The Impact of Unions on Non-union Wage Setting: Threats and Bargaining*

David A. Green† Ben M. Sand‡ Iain G. Snoddy§ Jeanne Tschopp¶

May 27, 2024

Abstract

In this paper we provide new estimates of the impact of unions on non-union wage setting. We allow the presence of unions to affect non-union wages both through the typically discussed channel of non-union firms emulating union wages in order to fend off the threat of unionisation and through a bargaining channel in which non-union workers use the presence of union jobs as part of their outside option. In our most complete model, we specify these channels in a search and bargaining framework that includes union formation and the possibility of non-union firm responses to the threat of unionisation. Our results indicate an important role played by union wage spillovers in lowering wages over the 1980-2010 period. We find that de-unionisation can account for 35% of the decline in the mean hourly wage between 1980 and 2010 in the US, with two-thirds of that effect being due to spillovers. Both the traditional threat and bargaining channels are operational, with the bargaining channel being more important.

*We are grateful to Thomas Lemieux, Craig Riddell, and Raffaele Saggio for their comments. We also greatly benefited from comments from participants at seminars at the University of Chicago, the University of Melbourne, Carleton University, the University of British Columbia, the University College London, the University of Mannheim, and the Society of Labor Economists annual meeting.

†University of British Columbia and Research Fellow, IFS, London david.green@ubc.ca

‡Department of Economics, York University bmsand@yorku.ca

§Analysis Group, Montreal iain.snoddy@analysisgroup.com

¶University of Bern jeanne.tschopp@unibe.ch

1 Introduction

Private sector unionisation in the United States is very nearly dead. In 2020, only 6.3% of private sector workers belonged to a union (U.S. Bureau of Labor Statistics (2021)). Recently, however, there have been some glimmers of revival, including successful unionisation drives at Amazon and Starbucks, raising questions about whether a resurrected union movement could significantly impact wage levels across the economy. While we can't look ahead to future changes in unionisation, we can use the de-unionisation over the last 50 years to better understand union impacts. The most direct impact of decreased unionisation, of course, comes from the shifting of workers out of higher-paid union jobs. But it also has the potential to alter wage setting in non-union jobs. These latter spillover effects are important since their existence would imply that the reach of unions is larger than it might first appear and larger than what is calculated based on standard shift-share decompositions. In this paper, we build a model of the impact of unions on wage setting in the non-union sector and use it in estimation based on Current Population Survey (CPS) data to re-assess the role of de-unionisation in movements in the wage structure in the U.S.

The idea that unions could impact non-union wage setting goes back at least to Lewis [1963]. The core idea raised in that book, and discussed in subsequent papers such as Rosen [1969], is that non-union firms raise their wages in response to the 'threat' that their workers will unionise, presumably imposing extra costs beyond direct wage increases.[1] Our model incorporates that threat effect plus an added union impact mechanism: a bargaining channel whereby the outside options of non-union workers and, through that, their bargained wages are affected by their ability to find high-paying union jobs. In a sense, both are threats, with one being the threat of workers leaving and finding a union job (what we will call a bargaining effect) and the other channel being the threat to unionise the non-union workplace (which we call a standard threat effect).

These two channels, however, have different implications for attempts to raise wages through policy tools. The threat effect is unique – it can only be harnessed by increasing unionisation. The bargaining effect, in contrast, is more general. It is about more good-paying jobs, which improve the outside options for workers in all other jobs. As noted in Beaudry et al. [2012] and Caldwell and Danieli [2021], this can significantly increase wages in a given location. Unions are one way to create such a higher wage option, but other policies, such as eliminating non-compete arrangements, could also have such an impact [Johnson et al., 2020]. Our model clarifies the difficulties inherent in identifying these two effects separately while controlling for selection into the union/non-union sectors. Part of this paper's contribution is to offer estimates of spillover effects through both channels, expanding our understanding of the impact of de-unionisation on the wage structure. Based on our estimates, we then assess how de-unionisation has contributed to changes in the wage structure in the U.S. over recent decades.

The existing literature estimates union wage spillovers by regressing non-union wages on the percent of organised workers in labour markets defined by location and/or industry.

[1] For instance, Starbucks recently offered wage increases to "company-operated stores" but not to "unionised stores, or to stores that may be in the process of unionising". The NLRB has designated the announcement as a threat, designed to have a "chilling effect" on impending union votes (New York Times, May 2022).

Evidence based on this approach is mixed and sensitive to the included control terms, with the preponderance of studies finding a small positive spillover effect.[2] In an important analysis, Farber [2005] carefully considers the role played by omitted variables, sequentially introducing industry and state-fixed effects. He finds great sensitivity in estimates to the source of variation used, providing some context for the disparity in estimates across earlier studies. When controlling for a wide range of potential omitted variables, his results indicate, at most, a small positive effect of union power on non-union wages.

In a recent paper, Fortin et al. [2021] estimate the impact of union threat effects on wage inequality, using industry$\times$state-level variation in the unionisation rate as an additional covariate in their distribution regression approach. They find positive effects of the unionisation rate operating primarily at the part of the wage distribution just below the median, and their counterfactual exercise indicates that spillovers double the measured impact of de-unionisation in increasing wage inequality in the U.S. While very useful, the paper shares with all of the early analyses a lack of an identification strategy for addressing the potential endogeneity of the union proportion – stemming from a lack of an effective instrument.[3] None of the papers in the literature even mention the twin problem of potential selectivity bias. As the proportion of unionised workers declines, the composition of non-union workers and firms will change.

In contrast to the existing literature on union spillovers that largely relies on reduced-form estimation, our approach formalises union spillovers in a search and bargaining framework, endogenising the union formation process and incorporating wage effects arising through differences in the bargaining process. In making clear what is being identified in the model and the variation used, we overcome the problems inherent in early studies of likely biases due to omitted characteristics and selection into the union sector, and we estimate an effect with a clear theoretical basis and interpretation.

Our model is based on that of Taschereau-Dumouchel [2020] (henceforth TD), whose work is informed by the contributions of Pissarides [1986], Açıkgöz and Kaymak [2014], and Krusell and Rudanko [2016], among others. The TD model is centred around union threat effects through the hiring channel. In the model, more skilled workers tend to dislike unionisation, and firms skew their hiring toward these workers to stack the unionisation vote. Though this effect is certainly interesting, we believe it is likely of second-order importance relative to a more direct firm response through raising wages to lessen the gains from unionising and direct union-busting actions, which raise the costs of unionisation. Our model focuses on these latter effects instead of the hiring channel.

Additionally, our framework is informed by papers, including Beaudry et al. [2012] (henceforth BGS), Tschopp [2017], Caldwell and Danieli [2021], Jarosch et al. [2024], and

[2] Both Freeman and Medoff [1981] and Donsimoni [1981] find a non-significant positive correlation between non-union wages and the proportion of unions. Conversely, Holzer [1982], Kahn [1980] and Dickens and Katz [1986] estimate large positive effects. Hirsch and Neufeld [1987] find a positive spillover effect at the industry level but insignificant effects at the local labour market level. Podgursky [1986] finds spillover effects exist only for large establishments, and Neumark and Wachter [1995] estimate a negative effect.

[3] Farber [2005] presents event studies of the enactment of Right to Work (RTW) laws in Idaho and Oklahoma. In an earlier version of Fortin et al. [2019], the authors extend Farber's analysis to include more states. They find evidence of reductions in non-union wages with the introduction of RTW laws but the estimates are poorly defined because few states switch RTW status in their time period.

Bassier [2022], which formalise the impact of changing alternative job prospects (outside options) on wages. Following BGS, we model local labour markets composed of industries and firms with workers able to transition between jobs in proportion to job prevalence. As in BGS, we will exploit cross-city, within-industry variation – in our case, to identify the effect of declining unionisation on non-union wages from 1980 to 2020.

Combining these elements, we derive an empirical specification which incorporates spillover effects operating through both the bargaining and standard threat channels, formalises selectivity, and makes it straightforward to see barriers to identification. Specifically, changes in outside options associated with the union sector may be correlated with unobserved local productivity shocks. As in BGS and Beaudry et al. [2014], we overcome this problem using Bartik-style instruments related to worker outside options. For non-union workers, outside options are related to the probability the worker could transit to a union job (which we allow to vary by industry and over time) times the expected wage the worker could get in that job. It also depends on expected wages in non-union jobs in the local economy and the probabilities of transitioning to those jobs. The Bartik instruments use versions of these outside options based on the start-of-period industry and union employment composition in a locality interacted with changes in industry growth, industry premia, and the probabilities of moving to different types of jobs defined by industry and union status at the national level.

The outside option for non-union workers identifies the bargaining channel for union effects. We get extra power to identify the bargaining effect because improvements in outside options have the same effect on bargained wages, whether they stem from reduced probabilities of finding a union job or a high-rent non-union job. That means we get identification from both unionisation changes and industrial structure shifts in both the non-union and union sectors. We argue that the validity of our Bartik instrument depends on a random walk-type assumption that we show implies an over-identifying restriction. We test that restriction and cannot reject it. We also show that within the context of the model, we identify the threat channel by the impact on non-union wages of the interaction of the probability a firm in a given industry×city cell would face a union election (which shows the size of the direct threat) with the outside option value for union workers (which captures the size of what the firm needs to respond to in order to prevent unionisation). We construct and implement similar Bartik instruments related to this component.

The results from our estimation point to the importance of both spillover channels. Between 1980 and 2010, the mean real wage in the U.S. fell 16% (holding composition in terms of education, experience, race and gender constant). A decomposition exercise based on our estimates shows that de-unionisation accounts for 35% of the decline. A third of that impact arises from a standard shift-share effect (because workers shifted away from higher-paying union jobs), with the other two-thirds from spillover channels. Unions have spillover effects on non-union wages, and they are sizeable. While both the traditional threat and bargaining effects show up significantly in our estimates, our decomposition exercise indicates that the spillover effects are almost entirely due to the latter. The threat probability was too low, even in 1980, to play a substantial role. As we point out earlier, the dominance of the bargaining channel means the effects of unions in raising non-union wages could also be achieved through other policies that raise average worker rents. The effect is not unique to unions.

Our estimates imply that spillovers roughly doubled the standard shift-share effect of unionisation over the long run. Perhaps surprisingly, we find that the spillover effect was smallest (though still sizeable) in the 1980s – the decade of the largest declines in unionisation. This is because those declines were offset by increases in the union wage premium, increasing the value of the outside option of non-union workers while the declining probability of finding union jobs reduced it. Our model provides an explanation for the increased wage premium in the 1980s, which echoes an argument in Farber [2005]. While both union and non-union wages faced downward pressures from technological change, trade, etc., the substantial reduction in the risk of being unionised in the decade meant that, in addition, non-union firms no longer had to pay higher wages in order to stave off unionisation. As a result, non-union wages fell faster than union wages. After 1990, the threat of unionisation stabilised at a low level, causing the union wage premium to decline, and the outside option effect of unions began to reflect the falling unionisation rate alone. The potential lesson for any re-unionisation efforts is that spillover effects onto non-union wages may arise through the traditional threat channel but the implied increase in non-union wages will dampen the bargaining channel. Union jobs would be more plentiful but not pay as high a premium over non-union jobs as before re-unionisation. Eventually, as the unionisation threat stabilised, the extent of spillover onto non-union wages would increase, but that could take time to realise fully.

Our work is also related to the substantial literature investigating patterns in declining unionisation, estimating both movements in the union wage premium and the role of declining unionisation in driving increasing wage inequality. Card et al. [2004] and Card et al. [2018b] provide comprehensive summaries of the research in this area following the early contribution of Freeman [1980]. Farber et al. [2021] provides the most comprehensive account of the relationship between union density and inequality in the U.S., introducing new survey data that allows them to push their analysis back to the 1930s. They find that increasing unionisation substantially impacted decreasing inequality after WWII, while the reversal in the unionisation trend had a smaller effect on increasing inequality in the last 50 years. Their estimates allow for spillover effects onto non-union wages, but they do not study spillovers directly. Our results imply that spillovers may have played an important role in their estimated inequality impacts from unions and explain why those impacts were less evident at the time of the big union decline in the 1980s.[4]

The remainder of the paper is organised as follows. In Section 2, we present our model. In Section 3, we derive our empirical specification and discuss the implementation and identification of challenges and solutions. We also present the construction of our key outside option variables and our instrumental variables. Section 4 describes the data and presents descriptive patterns that highlight our identifying variation. Section 5 contains our estimation results. In Section 6, we present a counterfactual exercise designed to demonstrate spillovers' impact on wage structure movements and the role played by our two channels. Section 7 contains conclusions.

[4]Other papers in this literature include an important contribution by DiNardo et al. [1996], which attributes 14% of the increase in wage inequality over 1979-1988 (for men) to declining unionisation. Extensions of this work are found in DiNardo and Lemieux [1997] and Fortin et al. [2021]. Further studies by Card [2001], Card et al. [2004], Gosling and Lemieux [2001], and Card et al. [2018b] extend the analyses by sector, gender, and across countries. See also recent studies by Farber et al. [2021] and Firpo et al. [2018].

2 The Model

2.1 Model Set-up

Our goal with our model is to derive an estimable specification for non-union wages that captures key channels through which those wages can be affected by changes in unionisation. Our model is based on that of Taschereau-Dumouchel [2020] (TD), which places union formation and wage setting in a search and bargaining model. Unions are able to bargain a higher wage because they can threaten to take the whole workforce out of production, while an individual, non-union worker can only threaten to withdraw her own labour. As mentioned in the introduction, TD focuses on firms responding to the threat of unionisation by altering the skill composition of their hiring while we focus on a response through paying higher wages. Through the rest of the paper, we will refer to non-union firms' wage responses to resist unionisation as standard threat effects (to reflect that these are what have been discussed in the previous literature).

In addition to standard threat effects, we allow for unionisation levels to affect non-union wages through a bargaining channel. Since unions can bargain higher wages for their members, having more unionised jobs in the local economy improves the outside option for all workers – even workers in firms not directly threatened with unionisation or workers in different industries – thus raising their wages. To investigate whether this channel has sizeable effects, we alter the TD model by having only one skill level but, following Beaudry et al. [2012](BGS), multiple industries.[5] We refer to any such effects as bargaining effects.

The model is an adjusted version of a standard Diamond-Mortensen-Pissarides bargaining model. In the model, there are C cities indexed by c, and we are interested in differences in non-union wages across cities with different unionisation levels. There are also I industries, indexed by i, which are assumed to produce tradeable goods with prices, p_i. Worker-firm matches die at an exogenous rate, δ^m, and all agents face a common discount rate, ρ. Firms face an additional probability of closing down, δ^e, with new firms born at the same rate to keep the number of firms fixed. Workers search for jobs while unemployed. The model is partial equilibrium in the sense that we treat the number of firms, the meeting rates between workers and vacancies, and the local employment rate as exogenous.[6] The model is centred on workers and firms (endogenously) ending up in one of three types of arrangements: simple non-union firms, non-union firms that emulate union wages, or union firms.

To understand the intuition underlying our model, it is helpful to go through its timing.

1. Firms are all born non-union.[7] At the time of birth they learn about their productivity and about the value of the idiosyncratic amenity that workers would create should they unionise the firm. The combination of the amenity and firm productivity, which are

[5] We bring differing skill levels back in through a model-consistent route in our empirical specification.

[6] Working in partial equilibrium in this way eliminates a channel through which de-unionisation could affect wages by lowering labour costs, causing firms to post more vacancies and, through that, increasing labour market tightness. This channel would have effects that are opposite to those of the channels we emphasize. We return to this channel in the empirical work.

[7] Although we specify the firm as the level at which workers become unionised in the model, our estimation exploits variation at the industry-city level. In that sense, firms can be thought of more accurately as establishments or, more generally, the level at which the unionisation vote occurs.

assumed to be independent, will determine which firms are unionised.

2. Following TD, firms first unilaterally determine their optimal level of employment and open vacancies to meet that target. The target employment will depend on the anticipated wage that will be bargained with workers, which, in turn, depends on which of the three arrangements is relevant to the firm. In steady state, the firm knows which arrangement it will experience.

3. Next, workers and firms meet according to a matching technology that is allowed to vary by the industry and union status of the previous job held by the worker. For example, a worker formerly in a unionised construction job may find it easier to be hired at a unionised car plant than another worker formerly a non-unionised retail employee.

4. Assuming the match-specific surplus is positive, the newly hired workers decide whether to unionise, which we assume is determined through a median voter model. Unions create firm-specific amenities, and locations differ in legislation that alters workers' unionisation costs. These imply that workers will not want to unionise some firms in some locations. We refer to those firms as 'simple non-union' firms.

5. In simple non-union firms, individual workers bargain wages with the firm. The worker outside option is the value of unemployment, while firms suffer the loss of the value of the marginal product of the individual worker if bargaining breaks down.

6. Without any firm response, the remaining firms would become unionised. However, the cost of unionising and the creation of union amenities implies that the size of the match-specific surplus will change once a firm is unionised. In fact, if the costs are larger than the amenity benefits, then the surplus will shrink, and the firm and the workers would jointly be better off if they remained non-union. In those situations, workers and firms bargain a wage with the outcome if bargaining breaks down being unionisation and, thus, the outside option values in the bargaining being the value to each of being in a unionised arrangement. The resulting wage is higher than the simple non-union wage, so it is worthwhile for workers not to unionise. This is the classic response to a unionisation threat seen in papers such as Rosen [1969], though, in the previous literature, the wage increase to avoid unionisation is treated as a unilateral firm decision while we treat it as a joint bargaining outcome. We refer to firms in this situation as 'emulating non-union' firms.

7. For a third set of firms, the balance of the benefits of amenities and the costs of unionisation are such that the surplus is larger if the firm is unionised. In that case, the firm bargains with the whole set of workers. As in TD, the worker outside option remains the value of being unemployed, but for firms, a breakdown in bargaining means a complete shutdown in production, which, as we said earlier, is why unions can bargain higher wages. We will refer to these firms as 'union' firms.

In this model, an increase in the cost of unionisation reduces the union threat and, thus, both the number of emulating non-union firms and the wages that they pay. It also reduces workers' outside options in simple and emulating non-union firms because there are fewer higher-paid union and emulating non-union firms for them to move to. The reduction in the outside option is relevant for workers in all sectors, lowering their bargained wage.

With this structure in mind, we next fill in the details needed to derive our estimating wage equations. A complete derivation of the model can be found in Appendices B – F.

2.2 Some Notation

We index union arrangements by τ with: $\tau = 1$, 2 and 3 corresponding to union firms, simple non-union firms, and emulating non-union firms, respectively. We index 'jobs' by j with $j = \{\tau, i\}$, i.e., jobs are combinations of union arrangements and industry. We use the subscript k for the potential destination jobs, with destination union status and industry, denoted τ' and i', respectively, such that $k = \{\tau', i'\}$.

2.3 Matching

Firms and workers operate in a labour market with frictions, meaning that workers and firms do not find each other and form a match perfectly easily. We assume that match formation depends on both the job type in which the vacancy is posted and the job type in which the worker was last employed. In particular, we employ a matching function of the form:

$$M_{kc|jc} = \theta_{jc} M(U_c, \Omega_c) \phi_{kc} \chi_{kc|j}(\varphi_{k|j}) \tag{1}$$

where $M_{kc|jc}$ is the number of matches of unemployed workers whose last job was of type j to vacancies of job type k in city c; θ_{jc} is the proportion of unemployed workers who were formerly in j; ϕ_{kc} is the proportion of vacancies that come from k; $M(U_c, \Omega_c)$ is the total number of matches observed in a city, with U_c being the total number of unemployed workers and Ω_c, the total number of vacancies in the city; and $\chi_{kc|j}(\varphi_{k|j})$ represents the specific frictional costs of moving from j to k. Thus, the number of matches of workers from j to vacancies of type k equals a purely mechanical component (the total number of matches times the proportion of unemployed workers who come from j and the proportion of vacancies of type k) times a component representing the fact that there are barriers to forming some j, k matches.[8] For example, a match between a worker who formerly worked in a unionised construction job and a vacancy posted by a unionised steel firm may be particularly easy to consummate while a match between that same worker and a non-union legal services firm may be less likely to actually happen. $\chi_{kc|j}(\varphi_{k|j})$ represents these frictional costs and takes the form:

$$\chi_{kc|j} = \frac{\varphi_{k|j}}{\sum_{k'} \eta_{k'c} \varphi_{k'|j}} \quad \forall k \tag{2}$$

where $\varphi_{k|j}$ represents the specific mobility frictions in moving from jobs of type j to jobs of type k (regardless of city). Assuming (as is standard) that $M(U_c, \Omega_c)$ is constant returns to

[8]Characterizing differential match rates as reflecting differential frictional costs follows Tschopp [2017]. Bassier [2022], alternatively, refers to differences in worker movements across firms as reflecting differences in 'consideration sets'. Caldwell and Harmon [2019] discusses differences based on personal networks.

scale (CRS), $M_{kc|jc}$ is also CRS. As we show in Appendix B, in steady state, θ_{jc}, ϕ_{kc} and $\chi_{kc|j}(\varphi_{k|j})$ all adjust to maintain a constant matching rate and sectoral composition.

The probability that a firm fills a vacancy of job type k is, $q^v_{kc} = \frac{M_{kc}}{\Omega_{kc}}$, where $M_{kc} = \sum_j M_{kc|jc}$ and $\Omega_{kc} = \phi_k \Omega_c$. It is straightforward to show that given the CRS assumption, in steady state, $q^v_{kc} = \frac{M_c}{\Omega_c} = q^v_c$. Hence, the probability that a firm fills a job is independent of the specific job and only depends on the local matching process. In a similar vein, the probability an unemployed worker from j makes a successful match with a vacancy in k equals $q^u_{kc|j} = q^u_c \eta_{kc} \chi_{kc|j}(\varphi_{k|j})$, where η_{kc} is the proportion of employment in job k in city c and $q^u_c = \frac{M_c}{U_c}$. Thus, the probability an unemployed worker who last worked at a job j matches to a vacancy in k is a function of the overall average probability unemployed workers make matches, the proportion of employment in job k, and the mobility friction, $\varphi_{k|j}$.

2.4 Firms

We assume the number of firms operating with type j jobs in city c is fixed, leaving endogenising firm formation for future work. For notational clarity, we drop the (firm-job-city-specific) subscript on firm employment and vacancies. All firms operating in a given industry have a common production function:

$$y_{fjc}(n) = \epsilon_{fic} n - \frac{1}{2}\sigma_i n^2,$$

where ϵ_{fic} is a firm-specific productivity draw, n is the number of employees, and $\sigma_i > 0$ is a parameter reflecting the potential span of control issues.[9] It will prove useful to write $\epsilon_{fic} = \epsilon_{ic} + u_{fic}$, where ϵ_{ic} corresponds to a sector-wide productivity shock (at the city level) and u_{fic} is a mean zero, firm-specific component. This specification implies that technology is common across cities within an industry but that there are differences across cities in comparative advantage in producing each good, captured in the ϵ_{ic}'s. We assume that the technology does not vary by union status. The literature on union effects on productivity seems to us to be inconclusive, and so we adopt an agnostic take in which unions affect firm activity by affecting wages (and employment) but not through technological adaptations.[10] We assume that the σ_i's are sufficiently smaller than 1 such that, combined with the assumption of a fixed number of firms in each ic cell, they imply that production of any good is spread across cities.

