The Union Threat*

Mathieu Taschereau-Dumouchel†
The Wharton School of the University of Pennsylvania

April 20, 2015

Abstract

This paper develops a search theory of labor unions in which the possibility of unionization distorts the behavior of nonunion firms. In the model, unions arise endogenously through a majority election. As unionized workers bargain collectively with the firm, unionization compresses the wage distribution and lowers profits. To prevent their own unionization, nonunion firms distort the skill composition of their workforce by over-hiring high-skill workers, who vote against the union, and under-hiring low-skill workers, who vote in its favor. Because of decreasing returns to labor, this change in hiring lowers output while reducing the dispersion of wages. In the calibrated economy, removing the threat of unionization, by freezing the union status of firms, reduces unemployment and increases output and the variance of wages. Removing, in addition, all unions from the economy leads to a larger increase in wage inequality but does not further affect output and unemployment. These results suggest that the threat that unionization exerts on nonunion firms, more than the fact that some firms are actually unionized, is the main channel through which unions affect output and unemployment in the U.S. economy. Finally, forcing all firms to be unionized lowers wage inequality and increases output, although less so than when the economy is union free.

JEL Classifications: J51, E24

*I am grateful to Nobuhiro Kiyotaki, Esteban Rossi-Hansberg and Oleg Itskhoki for their advice and support. I also thank Andy Abel, Zvi Eckstein, Lukasz Drozd, Pablo Fajgelbaum, Henry Farber, João Gomes, Thomas Lemieux, Edouard Schaal, Aleh Tsyvinski, as well as seminar participants at Booth, Boston University, Chicago, Chicago Fed, Cornell/PSU Macro Workshop, CREI, HEC Lausanne, IIES, McGill, Minneapolis Fed, Penn, Princeton, SED Ghent, St. Louis Fed, UCLA, Université de Montréal, UQAM, Wharton and Yale for helpful comments.

†Email: mathieu@wharton.upenn.edu; Address: 3620 Locust Walk, Philadelphia, PA 19104.

1 Introduction

As unions are now covering only 7% of private sector jobs in the United States, many observers have argued that their impact on the aggregate economy must be small.[1] In opposition to this view, this paper investigates how unions can nonetheless have a sizable impact on the macroeconomy through the influence they have on *nonunion* firms. Indeed, if unionization lowers profits, vulnerable nonunion firms would distort their behavior to prevent their own unionization. Through this channel, unions may influence employment, wages and output in nonunion firms and, as most firms are union free, in the aggregate economy.

To analyze the impact of this threat of unionization, this paper proposes a general equilibrium theory of endogenous union formation in which each firm hires multiple workers who differ in their productivity. In the model, unionization is a way for the workers to force the firm into a different wage setting mechanism. If a simple majority of the workers vote in favor of unionization, a union is created and wages are bargained collectively between the firm and all of its employees. If, instead, the vote fails to gather enough support, the firm remains union-free and wages are bargained individually between each worker and the firm.

By changing the scope of the wage bargaining, unionization generates two conflicts within the firm. First, as collective bargaining compresses the distribution of wages, high-productivity workers vote against the creation of the union while low-productivity workers vote in its favor. Unionization therefore creates a conflict between workers. Second, as collective bargaining allows the workers to extract a higher share of the production surplus, unionization increases the average wage and lowers profits, thereby creating a second conflict, this time between the firm and its workforce. The union threat affects the decisions of the firm through the interaction of these two conflicts: to prevent profits-reducing unionization, the firm hires more high-skill workers and fewer low-skill workers, thereby adjusting the outcome of the vote in its favor.[2]

In the theory, the distortion created by the union threat interacts with the decreasing returns in production to push the firm towards lower output and employment. As a result, the average marginal product of the workers increases which leads to a higher average wage. The union threat also affects wage inequality. As the firm over-hires high-skill workers, their marginal products decline which reduces their wage. Since the opposite happens to low-skill workers, nonunion firms pay a more compressed distribution of wages in response to the threat of unionization.

In the model, the labor market is subject to search frictions: it takes time for workers to be matched with vacancies. The resulting unemployment level is also affected by the union threat. In general equilibrium, as the threatened firms hire less, the unemployment rate goes up and it takes more time for workers to find jobs. As unemployment becomes less attractive, firms are able to extract a higher share of production and wages go down.

[1] See Hirsch and Macpherson (2003) and their database at `http://www.unionstats.com/`.

[2] Empirical studies have found that low-skill workers favor unions while high-skill workers do not (Farber and Saks, 1980). It is therefore natural that a firm that wants to avoid unionization modifies the skill composition of its workforce to increase its odds of remaining union free.

By providing a microfounded bargaining theory of unionization, the model is able to replicate important empirical facts associated with unions: *i)* union wages have a smaller variance and are on average higher than nonunion wages (Card et al., 2004), *ii)* the preference for unionization and the difference between union and nonunion wages decrease with skill (Farber and Saks, 1980), and *iii)* unionized firms are on average less profitable than their nonunion counterpart (Hirsch, 2004).

To quantify the impact of the union threat, I calibrate the economy to the private sector of the United States in 2005 and use the model to conduct three policy experiments in general equilibrium. The first experiment consists of removing the threat of unionization: the union status of all firms is fixed and cannot be changed anymore. As a result, nonunion firms stop distorting their behavior to prevent unionization. This first policy experiment captures the impact of the threat of unionization alone, as the union status of all firms remain unchanged. In the new equilibrium, the variance of log wages goes up by 4.3%, output increases by 1.7%, while the unemployment rate decreases by 3.3 percentage points. If, in addition to removing the threat of unionization, all union firms are forced to become union free, the variance of log wages goes up by an additional 3.8%, but output and unemployment are not further affected. This second policy experiment suggests that the threat of unionization alone, more than the fact that some firms are actually unionized, is the main channel through which unions affect output and unemployment in the U.S. economy. Finally, in the third policy experiment, all firms are forced to be unionized. Comparing this new equilibrium to the calibrated economy, the variance of log wages goes down by 27% while output and employment increase, although less so than when the economy is union free. These policy experiments confirm that the union threat has a substantial effect on output and unemployment while the impact of the union status of firms is mostly on wage inequality.

One contribution of the paper is to highlight and quantify ways in which unions affect wage inequality that are not taken into account by empirical estimators. Indeed, the policy experiments find that unions have a larger impact on wage inequality than the empirical literature suggests. For instance, Dinardo and Lemieux (1997) estimate that in 1988 unions reduced the variance of male wages by 3% in the United States. Similarly, Card (2001) finds that in 1993 unions lowered the variance of male wages by 5% while having little impact on the variance of women wages.[3] Finally, my own calculations using the classical estimator of Freeman (1980) on the calibrated economy find that unions reduce the variance of wages by only 0.4%, twenty times less than the policy experiments suggests.[4] This large difference between the empirical literature and the outcome of the policy experiments can be explained by the threat of unionization, as it induces *nonunion* firms to pay a more equal distribution of wages. Standard empirical estimators do not capture this

[3]Note that these estimates consider years in which the union coverage was higher than in 2005, the year of the calibration. Also, unlike the data considered in the calibration, they include workers from the public sector, where unionization rates are much higher than in the private sector. See Card et al. (2004) for a summary of the literature on unions and wage inequality.

[4]The Freeman estimator can be written (Card et al., 2004) as $V - V^n = U\Delta_v + U(1-U)\Delta_w^2$ where V is the observed variance of log wages, V^n is the variance of log wages without unions, U is the unionization rate, Δ_v is the difference in the variance of log union and nonunion wages and Δ_w is the difference between the mean log of union and nonunion wages. Using it on the data generated by the calibrated economy gives 0.4%.

channel.

The theory also provides an explicit mechanism to explain why regression discontinuity studies, such as DiNardo and Lee (2004), find little impact of unionization on firms. These studies compare firms before and after unionization. According to the theory, before unionization, these firms are actively distorting their behavior in response to the threat of unionization. As a result, regression discontinuity estimators only capture part of the impact of unions on firms.[5]

A large body of evidence suggests that the possibility of unionization distorts the behavior of firms. Holmes (1998) shows that firms prefer to locate their establishments in states with union-weakening laws. Firms also employ a wide array of techniques, legal and illegal, to prevent their own unionization (Dickens, 1983; Bronfenbrenner, 1994; Freeman and Kleiner, 1990). Matsa (2010) shows that union bargaining power influences corporate financing decisions. Corneo and Lucifora (1997) consider a model in which firms preemptively increase wages if they believe a union will force costly negotiations.

This paper is part of a literature that includes labor unions in search models. Pissarides (1986) finds that introducing a monopoly union with control over the wage in a search framework might lead to efficiency. Alvarez and Veracierto (2000) study the impact of many labor market policies in a search model. They find that unions who control hiring have adverse effects on unemployment and welfare. Ebell and Haefke (2006) and Delacroix (2006) investigate the interaction between union formation and product market regulations. Açikgöz and Kaymak (2014) estimate the impact of a rising skill premium on the decline of union membership in the United States. Boeri and Burda (2009) look into the impact of an endogenous bargaining regime on economic activity. Recently, Krusell and Rudanko (2012) have studied the dynamic problem of a monopoly union that sets wages with or without commitment. In contrast to the literature, this paper investigates the impact of the threat of unionization on decision makers and the macroeconomy.

The next section introduces the model. An explanation of how firms respond to the distortion created by the union threat follows. The model is then calibrated to the U.S. economy and policy experiments are conducted to evaluate the impact of unions. The last section concludes.

2 Model

2.1 Preferences and technology

The economy is populated by heterogeneous workers, each endowed with a skill $s \in \{1, \dots, S\}$, constant over time. The exogenous distribution of skills in the economy is N_s, with $N_s > 0$ for all s. Workers live forever, are risk-neutral and discount future consumption at the rate $0 < \gamma < 1$.

Firms combine the labor provided by workers of different skills into consumption goods. To do so, they use one of two production technologies, indexed by $j \in \{1, 2\}$.[6] A firm of type j that

[5] DiNardo and Lee (2004) discuss how the union threat may contribute to their results. Frandsen (2011) uses a regression discontinuity approach to estimate the impact of unionization on the full wage distribution.

[6] It is straightforward to extend the model to accommodate a more general distribution of firm types.

employs a (non-normalized) distribution of workers g_s produces goods according to the production function

$$F_j(g) = A_j \left[\left(\sum_{s=1}^{S} z_{j,s} g_s^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \right]^{\alpha_j} \tag{1}$$

where $A_j > 0$ is total factor productivity and $\sigma > 0$ is the elasticity of substitution between the different skills. The vector $z_{j,s} > 0$ represents the relative intensity of skill utilization in firm j and is normalized to sum to one. The parameter $0 < \alpha_j < 1$ describes the returns to scale of the production function.[7] To avoid cluttering the notation, the subscript j is often omitted when referring to a single firm.[8]

In the model, the notion of skill is only used to characterize some form of worker heterogeneity. The productivity of a worker of type s is determined in equilibrium by the supply (related to N_s) and the demand (related to $z_{j,s}$) for this skill.

2.2 Labor markets

There are S labor markets in which unemployed workers search for jobs and firms post vacancies. Workers only search in the labor market corresponding to their skill but firms are free to post multiple vacancies, at a unit cost κ, in multiple markets. This segmentation of the labor markets by skill groups allows the firm to control precisely the skill composition of its workforce and, through this channel, influence the unionization vote.[9] The presence of search frictions generates a gap between union and nonunion wage. This union wage gap, which is well documented in the empirical literature, influences wage inequality in equilibrium.

Since, in equilibrium, each type of firms posts vacancies in each market, a searching worker can be matched with firms that use different technologies. This implies that the uncertainty facing the unemployed workers is about whether they meet a firm or not and, if so, what type of firm they meet.

In a labor market where U unemployed workers are searching and V vacancies are posted, $m(U, V)$ random matches are created in a period. The matching function m is assumed to be strictly concave, strictly increasing and homogenous of degree one. By defining the labor market tightness $\theta = V/U$, the probability that a vacancy is filled is $q(\theta) = m(U, V)/V = m(1/\theta, 1)$ and the probability that an unemployed worker finds a job is $p(\theta) = m(U, V)/U = m(1, \theta)$. Since search requires no effort, all unemployed workers are searching. At the end of each period, a fraction δ of jobs are exogenously destroyed.

[7] Appendix D shows that including capital in the production function is straightforward and simply leads to a reinterpretation of the parameters.

[8] Acemoglu et al. (2001), Açikgöz and Kaymak (2014) and Dinlersoz and Greenwood (2012) investigate the link between technological changes and labor unions.

[9] The assumption of segmented markets is not necessary for the main mechanism to be active. As long as a firm has some control over the skills of the workers it hires the threat of unionization will influence its decision.

2.3 Firms

A firm that employed a distribution of workers g_{-1} during the previous period starts the current period with the distribution $(1-\delta)g_{-1}$. It then posts a schedule of vacancies v to maximize its expected discounted profits. Since the firm is posting a continuum of vacancies in each labor market, a law of large numbers implies that the number of successful matches is deterministic.

By defining the period profit $\pi(g) = F(g) - \sum_s w_s(g)g_s$, where $w_s(g)$ is the wage, the recursive problem of a firm is

$$J(g_{-1}) = \max_{v \geq 0} \pi(g) - \kappa \sum_{s=1}^{S} v_s + \gamma J(g) \tag{2}$$

subject to, for all s, the law of motion of the distribution of workers

$$g_s = g_{-1,s}(1-\delta) + v_s q(\theta_s)$$

which states that current workers were either with the firm last period or are newly hired.

At a steady state, we can simplify the problem of the firm substantially. At the beginning of a period, the firm has a fraction $1-\delta$ of its optimal distribution of workers. Since its hiring cost is linear, it immediately hires back to its optimal level. The constraint $v_s \geq 0$ is therefore never binding and we have the following lemma:

Lemma 1. *In a steady-state equilibrium, the firm's dynamic problem is equivalent to*

$$\max_{g} \pi(g) - \kappa \sum_{s=1}^{S} \frac{g_s}{q(\theta_s)} + \kappa(1-\delta)\gamma \sum_{s=1}^{S} \frac{g_s}{q(\theta_s)}. \tag{3}$$

Proof. All the proofs are in the appendix. □

This equation states that hiring an extra worker generates some profit (first term), that it requires posting some costly vacancies (second term) and that, since the additional worker remains with the firm with probability $1-\delta$, it lowers hiring costs next period (last term).

2.4 Workers

In each period, a worker is either employed or unemployed. Employed workers lose their jobs with probability δ, in which case they become unemployed. The lifetime discounted expected utility of a worker of type s who is matched with a firm of type j and who is earning a wage w is therefore

$$W_{j,s}^{e}(w) = w + \gamma \left[\delta W_s^u + (1-\delta) W_{j,s}^{e}(w_{j,s})\right] \tag{4}$$

where W_s^u is the lifetime utility of being unemployed and $w_{j,s}$ is the wage that a worker of type s expects to receive next period if there is no job separation. Since wages are bargained every period, the negotiations with the firm are over w only. Both parties take $w_{j,s}$ as given and it is determined in equilibrium.

At the beginning of a period, an unemployed worker finds a job with probability $p(\theta_s)$. The expected value of this job is $\mathbb{E}(W^e_{j,s})$, where the expectation is taken over all the vacancies, posted by different types of firms, in submarket s. If no job is found, the worker receives b_s from home production, with b_s increasing in s. The lifetime discounted utility of an unemployed worker is therefore

$$W^u_s = p(\theta_s)\mathbb{E}(W^e_{j,s}) + (1 - p(\theta_s))\left[b_s + \gamma W^u_s\right]. \tag{5}$$

By combining the last two equations we can characterize the utility gain provided by employment at wage w:

$$W^e_{j,s}(w) - b_s - \gamma W^u_s = w - c_{j,s} \tag{6}$$

where

$$c_{j,s} = b_s + \gamma(1-\delta)\frac{(1-\gamma)W^u_s - w_{j,s}}{1-\gamma(1-\delta)} \tag{7}$$

is defined as the net outside option of a worker s who is bargaining with a firm j. This convenient notation makes explicit the fact that the worker loses all potential future wages $w_{j,s}$ if the bargaining breaks down.

2.5 Wages

In the United States, the typical unionization process starts when a group of workers petition the National Labor Relation Board (NLRB) for a union recognition.[10] If there is sufficient interest from employees, the NLRB makes a ruling on whether the workers that would be covered by the union share a "community of interest". In practice, the coverage of the union is usually at the enterprise level (Traxler, 1994; Nickell and Layard, 1999).[11] Then, the NLRB organizes a vote at the work site and a simple majority is required for the union to be certified as the exclusive bargaining agent of the workers. All work related negotiations between the workers and the firm must then be conducted by the union.

The model integrates these features of the institutional environment. The sequence of events that occurs once a firm has hired its new workers is represented on figure 1. First, the workers vote to decide whether to form a union or not. Then, if the union vote is successful, wages are bargained *collectively*. The outcome of this bargaining is a wage schedule $w^u_s(g)$ and a profit function $\pi^u(g)$. Instead, if the union vote fails, wages are bargained *individually*. This generates the wage schedule

[10] See DiNardo and Lee (2004) for more details about the unionization process.

[11] A large fraction of the literature models unions at the level of the production function. For recent examples, see Ebell and Haefke (2006) and Dinlersoz and Greenwood (2012).

$w_s^n(g)$ and the profit function $\pi^n(g)$. Unionization is therefore a way for the workers to force the firm into a different wage setting mechanism.

Both individual and collective bargaining are modeled using Nash bargaining, but the surplus that is bargained over is different in both cases. In a union firm, the workers and the firm bargain collectively over the total surplus generated by *all* the workers. If an agreement on wages cannot be reached, the whole workforce leaves the firm and no production takes place. In a nonunion firm, each worker bargains individually with the firm over the *marginal* surplus he or she alone generates. If the bargaining fails, this specific worker goes to unemployment but the firm can still produce with the remaining workers. As we will see, this asymmetry between collective and individual bargaining interacts with the decreasing returns of the production function and has important consequences for profits and wages. It is the only difference between a union and a nonunion firm in the model.

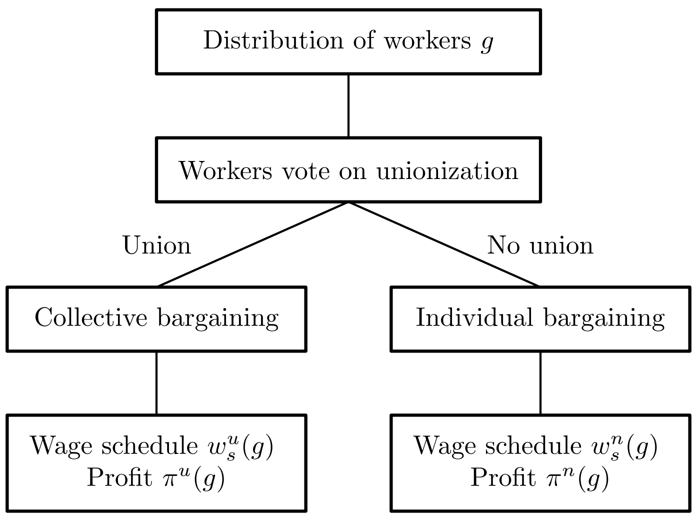


Figure 1: Timing of events in a firm after hiring

Collective bargaining

Collective bargaining is modeled as an n-player Nash bargaining between the firm and all its workers.[12] If an agreement on a wage w is reached, a worker s receives $W_s^e(w)$, otherwise he or she receives home production b_s today and starts the next period as unemployed, which has value γW_s^u. The net benefit of an agreement to a worker is therefore $W_s^e(w) - b_s - \gamma W_s^u$. On the firm side, if an agreement is reached production takes place and wages are paid. Otherwise, the firm loses its workers and needs to hire extensively next period to get back to its optimal size.

The following Lemma formalizes this collective bargaining problem.

