Inefficient Unemployment and Bargaining Friction*

C. Y. Kelvin Yuen†

Washington University in St. Louis

October 14, 2019

Abstract

This paper explores the possibility of privately inefficient job separations due to bargaining friction and its implications for the unemployment dynamics. I propose a simple specification of bargaining friction by including bargaining wedges in the standard Nash bargaining model. Such bargaining wedge arises when, for example, wages are determined by alternating offers bargaining, which is often used in the literature to generate real wage rigidity, or when there is asymmetric information about worker's productivity. I show that due to the misalignment between actual surpluses and bargaining surpluses, inefficient separations could be generated which would in turn induce inefficient unemployment. The existence of inefficient unemployment due to bargaining friction could potentially explain the excessive fluctuation of unemployment observed in the data. Quantitatively, I find that inefficient unemployment constitutes up to 54% of the total unemployment volatility in the calibrated model.

JEL classification: C78, E24, E32, J63, J64

Keywords: inefficient unemployment, bargaining friction, inefficient separation, unemployment volatility, alternating offers bargaining, asymmetric information

---

*I would like to thank Rody Manuelli, Yongseok Shin, Ping Wang for their comments and advice. I am also grateful to the participants at Midwest Macroeconomics Meetings in Madison, WI, Asian Meeting of the Econometric Society in Seoul, Korea, and internal seminars at WashU for helpful comments and discussions.

†Department of Economics, Washington University in St. Louis, One Brookings Drive, St. Louis, MO 63130. Email: kelvinyuen@wustl.edu

1 Introduction

How would the possibility of inefficient job separation affect the labor market dynamics? A job separation is efficient if an employment relationship is dissolved whenever the joint surplus from matching is negative[1]. If wages are fully flexible, then any inefficient separation can be avoided by wage renegotiation. In fact, this has been the thesis of the efficient-turnover literature. McLaughlin (1991), for instance, shows that the quit-layoff distinction is inconsequential when all separations are efficient. However, such Pareto improvement may be infeasible when there are labor market frictions. There is a literature showing that when the wage rate is somehow rigid to changes in the environment, inefficient job separations may exist. Hall and Lazear (1984), for example, show that inefficient job turnovers necessarily arise from the existence of asymmetric information about the productivity and the outside option of the worker, resulting in excessive layoffs and quits. Also, inefficient separations may result when renegotiation is costly (Antel, 1985) and when there exists firm-specific capital (Becker, 1962) so that the worker would demand a higher wage to reap the return of the investment. Empirically, Gielen and van Ours (2006) show that while inefficient quits are insignificant, almost half of all predicted layoffs are inefficient. They argue that the substantial amount of inefficient layoffs may be due to downward wage rigidity. More recently, by exploiting the quasi-experiment of a temporary unemployment insurance extension in Austria, Jäger et al. (2019) estimates that a vast majority of job separations are inefficient.

Standard search and matching models á la Mortensen and Pissarides (1994, henceforth MP) with endogenous separation have assumed that wages are determined by Nash bargaining. One implication of the flexible wage rule is that job turnovers are always privately efficient and mutually agreed. Hence, there is no involuntary layoffs and quits. As pointed out by the seminal critique by Shimer (2005), such models fail to explain the unemployment volatility observed in the data since wages are too flexible in the model. Subsequently, researchers have attempted to generate a larger response of unemployment to productivity change by incorporating real wage stickiness (e.g. Hall, 2005c and Gertler and Trigari, 2009). However, they largely maintain the assumption, without justification, that all employment pairs are efficient.

Recently, replacing Nash bargaining with alternating offers bargaining (AOB) in the otherwise

[1] In this paper, I focus on the private efficiency of job separations. There has been a large literature focusing on the social efficiency of search and matching models, see, e.g., Hosios (1990).

standard search and matching model, first proposed by Hall and Milgrom (2008), has achieved some success in explaining labor market dynamics. Due to the different bargaining protocol, wages become less sensitive to the labor productivity, generating real wage rigidity endogenously. In fact, Christinano et al. (2016) show that the AOB model outperforms both Nash bargaining and Calvo sticky wage models in matching the dynamic responses to macroeconomic shocks in the data[2]. While the real wage rigidity derived in the model can potentially produce inefficient job turnovers, the literature is mostly silent about the possibility of separation, and the efficiency thereof, in this class of models.

I construct an otherwise standard MP model with bargaining friction, where wages are determined by Nash bargaining with the existence of wedges between the actual surplus and the respective bargaining surplus. Such bargaining wedges can be considered as a reduced-form specification of some bargaining friction. In this model, workers and firms would still separate efficiently in response to a sufficiently low match-specific productivity, leading to a negative joint match surplus. However, I show that due to the misalignment between actual surpluses and the bargaining surpluses, inefficient separations could be generated when the match-specific productivity lies in certain range. Hence, there exists an interval of the productivity shock where inefficiency may arise. The inefficiency gap depends positively on the *net* bargaining wedge between the worker and the firm. If the net wedge is positive (resp. negative), there may be inefficient layoffs (resp. quits). As a result of the excessive job separations, inefficient unemployment may also arise.

To justify the bargaining wedge specification, I show that some bargaining protocols used in the literature to generate higher unemployment volatility belong to the same class. For example, the solution to the intra-period version of AOB model by Hall and Milgrom (2008), also adopted by Christinano et al. (2016), can be considered as a micro-foundation of the model considered in this paper. Here the inefficiency comes from the possibility of delay to reach an agreement, and the existence of flow benefit of the worker and flow cost of the firm during bargaining. Hall and Milgrom (2008) distinguish between outside-option payoff and disagreement payoff. This view is consistent with the existence of bargaining wedges when the actual surplus is different from the bargaining surplus. In this case, the net bargaining wedge is positive, resulting in some inefficient

[2] Christinano et al. (2016) show that while the Nash bargaining and the AOB models can produce similar likelihood functions with respect to matching the empirical impulse response functions, the Nash model requires implausibly high replacement ratio, as in Hagedorn and Manovskii (2008).

layoffs.

Such bargaining wedges may also arise when there is asymmetric information about the match-specific productivity. Kennan (2010) applies the neutral bargaining solution (NBS) of Myerson (1984) when the firm has private information about the productivity of the worker. However, he assumes the worker would always propose a pooling offer, resulting in efficient employment. By allowing the worker to choose optimally also the screening offer, there may be inefficient job separations. I show that in this case, the NBS is equivalent to the Nash bargaining solution when there is a positive net bargaining wedge and the firm has all the bargaining power due to private information.

It is worth noting that while I present Hall and Milgrom (2008) and Kennan (2010) as micro-foundations of the reduced-form bargaining wedge, the purpose of introducing the bargaining wedge specification goes beyond mere generalization of those bargaining models. In fact, by allowing endogenous and potentially inefficient separations, the existence of bargaining friction has important and novel implications for the volatilities of job separations and unemployment. First, there is the *small surplus channel*, as emphasized by Ljungqvist and Sargent (2017). The net bargaining wedge would effectively reduce the fundamental surplus if it is positive. As a result, a relatively small change in the labor productivity could generate large change in the job creation incentive, and thus the market tightness. I have shown also that if the net bargaining wedge is positive and relatively acyclical, then the worker's share of the match surplus will be countercyclical. This entails that wages are less sensitive to the aggregate conditions, as in Hall and Milgrom (2008). By allowing inefficient job separations I have unleashed a different and novel channel: the *separation channel*. While the workers' productivity varies with the aggregate productivity y_t, the bargaining wedge is relatively acyclical. As a result, a small change in y_t causes significant fluctuations in the cutoff productivity, and thus in the quantity of both efficient and inefficient job separations. The additional movement in job separations, coupled with the reduction of job creation incentive in recessions, generates excessive unemployment volatility.

Quantitatively, bargaining friction has large impact on the volatility of the unemployment rate. Specifically, I show that the unemployment volatility in the economy is increasing with the net bargaining wedge. As a result, by including a relatively small amount of bargaining friction, the model can generate a large labor market fluctuation without relying on unrealistic calibration as in Hagedorn and Manovskii (2008). By contrast, a model with homogenous match-specific productivity and exogenous separations is much less sensitive to bargaining friction. I show that in the calibrated model, inefficient unemployment constitutes 54% of the total unemployment volatility. Moreover, the peak of the unemployment rate in the U.S. during the Great Recession would have been 3 percentage points lower if all separations were efficient. Lastly, I find that while the existence of bargaining friction has little effect on job finding rate, it increases substantially the volatility of the separation rate, and that of the transition rates between unemployment and employment states. This shows that the model does not need the small surplus assumed by much of the literature, which mainly operates through the job finding rate, to generate sizable unemployment volatility.

This paper focuses on the job separations in the labor market[3]. The literature has been divided on whether job separation rate contributes to a significant fraction of the unemployment volatility. Hall (2005a) and Shimer (2012) argue that the job separation rate into unemployment is roughly acyclical over the business cycle. On the other hand, Elsby et al. (2009), Fujita and Ramey (2009), and Pissarides (2009) show that empirically job separations are countercyclical and contribute substantially to the unemployment volatility. More recently, Coles and Moghaddasi Kelishomi (2018) find, by relaxing the free entry condition, that job destructions are the main driver of unemployment volatility.

The literature on the unemployment volatility puzzle mostly assumes that separation rate is constant and exogenous. Recently, Ljungqvist and Sargent (2017) survey the literature and find that it is the so-called fundamental surplus that matters for the elasticity of market tightness. But they do not consider the case of endogenous separations. Mortensen and Nagypal (2007) and Fujita and Ramey (2012) show that by allowing endogenous separation, the MP model is able to generate a sizable, albeit not sufficient, amount of unemployment volatility. Hence, this paper adds to the literature of endogenous separations by considering also inefficient job separations.

Few papers to the best of my knowledge in the literature explore the possibility of privately inefficient separations in the standard MP environment. Hall (2005b) compares the performance of several types of models and concludes that the sticky-wage efficient-separations model of Hall (2005c) performs much better than its inefficient counterpart from the behavior of the flows in

[3] Specifically, in this paper I consider job separations as only the inflows from employment state into unemployment state. Other authors in the literature consider also separations into employment (i.e. job-to-job transitions) and inactive states.

the labor market. However, as pointed out by Robert Shimer in the discussion of the lecture, Hall (2005b) considers only the inefficiency arising from a constant wage model. In the model considered in this paper, wages can still be flexible, and inefficient separations are due to some bargaining friction. Blanchard and Galí (2010) generate inefficient unemployment fluctuations in a model with nominal and real wage rigidities. However, their model does not allow for endogenous separations. Also, there is a strand of literature focusing on the social efficiency of unemployment. For example, Guerrieri (2008) studies the social inefficiency of unemployment in a competitive search equilibrium due to asymmetric information.

Finally, this paper is related to the literature on inefficient bargaining. For example, Chatterjee and Samuelson (1987), Kennan and Wilson (1993), and the references therein find that incomplete information may lead to inefficient bargaining outcomes. Also, Busch and Wen (1995), and Anderlini and Felli (2001) find that inefficiencies may arise when there are transaction costs in the process of bargaining. To the best of my knowledge, this paper is the first to bring the notion of inefficient bargaining outcome to the equilibrium model of unemployment.

The rest of the paper is organized as follows. The baseline model is constructed in Section 2, along with the discussion of the efficiency of job separations and unemployment. In Section 3, I consider the AOB model by Hall and Milgrom (2008) and I show that it belongs to the same class of model. I introduce asymmetric information about the match-specific productivity, as in Kennan (2010), in Section 4. Then I discuss the implications of the existence of bargaining friction for the volatility of cyclical unemployment in Section 5. In Section 6, I calibrate the model to the US economy, and evaluate quantitatively the impacts of bargaining friction on the labor market dynamics. Section 7 concludes.

2 The model economy

In this section, I consider a standard search and matching models á la Mortensen and Pissarides (1994) with endogenous separations and ex-post heterogeneity. Different from their model, however, I include bargaining wedges during the Nash bargaining process. I show that the wedges would generate inefficient unemployment. Models with alternating offers bargaining and with asymmetric information will be shown later to belong to the same class. Lastly, I discuss how inefficient separations would induce a higher-than-efficient level of unemployment, or inefficient unemployment.

2.1 The MP environment

Workers and firms are risk-neutral and infinitely lived with common discount factor β. They randomly match in the job market subject to search frictions. Once matched, the firm-worker pair produces a flow value of output $p_t(\varepsilon_t)$, where ε_t is the idiosyncratic, or match-specific, productivity specific only to the pair and $p_t'(\cdot) > 0$. To model the persistence of the idiosyncratic shock, I follow Fujita and Ramey (2012) to assume a probability λ of switching productivity. Hence, with probability $1-\lambda$, we have $\varepsilon_{t+1} = \varepsilon_t$. Otherwise, ε_{t+1} will be drawn from the stationary distribution $G(\cdot)$ on a compact support $[\varepsilon_{\min}, \varepsilon_{\max}]$. There is a probability s of exogenous separation. Otherwise, in each period, firms and workers can decide endogenously whether to dissolve the employment relationship. There is also a standard matching function $m(u_t, v_t)$ where u_t and v_t are the measures of unemployed workers and vacancies respectively. The market tightness at time t is $\theta_t = \frac{v_t}{u_t}$. Hence, the probability of an unemployed worker meeting with a vacancy is $\phi(\theta_t) = \frac{m(u_t, v_t)}{u_t} = m(1, \theta_t)$, and that of a vacancy meeting with an unemployed worker is $q(\theta_t) = \frac{\phi(\theta_t)}{\theta_t}$.