At the beginning of each period, firms choose the optimal number of vacancies (and, so, optimal employment) given the wage (specified as a function of firm employment) they know will be bargained with their workers later. To simplify, we assume that the flow cost of hiring is linear in the number of vacancies posted. Since δ^m matches are randomly destroyed in each period, a firm which had n_{-1} workers in the previous period enters the current period with $(1 - \delta^m)n_{-1}$ workers. From this, it knows the number of vacancies, v, it must post in order to have n workers for production in the current period. Hence, the firm value function of filled positions is given by:

[9] The production function includes a firm-job-city-specific subscript to account for the variations in firm employment, which differs across firms depending on their union status.

[10] Hirsch and Link [1984] and Addison and Hirsch [1989] summarise the early research in this area which finds largely inconclusive and mixed evidence on the effect of unionisation on productivity.

$$\begin{aligned}\Pi_{fjc}(n_{-1}) = \max_{v} \quad & [p_i y_{fjc}(n) - w_{fjc}(n)n - \kappa v + \rho^e \Pi_{fjc}(n)] \\ \text{s.t.} \quad & n = n_{-1}(1-\delta^m) + q_c^v v,\end{aligned} \tag{3}$$

where $w_{fjc}(n)$ is the wage bargained at the firm for this type of job with n workers at the firm, and κ is the cost per vacancy posted. ρ^e is the firm effective discount rate, taking account of the firm death rate, i.e. $\rho^e = (1-\delta^e)\rho$. Note that we assume that union amenities are created by the union and, so, do not enter the cost function of the union firm.

2.5 Workers

The value of employment for a worker in a job of type j in firm f is given by:

$$V_{fjc}^E(w_{fjc}) = w_{fjc} + \psi_{fjc} + \rho[\delta V_{jc}^U + (1-\delta)V_{fjc}^E(w'_{fjc})] \tag{4}$$

where ψ_{fic} is a non-wage amenity for workers from being in a union in this particular firm and, so, equals zero in non-union firms. w'_{fjc} is the wage that will be paid by the firm in the next period if the job is not terminated and δ is the job destruction probability; i.e. $\delta = \delta^e + (1-\delta^e)\delta^m$. Following TD, we assume that workers and firms believe that the next period's wage will be set optimally and that they cannot affect it through actions during this period. Thus, in a steady state, the agents treat the predicted wage for the next period as a constant. V_{jc}^U is the value of unemployment for a worker formerly employed in a job of type j and is given by:

$$V_{jc}^U = b + \rho\left[q_c^u \sum_k T_{kc|j} V_{kc}^E(w'_{kc}) + (1-q_c^u)V_{jc}^U\right] \tag{5}$$

where, b is the flow value of being unemployed, w'_{kc} is the average wage in job type k next period, and $T_{kc|j}$ is the probability a worker formerly in job type j in city c finds a job of type k, conditional on making any match. Based on our discussion of matching rates, $T_{kc|j} = \eta_{kc}\chi_{kc|j}(\varphi_{k|j}) = \eta_{kc}\frac{\varphi_{k|j}}{\sum_{k'}\eta_{k'c}\varphi_{k'|j}}$. We do not model the source of the differences in $T_{kc|j}$ by k and j, treating them as exogenous facts from the workers' perspectives. Thus, this is a model of random search with probabilities of a worker meeting specific jobs given by $T_{kc|j}$.[11]

Equation (5) says that the value of unemployment is higher when b is higher, when the probability of making a match (q_c^u) is higher, and when the expected value of the match, $\sum_k T_{kc|j}V_{kc}^E$, is higher. In Appendix C.3, we show that in steady state, $\sum_k T_{kc|j}V_{kc}^E$ can be written as $\Gamma_c + \mathbb{B}_c \sum_k T_{kc|j} w_{kc}$, where $\Gamma_c > 0$ and $\mathbb{B}_c > 0$ are functions of the matching rates and model parameters, and w_{kc} are average wages across firms offering jobs of type k in city c. With a slight abuse of terminology, we will refer to $\sum_k T_{kc|j} w_{kc}$ as the 'outside option value' of the worker. This outside option is higher if the local economy has a greater concentration in high-wage jobs that the worker has a relatively high probability of matching with (i.e., with high associated $T_{kc|j}$ values).

[11]Our model, therefore, abstracts away from issues related to workers queueing for union jobs (Abowd and Farber [1982]). This queuing mechanism could imply an additional spillover channel whereby the existence of union firms drives down vacancy-filling rates in the non-union sector, pushing up wages. The prevalence of queueing is likely driven by union wage premia and the relative likelihood of finding union work such that queuing effects are likely to enter through the outside option channel in our model.

2.6 Wage Bargaining

2.6.1 Union Firms

Wages in union firms are given by the solution to the Nash Bargaining condition:

$$\beta S_{fjc} = (1-\beta)n(V^E_{fjc} - V^U_{jc}) \quad \forall j = \{1, i\}, \tag{6}$$

where β is the bargaining weight. Following TD, in a unionised setting, the firm surplus equals:

$$S_{fjc}(n) = [\pi_{fjc}(n) + \rho^e \Pi_{fjc}(n)] - [\pi_{fjc}(0) + \rho^e \Pi_{fjc}(0)] \quad \forall j = \{1, i\}, \tag{7}$$

where $\pi_{fjc}(n)$ denotes current-period profits, $\Pi_{fjc}(n)$ is the value of the firm with n workers, and $\pi_{fjc}(0)$ and $\Pi_{fjc}(0)$ are the flow profits and value of the firm with no workers, respectively, reflecting the fact that if bargaining breaks down, the union will remove all the workers. At the point of bargaining, the number of workers in the firm is fixed and the hiring cost is sunk. For this reason, current period recruitment costs do not appear in the surplus expression.

In Appendix C.2, we show that for a union firm:

$$S_{fjc} = p_i y_{fjc} - w_{fjc} n + \rho^e (1-\delta^m) \frac{\kappa}{q^v_c} n \quad \forall j = \{1, i\}. \tag{8}$$

That is, it equals the profits lost from a shutdown plus the additional cost of hiring the entire optimal workforce in the following period.

On the right-hand side of (6) is the sum of workers' surplus, which is given by the gain to employment for all workers hired by the firm. Since the workers are identical, we use a specification that focuses on the total surplus and assume that the union members will all get an equal share of the part of the surplus captured by the union. This ignores issues related to seniority, for example.[12] Note that workers' surplus will depend on the value of unemployment and, through that, on $\sum_k T_{kc|j} w_{kc}$, the outside option value of the worker.

In Appendix C.4, we solve for the steady state wage written as a function of firm size, then solve for optimal firm size, substituting it into the bargained wage equation to arrive at our expression for the union wage. That expression is a non-linear function of q^u_c and q^v_c, the matching rates for unemployed workers and vacancies, respectively. BGS show that in a steady state, these matching rates can be written as simple functions of the city employment rate, ER_c, and we substitute those functions. To get to our empirical specification, we linearize the resulting wage expression with respect to the vector, $\mathbf{x} = \{\psi, b, p_i, \sum_k T_{kc|j} w_{kc}, ER_c, \epsilon_{fic}, \psi_{fjc}\}$. We take the linear approximation around a point $\mathbf{x_0}$ where employment is equally spread across industries (which occurs when the national mobility frictions are constant, i.e. when $\varphi_{k|j} = \varphi$) and the employment rate takes the same value in all cities (see Appendix D).

The linearized union wage expression is:

$$w_{fjc} = \tilde{\gamma}_{0i} + \tilde{\gamma}_1 \sum_k T_{kc|j} w_{kc} + \tilde{\gamma}_2 ER_c + \tilde{\gamma}_3 \epsilon_{ic} + \tilde{\gamma}_3 u_{fic} - \tilde{\gamma}_4 \psi_{fjc} \quad \forall j = \{1, i\}, \tag{9}$$

[12] See Abraham and Medoff [1984, 1985] who present evidence of the importance of seniority for layoffs and promotions, and see Abraham and Farber [1988] for evidence that the seniority wage profile is steeper under collective bargaining.

where $\tilde{\gamma}_{0i}$ is a function of the price p_i and constant terms stemming from the expansion point values. $\tilde{\gamma}_1$, $\tilde{\gamma}_2$, $\tilde{\gamma}_3$, and $\tilde{\gamma}_4$ are all positive. Expressions for each, written as functions of underlying parameter values, are given in Appendix D. Thus, union wages are a positive function of productivity (captured in $\tilde{\gamma}_{0i}$, u_{fic}, and ϵ_{ic}), the workers' outside option value ($\sum_k T_{kc|j} w_{kc}$), the tightness of the labour market, as reflected in ER_c, and a negative function of union amenities (ψ_{fjc}). As pointed out in BGS, in a frictionless environment, the wage would only be a function of productivity and the union amenity. In particular, the value of a worker's outside option would not play a role in wage determination.

2.6.2 Simple Non-union Firms

In simple non-union firms, the firm bargains with each worker individually, yielding wages that satisfy the bargaining rule:

$$\beta S_{fjc} = (1-\beta)\left(V^E_{fjc} - V^U_{jc}\right) \quad \forall j = \{2, i\} \tag{10}$$

As with union workers, the worker outside option is the value of unemployment, though the size of that option can differ because union and non-union workers have potentially different probabilities of accessing jobs of various types.[13] For firms, the fact that they are bargaining with one worker at a time means the firm surplus equals profits at the current firm size minus profits the firm would attain if it lost this one worker. In Appendix C.2 (following TD), we show that this implies that:

$$S_{fjc}(n) \quad = \quad \frac{\partial \pi_{fjc}(n)}{\partial n} + \rho^e(1-\delta^m)\frac{\kappa}{q^v_c}, \quad \forall j = \{2, i\} \tag{11}$$

where $\frac{\partial \pi_{fjc}(n)}{\partial n} = p_i \frac{\partial y_{fjc}(n)}{\partial n} - w_{fjc}(n) - n\frac{\partial w_{fjc}(n)}{\partial n}$. This reflects the fact that a breakdown in bargaining would cost the firm the marginal profit from losing one worker plus the cost of having to hire one additional worker the following period.

Solving for the simple non-union wage as for the union wage and again taking a linearization leads to our simple non-union wage expression:

$$w_{fjc} = \gamma_{0i} + \gamma_1 \sum_k T_{kc|j} w_{kc} + \gamma_2 ER_c + \gamma_3 \epsilon_{ic} + \gamma_3 u_{fic} \quad \forall j = \{2, i\} \tag{12}$$

As with the union equation, γ_{0i} is a function of the price p_i and constant terms stemming from the expansion point values, and the other coefficients are all positive. Expressions for each of these coefficients are given in Appendix D. Importantly, $\tilde{\gamma}_{0i} > \gamma_{0i}$ and, so, union wages within an industry are on average higher, reflecting the fact that union wages are proportional to total product while non-union wages are proportional to the marginal product of a worker and the latter is smaller if there are span of control issues. More intuitively, unions can

[13] A referee correctly pointed out that union and non-union workers may also have different job separation rates. Appendix F shows that allowing for this implies that the coefficients in our linearized wage equations should differ between the union and non-union wage equations. Since we estimate the non-union wage equation on its own (i.e., not together with the union equation), our specification allows for such differences and no further adjustments are needed to take account of differences in separation rates.

bargain higher wages because they can threaten to withdraw the whole labour force, while a non-union worker can only threaten to withdraw her labour. In addition, $\tilde{\gamma}_3 > \gamma_3$, i.e., unions can capture a greater share of productivity shocks than non-union workers.

2.6.3 Emulating Non-union Firms

In emulating non-union firms, the bargaining situation fundamentally differs from the union and simple non-union cases. In those cases, the outside options have to do with the outcomes of separation if bargaining breaks down: the value of unemployment for workers and the loss of total or marginal production for firms. But in the emulating non-union firms, workers would choose to unionise if left to make that decision unilaterally (a decision we will return to in the next section), and the bargaining is over a wage that would make workers not want to pursue that choice. Thus, the outside options have to do with what each side would get in the union environment, which is the outcome if this bargaining breaks down.

We assume that in a steady state, firms know they will be an emulating non-union firm and that they choose their number of workers based on that. We also assume that they would not be allowed to change this number of workers in the union alternative. This is legally the case (firms cannot lay off workers for organizing a union) but, of course, firms do break this rule to a considerable extent, resulting in the filing of Unfair Labour Practices (ULP) complaints by workers to the NLRB [Bronfenbrenner, 2009]. We begin with an assumption that workers face a one-time cost of unionising a workplace, λ_c^*. In Section 2.7.1, we discuss the case where firms can act to raise λ_c^* through intimidation tactics.

Given these assumptions, the wage solves the bargaining problem:

$$\beta S_{fjc} = (1-\beta)n\left[V^E_{fjc} - (V^E_{f1ic} - \lambda_c^*)\right] \quad \forall j = \{3, i\}, \tag{13}$$

where the worker surplus on the right-hand side equals the difference between the value of being in an emulating non-union job relative to the value of being in a union job minus the cost of unionising, and where n denotes the number of workers in the emulating firm. The firm surplus, S_{fjc}, is given by:

$$S_{fjc} = \left[\pi_{fjc}(n) + \rho^e \Pi_{fjc}(n)\right] - \left[\pi^*_{f1ic}(n) + \rho^e \Pi^*_{f1ic}(n)\right] \quad \forall j = \{3, i\}, \tag{14}$$

where the first term in the brackets reflects the discounted profits following successful bargaining. The second term in the brackets captures the discounted profits if the negotiation breaks down, i.e., if the firm is unionised (but employs n workers). In Appendix C.2, we show that $S_{f3ic} = \frac{1}{1-\rho^e}\left[\pi_{f3ic}(n) - \pi^*_{f1ic}(n)\right]$, i.e. the firm surplus is given by the difference in profits when operating as an emulating non-union firm versus as a unionised firm.

Solving the bargaining problem yields the wage expression:

$$w_{fjc} = w_{f1ic} + \bar{\xi}\left[(\mathbb{A}\psi_{f1ic} - \lambda_c^*) - \Delta_{jc,1ic}\right] \quad \forall j = \{3, i\}, \tag{15}$$

where $\bar{\xi}$ and $\mathbb{A}$ are positive functions of model parameters and $\Delta_{jc,1ic}$ is a function of differences in transition rates to other firms between non-union workers and union workers that drops out in the linearization step, with the specific forms of each given in Appendix C.4. This expression says that the emulating non-union wage equals the union wage plus an adjustment that is positively related to any amenities the union would create and negatively related to the cost of unionisation. Substituting in our expression for the union wage, w_{f1ic}, and linearizing, we arrive at:

$$w_{fjc} = \tilde{\gamma}_{0i} + \tilde{\gamma}_1 \sum_k T_{kc|1i} w_{kc} + \tilde{\gamma}_2 ER_c + \tilde{\gamma}_3 \epsilon_{ic} + \tilde{\gamma}_3 u_{fic} - (\tilde{\gamma}_4 - \bar{\xi}\mathbb{A})\psi_{f1ic} - \bar{\xi}\lambda_c^* \quad \forall j = \{3, i\}, \tag{16}$$

where $\tilde{\gamma}_4 - \bar{\xi}\mathbb{A} > 0$. Notice that the emulating non-union wage depends on $\sum_k T_{kc|1i} w_{kc}$, the outside option value for union workers since when that is higher (when, for example, there is a lot of high wage jobs accessible to union workers), emulating firms are forced to pay a higher wage to prevent workers from unionising.

2.7 Union Arrangement Determination

Firms, in our model, are born non-union. The firm hires its optimal set of workers, who then decide whether to proceed toward unionisation. The standard NLRB certification procedure starts with determining whether enough workers are potentially interested in a union to proceed to an election. Legally, an election can be initiated if at least 30% of workers sign a card indicating an interest in having a union, though, in practice, unions do not seek to open an election unless they have cards signed by at least 50% of the workers [DiNardo and Lee, 2004]. Thus, we can think of the first step in the process as being determined in a median voter model, in which if the median worker does not want a union, then neither the union nor the firm puts further effort into union determination. The median voter will compare the value of being employed when the firm is non-union to the value when it is union minus the cost of unionisation, λ_c^*, i.e. the firm remains non-union without any further actions if:

$$V_{f2ic}^E - (V_{f1ic}^E - \lambda_c^*) > 0. \tag{17}$$

In Appendix E, we show that substituting expressions for the values of employment results in an index function:

$$I_{fic} = \mathbb{A}(w_{f1ic} - w_{f2ic}) + (\mathbb{A}\psi_{f1ic} - \lambda_c^*) + \rho\delta\mathbb{A}(V_{f1ic}^U - V_{f2ic}^U), \tag{18}$$

such that not enough workers sign cards to proceed with the certification process (or, in our terms, the firm remains a simple non-union firm) if $I_{fic} \leq 0$. Hence, whether workers want a union depends positively on the union-non-union wage differential, the difference between the outside options of their respective workers and the difference between union amenities and the cost of unionising. Those costs could include organizing efforts on the part of the union or individual workers and barriers erected by the firm to discourage workers. The level of both types of costs will vary with the local legal environment, which is very heterogeneous across states [Clark and Johnston, 1987] and, so, we index them by c.

Substituting in the linearized wage expressions and linearizing the values of unemployment results in a rewriting of the index function as:

$$I_{fic} = \alpha_{0i} + \alpha_1 \sum_k T_{kc|1i} w_{kc} - \alpha_2 \sum_k T_{kc|2i} w_{kc} + \alpha_3 ER_c + \alpha_4 \epsilon_{fic} + \alpha_5 \psi_{f1ic} - \lambda_c^*, \tag{19}$$

recalling that $\epsilon_{fic} = \epsilon_{ic} + u_{fic}$, and where the αs are constant terms obtained from the linear approximation. In particular, $\alpha_1 > 0$, $\alpha_2 > 0$, $\alpha_3 < 0$, $\alpha_4 > 0$ and $\alpha_5 > 0$ (see Appendix E).

Note that unionisation is increasing in the productivity shock because unions can capture a larger proportion of that shock.

Rearranging equation (19), we arrive at a threshold value for firm-specific union amenities, ψ^*_{fic}, such workers in firms with $\psi_{f1ic} \leq \psi^*_{fic}$ remain happily non-union. Specifically, as shown in Appendix E,

$$\psi^*_{fic} = \frac{1}{1-\tilde{\gamma}_4}\lambda_c - \Delta^*_{1ic,2ic} - \frac{\alpha_4}{\alpha_5}\epsilon_{fic}, \tag{20}$$

where $\lambda_c = \frac{1}{\mathbb{A}}\lambda^*_c$ is an annualized version of the fixed unionisation cost, $\Delta^*_{1ic,2ic} = \frac{\alpha_{0i}}{\alpha_5} + \frac{\alpha_1}{\alpha_5}\sum_k T_{kc|1i}w_{kc} - \frac{\alpha_2}{\alpha_5}\sum_k T_{kc|2i}w_{kc} + \frac{\alpha_3}{\alpha_5}ER_c$, and where $0 < \frac{1}{1-\tilde{\gamma}_4} < 1$, $\frac{\alpha_4}{\alpha_5} > 0$, $\frac{\alpha_{0i}}{\alpha_5} > 0$, $\frac{\alpha_1}{\alpha_5} > 0$, $\frac{\alpha_2}{\alpha_5} > 0$, $\frac{\alpha_3}{\alpha_5} < 0$. Notice that the value of the threshold varies with ϵ_{fic} – a point we return to below.

When $\psi_{f1ic} > \psi^*_{fic}$, the workers at a firm will want to unionise. But, as we discussed earlier, depending on the values of union amenities and costs of unionising, it may be the case that the surplus to the match is larger if the firm remains non-union. In that case, a new wage is bargained, ensuring workers no longer want to unionise. In discussing these middle, emulating firms, it is useful to define a second threshold, ψ^{**}_{fic}, such that if $\psi_{f1ic} > \psi^{**}_{fic}$, the total surplus is larger if the firm becomes union and the firm will be unionised.[14] In Appendix E, we show that $\psi^{**}_{fic} = \lambda_c$.

Recall that ψ^*_{fic} is a function of ϵ_{fic}. For a low enough value of ϵ_{fic}, $\psi^{**}_{fic} \leq \psi^*_{fic}$, i.e., there will be no emulating non-union firms. In essence, given that union wages rise faster with ϵ_{fic} than non-union wages, for a low enough value of ϵ_{fic}, there is not enough room between union and non-union wages to permit an intermediate, emulating non-union wage that can forestall unionisation. In Appendix E, we solve for a threshold value ϵ^*_{ic} such that if $\epsilon_{fic} \leq \epsilon^*_{ic}$ there are no non-union emulators in the industry-city cell.

To summarize, for values of $\epsilon_{fic} \leq \epsilon^*_{ic}$, firms will be non-union $\forall\, \psi_{f1ic} \leq \psi^*_{fic}$ and unionised if $\psi_{f1ic} > \psi^*_{fic}$. If $\epsilon_{fic} > \epsilon^*_{ic}$, firms will be non-union if $\psi_{f1ic} \leq \psi^*_{fic}$, emulators if $\psi^*_{fic} < \psi_{f1ic} \leq \psi^{**}_{fic}$, and unionised for amenity values above ψ^{**}_{fic}. Table 1 summarizes the various statuses described above.

2.7.1 Intimidation Responses

Recent decades have seen a high and rising level of illegal interventions by firms in union certification drives [Bronfenbrenner, 2009]. We can model this as firms responding to a unionisation threat by increasing the cost of unionising, λ^*_c, at a cost to themselves. For example, they could fire union organizers, implement mandatory information sessions and otherwise seek to intimidate workers.

In Appendix E.3, we set out the value function for a firm employing intimidation and compare it with the value function when it chooses the wage emulation response. We show that incorporating the possibility of intimidation in the model introduces an additional amenity threshold, ψ^b_{fic}, which determines the intimidation status. Specifically, when $\psi_{f1ic} \in$

[14] In standard union threat analyses, the firm is assumed to be the agent that decides on whether to pay an emulating wage to forestall unionisation. In that case, the equivalent of ψ^*_{fic} is set to make the median worker just indifferent between unionising or remaining non-union. Instead, we have framed the decision as a bargaining problem so that the threshold reflects a division of the surplus.

Table 1: Summary of Union Statuses

	$\epsilon_{fic} \leq \epsilon^*_{ic}$	$\epsilon_{fic} > \epsilon^*_{ic}$
$\psi_{f1ic} \leq \psi^*_{fic}$	Simple non-union	
$\psi_{f1ic} > \psi^*_{fic}$	Union	
$\psi_{f1ic} \leq \psi^*_{fic}$		Simple non-union
$\psi^*_{fic} < \psi_{f1ic} \leq \psi^{**}_{fic}$		Emulating non-union
$\psi_{f1ic} > \psi^{**}_{fic}$		Union

$[\psi^*_{fic}, \psi^b_{fic}]$, unionisation represents only a marginal gain for workers, leading the firm to prefer intimidation over allowing unionisation or responding through emulation tactics. However, as amenities increase and the benefits of unionisation grow, the cost of intimidation rises, making emulation increasingly advantageous for the firm. Therefore, beyond the threshold ψ^b_{fic}, it becomes optimal for the firm to co-opt workers by offering higher wages rather than relying on increasingly large threats. Finally, for $\psi_{f1ic} > \psi^{**}_{fic}$, the surplus of the firm in the emulation status turns negative, and allowing workers to unionise dominates any firm response.