Lemma 2. *If all the workers have the same bargaining power, and the firm has bargaining power*

[12] Nash bargaining with more than two players is microfounded in axiomatic bargaining theory (Roth, 1979) and in game theory (Krishna and Serrano, 1996).

$1-\beta_u$, the collective Nash bargaining problem can be written as

$$\max_{w} \left[\prod_{s=1}^{S} \left(W_s^e(w) - b_s - \gamma W_s^u\right)^{\frac{g_s}{n}} \right]^{\beta_u} \left[F(g) - \sum_{s=1}^{S} w_s g_s + (1-\delta)\kappa\gamma \sum_{s=1}^{S} \frac{g_s}{q(\theta_s)} \right]^{1-\beta_u} \tag{8}$$

where $n = \sum g_s$ is the total number of employed workers. Furthermore, the wage equation

$$w_s^u(g) - c_s = \frac{\beta_u}{n} \left(F(g) - \sum_{k=1}^{S} c_k g_k + \gamma(1-\delta)\kappa \sum_{k=1}^{S} \frac{g_k}{q(\theta_k)} \right) \tag{9}$$

solves this bargaining problem.

This collective bargaining problem is very similar to the usual 2-player bargaining. The first term between brackets in (8) can be interpreted as the surplus of the union; it takes the simple form a geometric average of all the workers' individual surpluses. The second term between brackets is the surplus of the firm. Its interpretation is straightforward; if negotiations break down, the firm loses the current period profit and pays a higher hiring cost tomorrow to compensate for the loss of the fraction $1-\delta$ of its current workforce that would have remained with it next period if negotiation had been successful.[13]

From the wage equation 9, it is straightforward to compute the one-period profit of a union firm employing the distribution of workers g

$$\pi^u(g) = (1-\beta_u)F(g) - (1-\beta_u)\sum_{s=1}^{S} c_s g_s - \beta_u(1-\delta)\kappa\gamma \sum_{s=1}^{S} \frac{g_s}{q(\theta_s)}. \tag{10}$$

One key advantage of using collective bargaining to model the union negotiation process is that there is no need to model an actual union organization explicitly. Here, a unionized firm is simply a firm in which wages are bargained collectively.[14] In particular, there is no union leader or union preferences to specify. The union is simply the collective of the workers bargaining together.

Individual bargaining

If the workers vote against the union, they each bargain individually with the firm. The surplus to split is, however, dependent on the outcome of the bargaining between the firm and all the *other* workers. Indeed, if one of the bargaining sessions breaks down and a worker leaves the firm, the marginal product of the remaining workers will change. As a result, these workers or the firm might want to reopen the bargaining. Stole and Zwiebel (1996a,b) provide the theoretical foundation for

[13] Açikgöz and Kaymak (2014) and Bauer and Lingens (2010) assume instead that the union maximizes the sum of the workers' surplus. With heterogeneous workers, this approach only pins down the *total* share of the surplus going to the workers, not how it is shared among them. It also requires specifying preferences for the union.

[14] Appendix E shows how one can explicitly model a union as an intermediary between the workers and the firms. If the workers bargain collectively with the union, the wage is the same as in equation 9. Equation 9 is therefore consistent with various assumptions about the nature of the union.

this type of bargaining.[15]

In this context, the firm's marginal gain from employing an extra worker of type s is[16]

$$\Delta_s^n(w) = \frac{\partial F(g)}{\partial g_s} - w_s(g) - \sum_{k=1}^{S} g_k \frac{\partial w_k(g)}{\partial g_s} + \gamma(1-\delta)\frac{\kappa}{q(\theta_s)}.$$

The first term is the extra output produced by the worker. The next one is simply the wage paid to the worker. The third term is the marginal effect of this worker on the wages of other members of the workforce. Finally, the last term is the expected vacancy costs saved from retaining, with probability $1-\delta$, this worker next period.

Defining $0 < \beta_n < 1$ as the bargaining power of a nonunion worker, Nash bargaining implies that the nonunion wage must solve the system of partial differential equations

$$\Delta_s^n(w) = \frac{1-\beta_n}{\beta_n}\left(W_s^e(w) - b_s - \gamma W_s^u\right), \tag{11}$$

for all s with the standard boundary conditions $\{\lim_{g_s \to 0} w_s^n(g) g_s = 0\}_{s=1}^{S}$. The solution to this system is characterized in the following lemma.

Lemma 3. *The wage schedule*

$$w_s^n(g) - c_s = \frac{\beta_n}{1-(1-\alpha)\beta_n} \underbrace{\frac{\alpha z_s}{g_s^{1/\sigma}} A \left(\sum_{k=1}^{S} z_k g_k^{\frac{\sigma-1}{\sigma}}\right)^{\frac{1-\sigma(1-\alpha)}{\sigma-1}}}_{\text{Marginal product of worker } s} - \beta_n c_s + \beta_n \gamma(1-\delta)\frac{\kappa}{q(\theta_s)} \tag{12}$$

solves the bargaining problem of a nonunion firm.

It follows directly that the one-period profit of the firm is

$$\pi^n(g) = \frac{1-\beta_n}{1-(1-\alpha)\beta_n} F(g) - (1-\beta_n)\sum_{s=1}^{S} c_s g_s - \beta_n(1-\delta)\kappa\gamma \sum_{s=1}^{S} \frac{g_s}{q(\theta_s)}. \tag{13}$$

Comparing collective and individual bargaining

The wage equations 9 and 12, from the collective and the individual bargaining, have a remarkably similar structure. They both consist of three terms: one related to production, one related to the outside option of the workers and one related to the hiring costs. They, however, differ in how these quantities influence wages. Indeed, the union wage is a function of the *average* characteristics of the firm's workforce, while the nonunion wage is a function of the *individual* characteristics of

[15] Bertola and Garibaldi (2001) show that a standard search model with wages set through a Stole and Zwiebel procedure is broadly consistent with the empirical "relationship between employer size, the mean and variance of employees' wages, and the character of gross job creation and destruction". Cahuc and Wasmer (2001), Elsby and Michaels (2013), Acemoglu and Hawkins (2014) also use Stole and Zwiebel bargaining in search models.

[16] See the proof of lemma 3 in the appendix for a derivation of this equation.

each worker. In particular, the union wage depends on the average production $F(g)/n$ while the nonunion wage is a function of the marginal product of the worker, as shown in (12). This asymmetry has two important consequences. First, the presence of a union influences the variance of wages within the firm. Second, the possibility of unionization creates a conflict between workers of different skills. Indeed, a worker with valuable characteristics, for instance a high marginal product, would rather bargain individually with the firm than share their advantage with the other employees.[17]

The following proposition shows that unionization also creates a second conflict, this time between the firm and its workers.

Proposition 1. *If the bargaining powers are equal ($\beta = \beta_n = \beta_u$) than the difference between nonunion and union average wages is*

$$\mathbb{E}_g(w^n) - \mathbb{E}_g(w^u) = -\frac{\beta(1-\beta)(1-\alpha)}{1-(1-\alpha)\beta}\frac{F(g)}{n} < 0$$

where $\mathbb{E}_g$ is the expectation across skills. Also, the difference in the profit of the firm is

$$\pi^n(g) - \pi^u(g) = \frac{(1-\beta)(1-\alpha)\beta}{1-(1-\alpha)\beta}F(g) > 0.$$

This proposition shows that, for any distribution g, the firm prefers to bargain individually, while the workers, on average, would rather be represented by a union. This conflict of preferences is a direct consequence of the decreasing returns to scale. Indeed, as $\alpha \to 1$, the differences in profits and in average wages go to zero. When bargaining individually, the firm considers producing with or without the *marginal* worker. Because of diminishing returns to labor, this marginal worker has a relatively small impact on the total production, limiting their possibility to bargain. The firm can then extract a large share of the total surplus. On the other hand, when the firm bargains with the union, the surplus is a function of the *total* production, which includes the relatively high output generated by the infra-marginal workers. By forming a union, the workers can thus extract a bigger share of these high marginal products, which lowers the firm's profit.

This feature of the model is consistent with evidence from Kleiner (2001) showing that firms generally oppose unions. Bronfenbrenner (1994) and Freeman and Kleiner (1990) also detail various tactics used by firms to prevent unionization. Hirsch (2004) summarizes the literature on union and profitability and concludes that union firms are in general less profitable than firms that are not unionized.

2.6 Voting procedure

When the union vote takes place, the distribution of workers is fixed and the workers know the wages they will get after either outcome of the vote. Each worker has random preferences

[17] Verna (2005) discuss the literature on the relationship between measures of productivity and pay in union firms. Consistent with the theory, pay is much more correlated with ability and performance in nonunion firms than in union firms.

over the union status of the firm. Specifically, a worker of skill s votes for a union if and only if $w^u_s(g) - w^n_s(g) > \epsilon$, where ϵ is a disutility cost of being part of a union. It has mean 0, is drawn independently across workers in each period and has CDF ϕ with $\phi(0) = 1/2$.

A law of large numbers applies when aggregating the workers of a given skill. Therefore, a fraction $\phi(w^u_s(g) - w^n_s(g))$ of workers of type s votes for unionization. Denoting by

$$V(g) = \sum_{s=1}^{S} g_s \phi(w^u_s(g) - w^n_s(g)) - \frac{1}{2}n \tag{14}$$

the excess number of workers in favor of unionization, a firm is unionized if and only if $V(g) > 0$.

Notice that even though the preference shocks ϵ are random, the outcome of the vote is fully deterministic. Therefore, at the moment of posting vacancies, the firm knows whether the workers will form a union or not. The firm is effectively deciding its union status.

2.7 Problem of a firm

Now that we have derived the wage schedules and outlined the voting procedure, we can go back to the problem of a firm. As shown in lemma 1, at a steady state, a firm solves

$$\max_{g} \mathcal{J}(g, w(g)) \tag{15}$$

with

$$\mathcal{J}(g, w(g)) = F(g) - \sum_{s=1}^{S} g_s w_s(g) - \kappa(1 - (1-\delta)\gamma) \sum_{s=1}^{S} \frac{g_s}{q(\theta_s)}$$

where

$$w(g) = \begin{cases} w^u(g) & \text{if } V(g) > 0 \\ w^n(g) & \text{if } V(g) \leq 0 \end{cases}$$

and where $V(g)$ is given by (14), $w^u(g)$ is given by (9) and $w^n(g)$ is given by (12).

2.8 Steady-state equilibrium

In a steady state, the flows in and out of unemployment in all sub-markets must be equal. This implies that the unemployment rate is

$$u = \frac{U_s(1 - p(\theta_s))}{N_s} = \frac{\delta(1 - p(\theta_s))}{\delta + p(\theta_s)(1-\delta)} \tag{16}$$

where U_s is the number of workers searching in market s.

We can now define a steady-state equilibrium.

Definition 1. A steady-state equilibrium is a set of value functions $\left\{W^e_{j,s}, W^u_s\right\}_{s\in\{1,\dots S\}}^{j\in\{1,2\}}$, labor market tightnesses $\{\theta_s\}_{s=1}^{S}$, distributions of workers $\{g_s^j\}_{s\in\{1,\ldots,S\}}^{j\in\{1,2\}}$ and wage schedules $\{w_s^j\}_{s\in\{1,\ldots,S\}}^{j\in\{1,2\}}$ such that,

1. the value functions $\left\{W_{j,s}^e, W_s^u\right\}_{s\in\{1,\ldots S\}}^{j\in\{1,2\}}$ solve (4) and (5);
2. the distribution $\left\{g_s^j\right\}_{s\in\{1,\ldots S\}}$ solves the optimization problem of firm j given by (15);
3. the wage schedule $\left\{w_s^j\right\}_{s\in\{1,\ldots S\}}$ solves the collective bargaining problem, it satisfies (9), if firm j is unionized or the individual bargaining problem, it satisfies (12), if firm j is not unionized;
4. unemployment is stationary in each labor market such that (16) is satisfied.

3 Economic forces at work

In an economy without the possibility of forming a union, all workers would be paid their individual bargaining wage $w^n(g)$ and the firm would simply hire to maximize its discounted profits. Let us denote this optimal distribution $g^{n*} = \text{argmax}_g \; \mathcal{J}(g, w^n(g))$. Instead, if forming a union is possible, we know from proposition 1 that the workers are likely to do so. If this happens, the firm is *constrained* by the unionization vote and hiring according to g^{n*} is not optimal.

A constrained firm will consider distorting g^{n*} to make the workers reject the union. Formally, this distorted distribution g^n maximizes $\mathcal{J}(g, w^n(g))$ subject to the voting constraint $V(g) \leq 0$. Because of this additional constraint, the possibility of unionization modifies the behavior of firms that are union free in equilibrium. It is through this channel that the threat of unionization affects the economy.

The constraint $V(g) \leq 0$ is also responsible for the existence of union firms in equilibrium. Since workers generally favor unions, the profits of the firm would then be $\mathcal{J}(g^{u*}, w^u(g^{u*}))$ where g^{u*} is the optimal distribution of workers under collective bargaining. If the voting constraint is strong enough, such that $\mathcal{J}(g^{u*}, w^u(g^{u*})) > \mathcal{J}(g^n, w^n(g^n)))$, the firm chooses to be unionized as an optimal reaction to the unionization threat.[18]

To understand the mechanisms at work, it is useful to first consider an equilibrium in which the union status of each firm is given exogenously, such that no union vote takes place. In this case, we can characterize how the firm hires, the wages it pays as well as the workers' preference for unionization. The full problem of a firm, in which its union status depends endogenously on the vote, can then be thought of as a deviation from the exogenous case.

3.1 Exogenous union status

We first consider the problem of a firm whose union status is exogenously given, such that the union threat has no impact on its behavior. Such a firm maximizes $\mathcal{J}(g, w^i(g))$ where $i = u$ if the

[18] It is possible to build examples in which $V(g^{u*}) < 0$ but this requires extreme parameters.

firm is unionized and $i = n$ otherwise. By defining

$$\mathrm{MC}_s^i = (1-\beta_i)c_s + (1-\gamma(1-\delta)(1-\beta_i))\frac{\kappa}{q(\theta_s)} \tag{17}$$

as the marginal cost paid to hire a worker s, the firm's optimal hiring decision g^{i*} solves

$$\mathrm{MC}_s^i = B_i \frac{\alpha A z_s}{(g_s^{i*})^{1/\sigma}} \left(\sum_{k=1}^{S} z_k (g_k^{i*})^{\frac{\sigma-1}{\sigma}} \right)^{\frac{1-\sigma(1-\alpha)}{\sigma-1}} \tag{18}$$

where

$$B_i = \begin{cases} 1-\beta_u & \text{if } i = u \\ \frac{1-\beta_n}{1-(1-\alpha)\beta_n} & \text{if } i = n \end{cases}$$

is the share of output retained by the firm. Equation 18 simply states that at the optimum the marginal revenue from hiring an extra worker of type s equals its marginal cost. Solving this equation, we find that the optimal hiring decision g^{i*} is

$$g_s^{i*} = (\alpha A B_i)^{\frac{1}{1-\alpha}} \left(\frac{z_s}{\mathrm{MC}_s^i} \right)^{\sigma} \left(\sum_{k=1}^{S} z_k \left(\frac{z_k}{\mathrm{MC}_k^i} \right)^{\sigma-1} \right)^{\frac{1-\sigma(1-\alpha)}{(\sigma-1)(1-\alpha)}}. \tag{19}$$

We see that workers who search in tight labor markets (θ_s large) or who have attractive outside options (c_s large) are expensive to hire (MC_s^i large) and the firm therefore relies less on them for production (g_s^{i*} small). All else equal, nonunion firms are also larger than union firms as they tend to hire more workers to lower their marginal products and thus pay lower wages.

The following proposition characterizes the wages paid by firms.

Proposition 2. *If W_s^u and θ_s are increasing in s, then the equilibrium wage schedules $w_s^u(g^{u*})$ and $w_s^n(g^{n*})$ are increasing in s and the union wage gap $w_s^u(g^{u*}) - w_s^n(g^{n*})$ is increasing in s.*[19]

This proposition is consistent with a large empirical literature that finds that the union wage gap in the U.S. declines with income (Card, 1996; Card et al., 2004). It characterizes the *observed* wages that are paid in equilibrium but not the workers' preferences about unionization. For those, we need to consider the counterfactual wages that the workers would receive if they would vote in favor or against unionization in a given firm.

Proposition 3. *If W_s^u and θ_s are increasing in s, then the counterfactual union wage gap $w_s^u(g^{i*}) - w_s^n(g^{i*})$ is increasing in s for $i \in \{u, n\}$.*

Proposition 3 is consistent with Farber and Saks (1980), who show that the desire to be unionized goes down with the position of the worker in the intra-firm earnings distribution, such that low-skill workers have a higher preference for unions than high-skill ones.

[19] In the calibrated economy, W_s^u and θ_s are increasing in s.

Finally, we can compare the profits of the firm under both the union and the nonunion scenarios.

Proposition 4. *An unconstrained firm strictly prefers to be union free if and only if*

$$\frac{B_n}{B_u} > \left(\frac{\sum_s z_s^\sigma \left(\mathrm{MC}_s^u\right)^{1-\sigma}}{\sum_s z_s^\sigma \left(\mathrm{MC}_s^n\right)^{1-\sigma}} \right)^{\frac{\alpha}{\sigma-1}}.$$

The left-hand side of this inequality compares the benefit of being union free to the benefit of being unionized, while its right-hand side compares the hiring costs under both union statuses. If $\beta_n = \beta_u$, then $\mathrm{MC}^n = \mathrm{MC}^u$ and $B_n > B_u$, this condition is automatically satisfied. Also, the firm prefers to be union free when unions are strong ($\beta_u \to 1$) and it would welcome a union if individual workers have strong bargaining power ($\beta_n \to 1$).

It is useful to consider a numerical example, presented in figure 2, to understand the behavior of the firm.[20] Panel (a) shows the two optimal distributions, g^{n*} and g^{u*}, that are chosen by nonunion and union firms. Panels (b) and (c) show the wages that workers would be considering if a vote were to take place. Panel (b) displays the wage schedules when the firm hires according to g^{n*} while panel (c) shows the same schedules when the firm employs g^{u*}. Consistent with the previous propositions, the wage schedules increase with s, while the union wage gap $w^u - w^n$ decreases with skill, indicating that low-skill workers have a higher preference for unions than high-skill ones.

In this example, if a vote were to take place in a firm hiring according to g^{n*}, a majority of the workers would vote for unionization. The distribution g^{n*} would therefore not be an optimal decision if, instead of being exogenous, the union status of the firm was decided by the workers' vote. We therefore need to understand how the firm behaves when this threat of unionization is binding.

3.2 Preventing unionization

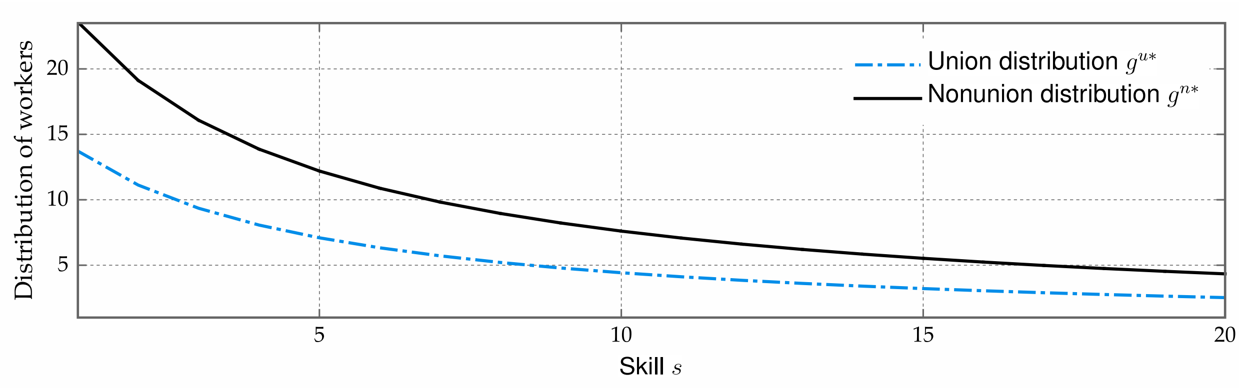
We now consider the problem of a firm whose union status is endogenously determined by the vote of its workers. The firm therefore compares its profits under two distributions of workers: the optimal one under which the workers unionize g^{u*} and the optimal one under which the workers reject the union g^n.