2.1.1 Firm's values

Firms post vacancies with a flow cost c. Let P_t be the (present discounted) value of production and W_t be the expected present value of wages during the time of the production[4]. Hence, a productive and working firm would receive a value of $P_t(\varepsilon_t) - W_t(\varepsilon_t)$. On the other hand, an unemployed worker earns a value of U_t. Once employed in a firm, the worker would receive the value of wage W_t and the value of working V_t. Therefore, the firm would fire the worker when $P_t(\varepsilon_t) \leq W_t(\varepsilon_t)$, and the worker would quit the job if $W_t(\varepsilon_t) + V_t(\varepsilon_t) \leq U_t$. Let $\chi_t(\varepsilon_t)$ is the indicator function of endogenous separation at time t, i.e.

$$\chi_t(\varepsilon_t) = \begin{cases} 1 & \text{when } \underbrace{P_t(\varepsilon_t) \leq W_t(\varepsilon_t)}_{\text{layoff}} \text{ or } \underbrace{W_t(\varepsilon_t) + V_t(\varepsilon_t) \leq U_t}_{\text{quit}} \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

[4] Here I follow Hall and Milgrom (2008) and Christinano et al. (2016) to use the discounted value of wages instead of the flow value of wages which would simplify the notations later on when I consider alternating offers bargaining.

Hence, the value of production is given by

$$P_t(\varepsilon_t) = p_t(\varepsilon_t) + \beta(1-s)\mathbb{E}_t \left[\begin{array}{c} \lambda\left(1-\chi_{t+1}(\varepsilon_{t+1})\right) P_{t+1}(\varepsilon_{t+1}) \\ +(1-\lambda)\left(1-\chi_{t+1}(\varepsilon_t)\right) P_{t+1}(\varepsilon_t) \end{array} \right] \tag{2}$$

Moreover, I follow Pissarides (2009) to include a fixed matching cost. Hence, free entry of the firms implies that

$$\beta q(\theta_t)\mathbb{E}_t\left[\left(1-\chi_{t+1}(\varepsilon_{t+1})\right)\left(P_{t+1}(\varepsilon_{t+1}) - W_{t+1}(\varepsilon_{t+1}) - k\right)\right] = c \tag{3}$$

where ε_{t+1} denotes the idiosyncratic productivity at time $t+1$ which is random at time t, and k is the fixed matching cost. The free entry condition states that the flow cost equals expected profit of the firm in the future. Note that there is a possibility of immediate separation, should the value of production be not enough to cover that of the wage, or the worker is rejecting the job offer.

2.1.2 Worker's values

Unemployed workers receive a flow value b of non-market activity, which may include such things as leisure and unemployment benefit. Hence, the value of an unemployed worker is given by

$$U_t = b + \beta\mathbb{E}_t \left[\begin{array}{c} \phi(\theta_t)\left(1-\chi_{t+1}(\varepsilon_{t+1})\right)\left(W_{t+1}(\varepsilon_{t+1}) + V_{t+1}(\varepsilon_{t+1})\right) \\ +\left[(1-\phi(\theta_t)) + \phi(\theta_t)\chi_{t+1}(\varepsilon_{t+1})\right] U_{t+1} \end{array} \right] \tag{4}$$

and the value of working is given by[5]

$$V_t(\varepsilon_t) = \beta\mathbb{E}_t \left\{ \begin{array}{c} \left[s + (1-s)\left(\lambda\chi_{t+1}(\varepsilon_{t+1}) + (1-\lambda)\chi_{t+1}(\varepsilon_t)\right)\right] U_{t+1} \\ +(1-s)\left[\begin{array}{c} \lambda\left(1-\chi_{t+1}(\varepsilon_{t+1})\right) V_{t+1}(\varepsilon_{t+1}) \\ +(1-\lambda)\left(1-\chi_{t+1}(\varepsilon_t)\right) V_{t+1}(\varepsilon_t) \end{array} \right] \end{array} \right\} \tag{5}$$

2.1.3 Match surplus and thresholds

The worker's surplus is given by

$$J_t(\varepsilon_t) = V_t(\varepsilon_t) + W_t(\varepsilon_t) - U_t \tag{6}$$

[5] Alternatively, the value of working can be interpreted as the difference between the value of being employed and the value of wages. See Appendix A for a discussion.

whenever it is positive, and zero otherwise. Similarly, the firm's surplus is given by

$$F_t(\varepsilon_t) = P_t(\varepsilon_t) - W_t(\varepsilon_t) \tag{7}$$

whenever it is positive, and zero otherwise. Finally, the joint surplus is given by

$$S_t(\varepsilon_t) = P_t(\varepsilon_t) + V_t(\varepsilon_t) - U_t \tag{8}$$

whenever it is positive, and zero otherwise. I assume the wage rule is such that the above surplus functions are all strictly increasing and continuous in ε_t (which would apply to the Nash bargaining model I consider in the next subsection). Hence, there exist cutoff values $\varepsilon_t^w, \varepsilon_t^f$ and ε_t^e such that

$$J_t(\varepsilon_t^w) = 0 \tag{9}$$

$$F_t\left(\varepsilon_t^f\right) = 0 \tag{10}$$

$$S_t(\varepsilon_t^e) = 0 \tag{11}$$

Intuitively, the worker will choose to quit the job if she observes a productivity $\varepsilon < \varepsilon_t^w$. Similarly, the firm would fire the worker if $\varepsilon < \varepsilon_t^f$. Finally, ε_t^e denotes the efficient cutoff productivity. Therefore, any job separation happening at a productivity $\varepsilon_t > \varepsilon_t^e$ is privately inefficient, since the employment pair would have had a positive joint surplus if they were to choose not to separate.

2.2 Nash bargaining with bargaining wedges (NBBW) and inefficient separation

Let $\eta \in [0,1]$ be the bargaining power of the worker. Wages are determined by maximizing the generalized Nash product

$$\begin{aligned} & \max_{W_t} \left(J_t(\varepsilon_t) - \Omega_t^w\right)^\eta \left(F_t(\varepsilon_t) - \Omega_t^f\right)^{1-\eta} \\ = & \max_{W_t} \left(V_t(\varepsilon_t) + W_t(\varepsilon_t) - U_t - \Omega_t^w\right)^\eta \left(P_t(\varepsilon_t) - W_t(\varepsilon_t) - \Omega_t^f\right)^{1-\eta} \end{aligned} \tag{12}$$

where Ω_t^w and Ω_t^f, which are exogenous to the worker and the firm, are the bargaining wedges of the worker and the firm respectively. A bargaining wedge is the difference between the actual surplus and the bargaining surplus. We can interpret the wedges as a reduced-form specification of bargaining friction, when the parties fail to recognize the actual surpluses when bargaining. If $\Omega_t^w > 0$, for example, then the actual surplus the worker receives ex-post will be higher than the surplus perceived when the worker is bargaining with the firm. Note that when $\Omega_t^w = \Omega_t^f = 0$, it reduces to the standard Nash bargaining problem. The solution to the above problem is given by

$$\eta\left(P_t(\varepsilon_t) - W_t(\varepsilon_t) - \Omega_t^f\right) = (1-\eta)\left(V_t(\varepsilon_t) + W_t(\varepsilon_t) - U_t - \Omega_t^w\right) \tag{13}$$

and by rearranging we have the wage rule

$$\begin{aligned} W_t(\varepsilon_t) &= \eta\left(P_t(\varepsilon_t) - \Omega_t^f\right) + (1-\eta)\left(U_t - V_t(\varepsilon_t) + \Omega_t^w\right) \\ &= \eta P_t(\varepsilon_t) + (1-\eta)\left(U_t - V_t(\varepsilon_t)\right) + (1-\eta)\,\Omega_t^w - \eta\Omega_t^f \end{aligned} \tag{14}$$

Note that other things equal, the wage is increasing with Ω_t^w and decreasing with Ω_t^f. In what follows I define

$$\Omega_t^{net} \equiv (1-\eta)\,\Omega_t^w - \eta\Omega_t^f \tag{15}$$

as the *net bargaining wedge* between the worker and the firm. It measures the weighted difference between the bargaining wedge of the worker and that of the firm. Hence, it represents the net distortionary effect of the bargaining wedges on the wage rule. It turns out that the value of the net bargaining wedge has important implications for the efficiency of separation as well, as demonstrated in the following proposition.

Proposition 1 *Suppose wages are determined by the NBBW model.*

(i) If $\Omega_t^{net} = 0$, *then* $\varepsilon_t^w = \varepsilon_t^e = \varepsilon_t^f$.

(ii) If $\Omega_t^{net} > 0$, *then* $\varepsilon_t^w < \varepsilon_t^e < \varepsilon_t^f$.

(iii) If $\Omega_t^{net} < 0$, *then* $\varepsilon_t^f < \varepsilon_t^e < \varepsilon_t^w$.

Proof. See Appendix B. ■

The intuition of Proposition 1 is summarized in Figure 1. When $\Omega_t^{net} = 0$, the three cutoff productivities are the same, hence job separation is always efficient in this case. This implies the classical result that employment and separation under standard Nash bargaining are always

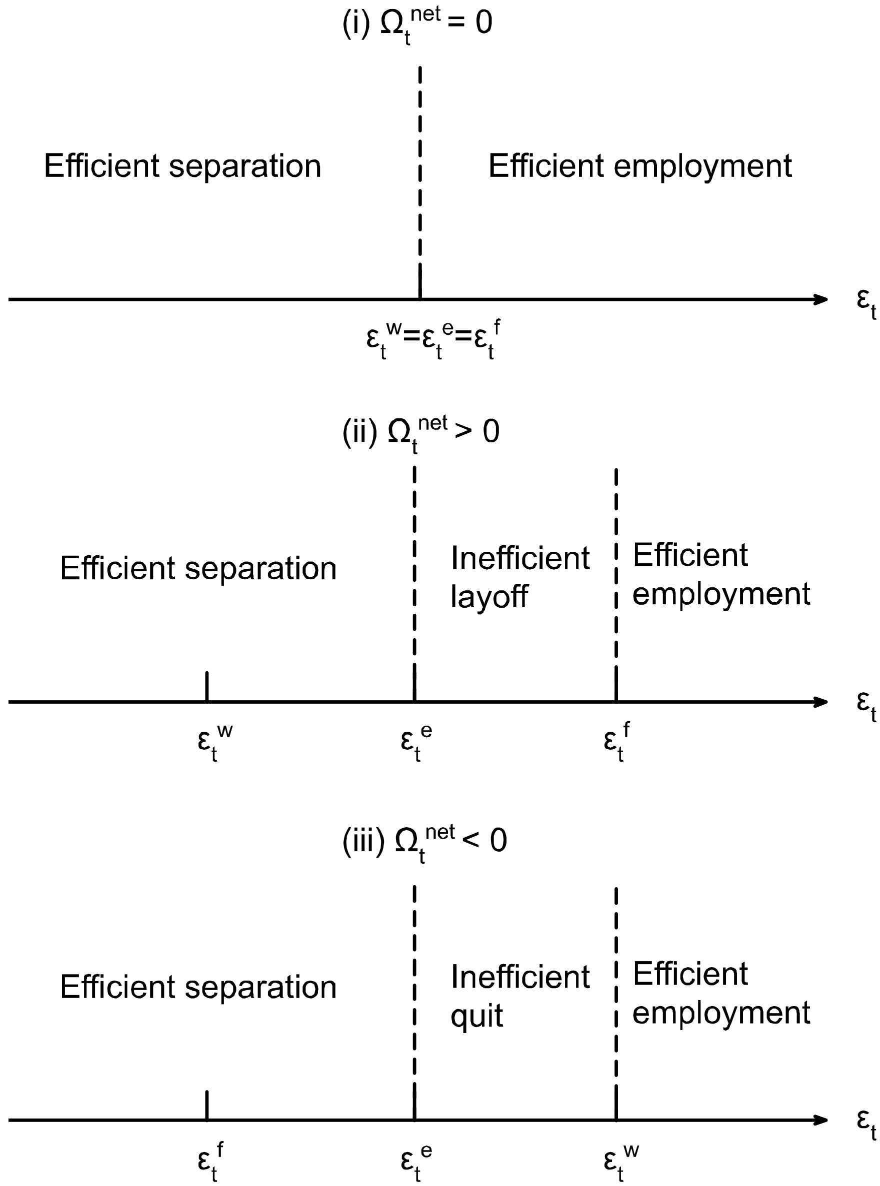

Figure 1: Efficiency of job separations

privately efficient. On the other hand, inefficient separation may arise when the net bargaining wedge is non-zero. When $\Omega_t^{net} > 0$, wages are excessively high relative to the standard bargaining solution. As a result, the cutoff productivity ε_t^f for layoff decision by the firm is larger than the efficient cutoff productivity ε_t^e. In this case, *inefficient layoff* would be possible. For any productivity that lies between ε_t^e and ε_t^f, the firm would choose to fire the worker even when the joint surplus is positive. It is only if the productivity is sufficient low (i.e. less than ε_t^e) would there be efficient separation. On the other hand, when $\Omega_t^{net} < 0$, wages are too low relative to the standard case. So the cutoff productivity ε_t^w for quitting decision by the worker is larger than the efficient cutoff productivity ε_t^e. Similar to the previous case, *inefficient quit* would occur when the productivity falls between ε_t^e and ε_t^w.

It may be interesting to see what factors affect the *inefficiency gaps* $\varepsilon_t^f - \varepsilon_t^e$ and $\varepsilon_t^w - \varepsilon_t^e$ in (ii) and (iii). To obtain a closed-form solution, I consider a special case when there is no persistence between productivity shocks (i.e. $\lambda = 1$), and the total productivity p_t is a product of aggregate productivity y_t and individual productivity ε_t.

Proposition 2 *If* $\lambda = 1$ *and* $p_t(\varepsilon_t) = y_t \varepsilon_t$, *then*

$$\begin{aligned} \textit{(i)}\ \varepsilon_t^f - \varepsilon_t^e &= \frac{\Omega_t^{net}}{y_t(1-\eta)} \\ \textit{(ii)}\ \varepsilon_t^w - \varepsilon_t^e &= \frac{-\Omega_t^{net}}{y_t \eta} \end{aligned}$$

Proof. See Appendix B. ■

Proposition 2 states that the inefficiency gap is proportional to the magnitude of the net bargaining wedge, and inversely proportional to the aggregate productivity. It also depends negatively on the bargaining power of the party that initiates the inefficient separation. When $\Omega_t^{net} > 0$, for example, the inefficiency gap for inefficient layoff depends negatively on the bargaining power of the firm.