While the model suggests the existence of one additional type of firm, it is not clear that there is any advantage in separating simple non-union workers and intimidated non-union workers in practice since we are just studying the overall non-union wage, and both of these subgroups receive the same wage. Moreover, nothing in our data allows us to identify the simple non-union workers from the workers who are remaining non-union only because of threats. Finally, the thresholds, ψ^*_{fic} and ψ^b_{fic}, are functions of the same variables – ϵ_{fic}, λ^*_c, the expected rents, and the employment rate – and, so, there is no impact on our empirical specifications of including or not including intimidation. Therefore, we will proceed as if there is no intimidation to simplify the exposition.

3 Empirical Specification

We are now in a position to set out our empirical specifications. We begin with the simple case of non-union wages in a context in which workers choose whether to unionise, and there is no firm response (i.e., $\tau = 3$ does not exist) because that generates a specification that is more similar to the existing literature and provides a straightforward venue for discussing identification issues.[15] We then proceed to our complete specification, incorporating an emulating wage response.

[15] In addition, as we describe in Section 4, we only have the data for the full specification for a subset of our years, and we want to present some results for the entire period.

In the simple situation, the non-union wage equation is given by (12). We work at the industry×city cell level and estimate our derived specification in first differences to eliminate any industry by city time-invariant characteristics. Thus, our base specification is given by:

$$\Delta \ln w_{2ict} = \Delta\gamma_{0it} + \gamma_1 \Delta E_{2ict} + \gamma_2 \Delta ER_{ct} + \gamma_3 \Delta\epsilon_{ict} + \gamma_3 \Delta\bar{u}_{2ict}. \tag{21}$$

where we have now introduced a time subscript; $\Delta x_t = x_t - x_{t-1}$; E_{2ict} is our shorthand notation for the outside option value for non-union workers and equals $\sum_k T_{kct|2i} w_{kct}$; and $\bar{u}_{2ict}$ is the mean value for the firm-specific productivity, which equals zero in a random sample of firms but, as we discuss below, not when there is systematic selection into union status. The error term is $\Delta\tilde{u}_{2ict} = \gamma_3 \Delta\epsilon_{ict} + \gamma_3 \Delta\bar{u}_{2ict}$. We view the different time periods (which are a decade apart in our data) as corresponding to different steady states with different draws on productivity shocks, amenity values, and the cost of creating a union.

Equation (21) says that non-union wages are increasing in the value of the outside option, fitting with results in, among others, Beaudry et al. [2012], Tschopp [2017], Caldwell and Danieli [2021], Jarosch et al. [2019], and Bassier [2022]. The value of the outside option varies across cities and is a function of three factors. The first is η_{kc}: the proportion of employment accounted for by a given job type, defined by the combination of union status and industry, in the city. The second is the wage rate paid in that job type in the city, w_{kc}. Importantly, to this point, we have assumed that workers are homogeneous, implying that wage differences across job types correspond to rents – differences in pay over and above what is required for the marginal worker to want to join that industry. Those rents are maintained because of the frictions in the labour market. It is important that we consider rents since wage differences across industries that correspond to compensating differentials or skill differentials cannot be the basis of bargaining a higher wage with your current employer. The third factor driving the outside option value is the ease with which an employee from job type j can transit to a job of type k. This includes the ease of moving from non-union to union jobs. Thus, for workers in job type j, a city will have higher wages if it has more jobs that workers can actually access that pay high rents.

We can write E_{2ict} as,

$$E_{2ict} = \sum_k T_{kct|2i} w_{kct} = \underbrace{\sum_{i'} T_{1i'ct|2i} w_{1i'ct}}_{\text{Union component}} + \underbrace{\sum_{i'} T_{2i'ct|2i} w_{2i'ct}}_{\text{Non-union component}} \tag{22}$$

where the first component of the outside option is associated with potential union jobs and the second term with potential non-union jobs. We will refer to the first component in equation (22) as $E_{1ct|2i}$ and the second component as $E_{2ct|2i}$. Our theory says that the two components should have an equal effect on bargained wages since it doesn't matter to the employer in what specific sector a worker's improved outside options arise. We use this fact as the basis of an over-identification test of the model in a specification in which we replace E_{2ict} with $E_{1ct|2i}$ and $E_{2ct|2i}$.

The wage is also predicted to increase in ER_c since a tighter labour market implies that workers can access their alternative options more easily. The specification also indicates that the regression should include a complete set of industry effects, reflecting differences in the output prices. This means that the relevant identifying variation for the estimated coefficients comes from across-city within-industry variation. Intuitively, we identify the impact of outside options by comparing wage changes in the same industry in two different cities that are experiencing different changes in the quality of outside employment prospects, holding the employment rate constant. For example, we could compare construction workers in Pittsburgh in the 1980s, when the decline of big steel meant a decline in the possibility of high-rent jobs, to construction workers in a city not substantially altering its sectoral composition, predicting larger wage declines in Pittsburgh.

3.1 Implementation and Identification Challenges

In estimating (21), we face a number of potential implementation and identification problems.

3.1.1 Worker Heterogeneity

To take the model to the data, we must confront the fact that while workers are homogeneous in our model, they are heterogeneous in our data. Our approach is to treat individuals as representing different bundles of efficiency units of work and to assume those bundles are perfect substitutes in production. We then interpret w_{f2ict} in equation (12) as the cost per effective labour unit. Let effective labour units be $\exp(H_l'\beta_t + a_l)$, where H_l and a_l capture observable and unobservable skills of worker l, respectively. Adding industry, city and time subscripts, workers' observed non-union log wages, $\ln w_{l2ict}$, are given by:

$$\ln w_{l2ict} = H_{lt}'\beta_t + \ln w_{2ict} + a_{lt}. \tag{23}$$

The $\ln w_{2ict}$ values are our object of interest. To obtain a measure of these, we estimate (23) capturing $\ln w_{2ict}$ as the coefficients on a complete set of job$\times$city fixed effects under an assumption that a_l is orthogonal to job$\times$city effects. Our specification of H_l includes a complete interaction of dummies for educational attainment, a quadratic in potential experience, and gender and race dummy variables. We estimate (23) using only non-union workers, separately by year. This allows for flexible changes in the returns to education and other observable characteristics over time. The estimated vector of coefficients on the job-city fixed effects are regression-adjusted, with average local job type wages, and we use these coefficients as the dependent variable in our regressions.[16]

3.1.2 Reflection Problem

As specified, the regression (21) incorporates a standard reflection problem. Its dependent variable is a job-city cell wage and the right-hand side includes E_{2ict}, which is a weighted average of local job type wages. In a related model, BGS show that one can replace the local wages, w_{kct}, with national-level, sector-specific wages, w_{kt}, in the outside option variable in a model-consistent way. As described earlier, it is important that the outside option variables are constructed from rents. To meet that requirement, we implement a similar exercise to constructing our dependent variable, regressing log wages on the same set of skill and demographic variables (H_l) plus a complete set of job-type dummy variables. We refer to

[16] After differencing (12), we divide both sides of the equation by a base wage, w_0, in order to transform the dependent variable to log differences.

the coefficients on those dummy variables as industry-union status-specific rents, ν_{kt}, and use them to form our instrumental variables (described next). The results are instruments that do not depend on local wage variation, breaking the link that forms the reflection problem.

3.1.3 Endogeneity

One advantage of deriving our empirical specification from a model is that it allows us to understand what is in the error term and what that implies for both endogeneity problems and solutions. In our case, the error term contains changes in industry×city productivity shocks, $\Delta\epsilon_{ict}$. We would expect that these productivity changes would be related to changes in the size of different jobs in the city, which, in turn, would alter the size of the outside option value, E_{2ict}. We would also expect them to be related to labour market tightness, which we capture with the ER_{ct} variable in our regression.

We respond to the endogeneity of E_{2ict} using a Bartik instrument strategy that can be best understood by writing out the definition of a slightly altered version of the outside option value more completely:[17]

$$\tilde{E}_{2ict} = \sum_{k \neq 2i} \eta_{kct} \frac{\varphi_{kt|2i}}{\sum_{k'} \eta_{k'ct}\varphi_{k't|2i}} \nu_{kt}, \tag{24}$$

Note, first, that we have replaced w_{kct} with ν_{kt}, which is the job type wage premium at the national level. The fact that they are at the national level breaks the local reflection problem. Using $\tilde{E}_{2ict}$ alters our outside option prediction slightly to say that we predict wages will be higher for workers in an industry i in city c if there are relatively high proportions of jobs in high-rent sectors to which the worker has effective access. Second, in constructing our instruments, we also take a 'leave one out' approach, dropping the job type defined by the combination $j = \{\tau, i\} = \{2, i\}$ to ensure the instrument is not getting its power from the sector we are focusing on.

Recall that with the inclusion of industry effects in our specification, the identifying variation is across cities within industry. That implies that the $\varphi_{kt|j}$ and ν_{kt} terms in (24) are not a concern in terms of being correlated with the cross-city variation in $\Delta\epsilon_{ict}$. However, the job shares (the η's) seem likely to be related to the size of sector-specific shocks (the ϵ's). To address this concern, we form predicted values for employment in a type k job cell in a city at the end of a decade using employment in that cell at the start of the decade times the growth rate in type k jobs at the national level. We then use those predicted employment levels to form predicted shares, $\hat{\eta}_{kct}$ and with those, we form the predicted end-of-decade option value:

$$\hat{E}_{2ict} = \sum_{k \neq 2i} \hat{\eta}_{kct} \frac{\varphi_{kt|2i}}{\sum_{k'} \hat{\eta}_{k'ct}\varphi_{k't|2i}} \nu_{kt} \tag{25}$$

Using that, we form our instrument for ΔE_{2ict}:

$$IV1_{2ict} = \hat{E}_{2ict} - \tilde{E}_{2ict-1} \tag{26}$$

The cross-city variation in this instrument comes from the η_{kct-1}'s since $\hat{\eta}_{kct}$ is a function of η_{kct-1} and all the other terms in (26) do not vary at the city level. Thus, the instrument's

[17] Recalling that $T_{kct|j} = \eta_{kct}\chi_{kct|j} = \eta_{kct}\frac{\varphi_{kt|j}}{\sum_{k'} \eta_{k'ct}\varphi_{k't|j}}$.

validity requires that the η_{kct-1}'s are independent of the relevant variation in the error term: cross-city variation in productivity growth. That is, we require an assumption that the productivity process follows a random walk (since, as BGS shows, the η_{jct}s can be written as functions of the ϵ_{ict}s). We can assess this assumption using an over-identification test, which we discuss when we present our results.[18] We also present indicative tests from Goldsmith-Pinkham et al. [2020]. We use $IV1_{2ict}$ in specifications where we only include ΔE_{2ict} on the right-hand side of the regression and construct union and non-union specific versions ($IV1_{1ct|2i}$ and $IV1_{2ct|2i}$) using only union or non-union sector ν's, φ's, and η's, respectively, when we include separate union and non-union outside option values.

We do not instrument for ΔER_{ct} even though there are clear reasons to assume it is correlated with our error term. We follow Stock and Watson [2011] in interpreting the employment rate as a control variable – a variable that is not of direct interest in its own right but is useful for picking up its own effect and those of correlated omitted variables. In our case, we view the employment rate as capturing its own effect plus the impact of general, local demand shifts. This allows us to isolate the outside option effects we care about from general demand effects. The required identifying assumption is that $\Delta\tilde{u}_{2ict}$ is conditionally mean independent of $IV1_{2ict}$, i.e., that the instrument is independent of the error term once we condition on the control variable ($E(\Delta\tilde{u}_{2ict}|IV1_{2ict}, \Delta ER_{ct}) = E(\Delta\tilde{u}_{2ict}|\Delta ER_{ct})$). Stock and Watson [2011] show that if this condition is met, then the coefficient on ΔE_{2ict} is consistent for the causal effect of ΔE_{2ict} on $\Delta \ln w_{2ict}$ while the coefficient on ΔER_{ct} does not have a causal interpretation. They also show that standard IV inference results, such as weak instrument tests, are valid under the conditional mean independence assumption. We have already argued that our instrument is independent of the error term. Treating ΔER_{ct} as a control strengthens the argument by ensuring that city-level productivity changes are being controlled for. In our estimation, we implement one specification to go one step further and include a complete set of city×time effects. This will soak up city-level productivity effects of the type represented by ΔER_{ct} and other potential confounders such as city-level changes in housing supply. Including these effects is demanding of the data and increases standard errors, but even so, we will see that their inclusion does not materially alter our key estimated coefficients.

3.1.4 Selection

Estimation of (21) also includes a classic selection problem. In particular, the conditional mean of the non-union wage for workers at firms that are actually observed to be non-union is given by:

$$E(w_{f2ict}|I_{fict} \leq 0) = \gamma_{0it} + \gamma_1 E_{2ict} + \gamma_2 ER_{ct} + \gamma_3 \epsilon_{ict} + \gamma_3 E(u_{fict}|I_{fict} \leq 0), \qquad (27)$$

where I_{fict} (given in (19)) is the index function capturing worker decisions on unionisation in a simple world with no firm response. Because the coefficient on the firm productivity

[18] As Goldsmith-Pinkham et al. [2020] and BGS point out, Bartik instruments are functions of the start of period values for the η_{ict}'s – the local industrial composition – and any combination of those values can be used as an instrument. BGS argue that in our case, one can find specific combinations within the theory by examining decompositions of the outside option variables that both have intuitive appeal and imply testable over-identifying restrictions.

term is positive in the index function, and recalling that ψ_{f1ic} and ϵ_{fic} are assumed to be independent, union firms tend to have higher productivity. Thus, the marginal firms that would be unionised in a city, c, but non-union in another city, c', where the costs of unionisation are higher, will be at the low end of the productivity range for union firms but the high end for non-union firms. This has implications for common specifications using the proportion of unionised workers to capture spillover effects since the estimated coefficient on the union proportion would be biased downward. Industry-city cells with higher unionisation rates would be ones with lower productivity among non-union firms.

We address selection through a generalized Heckman two-step approach (see Heckman [1979], Snoddy [2019]). The idea in this approach is that the error mean term in (27), $E(u_{fict}|I_{fict} \leq 0)$, creates an omitted variables bias that can be addressed by including a control function. In particular, since $E(u_{fict}|I_{fict} \leq 0)$ can be expressed as a non-linear function of the probability of selection (the probability of being non-union in our case), the relevant control function can be a polynomial in that probability or of exogenous variables that drive that probability.

Given these arguments, we examine potential selection effects using two sets of variables. First, we include a quadratic in ΔP_{ict}, the change in the proportion of unionised workers in the industry×city×time cell. In doing this, we are taking the model very seriously in the sense that it does not indicate a reason to include a function of the union proportion in its own right, so its inclusion can be interpreted as capturing selection effects. According to the model, access to union jobs enters as transition rates in the outside option values. Following Fortin et al. [2019], we also estimate specifications in which we proxy for costs of unionisation using NLRB data on certification elections. In particular, we calculate the number of workers involved in certification elections that resulted in union certification in each city over three-year windows around the years 1980-1990-2000-2010, divided by the number of non-union workers in the city. We view low values of this variable as reflecting higher costs of unionisation at either the stage of initiating an election or of winning the election. We include a quadratic in this variable as additional regressors to address selectivity.

3.2 Specification Including Wage Emulation Response

Having set out the simple specification, we now turn to our more complete specification incorporating wage responses to the threat of unionisation. Importantly, in our data, we can only see whether a worker is union or non-union, not the type of non-union firm they work for. As a result, our observed dependent variable (the non-union wage) is a weighted average of the wages in the two types of firms:

$$w_{ict}^{n} = P_{ict}^{ne} \cdot w_{3ict} + (1 - P_{ict}^{ne}) \cdot w_{2ict}, \tag{28}$$

where w_{ict}^{n} is the observed mean non-union wage in industry i in city c, and P_{ict}^{ne} is the probability a firm is an emulating non-union firm conditional on it being a non-union firm of either kind. For w_{2ict} and w_{3ict}, we substitute expectations over firms in the ic cell of the relevant firm level equations ((12) and (16), respectively). Taking those expectations implies including error mean terms reflecting selection of the same form as in equation (27). Thus, we obtain:

$$
\begin{aligned}
w^n_{ict} = \gamma_{0it} + \tilde{\gamma}_1 P^{ne}_{ict} E_{1ict} + \gamma_1(1 - P^{ne}_{ict}) E_{2ict} + \gamma_2 ER_{ct} + \gamma^*_{0it} P^{ne}_{ict} \\
- \gamma_6 P^{ne}_{ict} \lambda_{ct} + \mu_{ict} + \tilde{\gamma}_3 P^{ne}_{ict} \epsilon_{ict} + \gamma_3(1 - P^{ne}_{ict}) \epsilon_{ict},
\end{aligned} \tag{29}
$$

where $\gamma^*_{0it} = (\tilde{\gamma}_{0it} - \gamma_{0it})$, $\gamma_6 > 0$ and μ_{ict} is the error mean term capturing selection of firms into $\tau = 2$ or $\tau = 3$ status, $E_{1ict} = \sum_k T_{kct|1i} w_{kct}$ is the outside option for union workers, and $E_{2ict} = \sum_k T_{kct|2i} w_{kct}$ is the outside option for non-union workers.[19] Once again, we present the complete derivation and the form of μ_{ict} in Appendix E.2.2.

As with the simple specification, we difference at the decadal level, dividing through by a base wage so that we are working with log wages:

$$
\begin{aligned}
\Delta \ln w^n_{ict} = {} & \Delta\gamma_{0it} + \gamma_1 \Delta((1 - P^{ne}_{ict}) E_{2ict}) + \gamma_2 \Delta ER_{ct} \\
& + \tilde{\gamma}_1 \Delta(P^{ne}_{ict} E_{1ict}) + \Delta(\gamma^*_{0it} P^{ne}_{ict}) - \gamma_6 \Delta(P^{ne}_{ict} \lambda_{ct}) + (\tilde{\gamma}_3 - \gamma_3) \epsilon_{ict} \Delta P^{ne}_{ict} \\
& + \Delta\mu_{ict} + \gamma_3 \Delta\epsilon_{ict} + (\tilde{\gamma}_3 - \gamma_3) P^{ne}_{ict-1} \Delta\epsilon_{ict}.
\end{aligned} \tag{30}
$$

The three lines in equation (30) correspond to three sets of influences on the observed non-union wage. The first line contains the same factors as in the simple specification. In particular, as before, there are a complete set of industry effects (implying that our identifying variation is across cities within industries). We also expect non-union wages to increase in the outside option of non-union workers, E_{2ict}, and in labour market tightness, as represented by the employment rate. The only adjustment is that the effect of the outside option is, in principle, reduced as the size of the union threat increases and firms shift away from the simple non-union type.

The second line contains factors related to changes in the union threat (ΔP^{ne}_{ict}). The impact of an increase in the threat probability will be higher when union wages (and, as a consequence, emulating firm wages) are higher, and the terms on the second line correspond to reasons why union wages might be higher than non-union wages. Union wages are higher, in part, because of the basic bargaining environment emphasized in the model – workers have more bargaining power when they organize. That means that union workers are able to capture a bigger proportion of rents, as reflected in the $\Delta(\gamma^*_{0it} P^{ne}_{ict})$ terms (which corresponds to union wages capturing a bigger proportion of p_i, price differences across industries) and $(\tilde{\gamma}_3 - \gamma_3)\epsilon_{ict}$ (which corresponds to unions capturing a larger share of industry×city specific rents, ϵ_{ict}). In addition, union wages are higher because union workers' outside options are better than those of non-union workers. In our data, union workers are much more likely to access union jobs than non-union workers, and, as a result, their outside option value is larger and moves differently from that of non-union workers. If the union workers' outside option value were to be a significant determinant of the non-union wage, we would view this as strong evidence in favour of the threat of unionisation affecting non-union wage setting. In contrast to these forces, the effects of increases in the threat probability are mitigated in higher cost of unionising environments (as reflected in the $\gamma_6 \Delta(P^{ne}_{ict} \lambda_{ct})$ component).

[19] We assume that hiring firms can distinguish between whether a worker's previous job was union or non-union but not whether it was simple or emulating non-union. Based on this, workers from both types of non-union firms face the same outside option, E_{2ict}. We also show in Appendix E.2.2 that at reasonable values for the underlying structural parameters, $\gamma_2 \approx \tilde{\gamma}_2$, so we don't include interactions of ER_{ct} with P^{ne}_{ict}.

The third line of (30) contains changes in firm selection into non-union status and terms capturing changes in sectoral productivity. The presence of those terms implies that we face the same implementation challenges as in our simple specification. In particular, we again use functions of the probability of unionisation to address selection issues. As before, we face a potential endogeneity issue based on the relationship of changes in outside option values to the $\Delta\epsilon_{ict}$'s. We form the same outside option instruments, using industry wage premia to eliminate the reflection problem. The outside option value for union workers, E_{1ict}, and the instrument that matches it are generated in the same way as the outside option value for non-union workers except that the transition rates correspond to the probabilities of transiting to a type k job from a union rather than from a non-union job.

The key new addition in the full specification (apart from the inclusion of the union worker outside option) is the presence of P^{ne}_{ict} throughout the equation. The variable P^{ne}_{ict} corresponds to the proportion of non-union firms that are under direct threat of unionisation. We assume the size of that threat is proportional to the proportion of firms that experience successful new unionisation drives and, so, approximate P^{ne}_{ict} with the number of firms successfully unionised divided by the number of non-union firms in the industry-city cell in the previous four years. The idea is that when that proportion is higher, the threat of unionisation is more present, and a larger proportion of firms have to pay a higher wage to emulate union wages and dissuade their workers from unionising. This is the third of our unionisation variables. Recall that we capture movements in the probability that workers who are switching jobs can move into a union job (through the φ's), movements in the proportion of workers who are unionised (through P_{ict}), and now movements in the probability a firm will face a successful union campaign. Each represents a specific way de-unionisation affects observed mean non-union wages (through outside option values for the workers, selection effects, and the threat of unionisation for the firm, respectively). While all three are related, we will see that they move differently over time and by city.

In our implementation of (30), we also include a complete set of interactions of P^{ne}_{ict-1} with industry×time effects and of city×time effects to capture the $\Delta(\gamma^*_{0it}P^{ne}_{ict})$ and $\gamma_6\Delta(P^{ne}_{ict}\lambda_{ct})$ components of the specification. We work with P^{ne}_{ict-1} rather than ΔP^{ne}_{ict} to avoid endogeneity issues. We think of the interactions of the lagged proportion unionised with industry dummies as the equivalent of including a Bartik instrument that distributes national-level changes at the industry level to cities based on their initial levels of union activity at the local level.