The optimal nonunion distribution g^n solves a modified version of the first-order conditions 18 that takes into account the impact of the workers on the vote. The new conditions are

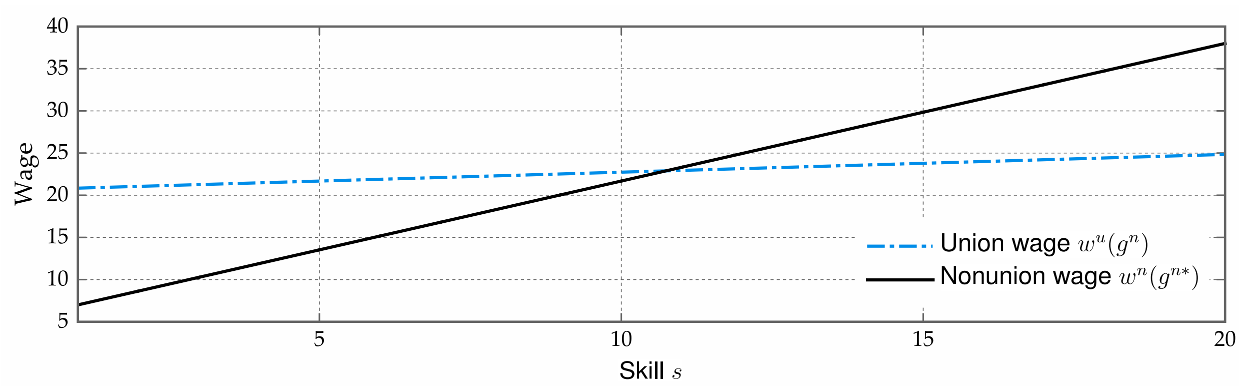
$$\mathrm{MC}_s^n + \lambda \frac{\partial V(g^n)}{\partial g_s^n} = B_n \frac{\alpha A z_s}{(g_s^n)^{1/\sigma}} \left(\sum_{k=1}^{S} z_k (g_k^n)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{1-\sigma(1-\alpha)}{\sigma-1}} \tag{20}$$

where $\lambda \geq 0$ is the Lagrange multiplier associated with the voting constraint. We see that the voting constraint effectively increases the marginal cost of hiring workers who vote in favor of the union.

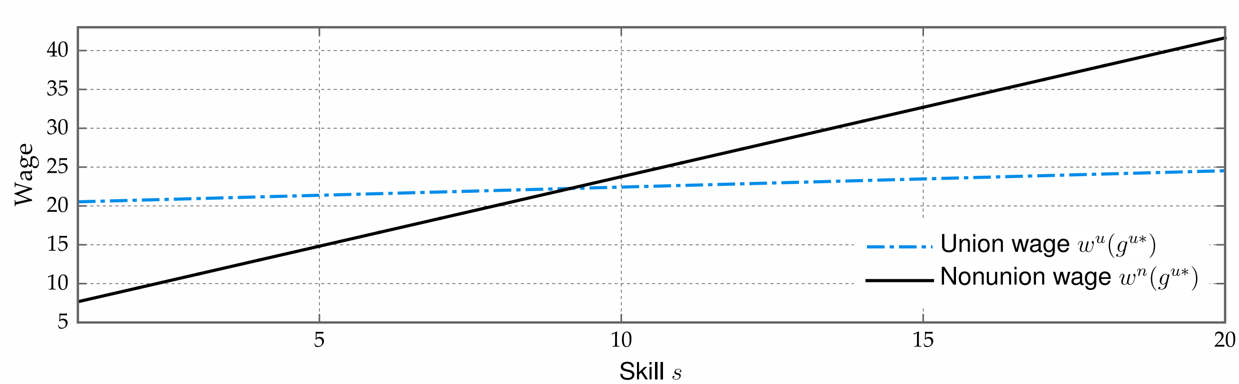
[20] Appendix A provides the details of these simulations. To emphasize only the behavior of the firm we keep the equilibrium variables θ and c fixed.



(a) Distributions of workers under exogenous union status



(b) Union and nonunion wages when the firm hires according to g^{n*}



(c) Union and nonunion wages when the firm hires according to g^{u*}

Figure 2: Example of the decision of an unconstrained firm (see appendix A for details)

The term related to the unionization constraint can be expanded to highlight the various mechanisms influencing the vote:

$$\frac{\partial V(g)}{\partial g_s} = \underbrace{\phi(\Delta_s(g)) - \frac{1}{2}}_{\text{(a)}} + \underbrace{g_s \frac{\partial \Delta_s(g)}{\partial g_s} \frac{\partial \phi(\Delta_s(g))}{\partial \Delta_s(g)}}_{\text{(b)}} + \underbrace{\sum_{s' \neq s} g_{s'} \frac{\partial \Delta_{s'}(g)}{\partial g_s} \frac{\partial \phi(\Delta_{s'}(g))}{\partial \Delta_{s'}(g)}}_{\text{(c)}} \tag{21}$$

where $\Delta_s(g)$ is short notation for the counterfactual union wage gap $w_s^u(g) - w_s^n(g)$ and where $\phi(\Delta_s)$ is, as before, the fraction of workers of skill s who vote for the union when the wage gap is Δ_s.

Each term in (21) represents one mechanism that the firm takes into account to prevent unionization.

(a) Fraction of voters for union As a fraction $\phi(\Delta_s)$ of workers of type s vote in favor of the union, adding an extra worker of type s directly increases the excess number of voters in favor of unionization by $\phi(\Delta_s) - 1/2$.

(b) Wages of workers in the same skill group Adding an extra worker of type s changes the union wage gap for all workers of type s. In particular, it lowers their marginal product which lowers their nonunion wage and makes a larger fraction of them vote in favor of unionization.

(c) Wages of workers in other skill groups Adding an extra worker of type s also changes the union wage gap for all workers of type $s' \neq s$. For instance, since union wages are determined by the average product, increasing the number of high-skill workers shifts the union wage schedule upward, leading some workers to change their vote in favor of unionization. Similarly, if the firm hires a lot of low-skill workers, their relatively low marginal product pushes the union wage schedule downward, which increases the number of workers against unionization.

In general, the problem of a firm constrained by the union vote must be solved numerically. By using simplifying assumptions, we can however derive some analytical results while keeping the main mechanisms of the model active.

Assumption 1. *The discount rate is $\gamma = 0$ and home production is $b_s = 0$. Also, there are only two types of workers, high-skill h and low-skill l, and the probabilities that vacancies are filled are such that $q(\theta_l) > q(\theta_h)$. Finally, in a union election, workers vote for the outcome that gives them the highest wage.*

Under these assumptions, agents behave as in a one-period economy and the heterogeneity across workers is limited to how expensive they are to hire. We also make the following simplifying assumptions about the firm.

Assumption 2. *The skill intensities are such that $z_h > z_l$ and the firm combines labor inputs using a Cobb-Douglas technology ($\sigma = 1$). Also, $z_h q(\theta_h) < z_l q(\theta_l)$ and the bargaining powers are such that $B_n > B_u$.*

These assumptions make the union threat active. When the inequality $B_n > B_u$ holds, the firm retains a higher share of output when it bargains individually with its workers and prefers to be union free. When the inequality $z_h q(\theta_h) < z_l q(\theta_l)$ holds, the firm would hire more low-skill than high-skill workers in an environment without the voting constraint, thereby giving low-skill workers the majority of the votes in an election. As these workers vote in favor of the union, the firm must distort its hiring decision, from g^{n*} to g^n, to prevent unionization. The firm does so by over-hiring high-skill workers, who vote against the union, and under-hiring low-skill workers, who vote in its favor. The following proposition details under what condition it decides to do so.

Proposition 5. *Under assumptions 1 and 2, the firm prefers to be union free but, if it hires without taking the vote into account, the majority of its workers would vote to form a union.*

Furthermore, if

$$\frac{B_n}{B_u} \geq \left(\frac{q(\theta_h)^{-1} + q(\theta_l)^{-1}}{\left(\frac{1}{z_l} q(\theta_l)^{-1}\right)^{z_l} \left(\frac{1}{z_h} q(\theta_h)^{-1}\right)^{z_h}} \right)^{\alpha} \tag{22}$$

and

$$\frac{\alpha \beta_n}{1-(1-\alpha)\beta_n} 2 z_h \geq \beta_u \tag{23}$$

it is optimal for the the firm to prevent unionization, otherwise it chooses to be unionized.

Equation 23 is a feasibility condition. If it is not satisfied, the firm cannot hire such that high-skill workers have the majority of the votes and also prefer to be in a union free firm.[21] Equation 22 is a profitability condition. Its left-hand side compares the share of the production surplus that the firm receives if it bargains individually with its workers to the share it receives under collective bargaining. It is a measure of the gain in profits associated with preventing unionization. The right-hand side of the equation measures the cost associated with preventing unionization. It depends on the amount of heterogeneity between workers, defined as $z_l q(\theta_l) - z_h q(\theta_h)$. When the heterogeneity is minimal, $z_l q(\theta_l) = z_h q(\theta_h)$, the right-hand side equals 1 and it is always profitable to prevent unionization. As heterogeneity grows larger, the right-hand side increases and preventing unionization becomes more costly.

To understand the link between worker heterogeneity and the cost of preventing unionization, consider the problem of a firm in which $z_l q(\theta_l)$ and $z_h q(\theta_h)$ are close to each other. In this case, without the voting constraint, the firm would hire a similar number of high-skill and low-skill workers. With the voting constraint, the firm therefore needs to hire only a few additional high-skill workers and a few less low-skill workers to prevent unionization. The distortion associated with this change is small and so is the cost of preventing unionization. As the gap between $z_l q(\theta_l)$ and $z_h q(\theta_h)$ becomes larger, so is the size of the distortion needed for the firm to win the union vote.

[21] The firm prevents unionization by hiring more high-skill workers and fewer low-skill workers. If inequality 23 does not hold, the firm will reach a point at which the nonunion wage of the high skill workers $w_s^n(g^n)$ is equal to their counterfactual union wage $w_s^u(g^n)$ while the low-skill workers still have a majority of the vote $g_l^n > g_h^n$. In this case, adding an extra high-skill worker would push their nonunion wage $w_s^n(g^n)$ under $w_s^u(g^n)$ and they would vote for the union. Removing a low-skill worker would push the union wage of the high skill workers $w_s^u(g^n)$ above their nonunion wage $w_s^n(g^n)$ and they would also vote for the union.

Figure 3 shows the implications of proposition 5 for the union status of the firm as a function of the bargaining powers β_u and β_n. For all β_u's under the thick black curve, assumption (2) is not satisfied and the firm prefers to be unionized. For the β_u's above this curve, the firm prefers to be union free but low-skill workers, who have the majority of the vote, prefer to be in a union firm. The firm must therefore modify its hiring policy in order for the union vote to fail. In the grey zone, the firm finds it optimal to do so (inequalities 22 and 23 are satisfied) and it hires according to g^n. In the white zone above the black curve, the firm prefers to let the union win; preventing unionization would be too costly. In this case, the firm hires according to g^{u*}.[22]

The following proposition describes the impact of the union threat on the firm.

Proposition 6. *Under assumptions 1 and 2, a binding voting constraint in a nonunion firm lowers profits, employment and output.*

In the grey zone of 3, the firm distorts the distribution of workers it hires to prevent unionization. This distortion hurts the marginal cost of producing one unit of goods which, because of decreasing returns, leads to a decline in output and employment.[23] Again, these declines are more important when the heterogeneity between workers $z_l q(\theta_l) - z_h q(\theta_h)$ is large.

The following proposition explains the impact of the union threat on wages.

Proposition 7. *Under assumptions 1 and 2, a binding voting constraint in a nonunion firm increases the average wage and decreases wage inequality, as defined by the ratio of the high-skill wage to the low-skill wage.*

As the firm reduces its size in response to the threat, the average marginal product of the workers increases which leads to a higher average wage. As the firm hires a higher ratio of high-skill to low-skill workers, the marginal product of high-skill workers falls relative to that of the low-skill workers. As a result, wage inequality decreases when a firm is subject to the union threat.

The results obtained under assumptions 1 and 2 hold more generally. To show this point, and to illustrate how the union threat affects firms that employ more than two types of workers, we now extend the numerical example of figure 2 to include the behavior of a constrained firm and show the results in figure 4. Once again, the firm compares the optimal distribution under which the workers unionize g^{u*} to the optimal distribution under which the workers reject the union g^n. These two distributions are plotted in panel (a) with, for comparison, the optimal nonunion distribution when the threat is absent g^{n*}. Panel (b) shows the wages that workers receive if they form a union or not, conditional on the firm hiring g^n. Notice that, as expected, low-skill workers vote in favor of the union while the opposite is true for high-skill workers. Panel (b) also depicts, for comparison, the nonunion wages $w^n(g^{n*})$ when there is no unionization constraint.

[22] In figure 3, the lower curve of the grey zone corresponds to inequality 22, while the upper curve corresponds to inequality 23.

[23] The voting constraint effectively increases the marginal cost of hiring a low-skill worker by a Lagrange multiplier and decreases the marginal cost of hiring a high-skill worker by the same multiplier. As a result of the concavity of the production function, the firm decreases the number of low-skill workers it employs by more than it increases the number of high-skill workers. Because of the complementarity between the two skill groups, the marginal product of high-skill workers goes down which leads to even less employment and less production.

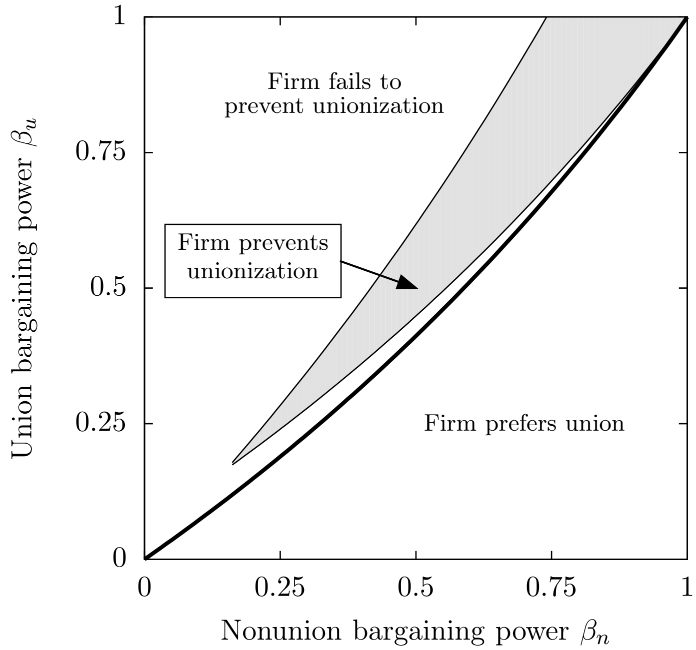
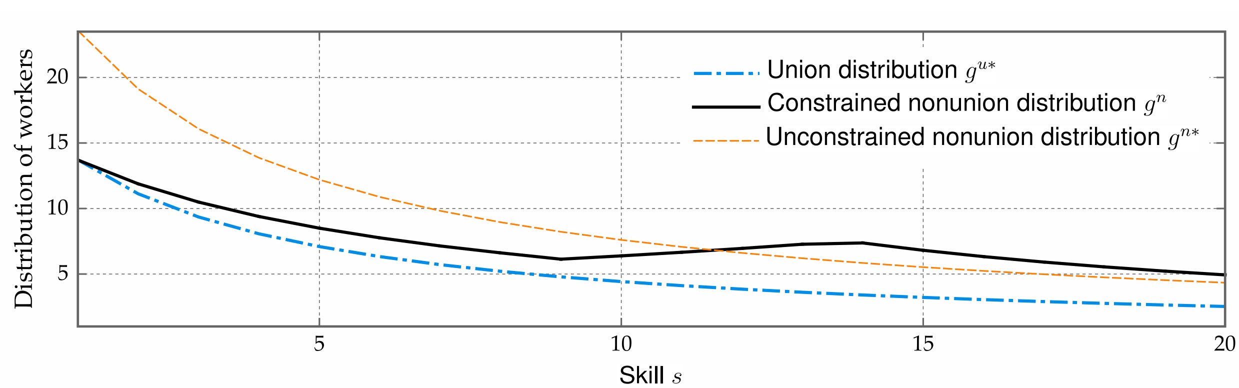
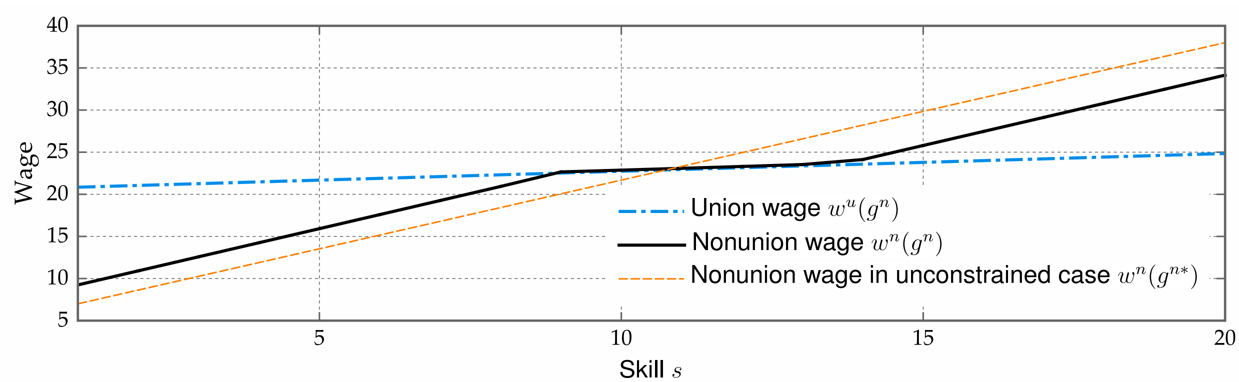


Figure 3: Firm status as a function of the bargaining powers



(a) Distributions of workers under endogenous union status



(b) Union and nonunion wages when the firm hires according to g^n

Figure 4: Example of the decision of a firm facing the union threat (see appendix A for details)

We see from panel (a) that, to prevent unionization, the firm distorts its hiring decision by hiring less low-skill workers and more high-skill workers. Doing so increases the marginal product of low-skill workers, which increases their wages. The opposite takes place at the top of the skill distribution and leads to lower wages for high-skill workers, as can be seen in panel (b). By reacting to the fact that its workers can unionize, this nonunion firm pays a more compressed wage distribution.[24]

As in proposition 6, the firm's output, employment and profit are lower under g^{n*} than under g^n. Also, the threat of unionization lowers the variance of wages and increases average wages, as in proposition 7.[25]

3.3 Impact of technology on unionization

When assumptions 1 and 2 hold, we can derive analytical results regarding the impact of technology on the union status of firms. In figure 5, we compare the union status of two firms, one with the same returns to labor α as in figure 3 and one with a smaller α. As α declines, the zone in which the firm prefers to be union free, in grey, moves down. Two mechanisms are at work. First, the asymmetry between individual and collective bargaining is more important when α is low. As explained after proposition 1, a small α makes individual bargaining more attractive to the firm. This effect pushes down the thick black curve and the lower curve of the grey zone (equation 22). Second, a smaller α lowers the share of the surplus that the workers receive when bargaining individually. As a result, the nonunion wage of high-skill workers can easily fall under their counterfactual union wage, which makes it hard for the firm to prevent unionization. This effect pushes down the upper curve of the grey zone (equation 23). We can see from the figure which firm is more likely to me unionized. For instance, if the bargaining powers are both in the middle of the parameter space, indicated by the $\star$ on the figure, a firm with a low α would be unionized while a firm with a high α would be union free. This prediction of the model is consistent with Hirsch and Berger (1984) who find that industries with lower labor shares are more unionized.