There are some implications following from Proposition 2. First, the effect by the net bargaining wedge is magnified in bad times, implying that the inefficiency gap tends to be larger in recession. Intuitively, job separation decision is made only when the total productivity is low enough. Hence, the labor market will be more sensitive to the distortionary effect during bad times, when the aggregate productivity is low. This generates more inefficient job separations and could potentially explain the mass layoffs observed in the recessions. Second, a higher bargaining power of the firm would be able to mitigate the distortion when the net bargaining power is positive, and similarly for that of the worker when net bargaining power is negative. Intuitively, when the weighted bargaining wedge of the worker is larger than that of the firm (i.e. $\Omega_t^{net} > 0$), the wage rule would be distorted in favor of the worker, which would entail a higher-than-efficient level of wage rate. A higher bargaining power of the firm would mean the firm would be able to bargain for a lower wage, hence offsetting part of the distortion by the bargaining wedge.

2.3 Worker flows and inefficient unemployment

I measure the worker flows as follows. Let $e_t(\varepsilon)$ be the distribution function of the employed workers, and u_t be the unemployment rate. Hence, $e_t = e_t(\varepsilon_{\max}) = 1 - u_t$ will be the employment rate. Also, denote ε_t^s as the separation cutoff value of the individual productivity. i.e.

$$\varepsilon_t^s = \max\left\{\varepsilon_t^w, \varepsilon_t^f\right\} \tag{16}$$

Note that Proposition 1 implies that $\varepsilon_t^s \geq \varepsilon_t^e$. Then the flow equation of the employment distribution is given by

$$\begin{aligned} e_{t+1}(\varepsilon) \quad &= \quad (1-s)\left\{\lambda e_t\left[G(\varepsilon) - G\left(\varepsilon_{t+1}^s\right)\right] + (1-\lambda)\left[e_t(\varepsilon) - e_t\left(\varepsilon_{t+1}^s\right)\right]\right\} \\ &\quad + u_t\phi(\theta_t)\left[G(\varepsilon) - G\left(\varepsilon_{t+1}^s\right)\right] \end{aligned} \tag{17}$$

Hence, the measure of employed workers having productivity less than or equal to ε at time $t+1$ comes from three groups of workers at time t, namely (i) those employed at time t who are switching productivity, and drawing a productivity from the interval $\left[\varepsilon_{t+1}^s, \varepsilon\right]$, (ii) those employed at time t who have a productivity in $\left[\varepsilon_{t+1}^s, \varepsilon\right]$ and not switching their productivity, and (iii) those unemployed at time t, having matched to a firm, and drawing a productivity from the interval $\left[\varepsilon_{t+1}^s, \varepsilon\right]$.

Similarly, the flow to unemployment is given by

$$u_{t+1} = e_t\left\{s + (1-s)\lambda G\left(\varepsilon_{t+1}^s\right)\right\} + e_t\left(\varepsilon_{t+1}^s\right)(1-s)(1-\lambda) + u_t\left[1 - \phi(\theta_t)\left(1 - G\left(\varepsilon_{t+1}^s\right)\right)\right] \tag{18}$$

Therefore, those employed workers at time t whose jobs are destroyed for exogenous reason and whose productivities are too low, and those unemployed at time t who are not able to match a firm and to draw a high enough productivity, will be unemployed at time $t+1$.

To see the effect of the inefficient job separation on the efficiency of unemployment, let's consider the steady state unemployment rate,

$$u = \frac{s + (1-s)\lambda G(\varepsilon^s)}{\phi(\theta)(1 - G(\varepsilon^s)) + s + (1-s)\lambda G(\varepsilon^s)} \tag{19}$$

and its efficient counterpart u^e, when ε^s is replaced by ε^e. Since $\varepsilon^s \geq \varepsilon^e$ and u is clearly increasing with ε^s. We have $u \geq u^e$.[6] This demonstrates that the existence of bargaining friction induces a higher level of unemployment than the standard model where all separations are efficient. Hereafter I shall call the extra level of unemployment the *inefficient unemployment*.

I derive the flow equations of efficient level of unemployment u_t^e by replacing ε_t^s in the flow equations (17) and (18) with ε_t^e. Hence, efficient unemployment corresponds to the level of unemployment when all separations are efficient. Inefficient unemployment will then be given by the difference $u_t^i = u_t - u_t^e$. This simple decomposition allows me to measure the contribution of unemployment volatility by the inefficient separations later in the quantitative analysis.

The EU transition and UE transition are given respectively as follows:

$$EU_{t+1} = e_t \left\{ s + (1-s)\lambda G\left(\varepsilon_{t+1}^s\right) \right\} + e_t\left(\varepsilon_{t+1}^s\right)(1-s)(1-\lambda) \tag{20}$$

$$UE_{t+1} = \phi(\theta_t)\left(1 - G\left(\varepsilon_{t+1}^s\right)\right) u_t \tag{21}$$

Finally, the job separation and job finding rates are given by

$$SR_{t+1} = \frac{EU_{t+1}}{e_t} = s + (1-s)\lambda G\left(\varepsilon_{t+1}^s\right) + \frac{e_t\left(\varepsilon_{t+1}^s\right)}{e_t}(1-s)(1-\lambda) \tag{22}$$

$$FR_{t+1} = \frac{UE_{t+1}}{u_t} = \phi(\theta_t)\left(1 - G\left(\varepsilon_{t+1}^s\right)\right) \tag{23}$$

As we will see later, a relatively acyclical level of bargaining friction would induce a higher volatility in the cutoff productivity ε_t^s and hence in the workers flow in the labor market. Quantitatively,

[6] Note that I have assumed the steady state value of the market tightness is fixed and determined how unemployment rate varies with the cutoff productivity.

however, the effect on the job finding rate is limited, since the cutoff productivity only affects the job finding rate through the distribution function $G(\cdot)$.

3 Alternating offers bargaining (AOB)

Hall and Milgrom (2008) apply the AOB model of Rubinstein (1982) and Binmore et al. (1986) to the MP environment, where they distinguish between the *outside-option payoff* and the *disagreement payoff*. In this bargaining situation, when the parties fail to reach an agreement, they continue to bargain through a sequence of offers and counter-offers, instead of quitting the negotiation immediately. In fact, at the unique subgame perfect equilibrium in the AOB model, both parties would move to an agreement immediately. Because workers and firms are mainly considering the disagreement payoff when bargaining, the wage solution can be shown to be less sensitive to productivity change than the standard Nash bargaining model. Hence, this generates real wage rigidity endogenously.

The model is described as follows. Following Christinano et al. (2016), I assume the whole bargaining process lasts within one period[7]. In each period, the firm would make a first offer W_t in the first subperiod[8]. If the worker accepts the offer, she will receive a value of $W_t + V_t$. Otherwise, the worker can propose a counter-offer W_t' in the next subperiod, when the firm can similarly choose to accept the counter-offer with the value $P_t - W_t'$ or reject the offer and propose another offer in the next round, and so on *ad infinitum*. In addition, there is a probability δ in each subperiod that the bargaining would break down and both parties would get their respectively outside-option payoff. While bargaining, the worker receives a flow benefit z_t, and the firm incurs a cost γ_t for maintenance. I assume, for simplicity, no party can leave voluntarily once entered into the negotiation.[9]

At the unique subgame perfect equilibrium, the firm would always propose the lowest wage that

[7] Christinano et al. (2016) assume the number of subperiods in a period is finite. Hence, the solution obtained here can be considered as the limiting case when the number of subperiods tends to infinity. The results below would still follow when number is finite instead. Hall and Milgrom (2008) assume bargaining lasts for multiple periods. The bargaining setting considered here is the intra-period version of their model.

[8] As in the NBBW model, I assume the worker and the firm are bargaining only over the present value of wages W_t, instead of over other values such that V_t and P_t. Theoretically, the firm can change V_t (and similarly the worker can also affect P_t) by choosing the timing of separation. I ignore this possibility since in all practical bargaining situation, workers and firms mainly bargain over wages only.

[9] Since the process of bargaining is costly, it can be shown that if one party chooses to initiate separation during the bargaining process, the party would not have entered the bargaining in the first place.

the worker would accept, and similarly the worker would always propose the highest wage that the firm would accept. Hence, we have the following indifference conditions:

$$W_t(\varepsilon_t) + V_t(\varepsilon_t) = \delta U_t + (1-\delta)\left[z_t + \left(W_t'(\varepsilon_t) + V_t(\varepsilon_t)\right)\right] \tag{24}$$

$$P_t(\varepsilon_t) - W_t'(\varepsilon_t) = (1-\delta)\left[-\gamma_t + \left(P_t(\varepsilon_t) - W_t(\varepsilon_t)\right)\right] \tag{25}$$

The worker receives the value of $W_t(\varepsilon_t) + V_t(\varepsilon_t)$ by accepting the offer. On the other hand, if the worker rejects the offer, there is a probability δ that the worker would be unemployed; otherwise, the worker earns a flow benefit z_t and proposes a counter-offer W_t'. Hence, (24) shows that the firm would choose W_t such that the worker is indifferent between accepting and rejecting the offer. The same is true for the firm, which is reflected in (25). Since the firm is assumed to make the first offer, the worker would accept it by indifference, and so W_t would be the wage solution in this model. In fact, it can be shown that it is a special case of the NBBW model considered in the previous section.

Proposition 3 *The unique subgame perfect Nash equilibrium in the AOB model is equivalent to the solution to the NBBW model with*

$$\begin{aligned}
\eta &= \frac{1-\delta}{2-\delta} \\
\Omega_t^w &= \frac{1-\delta}{\delta} z_t \\
\Omega_t^f &= -\frac{1-\delta}{\delta} \gamma_t
\end{aligned}$$

Proof. See Appendix B. ■

Intuitively, under the AOB protocol, while the outside-option payoff of the worker is still U_t, the disagreement payoff is

$$\sum_{i=0}^{\infty} (1-\delta)^i \left[\delta U_t + (1-\delta) z_t\right] = U_t + \frac{1-\delta}{\delta} z_t$$

which is larger than the outside-option payoff. Hence, the difference $\frac{1-\delta}{\delta} z_t$ would be the bargaining wedge. Similarly, the outside-option payoff of the firm is zero due to free entry, but the disagreement payoff is

$$\sum_{i=1}^{\infty} (1-\delta)^i (-\gamma_t) = -\frac{1-\delta}{\delta}\gamma_t$$

which would also be the bargaining wedge. Note that the bargaining power of the worker is no larger than $\frac{1}{2}$, which shows the first-mover advantage of the firm by proposing an offer first.

In this case, the net bargaining wedge is given by

$$\Omega_t^{net} = (1-\eta)\,\Omega_t^w - \eta\Omega_t^f = \frac{1}{2-\delta}\frac{1-\delta}{\delta}\left[z_t + (1-\delta)\,\gamma_t\right] > 0 \tag{26}$$

and so by Proposition 1, there could be inefficient layoffs. Here the inefficiency comes from the cost of delay. Although the negotiation happens within a period, hence there is no discounting, the probability of a breakdown, along with the flow benefit and cost during bargaining creates incentive problem for both parties.

Finally, the comparative static analysis of the inefficiency gap $\varepsilon_t^f - \varepsilon_t^e$ for the special case considered previously is summarized as follows.

Corollary 4 *If $\lambda = 1$ and $p_t(\varepsilon_t) = y_t\varepsilon_t$, then*

$$\begin{aligned} \frac{\partial\left(\varepsilon_t^f - \varepsilon_t^e\right)}{\partial z_t} &> 0 \\ \frac{\partial\left(\varepsilon_t^f - \varepsilon_t^e\right)}{\partial \gamma_t} &> 0 \\ \frac{\partial\left(\varepsilon_t^f - \varepsilon_t^e\right)}{\partial \delta} &< 0 \end{aligned}$$

In particular, if $\delta \to 1$, then $\varepsilon_t^f \to \varepsilon_t^e$.

Proof. See Appendix B. ■

The intuition of Corollary 4 is straightforward. If the flow benefit during bargaining increases, the worker would be more willing to disagree under AOB, and hence would be asking a higher wage than the standard Nash bargaining model. This could possibly generate more inefficient layoffs. Also, if the maintenance cost during bargaining increases, the firm would be more likely to agree with the worker, and hence would be willing to accept a lower wage than the standard model. Again it could generate more inefficient layoffs. Lastly, when the breakdown probability increases, the expected duration of disagreement decreases, and hence would reduce the effects of the flow benefit and the flow cost on the wage. This would mitigate the inefficiency.

4 Asymmetric information and neutral bargaining solution (NBS)

In this section, I introduce asymmetric information into the MP model. I consider the simple case when only the firm has private information about the match-specific productivity. This case has been studied by Kennan (2010), who employs the neutral bargaining solution (NBS) of Myerson (1984) to solve the model. However, he assumes the worker would always propose a pooling offer, resulting in efficient employment. Here I relax this assumption, which allows the possibility of inefficient separation.