We are concerned about the endogeneity of P^{ne}_{ict} because unions are expected to capture a larger share of rents than non-union workers, leading to an anticipated increase in union activity when there are positive rent shocks. To better understand our response to this endogeneity issue, discussing the process of determining the threat of unionisation is helpful. We view the threat facing a firm as depending on four factors. The first, as we just stated, is its level of rents. The remaining factors are the level of organizing activity of unions that operate in its sector or locality; the demographic makeup of its workforce, as some groups may be more easily organized [Card et al., 2018b]; and the regulatory environment. We construct one set of instruments based on the organizing activity in the firm's sector and location. We cannot use changes in union organizing at the city×industry level because we expect those to be related to $\Delta\epsilon_{ict}$ (the productivity shocks). However, union organisation changes at the national-sectoral and local levels can form the basis of instruments. We capture such changes by constructing a leave-one-out measure of the decadal growth rate in elections per establishment for the city (UA_{ct}) and at the national industry level (UA_{it}).[20] For the latter, the premise is that if national-level unions shift toward more activist leadership then election drives in the industries in which they operate will increase.[21] At the same time, organizational spillovers could create different recruiting environments in different cities (Holmes [2006]). Secondly, we use two indicators to proxy the regulatory environment: one for Right To Work states, RTW_{ct}, and another for whether the Republican Party controlled all three branches of the state legislature, R_{ct}, which we assign to cities based on the state in which they reside and average based on population shares for cities that cross state borders.

Having identified potentially exogenous factors that influence the threat of unionisation, we construct a generalised Bartik-type instrument incorporating these variables. We begin by regressing the local growth rate of P^{ne}_{ict} on UA_{ct}, UA_{it}, RTW_{ct}, R_{ct}, and the interaction of UA_{ct} and UA_{it}. The predicted values from this regression, $\hat{g}(P^{ne}_{ict})$, represent the estimated growth rate in the local sector of the threat of unionisation, driven by trends in union organizing efforts outside that local sector and modulated by local regulatory conditions. We form estimates of $\hat{P}^{ne}_{ict} = \hat{g}(P^{ne}_{ict}) \cdot P^{ne}_{ict-1}$ and construct a second set of instruments:

$$IV2_{1ict} = \hat{P}^{ne}_{ict} \cdot \hat{E}_{1ict} - P^{ne}_{ict-1} \cdot \tilde{E}_{1ict-1}$$
$$IV2_{2ict} = \left(1 - \hat{P}^{ne}_{ict}\right) \cdot \hat{E}_{2ict} - \left(1 - P^{ne}_{ict-1}\right) \cdot \tilde{E}_{2ict-1}.$$

where $\hat{E}_{1ict}$, $\tilde{E}_{1ict-1}$, $\hat{E}_{2ict}$, and $\tilde{E}_{2ict-1}$ are outside option terms analogous to those used in constructing IV1. Our identifying assumption is that our instruments are mean independent of $\Delta\epsilon_{ict}$. If that is the case, then the covariances of our instruments with $\Delta\epsilon_{ict}$ and $P^{ne}_{ict}\Delta\epsilon_{ict}$ are both zero, and those error components, presented in the third line of equation (30), do not induce bias in our coefficient estimates.[22]

We face one other new issue in the full specification. Because we do not have direct observations of the sectoral rents, we cannot form the variable corresponding to the extra rent capture term by unions, $(\tilde{\gamma}_3 - \gamma_3)\epsilon_{ict}\Delta P^{ne}_{ict}$. That variable, then, becomes part of the error term, and its effect will be reflected in the estimated coefficients on the right-hand side variables according to a standard omitted variables bias argument. More specifically, we can represent our IV regression of the vector of coefficients in (30) as $\hat{\xi} = (Z'X)^{-1}Z'y$, where X is the matrix of right-hand side variables, Z is the matrix of instruments, and $y = \Delta \ln w^n_{ict}$. Then, under the assumption that the instruments are mean independent of $\Delta\epsilon_{ict}$, the expectation of the estimated coefficients would equal $\xi + (\tilde{\gamma}_3 - \gamma_3) \cdot \alpha^*$, where $\alpha^* = (Z'X)^{-1}Z'(\epsilon_{ict}\Delta P^{ne}_{ict})$ and ξ is the vector of true parameter values. For example, the

[20]Note that since we leave out the specific city, for UA_{ct}, and the specific industry, for UA_{it}, of an *ic* observation when constructing the instrument values, these variables actually take different values for each location and industry.

[21]For instance, when John Sweeney became president of the AFL-CIO in 1995, he pledged to increase unionisation drives, allocating \$20 million to 'organize at a pace and scale that is unprecedented' (cited in Bronfenbrenner [1997], p. 196).

[22]For the $P^{ne}_{ict}\Delta\epsilon_{ict}$ term, we require, $\mathbf{E}[z'_m P^{ne}_{ict-1}\Delta\epsilon_{ict}] = 0, \forall m$, where $\mathbf{E}$ is the expectation operator and z_m is the mth column of the instrument matrix. By iterated expectations, this equals $\mathbf{E}[\mathbf{E}[\Delta\epsilon_{ict}|z'_m P^{ne}_{ict-1}]]$. Given our identifying assumption that the productivity shocks within a city-industry sector follow a random walk, $\mathbf{E}[\Delta\epsilon_{ict}|z'_m P^{ne}_{ict-1}] = 0$ and so, $\mathbf{E}[z'_m P^{ne}_{ict-1}\Delta\epsilon_{ict}] = 0$.

expectation of the estimated coefficient on $\Delta(P^{ne}_{ict}E_{1ict})$ equals $\tilde{\gamma}_1$ plus $(\tilde{\gamma}_3 - \gamma_3)$ times the element of α^* that corresponds to $\Delta(P^{ne}_{ict}E_{1ict})$. That is, the estimated coefficient is a biased estimate for $\tilde{\gamma}_1$ but corresponds to a total effect that includes both the outside option channel and the extra rent capture channel. The same is true of the other estimated coefficients in (30). Given our interest in estimating the total effect of de-unionisation on changes in the non-union wage, the fact that we capture at least part of the unmeasurable rent capture term, $(\tilde{\gamma}_3 - \gamma_3)\epsilon_{ict}\Delta P^{ne}_{ict}$, is useful, though it makes the decomposition of that effect into its component parts less precise because each component also reflects some of the rent capture.

It is worth asking whether the rent capture effect alters the other coefficients to a large degree, i.e., whether $(\tilde{\gamma}_3 - \gamma_3) \cdot \alpha^*$ is large. If we combine standard estimates of the elasticity of wages with respect to rents (which Card et al. [2018a] argue is about 0.1 in the recent literature) with standard findings that unions pay about a 20% wage premium (so, $\frac{\tilde{\gamma}_3}{\gamma_3} \approx 1.2$) and that 15% of workers are unionised then $(\tilde{\gamma}_3 - \gamma_3) \approx 0.02$. The α^* vector equals the vector of coefficients from an 2SLS regression of $\epsilon_{ict}\Delta P^{ne}_{ict}$ on the variables in the X matrix, using Z as instruments. We generate simulated versions of $\epsilon_{ict}\Delta P^{ne}_{ict}$ using the actual values of ΔP^{ne}_{ict} and random draws from a mean zero normal distribution with standard deviation equal to the standard deviation of the residuals from (30) to stand in for ϵ_{ict}.[23] We then run an instrumental variables regression of this simulated $\epsilon_{ict}\Delta P^{ne}_{ict}$ on all the right hand side variables in (30). We re-run this simulated regression 1000 times. The result of that exercise shows that the distributions of the elements of α^* all have means very close to zero and relatively small standard deviations that vary across the specific coefficients. In particular, the 95th percentile of the estimates of α^*_1 (the coefficient corresponding to $\Delta((1-P^{ne}_{ict})E_{2ict})$) is .0036 and of α^*_2 (the coefficient corresponding to $\Delta(P^{ne}_{ict}E_{1ict})$) is .027. Combining these with our calculation that $(\tilde{\gamma}_3 - \gamma_3) \approx 0.02$ implies that our estimated coefficients on these variables include rent capture components that amount to only 0.00007 and 0.0005, respectively. We will see that these amount to very small proportions of our estimated coefficients (which are on the order of about 0.6). We conclude that our estimates of these central parameters can be viewed as essentially only capturing the outside option channels. In contrast, the effects on the interactions of P^{ne}_{ict-1} with industry and time, while still near zero on average, show a larger range. Thus, the rent capture effect seems more likely to be loaded onto the latter estimated coefficients.

4 Data and Descriptive Patterns

Our analysis uses data from the Current Population Survey Merged Outgoing Rotation Groups for 1983-2020 and the CPS May extracts for 1978-1982. We are interested in comparisons across steady states over a medium-long time horizon, and, as such, we consider variation over 10-year periods. We pool observations across 3 years for each period to reduce statistical noise. We consider variation across 1980, 1990, 2000, 2010, and 2020 using the years 1978-1980, 1988-1990, 1998-2000, 2008-2010, and 2018-2020.

From this data, we keep all workers between the ages of 20-65 who do not report being

[23]Note that the actual ϵ_{ict} values in (30) are correlated with the elements of X, but we don't need to recreate those correlations in our simulated ϵ_{ict} draws because they only enter the estimator in correlations with the elements of Z, and ϵ_{ict} is independent of those elements by assumption.

in school either full-time or part-time. We follow Lemieux [2006] in constructing our wage data, working with weekly wages. We use an aggregated grouping of industry codes based on the 1980 industrial classification from the Census Bureau. We obtain a consistent industry classification using crosswalks provided by IPUMS and the Census Bureau that map the 1970, 1990, and 2000 industry codes to the 1980 classification. The result is a consistent classification system with 51 industries. Appendix J contains additional processing details.

We construct a set of cities with as consistent geographic boundaries as possible, given data limitations in the CPS. We are constrained by the number of SMSAs available in the May extract data and end up with 43 cities. Making use of the limited number of counties identified in the CPS, we can create a set of cities which are reasonably, though not always perfectly, consistent over time.[24] The final geographic definition we use pools data for these 43 cities and the remaining population. Specifically, we create additional regions comprising the remaining state population absent the population living in these 43 cities. In the end, our core geographic measure is composed of 93 areas that are fairly consistently defined over the course of the sample period.

Additionally, we use NLRB case data for the sets of three years associated with each of our decadal points to construct probabilities of firms facing successful union certification drives.[25] We use the county of the unit involved in the election to construct our geographic measures, aggregating counties to our city definition discussed above. In particular, we calculate the proportion of firms in an *ic* cell that experienced a successful unionisation drive in the previous 4 years. We view those probabilities as proxies for the proportion of non-union firms that are emulating non-union firms based on the idea that when more firms are being unionised, the threat of unionisation for the remaining non-union firms is greater. We also calculate the proportion of firms facing any union election (successful or not) under the argument that any level of union organizing activity will raise concern for non-union firms. Unfortunately, the election data ends before 2020, and so we estimate the full model over the years 1980, 1990, 2000 and 2010.

Central to our empirical work are the outside option terms characterising alternative job prospects in either the union or non-union sectors. As defined above, these terms are composed of the rents a worker would get in expectation when searching for a new job and are functions of the average wage rent paid in each possible job by city cell (w_{kct}), the proportion of workers in each cell in the city (η_{kct}), and the term that captures the difficulty with which a worker in a job of type j can move to a job of any other type, k ($\varphi_{k|j}$). For the rent component, we use our regression-adjusted wages in order to get as close as possible to rents rather than skill differentials since wage differences that reflect skill differentials cannot be used as an outside option in bargaining (a janitor cannot use the opening of new jobs for lawyers in town to bargain a better wage).

We compute the η_{kct}'s directly from the CPS data. We proxy the $\varphi_{k|j}$ terms with transition probabilities at the national level, estimated using the matched CPS. In particular,

[24] The metropolitan area definition used by the IPUMS identifies a general pattern of expanding metropolitan area definitions over time that we overcome to some extent, but not perfectly: https://usa.ipums.org/usa/volii/county_comp2b.shtml. Estimation using states as the geographic unit yields very similar results, suggesting that issues related to geographic definitions are not driving our results.

[25] We are grateful to Hank Farber for providing this data. We use data on certification elections in which a conclusive decision on certification was reached.

we calculate the proportion of workers in a given cell, j, in year t observed in each possible other cell in year $t+1$.[26] We do this for each of the three CPS years at each decade point (e.g., initial years 1988 – 1990 for the 1990 observation) and average over those three years. This is done separately at each decade point, allowing for changes in transitions over time. One complication in this is workers observed in the same cell in years t and $t+1$ since we can't observe whether they have moved to a different firm in the same cell. To the extent they do, the wage in that cell is part of their outside option with their current firm. We estimate the proportion of workers making such a transition by calculating the proportion of workers who are observed in the same cell in both years but have different values of a set of job characteristics, including how they are paid (hourly versus not hourly), worker class (private versus public), and sub-industry.

4.1 Descriptive Patterns

Before turning to estimation, we present key patterns in unionisation over our sample period. As is well known, the decline of unionisation in the United States from 1980 to 2019 (and for other rich world nations over a similar time frame, see Schmitt and Mitukiewicz [2012] and Lesch [2004]) has been remarkable.

In Figure 1, we plot the fraction of workers unionised at the city level over 1980-2019 for each city, highlighting a subset of cities with particularly large or small declines in unionisation. We also highlight the national average (the thick black dashed line in the figure). On average, about 25% of workers were unionised at the city level in 1980, but this number declines to 17% by 1990 and then to 13% by 2019. In cities like Detroit, Gary, and Pittsburgh, where the union sector played a much larger role in the 1980 economy, the declines are substantial: 21, 29, and 22 percentage points, respectively, by 2019. Smaller declines (under 10%) are observed in cities with low initial unionisation rates, such as Dallas and Rochester. Thus, there is a considerable range in the changes in unionisation across cities and, importantly, there is variation in the decade in which the declines occur. This will allow us to separate the effects of union declines from general trends.

Observed declines in unionisation rates are large, which naturally has implications for our outside option terms. In particular, as discussed in Section 2.5, outside options depend on the probability of a non-union worker getting a union job, $T_{1ict|2i}$, and wages in union jobs, w_{1ict}. Similarly, as unionisation rates fall, so does the threat of unionisation, P^{ne}_{ict}. The theoretical framework we present above suggests that shifts in these terms will impact non-union wages through the bargaining and threat channels.

As a first step in establishing the relevance of these channels, Figure 2 plots the 1980-2010 city-level change in log non-union wages against (a) changes in non-union-to-union transition rates, $\sum_{i'} T_{1i'ct|2i}$, averaged across industries in the given city, and (b) changes in

[26] Our framework assumes that bargaining effects operate only through the unemployment channel, that is, workers must first transition through unemployment to access other jobs. However, due to data limitations, our transition measures use transitions between sectors, which may or may not have included an intervening unemployment spell. Thus, the union outside option term may reflect on-the-job search dynamics. Formally modelling on-the-job search, or job laddering, is beyond the scope of this paper. As noted by Beaudry et al. [2012], it is not straightforward and is sensitive to the modelling of the search process and its relationship to wage determination.

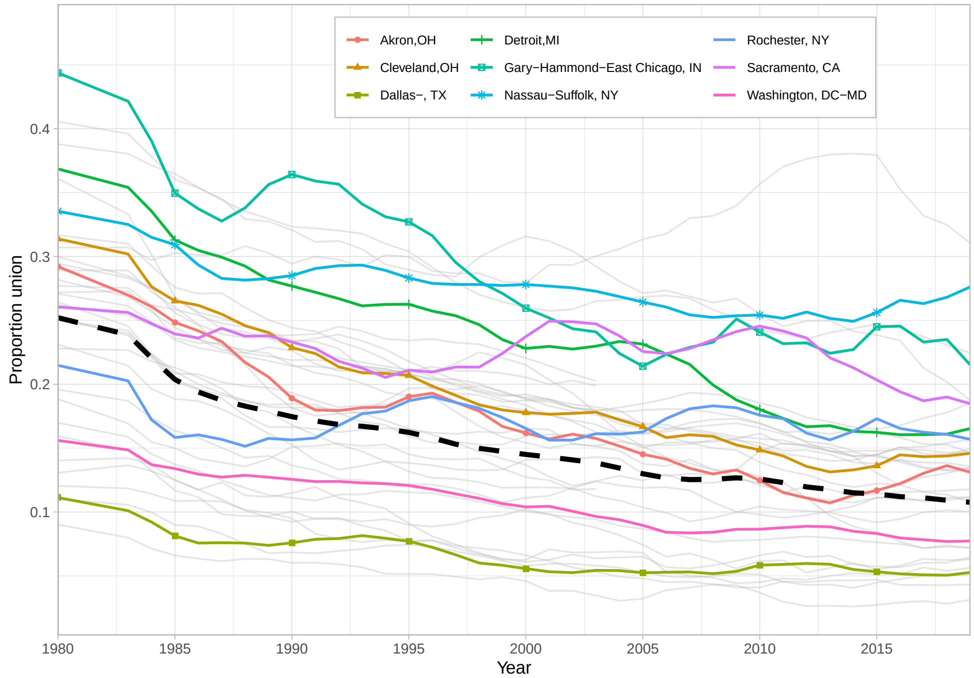


Figure 1: Union percentage over time

Notes: Data comes from the CPS. The figure denotes the fraction of unionised workers at the city level over the 1980-2019 period, calculated as a three-year moving average. The dashed black line refers to the level of unionisation at the national level.

the probability of a firm in an ic cell facing a union certification election, P^{ne}_{ict}, also aggregated at the city level. In both panels, we show the relationship between the change in non-union wages and the measure on the horizontal axis, after accounting for the effect of the other measure through residualization, to present the independent variation from each. The fact that we are working with residualized measures means that movements in the transition rates (that help identify bargaining effects) and the election probabilities (that help identify threat effects) have separate variation. We emphasize this in the figure by highlighting several cities. Some cities, like Chicago and Detroit, show declines in both election probabilities and transitions, while others, such as New York, Baltimore, and Washington, display quite different patterns. These figures suggest that we have sufficient variation to separately identify the bargaining and threat effects.

5 Estimation Results

In the first column of Table 2, we present OLS estimates of our simple specification equation (21). Recall that given the way we created the dependent variable, we are controlling for education, age, gender and race in a flexible way. The standard errors are clustered at the city×year level.[27] To avoid having small cells drive our results, we drop industry×union

[27]In Appendix I, we discuss recent papers on clustering and standard errors using Bartik Instruments. The proposed approach in Borusyak et al. [2022], in which data is aggregated to the level of shocks, is not possible

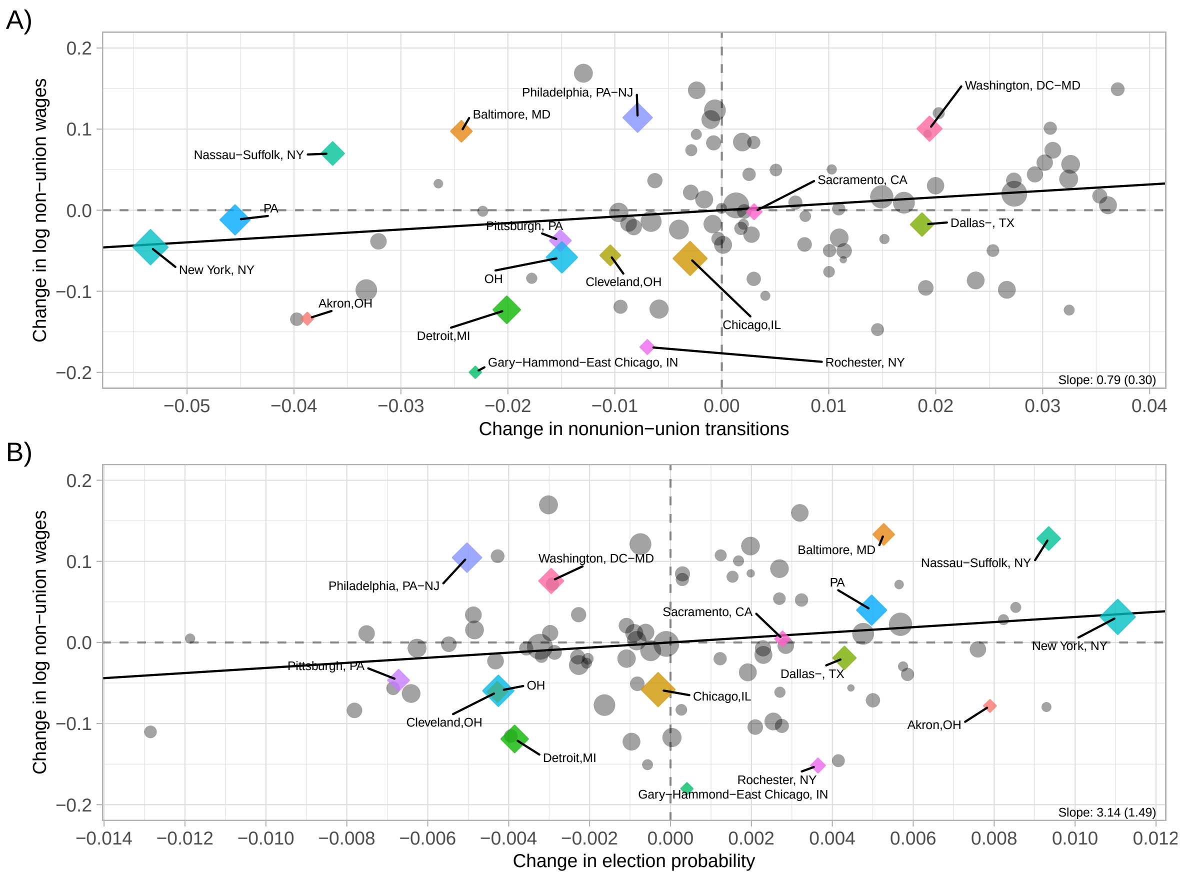


Figure 2: Wages and de-unionisation

Notes: Data from the CPS and the NLRB. In both panels, the y-axis denotes the change in regression-adjusted non-union wages at the city level between 2010 and 1980, and the marker size is relative to the size of the city in 1980. In panel A, the x-axis variable is the change in our measure of non-union-to-union transition rates, $\sum_{i'} T_{1i'ct|2i}$, aggregated to the city level. In panel B, the x-axis denotes the change in the probability of a firm facing a union certification election calculated from the NLRB data and aggregated to the city level. In the bottom right-hand corner of each panel we present slope coefficients (standard errors). Appendix J contains more information on our data construction.

status×city cells with fewer than 10 observations and weight observations by the square root of the cell size. We break the outside option term into its two components: the part related to finding non-union jobs ($E_{2ct|2i}$) and the part related to union jobs ($E_{1ct|2i}$). Note that the coefficients on these variables should be equal since what matters is the overall outside option value, and these are just components of that value. In other words, it should not matter to a firm whether a worker's outside option loses value because a high-rent unionised firm leaves town or a non-unionised firm in an industry that also pays high rents shuts down. The results in column (1) indicate that the option values associated with union and non-union

in our case where the Bartik instruments vary with each jct cell. Adao et al. [2020] argue that standard errors with Bartik instruments face a clustering problem because of correlations across observations with a similar base period composition of the shock exposure shares. Since that would correspond to industry shares in our case and we already control for time-varying industry effects, we argue that we do not face this issue.

jobs have positive and significant effects on non-union wages that are reasonably similar in size. In addition, the employment rate enters significantly but with a theoretically incorrect sign.