Figure 6 shows the impact of an increase in the intensity of the high-skill workers z_h on the union status of a firm. This increase lowers the heterogeneity $z_l q(\theta_l) - z_h q(\theta_h)$ between workers and makes it easier for the firm to prevent unionization. As a result, the firm finds it optimal to be union free for a larger range of β_u's and β_n's.

[24] We also see in panel (b) that $w^u(g^n)$ and $w^n(g^n)$ are very close to each other for workers with skills in the middle of the distribution. The firm would like to hire more of these workers: they vote against unionization and their relatively small marginal product has a small impact on the union wage schedule. However, hiring additional workers in this zone would lower their marginal product and push their nonunion wage under their union wage. They would therefore change their vote to support the union, which the firm wants to avoid.

[25] See appendix A for the details of these simulations.

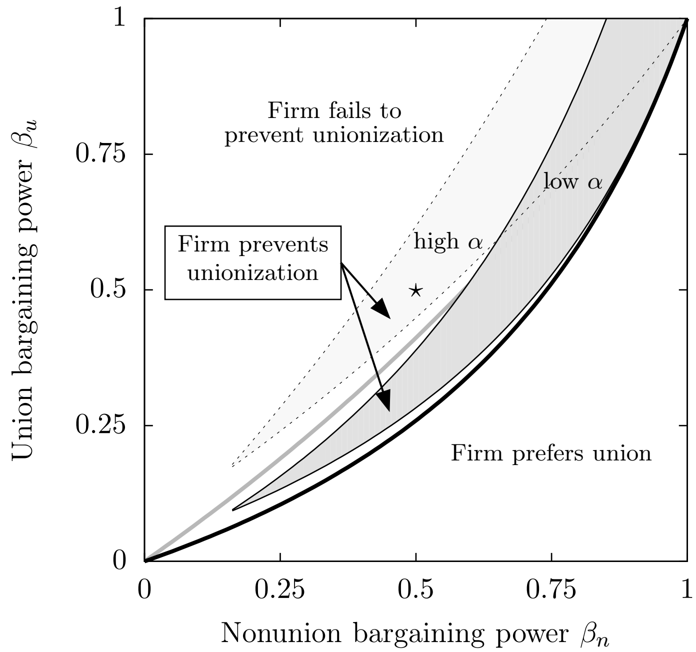


Figure 5: Impact of a decline in α on the status of a firm

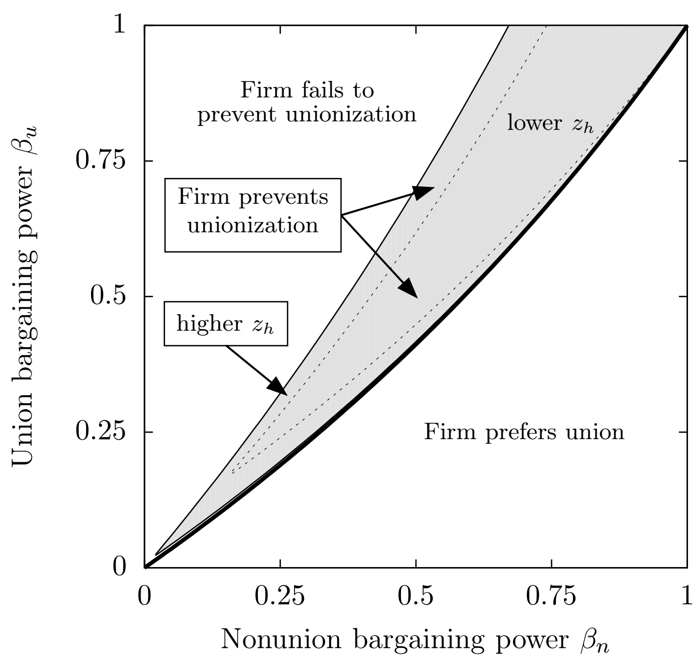


Figure 6: Impact of an increase in z_h on the status of a firm

4 Data and calibration

To evaluate the impact of the threat of unionization on the aggregate economy, we now use a simulated method of moments to calibrate the model to the private sector of the United States in 2005.

To reflect the typical duration of labor contracts the time period is set to one year.[26] All monetary amounts are measured in thousands of dollars. The discount rate is set to $\gamma = 0.95$ and the job destruction rate to $\delta = 0.113$, as in Pries and Rogerson (2005). For the matching function, I follow Krusell and Rudanko (2012) and use $m(U, V) = UV/(U + V)$ so that all probabilities are strictly between 0 and 1. The elasticity of substitution between skills is set to $\sigma = 1.5$ as in Krusell et al. (2000). For the cost of posting a vacancy, I follow the analysis of Silva and Toledo (2009) who find that training and vacancy posting costs amount to 69% of quarterly wages, which translates to $\kappa = 2.1$.[27] The number of skill groups is set to $S = 6$, which is enough to observe the impact of union policies across skills while keeping the computational complexity at a reasonable level. Finally, the random part of the workers preference for unionization ϵ is assumed to be a logistic random variable with mean 0 and scale parameter $1/\rho$, to be estimated. Table 1 summarizes the parameters. The details of the data work are in appendix B.

Parameter	Definition	Value	Source/reason
γ	Discount factor	0.95	5% annual interest rate
δ	Job destruction probability	0.113	Pries and Rogerson (2005)
σ	Skill elasticity of substitution	1.5	Krusell et al. (2000)
κ	Cost of posting a vacancy	2.1	Silva and Toledo (2009)
S	Number of skills	6	See text

Table 1: Parameters taken directly from the data or the literature

Each firm is endowed with one of two technologies, denoted by $j \in \{u, n\}$. In equilibrium, firms of type u are unionized, while firms of type n are not. The technologies of union and nonunion firms are (A_u, α_u, z_u) and (A_n, α_n, z_n), respectively. As lemma 4 in the appendix shows, it is equivalent to change the mass of firms of type j or their productivity A_j, so the mass of firms is normalized.

Skill distribution

In the theory, the skill index is only used to characterize some form of heterogeneity across workers. In the data, unions affect workers with different skills, as measured in various ways, very differently. For instance, the union wage gap and the preference for unionization vary in an important way with the skill of the workers. Because of this, the notion of skill is the relevant form of heterogeneity for the model.

[26] An alternative calibration using one month as the time period found similar effects of unions on the economy.

[27] Reasonable changes in parameter values have little influence on the impact of unions on the economy. In particular, I have experimented with a higher job destruction probability of $\delta = 0.4$ and lower vacancy costs equivalent to 14% of quarterly wage. The benchmark parameters offer the best fit of the model.

In the calibration, I first define a skill index from the data and then calibrate the technologies of the firms to match moments of the data. This way, the skill index and the technologies are consistently determined to make the model fit the distributions of workers.

I use data from the Merged Outgoing Rotation Groups of the Current Population Survey to build the skill index. To do so, I follow Card (1998) and regress log monthly *nonunion* wages on two types of variables. The first type includes variables intrinsic to each individual. The second type of variables depends on the job in which the individual currently works. I then use the predicted variable given by the OLS estimator of the individual characteristics alone as the skill index. Explicitly, denote by w_i the monthly wage of a worker i, who is working in industry $j(i)$. The regression is

$$\log w_i = \Lambda X_i^1 + \Psi X_{i,j(i)}^2 + \epsilon_i$$

and the skill index is given by the predicted values $\hat{s}_i = \exp\left(\hat{\Lambda} X_i^1\right)$. This way of constructing the index isolates the impact of variables intrinsically related to the individual from match-related factors that could also influence the wage. The individual characteristics X^1 are *age*, *education*, *occupation*, *race* and *sex*. The job-related characteristics X^2 are *industry* and the current *U.S. state* in which worker i lives.[28] Notice that even though the regression is run only on nonunion workers, the predicted values $\hat{s}_i$ are computed for *all* members of the labor force. The support of the distribution is split into six bins of equal size to generate the empirical skill distribution N_s.

Labor market tightness and the value of nonwork activities

In the United States, unemployment insurance programs are administered by the states and they vary considerably. In the model, b also takes into account home production and the value of the extra leisure provided by unemployment, two elements that are harder to quantify. Different numbers have been used in the literature. For instance, Hall and Milgrom (2008) use a flow value of non-work that equals 71% of productivity. Hagedorn and Manovskii (2008), on the other hand, use 95.5%. Because of the multi-worker production function used in this paper, setting b_s to be a certain fraction of productivity is inconvenient. Instead, I rely on the analysis of Hall (2009) and set b_s to be 85% of the average wage earned by workers of skill s.

Equation 16 is then used together with the observed unemployment rate in each skill bin to identify the labor market tightness schedule θ. Using the mean wage of union and nonunion workers together with the fact that, at the steady state, firms hire a fraction δ of their workforce every period, I compute the expected wage of a worker searching for a job. I then use equations 4 and 5 to identify the value of unemployment W^u for each of the skill bins.

[28]Including *U.S. state* as an individual characteristic instead has a minimal impact on the distribution. For *industry* and *occupation*, I use the variables generated by the NBER. Both are at the 3-digit level. I drop individuals with a skill index below the second percentile and above the 98th percentile.

Fraction of workers in favor of unionization

One of the moments that the estimation attempts to match is the fraction of workers of each skill bin voting in favor of unionization in each type of firm. Using data on workers who participated in union elections, Farber and Saks (1980) estimate a probit model to predict votes using worker and firm characteristics. I use their estimated coefficients to back out the fraction of workers of each skill voting for unionization.

Calibrated economy

To find the remaining parameters, the most important ones to quantify the threat of unionization, the estimation selects the vector $\xi = (\rho, \beta_n, \beta_u, \alpha_n, \alpha_u)$ to match: N_s, b_s, the labor shares in both union and nonunion firms, and the probability that a worker of type s votes for unionization in both union and nonunion firms. Appendix C explains the procedure.

The parameters are jointly estimated but we can provide some intuition about the moments of the data that matter the most in determining their value. Broadly speaking, the returns to labor parameters α_n and α_u are identified by the labor shares. The gap between the bargaining powers β_n and β_u is identified by the counterfactual union wage gaps in union and nonunion firms through their impact on the probability of voting for the union in both types of firms. The level of β_n and β_u is fitted to the level of wages through b_s.

Table 2 shows the parameter values that best fit the data. The estimation sets β_n to be larger than β_u. A few unmodeled features of the data that lower the bargaining position of the workers may explain this difference. First, there are costs to unionization: workers may have to pay dues or spend time organizing the union (Voos, 1983). The estimation captures these costs by lowering β_u. Second, consistent with evidence from Farber (1987), union workers might want the firm to hire more workers even if it leads to lower wages. In the model, since an increase in the bargaining power leads to higher wages and to lower employment, this preference would also be captured by a lower β_u. Finally, Bronfenbrenner (1994) and Freeman and Kleiner (1990) detail various tactics, some legal and some illegal, used by firms to prevent unionization. These tactics make it easier for firms to stay union free and they would also be captured by a low β_u. In the end, the estimation sets the distance between β_n and β_u to appropriately capture the magnitude of the unionization threat.[29]

Figure 7 shows the calibrated skill intensities z_u and z_n. We see that union firms rely relatively more on workers with average skills. This comes from the fact that, in the data, the distribution of union workers is concentrated in the middle of the skill distribution.

Figure 8 shows how the model fits the wage schedules, the distributions of workers employed by the firms, the distribution of workers in the labor force, the value of nonwork and the probability of voting for the union in both types of firm. As a consequence of the estimation strategy and of the way the firms' technologies are identified in the data, the model fits perfectly the distribution

[29] Appendix D provides an alternative calibration in which the two bargaining powers are equal. In this case, the impact of the union threat on the economy is much more important.

Parameter	Definition	Calibrated value
ρ	Preference for unionization parameter	1.06
β_n	Bargaining power of individual worker	0.44
β_u	Bargaining power of union	0.21
α_n	Returns to labor of nonunion firms	0.54
α_u	Returns to labor of union firms	0.49

Table 2: Estimated parameters

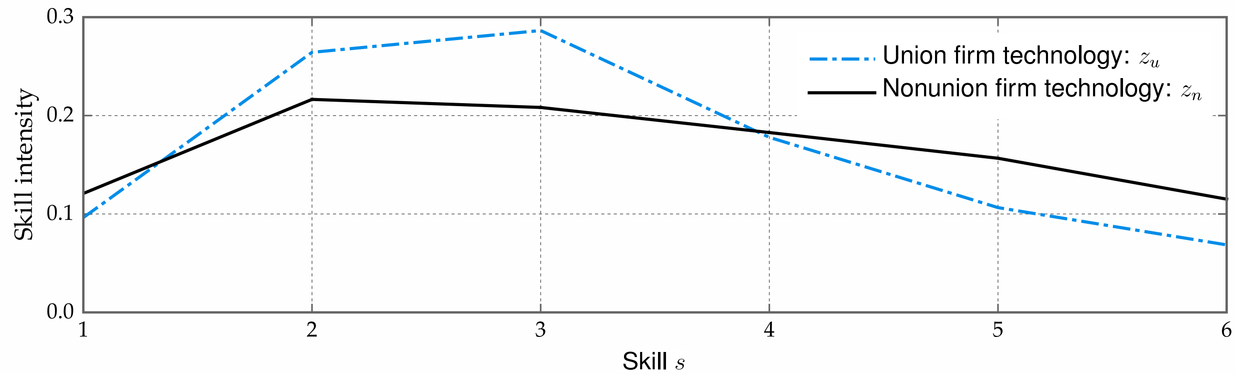


Figure 7: Calibrated skill intensities z_n and z_u

of workers in each firm, the unemployment rate and the distribution of workers in the labor force. The labor shares are fitted perfectly.[30] The model also fits the nonunion wage schedule quite well. The fit of the union wage schedule is, however, less precise, a consequence of the rigid structure imposed on union wages by equation 9. Union wages in the calibrated economy are more unequal than in the data; suggesting that the real equalizing effect of unions might be stronger than the one captured by the calibration.[31]

[30]The labor shares are 0.597 for the union firm and 0.613 for the nonunion firm. Because of the bargaining and the frictions, they differ from the parameters α_n and α_u.

[31]Since the search friction in the calibrated economy is quite small, the model is unable to quantitatively generate the full union wage gap, although its sign is fitted properly. As in Hornstein et al. (2011), the gap could be fitted better by lowering b. However, even though the observed wage gap is small in the calibrated economy, the quantity that matters for the unionization threat is the counterfactual union wage that workers expect to receive if they unionize. A proxy for this quantity enters the loss function, through the probability of voting for the union, and is fitted well.

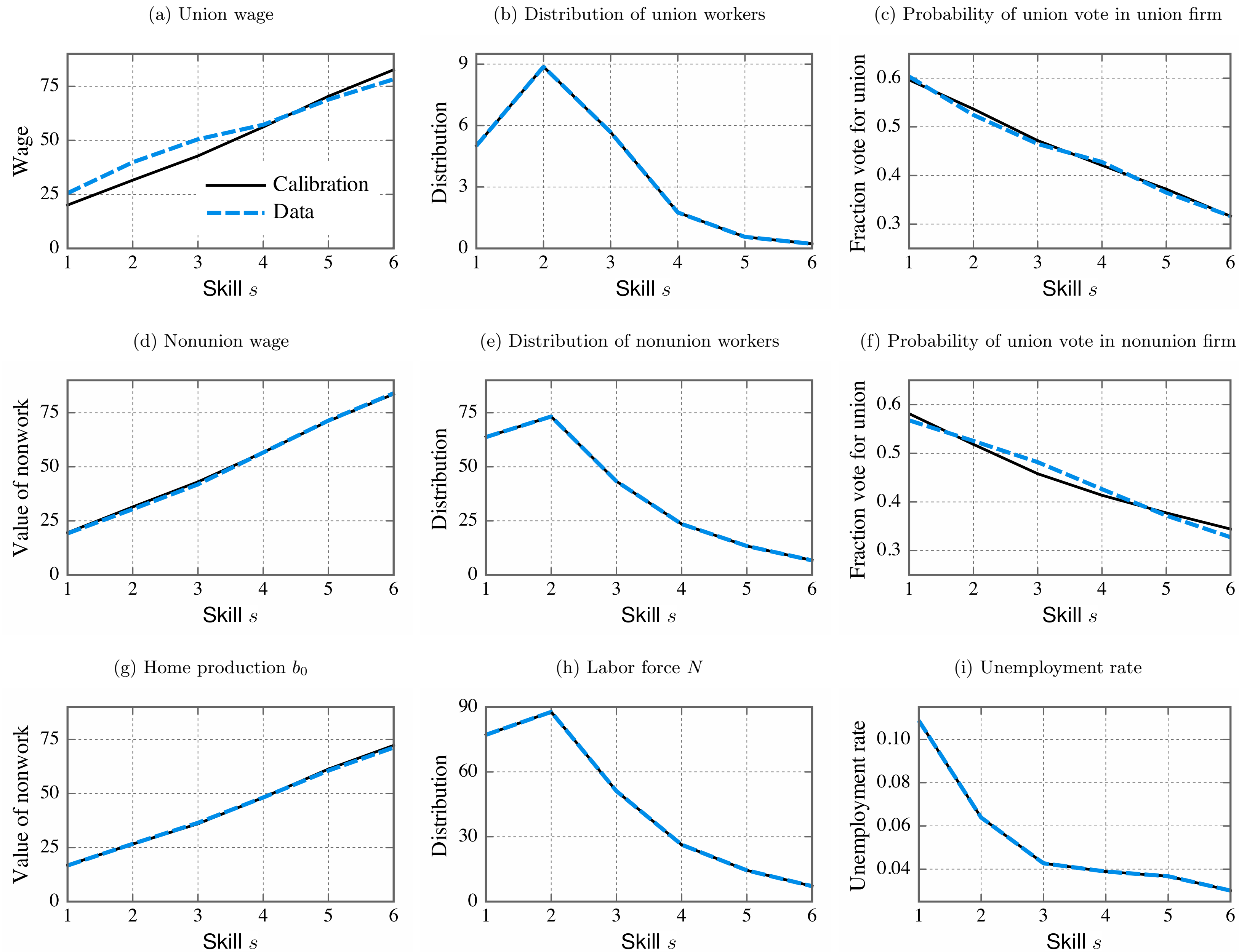


Figure 8: Fit of the calibrated model

5 Impact of unions on the economy

We now proceed to three policy experiments in general equilibrium.

1. Workers in nonunion firms cannot form a union anymore. Because of this, nonunion firms do not have to distort the distribution of workers they hire to prevent unionization. Union firms remain unionized.

2. Unions are illegal, so that all firms are union free.

3. Unions are mandatory, so that all firms are unionized.

The first experiment is designed to evaluate the impact of the union threat alone, without changing the union status of any firm. The second and third experiments evaluate how changes in union statuses influence the economy. Notice that in all these experiments the threat of unionization is inactive.

Figure 9 shows the new steady-state equilibria and how they differ from the calibrated economy. Consider first the nonunion wages shown in panel (a). We see that removing the threat of unionization leads to higher wage inequality by increasing high wages more than low ones. Since nonunion firms do not have to distort their hiring behavior to prevent unionization anymore, they hire fewer high-skill workers and more low-skill workers. Doing so changes the marginal product of each skill group, which influences wages. This change in wages also modifies the value of unemployment W_s^u, which then affects the wages paid by union firms, as shown by panel (b). The impact of the threat removal on union wages is purely through general equilibrium linkages.

As we can see on figure 9, the disappearance of the union threat also leads to an increase in the average wage and to a lower unemployment rate, as predicted by propositions 6 and 7.

Removing all unions, as in experiment 2, further amplifies the change in wages. In this case, in addition to the removal of the threat, the firms that were previously unionized switch to individual bargaining, which leads to a further increase in wage inequality.

