Unions in a Frictional Labor Market*

Per Krusell
Stockholm University, CEPR, and NBER
per.krusell@iies.su.se

Leena Rudanko
Boston University and NBER
rudanko@bu.edu

April 7, 2012

Abstract

We analyze a labor market with search and matching frictions where wage setting is controlled by a monopoly union. We take a benevolent view of the union in assuming it to care equally about employed and unemployed workers and we assume, moreover, that it is fully rational, thus taking job creation into account when making its wage demands. Under these assumptions, if the union is also able to fully commit to future wages it generates an efficient level of long-run unemployment. However, in the short run, it uses its market power to collect surpluses from firms with existing matches by raising current wages above the efficient level. These elements give rise to a time inconsistency. Without commitment, and in a Markov-perfect equilibrium, not only is unemployment well above its efficient level, but the union wage also exhibits stickiness which amplifies the responses of vacancy creation and unemployment to shocks. We consider extensions to partial unionization and collective bargaining between a labor union and an employers' association.
JEL classification: E02, E24, J51, J64
Keywords: Labor unions, frictional labor markets, time-inconsistency

*We are grateful to audiences at BU, IIES, the Society of Economic Dynamics meeting in Ghent, and in particular William Hawkins, Fabien Postel-Vinay, and Robert Shimer for comments. Mirko Fillbrunn provided research assistance.

1 Introduction

Labor unions play an important role in many labor markets in many countries. There is also a large literature within labor economics studying how union presence influences labor-market outcomes. Yet there is relatively little work studying the impact of this institution on the aggregate labor market when this market is described as having frictions and featuring unemployment due to these frictions. Since search and matching models have come to play a central role as a workhorse for macroeconomic labor-market analyses, this gap in the literature leaves open important questions. What is the impact of unions on aggregate unemployment, and wages? How do unions affect how strongly unemployment varies over the business cycle? What institutional settings are desirable, when considering implementing rules regarding union coverage or centralized bargaining between a union and employer representatives? A circumstance of particular relevance, especially for many European economies, is one where there is a nationwide union, or cooperation/agreements among unions representing different industries. In these cases, unions directly take on an aggregate role. To develop our understanding of the impact of a large union on the aggregate economy, this paper develops a dynamic model of unionized frictional labor markets. Using this model, we then examine, in turn, the union impact on wage setting in the long run, in response to shocks, and in settings where the institutional details differ.

Our first and arguably most general finding is that the degree to which the union can commit to future wage setting is qualitatively and quantitatively important for outcomes. We start from a rather benevolent presumption about the union: it cares equally about employed and unemployed workers. Second, we assume that the union is fully rational, thus taking job creation into account when making its wage demands. Current job creation helps currently unemployed workers, but comes about by lowering wage demands, hurting currently employed workers. The union thus sets wages trading these effects against each other. Under the additional assumption that the union is also able to fully commit to future wages, we show that the outcome is an efficient level of long-run unemployment. In the short run, however, unemployment is inefficiently high as the union uses its market power to collect surpluses on existing matches by raising current wages above the efficient level. More precisely, we demonstrate that labor-market tightness is inefficient in the very first period but set efficiently from the second period and on.

These elements give rise to a time inconsistency. That is, if a union had implemented a commitment plan yesterday but had the opportunity to revise it today, it would indeed revise it and lower labor-market tightness relative to the plan, thus benefitting again from the pre-existing matches. What, then, would the outcome be if one simply assumed that unions do not have commitment? We answer this question by analyzing Markov-perfect equilibria. In these equilibria, we show, unemployment is above its efficient level both in the short and in the long run. The longer is the horizon of commitment, the weaker is this effect. For an annual commitment horizon, the effects on unemployment are still quite sizable: unemployment without commitment is almost twice the efficient level, and the output loss is 4% of GDP per period.

An important reason macroeconomists have been interested in labor unions is the notion that unions create rigidity in wages, which may help reconcile the large variation in employment over the business cycle with macroeconomic theory. For example Blanchard and Fischer (1989) discuss unions in this context, offering an overview of the basic theories of union wage-setting. We build on these theories by incorporating them into a framework with an explicitly frictional labor market, which highlights the dynamic nature of the union problem. Interestingly, we find that economies with large unions—which are not able to commit to future wages—also display short-run wage stickiness, which amplifies the responses of unemployment and vacancy creation to shocks to labor productivity. To understand the source of this short-run stickiness, note that the union's incentive to distort wages up depends on the level of employment—the higher is employment, the stronger is the incentive to hike up wages to collect rents from the employed. When labor productivity increases, vacancy-creation increases, and employment thus begins to rise over time in response. During this transition path to higher employment, the union wage distortion strengthens. The dynamics of wages thus exhibit stickiness: in addition to wages increasing on impact in response to a shock, they continue to rise with employment, before reaching their full response. Symmetrically, in response to a negative shock, wages continue to fall as employment falls over time, reaching their full response only as employment does.

Throughout the analysis, our analytical work-horse, both for qualitative analysis of the different forces underlying equilibria and for numerical computation, is the Euler equation of the wage-setting union. This equation is readily compared to its efficient equivalent, and the Euler equation under commitment can also be compared to that without.

Although our focus is on a setting with universal union coverage, we also consider economies with less than full unionization of workers. In order to side-step the complex issue of how union objectives change over time as more or less workers are unionized, we consider the case of a constant unionization rate, where a fixed subset of workers are union members, and the remainder bargain individually with firms. In doing so, we assume that firms cannot discriminate workers based on union membership; thus, they search in an indirected manner and may end up being matched either with a union worker or with a non-union worker. A special case of this setting is one where the unionization rate is such that union and non-union wages are the same, and workers thus indifferent about being unionized or not. This outcome is possible if individual workers have strong bargaining power. We demonstrate that in this case, a law requiring universal coverage of union wages can be welfare-enhancing. However, if individual workers have low bargaining power, union members always earn higher wages than non-union workers, and outlawing unions would improve welfare.

Finally, we examine collective bargaining: a Nash bargaining game between a centralized labor union and an employers' association. This game leads to the same general conclusion as in our simple monopoly union case: under commitment, outcomes are inefficient only in the short run, and labor market tightness at the efficient level after the initial period. However, the direction of the inefficiency—whether market tightness is above or below the efficient level—depends on the relative bargaining strength of the union vis-a-vis the employers' association. We show an illustrative example where, under limited commitment, a union bargaining power close to (but strictly less than) one leads to an efficient outcome.

Within the literature on labor unions, this paper is most closely related to two strands: i) a set of papers considering the dynamic decision problem of a labor union when labor is subject to adjustment costs, and ii) a second set of papers incorporating labor unions into the Mortensen-Pissarides search and matching framework, work which has so far largely focused on static union problems.[1]

The first set of papers develops the idea that dynamic concerns become important in union decision making when labor markets are frictional. The most directly related papers in this vein are Lockwood and Manning (1989) and Modesto and Thomas (2001). These papers study labor markets where firms face adjustment costs to labor, and forward-looking unions take these adjustment costs into account in deciding on their wage demands.[2] Modesto and Thomas (2001) introduce the idea that the union's ability to commit to future wage demands matters in this setting and contrast the differences between outcomes with and without union commitment. The simple reduced-form adjustment cost framework allows these authors to derive closed-form results which speak to the level of union wage-demands, as well as the speed of adjustment in employment, both argued to be greater in a unionized labor market than a non-unionized one. We, on the other hand, study dynamic union decision-making within the context of the Mortensen-Pissarides search and matching model—the modern workhorse model of frictional labor markets—where such adjustment costs are endogenous. This allows us to study the impact of unions on equilibrium unemployment, vacancy creation, output, and welfare, something the adjustment-cost framework cannot directly address.

[1] Further recent work on labor unions in macroeconomics includes Acemoglu, Aghion, and Violante (2001), Alvarez and Shimer (2009), and Greenwood (2010). While these papers adopt varying approaches to modeling the labor market, they all depart from the Mortensen-Pissarides framework.

[2] Other papers which feature unions in settings where labor adjustment occurs slowly due to adjustment costs or otherwise, but focus on other issues, include Booth and Schiantarelli (1987), Card (1986), and Kennan (1988). Some authors have also sought to develop dynamic models of unions building on the insider-outsider theory of Lindbeck and Snower (1986), such as Huizinga and Schiantarelli (1992). In order to focus our analysis, we abstract from insider-outsider concerns here.

The second set of papers incorporate unions into models of labor markets with search frictions. Perhaps the closest in spirit to our paper in this group is Pissarides (1986), which first introduces a monopoly union into the Pissarides (1985) framework, and studies the impact on equilibrium outcomes in the labor market. Like most of the literature following it, that paper focuses on steady states, side-stepping the dynamic issues we highlight here. Mortensen and Pissarides (1999) proceed to incorporate a notion of wage-compression into the analysis, allowing for worker heterogeneity. Wage compression plays a central role in subsequent work on unions in the Mortensen-Pissarides framework. For example Garibaldi and Violante (2005) and Boeri and Burda (2009) study the effects of employment protection in a frictional labor market where a centralized labor union compresses wages. Some papers depart from the assumption of a centralized union, allowing firm-level unions instead, and also speak to the degree of unionization, such as Ebell and Haefke (2006), Acikgoz and Kaymak (2009), and Taschereau-Dumouchel (2011). Finally, Delacroix (2006) extends the framework of Ebell and Haefke (2006) to industry-level unions, illustrating the non-monotonic relationship between the degree of coordination in bargaining and economic performance discussed by Calmfors and Driffill (1988).[3]

Our paper is organized as follows. Section 2 analyzes the benchmark model—first a one-period model to set out notation and introduce the key elements, then an infinite-horizon model with commitment, and lastly the infinite-horizon model without commitment. Section 3 provides the quantitative analysis and Section 4 looks at extensions: partial unionization in Section 4.1 and collective bargaining in Section 4.2. Section 5 concludes.

2 The benchmark model

This section begins by describing the simple Mortensen-Pissarides search and matching environment we base our analysis on. We then introduce a monopoly union into that framework,

[3]Further examples of work on unions in a search framework include Mortensen (1989) on multiple equilibria and Burdett and Wright (1993) on the impact of unions under non-transferable utility.

and characterize its behavior. We consider extensions to partial unionization and collective bargaining later on.[4]

A frictional labor market Time is discrete and the horizon infinite. The economy is populated by a continuum of measure one identical workers, together with a continuum of identical capitalists who employ these workers. All agents have linear utility, and discount the future at rate $\beta < 1$. Capitalists have access to a linear production technology, producing z units of output per period for each worker employed.

The labor market is frictional, requiring capitalists seeking to hire workers to post vacancies. The measure of matches in the beginning of the period is denoted by $n \in [0, 1]$, leaving $1 - n$ workers searching for jobs. Searching workers and posted vacancies are matched according to a constant-returns-to-scale matching function $m(v, 1-n)$, where v is the measure of vacancies. With this, the probability with which a searching worker finds a job within a period can be written $\mu(\theta) = m(\theta, 1)$; the probability with which a vacancy is filled is $q(\theta) = m(1, \theta^{-1})$, where $\theta = v/(1 - n)$ is the labor market tightness. We will assume that $\mu'(\theta)$ is positive and decreasing and that $q'(\theta)$ is negative and increasing. With this, employment equals n plus the number of new matches, $\mu(\theta)(1 - n)$. Jobs are destroyed each period with probability δ. Thus, the measure of matches evolves over time according to the law of motion

$$n_{t+1} = (1 - \delta) \underbrace{(n_t + \mu(\theta_t)(1 - n_t))}_{\text{employed}_t} . \tag{1}$$

Notice that a worker separated after production at t may be reemployed in $t + 1$ and not need to suffer unemployment.

In addition to the market production technology, unemployed workers also have access to a home production technology, producing $b(< z)$ units of output per period.

[4] See Section 4.

