the wage. Christiano, Eichenbaum, and Trabandt (2016) is not an exotic calibration but rather the norm. In Gertler, Sala, and Trigari (2008), the marginal cost of hiring a worker represents a 3.3% of its quarterly wage, which implies $\tau = 0.996$. As a third example, Blanchard and Galí (2010) calibration implies $\tau = 0.989$. As a general rule, models where the cost of hiring a worker is around 0-10% its quarterly wage will have a hard time getting a τ significantly below one. Taken to the real world, a value of 5% implies that a firm values a worker that earns \$40.000 a year by \$500. It would rather look for another worker than face a one-time-off cost of \$501 to keep the incumbent.

The monopsony view. Proposition 1 presented a way to back up τ. It coincides with the wage markdown, and it provides a formula for it. All we need to know is the turnover rate $\tilde{\delta}$, which is an observable variable, the quit elasticity $\varepsilon_{\tilde{\delta}}$, which has been estimated by many papers, and β, the discount rate, which we assume to be 0.995, 2% annual.[7]

Sokolova and Sorensen (2021) provide a meta-study of 1,320 estimates the elasticity of labor supply from 53 studies. Out of those that compute a separation elasticity, the authors conclude that the best practice estimate for $\epsilon_{\tilde{\delta}}$ is 3.5, although they report a lot of variance across estimates. Together with a quarterly turnover rate of $\tilde{\delta} = 0.12$, it implies that $\tau = 0.87$, which is equivalent to say that workers get 87% of their marginal product, or that the wage represents 87% of the cost of hiring a worker, and the remaining 13% is the net cost of hiring him. This fraction is an order of magnitude larger than the one indirectly calibrated by previous papers.

Is this order of magnitude reasonable? The wage setting condition in steady state is $\omega = \frac{\kappa}{a}\tilde{\delta}2\epsilon_{\tilde{\delta}}$. As in many search models, the definition a vacancy, it's cost and the acceptance rate are objects hard to define and measure. One could interpret a vacancy as a cheap 'now hiring' add, with very low cost and also very low probability that anybody looking at this add ends up being employed. Another interpretation is that a vacancy is a formal job offer after several rounds of interviews and after discarding other candidates. The cost of this job offer is much higher but the likelihood that the candidate accepts the job is also high. Without taking a stance on what v, κ and a are, the term $\frac{\kappa}{a}$ has a clear interpretation: is the willingness to pay for a fully productive new worker. This value equals to 1.2 times the quarterly wage of this worker, or 15 weeks of wage, versus the less than 1 week of wage when we take standard calibration values.

Quantifying hiring costs is challenging. Empirical evidence on how firms recruit employees is still scarce, largely as a result of data limitations. In two papers using the same rich dataset of Swiss workers, Blatter, Muehlemann, and Schenker (2012) and Muehlemann and Strupler (2018) estimate that hiring costs of skilled workers range from 10 to 17 weeks, consistent with the findings reported here. Their dataset consists of a questionnaire of 4032 firms for the 2012 paper and 8874 firms for the 2018 paper administrated by the Swiss Federal Statistical Office and the Centre for Research in Economics of Education at the University of Bern. There, the human resources department filled out a questionnaire answering questions precisely related to the cost of hiring a worker. In particular, they were asked about average advertising costs, time spend in recruiting activities, time for the worker to become fully productive, and training costs spent per hire.[8]

An alternative way to assess how costly is to hire an employee is to observe its market

[7]In case vacancy costs are not linear but can be written as $\frac{\kappa}{1+\nu}\left(\frac{v_t}{(1-\bar{\delta})n_{t-1}}\right)^{1+\nu}(1-\bar{\delta})n_{t-1}$, then the effective discount rate is $\tilde{\beta} = \beta\left(1 - \frac{1}{1+\nu}\tilde{\delta}\right)$, so we should also know the convexity of the vacancy function. However, $\nu > 0$ has a small effect into τ

[8]See Blatter, Muehlemann, and Schenker (2012) Appendix A for the wording of the questionnaire.

Sector	$\epsilon_{\tilde{\delta}}$	$\tilde{\delta}$	τ
Art, accomodation & food	1.20	0.19	0.70
Wholesale, trade & transport	1.39	0.11	0.73
Education and healt	2.15	0.07	0.80
Manufacturing	2.29	0.07	0.81
Prof. business & financial services	3.91	0.12	0.88

Table 2: Separation elasticities from Bassier, Dube, and Naidu (2022), turnover is from the LEHD J2J dataset and implied τ by sector

price. Many firms opt for externalizing its hiring process to staffing firms, whose job is to find ideal candidates for every job opening. Indeed, one of the largest employment websites in the world, reports that external recruiters charge a commission of 15% to 30% of the hired employee's first-year salary, or equivalently from 8 to 16 weeks. Consistent with that the Staffing Industry Analysis survey of 300 North America staffing firms report that their median fee is 20%, or 10.4 weeks. A consideration with these values is that (i) they include a price markup which overestimates the cost of hiring an employee but (ii) they do not include the adaptation and training costs, which underestimates the cost of a new hire. Importantly, these three pieces of direct evidence point to the same order of magnitude as the value obtained using the model.

The model presented here only has one sector, but we can compute the importance of hiring costs by sector if we have estimates of the quit elasticity. Bassier, Dube, and Naidu (2022) provide such estimates for workers in the state of Oregon using LEHD data for 5 aggregated sectors. They estimate separation elasticities by comparing workers with similar work histories that moved to high vs low wage firms, and then compare the quit rate of both workers in the new firms. Using their preferred specification, they find a quit elasticity of 2.1, lower than the 3.5 best estimate proposed by Sokolova and Sorensen (2021). Table (2) shows their estimates by sector, together with the average turnover rates from 2000:2022 taken from the LEHD J2J data set and the implied value of the wage markdown and τ. There is substantial heterogeneity across sectors, and monopsony power is more prevalent in low-wage industries as expected. Despite high turnover rates, the elasticity with respect to the wage in such sectors is low. Professional business, on the other side, gets closer to a perfectly competitive market where workers are elastic to the wage and move often.

While a wage markdown of 70% in the Art, accommodation & food sector is big, but a credible estimate, it is hard to argue that firms devote 30% of their labor costs in to hiring workers. For ease the exposition, we have set up the problem such that vacancy costs are linear in vacancies, which equates marginal costs with average costs. If vacancy costs are convex in the vacancy rate, then the the 30% refers to the marginal cost, while the average costs can be significantly lower. If instead of κv_t we had $\frac{\kappa}{1+\nu}\left(\frac{v_t}{n_{t-1}}\right)^{1+\nu} n_{t-1}$, where ν is a measure of the convexity of the vacancy cost, assuming quadratic costs ($\nu = 1$) would implies that the wage bill represents 83% of the total cost of the workforce, in the most adverse scenario where the wage markdown is 0.7.[9]

5 Quantitative evaluation

Having derived the properties of the model, in this last section we calibrate it and evaluate it numerically. We performing several exercices and extensions, and compare it with a standard search model where wages are bargained and the real wage is rigid as in Blanchard and Galí (2010).

5.1 Calibration

A time period is a quarter. First, we discuss standard parameters. The discount factor is set to $\beta = 0.995$ to target a steady state real rate of two percent. Utility is log and the product market power is set at $\epsilon_p = 6$, which implies a markup of twenty percent over marginal cost. In the previous section, production was linear. Here we allow for decreasing returns to scale with $f(n_t) = An_t^\alpha$, normalizing $A = 1$ and we set $\alpha = 0.7$.