In the last policy experiment, a large fraction of the firms, those that were previously union free, are now forced to be unionized. As can be seen in panel (b), this massive switch from individual to collective bargaining leads to a substantial increase in the wage of low-skill workers while at the same time lowering the wage of high-skill workers.

Table 3 presents the variance of wages, the unemployment rate and total output in the three policy experiments. We see that removing the threat increases the variance of wages, lowers unemployment and increases output. If, in addition to removing the threat, all union firms are forced to become union free, the variance of log wages goes up by an additional 3.8%, but output and unemployment are not further affected.

Notice that the impact of unions on aggregates is substantial even though only 9% of the workers are unionized in the calibrated economy. To understand why, it is useful to evaluate the importance of the two channels through which unions affect the economy: *i)* the threat of unionization changes the behavior of nonunion firms, and *ii)* the union status of firms itself, through the nature of the

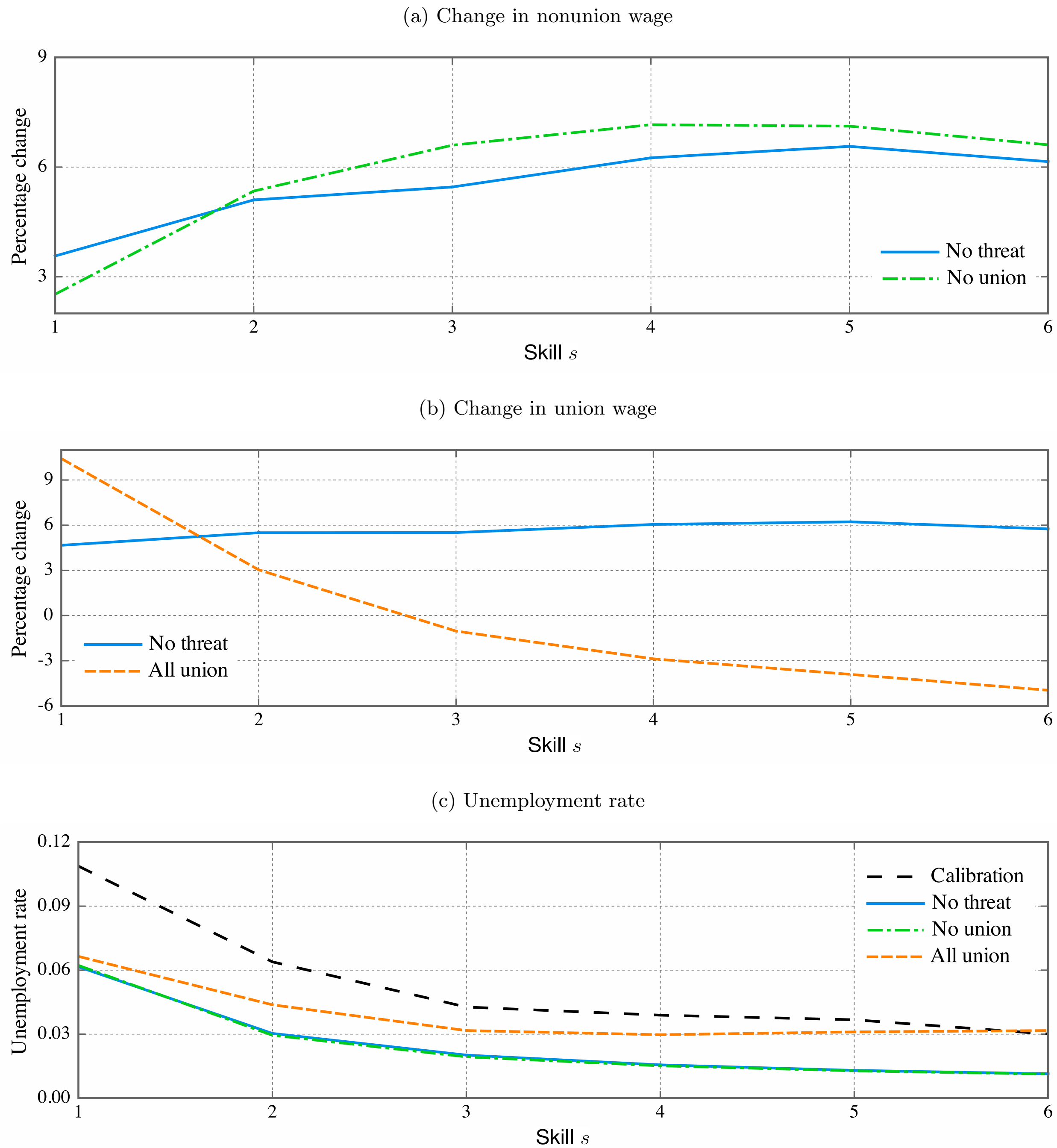


Figure 9: Impact of changes in union policies on wages and unemployment

bargaining, also affects the aggregates. From the results of table 3, we can conclude that the influence of the threat has the most impact. Indeed, output and unemployment react heavily to the removal of the threat (experiment 1) but are not further affected when, on top of that, the union status of firms is changed (experiment 2). The change in union status has, however, an additional effect on wage inequality.

Finally, in the last policy experiment, the switch from individual to collective bargaining pushes the variance of log wages down by 27%, a substantial reduction in inequality. Moreover, the removal of the voting constraint at the firm level is responsible for an increase in output and a decline in the unemployment rate. This shows once again that the threat of unionization, more so than the union status of firms, has a major negative influence on output and unemployment.[32]

	Calibration (level)	No threat	No unions	All unions
		(percentage change from calibration)		
Var(log wages)	0.18	+4.3%	+8.1%	-27.0%
Total output ($\times 10^9$)	1.5	+1.7%	+1.7%	+1.0%

	Calibration	No threat	No unions	All unions
		(all numbers are in levels)		
Unemployment rate	6.8%	3.5%	3.5%	4.5%
Unionization rate	9.0%	7.6%	0%	100%

Table 3: Impact of policy experiments on wages, unemployment and output

6 Conclusion

This paper proposes a general equilibrium theory of endogenous union formation to study the impact of unions on the economy. Unions are created by a majority vote within each firm. If a union is created, wages are bargained collectively otherwise each worker bargains his wage individually with the firm. This asymmetry in wage setting mechanisms causes unions to compress the wage distribution inside a firm and to lower its profit. A key mechanism in the theory is that, to prevent their own unionization, nonunion firms distort their hiring decisions in a way that also compresses wages and reduces employment and output. The main predictions of the theory are in line with facts documented in the empirical literature.

Policy experiments using an estimated version of the model show that removing the threat of unionization increases the variance of wages while also raising output and lowering unemployment. Outlawing unions completely amplifies these effects. Forcing all firms to be unionized, on the other hand, reduces wage inequality substantially while still improving output and unemployment.

For future research, the theory could be used to study the interaction between the rise in income inequality and the strong deunionization that has been observed in the United States during the last decades. In particular, the model could help us understand how a change in production technologies, or in the skill distribution, could impact the unionization rate and, through that channel, wage inequality. Finally, one could investigate empirically the impact of the union threat on firms. Changes in right-to-work legislations in the United States could be used to identify

[32]The search literature has documented that firms hire more under Stole and Zwiebel bargaining than under traditional Nash bargaining. From table 3, we see that this effect is of second-order importance here, when compared to the threat of unionization.

potential variations in wage distributions and employment in *nonunion* firms and therefore provide direct evidence of the impact of the threat on firm behavior.

References

Açıkgöz, O. T. and B. Kaymak (2014): "The Rising Skill Premium and Deunionization," *Journal of Monetary Economics*, 63, 37 – 50.

Acemoglu, D., P. Aghion, and G. L. Violante (2001): "Deunionization, technical change and inequality," *Carnegie-Rochester Conference Series on Public Policy*, 55, 229–264.

Acemoglu, D. and W. B. Hawkins (2014): "Search with multi-worker firms," *Theoretical Economics*, 9, 583–628.

Alvarez, F. and M. Veracierto (2000): "Labor-Market Policies in an Equilibrium Search Model," in *NBER Macroeconomics Annual 1999, Volume 14*, National Bureau of Economic Research, Inc, NBER Chapters, 265–316.

Bauer, C. J. and J. Lingens (2010): "Individual vs. Collective Bargaining in the Large Firm Search Model," *University of Munich Discussion Paper 11315.*

Bertola, G. and P. Garibaldi (2001): "Wages and the Size of Firms in Dynamic Matching Models," *Review of Economic Dynamics*, 4, 335 – 368.

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

Bronfenbrenner, K. (1994): "Employer Behavior in Certification Elections and First-Contract Campaigns: Implications for Labor Law Reform," in *Restoring the promise of American labor law*, ed. by S. Friedman, R. Hurd, R. Oswald, and R. Seeber, IRL Press, 75–89.

Cahuc, P., F. Marque, and E. Wasmer (2008): "A Theory of Wages and labor demand with intrafirm bargaining and matching frictions," *International Economic Review*, 48, 943–972.

Cahuc, P. and E. Wasmer (2001): "Labor Market Efficiency, Wages and Employment when Search Frictions Interact with Intrafirm Bargaining," *IZA Discussion Papers.*

Card, D. (1996): "The Effect of Unions on the Structure of Wages: A Longitudinal Analysis," *Econometrica*, 64, 957–979.

——— (1998): "Falling Union Membership and Rising Wage Inequality: What's the connection?" *National Bureau of Economic Research Working Paper Series*, No. 6520.

——— (2001): "The Effect of Unions on Wage Inequality in the U.S. Labor Market," *Industrial and Labor Relations Review*, 54, pp. 296–315.

Card, D., T. Lemieux, and W. Riddell (2004): "Unions and wage inequality," *Journal of Labor Research*, 25, 519–559.

Corneo, G. and C. Lucifora (1997): "Wage formation under union threat effects: Theory and empirical evidence," *Labour Economics*, 4, 265 – 292.

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

Dickens, W. T. (1983): "The Effect of Company Campaigns on Certification Elections: Law and Reality Once Again," *Industrial and Labor Relations Review*, 36, pp. 560–575.

DiNardo, J. and D. S. Lee (2004): "Economic Impacts of New Unionization on Private Sector Employers: 1984-2001," *Quarterly Journal of Economics*, 119, 1383–1441.

Dinardo, J. and T. Lemieux (1997): "Diverging Male Wage Inequality in the United States and Canada, 1981-1988: Do Institutions Explain the Difference?" *Industrial and Labor Relations Review*, 50, pp. 629–651.

Dinlersoz, E. M. and J. Greenwood (2012): "The Rise and Fall of Unions in the U.S." Working Paper 18079, National Bureau of Economic Research.

Ebell, M. and C. Haefke (2006): "Product Market Regulation and Endogenous Union Formation," IZA Discussion Papers 2222, Institute for the Study of Labor (IZA).

Elsby, M. W. L. and R. Michaels (2013): "Marginal Jobs, Heterogeneous Firms, and Unemployment Flows," *American Economic Journal: Macroeconomics*, 5, 1–48.

Farber, H. S. (1987): "The analysis of union behavior," in *Handbook of Labor Economics*, ed. by O. Ashenfelter and R. Layard, Elsevier, vol. 2 of *Handbook of Labor Economics*, chap. 18, 1039–1089.

Farber, H. S. and D. H. Saks (1980): "Why Workers Want Unions: The Role of Relative Wages and Job Characteristics," *The Journal of Political Economy*, 88, 349–369.

Frandsen, B. R. (2011): "Unionization and the Distribution of Employee Earnings: Evidence from Certification Elections," Working paper.

Freeman, R. B. (1980): "Unionism and the Dispersion of Wages," *Industrial and Labor Relations Review*, 34, 3–23.

Freeman, R. B. and M. M. Kleiner (1990): "Employer Behavior in the Face of Union Organizing Drives," *Industrial and Labor Relations Review*, 43, pp. 351–365.

Hagedorn, M. and I. Manovskii (2008): "The Cyclical Behavior of Equilibrium Unemployment and Vacancies Revisited," *American Economic Review*, 98, pp. 1692–1706.

HALL, R. E. (2009): "Reconciling Cyclical Movements in the Marginal Value of Time and the Marginal Product of Labor," *Journal of Political Economy*, 117, 281–323.

HALL, R. E. AND P. R. MILGROM (2008): "The Limited Influence of Unemployment on the Wage Bargain," *American Economic Review*, 98, 1653–1674.

HIRSCH, B. T. (2004): "What Do Unions Do for Economic Performance?" *Journal of Labor Research*, 25, 415–456.

HIRSCH, B. T. AND M. C. BERGER (1984): "Union Membership Determination and Industry Characteristics," *Southern Economic Journal*, 50, 665–679.

HIRSCH, B. T. AND D. A. MACPHERSON (2003): "Union Membership and Coverage Database from the Current Population Survey: Note," *Industrial and Labor Relations Review*, 56, 349–354.

HOLMES, T. J. (1998): "The Effect of State Policies on the Location of Manufacturing: Evidence from State Borders," *Journal of Political Economy*, 106, 667–705.

HORNSTEIN, A., P. KRUSELL, AND G. L. VIOLANTE (2011): "Frictional Wage Dispersion in Search Models: A Quantitative Assessment," *American Economic Review*, 101, 2873–98.

KLEINER, M. M. (2001): "Intensity of management resistance: understanding the decline of unionization in the private sector," *Journal of Labor Research*, 22, 519–540.

KRISHNA, V. AND R. SERRANO (1996): "Multilateral Bargaining," *Review of Economic Studies*, 63, 61–80.

KRUSELL, P., L. E. OHANIAN, J.-V. RÍOS-RULL, AND G. L. VIOLANTE (2000): "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis," *Econometrica*, 68, pp. 1029–1053.

KRUSELL, P. AND L. RUDANKO (2012): "Unions in a Frictional Labor Market," Working Paper 18218, National Bureau of Economic Research.

MATSA, D. A. (2010): "Capital Structure as a Strategic Variable: Evidence from Collective Bargaining," *Journal of Finance*, 65, 1197–1232.

NICKELL, S. AND R. LAYARD (1999): "Chapter 46 Labor market institutions and economic performance," in *Handbook of Labor Economic*, ed. by O. C. Ashenfelter and D. Card, Elsevier, vol. Volume 3, Part 3, 3029–3084.

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

PRIES, M. AND R. ROGERSON (2005): "Hiring Policies, Labor Market Institutions, and Labor Market Flows," *Journal of Political Economy*, 113, 811–839.

ROTH, A. E. (1979): *Axiomatic Models of Bargaining*, Springer-Verlag.

Silva, J. I. and M. Toledo (2009): "Labor Turnover Costs And The Cyclical Behavior Of Vacancies And Unemployment," *Macroeconomic Dynamics*, 13, 76–96.

Stole, L. A. and J. Zwiebel (1996a): "Intra-Firm Bargaining under Non-Binding Contracts," *Review of Economic Studies*, 63, 375–410.

——— (1996b): "Organizational Design and Technology Choice under Intrafirm Bargaining," *American Economic Review*, 86, 195–222.

Traxler, F. (1994): "Collective Bargaining: Levels and Coverage," in *Employment Outlook 1994*, OECD, 167–194.

Verna, A. (2005): "What Do Unions Do to the Workplace? Union Impact on Management and HRM Policies," *Journal of Labor Research*, 26, 415–449.

Voos, P. B. (1983): "Union Organizing: Costs and Benefits," *Industrial and Labor Relations Review*, 36, pp. 576–591.

A Parameters for the simulations in partial equilibrium

Table 4 contains the parameters used for the partial equilibrium simulations of figures 2 and 4 as well as for table 5.

Parameter	Interpretation	Value
S	Number of skills	20
δ	Probability of job destruction	0.05
γ	Discount rate	0.95
κ	Cost of posting a vacancy	3
β_n	Bargaining power of individual workers	$1/2$
β_u	Bargaining power of union workers	$1/2$
c_s	Outside option of workers	Linear from 1 to 5
θ_s	Labor market tightness	Linear from 1 to 10
ϕ	Shape of voting preference	$[1+\exp\{-50(w^u-w^n)\}]^{-1}$
A	Firm's total factor productivity	1000
α	Firm's return to scale parameter	0.7
z_s	Skill intensity	$1/S$
σ	Elasticity of substitution	1

Table 4: Parameters for the simulations

Table 5 shows some characteristics of this firm under the three following scenarios:

1. Exogenously unionized: The firm hires according to g^{u*}.

2. Endogenous union status: The firm compares profits under g^n and g^{u*}. It picks g^n.

3. Exogenously union free: The firm hires according to g^{n*}.

These scenarios can be thought of as policy environments in which unions are mandatory, allowed or forbidden.

	1. Exogenously unionized	2. Endogenous union status	3. Exogenously union free
Union status of the firm	Union	Nonunion	Nonunion
Firm discounted profit ($\times 10^4$)	0.9	1.4	1.5
Number of workers	108	151	185
Fraction of voters for union	66%	50%	66%
Mean of wages	22	20	18
Standard deviation of wages	1.2	7.1	9.4

Table 5: A firm under three union policies

First, let us compare the firm under scenarios 1 and 3. Consistent with propositions 1 and 2, profits are higher and wages are lower and more dispersed when the firm is union free. Also, the firm is larger in column 3 as it hires more workers to lower their marginal products, an important determinant of wages under individual bargaining.

In column 2, the firm is subject to the voting constraint. In this example, the best nonunion distribution of workers is more profitable than its best union counterpart and the firm is therefore union free. While the union status of the firm is the same in columns 2 and 3, the voting constraint has important implications for wages and the number of workers. First, it increases the cost of producing an extra unit of goods. Since the firm wants to equalize the marginal cost and the marginal revenue of production, the constraint pushes the firm to hire fewer workers to increase their marginal products. In turn, this increase in marginal products raises the average wage paid to the workers. Second, the constraint pushes the firm to hire more high-skill workers and fewer low-skill ones, which leads to a decline in wage dispersion. These results are consistent with propositions 6 and 7.

B Data appendix

CPS and BEA data

The data about individuals come from the Merged Outgoing Rotation Groups of the 2005 Current Population Survey (CPS) as it is made available by the National Bureau of Economic Research (NBER). I clean the sample by removing agricultural workers and individuals with an hourly wage higher than \$100 or lower than \$5. I also remove individuals younger than 16 or older than 65 or those who are out of the labor force. Finally, I remove public sector workers. Industry data come from the Bureau of Economic Analysis (BEA) Annual Industry Accounts available on the BEA's website.

Labor shares

I merge the data from the BEA and the CPS to compute a measure of the labor shares for union and nonunion firms. For each industry in the BEA dataset, I divide total workers' compensation by value added to get an estimate of the labor share in that industry. I then associate each worker in the CPS sample with the labor share of the industry in which he is currently working. I then average this variable separately over all union and nonunion workers and find a labor share of 0.597 for union firms and of 0.613 for nonunion firms.

Fraction of workers voting in favor of unionization by skill bin

I use the statistical analysis provided by Farber and Saks (1980). They use a dataset covering 29 union votes that took place between 1972 and September 1973 in 29 establishments. The establishments were located in Illinois, Indiana, Iowa, Missouri and Kentucky, and in manufacturing, transportation, wholesale trade, retail trade and services. The elections involved a total of 2788 workers. I used the regression results presented in their table 2 and then use their equation 9 to find the probability of voting for the union per skill bin. I set all coefficients to their estimated value and set all variables for which there is no equivalent in the model to their mean. I set all coefficients that are not statistically significant at the 95% level to zero.[33] Explicitly,

$$\begin{aligned}\text{Prob(Worker } s \text{votes for union)} &= \Phi((1+\alpha_1 T_{1s}+\alpha_2 T_{2s})(\gamma_0+\gamma_1\text{DEV}_s \\ &+\gamma_2\text{RDET}_s+\gamma_3\text{FIMP}_s+\gamma_4\text{PRO}_s+\delta X_s)+\gamma_5\text{DIFF}_s+\gamma_6(\text{DIFF}_s\times\text{DS}_s))\end{aligned} \tag{24}$$

where Φ is the CDF of a standard normal random variable and the other variables are defined by Farber and Saks.