Firms Capitalists operate production through firms, and these firms need to post vacancies in order to find workers, at a cost κ per vacancy. Competition drives profits from vacancy-creation to zero, with firms taking into account the union wage-setting behavior today and in the future. The zero-profit condition thus determines the current labor-market tightness according to current and future wages as follows:

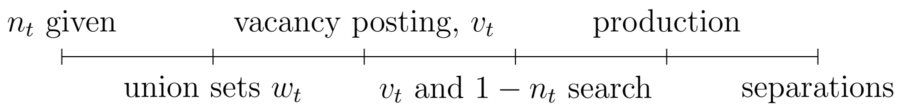
$$\kappa = q(\theta_t) \sum_{s=0}^{\infty} \beta^s (1-\delta)^s \big[z - w_{t+s} \big]. \tag{2}$$

The labor union Wages are set unilaterally by a labor union, with universal coverage. The union sets wages to maximize the welfare of all workers, but cannot discriminate among workers in doing so.[5] The union objective thus becomes

$$\sum_{t=0}^{\infty} \beta^t \big[\underbrace{\big(n_t + \mu(\theta_t)(1-n_t)\big)}_{\text{employed}_t} w_t + \underbrace{(1-n_t)(1-\mu(\theta_t))}_{\text{unemployed}_t} b \big] \tag{3}$$

The union takes as given the evolution of employment according to equation (1). It also internalizes the effect of its wage-setting decisions on hiring. Therefore, the union's problem is to choose a sequence of wages $\{w_t\}_{t=0}^{\infty}$ to maximize the objective (3) subject to the law of motion (1) and zero profit condition (2). Below, we will consider different assumptions regarding the union's commitment power.

Summarizing the events in period t, we have



Given the path of wages $\{w_t\}_{t=0}^{\infty}$, then, equation (2) determines the path of the vacancy-unemployment ratio $\{\theta_t\}_{t=0}^{\infty}$, which in turn determines the evolution of employment $\{n_t\}_{t=0}^{\infty}$.

[5]In particular, we rule out favorable treatment of "insiders" versus "outsiders" by assuming that all workers are paid the same wage.

2.1 A one-period example

To illustrate key forces at play, we first consider the impact of the union in a very simple setting: a one-period version of the above economy. Many of the features present here will be present in the subsequent analysis.

A natural starting point is the efficient benchmark—the output maximizing level of vacancy-creation a social planner would choose. Here the planner solves the problem

$$\max_{\theta} \big(\underbrace{n + \mu(\theta)(1-n)}_{\text{employed}}\big) z + \underbrace{(1-n)\big(1-\mu(\theta)\big)}_{\text{unemployed}} b - \underbrace{\theta(1-n)}_{\text{vacancies}} \kappa,$$

taking as given initial matches n. The planner's optimum is characterized by the first-order condition $-\kappa + \mu'(\theta)(z-b) = 0$, which pins down θ independent of n. For concreteness, consider the matching function $m(v,u) = vu/(v+u)$, such that $\mu(\theta) = \theta/(1+\theta)$. In this case the planner's optimum is given by $\theta^e = \sqrt{(z-b)/\kappa} - 1$, with labor-market tightness an increasing function of market productivity.

The union instead aims to maximize

$$\big(\underbrace{n + \mu(\theta)(1-n)}_{\text{employed}}\big) w + \underbrace{(1-n)\big(1-\mu(\theta)\big)}_{\text{unemployed}} b,$$

by choice of w and θ, subject to the free-entry condition of firms: $\kappa = q(\theta)[z-w]$. Using this free entry condition to solve for the wage, as $w = z - \kappa/q(\theta)$, and substituting into the union objective yields a maximization problem in θ only:

$$\begin{aligned}
&\max_{\theta} \big(\underbrace{n + \mu(\theta)(1-n)}_{\text{employed}}\big)\Big(z - \frac{\kappa}{q(\theta)}\Big) + \underbrace{(1-n)\big(1-\mu(\theta)\big)}_{\text{unemployed}} b \\
&= \max_{\theta} -\frac{n\kappa}{q(\theta)} + \underbrace{\big(n + \mu(\theta)(1-n)\big)z + (1-n)\big(1-\mu(\theta)\big)b - \theta(1-n)\kappa}_{\text{planner's objective}},
\end{aligned}$$

also taking as given n. The first line expresses the tradeoff the union faces in choosing θ: increasing θ increases employment, but at the cost of the lost wage income on new and existing workers required to raise θ.

Notice that this union objective differs from the planner's objective only by the initial term, $-\frac{n\kappa}{q(\theta)}$. To understand how the two objectives relate to each other, recall that the planner is concerned about all agents in the economy, while the union only cares about workers. The union objective thus equals the planner objective less the capitalists' share of incomes: firm profits from new matches—which are zero due to free entry—and firm profits from existing matches, i.e., $(z-w)n = \frac{n\kappa}{q(\theta)}$.

The union optimum is characterized by the first-order condition $-\kappa - \kappa \frac{n}{1-n} \frac{q'(\theta)}{q(\theta)^2} + \mu'(\theta)(z-b) = 0$, which implies that the union's choice of θ *depends on* n. In our example, the union optimum is given by $\theta = \sqrt{1-n}\sqrt{(z-b)/\kappa} - 1$. Labor-market tightness is thus again an increasing function of market productivity, but now decreases in initial employment. We can see that the union implements the socially optimal level of vacancy creation if initial employment is zero. But if initial employment is positive, the union has an incentive to raise wages above the level consistent with efficient vacancy creation, in order to collect surpluses from firms with existing matches.[6]

The one-period problem captures the essence of why a monopoly union chooses a sub-optimally low level of employment, and production. How does the argument just put forth play out in the infinite-horizon model—what is, for example, the effect on steady-state unemployment? That depends crucially, as we shall see below, on the extent to which the union can commit.

[6] Note that if the union could differentiate between employed and unemployed workers in setting wages, insiders and outsiders, it would attain efficient vacancy creation here. Wages for employed workers would be raised as high as possible, to z, leaving zero surplus for employers, while new hires would get a wage consistent with efficient vacancy creation.

2.2 The efficient benchmark and a recursive planner's problem

To characterize union wage-setting when the time horizon is infinite, we again begin with the efficient benchmark. The planner now chooses a sequence of $\{\theta_t\}_{t=0}^{\infty}$ to maximize

$$\sum_{t=0}^{\infty} \beta^t \Big[\underbrace{\big(n_t + \mu(\theta_t)(1-n_t)\big)}_{\text{employed}_t} z + \underbrace{(1-n_t)(1-\mu(\theta_t))}_{\text{unemployed}_t} b - \underbrace{\theta_t(1-n_t)}_{\text{vacancies}_t} \kappa \Big]$$

$$\text{s.t. } n_{t+1} = (1-\delta) \underbrace{\big(n_t + \mu(\theta_t)(1-n_t)\big)}_{\text{employed}_t},$$

with n_0 given.

For what comes later it will be useful to formulate problems recursively. Thus, we begin by writing the planner's problem recursively, and discussing efficient vacancy-creation in that context. We then compare these to outcomes in the unionized economy.

The recursive form for the planner's problem reads

$$V(n) = \max_{\theta} \big(n + \mu(\theta)(1-n)\big)z + (1-n)\big(1-\mu(\theta)\big)b - \theta(1-n)\kappa + \beta V\big(N(n,\theta)\big), \tag{4}$$

where $N(n,\theta) \equiv (1-\delta)\big(n + \mu(\theta)(1-n)\big)$. Notice that the state variable is n, the number of matches at the beginning of the period, and that the control variable—labor-market tightness θ—determines n' according to the law of motion $N(n,\theta)$.

The first-order condition, assuming an interior solution, is

$$\underbrace{\kappa}_{\text{vacancy cost}} = \underbrace{\mu'(\theta)}_{\text{increase in matches}} \underbrace{\big(z - b + \beta(1-\delta)V'\big(n'\big)\big)}_{\text{PDV surplus}}. \tag{5}$$

It equalizes the cost of an additional vacancy to its benefits: the increase in vacancies increases hiring, with each new worker delivering the flow surplus $z - b$ today together with a continuation value reflecting future flow surpluses. The envelope condition gives the marginal value of a beginning-of-period match, as

$$\underbrace{V'(n)}_{\text{value of match}} = (1 \underbrace{-\mu(\theta) + \theta\mu'(\theta))}_{\text{decrease in matches}} \underbrace{\left(z - b + \beta(1-\delta)V'(n')\right)}_{\text{PDV surplus}}. \tag{6}$$

An additional match has the same benefit as above: the flow surplus $z - b$ today and the corresponding continuation value. An increase in beginning-of-period matches increases the planner surplus by this benefit, but there is an additional effect as well, as the increase in existing matches hampers hiring today. To see this in the expression, note that the derivative of the matching function with respect to unemployment, $m_u(\theta, 1)$, equals $\mu(\theta) - \theta\mu'(\theta)$.

Eliminating the derivative of the value function in (5), we can write an Euler equation as

$$\underbrace{\frac{\kappa}{\mu'(\theta)}}_{\text{cost per match}} = z - b + \beta \underbrace{(1-\delta)\left(1 - \mu(\theta') + \theta'\mu'(\theta')\right)}_{\text{net match creation}'} \underbrace{\frac{\kappa}{\mu'(\theta')}}_{\text{cost per match}'}, \tag{7}$$

where θ is short for the optimal choice of θ given n. This equation states the efficiency condition for the Mortensen-Pissarides model. It can be interpreted as a variational calculation: a tradeoff between cost of creating a new job today and its benefits today and tomorrow, keeping the number of matches thereafter constant.

To understand equation (7), note that the cost of creating an additional match in a given period can be broken down into two factors: the cost of creating a vacancy, κ, times the measure of vacancies required to fill one job. Since an increase in vacancies by one unit gives an increase in labor market tightness of $1/(1-n)$ units and an increase in market tightness by one unit gives $(1-n)\mu'(\theta)$ new jobs, one new vacancy creates $\mu'(\theta)$ new jobs. Hence the cost of creating one new job is $\kappa/\mu'(\theta)$. The benefits include the market production output net of home production output today, $z - b$, as well as what is saved on vacancy creation costs next period. How much is saved? First, note that the net change in matches next period is not simply $1-\delta$. Although the share $1-\delta$ of newly created jobs survive to the next period, this increase in matches also shrinks the pool of unemployed, meaning that any planned vacancy-creation next period will yield fewer matches. For each worker now out of the unemployment pool, there is a decrease in new matches given by $m_u(\theta', 1)$, where m is the matching function. Since $m_u(\theta, 1) = \mu(\theta) - \theta\mu'(\theta)$, we thus obtain a net increase in matches next period of $(1 - \delta)(1 - \mu(\theta') + \theta'\mu'(\theta'))$. Finally, each additional match next period saves $\kappa/\mu'(\theta')$ consumption units.

Looking at the Euler equation, we notice a familiar feature of the benchmark search and matching model: it does not feature the state variable n explicitly. Only market tightness today and tomorrow appear, so that a natural solution is a constant tightness independently of n. It is straightforward to show that the Bellman equation is solved by a value function V that is linear in n, and that the efficient allocation thus features a constant θ^e, independent of n.

2.3 A union with commitment

This planner problem and the union problem are closely related. To see this, note first that the union can be thought of as simply choosing a sequence of vacancy-unemployment ratios, $\{\theta_t\}_{t=0}^{\infty}$, rather than a sequence of wages. This is because the union's choice of a sequence of wages $\{w_t\}_{t=0}^{\infty}$ determines, at each instant, the present value of wages workers expect to earn over an employment spell, as $W_t = \sum_{s=0}^{\infty} \beta^s (1-\delta)^s w_{t+s}$. The sequence of these present values $\{W_t\}_{t=0}^{\infty}$ then pins down the sequence $\{\theta_t\}_{t=0}^{\infty}$ through the zero-profit conditions. Intuitively, choosing higher wages (in present value) reduces firm profits from vacancy creation, thereby reducing θ_t. Conversely, given a sequence $\{\theta_t\}_{t=0}^{\infty}$, one can back out per-period wages by first using the zero-profit condition to find the present value of wages W_t each period, and then computing wages as $w_t = W_t - \beta(1 - \delta)W_{t+1}$.