For the labor market, we obtain employment data from FRED and LEHD, using data from 2000:2020. The unemployment rate target over this period is 5.8% and the separations rate is 10.6%. The ratio of employment-to-employment quits relative to total quits is 0.32, which coincides with the ratio of employment-to-employment hires relative to total hires. This means that in steady state, the exogenous separation rate is $\bar{\delta} = 7.2\%$ and the endogenous one is $\delta = 3.7\%$. We set the elasticity of the matching function to 0.5. The relative search effort of the employed worker is obtained from the Job Search Supplement of the Survey of Consumer Expectations by Faberman et al. (2022). The survey allows to directly observe the incidence of offer arrivals by employment status, and employed workers search efficiency is $s = 0.23$. While they document that conditional on searching for a job, employed workers are more likely to receive offers, not all employed workers actively seek for jobs. In this benchmark model, we are assuming that unemployed accept any job offer, which will be relaxed later. Knowing the ratio EE/UE, the search efficiency of employed workers, and the unemployed level, we can infer the probability that employed workers accept a job offer

[9]That is,

$$\frac{wn}{wn + \frac{\kappa}{2}\left(\frac{v}{n}\right)^2 n} = 0.83.$$

Parameter	Definition	Value
β	Discount factor	0.995
ϵ_p	Demand elasticity	6
$\bar{\delta}$	Exogenous turnover	0.072
s	Search intensity	0.23
α	Decreasing returns	0.7
η	Elasticity of the matching function	0.5

Table 3: Parameters used in the model

that pays the same as the current one. That is, $1 - F_\varepsilon(1) = 0.31$. While this model does not feature heterogeneity, this value lines up with Sorkin (2018) evidence that documents that 37% of employment-to-employment transitions see earning declines. Finally, we use the quit elasticity of labor supply to calibrate $f_\varepsilon(1)$, the density of the taste shocks at 1. We use $\varepsilon_{\tilde{\delta}} = 3.5$ as we just argued in the previous section, which implies that the contribution of hiring costs to marginal costs is 13%. Note that despite being a search model, we do not need to specify nor calibrate the unemployment benefits. In principle, a firm deviate by offering very low wages targeting the unemployed workers only. That firm knows that no employed worker will accept its offer and all workers will quit as soon as they receive another job offer fom another firm. If that was the case, the turnover rate of such firm would be 23% . We rule out this case by assuming that there is a minimum wage that prevents this deviation from being profitable, which implies that the derived wage Philips curve corresponds to a global optimum.

For the price rigidity parameter, we match the slope of the price Philips curve estimated by Gagliardone, Gertler, et al. (2023), who provide a mapping from marginal costs to inflation. They estimate this pass-through to be 0.05 using a rich administrative data-set of Belgian firms. This slope implies a coefficient κ_p of 120, which corresponds to a frequency of price adjustment of 5 quarters. This is in the upper end of value for price rigidity, and Gagliardone, Gertler, et al. (2023) get a frequencey of 3.3 quarters because they take into consideration the strategic price-seting behavior obtained from departing from the monopolistic CES case. For the wage rigidity parameter, we set it so for the firm is equally costly to change prices by 1% than it is to raise wages by 1%, which implies $\kappa_w = 122$.

5.2 Alternative model: bargaining with real wage rigidity

The model presented here does not have a clear benchmark to compare, since there is not an established model of search and wage determination. In Section 2 we briefly mention how the Philips wage curve is different from the one that assumes that unions set wages as

Target	Description	Value
u	Unemployment	0.058
$\tilde{\delta}$	Separations Rate	0.10
$\frac{EE}{UE}$	EE-UE ratio	0.32
$\epsilon_{\tilde{\delta}}$	Quit elasticity	3.5
$\frac{\epsilon_p}{\kappa_p}$	Price NKPC slope	0.05

Table 4: Targets

in Erceg, Henderson, and Levin (2000). A fairer comparison of the monopsonistic model is a model that features search frictions but Nash bargaining. We provide a standard model with real wages are rigid as in Blanchard and Galí (2010) or Krause and Lubik (2007), which has been recently used by Gagliardone and Gertler (2023). Since the model is standard, the details are left to the appendix.

Workers receive a flow value from being unemployed b and bargaining wage ς. Firms hire them by posting a vacancy at a cost κ, which meets a worker with probability $q(\theta)$. There is no job search, and workers quit exogenously with probability $\bar{\delta}$. Each worker produces one unit of 'labor services', sold at price ϑ_t. These services are used by a final output firm that uses the same production function $y_t = f(n_t)$ and set prices subject to Rotemberg rigidities.

Wages are Nash-bargained and subject to real rigidities. Let $J_t(\omega_t)$ be the value of a match for a firm and $H_t(\omega_t)$ the worker surplus when the negotiated real wage is ω_t. The wage that would arise in a flexible environment would be

$$\omega_t^{Nash} = \arg\max_{\omega} H_t(\omega)^{\varsigma} J_t(\omega)^{1-\varsigma}.$$

Real wage rigidities are introduced by assuming that the real wage does not fully adjust to the Nash negotiated wage, but instead the real wage is

$$\omega_t = \left(\omega_t^{Nash}\right)^{1-\gamma} \omega^{\gamma}.$$

where ω is the real wage in steady state. Under reasonable parametrizations, this behavior is consistent with rational behavior as it lies within the bargaining set, i.e. it is never above firm's reservation wage (the value to the firm of a worker) nor it is ever below worker's reservation wage (the flow value of unemployment). One way to interpret this wage setting protocol is as the firm providing some insurance to workers by offering a smoother real wage than would be the case under period-by-period Nash bargaining. [10]

[10] See Gertler and Trigari (2009) and Christiano et al. (2016) for formal models of real wage rigidity in a search and matching setting.

This model introduces three new parameters, (b, ς, γ). We target a replacement rate of $b/\omega = 0.7$ as proposed by R. E. Hall and Milgrom (2008). which implies $b = 0.36$. This includes not ony the unemployment benefits but also the leisure utility of not having to work. The literature on Nash bargaining has not converged on the right value for ς, and we set it to 0.5 as Gagliardone and Gertler (2023) to satisfy the Hosios condition. There is no concensus on what this parameter should be, and there is not empirical evidence to guide its calibration. Hagedorn and Manovskii (2008) argues it should be close to zero (0.05), while Gertler, Sala, and Trigari (2008) calibrate it to be close to one (0.9). A clear advantage of the monopsonistic model is that its calibration does not rely on parameters like this bargaining weight with is usually either set exogenously or calibrated but with a difficult interpretation. The wage rigidity parameter is set to 0.9.

The calibration of the bargaining model implies that the value of an employee, $\frac{\kappa}{q(\theta)}$ in steady state represents 11% of the worker wage, or 1.5 weeks of wage. This implies that the contribution of the wage into the cost of hiring a worker is $\tau = 0.993$, in line with the values obtained in Table 1.

5.3 Impulse Response

Having presented both models we compare its response to a demand shock. Assume that the central bank announces a drop of 0.5% on the interest rate on impact that reverts back to the steady state with persistence $\rho_m = 0.8$. This shock can also be interpreted as a drop in the household discount factor, making them more impatient. Figure 5 shows the impulse response of selected macroeconomic variables. The response of employment and inflation in both cases is similar, with the monopsonistic model being slightly more inflationary. The dynamics and drivers of the real wage and wage inflation are significantly different, which we comment below.