The variable of interest to us is DEV_s, which is defined as[34]

$$\text{DEV}_s = \frac{W_s - \bar{W}}{\sigma}$$

where W_s is the wage of workers of skill s, $\bar{W}$ is the average wage in the firm they work at and σ is the standard deviation of wages in the firm.

Plugging in the values of their table 1 and table 2 into equation 24, we get

$$\text{Prob(Worker } s \text{votes for union)} = \Phi\left(-0.0865 - 0.161\times\text{DEV}_s - 0.042\right).$$

This last equation gives us an estimate of how workers would vote if they were in the real economy. I use this equation together with the CPS data to construct panels (c) and (f) in figure 8.

[33] Setting the coefficients that are not statistically significant to their estimated value has minimal impact on the final equation.

[34] The equation in Farber and Saks (1980) also includes a shift coefficient denoted λ. Since this coefficient is not statistically significant in their analysis, they are set to zero.

C Algorithm to calibrate the model

This appendix contains the detailed algorithm used for the calibration. The values of some parameters taken directly from the data or the literature are shown in table 1.

The labor market tightness vector θ is identified directly by using the steady state equation 16 and the unemployment rate for each skill group provided by the CPS. The outside option vector W^u is also fixed to its data value, as the main text explains. The calibration is constructed in such a way that the vectors θ and W^u provided by the data are the equilibrium object in the calibrated economy. To find the remaining parameters, the estimation selects the vector $\xi = (\rho, \beta_n, \beta_u, \alpha_n, \alpha_u)$ to minimize the following loss function:

$$\begin{aligned}\text{Loss} = &\frac{1}{S}\sum_s \frac{|N_s - \hat{N}_s|}{N_s} + \frac{1}{S}\sum_s \frac{|b_s - \hat{b}_s|}{b_s} + \frac{|\text{LS}_n - \widehat{\text{LS}}_n|}{\text{LS}_n} + \frac{|\text{LS}_u - \widehat{\text{LS}}_u|}{\text{LS}_u} \\ &+ \frac{1}{S}\sum_s \frac{|\phi_s^n - \hat{\phi}_s^n|}{\phi_s^n} + \frac{1}{S}\sum_s \frac{|\phi_s^u - \hat{\phi}_s^u|}{\phi_s^u}\end{aligned}$$

where LS_j denotes the labor share and where ϕ_s^j denotes the probability that a worker of type s votes for unionization while working in a firm of type j, for $j \in \{u, n\}$. In the previous equation, a hat indicates that the variable is from the simulated economy, while the absence of a hat indicates that the variable is taken from the data.

The calibration algorithm for a given vector of parameters ξ is:

1. Find the technologies:
   (a) Guess two technology schedules z^u and z^n for the firms
   (b) Find the schedules c^u and c^n:
      i. Guess the net outside option schedules c^u and c^n for the firms
      ii. Given these schedules compute the hiring decision of the union and nonunion firms and compute the wage schedules.
      iii. Update c^u and c^n using equation 7.
      iv. Measure the distance between the new c^u and c^n and the old ones.
      v. If there is convergence, stop, or else use the new schedules and go back to step 1(b)ii.
   (c) Use lemma 4 to set A_n and A_u to match the total size of the nonunion and union sectors. Verify if the hiring decisions of the firms coincide with the distributions of workers in the data. If not, use equation 20 to back out new guesses z^u and z^n and go back to step 1b.
2. See if any firm wants to deviate from the tentative equilibrium (for instance, the union firm has a higher profit if it fights the union). If so, discard the current parameter vector ξ.[35]

[35] When considering deviations, firms take the equilibrium net outside option schedule c_j as given. Firms are also forced to pay wages higher than the workers' outside option; otherwise, they would quit.

3. Use equation 16 to reverse engineer the skill distribution $\hat{N}$ that makes θ an equilibrium outcome.

4. Use equation 5 to reverse engineer the outside option schedule $\hat{b^0}$ that makes W^u an equilibrium outcome.

5. Compute the loss function at point ξ.

Notice that this algorithm exploits the fact that the decisions of the firms do not depend on N and b directly but only through θ and W^u.

D Calibration with equal bargaining powers and capital

D.1 Model

This appendix contains an alternative calibration of the model in which both bargaining powers are equal: $\beta_n = \beta_u \equiv \beta$. Under the assumptions of the benchmark calibration of Section 4, the union threat is so strong when $\beta_n = \beta_u$ that only firms with unrealistically high returns to labor α can be union free. It is however possible to reinterpret the model differently, by adding capital, to get a realistic calibration.

To see this, let us modify the production function by adding capital as an input:

$$F_j(g, K) = \tilde{A}_j \left(\left[\left(\sum_{s=1}^{S} z_{j,s} g_s^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \right]^{\tilde{\alpha}_j} K^{1-\tilde{\alpha}_j} \right)^{\omega_j}$$

where ω_j determines the returns to scale of firm j, $\tilde{A}_j$ is total factor productivity and $\tilde{\alpha}_j$ is the firm's labor intensity. Assuming that there is a constant rental rate of capital r, the problem of the firm is now

$$J(g_{-1}) = \max_{g_s, K} F(g, K) - \sum_s w_s(g) g_s - rK - \kappa \sum_{s=1}^{S} \frac{g_s - g_{-1,s}(1-\delta)}{q(\theta_s)} + \gamma J(g).$$

Substituting the optimal level of capital back into the problem, the firm maximizes

$$\begin{aligned} J(g_{-1}) = \max_{g_s, K} \tilde{A}_j \left(1 - \omega_j(1-\tilde{\alpha}_j)\right) \left(\frac{\omega_j \tilde{A}(1-\tilde{\alpha}_j)}{r} \right)^{\frac{\left(1-\tilde{\alpha}_j\right)\omega_j}{1-\omega_j(1-\tilde{\alpha}_j)}} \left(\left(\sum_{s=1}^{S} z_{j,s} g_s^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \right)^{\frac{\tilde{\alpha}_j \omega_j}{1-\omega_j(1-\tilde{\alpha}_j)}} \\ - \sum_s w_s(g) g_s - \kappa \sum_{s=1}^{S} \frac{g_s - g_{-1,s}(1-\delta)}{q(\theta_s)} + \gamma J(g). \end{aligned}$$

By defining

$$A_j = \tilde{A}_j \left(1 - \omega_j(1-\tilde{\alpha}_j)\right) \left(\frac{\omega_j \tilde{A}(1-\tilde{\alpha}_j)}{r} \right)^{\frac{\left(1-\tilde{\alpha}_j\right)\omega_j}{1-\omega_j(1-\tilde{\alpha}_j)}}$$

and

$$\alpha_j = \frac{\tilde{\alpha}_j \omega_j}{1-\omega_j(1-\tilde{\alpha}_j)}$$

we see that the problem of the firm is exactly as in the benchmark model but with a different interpretation of the parameters. In particular, α_j now captures the labor intensity of the firm $\tilde{\alpha}_j$ as well as its returns to scale ω_j. This alternative model can be realistically calibrated when bargaining powers are equal.[36]

D.2 Calibration

This alternative calibration broadly follows the one of Section 4 but uses different values for a few parameters to improve the fit.[37] The time period is still a year. Since the average duration of employment is 2.5 years in the data, I follow Krusell and Rudanko (2012) and set the job destruction rate to $\delta = 0.4$. I take the vacancy posting cost of Silva and Toledo (2009) and set $\kappa = 0.17$. All the other parameters are as in the benchmark calibration. The value of nonwork b_s is still set to be 85% of the average wage earned by workers of skill s and the distribution of skills in the population is the same as in the benchmark calibration. Table 6 summarizes the parameters.

Parameter	Definition	Value	Source/reason
γ	Discount factor	0.95	5% annual interest rate
δ	Job destruction probability	0.4	Krusell and Rudanko (2012)
σ	Skill elasticity of substitution	1.5	Krusell et al. (2000)
κ	Cost of posting a vacancy	0.17	Silva and Toledo (2009)
S	Number of skills	6	See Section 4

Table 6: Parameters taken directly from the data or the literature

This calibration follows the same algorithm as the benchmark one but, since this model has fewer degrees of freedom, the loss function contains fewer moments. In particular, the calibration

[36] We are implicitly assuming that the firm can rent capital in a perfectly flexible way. In particular, if the bargaining with the workers breaks down, the firm can immediately adjust its capital stock.

[37] Using the benchmark parameters has little influence on the impact of unions on the economy but provides a worse fit to the data.

looks for the vector $\xi = (\rho, \beta, \alpha_n, \alpha_u)$ that minimizes[38]

$$\text{Loss} = \frac{1}{S}\sum_s \frac{|N_s - \hat{N}_s|}{N_s} + \frac{1}{S}\sum_s \frac{|b_s - \hat{b}_s|}{b_s} + \frac{|\text{LS}_n - \widehat{\text{LS}}_n|}{\text{LS}_n} + \frac{|\text{LS}_u - \widehat{\text{LS}}_u|}{\text{LS}_u}.$$

Parameter	Definition	Calibrated value
ρ	Preference for unionization parameter	0.46
β	Bargaining power	0.19
α_n	Returns to labor of nonunion firms	0.95
α_u	Returns to labor of union firms	0.51

Table 7: Calibrated parameters

Table 7 shows the parameter values that minimize the loss function. We see that the unique bargaining power β is similar to the bargaining power of union workers in the benchmark calibration. The parameter α_u is also similar to its value from section 4. The value of α_n needs, however, to be bigger for the nonunion firm to be able to prevent unionization. This gap between α_u and α_n can come from differences in returns to scale ω or in labor intensities $\tilde{\alpha}$ between the technologies of the union and nonunion firms.[39] Finally, figure 10 shows the calibrated skill intensities z_u and z_n. They are similar to those of the benchmark calibration.

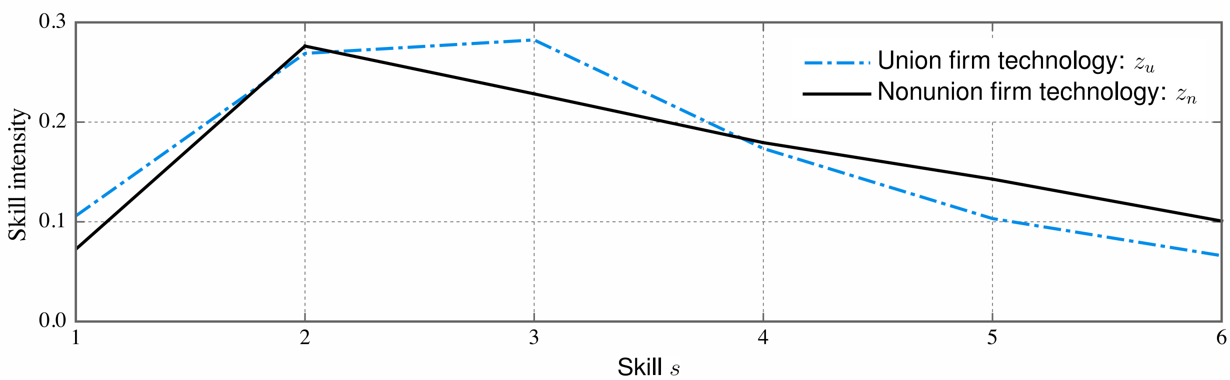


Figure 10: Calibrated skill intensities z_n and z_u

Figure 11 shows how the model fits the wage schedules, the distributions of workers employed by the firms, the distribution of workers in the labor force N, the value of nonwork b and the probability of voting for the union in both types of firm. As in the benchmark calibration, the model fits perfectly the distribution of workers in each firm, the unemployment rate and the distribution of workers in the labor force. The labor shares are also fitted very well.[40] The model also fits the

[38]This calibration is unable to properly fit the fractions of voter in favor of unionization, so I dropped these moments from the loss function. Alternatively, one could add them back and drop the labor shares. Since the labor shares, through the α's, play an important role in the theory this loss function seems more appropriate.

[39]Hirsch and Berger (1984) show that unionized firms have a lower labor share than nonunion firms.

[40]The labor share for the union firm is 0.597 in the calibrated economy and in the data. For the nonunion firm, the data and calibrated labor shares are both 0.613.

nonunion wage schedule quite well. The fit of the union wage schedule is less precise but it is better than in the benchmark calibration. Since the loss function does not target the fractions of workers for unionization, the fit on these dimensions is obviously worse than in the benchmark calibration. In particular, we see in panel (c) that union workers strongly favor unionization. The fraction of nonunion workers who vote for unionization, the important quantity for the union threat, is fitted better by the calibration, as seen in Panel (f).

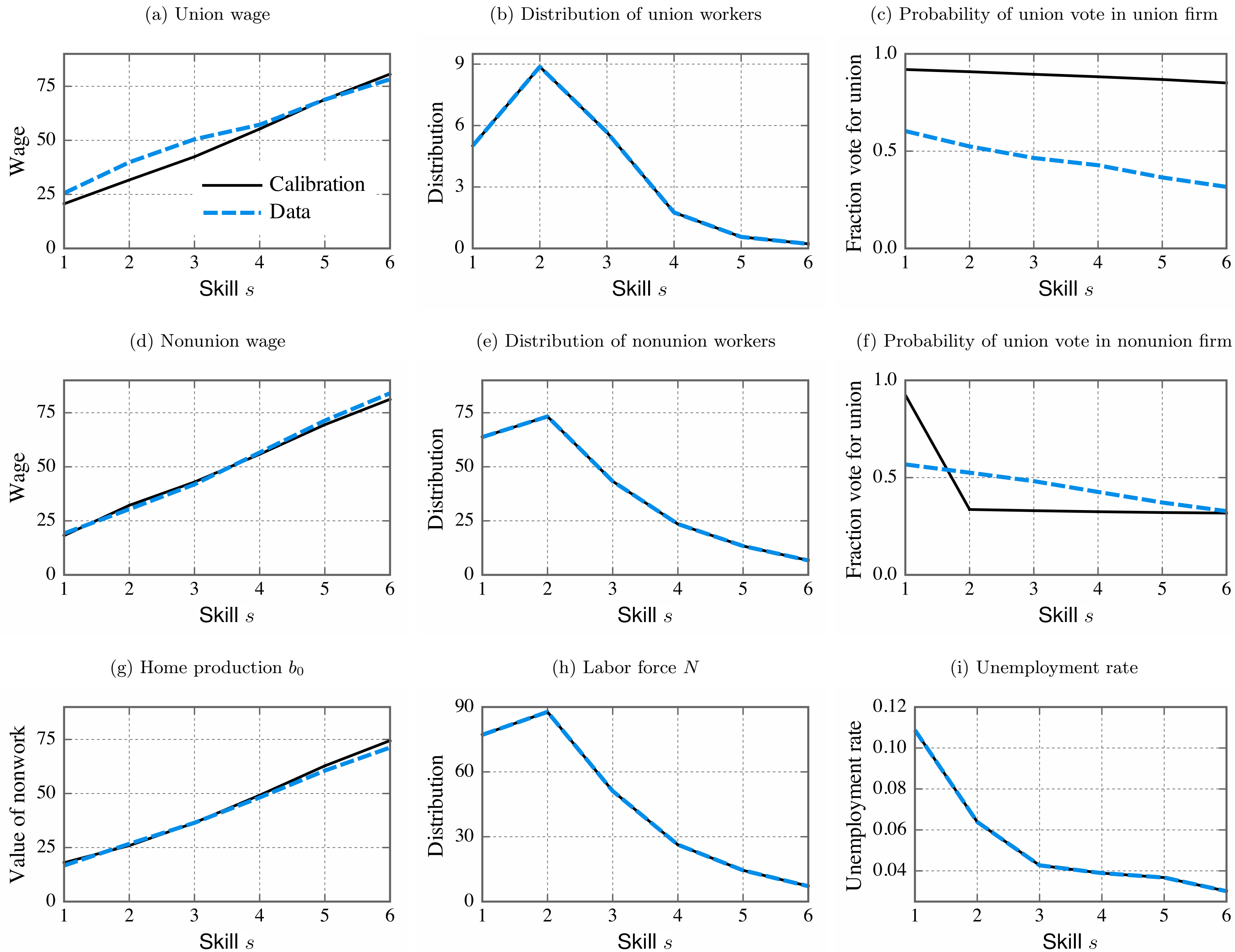


Figure 11: Fit of the calibrated model

D.3 Policy simulations

We now proceed to the three policy exercises detailed in Section 5. Figure 12 shows the new equilibria and how they differ from the calibrated economy. As in the benchmark calibration, removing the threat of unionization raises all wages and lowers unemployment.The effects are similar when there are no unions. Once again we see that the bulk of the impact of unions on the economy comes from the threat alone and not from the fact that some firms are actually unionized. In the last policy exercise, in which all firms are unionized, all wages go up, particularly so for low-skill workers. Unemployment also goes down and all workers gain from the change in policy.

Table 8 presents the variance of wages, the unemployment rate and total output under the three policies. As in the benchmark case, removing the union threat or outlawing unions leads to a large increase in output and in the variance of log wages. The unemployment rate also falls substantially. Unionizing all firms leads once again to lower wage inequality and to an increase in output. Interestingly, here the 'all union' exercise leads to a higher level of output and to a lower unemployment rate than the two other policy exercises. Once again, removing the threat of unionization is the main reason for the increase in output and the lower unemployment. The union status of firms does not matter much.

	Calibration (level)	No threat (percentage change from calibration)	No unions	All unions
Var(log wages)	0.19	+11.7%	+19.7%	-47.1%
Total output ($\times 10^9$)	1.47	+6.3%	+6.4%	+6.7%

	Calibration	No threat (all numbers are in levels)	No unions	All unions
Unemployment rate	6.8%	2.9%	2.9%	2.0%
Unionization rate	9.0%	3.9%	0%	100%

Table 8: Impact of policies on wages, unemployment and output

Notice that the impact of the union threat is stronger here than in the benchmark calibration of Section 4. Since, when $\beta_n > \beta_u$, the workers know that if they vote for the union their bargaining power will be lower, it is easier for firms to prevent unionization. Here, the workers have a lot of bargaining power if they unionize, and the firms must heavily distort their behavior to prevent unionization.