Using the zero-profit condition to eliminate the wage sequence in this way, the union objective becomes[7]

$$-\underbrace{\frac{n_0\kappa}{q(\theta_0)}}_{\text{surplus of initial matches}}+\sum_{t=0}^{\infty}\beta^t\Big[\underbrace{\big(n_t+\mu(\theta_t)(1-n_t)\big)}_{\text{employed}_t}z+\underbrace{(1-n_t)(1-\mu(\theta_t))}_{\text{unemployed}_t}b-\underbrace{\theta_t(1-n_t)}_{\text{vacancies}_t}\kappa\Big], \tag{8}$$

revealing an identical objective to that of the planner except for the first term. This term—familiar from equation (2)—summarizes the surplus accruing to capitalists. To see this, note that the capitalists share of the surplus, i.e., the present value of profits to firms, can be written as

$$\underbrace{n_0\sum_{t=0}^{\infty}\beta^t(1-\delta)^t[z-w_t]}_{\text{initial matches}}+\sum_{t=0}^{\infty}\beta^t\underbrace{[\mu(\theta_t)(1-n_t)\sum_{s=0}^{\infty}\beta^s(1-\delta)^s[z-w_{t+s}]-\theta_t(1-n_t)\kappa]}_{\text{vacancies in period } t}. \tag{9}$$

Here the first term captures the present value of profits to initial matches, and the second those to new vacancies created in periods $t=0,1,\ldots$. This objective reduces to representing initial matches only, however, as the free entry condition implies zero present value of profits to new vacancies.[8] Matches that exist at time zero, however, are due a strictly positive present value of profits, because these firms paid the vacancy cost in the past, anticipating positive profits in the future to make up for it. Moreover, using the free entry condition, this present value of profits can be written as $n\kappa/q(\theta_0)$.

The union objective (8) reflects the fact that while the planner maximizes the present value of output, the union only cares about the workers' share of that present value. In fact, the union will have an incentive to grab some of this present value from capitalists by raising wages—and this is exactly how the solutions to the two problems will differ. The union distorts vacancy creation in doing so, however, because these higher wages apply also to new vacancies.

[7]See Appendix A.

[8]Formally, can write the second term in (9) as $\sum_{t=0}^{\infty}\beta^t(1-n_t)\theta_t[q(\theta_t)\sum_{s=0}^{\infty}\beta^s(1-\delta)^s[z-w_{t+s}]-\kappa]$, which equals zero under the free entry condition (2).

Proposition 1. *When able to commit to future wages, the union attains efficient vacancy creation after the initial period. In the initial period, vacancy creation is efficient if* $n_0 = 0$ *and below the efficient level if* $n_0 > 0$.

Note that the union effectively solves the planner's problem in all but the initial period, so the solution equals the planner's solution θ^e in those periods. In the initial period the union instead solves the problem

$$\max_{\theta_0} -\frac{n_0\kappa}{q(\theta_0)} + \big(n_0 + \mu(\theta_0)(1-n_0)\big)z + (1-n_0)\big(1-\mu(\theta_0)\big)b - \theta_0(1-n_0)\kappa + \beta V\big(N(n_0,\theta_0)\big) \tag{10}$$

where $N(n,\theta) = (1-\delta)\big(n + \mu(\theta)(1-n)\big)$, n_0 is given, and V solves the planner problem (4) above.

Deriving the first-order condition for this first period is straightforward using the same methods as above. It becomes

$$\underbrace{\frac{\kappa}{\mu'(\theta_0)}}_{\text{cost of match}} - \underbrace{\frac{n_0}{1-n_0}\frac{q'(\theta_0)}{q(\theta_0)^2}\frac{\kappa}{\mu'(\theta_0)}}_{\text{tax on initial matches}} = z - b + \beta(1-\delta)\big(1-\mu(\theta^e) + \theta^e\mu'(\theta^e)\big)\underbrace{\frac{\kappa}{\mu'(\theta^e)}}_{\text{cost of match}'}. \tag{11}$$

Here we have used the fact that subsequent periods will entail the efficient level of market tightness θ^e (independent of the measure of matches). Using the efficiency condition (7), we can also write

$$\frac{1}{\mu'(\theta_0)}[1 - \frac{n_0}{1-n_0}\frac{q'(\theta_0)}{q(\theta_0)^2}] = \frac{1}{\mu'(\theta^e)}.$$

Because $q'(\theta) < 0$ and $\mu'(\theta)$ is decreasing, this equation implies (i) a lower value of θ_0 in the initial period than later on and (ii) the more initial matches, the stronger this effect. Intuitively, an increase in vacancy creation now includes the additional cost of giving up the the wage income the union could collect from firms with initial matches. Thus, in this first period market tightness now depends (negatively) on the measure of pre-existing matches. This is an important force in the present model, and a central mechanism behind unemployment dynamics when unions do not have commitment.

That the outcome in the initial period of the above problem differs from later ones reflects a time inconsistency problem in the union wage-setting problem. If the union were to re-optimize wages after the initial period, it would face a different objective and thus choose a different path of wages. While the union can thus get relatively close to the efficient outcome when it can commit, this immediate time-inconsistency begs the question: what happens if the union cannot commit to future actions? To study time-consistent union decision-making we next turn to a game-theoretic setting, which will be based on the kind of recursive formulation of the union problem we set up above.

2.4 A union without commitment

The union problem (10) suggests that if the union were to re-optimize at any date, its choice of θ (via its choice of wages) would depend on employment n. In particular, higher employment would imply a lower θ. How would outcomes be affected if the union could not commit to not re-optimizing? We study this question by focusing on (differentiable) Markov-perfect equilibria with n, the measure of matches in the beginning of the period, as a state variable. That n is a payoff- and action-relevant state variable should be clear from the analysis of the problem under commitment, where we argued that a higher n will cause a lower initial union choice for tightness.[9] In a Markov-perfect equilibrium, the union anticipates its future choices of θ to be a (decreasing) function of n, which we label $\Theta(n)$. Our task is now to characterize $\Theta(n)$.

[9] One can add states, representing histories of past behavior, but we do not consider such equilibria here.

This function $\Theta(n)$ solves a problem similar to (10), namely

$$\Theta(n) \equiv \arg\max_{\theta} -\frac{n\kappa}{q(\theta)} + \big(n + \mu(\theta)(1-n)\big)z + (1-n)(1-\mu(\theta))b - \theta(1-n)\kappa + \beta\tilde{V}\big(N(n,\theta)\big), \tag{12}$$

where the continuation value $\tilde{V}$ satisfies the recursive equation

$$\tilde{V}(n) = \big(n + \mu(\Theta(n))(1-n)\big)z + (1-n)(1-\mu(\Theta(n)))b - \Theta(n)(1-n)\kappa + \beta\tilde{V}\big(N(n,\Theta(n))\big). \tag{13}$$

Here, the union recognizes that its future actions will follow $\Theta(n)$, and this is reflected in the continuation value $\tilde{V}(n)$. Because $\Theta(n)$ will not in generally be efficient, this means that $\tilde{V}$ will not equal V, the continuation value under used in solving the commitment problem.

A *Markov-perfect equilibrium* is thus defined as a pair of functions $\Theta(n)$ and $\tilde{V}(n)$ solving (12)–(13) for all n. We will assume that these functions are differentiable and characterize equilibria based on this assumption. We discuss issues of existence and uniqueness/multiplicity of equilibria in Section 3 below.

The first-order condition for the choice of tightness reads

$$\underbrace{\kappa}_{\text{vacancy cost}} \underbrace{-\frac{n}{1-n}\frac{q'(\theta)}{q(\theta)^2}\kappa}_{\text{tax on initial matches}} = \underbrace{\mu'(\theta)}_{\text{increase in matches}} \underbrace{\big(z - b + \beta(1-\delta)\tilde{V}'(n')\big)}_{\text{PDV}} \tag{14}$$

and the equation paralleling the envelope condition in the commitment case—which now is not an envelope condition since the union does not agree with its future decisions—becomes

$$\begin{aligned}\underbrace{\tilde{V}'(n)}_{\text{value of match}} = {} & (1 \underbrace{-\mu(\theta) + \theta\mu'(\theta)}_{\text{decrease in matches}}) \underbrace{\big(z - b + \beta(1-\delta)\tilde{V}'(n')\big)}_{\text{PDV}} \\ & + \underbrace{\big(\Theta'(n)(1-n) - \theta\big)}_{\text{decrease in vacancies}} \underbrace{\Big(-\frac{n}{1-n}\frac{q'(\theta)}{q(\theta)^2}\kappa\Big)}_{\text{tax on initial matches}}.\end{aligned} \tag{15}$$

Equation (15) is derived by first, and straightforwardly, taking derivatives of equation (13), and then using equation (14) so as to arrive at a formulation that is close to the equivalent condition for the planner—equation (6). A feature of equation (15) not present in the planner's envelope condition is the presence of terms involving the derivative of Θ. This, again, is because the envelope theorem does not apply. Nevertheless, we can combine the two equations to eliminate $\tilde{V}'$: solve for it in equation (14) and insert it into equation (15) on both sides of the equation (evaluated at n and at n', respectively). Thus we obtain

$$\underbrace{\frac{\kappa}{\mu'(\theta)}[1-\frac{n}{1-n}\frac{q'(\theta)}{q(\theta)^2}]}_{\text{cost of match today}} = z - b + \beta(1-\delta)[\big(1-\mu(\theta')+\theta'\mu'(\theta')\big)\underbrace{\frac{\kappa}{\mu'(\theta')}[1-\frac{n'}{1-n'}\frac{q'(\theta')}{q(\theta')^2}]}_{\text{cost of match tomorrow}} + \underbrace{(\Theta'(n')(1-n')-\theta')}_{\text{decrease in vacancies}}\underbrace{(-\frac{n'}{1-n'}\frac{q'(\theta')}{q(\theta')^2}\kappa)}_{\text{tax on initial matches}}], \quad (16)$$

which is a *generalized Euler equation*. It is a functional equation in the unknown policy function Θ, where the derivative of Θ appears. The equation is written in a short-hand way: θ is short for $\Theta(n)$, θ' is short for $\Theta(N(n,\Theta(n)))$, and n' is short for $N(n,\Theta(n))$. Thus, the task is to find a function Θ that solves this equation for all n. In contrast to the case of the benevolent planner, or the commitment solution after period 0, n appears nontrivially in this equation and will generally matter for the tightness outcome—it is easily verified that a guess that Θ is constant will not solve this functional equation.

In terms of interpretation, this equation, like the planner's Euler equation (7), represents the tradeoff between creating matches today versus tomorrow, but now for the union instead of the planner. To understand the equation, recall that the union controls vacancy creation through its choice of wages. If the union wishes to increase vacancy creation today, it does so by reducing (the present value of) wages today. The cost of an additional match for the union differs from the cost for the planner, however. In addition to the increase in vacancy costs $\kappa/\mu'(\theta)$, the union also takes into account the loss of surpluses from existing matches that is associated with reducing wages in order to increase hiring, $\frac{\kappa}{\mu'(\theta)}\frac{n}{1-n}\frac{q'(\theta)}{q(\theta)^2}$ (in the present and in the future). The present value of an additional worker must therefore, from equation (14), be higher in in the unionized economy than what is efficient. This feature is present also in the Euler equation for the union with commitment, equation (11), though here this extra tax on matches appears in both periods, since the problem of the union today and tomorrow are symmetric, unlike in the commitment solution where tomorrow's union mechanically carries out the orders of today's plan.

But beyond this difference, here the union also takes into account its inability to commit to future wages: higher employment tomorrow will reduce vacancy creation, as the union will take advantage of its ability to collect surpluses from existing matches through higher wages. To see this, note that the measure of vacancies can be written as $\Theta(n)(1-n)$ and its derivative with respect to employment as $\Theta'(n)(1-n)-\Theta(n)$. Higher employment tomorrow reduces vacancy creation tomorrow, and that is a cost: it appears as the last, negative, term on the right-hand side of (16). It appears, in particular, since the current union regards next period's union as selecting an employment (or vacancy-creation) level that is too low compared to what it would select were it able to commit.

For the present model it is hard to establish, in general, that $\Theta(n)$ is indeed decreasing. In the one-period example of Section 2.1 we saw that Θ becomes a decreasing function of n, and in our numerically solved examples below, this feature is always present.[10] What is possible to show for the infinite-horizon case, however, is that whenever Θ is decreasing, steady-state unemployment is strictly below its efficient level:

Proposition 2. *If $\Theta(n)$ is decreasing in n, then the steady-state level of market tightness, θ, in the unionized economy (without commitment) is strictly below its efficient level.*

It follows that steady-state unemployment in the unionized economy is strictly above its efficient level.