Suppose now the firm has private information about ε_t and the worker has no way to verify it. Myerson (1984) proposes the NBS as a generalization of the Nash bargaining solution in the presence of private information. He shows that the NBS to the bargaining problem in this case always exists, and can be implemented by a Random Dictator mechanism. Specifically, suppose upon matching with the firm, there is a probability ν that the worker is chosen by the Dictator to propose a take-it-or-leave-it offer. Otherwise, the firm would be chosen to make the offer. Since the firm observes ε_t, its offer would extract all the match surplus $S_t(\varepsilon_t)$, leaving the worker with the value of unemployment and zero surplus. On the other hand, since the worker have no knowledge about the size of the surplus, the offer cannot depend on ε_t. If the worker demands a surplus of $S_t(\bar{\varepsilon})$, then there is a probability $G(\bar{\varepsilon})$ that the firm would dissolve the employment relationship, when the total surplus $S_t(\varepsilon_t)$ is no greater than the worker's demand. Hence, the worker chooses $\bar{\varepsilon}_t$ to maximizes the expected surplus:

$$\bar{\varepsilon}_t = \arg\max_{\varepsilon} \left(1 - G(\varepsilon)\right) S_t(\varepsilon) \tag{27}$$

Under the Random Dictator mechanism, the worker would receive the ν fraction of the above surplus and similarly for the firm as follows.

$$J_t(\varepsilon_t) = \nu(1 - G(\bar{\varepsilon}_t)) S_t(\bar{\varepsilon}_t) \tag{28}$$

$$F_t(\varepsilon_t) = \nu(1 - G(\bar{\varepsilon}_t))(S_t(\varepsilon_t) - S_t(\bar{\varepsilon}_t)) + (1 - \nu) S_t(\varepsilon_t) \tag{29}$$

Intuitively, with probability $\nu(1 - G(\bar{\varepsilon}_t))$, the worker is picked to make the offer $S_t(\bar{\varepsilon}_t)$, and the total surplus is enough (i.e. when $\varepsilon_t \geq \bar{\varepsilon}_t$) to cover the offer. Hence, the firm would get $S_t(\varepsilon_t) - S_t(\bar{\varepsilon}_t)$ in this case. On the other hand, with probability $1 - \nu$, the firm would propose to get the whole surplus, in which case the worker would get zero surplus. The Random Dictator mechanism stipulates that each party gets its corresponding expected payoff. It is clear that in the above specification, the strategies of the worker and the firm are incentive-efficient. Hence, the NBS can be implemented by the Random Dictator mechanism.

The following proposition follows immediately by the comparison with the bargaining solution in Section 2 in a similar manner to Proposition 3, and its proof is thus omitted.

Proposition 5 *The NBS to the asymmetric information model is equivalent to the solution to the NBBW model with*

$$\begin{aligned} \eta &= 0 \\ \Omega_t^w &= \nu(1 - G(\bar{\varepsilon}_t)) S_t(\bar{\varepsilon}_t) \\ \Omega_t^f &= -\nu(1 - G(\bar{\varepsilon}_t)) S_t(\bar{\varepsilon}_t) \end{aligned}$$

First, under asymmetric information, the effective bargaining power of worker is zero. The private information about the match-specific productivity effectively gives the firm all the bargaining power. This is in contrast with the complete information case, when the bargaining power is given by the probability of making offer (ν). It is because the worker, not knowing the actual size of the match surplus, is not able to bargain with the firm over the actual surplus. Instead, the worker demands a fixed amount of surplus which would then cause inefficiency.

Note that in this case the net bargaining wedge is given by the worker's surplus.

$$\Omega_t^{net} = \nu(1 - G(\bar{\varepsilon}_t)) S_t(\bar{\varepsilon}_t) \geq 0 \tag{30}$$

Since the worker demands the same surplus regardless of the realization of ε_t, inefficient layoff occurs when the realized total surplus is positive but lower than certain level so that the firm's surplus is negative. Since only the firm observes the productivity, the layoff decision is efficient if the total match surplus is in line with the firm's surplus. This happens if the firm is always making the offer, i.e. when $\nu = 0$.

5 Implications for the volatility of cyclical unemployment

Before turning to the quantitative analysis, I discuss the implications of bargaining friction for the unemployment volatility. I first compare the specification of the bargaining wedge with the concept of "fundamental surplus" in Ljungqvist and Sargent (2017). Then I discuss the cyclicality of worker's share of the match surplus in the NBBW model, and its implications for the unemployment volatility.

5.1 Relationship with the fundamental surplus in Ljungqvist and Sargent (2017)

Ljungqvist and Sargent (2017) show that in a steady state MP model with exogenous separations and homogenous productivity, the elasticity of market tightness θ with respect to productivity y depends crucially on a factor which they refer to as the fundamental surplus. They show that in order for the elasticity of market tightness to be larger, the fundamental surplus has to be small. Here I incorporate the bargaining wedges described previously into their model, and to see how they are related to the concept of fundamental surplus.

Following their derivation (details are in the Appendix B), it can be shown that the elasticity of market tightness in this case is given by

$$\epsilon_{\theta,y} = \Upsilon\left(\Omega^{net}\right) \frac{y}{y - z - \left[\frac{1-\beta(1-s)}{1-\eta}\right]\Omega^{net}} \tag{31}$$

where

$$\Upsilon\left(\Omega^{net}\right) = \frac{r + s + \eta\theta q\left(\theta\right)\left(1 + \beta q\left(\theta\right)\frac{\Omega^{net}}{c\eta}\right)}{\alpha\left(r+s\right) + \eta\theta q\left(\theta\right)\left(1 + \left(1-\alpha\right)\beta q\left(\theta\right)\frac{\Omega^{net}}{c\eta}\right)} \tag{32}$$

is the factor with an upper bound $\max\left\{\frac{1}{\alpha}, \frac{1}{1-\alpha}\right\}$ and $\alpha \equiv \frac{-q'(\theta)\theta}{q(\theta)}$ is the elasticity of the matching function. Here the fundamental surplus is $y - z - \left[\frac{1-\beta(1-s)}{1-\eta}\right]\Omega^{net}$. Therefore, the existence of bargaining wedges will increase the elasticity of market tightness if $\Omega^{net} > 0$. As in many cases studied by Ljungqvist and Sargent (2017), bargaining friction can also potentially increase the unemployment volatility through raising the elasticity of market tightness.

There are problems with the steady state analysis of unemployment volatility by Ljungqvist and Sargent (2017). First, as mentioned in detail in the Appendix of Christiano et al. (2016), steady state models are generally misleading about the dynamic effects of a persistent shock. This point is made similarly in Petrosky-Nadeau and Zhang (2017). Second, and more relevant to this paper, the sole focus on the elasticity of market tightness to explain unemployment dynamics replies on the assumption that the elasticity of the separation rate is zero. However, there is evidence that the separation rate is counter-cyclical[10]. As shown in these studies, the elasticity of the separation rate relevant to the unemployment dynamics is indeed non-trivial. In fact, as I will show in the quantitative analysis, the main effect of the bargaining friction is on the volatility of separation rate, which is ignored by the analysis in much of the literature.

In what follows, I shall refer to the above mechanism as the *small surplus channel*. In models with exogenous separations and homogenous productivity, the small surplus channel is the only channel through which bargaining friction can affect unemployment volatility.

5.2 Cyclicality of worker's share of the match surplus

In the standard Nash bargaining model, the worker's share of the match surplus, defined as

$$\phi_t(\varepsilon_t) = \frac{J_t(\varepsilon_t)}{S_t(\varepsilon_t)} \tag{33}$$

is equal to his/her (constant) bargaining power η, and thus is acyclical. Hence, the firm's share $1 - \phi_t(\varepsilon_t) = 1 - \eta$ is also unresponsive to aggregate fluctuation. This explains the flexibility of wages under Nash bargain, as the worker's surplus moves proportionally with the match surplus. Under the existence of bargaining friction, however, the worker's share can be responsive to the

[10] See, e.g., Elsby (2009), Fujita and Ramey (2009), and, more recently, Coles and Moghaddasi Kelishomi (2018). This fact is also shown in Table 1 in the next section.

business cycle. By rearranging the wage rule (14), we have the following[11].

Proposition 6 *The worker's share of the match surplus in the NBBW model is given by*

$$\phi_t(\varepsilon_t) = \eta + \frac{\Omega_t^{net}}{S_t(\varepsilon_t)}$$

Hence, if Ω_t^{net} *is relatively acyclical, then the worker's share is countercyclical (procyclical) if* $\Omega_t^{net} > 0 \ (< 0)$.

The intuition of the proposition can perhaps be better understood in the context of the AOB model discussed earlier. In this case, we have a constant $\Omega_t^{net} > 0$ and so the worker's share is countercyclical. Since the disagreement payoff for both parties is to continue for another round of bargain, the wage offers proposed would be less sensitive to the aggregate conditions[12]. This entails that, consistent with the findings in Hall and Milgrom (2008), wages move less than proportionally with the aggregate productivity. Equivalently, the worker's share of the match surplus increases in recessions. While the above is true for the AOB model, the NBBW model considered in this paper provides a more general framework to explain the increasing bargaining power of the workers in bad times. It is worth noting that the last result still holds even if Ω_t^{net} is either countercyclical, or procyclical but less so than the match surplus (in the sense that $\frac{\Omega_t^{net}}{S_t(\varepsilon_t)}$ is still countercyclical).

There are two consequences for the countercyclicality of the worker's share of the match surplus. First, this reduces the job creation incentive in the recession, compared with the standard Nash bargaining model. Also, a small change in y_t causes significant fluctuations in the cutoff productivity, and thus in the quantity of job separations. As a result, in recessions there would be more (inefficient) job separations, which only exists in the NBBW model. I call this effect the *separation channel*, which does not exist in homogenous models like that in Ljungqvist and Sargent (2017). Both the small surplus and the separation channels mentioned above could potentially generate excessive fluctuation of unemployment. In the next section, I will quantity these channels and show that Ω_t^{net} has much stronger effect on the separation channel.

[11] Clearly, if Ω_t^{net} is large or $S_t(\varepsilon_t)$ is small enough, it is possible that $\eta + \frac{\Omega_t^{net}}{S_t(\varepsilon_t)} > 1$. But in this case, the firm's surplus $F_t(\varepsilon_t)$ is negative and hence the firm would choose to fire the worker. The case when the expression is negative is similar. Therefore, the worker's share is properly defined when there is continuing employment.

[12] It should be noted that, due to the bargaining friction, a large part of the wage rigidity comes from the fact that many matches end before a large wage cut can be carried out. The wages of ongoing matches may in fact still be relatively procyclical.

6 Quantitative analysis

In this section, I calibrate the model to the US labor market. Details of the computation strategy are in Appendix D. I will use the calibrated model to evaluate the impact of bargaining friction on the unemployment dynamics. Motivated by the examples in the previous sections, I consider only the case when the net bargaining wedge is non-negative, in which case there may be inefficient layoffs.

6.1 Data

I consider monthly data from January, 1976 to December, 2016. The unemployment rate series are taken from the official rates from the Labor Force Statistics of the Current Population Survey (CPS). The EU and UE transition rates are also taken from the same survey. For the period from January, 1976 to December, 2005, the separation rates and job finding rates are taken from Fujita and Ramey (2009). For the remaining years, I calculate the rates by $SR_{t+1} = \frac{EU_{t+1}}{e_t}$ and $FR_t = \frac{UE_{t+1}}{u_t}$, and adjust them for time aggregation error as in Fujita and Ramey (2009). Labor productivity p_t is defined by real GDP per workers, where the real GDP data is from Bureau of Economic Analysis and the number of employed workers is from CPS. Finally, the vacancy data is from the updated series by Barnichon (2010).

A summary table of the business cycle statistics of the US labor market is given in Table 1. When calculating the business cycle statistics, all series are converted to quarterly data with simple averaging, as well as logged and HP filtered with smoothing parameter 1,600. In the table, σ_x^c denotes the cyclical volatility of the variable x, and $\epsilon_{x,p}$ denotes the elasticity of the variable x with respect to the labor productivity, as defined by

$$\epsilon_{x,p} = \frac{cov\left(x_t, p_t\right)}{var\left(p_t\right)} \tag{34}$$

It is worth noting that while the job finding rate is procyclical with an elasticity of 4.8 with respect to changes in labor productivity, the job separation rate is countercyclical with a comparable elasticity of -3.5. This shows that the separation margin of the labor market is indeed important at the business cycle frequency.

	p_t	u_t	FR_t	SR_t	UE_t	EU_t	v_t	θ_t
σ_x^c	0.014	0.11	0.085	0.065	0.044	0.057	0.128	0.233
$\epsilon_{x,p}$	1	-6.595	4.759	-3.503	-1.725	-2.818	8.006	14.599
$corr\left(x_t, x_{t-1}\right)$	0.865	0.939	0.838	0.701	0.444	0.651	0.919	0.936
Correlation matrix								
	p_t	u_t	FR_t	SR_t	UE_t	EU_t	v_t	θ_t
p_t	1	-0.870	0.802	-0.772	-0.561	-0.706	0.898	0.900
u_t		1	-0.921	0.746	0.646	0.662	-0.932	-0.980
FR_t			1	-0.617	-0.300	-0.524	0.883	0.916
SR_t				1	0.635	0.982	-0.774	-0.774
UE_t					1	0.617	-0.563	-0.612
EU_t						1	-0.706	-0.697
v_t							1	0.985
θ_t								1

Table 1: Summary statistics of the US labor market, 1976 - 2016

6.2 Specification and calibration

The matching function is of the standard Cobb-Douglas form.

$$m\left(u, v\right) = A u^{\alpha} v^{1-\alpha} \tag{35}$$

where α is the elasticity of the matching function and A measures the matching efficiency. This implies the probability of a vacancy meeting with an unemployed worker is $q\left(\theta\right) = \frac{m(u,v)}{v} = A\theta^{-\alpha}$ and that of the an unemployed worker meeting a vacancy is $\phi\left(\theta\right) = \frac{m(u,v)}{u} = A\theta^{1-\alpha}$.

Following Fujita and Ramey (2012), I assume $p_t\left(\varepsilon_t\right) = y_t \varepsilon_t$, where y_t is the aggregate productivity following the stochastic process

$$\ln y_t = \rho \ln y_{t-1} + \zeta_t \tag{36}$$

where ζ_t is identically and independently distributed normal white noise with mean μ_ζ and standard deviation σ_ζ. Also, the distribution function $G\left(x\right)$ is taken to be truncated log-normal with support $\left[0, \varepsilon_{\max}\right]$.

I use standard calibration values of parameters as far as possible. Each period in the model corresponds to one week. The discount factor is taken to be $\beta = 0.9992$ which is equivalent to an annual discount rate of 4%. The process of the aggregate productivity follows Hagedorn and Manovskii (2008). Specifically, we have $\rho = 0.9895$, $\mu_\zeta = 0$ and $\sigma_\zeta = 0.0034$. I set $s = 0.000675$, which is consistent with the mean value of the monthly rate of "other separations" in the Job Openings and Labor Turnover Survey (JOLTS)[13].