In the monopsonistic model, when firms face a demand shock they need to hire workers to satisfy it. They post more vacancies and raise wages to increase the vacancy yield and reduce quits. In partial equilibrium, prices increase by two channels. With higher employment, the marginal product of labor decreases if there are decreasing returns, and marginal vacancy costs increase if vacancy costs are convex. Contrary to standard models, it is not direct that in partial equilibrium raising wages increases marginal costs, because they reduce hiring costs. But that's not the case in general equilibrium, where all firms raising wages nullifies the effect of any individual firm. Moreover, all firms posting more vacancies mean that (i) each vacancy is less likely to meet with an applicant and (ii) more vacancies meet currently employed workers that may quit, which induces further vacancy creation. The net effect is

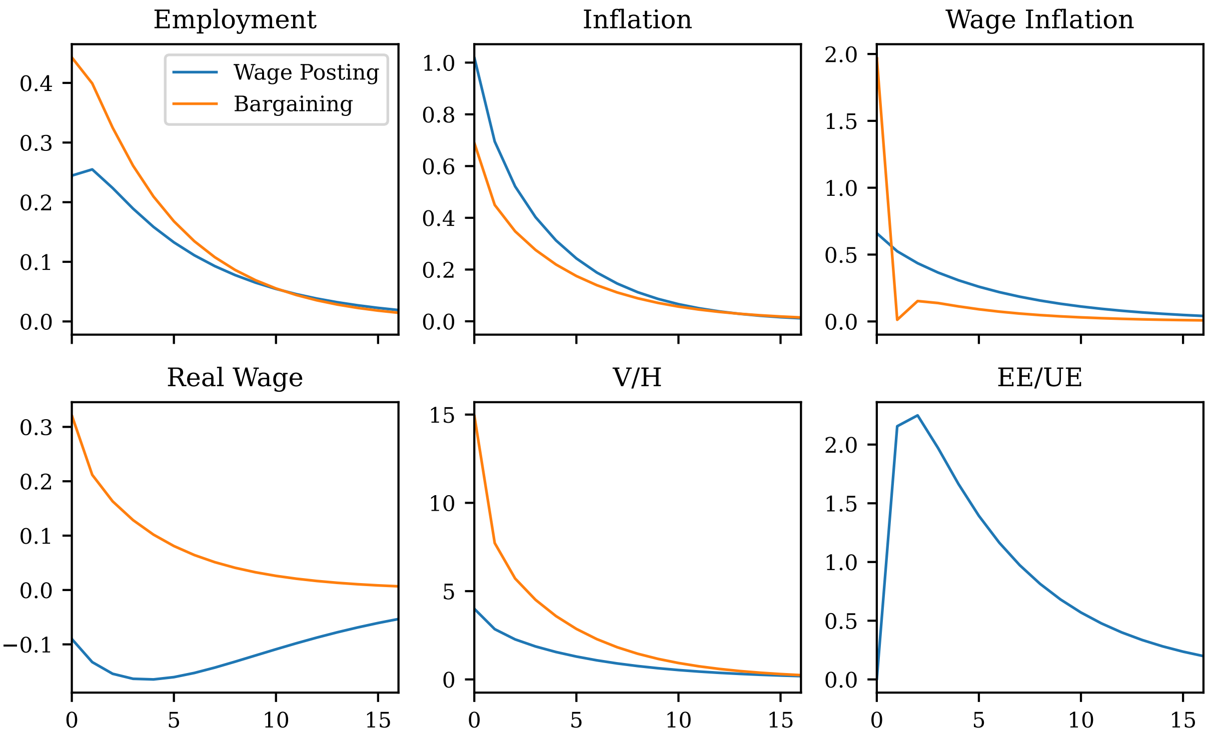

Figure 2: Response to a demand shock. All variables are presented as log deviations from steady state. Price and wage inflation are expressed in annual terms.

nominal wage inflation, more quits and hiring difficulties. Both effects are inflationary and pass it into prices. The effect on the real wage is ambiguous and depends on the relative slopes of the price and wage Phillips curves.

In the bargaining model with a two-layer economy, the mechanism is different. When final good demand rises, demand for labor services also increases. In order to induce vacancy creation, the real price of these services increases. This is equivalent to a productivity shock for the firm that provides labor services. As is general in those cases, the real wage increases, but given the real wage rigidity assumption, not as much as the Nash bargained solution.

In the monopsonistic model, the real wage decreases with the demand shock. Traditional models of nominal wage rigidity like Erceg, Henderson, and Levin (2000) can also share this feature, as exemplified in Lorenzoni and Werning (2023). But the mechanism is different. In models where firms take wages as given, if production has decreasing returns then increasing production raises nominal marginal costs, even if nominal wages do not rise. In Lorenzoni and Werning (2023) or Gagliardone and Gertler (2023), marginal cost raise sharply given the low substitutability between labor and oil. In the model presented here, this channel is also present, but its not the only, nor the main reason, why the real wage decreases. Figure 6 shows the evolution of marginal costs and its main components: the real wage, the marginal

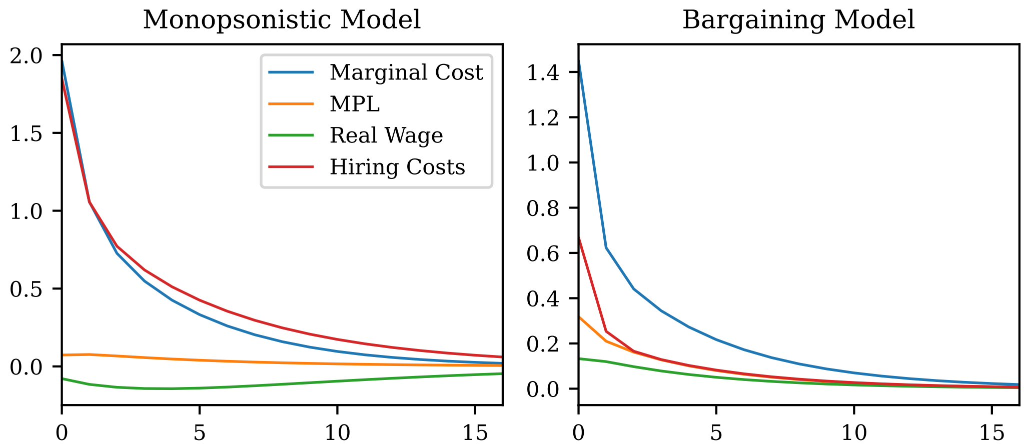

Figure 3: Split of marginal costs. In the monopsonistic model, the driver of inflation is the cost of hiring a worker, not the real wage, which falls.

product of labor, and the hiring costs, for both the monopsonistic model and the bargaining model. In the monopsonistic model, hiring costs are the driver of marginal costs, with a small effect of the decreasing returns to labor offset by a decrease in the real wage. Firms rise prices because they do not find workers, not because they have to pay them more (in real terms). In the bargaining model, the three effects are positive and add up, to have an almost identical response of marginal costs.

5.3.1 Costly Hire

As in Section 3.1, now we consider the case when vacancies are free to post, but hiring is costly. While at the firm level, vacancies are irrelevant for firms, this is not the case at the aggregate level, because aggregate vacancies determine poaching. The calibration is exactly the same as the previous case but now the relevance of hiring costs is amplified because the effective labor supply elasticity is divided by two. Figure 4 shows the response of a demand shock and compares both models.

The model with costly hire significantly mutes the price and wage inflation response to the shock, which implies a larger employment response. This is expected because the cost of hiring a worker is independent of the state of the labor market. C. A. Pissarides (2009) and Christiano, Eichenbaum, and Trabandt (2016) emphasize the role of fixed costs of hiring to generate wage inertia. However, these two papers have a particular formalization of such costs. They are paid *after* the match has been created but *before* the wage negotiation starts. If those costs were paid after the worker has bargained the wage, they would not affect the wage bargaining problem. The model presented here is not subject to that critique.