[10] We also have not been able to find an example where Θ is not decreasing.

3 Quantitative results: comparative statics and comparative dynamics

The previous section shows that the presence of the monopoly union affects the levels of unemployment, wages, and output in the economy. But are these effects quantitatively relevant? In this section we parameterize the model in order to study this question. We will also look at an extension with stochastic shocks to productivity and ask whether, in this model, shock amplification is significantly different than in the standard model.

3.1 Wages, unemployment, and output in steady state

How does the presence of the union in the labor market affect the levels of wages, unemployment, and output? The theory tells us that the answer hinges on the union's ability to commit to future wages. If the union can commit, the unionized economy attains efficiency in steady state. If the union cannot commit, the theory leads us to expect higher wages and unemployment, and consequently lower output, in the unionized economy than what would be efficient.

Calibration We adopt the benchmark parametrization used by Shimer (2005), but use an annual frequency for our calibration, motivated by the annual contracting practices commonly observed (we look at other frequencies below). We first set the time discount rate to correspond to a 5% annual rate of return, with $\beta = 1/1.05$. We normalize labor productivity to $z = 1$ and set $b = 0.4$. We depart from Shimer's specification slightly by adopting the matching function $m(v, u) = \mu_0 vu/(v + u)$, used by, e.g., den Haan, Ramey, and Watson (2000). This form is better suited for the discrete-time setting than a Cobb-Douglas functional form because it guarantees that matching probabilities remain between zero and one. We set the separation rate δ and matching function coefficient μ_0 following Shimer, but adjusting his numbers, 0.034 and 0.45, to the annual frequency. This implies $\delta = 0.34$

and $\mu_0 \approx 1$. With these parameters, a vacancy cost of $\kappa = 0.015$ guarantees a steady-state unemployment rate of approximately 5.4%, the calibration target used by Shimer.

Numerical solution technique The planner's problem, as well as the case of a union with commitment, can be solved almost in closed form. Solving for the union's behavior when it cannot commit is more challenging, however, with several issues to bear in mind: On the one hand, there are few results available on equilibrium existence for differentiable Markov-perfect equilibria. Moreover, differentiable equilibria may not be unique. And further, non-differentiable equilibria may exist as well.[11] Clearly, one needs to proceed with caution and be prepared to use several different solution techniques. The results we present in the tables and figures below use the methods in Krusell, Kuruscu, and Smith (2002) and rely on approximating the equilibrium function Θ with polynomials of increasingly high order. However, we have also used two other, very different methods, and they deliver very similar quantitative results. We discuss these issues in more detail in Appendix B.

Results Table 1 reports the steady-state levels of the wage w, the vacancy-unemployment ratio θ, unemployment, vacancies, and output. The table compares outcomes under efficiency, or in a unionized economy where the union can commit, to outcomes in a unionized economy where the union cannot commit. In the latter case, the union's incentive to raise wages leads to a 1.5 percent increase in wages,[12] which amounts to a 38 percent reduction in firm profits per employed worker. This reduction in profits leads to a 44 percent drop in the vacancy-unemployment ratio, composed of an over 70 percent increase in unemployment, and 40 percent drop in vacancies. The reduction in employment results in a 4 percent drop in per-period output.

As a robustness check, we provide results in Table 2 for a higher value of home production, with very similar results.

[11] For examples where no differentiable equilibria exist but there exists a non-differentiable equilibrium see, e.g., Krusell, Martin, and Rios-Rull (2010); for cases with a continuum of non-differentiable equilibria along with one or more differentiable equilibrium, see Krusell and Smith (2003) or Phelps and Pollak (1968).

[12] Relative to the wage which would implement the efficient level of vacancy creation, i.e., the efficient vacancy-unemployment ratio θ.

Table 1: Effect of Union on Level of Aggregate Variables

Level	Efficient	Union	Union impact
Unemployment	0.05	0.09	+71%
Vacancies	2.23	1.34	-40%
V-U ratio	5.93	3.33	-44%
Output	0.95	0.91	-4%

Notes: The table reports steady-state values.

Table 2: Effect of Union on Level of Aggregate Variables, higher b

Level	Efficient	Union	Union impact
Unemployment	0.05	0.09	+71%
Vacancies	2.23	1.34	-40%
V-U ratio	5.93	3.33	-44%
Output	0.95	0.91	-4%

Notes: The table reports steady-state values. Here $b = 0.5$ and, to maintain efficient unemployment at the original level, $\kappa = 0.0125$.

The role of the commitment horizon The recursive formulation assumes the union can commit to current period wages, but not beyond. This suggests that limited commitment becomes less of an issue as the period length increases. By adjusting the discount rate β, the separation rate δ, and the matching function coefficient μ_0, one can examine different period lengths, and hence different degrees of commitment. Table 3 reports the results. On the one hand, we can see that moving from the annual horizon to a shorter, biannual horizon, exacerbates the negative effects of limited commitment significantly: employment and output fall by as much as four percentage points, with unemployment increasing by about the same, from 9.2 to 12.9 percent. On the other hand, moving to the infinite horizon limit would increase employment and output by four percentage points, with unemployment falling from 9.2 to 5.4 percent.[13]

[13]It is intuitive that if the period length approaches zero, the lack of union commitment will lead to 100% unemployment. It is not exactly true that if the period length approaches infinity, the commitment solution obtains, however, as the commitment solution generally involves a time-varying policy, while here policy is fixed within a period.

Table 3: Role of Commitment Horizon

Level	6 months	12 months	Efficient
Unemployment	12.9%	9.2%	5.4%
Vacancies	0.37	1.34	2.23
V-U ratio	1.28	3.33	5.93
Output	0.87	0.91	0.95

Notes: The table reports steady-state values. First columns correspond to the baseline parameters, the third column has $\delta = 0.187$ and adjusts $\kappa = 0.047$ (to maintain unemployment fixed).

3.2 Welfare comparisons

The union maximizes the welfare of all workers in the economy, thus internalizing the general equilibrium effects of its wage demands. Nevertheless, the unionized economy generally departs from efficiency. Even in the simple one-period example, the unionized economy does not attain efficient vacancy creation because it cannot differentiate between new and existing workers when setting wages. The union seeks to redistribute from firms to workers by raising wages in existing matches, but these higher wages also distort vacancy creation. In the dynamic model, another source of inefficiency appears whenever the union lacks commitment: there is a long-run loss from suboptimal job creation.

How large are the welfare losses resulting from the labor union presence? To shed light on this question, we study the transitional dynamics of an economy with a labor union which cannot commit to future wages. Starting from steady state, we ask: 1) what would happen if the union gained commitment, and 2) how do these outcomes differ from the efficient response? Figure 1 illustrates the responses of the vacancy-unemployment ratio and employment in the two cases. As is clear from the pictures, the dynamics of market tightness θ reflect our analytical results above: with sudden commitment the union would maintain a low θ in the initial period—it is slightly above that of the no-commitment, steady-state starting point—but then have a fully efficient θ beginning in the consecutive period.[14] Consequently, the

[14] From the Euler equations for the commitment and no-commitment cases, a comparison reveals that it is not clear that the dynamics will be monotone. It turns out to be the case in the graph but it could have turned out instead that initial tightness would decrease slightly the first month before jumping up to the steady-state level.

dynamics are rather fast, in the sense that in a few periods the efficient and commitment-union economies both have unemployment very close to the efficient steady-state rate. The figure contrasts outcomes in the annual calibration (on the right) with those in a monthly calibration (on the left). The difference between the efficient response and the commitment-union response is naturally greater in the annual calibration, where the initial period is longer.

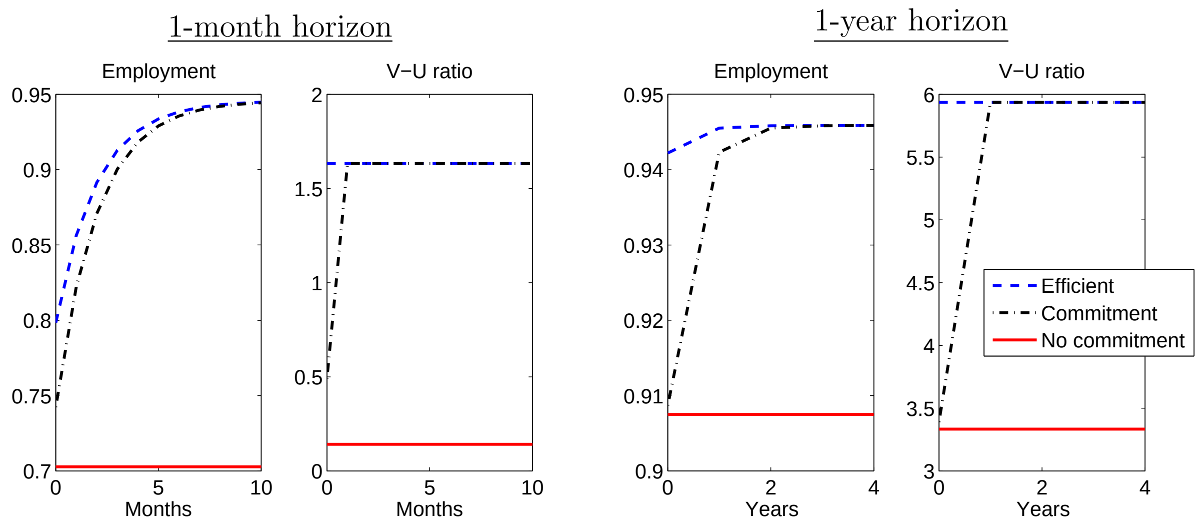


Figure 1: Adjustment Dynamics when Union Gains Commitment versus Efficient Response
Notes: The figure plots adjustment dynamics starting from the steady state where the union cannot commit. The figure shows: how the planner would respond, how the union would respond if it gained commitment, and how the union would respond if it did not.

How large are the effects on welfare? The present value of output at time zero on these three transition paths are as follows: in the annual calibration, the planner's response yields present value 2.54, attaining commitment gives present value 2.51, while remaining at no commitment gives 2.45. In terms of the per-period increase in output from attaining efficiency, these figures translate to 4.07%, while the increase from attaining commitment is 2.61%. For comparison, in the monthly calibration these numbers are 32.7% and 31.9%, respectively. Attaining commitment leads to non-trivial welfare gains in both cases, but the gains are much larger in the monthly calibration because the no-commitment outcome is substantially worse in that case. The difference between attaining commitment and attaining efficiency is larger in the annual calibration, however, as the initial adjustment period is longer in that case.

3.3 Aggregate shocks

An important motivation for macroeconomists to study labor unions is the idea that unions create rigidity in wages, which may help reconcile the large cyclical variation in employment observed with macroeconomic theory (see e.g. Blanchard and Fischer (1989)). What does our theory of unions imply about the responses of wages, vacancy creation, and unemployment to shocks? The dynamics of the model under efficiency are well known, but how do these dynamics change when the labor market has a monopoly union that cannot commit to future wages? To answer this question, one can study deterministic transitions to steady state. However, it appears more empirically interesting to compare economies that actually feature recurring fluctuations. The standard way of conducting this kind of analysis is that pioneered in Pissarides (1985) and revisited in Shimer (2005). A question of interest here is whether the amplification of shocks is stronger in the monopoly-union economy than in the basic model. As we shall see, it is.