On the worker side, I set $b = 0.71$, which follows the derivation of Hall and Milgrom (2008) and is based on a replacement ratio of 0.25. Also, the bargaining power of the worker is set to be 0.7, following Fujita and Ramey (2012) and Mortensen and Nagypal (2007). I assume Hosios condition, which implies $\alpha = \eta$. I set $k = 0.7$, which is close to the estimate in Christiano et al. (2016). Finally, I follow Fujita and Ramey (2012) to set the switching rate of match-specific productivity to be 0.85 and the flow cost of vacancy as $c = 0.17$.[14]

For each fixed value of the net bargaining wedge, the rest of the parameters are calibrated to match the labor market flows in the US economy. The maximum value of match-specific productivity $\varepsilon_{\max}$ is calibrated such that the mean value of the productivity per worker is 1. The coefficient of matching efficiency A is calibrated to fit the model to an average monthly job matching rate of 32%. Finally, I calibrate σ_ε to match an average monthly job separation rate 2% in the data. These calibrations imply a steady state unemployment rate of 5.9%, which is consistent with the long-run average of the data.

For later comparison, I consider also a homogenous version of the model where I assume workers are homogenous in match-specific productivity with no endogenous separation decision. Specifically, I set $\lambda = 0$ and $\chi_t(\varepsilon_t) = 0$ for all ε_t. The model is then calibrated similarly as above, except that now we have $\sigma_\varepsilon = 0$ and s is calibrated to match the average monthly job separation rate in the data.

A summary of the calibration values is given by Table 2. The third column shows the parameter values in the case of the standard model, where the net bargaining wedge is set to be zero. In this case, all job separations are privately efficient. The baseline calibration is shown in the third column, where I calibrate the net bargaining wedge to match the unemployment volatility in the

[13] In the JOLTS, the total separation rate is decomposed into quits, layoffs and discharges, and other separations. I consider other separations corresponding to exogenous separations in the model.

[14] To allow for clearer comparison of results across models, I keep most of the parameters fixed. Alternatively, one could also calibrate some of the above-mentioned parameters to match the moments in the data. For simplicity, I abstract from these alternative calibration strategies. It can be shown that modest changes in these parameters will not change the following results.

Parameters	Meaning	Standard	Baseline	Homogenous
b	flow value of unemployment	0.71		0.71
c	flow cost of posting vacancy	0.17		0.17
β	discount factor	0.99925		0.99925
ρ	persistence of the aggregate productivity	0.9895		0.9895
σ_ζ	standard deviation of the aggregate productivity	0.0034		0.0034
s	exogenous separation rate	0.000675		0.05
η	worker's bargaining power	0.7		0.7
α	elasticity of the matching function	0.7		0.7
λ	switching rate of match-specific productivity	0.085		0
k	fixed matching cost	0.7		0.7
$\varepsilon_{\max}$	maximum value of match-specific productivity	1.2092	1.1867	0.9987
σ_ε	standard deviation of match-specific productivity	0.3169	0.2745	0
A	coefficient of matching efficiency	0.1145	0.1165	0.1116

Table 2: Calibration values

data, see the next subsection for details. The last column shows the calibration of the homogenous model described above with $\Omega_{net} = 0$.

6.3 Unemployment volatility and bargaining friction

The first question to ask is how would the bargaining wedge affect unemployment volatility? It is well-known that the standard MP model fails to address the large unemployment volatility observed in the data. Can it be improved by adding bargaining friction to the model, thereby allowing inefficient job separations?

Figure 2 shows the relationship between unemployment volatility and the net bargaining wedge. The unemployment volatility is measured by the standard deviation of quarterly cyclical unemployment rate over a sample simulation of 1000 quarters. As we can see, unemployment volatility is monotonically increasing in the bargaining wedge. This is due to a combination of the small surplus and the separation channels described earlier: as the bargaining wedge increases, the wages are less sensitive to the aggregate conditions and the firm becomes more likely to fire the worker (inefficiently), hence there are more job separations in the economy, leading to higher level of unemployment fluctuation. In other words, while the workers' productivity varies with the aggregate productivity y_t, the bargaining wedge is relatively acyclical. As a result, a small change in y_t causes significant fluctuations in the cutoff productivity, and hence the quantity of job separations. The

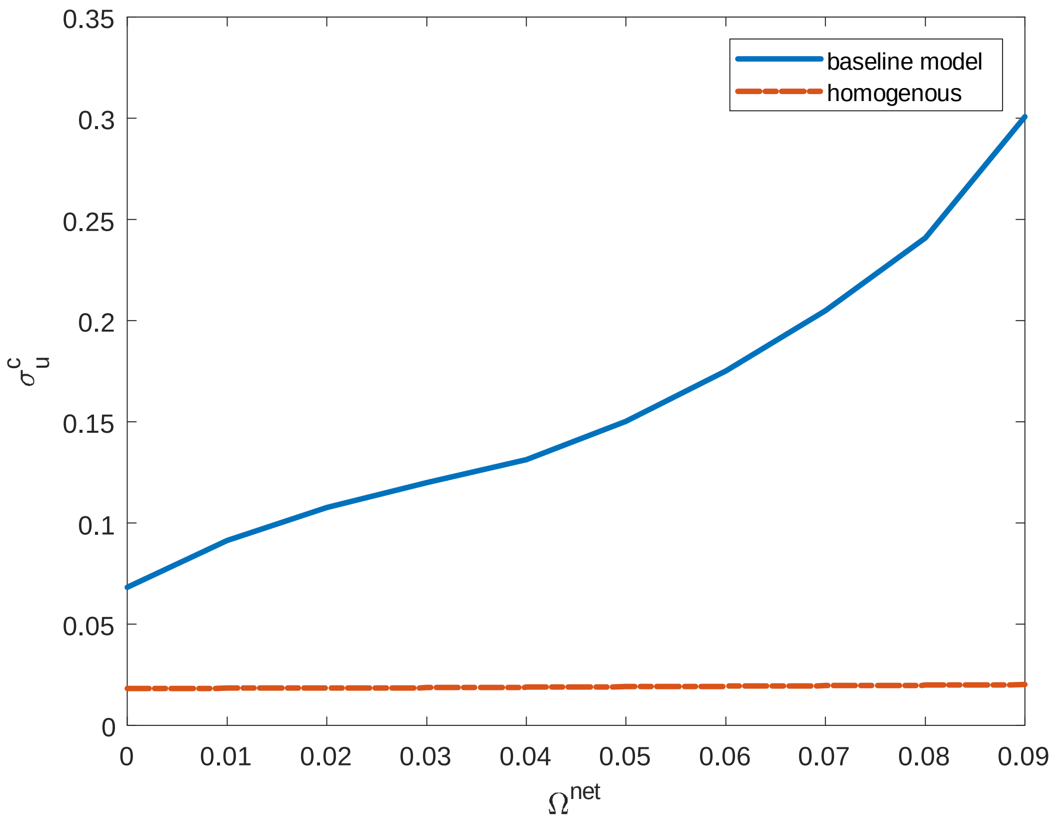

Figure 2: Unemployment volatility and bargaining friction

additional movement in job separations generates excessive unemployment volatility. Note that in principle, an increase in the bargaining wedge keeping other parameters unchanged would increase both the level of unemployment as well as its volatility. The procedure of the calibration described above, however, restricts the focus on the effect on unemployment volatility, since the steady state unemployment rate is kept fixed.

The average volatility of the cyclical unemployment rate in the data is about 11%. If all of the unemployment volatility not explained by the standard model is assumed to be caused by a constant bargaining friction, this would imply a bargaining wedge of 0.02. I will use this calibration as the baseline case. See the third column of Table 2 for the calibrated parameter values[15]. It can be argued, however, that the bargaining wedge is moving with the business cycle. Hence, later in the decomposition of unemployment, I will allow for different values of the bargaining wedge in expansions and in recessions.

To put the size of the bargaining wedge in perspective, the mean value of the worker's surplus in the simulation is about 3.28. Hence, the level of the bargaining wedge required to produce a realistic unemployment volatility is only 0.61% of the worker's surplus, which is hardly a large number. This shows that the unemployment dynamics is in fact very sensitive to the bargaining friction in the model, hence even a small value of the bargaining wedge can generate a large unemployment volatility.

Shown also in the same figure is the unemployment volatility in the homogenous model under different values of Ω^{net}, labeled as "homogenous". Recall that in the homogenous model, bargaining friction has effect on the unemployment volatility only through the small surplus channel, but not the separation channel. Two observations are in order. First, the homogenous model has a much lower level of unemployment volatility even without bargaining friction. This fact is shown also in Fujita and Ramey (2012). Second, it is clear that the homogenous model is much less sensitive to the bargaining friction. In fact, the unemployment volatility increases only from 1.83% to 2.04% as Ω^{net} increases from 0 to 0.09. This shows that in the range of values of Ω^{net} considered here, the small surplus channel is much weaker than the separation channel[16]. As a result, one would

[15] The business cycle statistics of the baseline calibration is shown and analyzed in Appendix C.

[16] Of course, the small surplus channel can become quantitatively significant when Ω^{net} is sufficiently large. Yet, the range of values of Ω^{net} that is consistent with a realistic (endogenous) separation rate is too low to induce any meaningful amplification of unemployment volatility.

need a much higher amount of bargaining friction to generate realistic unemployment fluctuation in the homogenous model. Therefore, we can conclude that much of the increase in unemployment volatility in the baseline model is due to the separation channel.

6.4 Inefficient unemployment

The next question is, how much of the unemployment volatility is explained by the inefficient unemployment? To answer this question, I first use the baseline calibration as described above. The bargaining wedge is identified by matching the average unemployment volatility in expansions and in recessions[17]. Then I match the unemployment rate in the model with the data, by choosing the error terms in the aggregate productivity process (36). The matching of the unemployment rate from 1976 to 2016 is given in panel (a) of Figure 3. Note that recessions defined by the NBER are in grey area. To see the fitness of the model, I have also shown the cyclical component of labor productivity in panel (b). In general, the model does well in matching the relationship between unemployment and cyclical productivity. Finally, I decompose the total unemployment into efficient and inefficient parts, as defined in Section 2.

Figure 4 shows the decomposition of the total unemployment into efficient and inefficient unemployment rates. Note first that the inefficient unemployment appears to be much smaller in value, but is as volatile as the efficient unemployment. Actually, the average inefficient unemployment rate is about 2%, compared to the average efficient unemployment of 4.4%. However, they have the relatively comparable standard derivations of 1.0% and 1.1% respectively. Also, in good times when the total unemployment rate is low, the inefficient-efficient unemployment ratio is also low relative to that in recessions. In fact, the inefficient unemployment rate is less than 1% around 2000. On the other hand, the level of inefficient unemployment rate can be as high as 3.5%, for example, in the Great Recession. Hence, the peak of the unemployment rate in the U.S. during the Great Recession would have been more than 3 percentage points lower if all separations were efficient. The efficient unemployment reflects the effect of efficient separations only, without changing the value of the bargaining wedge. Hence, this shows that inefficient separation is a significant source of unemployment volatility.

[17] See Appendix E for the unemployment volatility and the corresponding Ω_t^{net} in good and bad times. In Appendix E I have also considered the case when the bargaining wedge is constant over the business cycle.