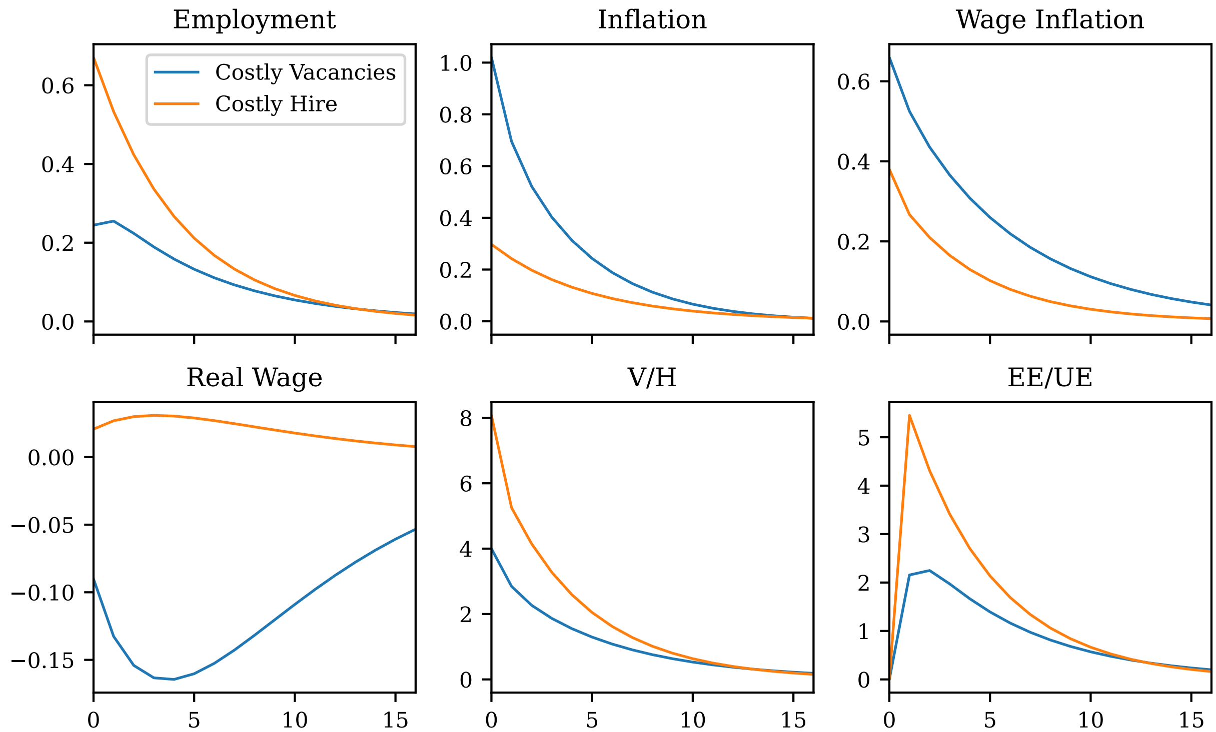

Figure 4: Impulse response of a demand shock comparing the model with costly vacancy costs and the model with costly hire.

If the cost of hiring workers is independent of the labor market, what drives the dynamics of prices and wages? Firms aim to reduce hiring needs by retaining more workers by paying them more. As we see in the bottom right panel, employment-to-employment transitions spike as poaching intensity increases. The increase in nominal wages also implies an increase in prices. The cost of hiring a worker is independent of labor market conditions, but the *net* cost of hiring does depend on it via the quit rate. This adds further pressure in to marginal costs, but is small compared with the effect of the increase in market tightness that operates in the model with costly vacancies. Hiring costs are not enough to overturn the positive real wage response.

5.4 Endogenous Labor supply

For simplicity, it has been assumed that unemployed workers accepted all job offers regardless of the wage. This implied that the aggregate supply of labor is independent of the wage as we see in equation (2). This is the norm in most search models except those that have an endogenous labor participation equation like Graves, Huckfeldt, and Swanson (2023). However, Faberman et al. (2022) document that the unemployed reject 50% of the best offer received over the last month. This could be the result of and heterogeneous distribution of outside options as in Burdett and Mortensen (1998), or as this paper assumes, idiosyncratic taste shocks.

Unemployed workers receive a flow b_t from being unemployed. When offered to work for a wage ω_t, they accept if $\omega_t \varepsilon_t \geq b_t$, the same way workers compare job offers. In the previous section, the symmetry assumption implied that only information about $f_\varepsilon(1)$ and $F_\varepsilon(1)$ was required to define the evolution of the economy, but now we need to make an assumption on the distribution of ε_t. We assume that follows a lognormal distribution with mean μ_ε and standard deviation σ_ε which we calibrate, as well as the value from being unemployed. The aggregate low of motion becomes

$$N_t = (1-\bar{\delta})N_{t-1} + V_t q(\theta_t) p_t^U \left(1 - F_\varepsilon\left(\frac{b_t}{\omega_t}\right)\right).$$

The inclusion of endogenous labor supply adds a new mechanism that was muted in the simplified model. We just saw that a demand shock lowers the real wage. Despite nominal wages being higher, unemployed workers are less willing to accept jobs, which exacerbates labor supply shortages. The appendix shows that the effect is not significative for demand shocks.

Labor supply shocks. The inclusion of an acceptance decision by the unemployed allows us to think about the effect of labor supply shocks, and compare it with the standard bargaining model. Figure 5 shows the response of such shock, that increases the value of unemployment by 5%, and compares it to the model with bargaining and real wage rigidities. While for demand shocks the reaction of employment and inflation were pretty similar, here they are starkly different.

In the monopsonistic model, a drop in the willingness to work by unemployed workers makes hiring more costly. Each vacancy sent is more likely to be turned off which pushes price and nominal wage inflation up, with a negative effect on the real wage. Vacancies take longer to fill and having trouble hiring unemployed workers, firms start competing among them for the employed ones.

In contrast, in a bargaining model with real wage rigidities a labor supply shock has very little effect on employment and inflation. The same would happen if we were considering a model with nominal rigidities like Gertler, Sala, and Trigari (2008) or the Erceg, Henderson, and Levin (2000) unions model. In all these models, the rate at which firms can hire workers is independent of the wage, as long as it is inside the bargaining bands (R. Hall (2005)). The disutility of work only affects the economy trough the wage bargaining condition, and if wages are rigid, then it does not effect the economy at all. While demand shocks behave similarly in the bargaining vs monopsonistic model, the implications for labor supply shocks are very different and help explain the post-covid inflation.

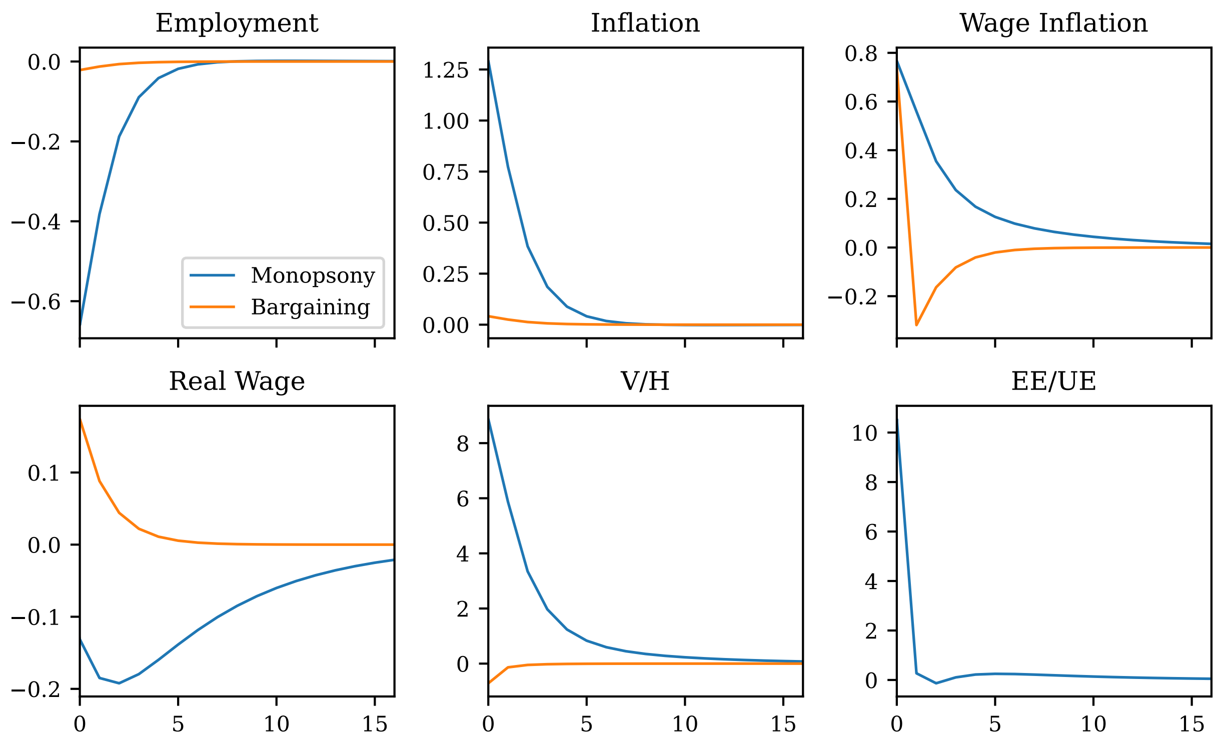

Figure 5: Response to a labor supply shock.