One could think of various kinds of shocks perturbing the economy over time. For purposes of illustration, the most obvious shock to consider is one to productivity z. It is straightforward to extend the setup above to allow z to follow a Markov process. A union that cannot commit to future wage setting in this environment will, as in the analysis above, play a dynamic game with its future counterparts, though the game here will be stochastic. As before, it is natural to focus on Markov-perfect equilibria. Thus, $\Theta(n, z)$ now depends on productivity, as

$$\begin{aligned}\Theta(n,z) \equiv \arg\max_{\theta} -\frac{n\kappa}{q(\theta)} + \big(n + \mu(\theta)(1-n)\big)z + (1-n)(1-\mu(\theta))b - \theta(1-n)\kappa \\ +\beta E_z \tilde{V}\big(N(n,\theta), z'\big),\end{aligned}$$

where the continuation value $\tilde{V}$ satisfies the recursive equation

$$\tilde{V}(n,z) = \big(n + \mu(\Theta(n,z))(1-n)\big)z + (1-n)(1-\mu(\Theta(n,z)))b - \Theta(n,z)(1-n)\kappa \\ +\beta E_z \tilde{V}\big(N(n,\Theta(n,z)),z'\big).$$

It is straightforward, along the lines above, to derive the generalized Euler equation for this case as well. It reads

$$\frac{\kappa}{\mu'(\theta)}\big(1 - \frac{n}{1-n}\frac{q'(\theta)}{q(\theta)^2}\big) = z - b + \beta(1-\delta)E_z\big[\big(1-\mu(\theta')+\theta'\mu'(\theta')\big)\frac{\kappa}{\mu'(\theta')}(1-\frac{n'}{1-n'}\frac{q'(\theta')}{q(\theta')^2}) \\ + \big(\Theta_n(n',z')(1-n')-\theta'\big)(-\frac{n'}{1-n'}\frac{q'(\theta')}{q(\theta')^2}\kappa)\big],$$

thus differing only in that there is an expectations operator in front of all the marginal future payoffs.

The model is calibrated as the deterministic economy, and our numerical method easily extended to cover the shock case.[15]

We first look at impulse responses. Figure 2 plots the impulse responses of wages, θ, unemployment and output, comparing the unionized economy (solid line), to the efficient outcome (dashed line). Note that the wage response in the efficient outcome refers to the wage which would implement the efficient allocation, i.e., which gives firms exactly the amount of surplus in matching to induce them to create the efficient measure of vacancies. Again, the right panel displays the annual calibration, while the left displays a monthly calibration meant to highlight the effects.

As can be observed, in the short run, the union acts so as to introduce "real wage stickiness" into the dynamics. This occurs because a positive productivity shock leaves the level of employment too low compared to what the higher productivity would imply. Although employment soon rises due to increased vacancy creation, the low initial employment level

[15]See the appendix for details. In brief, the numerical solution is recursive: one can first solve for deterministic dynamics in the state n and, as a function of that, for responses to z.

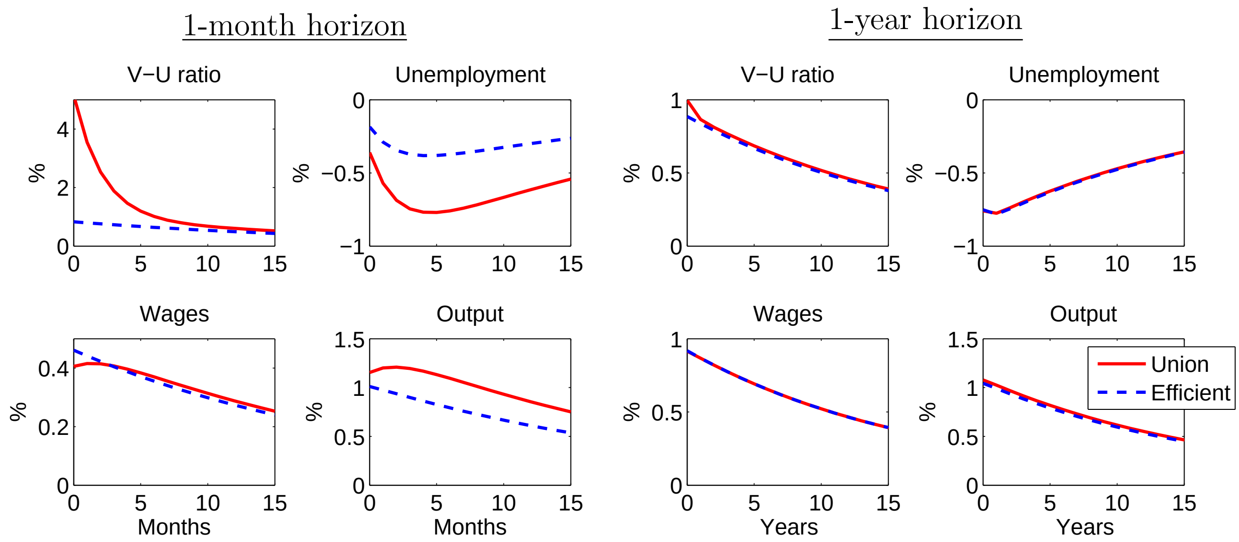


Figure 2: Impulse responses: efficient versus union

Notes: The figure plots impulse responses to a one percent positive productivity shock, when labor productivity follows and AR(1) process.

curbs the union's distortionary motive to raise wages. Wages thus appear sticky in the very short run, while vacancy creation consequently overshoots. We see, in line with insights in the literature following Shimer (2005), that the response of tightness to a wage change can be substantial. As a result, the percentage responses to a shock are generally to amplify responses (apart from wages, where the effects are roughly zero).[16] Generally speaking, the effect of the monopoly union on dynamics is stronger when employment is higher, because a large number of pre-existing matches gives the union incentives that are different from those in the efficient allocation.

Table 4 below reports simulated moments. As we can see, the dynamics are different from the efficient allocation and in a direction that helps us understand the data: productivity shocks are propagated throughout the economy with stronger amplification for market tightness and unemployment.

[16]The absolute responses vary, however, as there is amplification in unemployment and output but a certain dampening in θ and v.

Table 4: Effect of Union on Volatility of Aggregate Variables

	1-month horizon			1-year horizon		
Volatility	Efficient	Union	Union impact	Efficient	Union	Union impact
Unemployment	0.45%	0.91%	+100%	0.79%	0.80%	+1%
Vacancies	0.67%	3.15%	+370%	0.85%	0.93%	+9%
V-U ratio	0.87%	3.30%	+278%	0.88%	0.97%	+10%
Output	1.02%	1.37%	+33%	1.046%	1.082%	+3.5%

Notes: The table reports standard deviations of model variables relative to the standard deviation of labor productivity, based on simulated data from the model, logged and filtered. Productivity process is an AR(1), with the aggregated, logged and filtered series having standard deviation 1.62% and persistence 0.47 with an annual horizon, and 1.30% and 0.76 with a monthly horizon.

4 Extensions

Two extensions of the setting above are particularly relevant for understanding labor-market institutions: one where some workers are not unionized, and one where employers also act as an entity. We study partial unionization in Section 4.1 and collective bargaining in Section 4.2.

4.1 Partial unionization

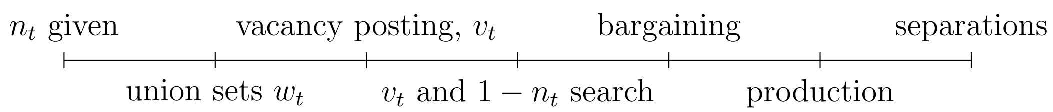
In most countries, workers are free to decide on becoming union members. The analysis above does not allow for such a choice. One difficulty in incorporating endogenous membership into the model has to do with the formulation of the union's objective function. How does membership evolve over time, and how does the endogeneity of the unionization rate affect the incentives of the union when setting wages? The evolution of the unionization rate over time is a focus of some recent work, e.g., Greenwood (2010). Here, we stay short of a full analysis but nevertheless examine how key labor-market variables depend on the unionization rate. This analysis offers some preliminary insights into the welfare consequences of policies such as forbidding unions or requiring universal coverage.

In the analysis below, we use α to denote the fraction of workers who belong to the union.

We treat α as exogenous, and assume that a worker's membership status is constant over time. We assume that the union's objective is to maximize the utility of its members. We also confine attention to steady states. In general, union workers may or may not earn higher wages than non-union workers. A particularly interesting steady state is a case which would make workers indifferent between being unionized and not, because this steady state can be interpreted as allowing workers to choose whether or not to become union members. As we will show, such a steady state exists for some parameter values.

Of course, we need to make clear how wages are determined for non-union workers. It is most natural here to simply adopt the standard assumption in the literature, i.e., one of decentralized Nash bargaining. We let the worker's bargaining share be denoted γ. We also need to make an assumption about whether the labor market is segmented by worker type—union vs. non-union—since firms in general are not indifferent about whom to meet. Our assumption is that the worker's union status is not observable ex ante so that matching is undirected. Moreover, we assume firms cannot discriminate based on union status later on either, with an identical separation probability for union and non-union workers.

Suppose, then, that only a share α of workers are unionized, while the rest bargain their wages bilaterally with firms. Both workers search in the same labor market, and firms learn the union status of workers only upon matching. At that time, bilateral bargaining occurs (also for previously matched firm-worker pairs).



In this labor market, the union recognizes the presence of the non-union workers when deciding on wage demands.

4.1.1 Analytical characterization

Beginning with a one-period example to gain intuition, we have that the union maximizes utility per member:

$$\underbrace{\left(n+\mu(\theta)(1-n)\right)}_{\text{employed}} w+\underbrace{(1-n)\big(1-\mu(\theta)\big)}_{\text{unemployed}} b,$$

by choice of w and θ, subject to the zero-profit condition: $\kappa = q(\theta)[\alpha(z-w)+(1-\alpha)(1-\gamma)S]$. Here $S \equiv z-b$ is the total surplus from a match between a firm and a non-union worker.

To see how the analysis compares to that with full unionization, we again use the zero-profit condition to substitute out the wage, as $w = z - \frac{\kappa}{\alpha q(\theta)} + \frac{1-\alpha}{\alpha}(1-\gamma)S$, in the union objective. This leads to a maximization problem in θ only:

$$\max_{\theta} \underbrace{\left(n+\mu(\theta)(1-n)\right)}_{\text{employed}}\Big(z-\frac{\kappa}{\alpha q(\theta)}+\frac{1-\alpha}{\alpha}(1-\gamma)S\Big)+\underbrace{(1-n)\big(1-\mu(\theta)\big)}_{\text{unemployed}} b,$$

also taking as given n. As before, increasing θ increases employment, but at the cost of the lost wage income on new and existing workers required to raise θ. The expression differs from the one before for two reasons. First, the union wage now has a more limited impact on vacancy-creation, because of the non-union workers among the pool of unemployed. Raising θ through the union wage thus requires giving up more wage income: $\frac{\kappa}{\alpha q(\theta)}$ is greater with $\alpha < 1$. This works to reduce θ and raise the union wage, compared to the fully unionized case. Second, the tradeoff between the union wage and θ also depends on the firms' surplus from matching with non-union workers, $(1-\gamma)S = (1-\gamma)(z-b)$, in proportion with their prevalence among the unemployed, $(1-\alpha)/\alpha$. If the non-union surplus is large (relative to the union surplus), the union can target a higher θ without giving up as much in wages, which works to raise both θ and the union wage. In the next section, we illustrate how these effects manifest themselves in labor market outcomes.

In a fully dynamic setting, non-union workers and firms operate according to the usual

Bellman equations:

$$
\begin{aligned}
U_t &= \mu(\theta_t)E_t + (1-\mu(\theta_t))(b+\beta U_{t+1}),\\
E_t &= w_t^n + \beta\delta U_{t+1} + \beta(1-\delta)E_{t+1},\\
J_t &= z - w_t^n + \beta(1-\delta)J_{t+1},
\end{aligned}
$$

where U_t is the value of an unemployed worker, E_t the value of an employed worker, J_t the value of a filled job, and w_t^n the wage of a non-union worker. Based on these equations, the non-union worker-firm match surplus, defined as $S_t = E_t + J_t - b - \beta U_{t+1}$, satisfies

$$S_t = z - b + \beta(1-\delta)(1-\mu(\theta_{t+1})\gamma)S_{t+1},$$

where bilateral wage bargains imply $J_t = (1-\gamma)S_t$, and $E_t - b - \beta U_{t+1} = \gamma S_t$.

The zero-profit condition reads

$$\kappa = q(\theta_t)[\alpha(\sum_{s=0}^{\infty}\beta^s(1-\delta)^s z - W_t) + (1-\alpha)(1-\gamma)S_t].$$

Clearly, firms realize that union workers require a present value of wages of W_t, while non-union workers yield the firm a net profit of $(1-\gamma)S_t$.