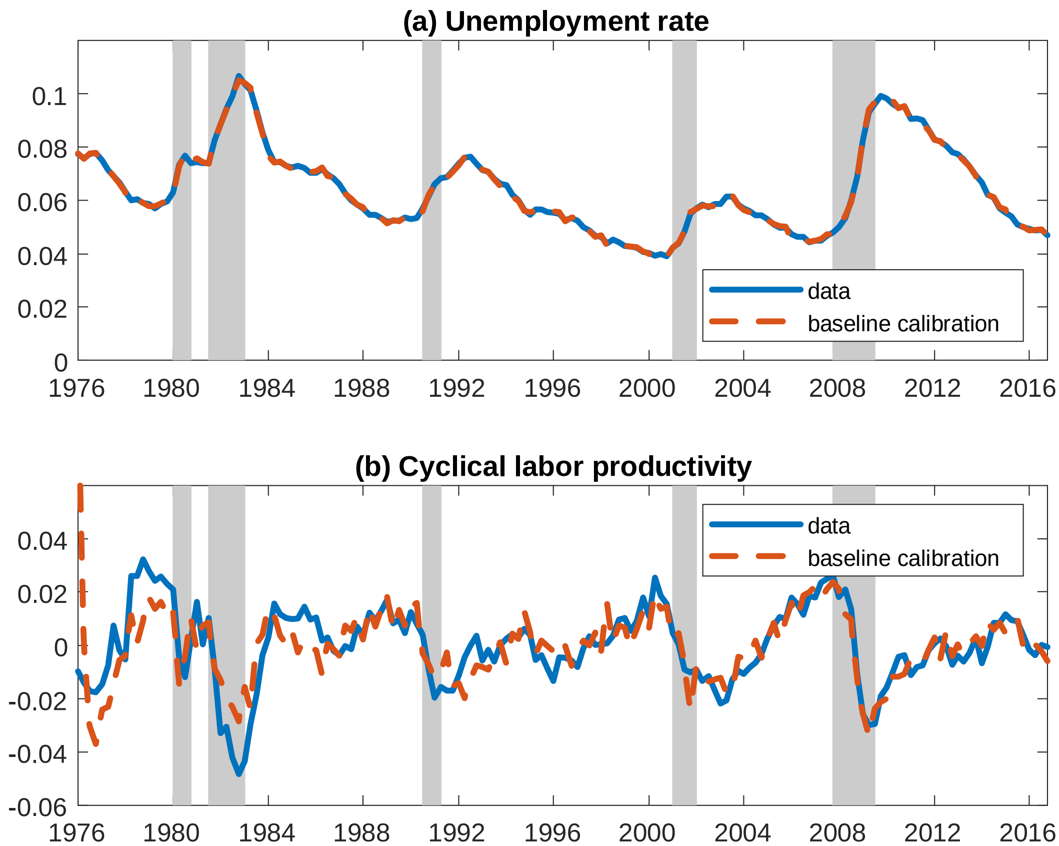

Figure 3: Unemployment and cyclical productivity: data vs. model

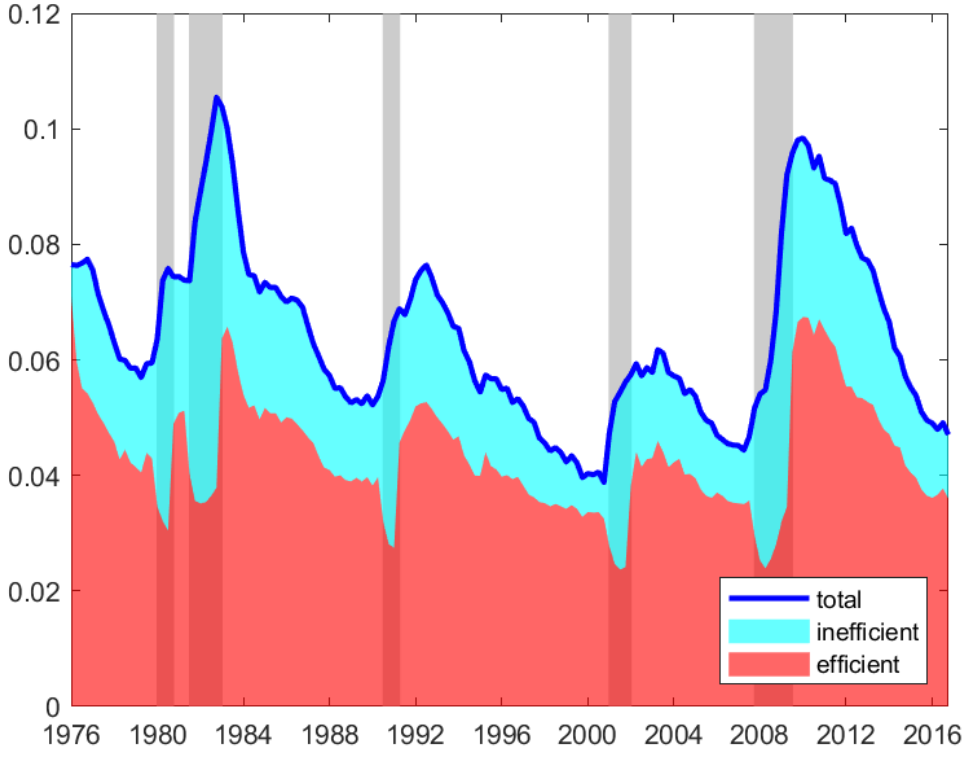

Figure 4: Decomposition of unemployment rate

To see quantitatively how much of the volatility is explained by the inefficient unemployment, I can write, by the decomposition of the unemployment rate,

$$\begin{aligned} var\left(u_t\right) &= cov(u_t, u_t^e + u_t^i) \\ &= cov\left(u_t, u_t^e\right) + cov\left(u_t, u_t^i\right) \end{aligned} \tag{37}$$

Hence, define

$$\gamma = \frac{cov\left(u_t, u_t^i\right)}{var\left(u_t\right)} \tag{38}$$

as the fraction of unemployment volatility explained by the inefficient job separations, and $1-\gamma$ would be that by efficient job separations. In the baseline case, I have $\gamma = 0.54$. Hence, over half of the unemployment volatility can be explained by inefficient job separations, and the remaining 46% is due to efficient separations.[18]

6.5 Worker flows

Recently, there has been increasing attention in the literature on the worker flows in the business cycle frequency[19]. How would the bargaining friction affect the labor market flows? To see this, I use the definition in Section 2 to compute that job finding and separation rates, and also the UE and EU transition rates.

Figure 5 shows the labor market flows for the standard model and for the baseline calibration in a simulation. Note that in both cases, I have calibrated the model to the mean values of the job finding and separation rates. Hence, the focus here is on the volatility of the flows. First, job finding rate is procyclical and the job separation rate is countercyclical. These are consistent with the evidence in Fujita and Ramey (2009). Next, the effect of the bargaining wedge on the job finding rate appears to be insignificant. The job finding rate in the baseline case is slightly more volatile than the standard model.

The main difference between the standard model and the baseline case lies in the volatility of separation rate and the transition flows between the two states. In fact, the cyclical volatility of

[18] By construction, it is assumed that inefficient separation is the explanation for all the unemployment volatility not explained by the standard model. Hence, this can be considered as an upper bound of the explanatory power of bargaining friction.

[19] See, e.g., Davis et al. (2012) and Fujita and Nakajima (2016).

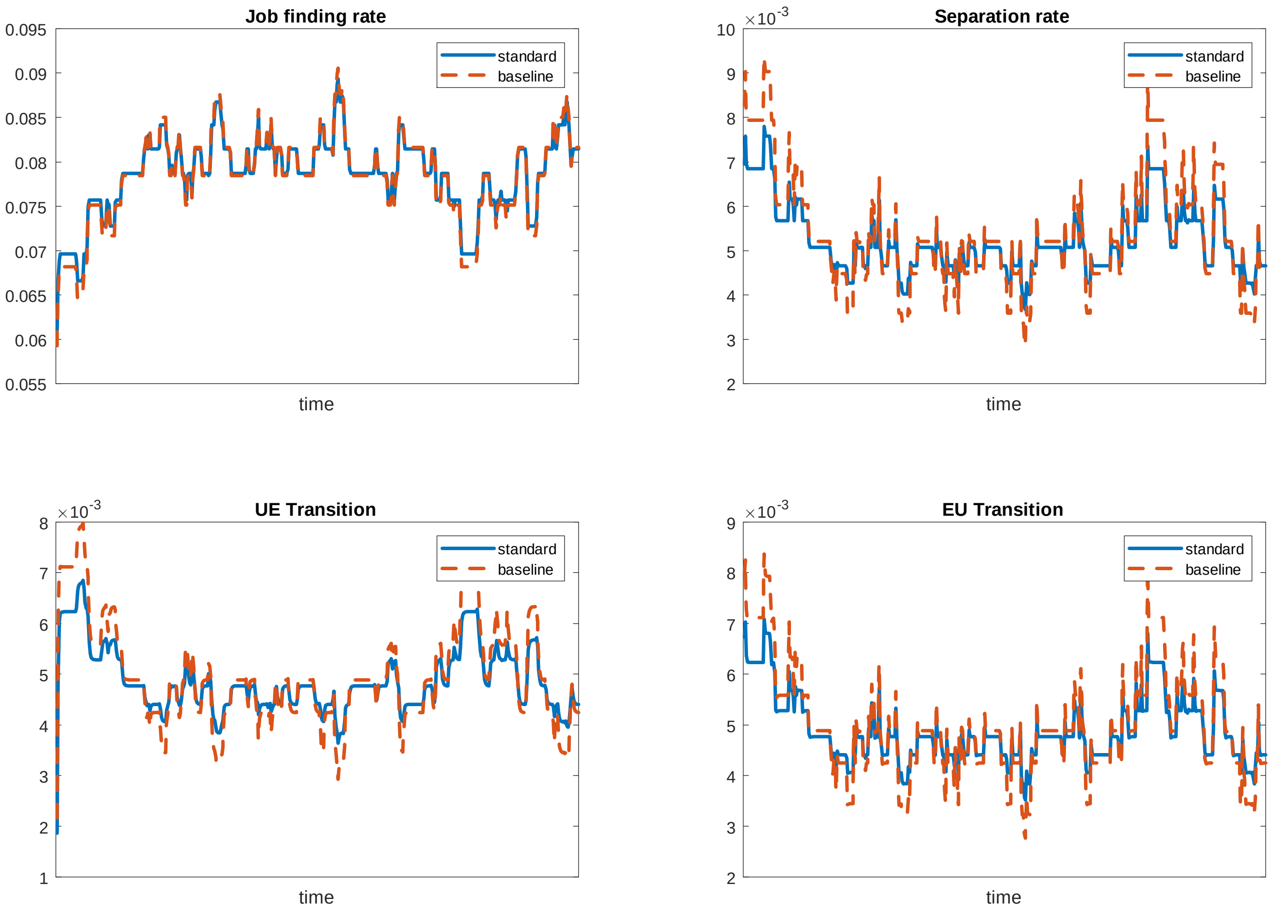

Figure 5: Labor market flows: standard model vs. baseline calibration

separation rate in the baseline case is 10.8% which is much higher than 6.2% in the standard model. The cyclical volatilities of the transition flows also increase considerably (8.4% vs. 4.8% for UE transition, and 10.3% vs. 5.9% for EU transition). On the other hand, the effect on job finding rate is quantitatively insignificant (2.5% vs. 2.2%). Hence, the existence of bargaining friction affects mainly the separation rate in the labor market.

Note that the above result implies that, at the business cycle frequency, the job separation rate is as important as job finding rate in driving the fluctuation of unemployment. This stands in contrast to the evidence in Shimer (2012), who finds that job separation rate is not significant in explaining unemployment volatility. However, Elsby et al. (2009) and Fujita and Ramey (2009) show that job separations are countercyclical and contribute substantially to the unemployment volatility. More recently, Coles and Moghaddasi Kelishomi (2018) estimate that job destructions have been the main driver of unemployment volatility if we relax the free entry assumption. This paper makes no attempt to resolve the debate on the relative importance of ins and outs of unemployment flows. The message here is clear: by allowing inefficient separations by bargaining friction, the separation rate and the worker flows/transitions become much more volatile, while the effect on job finding rate is insignificant.

This quantitative exercise shows that we do not need the small surplus assumption emphasized by Ljungqvist and Sargent (2017), which mainly operates through the job finding rate, to generate sizable unemployment volatility. Even a small amount of bargaining friction could have significant impact on the volatility of job separations, and thus on the unemployment volatility through the separation channel.

7 Conclusion

The contribution of this paper is threefold. First, I propose a simple specification of the bargaining friction by including bargaining wedges to the standard Nash bargaining model. It is shown that due to the misalignment between actual surpluses and the bargaining surpluses, inefficient separations could be generated when the match-specific productivity lies in certain range, which I refer to as the inefficiency gap. The bargaining wedge happens to have important implications to the efficiency of the separations and thus unemployment. Second, I have shown that the solutions to the alternating offers bargaining model and the simple asymmetric information model belong to the same class of the Nash bargaining model with some bargaining wedges. These micro-foundations are, by no means, exhaustive. I expect introducing a wage negotiation cost, for example, would also create some bargaining wedges. Instead of setting up a complicated bargaining environment to generate high unemployment volatility, this paper shows that one can simply introduce bargaining wedges to the standard Nash bargaining problem. In addition, the bargaining wedge specification goes beyond pure generalization of Hall and Milgrom (2008) and Kennan (2010) and give us important insights about inefficient separations and unemployment volatility. Lastly, I have shown quantitatively that bargaining friction can have significant impact on the volatility of the labor market. Even a small value of the bargaining wedge would induce a large unemployment fluctuation. In fact, if we calibrate the model to explain all of the unemployment volatility observed in the data, a decomposition exercise would imply that up to 54% of the cyclical fluctuation can be due to inefficient separations. This paper introduces a novel separation mechanism to generate unemployment volatility through job separations, which is different from the small surplus channel usually assumed in the literature.

The focus of this paper is only on the separation side of the labor market, which is usually ignored by the researchers in the literature by assuming a constant separation rate. Some more work can be done to make the job finding rate more volatile. Ljungqvist and Sargent (2017) survey the literature and they find that the "fundamental surplus" is the key driver of the elasticity of market tightness with respect to productivity. Their focus is on the matching side of the labor market and they again ignore the separation side. I believe a good model of the unemployment dynamics requires a better understanding of both sides of the story. To this end, this paper demonstrates that inefficient separations can be an important part of the model.

Further research can be done to improve the business cycle performance of the model. For example, the common failure of the MP model to produce enough volatility of the vacancy, and hence market tightness, also exists in the presence of bargaining friction. To address the problem, one can model the vacancy creation process by relaxing the free entry condition in the spirit of Fujita and Ramey (2007) and Coles and Moghaddasi Kelishomi (2018). Another extension of the model would be to consider also the on-the-job search. This would likely produce the Beveridge curve observed in the data, as shown in Fujita and Ramey (2012).

Finally, a more detailed identification of inefficient job separations can also be a worthwhile undertaking for future investigation. While, in the quantitative analysis, I identify inefficient job separations by using aggregate labor market statistics, a more direct way would be to identify them from the micro level data, which I will leave for further research. I believe the framework developed in this paper can be potentially useful to investigate both the quantity and the cause of inefficient job separations in the labor market.

References

Anderlini, Luca, & Felli, Leonardo. 2001. Costly Bargaining and Renegotiation. *Econometrica*, **69**(2), 377–411.

Antel, John J. 1985. Costly Employment Contract Renegotiation and the Labor Mobility of Young Men. *American Economic Review*, **75**(5), 976–991.

Barnichon, Regis. 2010. Building a composite Help-Wanted Index. *Economics Letters*, **109**(3), 175–178.

Becker, Gary S. 1962. Investment in Human Capital: A Theoretical Analysis. *Journal of Political Economy*, **70**(5, Part 2), 9–49.

Binmore, Ken, Rubinstein, Ariel, & Wolinsky, Asher. 1986. The Nash Bargaining Solution in Economic Modelling. *The RAND Journal of Economics*, **17**(2), 176.

Blanchard, Olivier, & Galí, Jordi. 2010. Labor Markets and Monetary Policy: A New Keynesian Model with Unemployment. *American Economic Journal: Macroeconomics*, **2**(2), 1–30.

Busch, Lutz-Alexander, & Wen, Quan. 1995. Perfect Equilibria in a Negotiation Model. *Econometrica*, **63**(3), 545.

Chatterjee, Kalyan, & Samuelson, Larry. 1987. Bargaining with Two-sided Incomplete Information: An Infinite Horizon Model with Alternating Offers. *The Review of Economic Studies*, **54**(2), 175.

Christiano, Lawrence J., Eichenbaum, Martin S., & Trabandt, Mathias. 2016. Unemployment and Business Cycles. *Econometrica*, **84**(4), 1523–1569.