6 Concluding remarks

This paper has presented a model of monopsony in the New Keynesian framework. Compared with alternative models of the labor market, it offers several advantages. It is more realistic and easier to interpret, since firms are the ones setting prices and wages. It does not rely on wage-setting protocols for which we hardly have evidence of. To calibrate the model all we need is the quit elasticity of labor supply, which has been widely studied in the labor economics literature. The implied cost of hiring a worker is more in line with empirical evidence than other models for which it is negligible.

These advantages come at a cost. Solving wage posting models out of steady state is notoriously difficult and we have made some simplifying assumptions to be able to do it analytically. Firms are symmetric, workers are myopic and nominal frictions are à la Rotemberg. Generalizing these simplifying assumptions would not change the main results and message of the model but it would be valuable. It is left for future research. In a steady state environment, de la Barrera (2023) solves a wage posting model with heterogeneous firms and forward-looking workers subject to idiosyncratic taste shocks.

We have highlighted the importance of hiring frictions in determining not only wages but also prices, an element that is not present by construction in models without search frictions and neglected by the calibrations in those that have them. Monopsony increases the importance of those marginal hiring costs and we provided a sufficient statistics for it; the wage markdown. An average worker costs around 15 weeks of its salary, consistent with market estimates.

The theoretical model presented here can inspire several empirical questions that would corroborate the model implications. While the monopsonistic model and alternative models behave similarly when shocked with demand shocks, the price and wage dynamics of supply shocks are significantly different. A challenge is to identify labor supply shocks. In this line, Autor, Dube, and Mcgrew (2023) finds that since the onset of the pandemic, regions were the market were tighter saw a bigger wage increase but also a price increase of the same magnitude.

The main goal of this paper is to rethink the labor market in macroeconomic models. Firms that post wages and prices is a more realistic assumption which comes at its costs. More research on this topic should be done to overcome those and better understand what drives the dynamics of nominal prices, wages and the real wage.

References

Albrecht, James, Carlos Carrillo-Tudela, and Susan Vroman (2018). "On-the-job search with match-specific amenities". In: *Economics Letters* 162, pp. 15–17. ISSN: 01651765. DOI: 10.1016/j.econlet.2017.10.013.

Autor, David, Arindrajit Dube, and Annie Mcgrew (2023). *The Unexpected Compression: Competition at Work in the Low Wage Labor Market.* URL: http://www.nber.org/data-appendix/w31010.

Barrera, Marc de la (2023). *Wage posting and heterogeneous job preference.*

Bassier, Ihsaan, Arindrajit Dube, and Suresh Naidu (Mar. 2022). "Monopsony in Movers: The Elasticity of Labor Supply to Firm Wage Policies". In: *Journal of Human Resources* 57 (2), pp. 1–52.

Berger, David, Kyle Herkenhoff, and Simon Mongey (Apr. 2022a). "Labor Market Powe". In: *American Economic Review* 112 (4), pp. 1147–1193. ISSN: 19447981. DOI: 10.1257/aer.20191521.

Berger, David, Kyle Herkenhoff, and Simon Mongey (Apr. 2022b). "Labor Market Powe". In: *American Economic Review* 112 (4), pp. 1147–1193. ISSN: 19447981. DOI: 10.1257/aer.20191521.

Blanchard, Olivier and Jordi Galí (2010). "Labor markets and monetary policy: A new keynesian model with unemployment". In: *American Economic Journal: Macroeconomics* 2 (2), pp. 1–30. ISSN: 19457715. DOI: 10.1257/mac.2.2.1.

Blatter, Marc, Samuel Muehlemann, and Samuel Schenker (2012). "The costs of hiring skilled workers". In: *European Economic Review* 56 (1), pp. 20–35.

Bloesch, Justin and Birthe Larsen (2023). *When do Firms Profit from Wage Setting Power? New vs. Classical Monopsony.*

Burdett, Kenneth and Dale T. Mortensen (1998). "Wage Differentials , Employer Size , and Unemployment". In: *International Economic Review* 39 (2), pp. 257–273.

Cerrato, Andrea and Giulia Gitti (2022). *Inflation Since COVID: Demand or Supply.*

Christiano, Lawrence J., Martin S. Eichenbaum, and Charles L. Evans (2005). "Nominal rigidities and the dynamic effects of a shock to monetary policy". In: *Journal of Political Economy* 113 (1), pp. 1–45. ISSN: 00223808. DOI: 10.1086/426038.

Christiano, Lawrence J., Martin S. Eichenbaum, and Mathias Trabandt (2016). "Unemployment and Business Cycles". In: *Econometrica* 84 (4), pp. 1523–1569. ISSN: 0012-9682. DOI: 10.3982/ecta11776.

Coles, Melvyn G. and Dale T. Mortensen (2016). "Equilibrium Labor Turnover, Firm Growth, and Unemployment". In: *Econometrica* 84 (1), pp. 347–363. ISSN: 0012-9682. DOI: 10.3982/ecta10700. URL: https://doi.org/10.3982/ECTA10700.

Erceg, Christopher J., Dale W. Henderson, and Andrew T. Levin (2000). "Optimal monetary policy with staggered wage and price contracts". In: *Journal of Monetary Economics* 46, pp. 281–313.

Faberman, Jason et al. (2022). "Job Search Behavior among the Employed and Non-Employed". In: *Econometrica* 90 (4), pp. 1743–1779. ISSN: 0898-2937. DOI: 10.3386/w23731.

Gagliardone, Luca and Mark Gertler (2023). "Oil Prices, Monetary Policy and Inflation Surges". In: *NBER Working Paper*.

Gagliardone, Luca, Mark Gertler, et al. (2023). "Anatomy of the Phillips Curve: Micro Evidence and Macro Implications". In: *NBER Working Paper*. URL: http://www.nber.org/papers/w31382.

Galí, Jordi (2011). *Unemployment Fluctuations and Stabilization Policies*. Ed. by Karl Gunnar. Zeuthen Lecture Book Series. ISBN: 9781626239777.

Gertler, Mark, Luca Sala, and Antonella Trigari (2008). "An Estimated Monetary DSGE Model with Unemployment and Staggered Nominal Wage Bargaining". In: *Journal of*

Money, Credit and Banking 40 (8), pp. 1713–1764. ISSN: 00222879. DOI: 10.1111/j.1538-4616.2008.00180.x.

Graves, Sebastian, Christopher Huckfeldt, and Eric T Swanson (2023). *The Labor Demand and Labor Supply Channels of Monetary Policy.*

Grigsby, John, Erik Hurst, and Ahu Yildirmaz (2021). “Aggregate Nominal Wage Adjustments: New Evidence from Administrative Payroll Data”. In: *American Economic Review* 111 (2), pp. 428–471. ISSN: 19447981. DOI: 10.1257/AER.20190318.