Using the free-entry condition to substitute out wages in the union objective yields

$$
\begin{aligned}
&\sum_{t=0}^{\infty}\beta^t[(n_t + \mu(\theta_t)(1-n_t))z + (1-\mu(\theta_t))(1-n_t)b - \theta_t(1-n_t)\frac{\kappa}{\alpha} + \frac{1-\alpha}{\alpha}(1-\gamma)\mu(\theta_t)(1-n_t)S_t]\\
&-\frac{n_0\kappa}{\alpha q(\theta_0)} + \frac{1-\alpha}{\alpha}(1-\gamma)n_0 S_0.
\end{aligned}
$$

The objective here is the dynamic extension of the objective in the one-period example above, with terms slightly reordered. Also, S_t is of course not exogenous here (it was equal to $z-b$ in the one-period model) because it depends on future surpluses as well.

We consider the case of no commitment again and extend the Markov-perfect equilibrium definition to cover a general value of α. The equilibrium will have the functions $\Theta(n)$, for market tightness, $\tilde{V}(n)$, for the indirect utility of union members, and $S(n)$, the total surplus of the match between a firm and a non-union worker. Note that no new state variable is needed. The functions satisfy the following functional equations:

$$S(n) = z - b + \beta(1-\delta)(1-\mu(\Theta(n))\gamma)S(N(n,\Theta(n))),$$

$$\begin{aligned}\tilde{V}(n) = {} & (n+\mu(\Theta(n))(1-n))z + (1-\mu(\Theta(n)))(1-n)b - \Theta(n)(1-n)\frac{\kappa}{\alpha} \\ & + \frac{1-\alpha}{\alpha}(1-\gamma)\mu(\Theta(n))(1-n)S(N(n,\Theta(n))) + \beta\tilde{V}(N(n,\Theta(n))),\end{aligned}$$

and

$$\begin{aligned}\Theta(n) = \arg\max_{\theta} & (n+\mu(\theta)(1-n))z + (1-\mu(\theta))(1-n)b - \theta(1-n)\frac{\kappa}{\alpha} - \frac{n\kappa}{\alpha q(\theta)} \\ & + \frac{1-\alpha}{\alpha}(1-\gamma)(n+\mu(\theta)(1-n))S(N(n,\theta)) + \beta\tilde{V}(N(n,\theta)).\end{aligned}$$

It is straightforward to derive the functional first-order condition for the union here, but it is more complex since it contains both $S(N(n,\theta))$ and $S'(N(n,\theta))$, which cannot be eliminated with simple substitution. We therefore proceed directly to the quantitative analysis.

4.1.2 Quantitative results

We calibrate as in the benchmark case and vary α and γ to illustrate the workings of the model. The numerical analysis uses the same methods as above, with the mere difference that there is an additional unknown function S.[17]

Figure 3 plots the vacancy-unemployment ratio, wages, unemployment, and output as a

[17] For details, see Appendix (B).

function of the unionization rate α when worker bargaining power γ has a relatively low value.[18] In this case, firms pay non-union workers lower wages than union workers, making non-union workers more profitable to firms. The figure contrasts outcomes with the model calibrated to a monthly, versus annual, frequency. The annual horizon, on the right, illustrates the first mechanism discussed: higher unionization increases employment and output, as the union internalizes the effects of its wage demands on the labor market. In this case greater unionization brings the economy closer to efficiency.

With a monthly horizon, as depicted on the left, the union's commitment problem is more severe (reflected in lower vacancy creation than on the right). Here the second mechanism discussed becomes dominant for vacancy creation: the presence of non-union workers in the pool of unemployed helps mitigate the adverse effects of the commitment problem. As a result, greater unionization reduces employment and output, taking the economy farther away from efficiency.

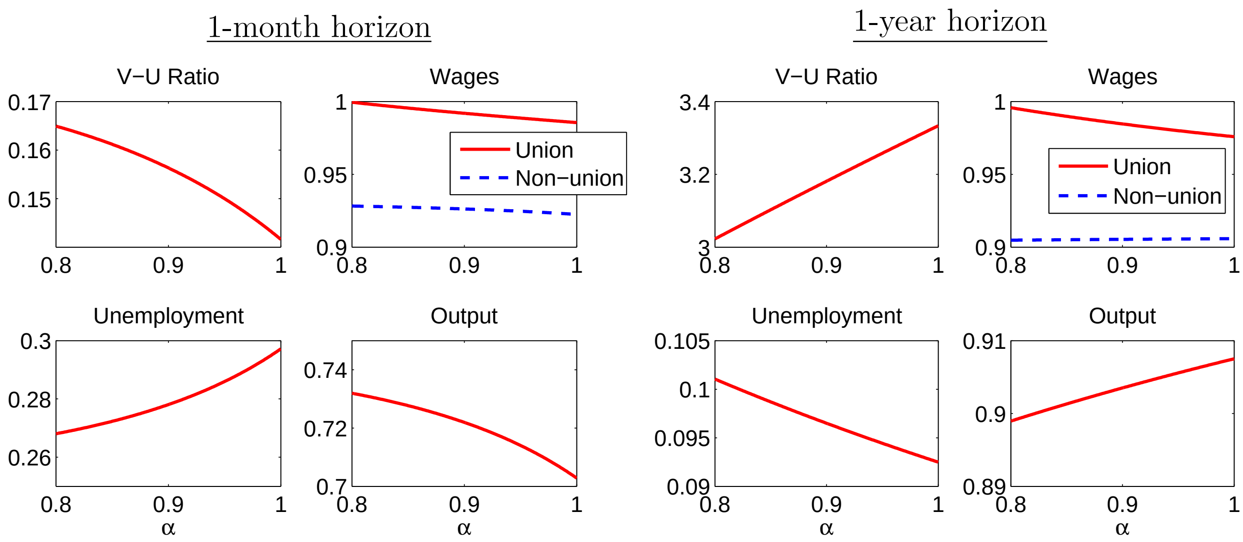


Figure 3: Labor markets when non-members are poor bargainers

Notes: The figure plots outcomes as a function of the unionization rate α for $\gamma = 0.7$.

Figure 4 turns to the case where workers are good bargainers on their own, showing that a higher unionization rate now globally encourages vacancy-creation, leading to lower unem-

[18]The specific value is 0.7; qualitatively, the graphs do not change if γ is lowered further.

ployment and greater output.

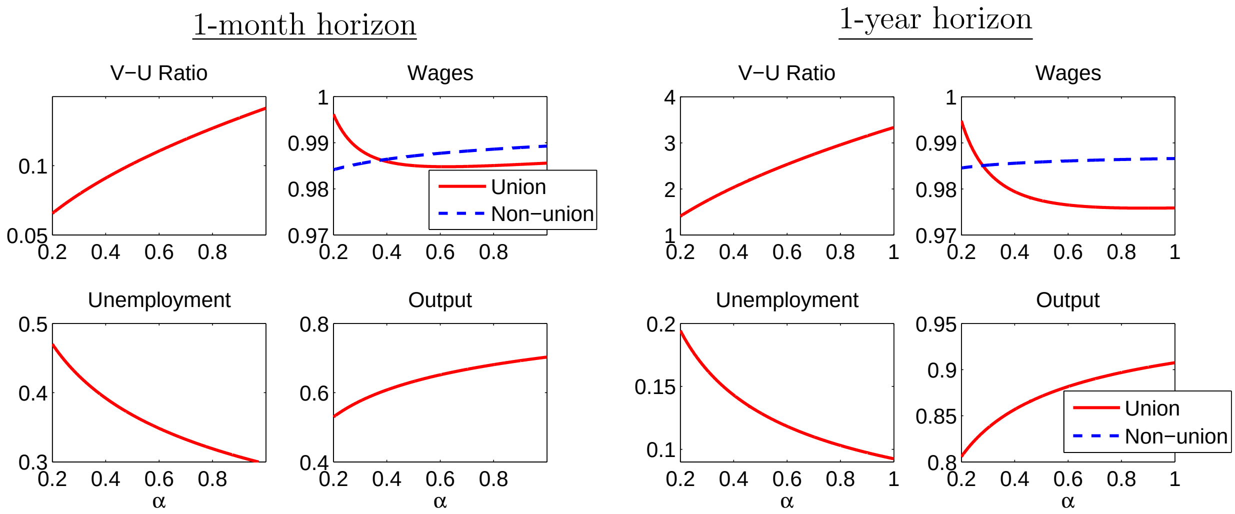


Figure 4: Labor markets when non-members are good bargainers
Notes: The figure plots outcomes as a function of the unionization rate α for $\gamma = 0.95$.

Interestingly, the figure shows that there is a level of unionization such that workers earn the same wages whether unionized or not. That is, if given the choice between becoming a union member or not, they would be indifferent. This steady state can be interpreted as the equilibrium outcome when workers can choose, at time zero, whether or not to be unionized. For these parameter values, the steady-state union wage can be non-monotonic in α. For very low unionization rates the union wage is high, and falling in α (following the first mechanism above), but eventually starts to rise again. This last part can be understood by noting that, when non-union workers earn high wages (relative to union workers), the union may find it optimal to moderate its wage claims to prevent employment from falling. As the unionization rate rises, these non-union workers become less important for vacancy creation, however, allowing the union to raise the union wage.

The previous example demonstrates that *requiring* all workers to be unionized—or covered by the union wage—can be welfare improving. Interpreting the intermediate value of α where wages are the same for union and non-union workers as an equilibrium where workers can choose membership status: for that value of α, the outcome with forced union membership would be better (and outlawing unions worse), in a steady-state sense. For the economy depicted first, the situation is of course the reverse. There, workers would all choose to become unionized, leading to $\alpha = 1$, but outlawing unions would be a good idea.

4.2 Collective bargaining

We can generalize the monopoly union framework to collective bargaining between a labor union and an employers' association, using a "right-to-manage" approach. Right-to-manage refers to firms having the right to decide on hiring independently, taking as given wages that are centrally bargained between the labor union and the employers' association. In the Mortensen-Pissarides framework this translates to hiring being determined by the usual zero-profit condition, given centrally bargained wages. Proceeding directly to the fully dynamic model, we adopt the same union objective in equation (3), and assume the employers' association maximizes the present value of profits accruing to firms in equation (9). We look at both (joint) commitment and lack thereof.

4.2.1 Commitment

We now denote the bargaining power of the labor union vis a vis the employers' association by γ. With commitment to future wages, the collective bargaining problem solves the problem

$$\max_{\{w_t,\theta_t\}_{t=0}^{\infty}} \{\sum_{t=0}^{\infty} \beta^t \left[\left(n_t + \mu(\theta_t)(1-n_t)\right)w_t + (1-n_t)(1-\mu(\theta_t))b\right]\}^{\gamma}\{n_0 \sum_{t=0}^{\infty} \beta^t (1-\delta)^t [z - w_t]\}^{1-\gamma}$$

subject to the law of motion (1) and the zero-profit condition (2). Note that, as before, the zero-profit condition implies that the employers' association objective reduces to representing initial matches only.[19]

[19] One could go into more detail in specifying alternative threat points in this bargaining problem, but we refrain to simply outlining the broader approach here.

To simplify, this bargaining problem can then be rewritten as a choice of a sequence of θ_t's. Using the zero-profit condition, we arrive at

$$\max_{\{\theta_t\}_{t=0}^{\infty}} \{-\frac{n_0\kappa}{q(\theta_0)} + \sum_{t=0}^{\infty} \beta^t \big[\big(n_t + \mu(\theta_t)(1-n_t)\big)z + (1-n_t)(1-\mu(\theta_t))b - \theta_t(1-n_t)\kappa\big]\}^{\gamma}\{\frac{n_0\kappa}{q(\theta_0)}\}^{1-\gamma}$$

subject to the law of motion (1).