Table 2: Non-Union Wages and Outside Options: OLS and 2SLS Estimates

	OLS	2SLS			
	(1)	(2)	(3)	(4)	(5)
$\Delta E_{2ct\|2i}$	1.06***	0.54***			
	(0.022)	(0.11)			
$\Delta E_{1ct\|2i}$	0.99***	0.56***			
	(0.026)	(0.10)			
ΔE_{2ict}			0.59***	0.60***	0.47**
			(0.094)	(0.083)	(0.22)
ΔER_{ct}	-0.36***	0.25*	0.18		
	(0.077)	(0.14)	(0.13)		
Obs.	9024	9024	9024	9024	9024
R^2	0.67				
Year × Ind.	Yes	Yes	Yes	Yes	Yes
Year × City					Yes
Instrument set:		$IV1_{2ct\|2i}$	$IV1_{2ct\|2i}$	$IV1_{2ct\|2i}$	$IV1_{2c\|2it}$
		$IV1_{1ct\|2i}$	$IV1_{1ct\|2i}$	$IV1_{1ct\|2i}$	$IV1_{1c\|2it}$
First-Stage p-Stat.:					
$\Delta E_{2ct\|2i}$		0.000			
$\Delta E_{1ct\|2i}$		0.000			
ΔE_{2cit}			0.000	0.000	0.000
Over-id. p-val		.	0.267	0.551	0.021

Notes: This table displays results from the estimation of equation (21) via OLS (column 1) and 2SLS (columns 2 - 5). The dependent variable is the decadal change in the regression adjusted average hourly wage of non-union workers in an industry-city cell, using CPS data from 1980-2019 across 50 industries and 93 cities. Standard errors, in parentheses, are clustered at the city-year level.

In column (2), we present the results of 2SLS estimation using the $IV1_{1ct|2i}$ and $IV1_{2ct|2i}$ instruments. The p-values for the Sanderson-Windmeijer test statistics for weak instruments [Sanderson and Windmeijer, 2016], found at the bottom of the table, are less than 0.001 in all cases, indicating that we do not face weak instrument problems. The results in column (2) again show positive and significant outside option effects, with the estimated coefficients for the two components being very similar in size. This is strongly supportive of the model since there is no mechanical reason why the two outside option terms should have similar-sized effects. In the context of the model, in which the composition of local employment does not determine productivity within a specific industry, the significance of these effects implies that wages are partly driven by bargaining responses to rents in the local economy. Following BGS, we view the fact that we get these estimates while controlling for industry-specific trends and the local employment rate as reinforcing this interpretation.[28] In Appendix K, we report the results of the estimation of the union wage specification (9). Because unionisation has declined so much over the course of time, we have lost about two-thirds of our industry×city cells when we switch to union wages. As a result, the overall estimates have the right sign but are not well defined when we use the total union sample. However, we were concerned that our model does not apply well to wage setting in the public sector and, indeed, we find that our model performs well for private sector union wages. The estimated outside option coefficients are quite similar in size to those for the non-union specification. While union wages are not our focus, we view the fact that our model performs well for private sector union wages as a useful consistency check for our model.

To understand the magnitude of our estimated outside option coefficients, recall that γ_1 represents the impact of a one-dollar increase in the value of the outside option for a non-union worker in industry i in city c on that worker's wage. For context, the decline in the probability of non-union workers in the motor vehicles and equipment sector transitioning to any union job between 1980 and 1990 led to a 3.4% decrease in the value of their outside option.[29] This, in turn, implies a 1.85% decline in the mean wage (approximately 20 cents) of non-union motor vehicle workers from reduced access to union jobs during that decade. This, though, is only the immediate impact of the change in transition rates into union jobs since we are holding all the wages in other industries constant when computing the new outside option value. If we consider the change in transition rates for non-union motor vehicle workers as part of a shift to a new steady state, reaching that steady state will involve further wage adjustments. The decrease in motor vehicle workers' wages will weaken the outside option of workers in other sectors who tend to have high transition rates into automotive jobs, thus lowering their wages. That, in turn, further reduces the outside option value for motor vehicle workers, leading to additional wage declines, and so on. In the end, the total impact of a one-dollar increase in the outside option value for the mean non-union wage in industry i is $\frac{\gamma_1}{(1-\gamma_1)}$.[30] Therefore, our estimated initial impact of a one-unit change in the value of the outside option of 0.55 (the average of the two estimates in column (2)) becomes 1.22 once we include feedback loops of the spillovers. The total decline in the non-union motor vehicle wage becomes 4.1% (or about 45 cents). In Section 6, we extend this counterfactual exercise to all sectors to calculate the total impact of de-unionisation on wages.

In column (3), we impose the restriction that the coefficients on the two components of the outside option take the same value in an 2SLS estimation. Not surprisingly, given the similarity of the coefficients in column (2), we cannot reject the restriction at any standard level of significance (the p-value associated with testing the restriction is given at the bottom

[28] Previous papers on union spillovers have used the proportion of unionised employment as the key spillover variable. We ran versions of our regressions in columns (1) and (2) but also including the proportion union as an added regressor. When we did so, the proportion union variable had a small and statistically insignificant effect, whether we estimated by OLS or 2SLS. This suggests that any threat effects captured by the union proportion variable occur through the bargaining channel reflected in our outside options variables.

[29] This figure was calculated by recomputing the value of outside options for this sector using our estimates of the 1980 mobility frictions, $\varphi_{k80|i2}$, instead of the 1990 ones. The difference between this value and the observed value of outside options is 3.4% when averaged across cities.

[30] Here, we have assumed that $\tilde{\gamma}_1 = \gamma_1$ so that the spillover effects on union and non-union wages are the same.

of column (3)). As described earlier, we view this as an over-identifying test of our model since there is no mechanical reason why the two terms should have the same effect (and they are based on quite different variation – in non-union versus union industrial proportions and wage premia) but theoretically their effects should be identical.

Apart from checking the validity of our instruments using the over-identification tests, we also follow the advice in Goldsmith-Pinkham et al. [2018] about checking patterns and correlations for further suggestive evidence that the exogeneity requirements are met in our case. These checks are weakened to some extent by the fact that our situation differs from the classic Bartik case because our key endogenous variables (the outside option terms) vary at the ic level rather than just the c level. They also include different national-level variables in the same expression, requiring us to make restrictive assumptions (such as that the industry wage premia are the same in the union and non-union sectors) in order to fit into the Goldsmith-Pinkham et al. [2018] framework. When we do that, the Rotemberg weights for each of our instruments – weights showing which industries are the main drivers behind the variation in our instruments – point to the top five weighted industries being mining, motor vehicles and equipment, retail trade, construction; and lumber and wood products. Apart from retail trade, this list is reassuring because it consists of sectors with high wage premia that, at least at one time, were highly unionised. Thus, they seem like a good set of industries for identifying the impacts of variation in access to high-rent jobs. Goldsmith-Pinkham et al. [2018] also suggests looking for correlates of the baseline industry proportions on which our instruments are built to see if those suggest possible issues. In our case, we would be worried about correlations with variables that might predict growth in city-level productivity. Given that we control education, age, gender, and employment rate, we do not have candidates for other variables that could fit this bill.

In the last column of Table 2, we present the results from a specification in which we drop the ΔER_{ct} variable. As described earlier, we derived our model under partial equilibrium assumptions, including treating labour market tightness as fixed. However, de-unionisation could affect labour market tightness if, for example, firms that de-unionise face lower wage costs and post more vacancies as a result. By not controlling for changes in the employment rate, we allow any such effects to show up in the estimated outside option coefficients – though at the cost of using a specification that is not strictly interpretable under our theory. The estimated coefficient on the outside option value is very similar to what we obtain in the previous column, controlling for ΔER_{ct}. This suggests that the indirect effects of de-unionisation through labour market tightness are unlikely to be large.

5.1 Controlling for Selectivity

Section 3.1.4 clarifies that there is likely selectivity into the union sector based on productivity draws. In Table 3, we present results using our generalized Heckman two-step approach. We present results from the specifications using the quadratic in the change in unionisation proportions in ic cells to control for selection in columns (1) - (3) of Table 3, with column (1) containing OLS estimates and IV estimates in columns (2) and (3). A test of the hypothesis that the parameters in the quadratic equal zero is not rejected at any standard significance level, and the estimates for the key covariates change very little from Table 2. Columns (4) - (6) contain the results when we use the NLRB variables to address selection. Because the elections data series does not extend to 2020, we estimate this specification using the 1980 through 2010 data. Here, we again do not reject the null hypothesis of no selectivity effects, though the associated p-values are much lower than with the union proportion quartic. The main estimated effects are similar in both approaches. In both cases, the over-identification test that the union and non-union outside option values should have the same effect is satisfied at any standard level of significance. We conclude that selectivity is not a central issue driving our results.

Table 3: Non-Union Wages and Outside Options: Controlling for Selectivity

	OLS	2SLS		OLS	2SLS	
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta E_{2ct\|2i}$	1.06***	0.53***		1.03***	0.58***	
	(0.022)	(0.11)		(0.023)	(0.100)	
$\Delta E_{1ct\|2i}$	0.99***	0.56***		0.97***	0.62***	
	(0.027)	(0.10)		(0.026)	(0.089)	
ΔE_{2ict}			0.58***			0.66***
			(0.099)			(0.080)
ΔER_{ct}	-0.36***	0.25*	0.19	-0.41***	0.49**	0.33*
	(0.077)	(0.15)	(0.13)	(0.100)	(0.22)	(0.17)
Obs.	9024	9024	9024	6284	6284	6284
R^2	0.67			0.68		
Year × Ind.	Yes	Yes	Yes	Yes	Yes	Yes
Instrument set:		$IV1_{2ct\|2i}$	$IV1_{2ct\|2i}$		$IV1_{2ct\|2i}$	$IV1_{2ct\|2i}$
		$IV1_{1ct\|2i}$	$IV1_{1ct\|2i}$		$IV1_{1ct\|2i}$	$IV1_{1ct\|2i}$
First-Stage p-Stat.:						
$\Delta E_{1ct\|2i}$		0.000			0.000	
$\Delta E_{2ct\|2i}$		0.000			0.000	
ΔE_{2cit}			0.000			0.000
Over-id. p-val		.	0.320		.	0.136
Selection Controls						
ΔP_{ict} Quadratic	Yes	Yes	Yes	No	No	No
Election Vars.	No	No	No	Yes	Yes	Yes
Joint Tests:						
p-val	0.71	0.95	0.92	0.46	0.14	0.07
F-Stat	0.53	0.19	0.23	0.96	1.63	2.00

Notes: This table displays results from the estimation of equation (21) via OLS (columns 1 and 4) and 2SLS (columns 2, 3, 5, and 6). The dependent variable is the decadal change in the regression adjusted average hourly wage of non-union workers in an industry-city cell, using CPS data from 1980-2019 across 50 industries and 93 cities. Standard errors, in parentheses, are clustered at the city-year level.

5.2 Emulation Specification

In Table 4, we present the results of our complete specification (30), which captures both the bargaining and wage emulation channels. Recall that because we don't observe whether a given non-union worker is in a simple or emulating type firm, our specification involves interactions of the probability a non-union firm is or is not an emulator with the outside options of union and non-union workers, respectively. In response, we use our $IV2$ instruments ($IV2_{1ict}$ and $IV2_{2ict}$) that make use of both the factors driving the unionisation threat and the factors related to worker outside options. We proxy the proportion of emulators using the number of union drive wins per non-union establishment, P^{ne}_{ict}. In the first column, we include a complete set of industry-by-time effects. In the second and third, we add in interactions of P^{ne}_{ict-1} with both a complete set of industry-by-time effects and a complete set of city-by-time effects. The theory indicates the need for both of the latter sets of effects. The identifying variation continues to be across cities within industries and we present the more limited specification in column (1) to highlight that point. In all columns, we include a quadratic in the change in the proportion of the workers in the ic cell who are unionised to control for selection.

The estimated coefficient on the first term in our complete specification (the probability of not being a union emulator times the outside option for a non-union worker) is slightly larger than the estimated coefficient on the outside option value in the simple specification with selection controls in Table 3. For the second term (the union worker's outside option value times the probability that the firm is an emulator), the theory implies that the coefficient is the effect of an increase in the outside option value on a union worker's wage, and it turns out to be slightly larger than the effect for non-union workers. Our estimates are quite robust to whether we include the effects involving interactions of P^{ne}_{ict} and whether we control for ΔER_{ct}.

The key result from these estimates is that both the bargaining channel (as captured in the coefficient in the first row) and the standard threat channel (represented by the coefficient in the second row) matter. If $\gamma_1 = \tilde{\gamma}_1$ (a restriction we cannot reject given our estimates) and the changes in the outside option values for union and non-union workers were the same, then our estimate of the impact of a one dollar increase in the value of the outside option on the non-union wage is slightly larger than in the simple case (1.9 instead of 1.2). We will provide evidence on the relative size of the bargaining and standard threat channels in a counterfactual exercise after investigating heterogeneity across different demographic and skill groups.

The significance of the union worker outside option is, in some ways, remarkable. Its effect is identified relative to the non-union worker outside option because of differential changes in transition rates for union and non-union workers. In Table 5, we present the mean probabilities, separately, that union and non-union workers transit to a union job by the following year for each of our sample years using the national level data (i.e., mean values of $\varphi_{1i't|1i}$ and $\varphi_{1i't|2i}$). These show a strong decline in the probability of accessing a union job for non-union workers (from .24 in 1980, to .091 in 2000, and .07 in 2020) but higher levels that don't decline as fast for union workers (where the probability is .273 in 1980, .197 in 2000, and .168 in 2020). The impact of these differences on local outside option values is mediated through their interactions with changes in local industrial composition (the

Table 4: Non-Union Wages and Outside Options: Including Wage Emulation Effects

	2SLS		
	(1)	(2)	(3)
$\Delta\left((1-P^{ne}_{ict})\cdot E_{2ict}\right)$	0.67***	0.64***	0.65***
	(0.073)	(0.10)	(0.094)
$\Delta\left(P^{ne}_{ict}\cdot E_{1ict}\right)$	0.80***	0.76***	0.73***
	(0.094)	(0.26)	(0.27)
ΔER_{ct}	0.31*	0.37*	
	(0.16)	(0.22)	
Obs.	5958	5958	5958
Year × Ind.	Yes	Yes	Yes
P^{ne}_{ict-1}× Ind. × Year		Yes	Yes
P^{ne}_{ict-1}× City × Year		Yes	Yes
Instrument set:			
	$IV2_{2ict}$	$IV2_{2ict}$	$IV2_{2ict}$
	$IV2_{1ict}$	$IV2_{1ict}$	$IV2_{1ict}$
Select controls			
ΔP_{ict} Quadratic	Yes	Yes	Yes
First-Stage p-Stat.:			
$\Delta\left((1-P^{ne}_{ict})\cdot E_{2ict}\right)$	0.000	0.000	0.000
$\Delta\left(P^{ne}_{ict}\cdot E_{1ict}\right)$	0.000	0.000	0.000

Notes: This table displays results from the estimation of equation (30) via 2SLS. The dependent variable is the decadal change in the regression adjusted average hourly wage of non-union workers in an industry-city cell, using CPS data from 1980-2019 across 50 industries and 93 cities. Standard errors, in parentheses, are clustered at the city-year level.

η_{kct}'s) and changes in wage premia for different job types, or, for our instruments, changes in national level job rents (the ν_{kt}'s). It is the variation in our instruments, $IV2_{1ict}$ and $IV2_{2ict}$, that is most relevant for our identification, and the differences in transition rates underlying each translate into a correlation between the instruments of only 0.16 across industry×city cells.[31] In Appendix H, we present results from a quasi-reduced form specification in which we regress $\Delta \ln w^{n}_{ict}$ on ΔE_{1ict}, ΔE_{2ict}, and ΔP^{ne}_{ict} separately, including all the same controls and using the same instruments as for equation (30). We find that all three elements enter significantly, supporting the argument that our instruments for the two outside option values have identifying variation relative to each other.

Table 5: Transitions to Union Jobs

	(1)	(2)
Year	Non-Union-to-Union	Union-to-Union
1980	0.240	0.273
1990	0.126	0.247
2000	0.091	0.197
2010	0.071	0.167
2020	0.070	0.168

This table reports transition probabilities for non-union and union workers into union jobs. The transition probabilities exclude same-job transitions for union workers. The data comes from matched CPS data, described in the main text.

5.3 Heterogeneity in Spillover Effects

There is considerable heterogeneity in the unionisation experience. Farber et al. [2021] and others show that unionisation has, historically, been particularly prevalent among low-skilled men (although, as highlighted by Card et al. [2018b], there has been a remarkable rise in the share of unionised jobs held by women). The decline in unionisation might then be expected to have had a bigger impact on groups with higher initial unionisation rates, and so it would be useful to know if they have a bigger or smaller reaction to outside options in their wage setting.

In Table 6, we present estimates of the bargaining and standard threat effect coefficients (γ_1 and $\tilde{\gamma}_1$ in equation (30)) for a set of sub-populations defined by gender, age, and education. Each row corresponds to estimates for a different sub-sample. We calculate the transition rates from any job type to any other job type, $\varphi_{kt|j}$, for the specific population being examined and, based on those transition rates, calculate outside option values for each sub-sample. The 3rd and 4th columns of the table also show the p-values from SW weak

[31] In comparison, the correlation between changes in the outside option variables, ΔE_{1ict} and ΔE_{2ict}, is .86. This is much larger than the correlation between the instruments because the outside option values use the local wages, w_{1ict} and w_{2ict} and the local changes in the job type shares. Since these tend to move together at the local level in a way that the national level ν_{kt}'s and the start of period η_{kct}'s do not, the outside options are much more correlated than the instruments that actually generate our estimated effects.

Table 6: Subsample Analysis - Coefficient Estimates on Outside Options

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Coefficient		First-stage p-value			1980	
Sample	γ_1	$\tilde{\gamma}_1$	(1)	(2)	N	Union Prop.	Union Prem.
Men	0.53**	0.71**	0.00	0.00	4545	0.32	0.14
	(0.15)	(0.22)					
Women	0.64**	0.75**	0.00	0.00	3700	0.18	0.17
	(0.08)	(0.10)					
Age 20–35	0.59**	0.66**	0.00	0.00	4021	0.23	0.17
	(0.11)	(0.16)					
Age 36–55	0.46**	0.59**	0.00	0.00	4133	0.29	0.13
	(0.18)	(0.20)					
$\leq$ HS	0.45**	0.61**	0.00	0.00	4081	0.30	0.18
	(0.10)	(0.15)					
$>$ HS	0.38**	0.42*	0.00	0.00	4283	0.21	0.11
	(0.19)	(0.22)					
Men Young/Low skill	0.62**	0.57*	0.00	0.00	2648	0.35	0.20
	(0.13)	(0.29)					
Men Young/High Skill	0.06	0.16	0.03	0.01	2446	0.20	0.08
	(0.61)	(0.64)					
Men Old/Low Skill	0.56**	1.00**	0.00	0.00	2422	0.42	0.13
	(0.16)	(0.27)					
Men Old/High Skill	–0.92	0.28	0.03	0.00	2309	0.21	–0.02
	(1.07)	(0.71)					

Notes: This table displays results from the estimation of equation (30) via 2SLS on separate subsamples.

instrument tests for the instruments corresponding to the two outside option terms. In all cases, these p-values are 0.03 or less, implying the absence of weak instrument problems.

The first two rows contain separate results for men and women. These indicate that both the bargaining and standard threat effects are larger for women, implying that declines in unionisation would have a more negative effect on non-union wages for women than men. Effects are also stronger for younger workers and for people whose highest level of education is high school graduation or less. In the last four rows, we delve deeper into skill-related differences for males, using an approach from Card [2009] for creating skill groups. In this method, weights are generated for each person that corresponds to their contribution to four groups: young, low educated; young, high educated; old, low educated; and old, high educated.[32] We focus on men since they suffered the largest declines in unionisation. The last two columns of these rows show that the young/low-skilled and old/low-skilled men had particularly large values of the union wage premium and the proportion unionised. Thus, these are groups where we would expect both the union threat and bargaining spillover effects to be particularly large and, indeed, the estimated effects are large relative to other groups – particularly more skilled workers.

We have also examined potential heterogeneity between public and private sector unions. Thus far, we have included the public sector, both in the construction of our outside option terms and as an observation on the left-hand side of the wage equation. Card et al. [2018b] however outline the marked difference in unionisation between the private and public sectors since 1980 such that unionisation is now 5 times higher in the public sector. To the extent that wage setting is different in the public sector, these shifts in composition could be driving some of our results. In Appendix K, we present results excluding the public sector both in the construction of the dependent variable and in the construction of our outside option variables and associated instruments. Our results are robust to these changes, though our estimated spillover effects are slightly larger in the simple specification.

6 Counterfactual Exercise

Our results thus far indicate a significant relationship between the quality of job opportunities in both the non-union and union sectors and non-union wage setting. However, the exact magnitude of the estimated effects remains unclear. In this section, we pursue a counterfactual exercise, asking what path mean wages in a typical city would have followed if unionisation rates and union wage premia had remained at their 1980 levels. This both provides a way of characterizing the size of our estimated effects and some insight into whether de-unionisation played an important role in wage changes over the last four decades.

[32] In particular, people are assigned an age weight for each of two categories – young (with the weight generated from a quadratic kernel centred on age 27.5 with a 20-year bandwidth) and old (using a quadratic kernel centred on age 50 with a 20-year bandwidth). They are also assigned a weight for the low-educated group and for the high-educated group using Card [2009]'s efficiency weights. The low-educated group puts a weight of 1 on high school graduates and smaller weights on adjacent education categories, while the high-educated group puts a weight of 1 on those with a BA. The four skill groups are formed by multiplying the weights for the age groups with the weights for the education groups.

6.1 Loss of Union Power and Movements in the Average Wage

Our focus is on changes in total mean wages at the city level, expressed as the weighted average of non-union and union mean wages, with the weight being the proportion unionised at the city level, P_{ct}^u:

$$w_{ct} = P_{ct}^u \cdot w_{ct}^u + (1 - P_{ct}^u) \cdot w_{ct}^n, \tag{31}$$

where w_{ct}^u is the mean union wage and w_{ct}^n is the mean non-union wage in city c at time t. We use residualized industry-city wages from our regressions to abstract from the confounding effects of changes in education, age, and other factors, combined with local industrial shares, to create city-level wages.[33]

Changes in union strength affect average city wages through four channels:[34]

1. Union Proportion (P_{ct}^u): This is the most direct effect, representing the shift from higher-paid union jobs to lower-paid non-union jobs, holding sector wages constant. This is the "between" component in standard decompositions.

2. Probability of a Non-union Firm Being Unionised (P_{ict}^{ne}): This captures part of the classic threat effect, representing the changes in the likelihood of a non-union firm becoming unionised.