Hagedorn, Marcus and Iourii Manovskii (Sept. 2008). “The cyclical behavior of equilibrium unemployment and vacancies revisited”. In: *American Economic Review* 98 (4), pp. 1692–1706. ISSN: 00028282. DOI: 10.1257/aer.98.4.1692.

Hall and Alan B. Krueger (2012). “Evidence on the incidence of wage posting, wage bargaining, and on-the-job search”. In: *American Economic Journal: Macroeconomics* 4 (4), pp. 56–67. ISSN: 19457707. DOI: 10.1257/mac.4.4.56.

Hall, Robert (2005). “Employment fluctuations with equilibrium wage stickiness”. In: *American Economic Review* 95 (1), pp. 50–65. ISSN: 00028282. DOI: 10.1257/0002828053828482.

Hall, Robert E and Paul R Milgrom (2008). “The Limited Influence of Unemployment on the Wage Bargain”. In: 6010, pp. 1653–1674.

Huo, Zhen and José Víctor Ríos-Rull (2020). “Sticky wage models and labor supply constraints”. In: *American Economic Journal: Macroeconomics* 12 (3), pp. 284–318. ISSN: 19457715. DOI: 10.1257/mac.20180290.

Krause, Michael U., David Lopez-Salido, and Thomas A. Lubik (2008). “Inflation dynamics with search frictions: A structural econometric analysis”. In: *Journal of Monetary Economics* 55 (5), pp. 892–916. ISSN: 03043932. DOI: 10.1016/j.jmoneco.2008.04.004.

Krause, Michael U. and Thomas A. Lubik (2007). “The (ir)relevance of real wage rigidity in the New Keynesian model with search frictions”. In: *Journal of Monetary Economics* 54 (3), pp. 706–727. ISSN: 03043932. DOI: 10.1016/j.jmoneco.2005.12.001.

Lorenzoni, Guido and Iván Werning (2023). “Wage Price Spirals”. In.

Manning, Alan (2003). *Monopsony in motion : imperfect competition in labor markets.* ISBN: 0691113122.

— (2006). “A Generalized Model of Monopsony”. In: *Economic Journal* 116, pp. 84–100.

— (2021). “Monopsony in Labor Markets: A Review”. In: *ILR Review* 74 (1), pp. 3–26. ISSN: 2162271X. DOI: 10.1177/0019793920922499. URL: https://doi.org/10.1177/0019793920922499.

Moscarini, Giuseppe and Fabien Postel-Vinay (2023). *The Job Ladder: Inflation vs. Reallocation.*

Muehlemann, Samuel and Mirjam Strupler (2018). "Hiring costs and labor market tightness". In: *Labour Economics* 52 (September 2016), pp. 122–131. ISSN: 0927-5371. DOI: 10.1016/j.labeco.2018.04.010. URL: https://doi.org/10.1016/j.labeco.2018.04.010.

Pissarides, Christopher (2017). *Equilibrium Unemployment Theory*. second. The MIT Press.

Pissarides, Christopher A (2009). "The Unemployment Volatility Puzzle: Is Wage Stickiness the Answer?" In: *Econometrica* 77 (5), pp. 1339–1369. ISSN: 0012-9682. DOI: 10.3982/ECTA7562. URL: https://doi.org/10.3982/ECTA7562.

Robinson, Joan (1933). *The Economics of Imperfect Competition*. ISBN: 9780333102893.

Salop, Steven C. (1979). "A Model of the Natural Rate of Unemployment". In: *American Economic Review* 69 (1), pp. 117–125.

Schmitt-Grohé, Stephanie and Martín Uribe (2005). "Optimal Fiscal and Monetary Policy in a Medium-Scale Macroeconomic Model". In: *Annual* 20, pp. 383–425.

Silva, José Ignacio and Manuel Toledo (2009). "Labor turnover costs and the cyclical behavior of vacancies and unemployment". In: *Macroeconomic Dynamics* 13 (SUPPL. S1), pp. 76–96. ISSN: 13651005. DOI: 10.1017/S1365100509080122.

Sokolova, Anna and Todd Sorensen (2021). "Monopsony in Labor Markets: A Meta-Analysis". In: 74 (January), pp. 27–55. DOI: 10.1177/0019793920965562. URL: https://doi.org/10.1177/0019793920965562.

Sorkin, Isaac (2018). "Ranking Firms Using Revealed Preference". In: *Quarterly Journal of Economics*, pp. 1331–1393. DOI: 10.1093/qje/qjy001.Advance.

Thomas, Carlos (2011). "Search Frictions, Real Rigidities, and Inflation Dynamics". In: *Journal of Money, Credit and Banking* 43 (6), pp. 1131–1164. ISSN: 00222879. DOI: 10.1111/j.1538-4616.2011.00420.x.

A The firm problem

A.1 Non-linear problem

We start with the firm problem, allowing for a general production function $y_t = f(n_t)$

$$J_t(x_{t-1}) = \max_{p_t, w_t, v_t} \frac{p_t}{P_t}\mathcal{D}_t\left(\frac{p_t}{P_t}\right) - \frac{w_t}{P_t}n_t - \kappa_v v_t - \frac{\kappa_p}{2}\left(\frac{p_t}{p_{t-1}} - 1\right)^2 Y_t - \frac{\kappa_w}{2}\left(\frac{w_t}{w_{t-1}} - 1\right)^2 N_t + E_t[\Lambda_{t,t+1}J_{t+1}(x_t)]$$

subject to:

$$n_t = \left(1 - \tilde{\delta}_t\left(\frac{w_t}{W_t}\right)\right)n_{t-1} + a_t\left(\frac{w_t}{W_t}\right)v_t$$

$$y_t = f(n_t)$$

Let μ_t be the Lagrange multiplier of the law of motion for employment and λ_t be the Lagrange multiplier of the constraint on the production function. The former is the value of a worker and the later the real marginal cost. Taking a first order condition with respect to v_t we get:

$$\mu_t = \frac{\kappa_v}{a_t\left(\frac{w_t}{W_t}\right)}$$

The first order condition with respect to the wage is:

$$\frac{1}{P_t}n_t + \kappa_w\left(\frac{w_t}{w_{t-1}} - 1\right)N_t\frac{1}{w_{t-1}} - \kappa_w E_t\left[\Lambda_{t,t+1}\left(\frac{w_{t+1}}{w_t} - 1\right)\frac{w_{t+1}}{w_t^2}N_{t+1}\right]$$
$$= \mu_t\left(-\tilde{\delta}'_t\left(\frac{w_t}{W_t}\right)\frac{1}{W_t}n_{t-1} + a'_t\left(\frac{w_t}{W_t}\right)v_t\right)$$

Multiply the equation by W_t, divide it by N_t, and apply symmetry so $w_t = W_t$ and $n_t = N_t$ to obtain:

$$\omega_t + \kappa_w\left(\Pi_t^w - 1\right)\Pi_t^w - \kappa_w E_t\left[\Lambda_{t,t+1}\left(\Pi_{t+1}^w - 1\right)\Pi_{t+1}^w\frac{N_{t+1}}{N_t}\right] = \frac{\kappa_v}{a_t(1)}\left(-\tilde{\delta}'_t(1)\frac{N_{t-1}}{N_t} + a'_t(1)\frac{V_t}{N_t}\right)$$

Plug $\delta'_t(1)$ and $a'_t(1)$

$$\omega_t + \kappa_w\left(\Pi_t^w - 1\right)\Pi_t^w - \kappa_w E_t\left[\Lambda_{t,t+1}\left(\Pi_{t+1}^w - 1\right)\Pi_{t+1}^w\frac{N_{t+1}}{N_t}\right] = \frac{\kappa_v}{a_t(1)}\left(\frac{V_t}{N_t}q(\theta_t)p_t^E f_\varepsilon(1) + q(\theta_t)f_\varepsilon(1)\frac{V_t}{N_t}\right)$$