For thinking about how the solution differs from the monopoly union case, it is useful to note that future values of θ_t only enter the union objective, not the employers' association objective. Given this, one could equally well follow the earlier approach of reformulating the objective as

$$\{-\frac{n_0\kappa}{q(\theta_0)} + \big(n_0 + \mu(\theta_0)(1-n_0)\big)z + (1-n_0)\big(1-\mu(\theta_0)\big)b - \theta_0(1-n_0)\kappa + \beta V(\,\cdot\,)\}^{\gamma}\{\frac{n_0\kappa}{q(\theta_0)}\}^{1-\gamma}$$

where $V(n)$ solves the recursive form of the planner's problem in equation (4). The solution to this planner's problem has θ constant at the efficient level, with $V(n)$ linear and increasing in n. The bargaining problem gives a different θ_0 in the initial period, however, depending on the bargaining power of the union vis-à-vis the employers' association. The employers' association moderates union wage demands, which translates into increased hiring. In fact, one can show that as union power γ declines, θ_0 increases from the monopoly union level.

Proposition 3. *When able to commit to future wages, collective bargaining attains efficient vacancy creation after the initial period. In the initial period, vacancy creation is efficient if* $n_0 = 0$*. If* $n_0 > 0$*, vacancy creation increases as union bargaining power* γ *falls, and is generically inefficient.*

4.2.2 Without commitment

We can adapt the right-to-manage formulation to the case of no commitment to future wages as follows. As before, we have an accounting equation for the continuation value

$$\tilde{V}(n) = \big(n + \mu(\Theta(n))(1-n)\big)z + (1-n)(1-\mu(\Theta(n)))b - \Theta(n)(1-n)\kappa + \beta\tilde{V}\big(N(n,\Theta(n))\big),$$

where

$$\begin{aligned}\Theta(n) := \arg\max_{\theta}\{-\frac{n\kappa}{q(\theta)} &+ \big(n + \mu(\theta)(1-n)\big)z + (1-n)(1-\mu(\theta))b \\ &- \theta(1-n)\kappa + \beta\tilde{V}\big((1-\delta)(n+\mu(\theta)(1-n))\big)\}^{\gamma}\{\frac{n\kappa}{q(\theta)}\}^{1-\gamma}.\end{aligned}$$

This differs from the monopoly union case only in that the choice of $\Theta(n)$ is now determined based on the bargaining problem instead of maximizing the union objective alone.

We proceed immediately to a numerical illustration, computed as in Section 4.1. Figure 5 plots the outcomes for key labor-market variables as a function of the union bargaining power γ, over a range where steady-state unemployment takes on values both above and below the efficient level. As the the figure shows, the stronger is union bargaining power, the higher union wages, leading to higher unemployment and lower output. Moreover, the collective bargaining outcome is efficient for an intermediate value of γ.

5 Conclusions

This paper studies the impact of labor unions on the aggregate economy, when labor markets are modeled as frictional. In particular, we study the forward-looking decision problem of a benevolent, centralized labor union setting wages over time. Our results highlight the dynamic nature of optimal wage policy in this context, and the role of commitment in determining outcomes. If the union can commit, then it attains efficient vacancy creation in the

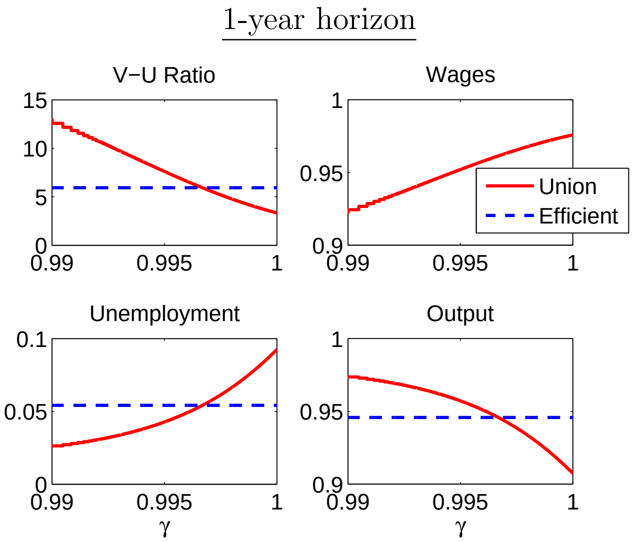


Figure 5: Labor market with collective bargaining

Notes: The figure plots outcomes as a function of the labor union bargaining power γ.

long run, but has an incentive to raise wages in the short run to collect rents from firms with existing workers, reducing vacancy creation. This wage policy is clearly time-inconsistent. If the union cannot commit, wages and unemployment are (quantitatively significantly) above the efficient level also in the long run. Moreover, in this case the union presence also leads to short-run stickiness in wages, amplifying cyclical fluctuations in vacancy-creation and unemployment.

Our analysis assumes the union to treat all workers identically, both in terms of the wage assigned and how workers enter the union objective. One could also consider a union which sets the wages of insiders (currently employed) differently from outsiders (currently unemployed). Insiders would then be paid their productivity z, leaving firms with zero present value of profit for the future. In a one period model, the union would nevertheless attain efficient hiring, as the wages of new workers could be set independently from the wages of existing workers. In a multi-period model, outcomes would depend on the ability to commit: If the union could commit, it would set wages equal to productivity for the initially matched, but implement a wage leading to efficient hiring for all new workers. Employment would then evolve efficiently from the initial period on. If the union could not commit, it would still set wages equal to productivity for the initially matched, but do the same also for all new matches after the initial period of the match. This would dampen vacancy-creation significantly, as firms would only make positive profits in the initial period of the match (and these profits would be bounded from above by $z - b$). Thus, allowing different wages for insiders and outsiders only underscores the role of commitment for outcomes.

Alternatively, one could also consider a union which places differing weights on insiders and outsiders in its objective function, but is required to treat the two groups identically in setting wages. An extreme example would be a union which, in each period, only cares about the currently employed. In this case there would be no vacancy creation at all, as the union would always set the wage equal to productivity, and unemployment would converge to 100 percent. Increasing the weight on the unemployed in the union objective, outcomes would likely approach our egalitarian benchmark specification.

Clearly the union objective, as well as how the union treats different workers, are important determinants of outcomes in unionized labor markets. Our modeling approach reflects a view that "fairness" is a concern for unions. It is conceivable that such a norm has emerged as a means for the union to overcome some of the additional commitment problems described above. Albeit extremely interesting, these issues are challenging to analyze and best left for future work.

References

Acemoglu, D., P. Aghion, and G. Violante (2001): "Deunionization, Technical Change and Inequality," *Carnegie Rochester Conference Series on Public Policy*, 55(1), 229–264.

Acikgoz, O., and B. Kaymak (2009): "The Rising Skill Premium and Deunionization in the United States," Unpublished Manuscript, Universite de Montreal.

Alvarez, F., and R. Shimer (2009): "Unions and Unemployment," Unpublished Manuscript, University of Chicago.

Blanchard, O. J., and S. Fischer (1989): *Lectures on Macroeconomics*. The MIT Press.

Boeri, T., and M. C. Burda (2009): "Preferences for Collective versus Individualised Wage Setting," *Economic Journal*, 119, 1440–1463.

Booth, A., and F. Schiantarelli (1987): "The Employment Effects of a Shorter Working Week," *Economica*, 54, 237–248.

Burdett, K., and R. Wright (1993): *Search, Matching and Unions*vol. Panel Data and Labor Market Dynamics of *Contributions to Economic Analysis*, pp. 411–426. North-Holland.

Calmfors, L., and J. Driffill (1988): "Bargaining structure, corporatism and macroeconomic performance," *Economic Policy*, 88, 13–61.

Card, D. (1986): "Efficient Contracts with Costly Adjustment: Short-Run Employment Determination for Airline Mechanics," *American Economic Review*, 86(5), 1045–71.

Carroll, C. D. (2006): "The Method of Endogenous Gridpoints for Solving Dynamic Stochastic Optimization Problems," *Economics Letters*, pp. 312–20.

Delacroix, A. (2006): "A multisectorial matching model of unions," *Journal of Monetary Economics*, 53, 573–596.

den Haan, W. J., G. Ramey, and J. Watson (2000): "Job Destruction and Propagation of Shocks," *American Economic Review*, 90(3), 482– 498.

Ebell, M., and C. Haefke (2006): "Product Market Regulation and Endogenous Union Formation," Unpublished Manuscript.

Garibaldi, P., and G. L. Violante (2005): "Employment Effects of Severance Payments with Wage Rigidities," *Economic Journal*, 115, 799–832.

GREENWOOD, J. (2010): “The Up and Down of Unions,” Unpublished Manuscript, University of Pennsylvania.

HUIZINGA, F., AND F. SCHIANTARELLI (1992): “Dynamics and Asymmetric Adjustment in Insider-Outsider Models,” *Economic Journal*, 102, 1451–66.

KENNAN, J. (1988): “Equilibrium Interpretations of Employment and Real Wage Rigidity,” *NBER Macroeconomics Annual*, 3, 157–216.

KRUSELL, P., B. KURUSCU, AND A. SMITH (2002): “Equilibrium Welfare and Government Policy with Quasi-geometric Discounting,” *Journal of Economic Theory*, 105, 42–72.

KRUSELL, P., F. MARTIN, AND V. RIOS-RULL (2010): “Time Consistent Debt,” Unpublished Manuscript.

KRUSELL, P., AND A. SMITH (2003): “Consumption–Savings Decisions with Quasi–Geometric Discounting,” *Econometrica*, 71, 365–375.

LINDBECK, A., AND D. J. SNOWER (1986): “Wage Setting, Unemployment, and Insider-Outsider Relations,” *American Economic Review*, 76, 235–239.

LOCKWOOD, B., AND A. MANNING (1989): “Wage-Employment Bargaining with Employment Adjustment Costs,” *Economic Journal*, 99, 1143–58.

MODESTO, L., AND J. P. THOMAS (2001): “An analysis of labor adjustment costs in unionized economies,” *Labour Economics*, 8, 475–501.

MORTENSEN, D. (1989): “On the Persistence and Indeterminacy of Unemployment in Serch Equilibrium,” *Scandinavian Journal of Economics*, 91, 347–370.

MORTENSEN, D., AND C. PISSARIDES (1999): *New developments in models of search in the labor market*chap. 39, pp. 2567–2627, Handbook of Labor Economics. Elsevier.

PHELPS, E., AND R. POLLAK (1968): “On Second-Best National Saving and Game-Equilibrium Growth,” *Review of Economic Studies*, 35, 185–199.

Pissarides, C. (1985): "Short-Run Equilibrium Dynamics of Unemployment, Vacancies, and Real Wages," *American Economic Review*, 75(4), 676–690.

——— (1986): "Trade Unions and the Efficiency of the Natural Rate of Unemployment," *Journal of Labor Economics*, 4(4), 582–595.

Shimer, R. (2005): "The Cyclical Behavior of Equilibrium Unemployment and Vacancies," *American Economic Review*, pp. 25–49.

Taschereau-Dumouchel, M. (2011): "The Union Threat," Unpublished Manuscript, Princeton University.

A Proofs

Proof of relationship between union and planner objectives For the benchmark model, we need to show that

$$\sum_{t=0}^{\infty} \beta^t \big(n_t + h_t\big) w_t = \sum_{t=0}^{\infty} \beta^t \big[\big(n_t + h_t\big) z - \theta_t (1 - n_t)\kappa\big] - \frac{n_0 \kappa}{q(\theta_0)}, \tag{17}$$

where h_t stands for newly hired workers, i.e., $h_t = \mu(\theta_t)(1 - n_t)$.