3. Probability of Finding a Union Job: Changes in transition rates, $T_{kct|j}$, which combine how changes in mobility frictions, $\varphi_{kt|j}$, and job shares, η_{kct}, impact outside options, affecting wages through the bargaining and emulation channels. As noted in Table 5 of Section 5.2, the probability of accessing a union job for non-union workers, $\varphi_{1it|2i}$, fell throughout the period we study. When linking changes in transition rates to the decline in union power, we do not want to attribute all of the changes in job shares to union effects. Instead, we assume that shifts in the industrial distribution for non-union workers capture changes in the overall economy, while a change in the industrial distribution for union workers relative to what happens for non-union workers is a union decline effect. We denote these relative job shares as η_{1ict}^* (the difference between the actual growth in the industrial share in industry i in the union sector and what would have happened if it had grown at the non-union rate) and transition rates using these shares as $T_{kct|j}^*$.

4. Union Wage Premium: Declines in union bargaining power could reduce the union wage premium ($w_{ct}^u - w_{ct}^n$), lowering the value of the outside option of finding a union job. This could arise because unions become less effective at unifying worker resistance during bargaining or become afraid to threaten the withdrawal of the whole workforce in a new policy environment.

[33] We set the wage level to correspond to the mean wage across all worker types. In particular, mean wages correspond to the wages of white workers, holding the proportion of education×gender groups at their 1980 levels.

[34] A fifth channel, selection effects, could theoretically increase observed non-union wages by changing the productivity composition of non-union firms. However, we find no substantial evidence of this effect, so it is not included in our decompositions.

Figure 3 plots the percentage change in these key drivers relative to their 1980 values, aggregated across cities using city populations as weights. Thus, the trends shown depict the movements of each component for an average city. The trend in the probability of unionisation (P^u_{ct}) is labelled as 'Proportion Union' in the figure. In the line labelled 'Transitions' we present the movement in the national level probabilities of a non-union worker in any industry transiting to a union job in any industry ($\varphi_{1i't|2i}$), averaged across industries. We present this series rather than the local transition probabilities ($T^*_{kct|j}$) to provide an unadulterated look at the main driving force in the transition rates. This force is obviously related to changes in P^u_{ct}, though one could imagine that it could decline faster than the overall union proportion (if older union workers keep their jobs but new job searchers have difficulty getting into a union job) or slower (if the proportion declines quickly because union workers suddenly start taking early retirement). In fact, the figure shows that the two proportions fell since the 1980s, but the probability of entering a union job declines faster. Notably, P^{ne}_{ict} (the probability a non-union firm is successfully unionised), labelled as 'Threat' in the figure, fell the fastest of any of the unionisation measures, particularly in the 1980s when the policy environment was strongly against unionisation.[35]

Perhaps the most interesting line in Figure 3 corresponds to the union wage premium (the red line). The premium actually increases in the 1980s before showing a sizeable decline in the 1990s and a smaller one thereafter. Both Card [2001] and Farber et al. [2021] have highlighted the seemingly odd result: the union wage premium did not decline during the 1980s when union power fell substantially.[36] Our model (echoing an argument in Farber [2005]) provides an explanation for the increase in the premium in the 1980s in our data and, potentially, the longer-term stability in the premium demonstrated in Farber et al. [2021] based on the emulation channel. Recall that the observed mean non-union wage equals a weighted average of the simple non-union wage (w_{2ict}) and the emulation wage (w_{3ict}). The weights are $(1 - P^{ne}_{ict})$ and P^{ne}_{ict}, respectively. Suppose that larger forces (trade, technological change, etc) drive down both w_{2ict} and the union wage, w_{1ict}, to the same extent. If the threat of unionisation declines simultaneously, the observed non-union wage will fall further because there will be fewer emulating firms, and the emulation wage they have to pay won't be as high. This pattern of faster decline in mean observed wages in the non-union sector is what we observe in the 1980s. It is striking that this is the decade in which the union threat fell fastest relative to other unionisation probabilities.[37]

6.2 Overall Decomposition

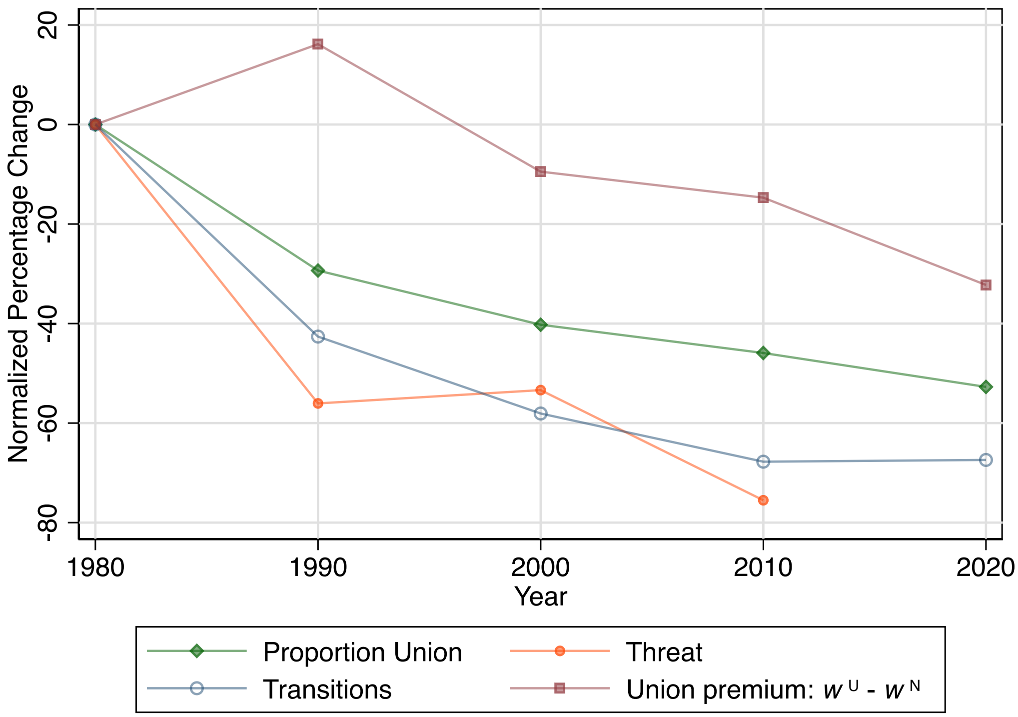
We present our decomposition of the overall trend in average city wages in Figure 4. The bottom line in Figure 4 is the actual trend in an average city's (residualized) mean wage.

[35] It is worth noting that the probability of a firm facing a union election was small even in 1980 (on the order of 4%).

[36] Farber et al. [2021] plot union wage premiums over an extended time period. Their plot differs from ours in showing a flat premium over the 1980s but is similar in showing a decline after 1990. Their estimates are based on family income and do not include controls for education that are part of our estimation.

[37] In Appendix K, we report on a rough check on this argument in which we regress changes in the union wage premia in industry × city cells on changes in our union threat variable, P^{ne}_{ict}. As our theory predicts, the union threat effect is negative and statistically significant.

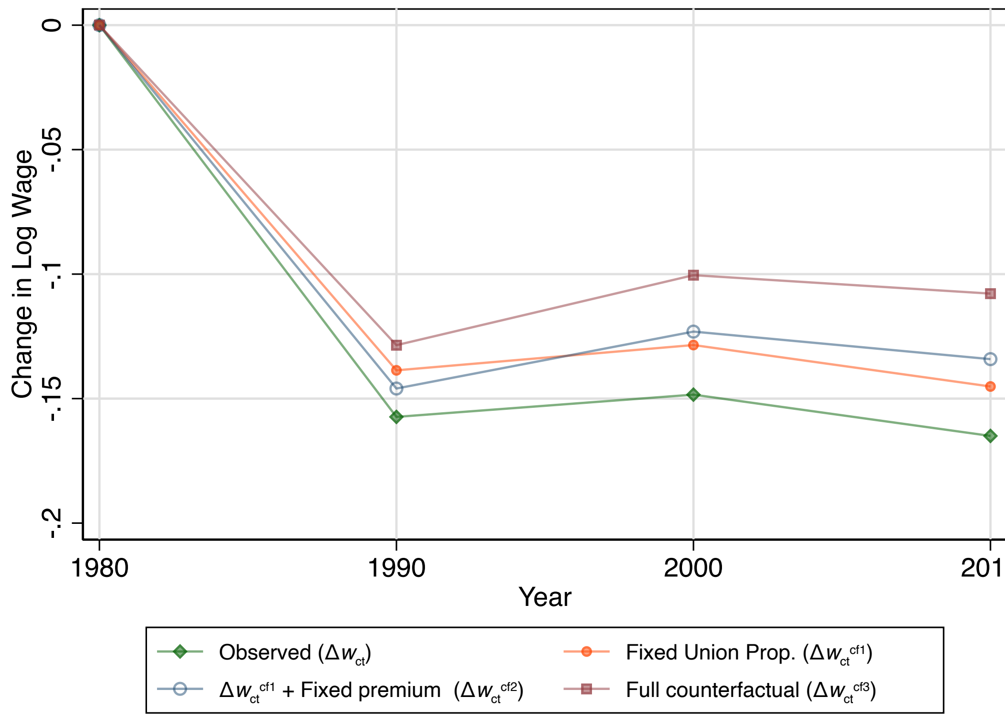
Figure 3: Components of Decomposition



Notes: Data from the CPS and the NLRB. Each series is represented as a percentage change from the corresponding 1980 level. Proportion union, union premium and transitions are constructed from the CBS data, discussed in Appendix J. The threat of union election comes from NLRB data and is described in detail in Appendix J.

It depicts an overall real wage trend that is strongly decreasing between 1980 and 1990 – falling by 15.7% in that decade – followed by a see-saw pattern of mild increases in the 1990s and declines in the 2000s.[38]

Figure 4: Average Wage Decomposition



Notes: Data from the CPS and the NLRB. Each series is represented as a log change from the corresponding 1980 level. Wage data is from union and non-union workers and is adjusted for worker characteristics. The 'Observed' wage series represents the national average of city-industry wages using the size of the city-industry in 1980 as fixed weights. 'Fixed Union Proportion' holds the proportion of union workers fixed at the 1980 levels. 'Full Counterfactual' also holds the threat, union premium, and union transitions at 1980 levels. Details of the series construction are described in the main text.

To understand the components in our decomposition, we use (31) to write the change in the city-level mean wage between period t and 1980 as follows:

$$\Delta w_{ct} = \underbrace{\Delta P^u_{ct} \cdot (w^u_{ct} - w^n_{ct})}_{\text{Change in Union Proportion}} + \overbrace{\underbrace{P^u_{c80} \cdot \Delta(w^u_{ct} - w^n_{ct})}_{\text{Change in Wage Differential}} + \underbrace{\underbrace{\Delta w^n_{ct}}_{\text{Change in Non-Union Wages}}}_{\Delta w^{\text{cf2}}_{ct}}}^{\Delta w^{\text{cf1}}_{ct}}. \tag{32}$$

The first component of our decomposition is formed by setting $\Delta P^u_{ct} = 0$ (i.e., holding the union proportion at its 1980 value while allowing other factors that determine wage changes

[38] We end our figure in 2010 because we only have data on one element of our decomposition – the part related to union elections – up to that year.

to vary). We denote this counterfactual wage series as $\Delta w_{ct}^{\text{cf1}}$ in Figure 4. This line shows that the decline in unionisation contributed to a 0.019 log-point drop in the mean wage in the 1980s, accounting for about 12% of the overall drop in the mean wage during that decade, with a similar effect on the drop from 1980 to 2010.

In examinations of the impact of unions on mean wage movements, authors often combine this first 'shifting weights' component with changes in the union wage premium. Thus, we form a second component by additionally setting $\Delta(w_{ct}^u - w_{ct}^n) = 0$, i.e., holding the union wage premium at its 1980 level. We refer to this counterfactual as $\Delta w_{ct}^{\text{cf2}}$ in the figure. As highlighted in Figure 3, the wage premium increased in the 1980s but declined thereafter. As a result, the impact of the union premium offsets the effect of de-unionisation in the 1980s but reinforces it in later decades. These two forces together account for a 3.1% drop in mean wages between 1980 and 2010. In a similar vein, Card et al. [2004] calculate that a standard shift-share analysis incorporating both declines in the unionisation rate and the union wage premium implies a drop in the mean US wage by 2.6% between 1984 and 2001.

A standard decomposition stops at this point. However, our estimates imply that de-unionisation affected the remaining component (the change in mean non-union wages) through both the bargaining and threat channels. To account for these effects, we return to our non-union wage specification, (30), which indicates that changes in the mean non-union wage are driven by changes in (1) industry wage premia in the non-union sector (γ_{0it}), (2) changes in outside option values (E_{1ict} and E_{2ict}), and (3) changes in the threat probability (P_{ict}^{ne}) (as well as ΔER_{ct} and other factors captured in the error term). De-unionisation affects the mean non-union wage through changes in (2) and (3) with underlying changes in the probability of finding a union job (transition rates, $T_{kct|j}$), and the relative value of union work (wage premium, $\nu_{1it} - \nu_{2it}$).[39] We denote a counterfactual non-union wage as if changes in these factors did not occur as $w^n_{ct|P^{ne}_{ic80}, T^*_{kc80|j}, \nu_{1i80}-\nu_{2i80}}$. Thus, non-union wage trends can be decomposed as:

$$\Delta w_{ct}^{\text{cf2}} \equiv \Delta w_{ct}^n = \underbrace{\left[w_{ct}^n - w^n_{ct|P^{ne}_{ic80}, T^*_{kc80|j}, \nu_{1i80}-\nu_{2i80}} \right]}_{\text{Non-Union Spillover Effect}} + \underbrace{\left[w^n_{ct|P^{ne}_{ic80}, T^*_{kc80|j}, \nu_{1i80}-\nu_{2i80}} - w_{c80}^n \right]}_{\Delta w_{ct}^{\text{cf3}}}. \tag{33}$$

To estimate $w^n_{ct|P^{ne}_{ic80}, T^*_{kc80|j}, \nu_{1i80}-\nu_{2i80}}$, we use our estimated non-union wage equation and plug in 1980 values for the indicated components. However, these initially estimated wages are only first-round effects of de-unionisation. If the counterfactual wages in a particular ic cell are higher than observed, outside options for other workers would also be higher. Thus, we create a second round of counterfactual outside option values using the first round of counterfactual wages and then form a second round of counterfactual wages using updated outside options. We iterate this process until the predicted wages change by less than 0.1 percent. This estimates the complete feedback loop inherent in bargaining schemes, ensuring that the union premia used in the outside option terms are consistent with the premia calculated from the set of counterfactual wages.

[39] We use national-level industrial premia differences ($\nu_{1ict} - \nu_{2ict}$) as drivers of outside option changes due to de-unionisation, not local wage premia ($w_{1ict} - w_{2ict}$). The former corresponds to ($\tilde{\gamma}_{0it} - \gamma_{0it}$) in our wage specifications and are treated as exogenous factors, while the latter are determined endogenously through spillovers within our model.

We refer to w_{ct}^{cf3} as the 'full counterfactual' in Figure 4 since it incorporates all the paths through which de-unionisation could affect the non-union mean wage. The last spillover component adds a further 2.6% over the full period, approximately doubling the estimated effect from the standard decomposition alone. Previous estimates of spillover effects of de-unionisation on non-union wages based on a regression of mean non-union wages on the union proportion range from large (Holzer [1982] and Denice and Rosenfeld [2018] – with estimated effects that would over-explain the decline in real non-union wages between 1980 and 2010) to near-zero effects (Farber [2005]) to negative effects (Neumark and Wachter [1995]). Our estimates are closer to (though somewhat larger than) those in Farber [2005]. A related literature focuses on the impacts of inequality rather than wage levels. Fortin et al. (2021) find that taking account of spillovers roughly doubles the estimated 'shift-share' impact of de-unionisation on wage inequality over the 1979-2017 period.

Over the full 1980-2010 period, the three components together imply that de-unionisation can account for 34.6% of the total decline in the mean wage. To provide further context for the size of our counterfactual effects, Autor et al. [2013]'s estimates of the impact of the China trade shock on the wages of non-manufacturing workers (their estimated effect on manufacturing wages is zero) amounts to a 0.009 decline between 1990 and 2000, and a 0.014 decline between 2000 and 2007. Together, these are approximately the same size as our estimated effect of de-unionisation on non-union wages alone from 1980 to 2010 and about a third of our estimate of the total de-unionisation effect. Over this same period, the US federal minimum wage fell by 13% in real terms. If we take a relatively extreme estimate of minimum wage spillovers and assume that the wages of workers up to 1.2 times the minimum wage shift when the minimum wage shifts then using the fact that 18% of workers earned under 1.2 times the minimum wage in 1980 (Hardy et al. [2023]), the decline in the minimum wage would account for a 2.3% decline in the mean wage – again, about a third of our total de-unionisation effect.

It is worth noting that, in our model, changes in the union wage premium (the second decomposition component) are driven by three factors. The first is changes in the difference in industrial wage premia between the union and non-union sectors ($\tilde{\gamma}_{0it}$ - γ_{0it}). This arises in our model because unions capture a different share of industry price movements over time, and that share may change as unions become weaker. We view that mechanism as an exogenous force, which our model does not explain. On the other hand, the second and third factors (relative changes in outside option values between the sectors and the union threat probability) are forces within our model. In that sense, we can estimate how much of the second decomposition component is determined by the threat and spillover effects from our model. It turns out those amount to most of the second component. Thus, changes in the threat and spillover effects stemming from de-unionisation drive about two-thirds of the total impact of de-unionisation on the decline in the overall mean wage from 1980 to 2010.

In the top panel of Table 7, we present our counterfactual analysis for various sub-groups. The first row shows the total mean wage decline between 1980 and 2010 for each subgroup. The second row displays the standard shift-share effect, the third shows changes in union wage premia, and the fourth shows spillover effects. The second panel further decomposes the spillover component, which we discuss in the next section. The sixth row sums the spillover, wage premium, and shift-share effects, while the last row shows this total as a proportion of the total wage decline, indicating how much the decline would have been reduced if union-related factors had remained at 1980 levels.

Columns 2 and 3 show that men experienced a decline in the mean real wage between 1980 and 2010, which was over double that experienced by women, and also had a much larger loss in unionisation and, as a result, larger spillover effects. In the end, though, the proportion of the overall wage decline explained by de-unionisation is quite similar for men and women. Low-educated male workers in both our young and old age groups experienced similar impacts from de-unionisation, and for both, the effects are sizeable: 30.7% for the former group and 34% for the latter. For more educated, young workers, de-unionisation had essentially no impact on their mean wage movements. For the highly educated, older workers, the spillover effects actually imply increases in mean wages. This arises because union jobs for this education group became more concentrated in higher-paying (public sector) jobs, implying increased average union wages that more than offset declines in the probability of getting a union job in their outside option term.

Table 7: Outside Options Contribution to Changing Wages - Subsample Analysis

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	All	Men	Women	Men			
				Young Low Skill	Young High Skill	Old Low Skill	Old High Skill
	1980-2010						
(1) Observed	–0.165	–0.223	–0.091	–0.322	–0.109	–0.288	–0.089
(2) Union Prop.	–0.020	–0.032	–0.007	–0.075	–0.003	–0.061	0.002
(3) Union Premium	–0.011	–0.002	–0.012	0.030	–0.004	0.036	0.000
(4) Non-Union Spillovers	–0.026	–0.044	–0.016	–0.054	–0.005	–0.072	–0.001
(4a) Fixed Threat	–0.005	–0.004	–0.006	–0.003	–0.005	–0.004	–0.007
(4c) Fixed Transitions	–0.020	–0.039	–0.008	–0.048	–0.002	–0.074	0.001
(4c) Fixed Union Prem.	–0.001	–0.001	–0.001	–0.003	0.002	0.006	0.005
(6) Total	–0.057	–0.078	–0.034	–0.099	–0.012	–0.098	0.002
(7) Total/Observed	0.346	0.349	0.377	0.307	0.107	0.340	–0.018

Notes: This table displays results from the decomposition for union and nonunion workers from 1980-2010. Each column contains the decomposition results for a different subsample. All figures are log changes from 1980 levels. Details described in main text.

6.3 Decomposing Non-union Wages

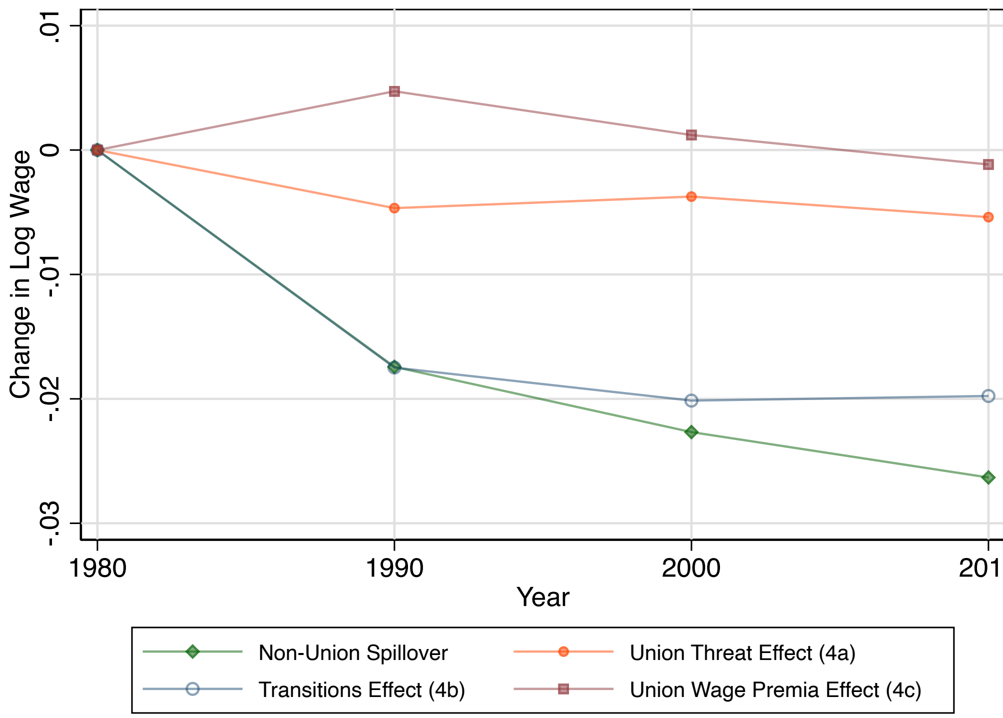
We next turn to decomposing the effect of de-unionisation on non-union wages, which, of course, is the focus of our estimation. To do so, we start with (33) and further decompose the Non-union Spillover Effect into its sub-components.

$$
\begin{aligned}
\text{Non-Union Spillover Effect} = & \underbrace{\left[w^n_{ct} - w^n_{ct|P^{ne}_{ic80}, T_{kct|j}, \nu_{1it}-\nu_{2it}}\right]}_{\text{Union Threat (4a)}} \\
& + \underbrace{\left[w^n_{ct|P^{ne}_{ic80}, T_{kct|j}, \nu_{1it}-\nu_{2it}} - w^n_{ct|P^{ne}_{ic80}, T^*_{kc80|j}, \nu_{1it}-\nu_{2it}}\right]}_{\text{Transitions (4b)}} \\
& + \underbrace{\left[w^n_{ct|P^{ne}_{ic80}, T^*_{kc80|j}, \nu_{1it}-\nu_{2it}} - w^n_{ct|P^{ne}_{ic80}, T^*_{kc80|j}, \nu_{1i80}-\nu_{2i80}}\right]}_{\text{Union Wage Premia (4c)}}
\end{aligned} \tag{34}
$$

In Figure 5, we present each of the elements of our counterfactual non-union spillover effect. The line with diamonds corresponds to the total counterfactual effect of holding all the de-unionisation components constant at 1980 levels on the non-union wage. It says that all the factors combined resulted in a decline in the non-union wage of about 1.8 percentage point in the 1980s, rising to 2.6 percentage points by 2010. The remaining lines on the figure show the contribution to the full counterfactual of its constituent parts, and the sum of the points on those lines in a given year equals the total counterfactual effect. Holding the union wage premia to their 1980 values (shown in the line labelled 'Union Wage Premia Effect') would have resulted in a decrease in the non-union mean wage in the 1980s because increased premia in that decade increased the value of outside options. As we described earlier, our model provides an explanation for why union premia would increase exactly when union power is being most substantially reduced, stemming from the reduction in the need for some non-union firms to emulate union wages since they no longer fear their shop being unionised. In contrast, the large decline in the probability a non-union worker could find a union job in the 1980s (shown in the 'Transitions Effect' line) implied a substantial decline in the non-union wage. In fact, because the threat and wage premium effects happen to offset each other in that decade, the reduction in transition rates equals the size of the total spillover effect. In subsequent decades, the union wage premia decline and the transition effect stabilizes somewhat so that over the 1980 to 2010 period, the decline in the transition rates accounts for about 75% of the total spillover effect.