Rearrange to get the non-linear wage Philips curve:

$$(\Pi_t^w - 1)\Pi_t^w = \frac{1}{\kappa_w}\left(2\frac{\kappa_v}{a_t(1)}\frac{V_t}{N_t}q(\theta_t)p_t^E f_\varepsilon(1) - \omega_t\right) + E_t\left[\Lambda_{t,t+1}\left(\Pi_{t+1}^w - 1\right)\Pi_{t+1}^w\frac{N_{t+1}}{N_t}\right] \tag{18}$$

Now, we can get the acceptance elasticity by noting that:

$$\frac{q(\theta_t)p_t^E f_\varepsilon(1)}{a_t(1)} = \frac{a'_t(1)\frac{1}{W}}{a_t(1)}W = \epsilon_{a,t}$$

to relate it to the quit elasticity, we can use the fact that in steady state, $a(1)V = \tilde{\delta}(1)N$

$$\frac{q(\theta)p^E f_\varepsilon(1)}{a(1)} = \frac{\frac{q(\theta)p^E f_\varepsilon(1)V}{N}}{\tilde{\delta}(1)} = \frac{\tilde{\delta}'(1)\frac{1}{W}}{\tilde{\delta}}W = \epsilon_{\tilde{\delta}}$$

Out of the steady state, both elasticities do not coincide. Instead, we have that:

$$\epsilon_{\tilde{\delta},t} \equiv \frac{\frac{V_t q(\theta_t) p_t^E f_\varepsilon(1)}{N_{t-1}}}{\tilde{\delta}_t(1)} = \frac{V_t q(\theta_t) p_t^E f_\varepsilon(1)}{a_t(1) V_t \left(1 - \frac{N_t - N_{t-1}}{N_t - (1-\tilde{\delta}_t(1)) N_{t-1}}\right)} = \left(1 + \frac{1}{\tilde{\delta}_t(1)} \frac{N_t - N_{t-1}}{N_{t-1}}\right) \epsilon_{t,a}$$

With that, we get that:

$$(\Pi_t^w - 1)\Pi_t^w = \frac{1}{\kappa_w}\left(2\kappa_v \epsilon_{a,t} \frac{V_t}{N_t} - \omega_t\right) + E_t\left[\Lambda_{t,t+1}\left(\Pi_{t+1}^w - 1\right)\Pi_{t+1}^w \frac{N_{t+1}}{N_t}\right]$$

The first order condition for prices is standard for a model with Rotemberg rigidities.

$$\frac{1}{P_t}\mathcal{D}_t\left(\frac{p_t}{P_t}\right) + \frac{p_t}{P_t}\mathcal{D}_t'\left(\frac{p_t}{P_t}\right)\frac{1}{P_t} - \kappa_p\left(\frac{p_t}{p_{t-1}} - 1\right)\frac{1}{p_{t-1}} Y_t +$$
$$E_t\left[\Lambda_{t,t+1}\kappa_p\left(\frac{p_{t+1}}{p_t} - 1\right)\frac{p_{t+1}}{p_t^2} Y_{t+1}\right] - \lambda_t \mathcal{D}_t'\left(\frac{p_t}{P_t}\right)\frac{1}{P_t} = 0$$

Multiply by $\frac{P_t}{Y_t}$ and apply symmetry:

$$1 - \epsilon_p - \kappa_p\left(\Pi_t^p - 1\right)\Pi_t^p + E_t\left[\Lambda_{t,t+1}\kappa_p\left(\Pi_{t+1}^p - 1\right)\Pi_{t+1}^p \frac{Y_{t+1}}{Y_t}\right] + \epsilon_p \lambda_t = 0$$

And rearrange to get equation (10)

$$(\Pi_t^p - 1)\Pi_t^p = \frac{1}{\kappa_p}(1 - \epsilon_p + \epsilon_p \lambda_t) + E_t\left[\Lambda_{t+1}(\Pi_{t+1}^p - 1)\Pi_{t+1}^p \frac{Y_{t+1}}{Y_t}\right]$$

Finally λ_t, marginal costs, are obtained from the first order condition with respect to n_t.

$$-\frac{w_t}{P_t} - \mu_t + \lambda_t f'(n_t) + E_t\left[\Lambda_{t,t+1}\left(1 - \tilde{\delta}_{t+1}\left(\frac{w_{t+1}}{W_{t+1}}\right)\right)\mu_{t+1}\right]$$

and rearranging we get that marginal costs are:

$$\lambda_t = \frac{\omega_t + \frac{\kappa_v}{a_t(1)} - \Lambda_{t,t+1}(1 - \tilde{\delta}_t(1))\frac{\kappa_v}{a_{t+1}(1)}}{f'(n_t)}$$

A.2 Steady state

We drop the t subindex to denote steady state variables. The pricing equation in steady state is:

$$1 = \mathcal{M}_p \frac{\omega + (1 - \beta(1 - \tilde{\delta}))\frac{\kappa_v}{a}}{f'(n)}$$

The wage setting condition, using the fact that in steady state $\epsilon_a = \epsilon_{\tilde{\delta}}$, and $\frac{V}{N} = \frac{\tilde{\delta}}{a}$ we get:

$$\omega = \frac{\kappa_v}{a} 2\tilde{\delta}\epsilon_{\tilde{\delta}}$$

Combining both equations we get:

$$1 = \mathcal{M}_p \frac{\omega + (1 - \beta(1 - \tilde{\delta}))\frac{1}{2\tilde{\delta}\epsilon_{\delta}}\omega}{f'(n)}$$

and the wage is:

$$\omega = \frac{1}{\mathcal{M}_p} \frac{2\tilde{\delta}\epsilon_{\tilde{\delta}}}{2\tilde{\delta}\epsilon_{\tilde{\delta}} + 1 - \tilde{\beta}} f'(n)$$

The real wage is below the marginal product of labor for two reasons. Market power adds a markup over marginal costs, and search frictions add a wage markdown $\mathcal{M}_w = \frac{2\tilde{\delta}\epsilon_{\tilde{\delta}}}{2\tilde{\delta}\epsilon_{\tilde{\delta}}+1-\tilde{\beta}}$. As proposition 1 shows, this markdown also coincides with the wage share over the total cost of hiring a worker:

$$\tau \equiv \frac{\omega}{\omega + (1 - \tilde{\beta})\frac{\kappa_v}{a}} = \frac{2\tilde{\delta}\epsilon_{\tilde{\delta}}}{2\tilde{\delta}\epsilon_{\tilde{\delta}} + 1 - \tilde{\beta}}$$

A.3 Linearizing the Philips cuves

Here we derive Proposition 2. We start with the wage Philips curve. First we start noting that the term inside the parenthesis in equation 18 can be written as:

$$\frac{\kappa_v}{a_t(1)} \frac{V_t}{N_t} q(\theta_t) 2 p_t^E f_\varepsilon(1) = \frac{\kappa_v}{a_t(1)} \frac{N_t - (1 - \bar{\delta})N_{t-1}}{N_t} 2 \frac{p_t^E}{p_t^U} f_\varepsilon(1)$$

after using the law of motion of aggregate employment $N_t = (1 - \bar{\delta})N_{t-1} + V_t q(\theta_t) p_t^U$ and substituting $V_t q(\theta_t)$. Also, we have that:

$$\frac{EE_t}{UE_t} = \frac{V_t q(\theta_t) p_t^E (1 - F_\varepsilon(1))}{V_t q_t p_t^U} = \frac{p_t^E}{p_t^U}(1 - F_\varepsilon(1)) \tag{19}$$

and this $\hat{p}_t^E - \hat{p}_t^U = \hat{EE}_t - \hat{UE}_t$ and $\hat{a}_t(1) = \hat{H}_t - \hat{V}_t$. With that, we get the wage Philips curve linearized and expressed as a function of labor market variables. The time to fill $\hat{V}_t - \hat{H}_t$, employment growth, the $\frac{EE}{UE}$ ratio, and the deviation from the real wage.