First, note that the law of motion for employment implies that $n_t = (1-\delta)^t n_0 + \sum_{k=0}^{t-1}(1-\delta)^{t-k} h_k$, so we can write $n_t + h_t = (1-\delta)^t n_0 + \sum_{k=0}^{t}(1-\delta)^{t-k} h_k$. Using this identity, the left hand side of equation (17) can then be written as

$$\sum_{t=0}^{\infty} \beta^t \big(n_t + h_t\big) w_t = n_0 \sum_{t=0}^{\infty} \beta^t (1-\delta)^t w_t + \sum_{t=0}^{\infty} \beta^t \sum_{k=0}^{t} (1-\delta)^{t-k} h_k w_t. \tag{18}$$

The first term on the right of equation (18) can be written, using the free entry condition, as

$$-\frac{n_0\kappa}{q(\theta_0)} + n_0\sum_{t=0}^{\infty}\beta^t(1-\delta)^t z.$$

The second term can be written, rearranging and using the free entry condition, as

$$\begin{aligned}
\sum_{t=0}^{\infty}\beta^t\sum_{k=0}^{t}(1-\delta)^{t-k}h_k w_t &= \sum_{k=0}^{\infty}\beta^k h_k\sum_{t=k}^{\infty}\beta^{t-k}(1-\delta)^{t-k}w_t \\
&= -\sum_{k=0}^{\infty}\beta^k h_k\frac{\kappa}{q(\theta_k)} + \sum_{k=0}^{\infty}\beta^k h_k\sum_{t=k}^{\infty}\beta^{t-k}(1-\delta)^{t-k}z \\
&= -\sum_{t=0}^{\infty}\beta^t h_t\frac{\kappa}{q(\theta_t)} + \sum_{t=0}^{\infty}\beta^t\sum_{k=0}^{t}(1-\delta)^{t-k}h_k z.
\end{aligned}$$

These two terms combine into

$$\sum_{t=0}^{\infty}\beta^t[(n_t+h_t)z - \theta_t(1-n_t)\kappa] - \frac{n_0\kappa}{q(\theta_0)}$$

i.e., the right hand side of equation (17). To see this, note that $h_t/q(\theta_t) = \theta_t(1-n_t)$, and

$$n_0\sum_{t=0}^{\infty}\beta^t(1-\delta)^t z + \sum_{t=0}^{\infty}\beta^t\sum_{k=0}^{t}(1-\delta)^{t-k}h_k z = \sum_{t=0}^{\infty}\beta^t(n_t+h_t)z.$$

With partial unionization, the free entry condition changes, affecting this derivation. The free entry condition now implies that the present value of wages W_t satisfy

$$W_t = \sum_{k=0}^{\infty}\beta^k(1-\delta)^k z - \frac{\kappa}{\alpha q(\theta_t)} + \frac{1-\alpha}{\alpha}(1-\gamma)S_t.$$

Using this new free entry condition, the first and second terms on the right of equation (18) can be written, respectively, as

$$-\frac{n_0\kappa}{\alpha q(\theta_0)} + n_0\sum_{t=0}^{\infty}\beta^t(1-\delta)^t z + n_0\frac{1-\alpha}{\alpha}(1-\gamma)S_0,$$

and

$$-\sum_{t=0}^{\infty}\beta^t h_t\frac{\kappa}{\alpha q(\theta_t)} + \sum_{t=0}^{\infty}\beta^t\sum_{k=0}^{t}(1-\delta)^{t-k}h_k z + \frac{1-\alpha}{\alpha}(1-\gamma)\sum_{t=0}^{\infty}\beta^t h_t S_t.$$

These terms now combine into

$$\sum_{t=0}^{\infty}\beta^t[(n_t+h_t)z - \theta_t(1-n_t)\frac{\kappa}{\alpha} + \frac{1-\alpha}{\alpha}(1-\gamma)h_t S_t] - \frac{n_0\kappa}{\alpha q(\theta_0)} + \frac{1-\alpha}{\alpha}(1-\gamma)n_0 S_0.$$

□

Proof of Proposition 1 The union objective can be written in terms of the planner's value function, as in equation (3). The planner problem is standard, and known to have a linear solution $V(n)$, with the planner's choice of θ constant, independent of n. The union objective differs in the initial period by the $-n_0\kappa/q(\theta_0)$ term, however, which implies that the initial θ_0 is below the planner's choice, and this difference is greater the greater is n_0. □

Proof of Proposition 2 Consider a steady state of the unionized economy, where $\Theta'(n) = -c$ for some $c > 0$. Using this fact, steady-state employment can be written as $n = (1-\delta)\mu(\theta)/(1-(1-\delta)(1-\mu(\theta)))$, equation (16) implies that the steady-state θ satisfies the equation

$$1 = \mu'(\theta)\frac{z-b}{\kappa} + \beta(1-\delta)\big(1-\mu(\theta)+\theta\mu'(\theta)\big) - \Delta(\theta), \tag{19}$$

where $\Delta(\theta) \equiv 0$ in the efficient outcome, and

$$\Delta(\theta) \equiv -\frac{1-\delta}{\delta}\mu(\theta)\frac{q'(\theta)}{q(\theta)^2}[1-\beta(1-\delta)(1-\mu(\theta)-\frac{\mu'(\theta)\delta c}{1-(1-\delta)(1-\mu(\theta))})]$$

in the unionized economy. The term $\Delta(\theta)$ thus captures the union distortion. Under efficiency, the right-hand side of equation (19) is strictly decreasing in θ pinning down a unique steady-state θ (as long as $\mu'(0)\frac{z-b}{\kappa} + \beta(1-\delta) > 1$).[20] Because the union distortion $\Delta(\theta)$ is strictly positive for any $\theta > 0$, the unionized economy must have lower steady-state θ. □

Proof of Proposition 3 Similarly to Proposition 1, this result follows from writing the union problem in terms of the planner's value function, as in the text. □

B Numerical approach

This section discusses our numerical solution approach in the case where the union cannot commit to future wages. We begin with the benchmark model in Section 2, and then turn to the extensions.

B.1 Solving the benchmark model

As discussed in Section 3, solving the no-commitment union problem requires special care. With this in mind, we tried several different numerical approaches, comparing results across methods. We begin with an overview of the methods tried, before discussing the conclusions.

Local polynomial approximation approach to solving the generalized Euler equation Our baseline solution method is that outlined in Krusell, Kuruscu, and Smith (2002), based on solving the generalized Euler equation (16). This equation is a functional equation in $\Theta(n)$, defined over a range of values of n encompassing the steady state value of n. This approach amounts to calculating a Taylor polynomial approximating $\Theta(n)$ around its steady state. Calculating a k^{th} order polynomial involves first analytically differentiating the Euler equation k times with respect to n, acknowledging that $\Theta(n)$ is a function

[20] Note that $m_u(v, u) = \mu(\theta) - \mu'(\theta)\theta$, an expression which is reasonable to assume to be increasing in θ.

of n, and that $N(n, \Theta(n))$ is one as well. This yields $k + 1$ equations, which pin down the $k + 1$ coefficients in the polynomial. Evaluating the equations in steady state, with $n = \mu(\theta)(1-\delta)/(\delta+\mu(\theta)(1-\delta))$, the unknowns become the steady state values of $\theta, \theta', \theta'', ...$ up to the $k + 1$ derivative. Setting the last derivative to zero, the system determines these derivatives up to the k^{th} order. We first calculate the analytical derivatives, and the equations they yield, in Mathematica. We then turn to Matlab, solving for these derivatives (which determine the coefficients of the Taylor polynomial) using a non-linear equation solver. In practice, solving this system of equations can require a good initial guess, so we approach the problem iteratively, starting with a 0^{th} order Taylor polynomial and proceeding to successively higher-order polynomials, using the results from the previous step as initial guesses.

Global polynomial approximation approach to solving the generalized Euler equation As a functional equation, one can also look for a global solution to the Euler equation by approximating the solution $\Theta(n)$ with a cubic spline over some range of n's. Here we selected a grid on n, with the unknowns being the values of $\Theta(n)$ on that grid. These values determine the spline coefficients, which can be used to evaluate the Euler equation on the grid (and at intermediate points). This problem involves using a non-linear equation solver to find the values of $\Theta(n)$ on the grid, to minimize Euler equation errors.

Iterative approach to solving the generalized Euler equation One can also approach solving the Euler equation globally with an iterative approach. One way to do this is iterating backward, for example from a function $\Theta(n)$ which solves the final period optimization problem of a finite horizon union problem, with each step updating the values of $\Theta(n)$ on a fixed grid of n. In each step, for each grid point of n, we use the current set of $\Theta(n)$ to find n' next period, and then evaluate the right-hand-side of the Euler equation at these points using a cubic spline and the current set of $\Theta(n)$. One can then calculate a revised set of values of $\Theta(n)$, as the values of θ on the left-hand-side of the Euler equation.

Carroll's (2006) iterative approach to solving the generalized Euler equation One could also implement the iteration in the style of Carroll (2006), on an endogenous grid.

Here we first rewrite the Euler equation with $N(n)$ as the unknown function instead of $\Theta(n)$. In doing so, the equation will have three successive values $\{n_{t-1}, n_t, n_{t+1}\}$, instead of the two $\{\theta_t, \theta_{t+1}\}$. At each iteration, we have for a grid of n_t, and corresponding values of $n_{t+1} = N_t(n_t)$. With these we can use the Euler equation to calculate the corresponding values of n_{t-1}. This gives a new grid on n_{t-1}, over which we have corresponding values $n_t = N_{t-1}(n_{t-1})$.

Value function iteration Finally, one can also use a value function iteration approach. Starting from a guess for $\tilde{V}$, at each step we first solve the maximization problem determining the optimal $\Theta(n)$ on a grid of n, and then calculate the preceding period's value of $\tilde{V}$ using the recursive equation. A natural starting point is a value of $\tilde{V}$ consistent with the final period of a finite-horizon problem. Here the recursive equation is not a contraction, however, so there is no guarantee of convergence.

Conclusions Each of these methods shows convergence toward very similar results, which is reassuring. In particular, they deliver functions $\Theta(n)$ which are quite similar. Moreover, the steady states we find are all stable. But each method also exhibits signs of numerical instability. To some extent we would anticipate this, because the recursive expressions need not be contractionary, and therefore the iterations may not converge from arbitrary initial guesses. Moreover, even the non-iterative approach to solving the Euler equation may be sensitive to numerical error. It is possible that these numerical issues are related to the presence of multiple equilibria, which confuse the algorithms. The fact that these varied numerical methods nevertheless show signs of convergence to very similar outcomes supports the idea that the equilibrium we study exists and is the relevant one to study. The infinite-horizon model using the concept of a differentiable Markov-perfect equilibrium thus delivers very similar intuition to the one-period example we started with, supporting it as the natural candidate to consider.

B.2 Solving the extensions

In addition to the basic non-stochastic no-commitment union problem discussed above, we also consider extensions to allow: i) aggregate shocks, ii) partial unionization, and iii) collective bargaining. We describe below how we extended our numerical methods in order to compute solutions in these cases as well.

Aggregate shocks Our baseline solution method can be extended to allow aggregate shocks by treating $\Theta(n, z)$ as a function of z also. We approximate this function again as a k^{th} order polynomial in n, but include also a linear, and quadratic term in z, as well as an interaction term. The coefficients of the polynomial in n are the same as in the non-stochastic case. Finding the terms involving z requires differentiating the generalized Euler equation with respect to z and proceeding with the same approach as described for n above. (It is important not to stop at just a linear term in z here, as the coefficient on z sharpens as more terms are added.)

To evaluate this procedure, we compare the results in the case of a fully persistent shock to the transitional dynamics to a permanent shock calculated using our baseline approach for non-stochastic problems.

Partial unionization and collective bargaining We extend our baseline solution method to these cases. There is no generalized Euler equation here, so we need to alter the approach somewhat. For simplicity, we describe how we do this in the context of the collective bargaining problem, which is slightly more straightforward.

In the collective bargaining problem, the first order condition involves $\tilde{V}'(\cdot)$, as before, but now also $\tilde{V}(\cdot)$, which prevents us from simply eliminating these functions to arrive at a generalized Euler equation. We can still implement the basic approach by allowing these $\tilde{V}(\cdot)$ to remain in the first order condition as we successively differentiate it k times (analytically). We simply need to use the recursive equation for $\tilde{V}(\cdot)$ to compute the successively higher order derivatives of $\tilde{V}(\cdot)$ which will show up in these $k+1$ equations. As before, in doing so we acknowledge the law of motion $N(n, \Theta(n))$ as we proceed with taking derivatives.

In the partial unionization problem, the approach is similar, but in addition to needing to calculate derivatives of $\tilde{V}(\cdot)$ based on the recursive equation for $\tilde{V}(\cdot)$, one also needs to calculate derivatives of $S(\cdot)$ based on the recursive equation for $S(\cdot)$.

To evaluate this procedure, we compare the results in these extensions with our baseline model in the special cases where, in the case of collective bargaining, the union has full bargaining power, and in the case of partial unionization, the unionization rate is one.