Coles, Melvyn G, & Moghaddasi Kelishomi, Ali. 2018. Do Job Destruction Shocks Matter in the Theory of Unemployment. *American Economic Journal: Macroeconomics*, **10**(3), 118–136.

Davis, Steven J., Faberman, R. Jason, & Haltiwanger, John. 2012. Labor market flows in the cross section and over time. *Journal of Monetary Economics*, **59**(1), 1–18.

Elsby, Michael W. L, Michaels, Ryan, & Solon, Gary. 2009. The Ins and Outs of Cyclical Unemployment. *American Economic Journal: Macroeconomics*, **1**(1), 84–110.

Fujita, Shigeru, & Nakajima, Makoto. 2016. Worker flows and job flows: A quantitative investigation. *Review of Economic Dynamics*, **22**(oct), 1–20.

Fujita, Shigeru, & Ramey, Garey. 2007. Job matching and propagation. *Journal of Economic Dynamics and Control*, **31**(11), 3671–3698.

Fujita, Shigeru, & Ramey, Garey. 2009. The Cyclicality of Separation and Job Finding Rates. *International Economic Review*, **50**(2), 415–430.

Fujita, Shigeru, & Ramey, Garey. 2012. Exogenous versus Endogenous Separation. *American Economic Journal: Macroeconomics*, **4**(4), 68–93.

Gertler, Mark, & Trigari, Antonella. 2009. Unemployment Fluctuations with Staggered Nash Wage Bargaining. *Journal of Political Economy*, **117**(1), 38–86.

Gielen, Anne C., & van Ours, Jan C. 2006. Why Do Worker-Firm Matches Dissolve? *IZA Discussion Papers*.

Guerrieri, Veronica. 2008. Inefficient Unemployment Dynamics under Asymmetric Information. *Journal of Political Economy*, **116**(4), 667–708.

Hagedorn, Marcus, & Manovskii, Iourii. 2008. The Cyclical Behavior of Equilibrium Unemployment and Vacancies Revisited. *American Economic Review*, **98**(4), 1692–1706.

Hall, Robert. 2005a. Job Loss, Job Finding, and Unemployment in the U.S. Economy Over the Past Fifty Years. *Pages 101–137 of: NBER Macroeconomics Annual*, vol. 20.

Hall, Robert E. 2005b. Employment Efficiency and Sticky Wages: Evidence from Flows in the Labor Market. *Review of Economics and Statistics*, **87**(3), 397–407.

Hall, Robert E. 2005c. Employment Fluctuations with Equilibrium Wage Stickiness. *American Economic Review*, **95**(1), 50–65.

Hall, Robert E., & Lazear, Edward P. 1984. The Excess Sensitivity of Layoffs and Quits to Demand. *Journal of Labor Economics*, **2**(2), 233–257.

Hall, Robert E., & Milgrom, Paul R. 2008. The Limited Influence of Unemployment on the Wage Bargain. *American Economic Review*, **98**(4), 1653–1674.

Hosios, Arthur J. 1990. On the Efficiency of Matching and Related Models of Search and Unemployment. *The Review of Economic Studies*, **57**(2), 279.

Jäger, Simon, Schoefer, Benjamin, & Zweimüller, Josef. 2019 (jan). *Marginal Jobs and Job Surplus: A Test of the Efficiency of Separations*. Tech. rept. 5. National Bureau of Economic Research, Cambridge, MA.

Kennan, John. 2010. Private Information, Wage Bargaining and Employment Fluctuations. *Review of Economic Studies*, **77**(2), 633–664.

Kennan, John, & Wilson, Robert. 1993. Bargaining with Private Information. *Journal of Economic Literature*, **31**(1), 45–104.

Ljungqvist, Lars, & Sargent, Thomas J. 2017. The Fundamental Surplus. *American Economic Review*, **107**(9), 2630–2665.

McLaughlin, Kenneth J. 1991. A Theory of Quits and Layoffs with Efficient Turnover. *Journal of Political Economy*, **99**(1), 1–29.

Mortensen, Dale, & Nagypál, Éva. 2007. Labor-market Volatility in Matching Models with Endogenous Separations. *Scandinavian Journal of Economics*, **109**(4), 645–665.

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

Myerson, Roger B. 1984. Two-Person Bargaining Problems with Incomplete Information. *Econometrica*, **52**(2), 461.

Petrosky-Nadeau, Nicolas, & Zhang, Lu. 2017. Solving the Diamond-Mortensen-Pissarides model accurately. *Quantitative Economics*, **8**(2), 611–650.

Pissarides, Christopher A. 2009. The Unemployment Volatility Puzzle: Is Wage Stickiness the Answer? *Econometrica*, **77**(5), 1339–1369.

Rubinstein, Ariel. 1982. Perfect Equilibrium in a Bargaining Model. *Econometrica*, **50**(1), 97.

Shimer, Robert. 2005. The Cyclical Behavior of Equilibrium Unemployment and Vacancies. *American Economic Review*, **95**(1), 25–49.

Shimer, Robert. 2012. Reassessing the ins and outs of unemployment. *Review of Economic Dynamics*, **15**(2), 127–148.

Tauchen, George. 1986. Finite state markov-chain approximations to univariate and vector autoregressions. *Economics Letters*, **20**(2), 177–181.

Appendix

A Flow value and present value of wages

The original Mortensen-Pissarides model uses the flow value of wages in defining the value of being employed, which in the model considered here is given by

$$E_t(\varepsilon_t) = w_t(\varepsilon_t) + \beta \mathbb{E}_t \left\{ \begin{array}{c} \left[s + (1-s) \left(\begin{array}{c} \lambda \chi_{t+1}(\varepsilon_{t+1}) \\ + (1-\lambda) \chi_{t+1}(\varepsilon_t) \end{array} \right) \right] U_{t+1} \\ + (1-s) \left[\begin{array}{c} \lambda \left(1 - \chi_{t+1}(\varepsilon_{t+1})\right) E_{t+1}(\varepsilon_{t+1}) \\ + (1-\lambda) \left(1 - \chi_{t+1}(\varepsilon_t)\right) E_{t+1}(\varepsilon_t) \end{array} \right] \end{array} \right\}$$

If we also define the present value of wages as

$$W_t(\varepsilon_t) = w_t(\varepsilon_t) + \beta(1-s) \mathbb{E}_t \left[\begin{array}{c} \lambda \left(1 - \chi_{t+1}(\varepsilon_{t+1})\right) W_{t+1}(\varepsilon_{t+1}) \\ + (1-\lambda) \left(1 - \chi_{t+1}(\varepsilon_t)\right) W_{t+1}(\varepsilon_t) \end{array} \right]$$

Then it is easy to see that

$$V_t(\varepsilon_t) = E_t(\varepsilon_t) - W_t(\varepsilon_t)$$

Moreover, the value of a matched firm is

$$\begin{aligned} F_t(\varepsilon_t) &= P_t(\varepsilon_t) - W_t(\varepsilon_t) \\ &= p_t(\varepsilon_t) - w_t(\varepsilon_t) \\ &\quad + \beta(1-s) \mathbb{E}_t \left[\lambda \left(1 - \chi_{t+1}(\varepsilon_{t+1})\right) F_{t+1}(\varepsilon_{t+1}) + (1-\lambda) \left(1 - \chi_{t+1}(\varepsilon_t)\right) F_{t+1}(\varepsilon_t) \right] \end{aligned}$$

which is consistent with the standard MP model with flow wages. Hence, all the analytical results still hold if wages are expressed in terms of present discounted value $W_t(\varepsilon_t)$ instead of flow value $w_t(\varepsilon_t)$.

B Derivations

B.1 Proof of Proposition 1

By putting $\varepsilon_t = \varepsilon_t^e$ in (13), we have

$$V_t\left(\varepsilon_t^e\right) + W_t\left(\varepsilon_t^e\right) - U_t = \left(1-\eta\right)\Omega_t^w - \eta\Omega_t^f \tag{B.1}$$

$$P_t\left(\varepsilon_t^e\right) - W_t\left(\varepsilon_t^e\right) = -\left[\left(1-\eta\right)\Omega_t^w - \eta\Omega_t^f\right] \tag{B.2}$$

Hence, if $\left(1-\eta\right)\Omega_w - \eta\Omega_f = 0$, then from (B.1) and (B.2), clearly $\varepsilon_t^w = \varepsilon_t^e = \varepsilon_t^f$. On the other hand, if $\left(1-\eta\right)\Omega_t^w - \eta\Omega_t^f > 0$, then from (B.1) and using the fact that $V_t\left(\varepsilon_t\right)$ and $W_t\left(\varepsilon_t\right)$ are strictly increasing, we have $\varepsilon_t^e > \varepsilon_t^w$. Also, from (B.2) and using the fact that $P_t\left(\varepsilon_t\right) - W_t\left(\varepsilon_t\right)$ is also strictly increasing in ε_t, we have $\varepsilon_t^f > \varepsilon_t^e$. Combining, we have $\varepsilon_t^w < \varepsilon_t^e < \varepsilon_t^f$. Similarly, if $\left(1-\eta\right)\Omega_t^w - \eta\Omega_t^f < 0$, then again from (B.1) and (B.2) respectively, we have $\varepsilon_t^e < \varepsilon_t^w$ and $\varepsilon_t^f < \varepsilon_t^e$.

B.2 Proof of Proposition 2

Notice first that when $\lambda = 1$, the value of working V_t is independent of the current idiosyncratic shock. Putting $\varepsilon_t = \varepsilon_t^f$ into (2) and using the definition of ε_t^f, we have

$$W_t\left(\varepsilon_t^f\right) = P_t\left(\varepsilon_t^f\right) = p_t\left(\varepsilon_t^f\right) + \beta\left(1-s\right)\mathbb{E}_t\left(1-\chi_{t+1}\left(\varepsilon_{t+1}\right)\right)P_{t+1}\left(\varepsilon_{t+1}\right) \tag{B.3}$$

Also, putting $\varepsilon_t = \varepsilon_t^e$ into (2) and using the definition of ε_t^e, we have

$$U_t - V_t = P_t\left(\varepsilon_t^e\right) = p_t\left(\varepsilon_t^e\right) + \beta\left(1-s\right)\mathbb{E}_t\left(1-\chi_{t+1}\left(\varepsilon_{t+1}\right)\right)P_{t+1}\left(\varepsilon_{t+1}\right) \tag{B.4}$$

Combining (B.3) and (B.4) yields

$$V_t + W_t\left(\varepsilon_t^f\right) - U_t = p_t\left(\varepsilon_t^f\right) - p_t\left(\varepsilon_t^e\right)$$

Now using the bargaining solution (13) and rearranging, we have

$$p_t\left(\varepsilon_t^f\right) - p_t\left(\varepsilon_t^e\right) = \frac{1}{1-\eta}\left[\left(1-\eta\right)\Omega_t^w - \eta\Omega_t^f\right]$$

Hence, the result (i) follows by putting $p_t(\varepsilon_t) = \bar{p}_t \varepsilon_t$. Similarly, by putting $\varepsilon_t = \varepsilon_t^w$ into (2) again, we have

$$P_t(\varepsilon_t^w) = p_t(\varepsilon_t^w) + \beta(1-s)\,\mathbb{E}_t\left(1 - \chi_{t+1}(\varepsilon_{t+1})\right) P_{t+1}(\varepsilon_{t+1})$$

Now using (B.4) and the definition of ε_t^w, we have

$$\begin{aligned} W_t(\varepsilon_t^w) - P_t(\varepsilon_t^w) &= U_t - V_t - P_t(\varepsilon_t^w) \\ &= p_t(\varepsilon_t^e) - p_t(\varepsilon_t^w) \end{aligned}$$

Finally, using (13) yields

$$p_t(\varepsilon_t^e) - p_t(\varepsilon_t^w) = \frac{1}{\eta}\left[(1-\eta)\,\Omega_w - \eta\Omega_f\right]$$

Hence, the result (ii) follows by noting $p_t(\varepsilon_t) = y_t \varepsilon_t$.

B.3 Proof of Proposition 3

Solving (24) and (25), the unique SPNE wage is given by

$$W_t(\varepsilon_t) = \frac{1}{2-\delta}\left[(1-\delta)\,P_t(\varepsilon_t) + U_t - V_t(\varepsilon_t) + \frac{(1-\delta)^2}{\delta}\gamma_t + \frac{1-\delta}{\delta}z_t\right]$$

Rearranging, we have

$$\frac{1-\delta}{2-\delta}\left[P_t(\varepsilon_t) - W_t(\varepsilon_t) - \left(-\frac{1-\delta}{\delta}\gamma_t\right)\right] = \frac{1}{2-\delta}\left[V_t(\varepsilon_t) + W_t(\varepsilon_t) - U_t - \frac{1-\delta}{\delta}z_t\right]$$

Hence the result follows by comparing with (13).