The last de-unionisation factor is the threat probability, which is captured by the line labelled 'Union Threat Effect'. What is most noteworthy about this effect is its size. While our estimates show clear evidence of the standard emulation threat effect, their actual impact on non-union wage movements was small. This occurs mainly because the threat probabilities themselves are small, even in 1980. A small threat effect means that the sizeable spillover effect that emerges by 2010 in Figure 4 is almost completely accounted for by the bargaining channel. This has potentially important implications for policymaking aimed at raising wages since the threat effect can only be harnessed by increasing unionisation. But the bargaining channel is not unique to unions - any policy that pushes up the outside option value for workers (such as eliminating non-compete clauses (Johnson et al. [2020] or expanding commuting options Hafner [2022]) can have this effect, and our results imply that this channel can be powerful. This is reminiscent of the results in Caldwell and Danieli [2021], who show that wages are increasing in their index of the value of outside options. Their index increases when workers have greater probabilities of transferring to other occupations and job opportunities. Our result is driven by decreases in the probability a worker can transfer to a union job.

Figure 5: Decomposition components: Non-union workers



Notes: Data from the CPS and the NLRB. Each series is represented as a log change from the corresponding 1980 level. Wage data is for non-union workers and is adjusted for worker characteristics. Each series corresponds to a decomposition component described in the main text.

The second panel in Table 7 shows the components of the non-union spillover effect for the full 1980-2010 sample period for different sub-groups. From this, we can see that the spillover effect for men is over double that for women, with most of that accounted for by differences in the transition effects. Similarly, the main reason that low-educated non-union workers were more affected by de-unionisation was because of a reduced chance of individual workers finding a union job rather than because of the reduced probability that their firm would be unionised.

7 Conclusion

In this paper, we provide new estimates of the impact of unions on non-union wage setting. We allow the presence of unions to affect non-union wages both through the typically discussed channel of non-union firms emulating union wages in order to fend off the threat of unionisation and through a bargaining channel in which non-union workers use the presence of union jobs as part of their outside option. We specify these channels in a search and bargaining model that includes union formation and the possibility of non-union firms responding to the threat of unionisation. By formalising wage setting and union formation, we derive a specification grounded in theory that provides guidance on what to control for, how to interpret our coefficients, and what is in the error term. Based on that, we derive a set of instruments and a model-based over-identification test, the values for which imply that our identification strategy is appropriate for this data.

Our estimates indicate that de-unionisation in the US after 1980 substantially affected non-union wages, particularly, and the wage structure in general. In a decomposition exercise, holding the probability a worker can find a union job, the probability a firm faces a unionisation drive, and union wage premia constant at their 1980 levels would have undone 35% of the 16% decline in the mean (composition constant) real wage in a typical city in the US between 1980 and 2010. While we find evidence for the spillover effects of unions on non-union wage setting through the traditional threat and bargaining channels, the latter dominates. That is important for policymakers looking for tools to help in raising wages. The union threat channel can only be implemented by increasing union power. However, the bargaining channel is not specific to unions. Any policy that raises worker outside option values will raise wages for a wide set of workers (Beaudry et al. [2012], Caldwell and Danieli [2021]). Unions are just one mechanism for doing that – though our estimates indicate a powerful and direct one. Finally, it is worth noting that what we have examined in this paper is only one path through which unions can affect labour market outcomes. When unions are stronger, there is also the possibility of their impacting elections and policy-making, shifting policy on labour market regulation and minimum wages that would have their own effects on the wage structure.

References

J. M. Abowd and H. S. Farber. Job queues and the union status of workers. *ILR Review*, 35(3):354–367, 1982.

K. G. Abraham and H. S. Farber. Returns to seniority in union and nonunion jobs: a new look at the evidence. *ILR Review*, 42(1):3–19, 1988.

K. G. Abraham and J. L. Medoff. Length of service and layoffs in union and nonunion work groups. *ILR Review*, 38(1):87–97, 1984.

K. G. Abraham and J. L. Medoff. Length of service and promotions in union and nonunion work groups. *ILR Review*, 38(3):408–420, 1985.

Ö. T. Açıkgöz and B. Kaymak. The rising skill premium and deunionization. *Journal of Monetary Economics*, 63:37–50, 2014.

R. Adao, M. Kolesar, and E. Morales. Shift-share designs: Theory and inference. *The Quarterly Journal of Economics*, 134(4):1949–2010, 2020.

J. T. Addison and B. T. Hirsch. Union effects on productivity, profits, and growth: Has the long run arrived? *Journal of Labor Economics*, 7(1):72–105, 1989.

D. H. Autor, D. Dorn, and G. H. Hanson. The china syndrome: Local labor market effects of import competition in the united states. *The American Economic Review*, 103(6): 2121–2168, 2013.

I. Bassier. Collective bargaining and spillovers in local labour markets. Discussion Paper 1895, Centre for Economic Performance, 2022.

P. Beaudry, D. A. Green, and B. Sand. Does industrial composition matter for wages? a test of search and bargaining theory. *Econometrica*, 80(3):1063–1104, 2012.

P. Beaudry, D. A. Green, and B. M. Sand. Spatial equilibrium with unemployment and wage bargaining: Theory and estimation. *Journal of Urban Economics*, 79:2–19, 2014.

K. Borusyak, P. Hull, and X. Jaravel. Quasi-experimental, shift-share research designs. *The Review of Economic Studies*, 89:181–213, 2022.

K. Bronfenbrenner. The role of union strategies in nlrb certification elections. *Industrial and Labor Relations Review*, 50(2):195–212, 1997.

K. Bronfenbrenner. No holds barred: The intensification of employer opposition to organizing. Briefing Paper 235, Economic Policy Briefing Paper, 2009.

S. Caldwell and O. Danieli. Outside options in the labor market. Technical report, Tel Aviv University, 2021.

S. Caldwell and N. Harmon. Outside options, bargaining, and wages: Evidence from coworker networks. Technical report, University of California at Berkeley, 2019.

D. Card. The effect of unions on wage inequality in the us labor market. *ILR Review*, 54 (2):296–315, 2001.

D. Card. Immigration and inequality. *American Economic Review*, 99(2):1–21, 2009.

D. Card, T. Lemieux, and W. C. Riddell. Unions and wage inequality. *Journal of Labor Research*, 25(4):519–559, 2004.

D. Card, A. R. Cardoso, J. Heining, and P. Kline. Firms and labor market inequality: Evidence and some theory. *Journal of Labor Economics*, 36(S1):s13–s70, 2018a.

D. Card, T. Lemieux, and W. C. Riddell. Unions and wage inequality: The roles of gender, skill and public sector employment, 2018b. NBER working paper w25313.

G. Clark and K. Johnston. The geography of us union elections 1: the crisis of us unions and a critical review of the literature. *Environment and Planning A*, 19:33–57, 1987.

P. Denice and J. Rosenfeld. Unions and nonunion pay in the united states, 1977-2015. *Sociological Science*, 5:541–561, 2018.

W. Dickens and L. F. Katz. Interindustry wage differences and industry characteristics. Technical report, National Bureau of Economic Research Cambridge, Mass., USA, 1986.

J. DiNardo and D. S. Lee. Economic impacts of new unionization on private sector employers: 1984-2001. *Quarterly Journal of Economics*, 119(4):1383–1441, 2004.

J. DiNardo and T. Lemieux. Diverging male wage inequality in the united states and canada, 1981–1988: Do institutions explain the difference? *ILR Review*, 50(4):629–651, 1997.

J. DiNardo, N. M. Fortin, and T. Lemieux. Labor market institutions and the distribution of wages, 1973-1992: A semiparametric approach. *Econometrica*, 64(5):1001–1044, 1996.

M.-P. Donsimoni. Union power and the american labour movement. *Applied Economics*, 13 (4):449–464, 1981.

H. S. Farber. Nonunion wage rates and the threat of unionization. *ILR Review*, 58(3): 335–352, 2005.

H. S. Farber, D. Herbst, I. Kuziemko, and S. Naidu. Unions and inequality over the twentieth century: New evidence from survey data. *The Quarterly Journal of Economics*, 136(3): 1325–1385, 2021.

S. Firpo, N. Fortin, and T. Lemieux. Decomposing wage distributions using recentered influence function regressions. *Econometrics*, 6(2):28, 2018.

N. M. Fortin, T. Lemieux, and N. Lloyd. Labor market institutions and the distribution of wages: The role of spillover effects. Technical report, University of British Columbia, 2019. *Working Paper*.

N. M. Fortin, T. Lemieux, and N. Lloyd. Labor market institutions and the distribution of wages: The role of spillover effects. *Journal of Labor Economics*, 39(S2):S369–S412, 2021. doi: 10.1086/712923. URL https://doi.org/10.1086/712923.

R. B. Freeman. Unionism and the dispersion of wages. *ILR Review*, 34(1):3–23, 1980.

R. B. Freeman and J. L. Medoff. The impact of the percentage organized on union and nonunion wages. *The Review of Economics and Statistics*, pages 561–572, 1981.

P. Goldsmith-Pinkham, I. Sorkin, and H. Swift. Bartik instruments: What, when, why, and how, 2018. NBER working paper w24408.

P. Goldsmith-Pinkham, I. Sorkin, and H. Swift. Bartik instruments: What, when, why, and how. *American Economic Review*, 110(8):2586–2624, August 2020. doi: 10.1257/aer.20181047. URL https://www.aeaweb.org/articles?id=10.1257/aer.20181047.

A. Gosling and T. Lemieux. Labour market reforms and changes in wage inequality in the united kingdom and the united states, 2001. NBER working paper w8413.

F. Hafner. The equilibrium effects of workers' outside employment options: Evidence from a labor market integration. Technical report, Alto University, 2022.

B. Hardy, S. Holla, E. Krause, and J. Ziliak. Inequality in the united states: 1975-2022. Technical report, Institute for Fiscal Studies, 2023.

J. J. Heckman. Sample selection bias as a specification error. *Econometrica*, 47(1):153–161, 1979.

B. T. Hirsch and A. N. Link. Unions, productivity, and productivity growth. *Journal of Labor Research*, 5(1):29–37, 1984.

B. T. Hirsch and J. L. Neufeld. Nominal and real union wage differentials and the effects of industry and smsa density: 1973-83. *The Journal of Human Resources*, 22(1):138–148, 1987.

T. J. Holmes. Geographic spillover of unionism. *NBER Working Paper Series*, 2006. URL https://api.semanticscholar.org/CorpusID:3105553.

H. J. Holzer. Unions and the labor market status of white and minority youth. *ILR Review*, 35(3):392–405, 1982.

G. Jarosch, J. S. Nimczik, and I. Sorkin. Granular search, market structure, and wages, 2019. NBER working paper w26239.

G. Jarosch, J. S. Nimczik, and I. Sorkin. Granular search, market structure, and wages. *Review of Economic Studies*, 2024.

M. Johnson, K. Lavetti, and M. Lipsitz. The labor market effects of legal restrictions on worker mobility, 2020. Available at SSRN: https://ssrn.com/abstract=3455381 or http://dx.doi.org/10.2139/ssrn.3455381.

L. M. Kahn. Union spillover effects on organized labor markets. *The Journal of Human Resources*, 15(1):87–98, 1980.

P. Krusell and L. Rudanko. Unions in a frictional labor market. *Journal of Monetary Economics*, 80:35–50, 2016.

T. Lemieux. Increasing residual wage inequality: Composition effects, noisy data, or rising demand for skill? *American Economic Review*, 96(3):461–498, 2006.

H. Lesch. Trade union density in international comparison. In *CESifo Forum*, number 4 in 5, pages 12–18, 2004.

H. G. Lewis. *Unionism and relative wages in the United States: an empirical inquiry.* University of Chicago press, 1963.

D. Neumark and M. L. Wachter. Union effects on nonunion wages: Evidence from panel data on industries and cities. *ILR Review*, 49(1):20–38, 1995.

C. A. Pissarides. Trade unions and the efficiency of the natural rate of unemployment. *Journal of Labor Economics*, 4(4):582–595, 1986.

M. Podgursky. Unions, establishment size, and intra-industry threat effects. *ILR Review*, 39 (2):277–284, 1986.

S. Rosen. Trade union power, threat effects and the extent of organization. *The Review of Economic Studies*, 36(2):185–196, 1969.

E. Sanderson and F. Windmeijer. A weak instrument f-test in linear iv models with multiple endogenous variables. *Journal of Econometrics*, 190(2):212–221, 2016.

J. Schmitt and A. Mitukiewicz. Politics matter: changes in unionisation rates in rich countries, 1960–2010. *Industrial Relations Journal*, 43(3):260–280, 2012.

I. G. Snoddy. Learning about selection: An improved correction procedure. Technical report, University of British Columbia, 2019. *Working Paper*.

J. H. Stock and M. M. Watson. Introduction to econometrics, 3rd international edition, 2011.

M. Taschereau-Dumouchel. The Union Threat. *The Review of Economic Studies*, 87(6): 2859–2892, 05 2020. ISSN 0034-6527.

J. Tschopp. Wage formation: Towards isolating search and bargaining effects from the marginal product. *The Economic Journal*, 127(603):1693–1729, 2017.

Online Appendix: The Impact of Unions on Non-union Wage Setting: Threats and Bargaining*

David A. Green† Ben M. Sand‡ Iain G. Snoddy§ Jeanne Tschopp¶

May 27, 2024

Contents

*We are grateful to Thomas Lemieux, Craig Riddell, and Raffaele Saggio for their comments. We also greatly benefited from comments from participants at seminars at the University of Chicago, the University of Melbourne, Carleton University, the University of British Columbia, the University College London, the University of Mannheim, and the Society of Labor Economists annual meeting.

†University of British Columbia and Research Fellow, IFS, London `david.green@ubc.ca`

‡Department of Economics, York University `bmsand@yorku.ca`

§Analysis Group, Montreal `iain.snoddy@analysisgroup.com`

¶University of Bern `jeanne.tschopp@unibe.ch`

A Main Notation

1. Indices:
    - τ: union status; $\tau = 1$ for union firms, $\tau = 2$ for simple non-union firms, and $\tau = 3$ for emulating non-union firms
    - i: industry
    - $j = \{\tau, i\}$: (source) job, defined as an industry-union status combination
    - $k = \{\tau', i'\}$: (destination) job, where τ' and i' denote the union status and industry
    - f: firm
2. Employment and wages:
    - n_{fjc}: firm f employment in job j and city c
    - v_{fjc}: firm f vacancies in job j and city c
    - N_c: employment in city c
    - N_{jc}: job j employment in city c
    - $N_{\tau c}$: union status τ employment in city c
    - N_{ic}: industry i employment in city c
    - η_{jc}: job j employment share within city c
    - $\eta_{\tau c}$: union status τ employment share within city c
    - w_{fjc}: wage in firm f, job j and city c
3. Matches, matching probabilities and transition probabilities:
    - $M_{kc|jc}$: number of matches of unemployed workers whose last job of type j to vacancies of job type k in city c
    - θ_{jc}: proportion of unemployed workers who were formerly in j
    - ϕ_{kc}: proportion of vacancies that come from k
    - M_c: number of matches observed in city c
    - U_c: number of unemployed workers in city c
    - Ω_c: number of vacancies in city c
    - $\chi_{kc|j}$: frictional cost of moving from j to k in city c
    - $\varphi_{k|j}$: nation-wise mobility friction in moving from job j to job k
    - q_c^v: probability a firm fills a vacancy in city c
    - q_{kc}^v: probability a firm fills a vacancy of job type k in city c
    - q_c^u: probability an unemployed worker makes a match in city c

- $q^u_{kc|j}$: probability an unemployed worker who was last employed at job j makes a match to a vacancy in k in city c
- $T_{kc|jc}$: probability that a worker formerly employed in job j in city c finds a job of type k, conditional on making any match

4. Discounted expected values and surpluses:

- V^E_{fjc}: worker expected value of employment in firm f, job j and city c
- V^U_{jc}: expected value of unemployment for a worker formerly employed in a job of type j in city c
- Π_{fjc}: firm f expected value of filled positions in job j and city c
- S_{fjc}: firm f surplus in job j and city c

5. Parameters:

- ψ_{fjc}: firm f, job j and city c amenity, set to zero for non-union firms (if $\tau = 2, 3$)
- ϵ_{fic}: firm-industry-city productivity shock
- σ_i: industry-specific span-of-control issues
- κ: flow cost per vacancy
- Υ: matching elasticity
- ρ: discount factor
- δ^m: death rate of matches
- δ^e: death rate of firms
- $\delta = \delta^e + (1 - \delta^e)\delta^m$: job destruction rate
- $\rho^e = (1 - \delta^e)\rho$: firms' effective discount factor

B Matching Process

Recall from the paper that we write the probability that a worker formerly employed in a job of type j (defined by industry and union status) matches with a vacancy of job type k as,

$$\begin{aligned} q^u_{kc|j} &= q^u_c T_{kc|j} \\ &= q^u_c \chi_{kc|j}(\varphi_{k|j})\eta_{kc} \\ &= q^u_c \frac{\varphi_{k|j}}{\sum_{k'} \eta_{k'c}\varphi_{k'|j}} \eta_{kc} \end{aligned}$$

where, q^u_c is the probability that any unemployed worker matches with a vacancy in any job type in city c, $T_{kc|j}$ is the probability a worker formerly in job type j in city c finds a job of type k, conditional on making any match, η_{kc} is the proportion of jobs in city c that are in industry k, $\chi_{kc|j}(\varphi_{k|j})$ is a term that captures the relative difficulty of moving from j to k in city c, and $\varphi_{k|j}$ is a specific mobility friction in moving from jobs of type j to jobs of type k (regardless of the city).

In this appendix, we demonstrate that, assuming that the city level matching function is constant returns to scale, this writing of the j, k specific matching rate as the product of the city level matching rate, the proportion of jobs in sector k in the city, and a factor reflecting the relative difficulty of moving from j to k is consistent with a steady state in which the size of each sector is stable. Furthermore, we can write q_c^u as a function of the city level employment rate, ER_c. In our empirical specification, this allows us to capture city level tightness with ER_c and use the relative probability of a worker coming from job type j matching with a firm of job type k in constructing outside option values.

B.1 Matching Function

Firms and workers operate in a labour market that includes frictions, meaning that workers and firms do not find each other and form a match perfectly easily. We assume that match formation depends both the job type (defined by industry and union sector) in which the vacancy is posted and the job type in which the worker was last employed. That is, we write, the number of effective matches between unemployed workers in city c who were previously employed in job j and vacancies posted by firms in job type k as $M_{kc|jc} = M(U_{kc|jc}, \Omega_{kc|jc})$, where $U_{kc|jc}$ is the number of unemployed workers who will find a job in k (in city c) and who were previously employed in job j (in city c), and $\Omega_{kc|jc}$ is the number of vacancies created by firms in job k for workers previously employed in job j, both in city c.

To generate the form for our matching function, we start with the total number of matches observed in a city, $M(U_c, \Omega_c)$, written (as is standard) as a function of the total number of unemployed workers, U_c, and the total number of vacancies, Ω_c in the city. If the frictions – what Tschopp (2017) refers to as the costs of making a match – are the same regardless of the job type, j, the worker is coming from and the job type, k, of the vacancy then the total number of vacancies involving a specific j, k combination would equal $\Omega_{kc|jc} = \theta_{jc}\Omega_c\phi_{kc}$, where, θ_{jc} denotes the proportion of city-specific matches involving workers that were previously employed in job j, and ϕ_{kc} is the proportion of matches that will be in job k. Similarly, the number of unemployed workers involved in matches of this type would equal $U_{kc|jc} = \theta_{jc}U_c\phi_{kc}$. But, in fact, it may be particularly difficult for a match to form for certain j, k pairs. For example, a match between a worker who formerly worked in a unionised construction job and a vacancy posted by a union firm in the steel industry may be particularly easy to consummate while a match between that same worker and a non-union legal services firm may be less likely to actually happen. Given that, we assume that there is an additional component capturing the relative costs of creating effective matches, $\chi_{kc|j}(\varphi_{k|j})$, which we write as a function of $\varphi_{k|j}$, representing the specific mobility frictions in moving from jobs of type j to jobs of type k (regardless of city).[1] Then, the effective number of unemployed workers relevant for matches in job type k involving workers from job type j is actually, $U_{kc|jc} = \theta_{jc}U_c\phi_{kc}\chi_{kc|j}(\varphi_{k|j})$, and the effective number of vacancies is, $\Omega_{kc|jc} = \theta_{jc}\Omega_c\phi_{kc}\chi_{kc|j}(\varphi_{k|j})$

[1] In this characterization of differential match rates as reflecting differential frictional costs, we follow Tschopp (2017). Bassier (2022), alternatively, refers to observed differences in worker movements across firms as reflecting differences in 'consideration sets'.

Given this, we write:

$$M_{kc|jc}(U_{kc|jc}, \Omega_{kc|jc}) = \theta_{jc} M(U_c, \Omega_c) \phi_{kc} \chi_{kc|j}(\varphi_{k|j}), \tag{1}$$

Assuming (as is standard) that $M(U_c, \Omega_c)$ is constant returns to scale (CRS), $M_{kc|jc}$ is also CRS. We will assume that $\chi_{kc|j}(\varphi_{k|j})$ takes the specific form:

$$\chi_{kc|j} = \frac{\varphi_{k|j}}{\sum_{k'} \eta_{k'c} \varphi_{k'|j}} \quad \forall k \tag{2}$$

Given all of this, the structure of our matching function implies that the number of effective matches depends on an origin (previous job) component, a destination (next job) component, and a component reflecting the ability of a worker to switch from the origin to the destination job.

Mobility frictions across jobs create a wedge between potential and realized matches, as captured by $\chi_{kc|j}(\varphi_{k|j})$, and the larger the frictions the higher the number of meetings that are required to reach a given level of effective matches. Since $M(U_c, V_c)$ yields effective local matches, the variables θ_{jc}, ϕ_{kc} and $\chi_{kc|j}(\varphi_{k|j})$ all adjust to guarantee a consistent matching process in steady state.