$$\pi_t^w = \frac{\omega}{\kappa_w} \left(\hat{V}_t - \hat{H}_t + \frac{1 - \bar{\delta}}{\bar{\delta}} \Delta \hat{N}_t + \hat{EE}_t - \hat{UE}_t - \hat{\omega}_t \right) + \beta E_t \pi_{t+1}^w$$

and using that $\omega = \frac{\tau}{\mathcal{M}_p}$ we can express the slope of the curve as a function of product and labor market power.

$$\pi_t^w = \frac{1}{\kappa_w}\frac{\tau}{\mathcal{M}_p}\left(\hat{V}_t - \hat{H}_t + \frac{1-\bar{\delta}}{\bar{\delta}}\Delta\hat{N}_t + \hat{EE}_t - \hat{UE}_t - \hat{\omega}_t\right) + \beta E_t \pi_{t+1}^w$$

The price Philips curve expressed is:

$$\pi_t^p = \frac{\epsilon_p}{\kappa_p}\hat{\lambda}_t + \beta E_t[\pi_{t+1}^p]$$

Linearizing λ_t:

$$\hat{\lambda}_t = \tau\hat{\omega}_t + (1-\tau)\left(\frac{1}{1-\tilde{\beta}}(\hat{V}_t - \hat{H}_t) - \frac{\tilde{\beta}}{1-\tilde{\beta}}\left(\hat{V}_{t+1} - \hat{H}_{t+1} - \frac{\tilde{\delta}}{1-\tilde{\delta}}\hat{\tilde{\delta}}_{t+1}\right)\right)$$

And we get the Equation 13

$$\pi_t^p = \frac{\epsilon_p}{\kappa_p}\left(\tau\hat{\omega}_t + (1-\tau)\left(\frac{1}{1-\tilde{\beta}}(\hat{V}_t - \hat{H}_t) - \frac{\tilde{\beta}}{1-\tilde{\beta}}\left(\hat{V}_{t+1} - \hat{H}_{t+1} - \frac{\tilde{\delta}}{1-\tilde{\delta}}\hat{\tilde{\delta}}_{t+1}\right)\right)\right) + \beta E_t[\pi_{t+1}^p]$$

B Bargaining Model

The household block is the same and the labor market does not have on the job search, so the quit rate is $\bar{\delta}$. There are two layers in the economy. Price setters buy 'labor services' at price ϑ_t and produce a differentiated good which price under Rotemberg rigiditites.

Firms The value of a filled job is given by:

$$J_t = \vartheta_t - \omega_t + (1-\bar{\delta})E_t[\Lambda_{t,t+1}J_{t+1}]$$

To post a vacancy, firms pay κ_v, and vacancies are filled with probability $q(\theta_t)$. Once filled, the worker starts producing at t.

$$V_t = -\kappa_v + q(\theta_t)J_t + E_t[\Lambda_{t,t+1}V_{t+1}]$$

The free entry condition sets $V_t = 0$, and therefore the value of a filled job is:

$$J_t = \frac{\kappa_v}{q(\theta_t)}$$

Workers

If a firm finds a worker with probability $q(\theta_t)$, then workers meet firms with probability $\theta_t q(\theta_t)$

$$V_t^u = b_t + E_t\left[\Lambda_{t,t+1}\left(\theta_t q(\theta_t) V_{t+1}^e + (1-\theta_{t+1}q(\theta_{t+1}))V_t^u\right)\right]$$

$$V_t^e = \omega_t + E_t[\Lambda_{t,t+1}(\bar{\delta}V_t^u + (1-\bar{\delta})V_t^e]$$

Let ς be the Nash bargaining weight of workers. The real wage is chosen to maximize:

$$\max_{\omega^{Nash}} (V_t^e - V_t^u)^{\varsigma} J_t^{1-\varsigma}$$

the first order condition is:

$$\frac{J_t}{V_t^e - V_t^u} = \frac{1-\varsigma}{\varsigma}$$

And using the expressions for J_t, V_t^u and V_t^e we get:

$$\omega_t^{Nash} = \varsigma\vartheta_t + (1-\varsigma)b_t + \kappa E_t\left[\Lambda_{t,t+1}\theta_{t+1}\right]$$

The free entry condition is:

$$\frac{\kappa}{q(\theta_t)} = \vartheta_t - \omega_t + E_t\Lambda_{t,t+1}(1-\bar{\delta})\frac{\kappa}{q(\theta_{t+1})}$$

C Labor supply and demand shocks

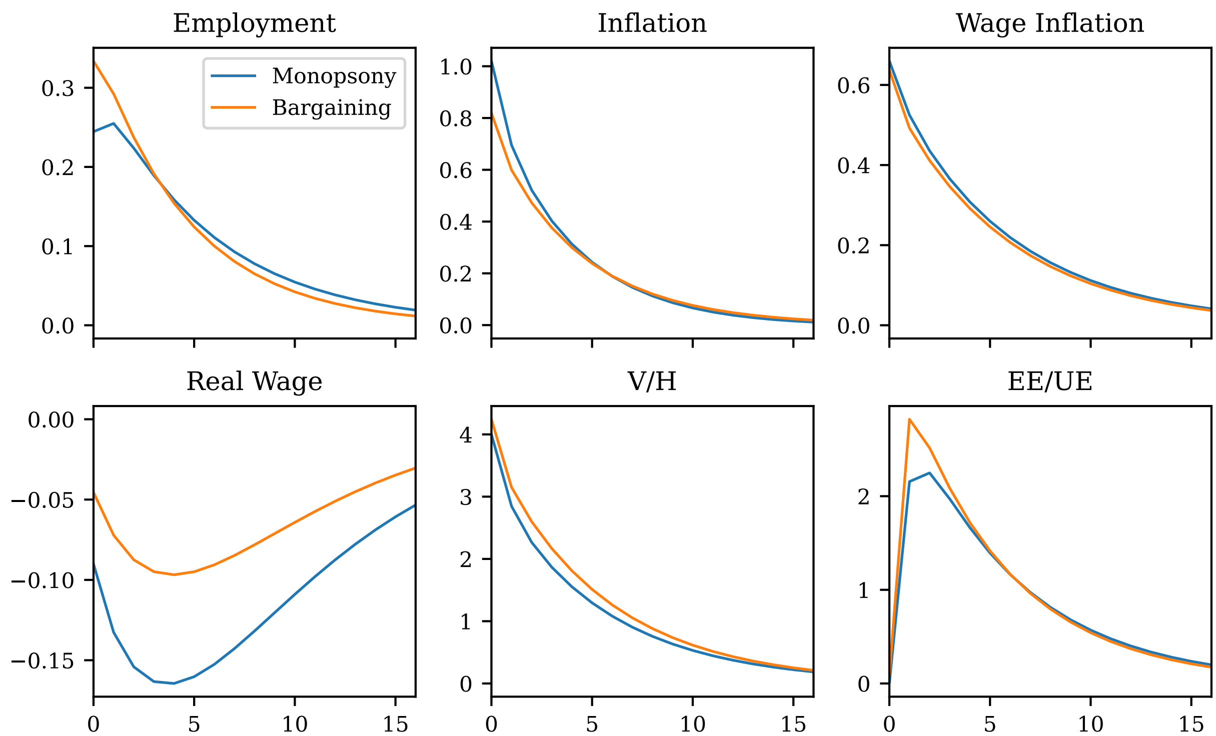

Figure 6: Difference between the benchmark model and the model that includes endogenous labor supply.