B.4 Proof of Corollary 4

From Proposition 2 and 3,

$$\varepsilon_t^f - \varepsilon_t^e = \frac{1}{y_t}\left[\frac{1-\delta}{\delta}z_t + \frac{(1-\delta)^2}{\delta}\gamma_t\right]$$

Differentiating, we have

$$\begin{aligned}
\frac{\partial\left(\varepsilon_t^f-\varepsilon_t^e\right)}{\partial z_t} &= \frac{1}{y_t}\left[\frac{1-\delta}{\delta}\right] \\
\frac{\partial\left(\varepsilon_t^f-\varepsilon_t^e\right)}{\partial \gamma_t} &= \frac{1}{y_t}\left[\frac{(1-\delta)^2}{\delta}\right] \\
\frac{\partial\left(\varepsilon_t^f-\varepsilon_t^e\right)}{\partial \delta} &= -\frac{1}{y_t}\left[\frac{z_t+\left(1-\delta^2\right)\gamma_t}{\delta^2}\right]
\end{aligned}$$

B.5 Fundamental surplus in Ljungqvist and Sargent (2017)

Consider the steady state MP model with exogenous separation and homogenous productivity in Ljungqvist and Sargent (2017). In their notations, the value of a filled vacancy is given by

$$J=y-w+\beta\left[sV+(1-s)\,J\right] \tag{B.5}$$

where y is the productivity, w is the flow wage, $\beta=(1+r)^{-1}$ is the discount factor, and V is the value of an unfilled vacancy. Free entry condition of the firms implies $V=0$ and

$$c=\beta q\left(\theta\right)J \tag{B.6}$$

Combining (B.5) and (B.6), we have

$$w=y-\frac{r+s}{q\left(\theta\right)}c \tag{B.7}$$

On the worker's side, the values of being employed and unemployed are given respectively by

$$E = w+\beta\left[sU+(1-s)\,E\right] \tag{B.8}$$

$$U = z+\beta\left\{\theta q\left(\theta\right)E+\left[1-\theta q\left(\theta\right)\right]U\right\} \tag{B.9}$$

Combining (B.8) and (B.9) yields

$$E-U=\frac{w-z}{1-\beta\left(1-s\right)+\beta\theta q\left(\theta\right)} \tag{B.10}$$

Suppose the wage rate is determined by the NBBW model, i.e. the Nash product becomes

$$(E - U - \Omega^{w})^{\eta} \left(J - \Omega^{f}\right)^{1-\eta}$$

The solution is given by

$$(1-\eta)(E-U) = \eta J + \Omega^{net}$$

which using (B.5), (B.6), and (B.10) implies that

$$w = (1-\eta) z + \eta (y + \theta c) + \frac{\Omega^{net} (r + s + \theta q(\theta))}{1+r} \tag{B.11}$$

Now equating (B.7) with (B.11) and rearranging, we have

$$\frac{1-\eta}{c}\left[y - z - \frac{\Omega^{net}(r+s)}{(1-\eta)(1+r)}\right] = \frac{r+s}{q(\theta)} + \eta\theta + \frac{\theta q(\theta)}{c(1+r)}\Omega^{net}$$

Hence, implicitly differentiating and rearranging, we get

$$\begin{aligned} \frac{d\theta}{dy} &= \frac{r + s + \eta\theta q(\theta)\left(1 + \beta q(\theta)\frac{\Omega^{net}}{c\eta}\right)}{\alpha(r+s) + \eta\theta q(\theta)\left(1 + (1-\alpha)\beta q(\theta)\frac{\Omega^{net}}{c\eta}\right)} \frac{\theta}{y - z - \frac{1-\beta(1-s)}{(1-\eta)}\Omega^{net}} \\ &= \Upsilon\left(\Omega^{net}\right) \frac{\theta}{y - z - \frac{1-\beta(1-s)}{(1-\eta)}\Omega^{net}} \end{aligned}$$

where $\alpha = \frac{-\theta q'(\theta)}{q(\theta)}$ is the elasticity of the matching function. Hence, the result follows by noting that $\epsilon_{\theta,y} = \frac{y}{\theta}\frac{d\theta}{dy}$.

C Business cycle statistics of the model (baseline calibration)

	p_t	u_t	FR_t	SR_t	UE_t	EU_t	v_t	θ_t
σ_x^c	0.011	0.11	0.025	0.108	0.084	0.102	0.052	0.065
$\epsilon_{x,p}$	1.000	-8.978	2.245	-8.713	-6.608	-8.167	-3.171	5.810
$corr\left(x_t, x_{t-1}\right)$	0.807	0.859	0.821	0.651	0.856	0.626	0.800	0.822
Correlation matrix								
	p_t	u_t	FR_t	SR_t	UE_t	EU_t	v_t	θ_t
p_t	1.000	-0.926	0.984	-0.899	-0.871	-0.886	-0.684	0.992
u_t		1.000	-0.950	0.858	0.990	0.838	0.903	-0.940
FR_t			1.000	-0.879	-0.902	-0.862	-0.728	0.997
SR_t				1.000	0.793	0.999	0.657	-0.895
UE_t					1.000	0.771	0.951	-0.887
EU_t						1.000	0.634	-0.879
v_t							1.000	-0.704
θ_t								1.000

Table C.1: Business cycle statistics (baseline calibration)

The business cycle statistics of the baseline calibration is given in Table C.1. Comparing with the Table 1, we can see that the model matches well for the correlation matrix and the autocorrelation of the variables. A few significant differences can be observed. First, the vacancy of job finding rate and the market tightness in the model are much lower than those in the data. This is related to Shimer (2005)'s finding that the MP model fails to generate enough volatility of the market tightness. This problem can most likely be solved by relaxing the free entry condition, as shown in Coles and Moghaddasi Kelishomi (2018). Second, the model fails to produce the Beveridge curve (i.e. negative relationship between unemployment and vacancy). This is due to the fact that the volatility of separation rate in the model is too high compared with the data. As demonstrated in Fujita and Ramey (2012), this likely can be solved by introducing on-the-job search. Instead of matching all the summary statistics, I abstract from these complications by looking exclusively how inefficient separations affect the labor market dynamics.

D Computation strategy

First, the stochastic processes of the aggregate and idiosyncratic productivity are discretized as follows. The AR(1) stochastic process (36) is approximated by a finite state markov-chain with N states and a transition probability matrix $[\pi_{i,j}]$ by using the Tauchen (1986) procedure. Also, I approximate the truncated log-normal distribution with K points of support $\{\varepsilon_1, ..., \varepsilon_K\}$, with $G(\varepsilon_K) = 1$. Let $g(\cdot)$ be the probability mass function of the discretized distribution. In the computation exercise, I set $N = 10$ and $K = 400$.

Next, I solve the model by using a fixed point algorithm. Since only the net bargaining wedge matters, I assume $\Omega^f = -\Omega^w \equiv \Omega$. In this case, we have $\Omega^{net} = \Omega$. Then I use a standard iteration procedure to find a fixed point. Specifically, given initial guess of $\{P_0, W_0, V_0, U_0, \theta_0, \chi_0\}$, I update the value functions and the market tightness as follows (derived from equations (1) - (5) and (14)).

$$
\begin{aligned}
\theta_{n+1}(y_i) &= \left\{ \begin{array}{c} \left(\frac{\beta A}{c}\right) \sum_{j=1}^{N} \sum_{k=1}^{K} g(\varepsilon_k)\, \pi_{i,j} \\ \left[1-\chi_n(y_j,\varepsilon_k)\right]\left[P_n(y_j,\varepsilon_k) - W_n(y_j,\varepsilon_k) - k\right] \end{array} \right\}^{\frac{1}{\alpha}} \\
P_{n+1}(y_i,\varepsilon_l) &= y_i\varepsilon_l + \beta(1-s)\sum_{j=1}^{N} \pi_{i,j} \left[\begin{array}{c} \lambda \sum_{k=1}^{K} g(\varepsilon_k)\left[1-\chi_n(y_j,\varepsilon_k)\right] P_n(y_j,\varepsilon_k) \\ +(1-\lambda)\left[1-\chi_n(y_j,\varepsilon_l)\right] P_n(y_j,\varepsilon_l) \end{array} \right] \\
V_{n+1}(y_i,\varepsilon_l) &= \beta \sum_{j=1}^{N} \pi_{i,j} \left\{ \begin{array}{c} \left[s + (1-s) \left[\begin{array}{c} \lambda \sum_{k=1}^{K} g(\varepsilon_k)\, \chi_n(y_j,\varepsilon_k) \\ +(1-\lambda)\,\chi_n(y_j,\varepsilon_l) \end{array} \right] \right] U_n(y_j) \\ +(1-s) \left[\begin{array}{c} \lambda \sum_{k=1}^{K} g(\varepsilon_k)\left[1-\chi_n(y_j,\varepsilon_k)\right] V_n(y_j,\varepsilon_k) \\ +(1-\lambda)\left[1-\chi_n(y_j,\varepsilon_l)\right] V_n(y_j,\varepsilon_l) \end{array} \right] \end{array} \right\} \\
U_{n+1}(y_i) &= b + \beta \sum_{j=1}^{N} \sum_{k=1}^{K} g(x_k)\, \pi_{i,j} \\
&\quad \left[\begin{array}{c} A\theta_n(y_i)^{1-\alpha}\left[1-\chi_n(y_j,\varepsilon_k)\right]\left[W_n(y_j,\varepsilon_k) + V_n(y_j,\varepsilon_k)\right] \\ +\left(1 - A\theta_n(y_i)^{1-\alpha} + A\theta_n(y_i)^{1-\alpha}\chi_n(y_j,\varepsilon_k)\right) U_n(y_j) \end{array} \right] \\
W_{n+1}(y_i,\varepsilon_l) &= \eta P_{n+1}(y_i,\varepsilon_l) + (1-\eta)\left(U_{n+1}(y_i) - V_{n+1}(y_i,\varepsilon_l)\right) + \Omega \\
\chi_{n+1}(y_i,\varepsilon_l) &= \left\{ \begin{array}{ll} 1 & \begin{array}{c} P_{n+1}(y_i,\varepsilon_l) \le W_{n+1}(y_i,\varepsilon_l) \text{ or} \\ W_{n+1}(y_i,\varepsilon_l) + V_{n+1}(y_i,\varepsilon_l) \le U_{n+1}(y_i) \end{array} \\ 0 & \qquad\qquad \text{otherwise} \end{array} \right.
\end{aligned}
$$

for all $i = 1, ..., N$ and $l = 1, ..., K$. The error of iteration is defined by

$$\max\{\|P_{n+1} - P_n\|, \|V_{n+1} - V_n\|, \|U_{n+1} - U_n\|, \|\theta_{n+1} - \theta_n\|\}$$

where $\|\cdot\|$ is the entrywise matrix norm. The iteration continues until the error is less than the tolerance level. In the computation exercise, I set the tolerance level to be 10^{-6}. From the solution of $\{P, W, V, U\}$ from above, I then find the cutoff productivities $\{\varepsilon^w, \varepsilon^f, \varepsilon^e\}$ from the definitions (9) - (11).

To compute the unemployment volatility, I simulate a time series of unemployment by using the flow equations (16) - (18). I choose the sample size to be $12,000$ weeks, which is equivalent to 1000 quarters. The weekly series to transformed to quarterly frequency by calculating the simple average in a quarter. Finally, the log deviation of the quarterly series computed by using a hp-filter with a coefficient of 1600, and the cyclical part is used to measure the unemployment volatility.

E Identification of the bargaining wedge

E.1 Baseline identification

	Unemployment volatility	Corresponding Ω^{net}
Expansions	0.104	0.0176
Recessions	0.139	0.0441

Table E.1: Baseline identification of the bargaining wedge

In the decomposition of the unemployment rate, the bargaining wedge is identified by matching the unemployment volatility in expansions and in recessions respectively. The monotonicity of unemployment volatility and bargaining friction in Figure 2 allows us to match the unemployment volatility with some corresponding bargaining wedge. Table E.1 shows the unemployment volatility in good and bad times and the corresponding bargaining wedges.

E.2 Constant bargaining wedge

Here I consider another identification of the bargaining wedge by assuming that the bargaining wedge is constant over the business cycle. As mentioned before, in the data, the average unemployment volatility is about 11%, entailing a bargaining wedge of 0.02. I then perform the same procedure to decompose the unemployment rate. Figure E.1 shows the results.

We can see a similar decomposition as in the case of cyclical bargaining wedge. A notable difference is that the efficient unemployment drops during the recessions in the baseline case, while it increases in the case of constant bargaining wedge. In this case, we have $\gamma = \frac{cov(u_t, u_t^i)}{var(u_t)} = 0.44$. Hence, an acyclical bargaining wedge alone can explain about 44% of the unemployment volatility in the data. Adding cyclicality to the bargaining wedge increases the explanatory power by another 10%.

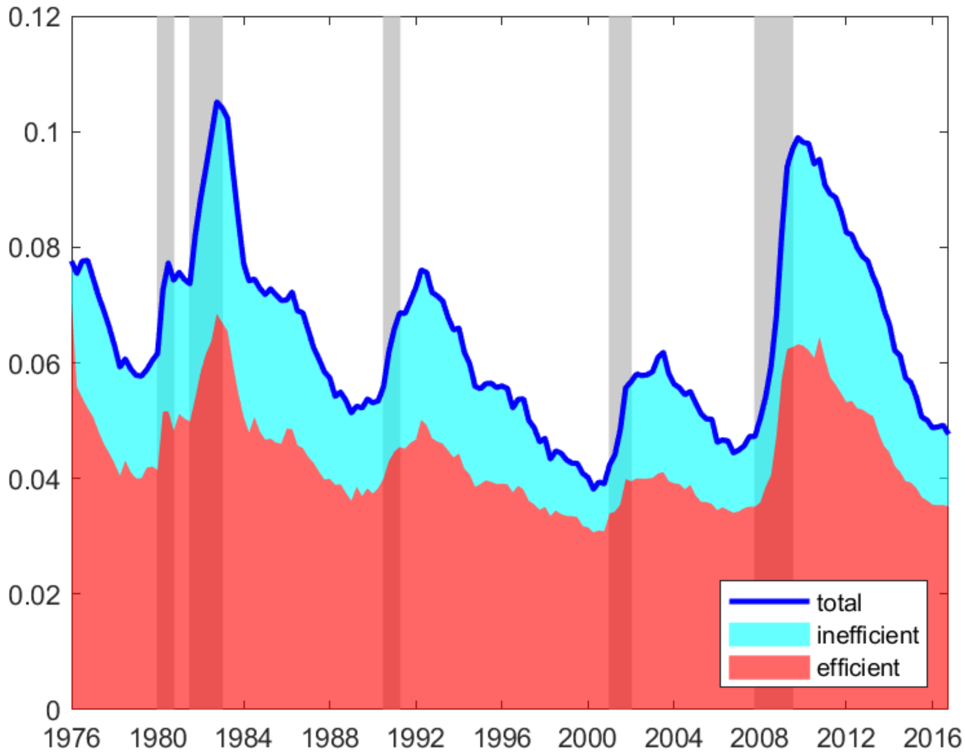

Figure E.1: Decomposition of unemployment rate (constant bargaining wedge)