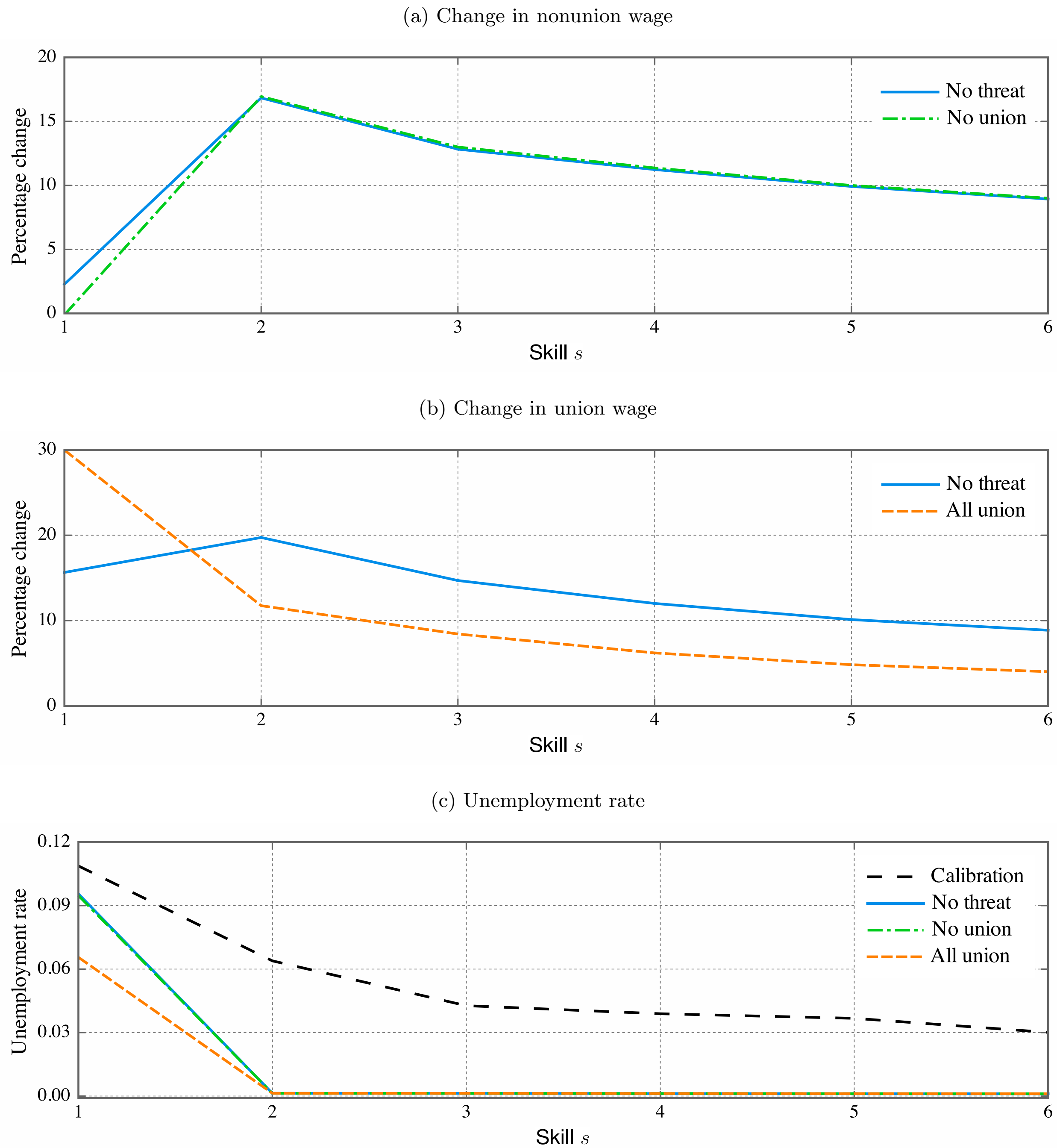


Figure 12: Impact of changes in union policies on wages and unemployment

E Alternative procedures for the collective bargaining

This appendix solves two alternative bargaining procedures for the union wages. I introduce a 'union' as an intermediary between the workers and the firm. Notice that this is different from the benchmark model in which union wages are determined by an n-person bargaining between the firm and the workers. I keep the framework as simple as possible to make the exposition transparent.

In particular, I assume that all jobs are destroyed at the end of the period and that all reservation values are zero. It is straightforward to add them back.

The bargaining now takes place in two stages. In the second stage, a union bargains with the firm on how to split the surplus generated by production. In the first stage, the workers decide on how to split their share of the surplus.

Suppose that, in this second stage, the firm bargains with a risk-neutral union. The bargaining problem is

$$\max_{T} T^{\beta} \left(F(g) - T\right)^{1-\beta}$$

where T is the transfer between the firm and the union. The solution is the standard outcome of Nash-bargaining: the union keeps a fraction β of the joint surplus $T(g) = \beta F(g)$.

In stage 1, the workers have to split this surplus among them. There are two ways to model this. We can assume that each worker enters a one-on-one negotiation with the union or we can assume that there is collective bargaining between the workers and the union. I explore both of these cases.

One-on-one negotiation between the workers and the union

In this first case, both parties know that if the worker walks away the union will extract a smaller amount from the firm in the second stage. Let us assume that all workers have the same bargaining power and that they bargain with a union leader who captures what's left of the surplus. The surplus of a worker from agreeing to stay in the union is, under our assumptions, simply w_s. The surplus of the union leader is

$$\frac{\partial T(g)}{\partial g_s} - w_s(g) - \sum_k g_k \frac{\partial w_k(g)}{\partial g_s}$$

where we see that the union internalizes the fact that, if this worker walks away, all the other negotiated wages may change. This is the Stole and Zwiebel (1996a) bargaining. This problem is very similar to the one encountered with the nonunion individual bargaining. We need to solve the following system of equations:

$$\frac{\partial T(g)}{\partial g_s} - w_s(g) - \sum_k g_k \frac{\partial w_k(g)}{\partial g_s} = \frac{1-\epsilon}{\epsilon} w_s$$

where ϵ is the bargaining power of the workers. The solution to this system is very similar to the one of the individual bargaining (I'm using the same boundary conditions):

$$w_s = \frac{\epsilon}{1-\epsilon(1-\alpha)} \frac{\beta \alpha z_s}{g_s^{1/\sigma}} A \left(\sum_k z_k g_k^{\frac{\sigma-1}{\sigma}} \right)^{\frac{1-\sigma(1-\alpha)}{\sigma-1}}$$

We see that, under these assumptions, this new union wage schedule is structurally identical to nonunion wages. In particular, inequality in log-wages is the same whether the firm is unionized or not, for a given distribution of workers. Therefore, this model does not seem appropriate to discuss the union-generated wage compression observed in the data (Frandsen, 2011). Furthermore, if we assume that the union leader gets a negligible share of the surplus ($\epsilon \to 1$), all workers would always vote for in favor of unionization. There would be no union free firms in the economy.

Collective bargaining between all workers and the union

We now consider the other way of splitting the surplus extracted from the firm. Here all workers and the union negotiate in a single bargaining session. Under our assumptions, this problem is

$$\max_w \left(\prod_s (w_s)^{\frac{g_s}{n}}\right)^{\epsilon} \left(T(g) - \sum_s w_s g_s\right)^{1-\epsilon}$$

where again $0 < \epsilon < 1$ denotes the bargaining power of the workers. This problem is very similar to the actual union bargaining solved in the core of the paper except that we are using the union surplus $T(g)$ instead of the production function $F(g)$. But, since the second stage bargaining yielded $T(g) = \beta F(g)$, this difference is minimal. In fact, we find that the wage with this procedure is

$$w_s(g) = \frac{\epsilon}{n} \beta F(g).$$

If we send the bargaining of the workers with the union to $\epsilon \to 1$, we find exactly the union wage equation from the core of the text (if we impose the same simplifying assumptions there too). This goes to show that adding an actual union as an intermediary between the workers and the firms does not affect the structure of wages.

F Proofs

This appendix contains the proofs from the previous sections.

Lemma 1. *In a steady-state equilibrium, the firm's dynamic problem is equivalent to*

$$\max_g \pi(g) - \kappa \sum_{s=1}^{S} \frac{g_s}{q(\theta_s)} + \kappa(1-\delta)\gamma \sum_{s=1}^{S} \frac{g_s}{q(\theta_s)} \tag{25}$$

Proof. First, the constraint $v_s \geq 0$ are never binding in a steady-state equilibrium. To see why, suppose that in such an equilibrium a firm's optimal distribution of workers is given by g_s^*. Two events might move the firm away from g_s^*. First, every period, it loses a fraction δ of its workers. Second, if one of the wage bargaining sessions breaks down without an agreement, the firm loses additional workers.[41] In both of these cases, the firm has to hire a positive number of workers in

[41] This does not happen in equilibrium but the value function needs to be defined along these paths to correctly

the next period to replace those that have been lost. Therefore, $v_s > 0$ in all markets s such that $g_s^* > 0$ and $v_s = 0$ elsewhere. The firm's problem is

$$J\left(g_{-1}\right) = (1-\delta)\kappa \sum_{s=1}^{S} \frac{g_{-1,s}}{q(\theta_s)} + \max_{g} \left\{ \pi(g) - \kappa \sum_{s=1}^{S} \frac{g_s}{q(\theta_s)} + \gamma J\left(g\right) \right\}.$$

The term that is maximized is constant with respect to g_{-1}. Denote that constant by B. Then, in particular

$$J\left(g\right) = (1-\delta)\kappa \sum_{s=1}^{S} \frac{g_s}{q(\theta_s)} + B.$$

and the firm solves

$$\max_{g} \pi(g) - \kappa \sum_{s=1}^{S} \frac{g_s}{q(\theta_s)} + \gamma \left((1-\delta)\kappa \sum_{s=1}^{S} \frac{g_s}{q(\theta_s)} + B \right)$$

which is the result. □

Lemma 2. *If all the workers have the same bargaining power, and the firm has bargaining power $1-\beta_u$, the collective Nash bargaining problem can be written as*

$$\max_{w} \left[\prod_{s=1}^{S} \left(W_s^e(w) - b_s - \gamma W_s^u \right)^{\frac{g_s}{n}} \right]^{\beta_u} \left[F(g) - \sum_{s=1}^{S} w_s g_s + (1-\delta)\kappa\gamma \sum_{s=1}^{S} \frac{g_s}{q(\theta_s)} \right]^{1-\beta_u}$$

where $n = \sum g_s$ is the total number of employed workers. Furthermore, the wage equation

$$w_s^u(g) - c_s = \frac{\beta_u}{n} \left(F(g) - \sum_{k=1}^{S} c_k g_k + \gamma(1-\delta)\kappa \sum_{k=1}^{S} \frac{g_k}{q(\theta_k)} \right)$$

solves this bargaining problem.

Proof. Axiomatic bargaining theory (Roth, 1979) (see also Krishna and Serrano (1996)) tells us that the solution to an n-players bargaining problem is the payoff that maximizes the geometric average of the n surpluses and where the average weights can be interpreted as bargaining powers. We therefore look at the surpluses of each player and then compute this average.

Consider the firm's surplus from agreeing on a wage schedule w. At this point, the distribution g is fixed and the hiring cost is sunk. In a steady state, the difference in discounted profits for the firm, denoted by $\Delta^u(w)$, is

$$\Delta^u(w) = \left[\pi(g, w) + \gamma J\left(g\right)\right] - \left[\pi(0) + \gamma J(0)\right] \tag{26}$$

specify the bargaining problems.

where the first term between brackets is discounted profits if an agreement is reached and $\pi(0) + \gamma J(0)$ is the firm's discounted profit if negotiations break down. In such a case, the firm has no workers, it produces nothing and pays no wages. Therefore, $\pi(0) = 0$. $J(0)$ is the value function of a firm that starts the period with no workers. Because of risk-neutrality and the linear vacancy cost, it hires back to its steady state optimal level g^* right away (we have seen in the proof of lemma 1 that, at a steady state, g does not depend on g_{-1}). Therefore,

$$J(0) = \pi(g^*, w^*) - \kappa \sum_{s=1}^{S} \frac{g_s^*}{q(\theta_s)} + \gamma J\left(g^*\right)$$

where w^* is the equilibrium wage schedule for this firm. We can therefore rewrite equation 26 as

$$\Delta^u(w) = \pi(g, w) + \gamma J\left(g\right) - \gamma \left(\pi(g^*, w^*) - \kappa \sum_{s=1}^{S} \frac{g_s^*}{q(\theta_s)} + \gamma J\left(g^*\right) \right).$$

But the firm's value function is

$$J(g) = \pi(g^*, w^*) - \kappa \sum_{s=1}^{S} \frac{g_s^* - (1-\delta) g_s}{q(\theta_s)} + \gamma J(g^*) \tag{27}$$

and therefore the firm's surplus from agreeing on a wage w is

$$\Delta^u(w) = \pi(g, w) + (1-\delta)\gamma\kappa \sum_{s=1}^{S} \frac{g_s}{q(\theta_s)}.$$

On the workers' side, the net benefit of an agreement is $W_s^e(w) - b_s - \gamma W_s^u$. Assume now that all workers have the same bargaining power and consider the discrete case in which there are $h_s \in \mathbb{N}$ workers of type s who all have mass $\epsilon > 0$ such that $h_s \times \epsilon \to g_s$ as we move to the continuum. The bargaining problem with *equal* bargaining power is

$$(W_1^e - b_1 - \gamma W_1^u)^{h_1} \times \cdots \times (W_i^e - b_i - \gamma W_i^u)^{h_i} \times \cdots \times (W_S^e - b_S - \gamma W_S^u)^{h_S} \times \Delta^u.$$

Since the bargaining power of the firm is $1 - \beta_u$ and that bargaining powers must sum to 1 we get

$$(W_1^e - b_1 - \gamma W_1^u)^{\frac{\beta_u \epsilon h_1}{\epsilon H}} \times \cdots \times (W_S^e - b_S - \gamma W_S^u)^{\frac{\beta_u \epsilon h_S}{\epsilon H}} \times (\Delta^u)^{1-\beta_u}$$

where $H = \sum_s h_s$. Taking the limit to the continuum, $\frac{\epsilon h_i}{\epsilon H} \to \frac{g_i}{n}$ for all i and we find equation 8.

When the surplus from the match is positive, the bargaining problem is defined on a convex set and is strictly concave. First-order conditions are therefore necessary and sufficient and yield the wage equation 9. $\square$

Lemma 3. *The wage schedule*

$$w_s^n(g) - c_s = \frac{\beta_n}{1-(1-\alpha)\beta_n} \frac{\alpha z_s}{g_s^{1/\sigma}} A \left(\sum_{k=1}^{S} z_k g_k^{\frac{\sigma-1}{\sigma}} \right)^{\frac{1-\sigma(1-\alpha)}{\sigma-1}} - \beta_n c_s + \beta_n \gamma (1-\delta) \frac{\kappa}{q(\theta_s)} \tag{28}$$

solves the bargaining problem of a nonunion firm.

Proof. The Stole and Zwiebel (1996a,b) solution to the bargaining problem is the wage function that gives the worker a share β_n of the joint surplus. The bargaining takes place when all vacancies have been posted and the vacancy costs are therefore sunk. When bargaining with a single worker, the firm compares two scenarios. Either an agreement is reached, in which case production takes place as planned, or the negotiations break down and the firm produces without this individual worker. In this last case, that worker departs from the firm and additional vacancies will have to be posted in the next period for the firm to go back to its optimal distribution of workers. In equilibrium, an agreement is always reached.

To solve the problem, assume that each worker has size h. We will take the limit as $h \to 0$. The marginal discounted profit from hiring a worker of type s is proportional to

$$\begin{aligned} \Delta_s^n(w) = & F(g) - \sum_{k=1}^{S} w_k(g) g_k - \Bigg(F(\dots, g_s - h, \dots) - \sum_{k \neq s} w_k(\dots, g_s - h, \dots) g_k \\ & - w_s(\dots, g_s - h, \dots)(g_s - h) - h\gamma(1-\delta)\frac{\kappa}{q(\theta_s)} \Bigg) \end{aligned}$$

where the notation $(\dots, g_s - h, \dots)$ makes explicit the fact that we are considering the distribution g without a measure h of its sth member. Δ_s^n is simply the difference between the value of the firm with and without an agreement. Notice that in the latter case, the firm loses value since it faces an additional hiring cost in the next period to get back to its equilibrium size.

A solution to the Stole and Zwiebel bargaining is a wage vector w that solves

$$\frac{\beta_n}{1-\beta_n} \Delta_s^n(w) = (W_s^e(w) - W_s^u) h$$

where the right-hand side is the worker's surplus. By dividing Δ_s^n by h and taking the limit $h \to 0$, we get

$$\lim_{h \to 0} \frac{\Delta_s^n(w)}{h} = \frac{\partial F(g)}{\partial g_s} - \sum_{k=1}^{S} g_k \frac{\partial w_k(g)}{\partial g_s} - w_s(g) + \gamma(1-\delta)\frac{\kappa}{q(\theta_s)}.$$

Therefore, a solution must solve the following system of partial differential equations:

$$\frac{\partial F(g)}{\partial g_s} - \sum_{k=1}^{S} g_k \frac{\partial w_k(g)}{\partial g_s} - w_s(g) + \gamma(1-\delta)\frac{\kappa}{q(\theta_s)} = \frac{1-\beta_n}{\beta_n}(w_s(g) - c_s)$$

for all $s \in \mathcal{S}$. General solutions to this system are of the form

$$w_s^n(g) - c_s = \frac{\beta_n}{1-\beta_n(1-\alpha)} \frac{\alpha z_s}{g_s^{\frac{1}{\sigma}}} A \left(\sum_{k=1}^{S} z_k g_k^{\frac{\sigma-1}{\sigma}} \right)^{\frac{1-\sigma(1-\alpha)}{\sigma-1}} - \beta_n c_s + \beta_n \gamma(1-\delta) \frac{\kappa}{q(\theta_s)} + C_s g_s^{-\frac{1}{\beta_n}}$$

where C_s is a constant term that could depend on $\{g_j\}_{j \neq s}$. To fix the constants, I use the convenient boundary conditions

$$\left\{ \lim_{g_s \to 0} w_s^n(g) g_s = 0 \right\}_{s=1}^{S}$$

which guarantees that $C_s = 0$ for all s.[42] □

Proposition 1. *If the bargaining powers $\beta \equiv \beta_n = \beta_u$ are equal than the difference in average wage is*

$$\mathbb{E}_g(w^n) - \mathbb{E}_g(w^u) = -\frac{\beta(1-\beta)(1-\alpha)}{1-(1-\alpha)\beta} \frac{F(g)}{n} < 0$$

where $\mathbb{E}_g$ is the expectation across skills. Also, the difference in the profit of the firm is

$$\pi^n(g) - \pi^u(g) = \frac{(1-\beta)(1-\alpha)\beta}{1-(1-\alpha)\beta} F(g) > 0.$$

Proof. From equations 9 and 12:

$$\frac{\sum_s w^n(g_s) g_s}{\sum_s g_s} = \frac{1}{n} \left[\frac{\beta}{1-(1-\alpha)\beta} \alpha F(g) + (1-\beta) \sum_s c_s g_s + \beta\gamma(1-\delta)\kappa \sum_s \frac{g_s}{q(\theta_s)} \right]$$

and

$$\frac{\sum_s w^u(g_s) g_s}{\sum_s g_s} = \frac{1}{n} \left(\beta F(g) + (1-\beta) \sum_s c_s g_s + \beta\gamma(1-\delta)\kappa \sum_s \frac{g_s}{q(\theta_s)} \right)$$

Taking the difference yields the first result. Subtracting equation 10 from equation 13 yields the second result. □

Proposition 2. *If W_s^u and θ_s are increasing in s, then the equilibrium wage schedules $w_s^u(g^{u*})$ and $w_s^n(g^{n*})$ are increasing in s and the union wage gap $w_s^u(g^{u*}) - w_s^n(g^{n*})$ is decreasing in s.*

Proof. We first start with the union wage. From equation 9, we can write $w_s^u(g^{u*}) = c_s^u + D$ where D is a constant that does not depend on s. Combining with equation 7, we get $w_s^u(g^{u*}) = (1-\gamma(1-\delta))(b_s + D) + \gamma(1-\gamma)(1-\delta)W_s^u$. Since W_s^u is increasing in s so is $w_s^u(g^{u*})$.

[42] Cahuc et al. (2008) study a similar bargaining problem with general production functions.

For the nonunion wage, by combining equations 12 and 18, we find

$$w_s^n(g^{n*}) = c_s^n + \frac{\beta_n}{1-\beta_n}\frac{\kappa}{q(\theta_s)}.$$

Using equation 7 once again yields

$$w_s^n(g^{n*}) = (1-\gamma(1-\delta))\left(b_s + \frac{\beta_n}{1-\beta_n}\frac{\kappa}{q(\theta_s)}\right) + \gamma(1-\gamma)(1-\delta)W_s^u.$$

Since W_s^u and θ_s are increasing in s so is $w_s^n(g^{n*})$.

For the union wage gap, notice that

$$w_s^u(g^{u*}) - w_s^n(g^{n*}) = (1-\gamma(1-\delta))\left(D - \frac{\beta_n}{1-\beta_n}\frac{\kappa}{q(\theta_s)}\right).$$

Since θ_s is increasing in s, the union wage gap is decreasing in s. □

Proposition 3. *If W_s^u and θ_s are increasing in s, then the counterfactual union wage gap $w_s^u(g^{i*}) - w_s^n(g^{i*})$ is decreasing in s for $i \in \{u, n\}$.*

Proof. We first start with the unionized firm. This firm hires according to g_s^{u*}. From lemma 2, we know that $w_s^u(g^{u*}) = c_s^u + D$ and that c_s^u is increasing in s. Consider now the off-equilibrium nonunion wage that the union workers *would* get if they voted against the union. From equation 12 together with the first-order condition of an unconstrained firm we have

$$w_s^n(g^{u*}) = (1-\beta_n)c_s^u + \frac{\beta_n}{1-(1-\alpha)\beta_n}\frac{\mathrm{MC}_s^u}{1-\beta_u} + \beta_n\gamma(1-\delta)\frac{\kappa}{q(\theta_s)}.$$

Using the definition of MC^u, it is straightforward to show that

$$w_s^n(g^{u*}) = c_s^u + c_s^u\frac{\beta_n^2(1-\alpha)}{1-(1-\alpha)\beta_n} + \frac{\beta_n}{1-(1-\alpha)\beta_n}\frac{\kappa}{q(\theta_s)}\underbrace{\left(\frac{1}{1-\beta_u} - (1-\alpha)\beta_n\gamma(1-\delta)\right)}_{>0}.$$

Since W_s^u and θ_s are increasing, both $w_s^u(g^{u*})$ and $w_s^n(g^{u*})$ are increasing in s and

$$\begin{aligned} w_s^u(g^{u*}) - w_s^n(g^{u*}) = D - \Bigg(& c_s^u\frac{\beta_n^2(1-\alpha)}{1-(1-\alpha)\beta_n} \\ & + \frac{\beta_n}{1-(1-\alpha)\beta_n}\frac{\kappa}{q(\theta_s)}\left(\frac{1}{1-\beta_u} - (1-\alpha)\beta_n\gamma(1-\delta)\right)\Bigg). \end{aligned}$$

so that the union wage gap $w_s^u(g^{u*}) - w_s^n(g^{u*})$ is decreasing in s.

The nonunion firm hires according to g_s^{n*}. From the proof of the previous proposition, we know that

$$w_s^n(g^{n*}) = c_s^n + \frac{\beta_n}{1-\beta_n}\frac{\kappa}{q(\theta_s)}.$$

Furthermore, from equation 9, we have that $w_s^u(g^{n*}) = c_s^n + D'$ where D' is a constant that does not depend on s. Therefore, the union wage gap is

$$w_s^u(g^{n*}) - w_s^n(g^{n*}) = D' - \frac{\beta_n}{1-\beta_n}\frac{\kappa}{q(\theta_s)}.$$

Since θ_s is increasing, the union wage gap decreases with s. □

Proposition 4. *An unconstrained firm strictly prefers to be union free if and only if*

$$\frac{B_n}{B_u} > \left(\frac{\sum_s z_s^\sigma \left(\mathrm{MC}_s^u\right)^{1-\sigma}}{\sum_s z_s^\sigma \left(\mathrm{MC}_s^n\right)^{1-\sigma}}\right)^{\frac{\alpha}{\sigma-1}}.$$

Proof. We need to compare the value of a firm under g^{u*} and g^{n*}, for a given vector c_s. From lemma 1, we know that we can compare

$$\pi(g^{i*}) - \kappa(1-(1-\delta)\gamma)\sum_{s=1}^{S}\frac{g_s^{i*}}{q(\theta_s)}.$$

Using equation 10 and 13 together with equation 19 and after simplification we can now compare

$$(AB_i)^{\frac{1}{1-\alpha}}\alpha^{\frac{\alpha}{1-\alpha}}(1-\alpha)\left(\sum_{s=1}^{S} z_s\left(\frac{z_s}{\mathrm{MC}_s^i}\right)^{\sigma-1}\right)^{\frac{\alpha}{(1-\alpha)(\sigma-1)}}$$

for $i \in \{u, n\}$. A few simplifications yield the result. □

Proposition 5. *Under assumptions 1 and 2, the firm prefers to be union free but, if it hires without taking the vote into account, the majority of its workers would vote to form a union.*

Furthermore, if

$$\frac{B_n}{B_u} \geq \left(\frac{q(\theta_h)^{-1} + q(\theta_l)^{-1}}{\left(\frac{1}{z_l}q(\theta_l)^{-1}\right)^{z_l}\left(\frac{1}{z_h}q(\theta_h)^{-1}\right)^{z_h}}\right)^{\alpha}$$

and

$$\frac{\alpha\beta_n}{1-(1-\alpha)\beta_n}2z_h \geq \beta_u$$

it is optimal for the the firm to prevent unionization, otherwise it chooses to be unionized.

Proof. Under the assumptions, proposition 4 shows that the firm prefers g^{n*} to g^{u*} if and only if $B_n > B_u$. From equation 19 we also have that $g_l^{n*} > g_h^{n*}$. We need to show that

$$w_l^u(g^{n*}) > w_l^n(g^{n*})$$

which is equivalent to

$$\beta_u > \frac{\alpha\beta_n}{(1-(1-\alpha)\beta_n)}\left((1-z) + z\frac{q(\theta_h)}{q(\theta_l)}\right)$$

which is true under our assumptions. The low-skill workers therefore prefer to be unionized and they have the majority of the vote. The union threat therefore binds.

We now consider how the firm prevents unionization. The firm solves

$$\max_{g_l,g_h} \frac{1-\beta_n}{1-(1-\alpha)\beta_n} A\left(g_l^{1-z} g_h^z\right)^\alpha - \kappa\left(\frac{g_h}{q(\theta_h)} + \frac{g_l}{q(\theta_l)}\right)$$

under three possible sets of constraints. Either: (1) subject to $w_h^n \geq w_h^u$ and $g_h \geq g_l$, or (2) subject to $w_l^n \geq w_l^u$ and $g_l \geq g_h$, or (3) subject to $w_h^n \geq w_h^u$ and $w_l^n \geq w_l^u$. The problem of the firm must fall into one of these three cases.

We can disregard the constraint set (3) right away as there is no distribution of workers such that both nonunion wages are higher than their union counterpart. In the constraint set (2) the first constraint cannot bind. Indeed, it would imply

$$\begin{aligned}
\frac{\alpha\beta_n}{1-(1-\alpha)\beta_n}(1-z)\frac{1}{g_l}F(g) &= \frac{\beta_u}{g_l+g_h}F(g) \\
\frac{\alpha\beta_n}{1-(1-\alpha)\beta_n}z\frac{1}{g_h}F(g) &< \frac{\beta_u}{g_l+g_h}F(g)
\end{aligned}$$

where the second inequality is $w_h^n < w_h^u$, otherwise we are is constraint set (3). For both these equations to be satisfied requires

$$(1-z)\frac{1}{g_l} > z\frac{1}{g_h}$$

which is impossible as $z > 1/2$ and the second constraint in (2) must hold. Let us show that the second constraint also cannot bind. If it were binding, the first constraint would be

$$\begin{aligned}
\frac{\alpha\beta_n}{1-(1-\alpha)\beta_n}(1-z)\frac{1}{g_l}F(g) &\geq \frac{\beta_u}{g_l+g_h}F(g) \\
\frac{\alpha\beta_n}{1-(1-\alpha)\beta_n} &\geq \frac{\beta_u}{2(1-z)}
\end{aligned}$$

but $\frac{\alpha\beta_n}{1-(1-\alpha)\beta_n} < \beta_u$ by assumption and since $z > 1/2$ this equality is never satisfied. So neither the first nor the second constraint binds but then we know that the union wins the vote. We can therefore disregard the constraint set (2) as an appropriate strategy for the firm.

We know turn to (1), which is the interesting case. Let us consider the case in which only the second constraint binds. In this case, the first-order conditions yields

$$g_s^n = \left(\frac{1-\beta_n}{1-(1-\alpha)\beta_n}\left(\frac{\kappa}{q(\theta_l)} + \frac{\kappa}{q(\theta_h)}\right)^{-1} \alpha A\right)^{\frac{1}{1-\alpha}}$$

and the firm prefers this strategy over unionization when

$$J(g^n) \geq J(g^{u*})$$
$$\left(\frac{1-\beta_n}{1-(1-\alpha)\beta_n}\right)\left(\left(\frac{1-z}{\kappa}q(\theta_l)\right)^{1-z}\left(\frac{z}{\kappa}q(\theta_h)\right)^{z}\right)^{-\alpha} \geq (1-\beta_u)\left(\frac{\kappa}{q(\theta_h)}+\frac{\kappa}{q(\theta_l)}\right)^{\alpha}$$

which is equivalent to the profitability condition 22. We need to verify that high-skill workers vote against the union. This is true if

$$w_h^n(g^n) \geq w_h^u(g^n)$$
$$\frac{\beta_n}{1-\beta_n}z_s\left(\frac{\kappa}{q(\theta_l)}+\frac{\kappa}{q(\theta_h)}\right) \geq \frac{\beta_u}{2}\frac{1}{\alpha}\frac{1-(1-\alpha)\beta_n}{1-\beta_n}\left(\frac{\kappa}{q(\theta_l)}+\frac{\kappa}{q(\theta_h)}\right)$$

which is equivalent to the feasibility condition 23.

We need to rule out the other possibilities in constraint set (1). First, note that if both constraints binds

$$w_h^n(g^n) = \frac{\alpha\beta_n}{1-(1-\alpha)\beta_n}z_h\frac{1}{g}F(g) = \frac{\beta_u}{2g}F(g) = w_h^u$$

which is, in general, not true. Finally, let us consider the case in which only the first constraint in (1) binds. The first-order conditions yield

$$\begin{aligned}\frac{1-\beta_n}{1-(1-\alpha)\beta_n}(1-z)\alpha F(g)-\kappa\frac{g_l}{q(\theta_l)} &= \gamma g_l\left(\frac{\alpha\beta_n}{1-(1-\alpha)\beta_n}z\right)\\ \frac{1-\beta_n}{1-(1-\alpha)\beta_n}z\alpha F(g)-\kappa\frac{g_h}{q(\theta_h)} &= \gamma g_h\left(\frac{\alpha\beta_n}{1-(1-\alpha)\beta_n}z-\beta_u\right).\end{aligned}$$

Plugging into the constraint to find the Lagrange multiplier γ and simple algebra shows that the second constraint, $g_h \geq g_l$, also requires the feasibility condition 23. Finally, straightforward algebra shows that this strategy is never more profitable than the one from constraint set (1) in which only the second constraint binds. □

Proposition 6. *Under assumptions 1 and 2, a binding voting constraint in a nonunion firm lowers profits, employment and output.*

Proof. It is obvious that the voting constraint lowers profits. We know from the previous proof that

$$g_s^n = \left(\frac{1-\beta_n}{1-(1-\alpha)\beta_n}\left(\frac{\kappa}{q(\theta_l)}+\frac{\kappa}{q(\theta_h)}\right)^{-1}\alpha A\right)^{\frac{1}{1-\alpha}}.$$

Plugging into the production function, we are interested in showing

$$F(g^{n*}) \geq F(g^n) \Leftrightarrow \left(\frac{1-z}{\frac{\kappa}{q(\theta_l)}}\right)^{1-z}\left(\frac{z}{\frac{\kappa}{q(\theta_h)}}\right)^{z} \geq \left(\frac{\kappa}{q(\theta_l)}+\frac{\kappa}{q(\theta_h)}\right)^{-1}$$

By the means inequality, this inequality is always satisfied, with equality if and only if $(1-z)q(\theta_l) = zq(\theta_h)$. For employment, we consider the inequality

$$\sum_s g_s^{n*} \geq \sum_s g_s^n$$

$$\frac{(zq(\theta_h) + (1-z)q(\theta_l))}{\frac{2}{\frac{1}{q(\theta_l)}+\frac{1}{q(\theta_h)}}} \geq \left(\frac{1}{2}\frac{2}{\frac{1}{q(\theta_l)}+\frac{1}{q(\theta_h)}}\left((1-z)^{-1}q(\theta_l)^{-1}\right)^{1-z}\left(z^{-1}q(\theta_h)^{-1}\right)^z\right)^{\frac{\alpha}{1-\alpha}}$$

By the means inequality, a sufficient condition for this inequality to hold is

$$\frac{(zq(\theta_h) + (1-z)q(\theta_l))}{\frac{2}{\frac{1}{q(\theta_l)}+\frac{1}{q(\theta_h)}}} \geq \left(\frac{1}{2}\frac{2}{\frac{1}{q(\theta_l)}+\frac{1}{q(\theta_h)}}\left(\frac{1}{q(\theta_h)}+\frac{1}{q(\theta_l)}\right)\right)^{\frac{\alpha}{1-\alpha}}.$$

The inequality simplifies to

$$(1-z)\frac{q(\theta_l)}{q(\theta_h)} + z\frac{q(\theta_h)}{q(\theta_l)} \geq 1.$$

For the lowest possible $q(\theta_l)$, $q(\theta_l) = zq(\theta_h)(1-z)^{-1}$, this inequality is satisfied. As $q(\theta_l)$ increases, so does the right-hand side of the inequality, so it is satisfied for any $q(\theta_l)$ that satisfies our assumptions.

□

Proposition 7. *Under assumptions 1 and 2, a binding voting constraint in a nonunion firm increases the average wage and decreases wage inequality, as defined by the ratio of the high-skill wage to the low-skill wage.*

Proof. For the average wage, we need to show that

$$\mathbb{E}\left(w\left(g^n\right)\right) \geq \mathbb{E}\left(w\left(g^{n*}\right)\right)$$

$$\frac{\beta_n}{1-\beta_n}\frac{1}{2}\left(\frac{\kappa}{q(\theta_l)}+\frac{\kappa}{q(\theta_h)}\right) \geq \frac{\beta_n}{1-\beta_n}\frac{\kappa}{(1-z)q(\theta_l)+zq(\theta_h)}$$

which is equivalent to

$$z\frac{q(\theta_h)}{q(\theta_l)} + (1-z)\frac{q(\theta_l)}{q(\theta_h)} \geq 1$$

which, as explained in the previous proof, is true. For the ratio of wages, we know that

$$\frac{w_h\left(g^{n*}\right)}{w_l\left(g^{n*}\right)} = \frac{q(\theta_l)}{q(\theta_h)}$$

and

$$\frac{w_h\left(g^n\right)}{w_l\left(g^n\right)} = \frac{z}{1-z}$$

so that

$$\frac{w_h\left(g^{n*}\right)}{w_l\left(g^{n*}\right)} > \frac{w_h\left(g^n\right)}{w_l\left(g^n\right)}$$

is true under our assumptions. $\square$

Lemma 4. *Consider two firms, identified by the subscripts 1 and 2, that have identical technologies except for $A_1 \neq A_2$. In equilibrium, if g_1 solves the problem of firm 1, then $g_2 = (A_2/A_1)^{\frac{1}{1-\alpha}} g_1$ solves the problem of firm 2. Also, both firms have the same union status and pay the same wages.*

Proof. Assume first that the equilibrium schedules c_1 and c_2 are identical and denote that schedule by c. This result will be shown later in the lemma. We can write the problem of firm $j \in \{1, 2\}$ as

$$\max_g \mathcal{J}(A_j, w(A_j, g), g)$$

such that

$$w(A_j, g) = \begin{cases} w^u(A_j, g) & \text{if } V(A_j, g) > 0 \\ w^n(A_j, g) & \text{if } V(A_j, g) \leq 0 \end{cases}$$

where w^u is the union wage function, w^n is the nonunion wage function and V is the excess number of workers for unionization.

The proof proceeds by showing that if g_1 solves the FOC of firm 1 then

$$g_2 = \left(\frac{A_2}{A_1}\right)^{\frac{1}{1-\alpha}} g_1$$

solves the FOC of firm 2.

We therefore start with the FOC of firm 1 given by equations 20 and 21. First, notice that

$$\begin{aligned} \frac{A_1}{(g_{1,s})^{1/\sigma}} \left(\sum_{k=1}^{S} z_k g_{1,k}^{\frac{\sigma-1}{\sigma}}\right)^{\frac{1-\sigma(1-\alpha)}{\sigma-1}} &= \frac{A_1 \left(\frac{A_1}{A_2}\right)^{\frac{1-\sigma(1-\alpha)}{\sigma(1-\alpha)}}}{\left(\left(\frac{A_1}{A_2}\right)^{\frac{1}{1-\alpha}} g_2\right)^{1/\sigma}} \left(\sum_{k=1}^{S} z_k g_{2,k}^{\frac{\sigma-1}{\sigma}}\right)^{\frac{1-\sigma(1-\alpha)}{\sigma-1}} \\ &= \frac{A_2}{(g_{2,s})^{1/\sigma}} \left(\sum_{k=1}^{S} z_k g_{2,k}^{\frac{\sigma-1}{\sigma}}\right)^{\frac{1-\sigma(1-\alpha)}{\sigma-1}}. \end{aligned}$$

This also implies that $w^n(A_1, g_1) = w^n(A_2, g_2)$. It is also straightforward to show that $F(A_1, g_1)/n_1 = F(A_2, g_2)/n_2$ such that $w^u(A_1, g_1) = w^u(A_2, g_2)$. We have so far shown that the terms not multiplied by the Lagrange multiplier in equation 20 are the same, which completes the proof if firm 1 is unconstrained ($\lambda^1 = 0$). In which case, firm 2 is also unconstrained.

We now consider the derivatives in equation 21. Notice that, for any $s' \neq s$, we have

$$g_{1,s'} \frac{\partial w^n_{s'}(A_1, g_1)}{\partial g_{1,s}} = \frac{\alpha\beta_n(1-\sigma(1-\alpha))}{(1-\beta_n(1-\alpha))\sigma} A_1 \left(\sum_{k=1}^{S} z_k g_{1,k}^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\alpha\sigma-2\sigma+2}{\sigma-1}} z_s g_{1,s}^{-1/\sigma} z_{s'} g_{1,s'}^{\frac{\sigma-1}{\sigma}}$$

$$= \frac{\alpha\beta_n(1-\sigma(1-\alpha))}{(1-\beta_n(1-\alpha))\sigma} A_2 \left(\sum_{k=1}^{S} z_k g_{2,k}^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\alpha\sigma-2\sigma+2}{\sigma-1}} z_s g_{2,s}^{-1/\sigma} z_{s'} g_{2,s'}^{\frac{\sigma-1}{\sigma}} = g_{2,s'} \frac{\partial w^n_{s'}(A_2, g_2)}{\partial g_{2,s}}.$$

Similarly,

$$
\begin{aligned}
g_{1,s}\frac{\partial w_s^n(A_1,g_1)}{\partial g_{1,s}} &= \frac{\alpha\beta_n z_s A_1}{1-\beta_n(1-\alpha)}\left(-\frac{1}{\sigma}\left(\sum_{k=1}^{S} z_k g_{1,k}^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\alpha\sigma-\sigma+1}{\sigma-1}} g_{1,s}^{-\frac{1}{\sigma}}\right. \\
&\quad \left. +g_{1,s}^{\frac{\sigma-2}{\sigma}}\frac{1-\sigma(1-\alpha)}{\sigma}\left(\sum_{k=1}^{S} z_k g_{1,k}^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\alpha\sigma-2\sigma+2}{\sigma-1}} z_s\right) \\
&= g_{2,s}\frac{\partial w_s^n(A_2,g_2)}{\partial g_{2,s}}.
\end{aligned}
$$

Similar computations yield that for any $s' \in \{1,\ldots,S\}$

$$
g_{1,s'}\frac{\partial w_{s'}^u(A_1,g_1)}{\partial g_{1,s}} = g_{2,s'}\frac{\partial w_{s'}^u(A_2,g_2)}{\partial g_{2,s}}.
$$

Combining these results, it follows that $V(A_1,g_1) = V(A_2,g_2)$ and that

$$
\frac{\partial V(A_1,g_1)}{\partial g_{1,s}} = \frac{\partial V(A_2,g_2)}{\partial g_{2,s}}
$$

for all s. This completes the proof since, if firm 1 is constrained, there exists a $\lambda^2 = \lambda^1 \geq 0$ such that g_2 solves the problem of firm 2 and $V(A_2,g_2) = 0$. Notice that firm 2 is also constrained. Notice also that since the two firms have the same union status and are paying the same wages, we find $c_1 = c_2$, which justifies our initial assumption. $\square$