B.2 Matching in Steady State

In steady state, on one side, the number of effective matches in job k must equal the number of type k jobs that are destroyed, i.e. $\delta N_{kc} = M_{kc}$, where δ is the job destruction rate, N_{kc} is employment in job k and city c and $M_{kc} = \sum_j M_{kc|jc}$. Note that the job destruction rate is given by $\delta = \delta^e + (1 - \delta^e)\delta^m$, where δ^e is the death rate of firms and δ^m is the death rate of matches. Summing across jobs, this steady state condition implies $\delta N_c = M_c$, where N_c is local employment and M_c is short form for $M(U_c, V_c)$. Consequently, $\frac{\delta N_{kc}}{\delta N_c} = \frac{M_{kc}}{M_c}$. Hence, the proportion of employment observed to be in industry k, η_{kc}, must equal $\frac{M_{kc}}{M_c}$, and it follows that, in steady state, the proportion of effective matches that will be in job k equals the employment share in that job, i.e. $\phi_{kc} = \eta_{kc}$.

The second steady state requirement is that effective matches must aggregate properly in a way that (i) $\sum_k M_{kc|jc} = M_{c|jc}$, where $M_{c|jc}$ denotes the number of city-specific effective matches made with workers previously employed in j, and (ii) $\sum_j M_{kc|jc} = M_{kc}$. Proper aggregation determines both θ_{jc} and $\chi_{kc|j}(\varphi_{k|j})$.

In particular, using equation (1) and the steady state condition that $\phi_{kc} = \eta_{kc}$, along with the fact that $\theta_{jc} = M_{c|jc}/M_c$, we have that

$$\begin{aligned}
\sum_k M_{kc|jc} &= M_c \theta_{jc} \sum_k \eta_{kc} \chi_{kc|j}(\varphi_{k|j}) \\
&= M_{c|jc} \sum_k \eta_{kc} \chi_{kc|j}(\varphi_{k|j}),
\end{aligned}$$

and, hence, $\sum_k M_{kc|jc}$ equates $M_{c|jc}$ if $\sum_k \eta_{kc} \chi_{kc|j}(\varphi_{k|j}) = 1$. This condition naturally holds given our assumed form for $\chi_{kc|j}(\varphi_{k|j})$. Therefore, $\chi_{kc|j}$ can be interpreted as a relative mobility friction from job j to job k, relative to the friction of moving anywhere else in the economy. When the frictions of moving across jobs is identical across origin and destination jobs (e.g. moving across any two jobs is equally costly to all workers), then $\chi_{kc|j}$ simplifies to $\chi_{kc|j} = 1$ and the wedge between potential and effective matches disappears.

Finally, using again equation (1) and $\phi_{kc} = \eta_{kc}$, we have that

$$\begin{aligned}\sum_j M_{kc|jc} &= M_c \eta_{kc} \sum_j \theta_{jc} \chi_{kc|j}(\varphi_{k|j}) \\ &= M_{kc} \sum_j \theta_{jc} \chi_{kc|j}(\varphi_{k|j}),\end{aligned}$$

and, thus, $\sum_j M_{kc|jc}$ equals M_{kc} (so that we have a consistent summing up of matches in k with respect to the specific matches in k for workers coming from all the other possible job types) if $\sum_j \theta_{jc} \chi_{kc|j}(\varphi_{k|j}) = 1$, or, using equation (2), if

$$\sum_j \theta_{jc} \frac{\varphi_{k|j}}{\sum_{k'} \eta_{k'c} \varphi_{k'|j}} = 1 \quad \forall k \tag{3}$$

Equation (3) yields a set of K equations, where K is the cardinality of jobs in the economy, which, given frictions, steady state job shares and the fact that the number of jobs is fixed in the economy, jointly determine the set of $\{\theta_{jc}\}_j$, within city c.

B.3 Implied Job Filling and Job Finding Probabilities

The probability that a firm in job k fills a job is given by the ratio between effective matches and vacancies, i.e. $q^v_{kc} = \frac{M_{kc}}{\Omega_{kc}}$, where $\Omega_{kc} = \sum_j \Omega_{kc|jc}$. In steady state, since $\sum_j \theta_{jc} \chi_{kc|j}(\varphi_{k|j}) = 1$, this ratio simplifies to

$$\begin{aligned} q^v_{kc} &= \frac{\eta_{kc} M_c}{\sum_j \Omega_{kc|jc}} \\ &= \frac{\eta_{kc} M_c}{\Omega_c \eta_{kc} \sum_j \theta_{jc} \chi_{kc|j}(\varphi_{k|j})} \\ &= \frac{M_c}{\Omega_c} \\ &= q^v_c, \end{aligned} \tag{4}$$

where $q^v_c = \frac{M_c}{\Omega_c}$. Hence, in steady state, the probability that a firm fills a job is independent of the job and only depends on the local matching process.

Similarly, the probability that a worker previously employed in j finds a job in k is given by the ratio of effective matches and unemployed workers, i.e. $q^u_{kc|jc} = \frac{M_{kc|jc}}{U_{c|jc}}$, where

$U_{c|jc} = \sum_k U_{kc|jc}$. In steady state, since $\sum_k \eta_{kc}\chi_{kc|j}(\varphi_{k|j}) = 1$, this ratio simplifies to

$$\begin{aligned} q^u_{kc|jc} &= \frac{\theta_{jc} M_c \eta_{kc}\chi_{kc|j}(\varphi_{k|j})}{\sum_k U_{kc|jc}} \\ &= \frac{\theta_{jc} M_c \eta_{kc}\chi_{kc|j}(\varphi_{k|j})}{\theta_{jc} U_c \sum_k \eta_{kc}\chi_{kc|j}(\varphi_{k|j})} \\ &= \frac{M_c}{U_c}\eta_{kc}\chi_{kc|j}(\varphi_{k|j}) \\ &= q^u_c \eta_{kc}\chi_{kc|j}(\varphi_{k|j}) \\ &= q^u_c T_{kc|j}, \end{aligned} \tag{5}$$

where $q^u_c = \frac{M_c}{U_c}$, and where we define $T_{kc|j} = \eta_{kc}\chi_{kc|j}(\varphi_{k|j})$. Hence, the probability that a worker previously employed in j finds a job in k depends on the proportion of jobs in k in their city, the mobility friction $\varphi_{k|j}$ and all the job shares in the economy (through $\chi_{kc|j}$).

In steady state, as well, we can write, $q^u_c = \frac{M_c}{U_c} = \frac{\delta N_c/L_c}{U_c/L_c} = \frac{\delta ER_c}{1-ER_c}$. That is, the local matching rate for unemployed workers can be written as a function of the employment rate, ER_c and parameters of the matching function. Note that this does not require an assumption on the form of the matching function. We can get a similar expression for q^v_c, the probability a vacancy matches to a worker if we assume the matching function is CRS. For example, assume the matching function takes a Cobb-Douglas form. In that case, in steady state:

$$\frac{J_c}{L_c} - ER_c = \left[\delta ER_c (1 - ER_c)^{-\Upsilon}\right]^{\frac{1}{1-\Upsilon}}, \tag{6}$$

where Υ is the matching elasticity and J_c is the number of jobs that are created in city c and L_c is the population in city c.[2] Therefore,

$$\begin{aligned} q^v_c &= \frac{M_c}{V_c} \\ &= \frac{\delta N_c/L_c}{(J_c - N_c)/L_c} \\ &= \frac{\delta ER_c}{[\delta ER_c(1-ER_c)^{-\Upsilon}]^{\frac{1}{1-\Upsilon}}} \\ &= \left[\frac{1-ER_c}{\delta ER_c}\right]^{\frac{\Upsilon}{1-\Upsilon}} \end{aligned} \tag{7}$$

[2] In fact,

$$\begin{aligned} M_c &= M(U_c, \Omega_c) \\ \delta N_c &= M\left[(L_c - N_c), (J_c - N_c)\right] \\ \delta N_c/L_c &= M\left(\left[1 - (N_c/L_c)\right], \left[(J_c/L_c) - (N_c/L_c)\right]\right) \\ \delta ER_c &= M((1-ER_c), [(J_c/L_c) - ER_c]) \\ \delta ER_c &= (1-ER_c)^{\Upsilon}\left[(J_c/L_c) - ER_c\right]^{1-\Upsilon} \end{aligned}$$

C Deriving Wages and Firm Sizes

We consider an environment with C local economies and I industries. To start, we assume that workers are homogeneous in terms of preferences and skills.

In our model, workers and firms can (endogenously) end up in one of three types of arrangements: simple non-union firms; non-union firms that emulate union wages; or union firms. When workers and firms meet, there is a match-specific surplus that depends on firm productivity and worker and firm outside options. Following Taschereau-Dumouchel (2020) (hereafter, TD), if the firm is non-union, a wage is established in bargaining between individual workers and the firm. In each period, non-union workers can vote to unionise, and if they do, then the bargaining is between the whole set of workers and the firm. However, we can divide non-union jobs into two types. The first are 'simple' non-union jobs in which workers have no incentive to unionise. The second corresponds to an intermediate set of productivity, amenity and cost of unionisation values such that workers would choose to unionise in a world where firms cannot respond but, instead, the workers and firm will agree to remain in a non-union arrangement with a higher wage that eliminates the desire of workers to unionise. For the 'simple' non-union and union states, the relevant value, if bargaining breaks down, is an option outside the match (non-production for firms and unemployment for workers). For the state where non-union firms emulate union wages, the breakdown values are those associated with continuing the match but in a unionised arrangement. It is the threat of that unionised alternative that allows workers to bargain a better wage (what we call an emulation wage) while remaining non-union.

We will focus on steady states. In the steady state, it is clear what arrangement each firm operates under and, so, we will write the relevant value functions as reflecting the firm and workers continuing with the arrangement.

The timing in the model is the following. Firms first choose their optimal employment levels through choosing their number of vacancies. They do this knowing what the outcome of the bargaining process will be and, in steady state, which type of union status they will face. Next, the workers in the firm decide whether to unionise. After that, one of the three types of wage bargaining takes place as described above.

C.1 Firm Problem

Let j denote a job, defined as an industry-union status object, i.e. $j = \{\tau, i\}$, where τ denotes the union status and i is an industry. We let $\tau = 1$ for union, $\tau = 2$ for simple non-union, and $\tau = 3$ for the emulating non-union status. We use the subscript k for potential destination jobs, with destination union status and industry denoted τ' and i', respectively, such that $k = \{\tau', i'\}$. Whenever possible, we work with equations at the job level, using the union status type and industry subscripts only where needed. For notational clarity, we also drop the (firm-job-city-specific) subscript on firm employment and vacancies.

The production technology of a firm f in job j and city c takes the following form:

$$y_{fjc}(n) = \epsilon_{fic} n - \frac{1}{2}\sigma_i n^2,$$

where ϵ_{fic} is a firm-specific productivity draw, n denotes firm employment, and σ_i captures industry-specific span-of-control issues. For later purposes, it is useful to decompose the productivity draw as $\epsilon_{fic} = \epsilon_{ic} + u_{fic}$, where ϵ_{ic} is a local sector-wide productivity term and u_{fic} is a mean zero firm-specific component. Note that the technology is not a function of τ, i.e., union status alters production only by altering employment levels.

At the beginning of each period, firms choose the optimal number of vacancies (and, so, optimal employment) given the wage (specified as a function of firm employment) they know will be bargained with their workers later. To simplify, we assume that κ, the flow cost of hiring, is linear in the number of vacancies posted. Since δ^m matches are randomly destroyed in each period, a firm which had n_{-1} workers in the previous period enters the current period with $(1-\delta^m)n_{-1}$ workers. From this, it knows the number of vacancies, v it must post in order to have n workers for production in the current period. It then posts v vacancies to produce with n workers in the current period. Hence, the firm value function of filled positions is given by:

$$\begin{aligned} \Pi_{fjc}(n_{-1}) = \max_{v} \quad & [p_i y_{fjc}(n) - w_{fjc}(n)n - \kappa v + \rho^e \Pi_{fjc}(n)] \\ \text{s.t.} \quad & n = n_{-1}(1-\delta^m) + q_c^v v, \end{aligned} \tag{8}$$

where p_i is the price of the industry i good and w_{fjc} denotes the wage offered by firm f for job j in city c. ρ^e is a firm effective discount rate; in particular $\rho^e = (1-\delta^e)\rho$, where δ^e is the death rate of firms and ρ the discount factor. δ^m is the destruction rate of matches and q_c^v is the local probability a vacancy meets a worker. Note that we assume that firms are producing tradable goods that are sold on the national market at the given price, p_i. We assume that union amenities are created by the union itself and, so, do not enter the cost function of the union firm.

The first order condition (FOC) in problem (8) is:[3]

$$\frac{\partial \pi_{fjc}(n)}{\partial n} = [1 - \rho^e(1-\delta^m)] \frac{\kappa}{q_c^v}, \tag{9}$$

where $\pi_{fjc}(n) = [p_i y_{fjc}(n) - w_{fjc}(n)n]$ denotes current-period profits. Hence, in steady-state, marginal profits equate the marginal cost of creating a new vacancy, adjusted for the fact that a portion $(1-\delta^m)$ of workers staying with the firm lowers hiring costs. From the FOC it is clear that, since recruitment costs are linear in employment, the dynamic problem of the firm is equivalent to a static problem. In consequence, the firm starts with $(1-\delta^m)$ of its optimal employment at the beginning of the period and instantly hires back to its optimal level.

The closed-form solution for optimal firm size depends on the bargained wage equation. Given that, we first derive that equation and then return to the derivation of optimal firm size for each of the three union status arrangements.

[3]The FOC is:

$$\frac{\partial \Pi_{fjc}(n_{-1})}{\partial n} = 0$$

where, due to the enveloppe theorem,

$$\frac{\partial \Pi_{fjc}(n_{-1})}{\partial n} = \frac{\partial \pi_{fjc}(n)}{\partial n} - \frac{\kappa}{q_c^v} + \rho^e \frac{\partial \Pi_{fjc}(n)}{\partial n} = \frac{\partial \pi_{fjc}(n)}{\partial n} - \frac{\kappa}{q_c^v} + \rho^e \frac{\kappa(1-\delta^m)}{q_c^v}$$

C.2 Firm Surplus

In order to derive the bargained wage, we first stipulate firm and worker surpluses for the various union arrangements.

C.2.1 Union Firms

As we will see, productivity, amenity and unionisation cost values imply that in some firms it is so beneficial for the workers to unionise that there is no advantage to the firm to try to resist. As mentioned earlier, in steady state, these firms will also be union firms in subsequent periods.

The match surplus for a union firm is given by the difference between the discounted profits from a successful and a failed bargain:

$$S_{fjc}(n) = [\pi_{fjc}(n) + \rho^e \Pi_{fjc}(n)] - [\pi_{fjc}(0) + \rho^e \Pi_{fjc}(0)] \quad \forall j = \{1, i\} \tag{10}$$

where, $\pi_{fjc}(0)$ and $\Pi_{fjc}(0)$ are the flow profits and value of the firm with no workers, respectively. At the point of bargaining, the number of workers in the firm is fixed and the hiring cost is sunk. For this reason, the recruitment cost does not appear in the current period.

A successful outcome allows the firm to produce with n workers. In the next period, in steady state, the firm replaces its lost workforce $\delta^m n$, implying recruitment costs of $\frac{\kappa}{q_c^v}\delta^m n$. Hence, the first term in brackets in equation (10) can be rewritten as:

$$\pi_{fjc}(n) + \rho^e \Pi_{fjc}(n) = \pi_{fjc}(n) + \rho^e \left[\pi_{fjc}(n) - \delta^m \frac{\kappa}{q_c^v} n + \rho^e \Pi_{fjc}(n)\right] \quad \forall j = \{1, i\} \tag{11}$$

If the firm and the union fail to reach an agreement, the firm loses its entire workforce, fails to produce and faces the cost $\frac{\kappa}{q_c^v} n$ of recruiting the entire workforce again in the next period. In the absence of any transitional dynamics, the firm jumps back to steady state in the following period. Therefore, the second term in brackets in equation (10) is given by:

$$\pi_{fjc}(0) + \rho^e \Pi_{fjc}(0) = \rho^e \left[\pi_{fjc}(n) - \frac{\kappa}{q_c^v} n + \rho^e \Pi_{fjc}(n)\right] \quad \forall j = \{1, i\} \tag{12}$$

where $\pi_{fjc}(0) = 0$.

Subtracting equation (12) from equation (11), we obtain:

$$S_{fjc}(n) = \pi_{fjc}(n) + \rho^e (1 - \delta^m) \frac{\kappa}{q_c^v} n \quad \forall j = \{1, i\} \tag{13}$$

C.2.2 Simple Non-union Firms

For simple non-union firms, the surplus relates to the loss in value from losing just one worker. Following TD, we first calculate the effect of losing h marginal units of labour and then send h to zero in order to get the marginal contribution of a single worker. In doing this, we make use of the expression $n-h$ to refer to the removal of h workers from the number of hires n and write the firm's surplus from having n versus $n-h$ workers as

$$S_{fjc}(n) = [\pi_{fjc}(n) + \rho^e \Pi_{fjc}(n)] - [\pi_{fjc}(n-h) + \rho^e \Pi_{fjc}(n-h)] \quad \forall j = \{2, i\} \tag{14}$$

where, as before, hiring costs are sunk at the time of bargaining in the current period. The first term in the brackets captures the discounted profits from a successful bargain and is given by equation (11) (except that we are now dealing with $j = \{2, i\}$ type firms).

If negotiations fail, the firm loses h workers in the current period and produces with $n-h$ workers. In the next period, the firm immediately moves back to its optimal employment size, implying that it needs to post $\frac{n-(1-\delta^m)(n-h)}{q_c^v}$ vacancies.

Therefore, the second term in the brackets in equation (14) can be rewritten as:

$$\begin{aligned} \pi_{fjc}(n-h) + \rho^e \Pi_{fjc}(n-h) &= \pi_{fjc}(n-h) + \rho^e \left[\pi_{fjc}(n) - \delta^m \frac{\kappa}{q_c^v}(n-h) - \frac{\kappa}{q_c^v} h + \rho^e \Pi_{fjc}(n) \right] \\ \forall \quad & j = \{2, i\} \end{aligned} \tag{15}$$

Substituting equations (11) and (15) into (14), and rearranging yields:

$$S_{fjc}(n) = \pi_{fjc}(n) - \pi_{fjc}(n-h) + \rho^e (1-\delta^m) \frac{\kappa}{q_c^v} h \quad \forall j = \{2, i\} \tag{16}$$

Dividing by h and taking the limit $\lim_{h \to 0}$ yields the following expression for the firm surplus:

$$\lim_{h \to 0} \frac{S_{fjc}(n)}{h} = \lim_{h \to 0} \left[\frac{\pi_{fjc}(n) - \pi_{fjc}(n-h)}{h} \right] + \rho^e (1-\delta^m) \frac{\kappa}{q_c^v} \quad \forall j = \{2, i\} \tag{17}$$

Therefore:

$$S_{fjc}(n) \;\; = \;\; \frac{\partial \pi_{fjc}(n)}{\partial n} + \rho^e (1-\delta^m) \frac{\kappa}{q_c^v}, \quad \forall j = \{2, i\} \tag{18}$$

where $\frac{\partial \pi_{fjc}(n)}{\partial n} = p_i \frac{\partial y_{fjc}(n)}{\partial n} - w_{fjc}(n) - n \frac{\partial w_{fjc}(n)}{\partial n}$.

C.2.3 Emulating Non-union Firms

In the third type of firms, workers and the firm bargain under the threat of unionisation. We assume that, in steady state, the firm knows it is an emulating type and picks its employment level optimally given that. We also assume that it cannot alter that employment level if bargaining breaks down and it is unionised. This fits with a legal environment in which firms are not allowed to punish workers for unionising by laying workers off.

Given this, the match surplus of an emulating non-union firm is given by:

$$S_{fjc} = [\pi_{fjc}(n) + \rho^e \Pi_{fjc}(n)] - \left[\pi^*_{f1ic}(n) + \rho^e \Pi^*_{f1ic}(n) \right] \quad \forall j = \{3, i\} \tag{19}$$

where the first term in the brackets reflects the discounted profits following successful bargaining and where n is employment in the emulating non-union firm. The second term in the brackets captures the discounted profits if the negotiation breaks down, with $\pi^*_{f1ic}(n)$ denoting current-period profits and $\Pi^*_{f1ic}(n)$ denoting the discounted value, both under the threat; i.e., if the firm is unionised but employs n workers.

As before, in steady state, $\Pi_{fjc}(n) = \pi_{fjc}(n) - \delta^m \frac{\kappa}{q^v_c} n + \rho^e \Pi_{fjc}(n)$, such that $\Pi_{fjc}(n) = \frac{1}{1-\rho^e}\left[\pi_{fjc}(n) - \delta^m \frac{\kappa}{q^v_c} n\right]$. Therefore, the first term in equation (19) can be rewritten as:

$$\begin{aligned}
\pi_{fjc}(n) + \rho^e \Pi_{fjc}(n) &= \pi_{fjc}(n) + \frac{\rho^e}{1-\rho^e}\left[\pi_{fjc}(n) - \delta^m \frac{\kappa}{q^v_c} n\right] \\
&= \frac{1}{1-\rho^e}\pi_{fjc}(n) - \frac{\rho^e}{1-\rho^e}\delta^m \frac{\kappa}{q^v_c} n \quad \forall j = \{3, i\} \qquad (20)
\end{aligned}$$

If bargaining breaks down, the workforce unionises and the firm pays the union wage w_{f1ic}. Employment, however, remains unaltered. Given all of this, the discounted profits from a failed bargain are:

$$\pi^*_{f1ic}(n) + \rho^e \Pi^*_{f1ic}(n) = \frac{1}{1-\rho^e}\pi^*_{f1ic}(n) - \frac{\rho^e}{1-\rho^e}\delta^m \frac{\kappa}{q^v_c} n \qquad (21)$$

where $\pi^*_{f1ic}(n) = p_i y_{f3ic}(n) - w_{f1ic} n$. Current profits from a failed and a successful bargain are similar except that the former are computed using the union wage. Moreover, since optimal firm size remains fixed regardless of the outcome of the negotiation, discounted profits involve discounted recruitment costs that are identical in equations (20) and (21).

Subtracting equation (21) from (20) the surplus of an emulating non-union firm is given by:

$$\begin{aligned}
S_{fjc} &= \frac{1}{1-\rho^e}\left[\pi_{fjc}(n) - \pi^*_{f1ic}(n)\right] \\
&= \frac{1}{1-\rho^e}(w_{f1ic} - w_{fjc})n \quad \forall j = \{3, i\} \qquad (22)
\end{aligned}$$

Hence, the surplus of an emulating non-union firm exclusively depends on the wage differential between the union and emulating non-union status, and on its optimal firm size.

C.3 Worker Surplus

C.3.1 Expected Values of Employment and Unemployment

A worker can be either employed or unemployed during each time period. We start by deriving the expected values of employment and unemployment. We will derive worker surplus in the different union and non-union arrangements in the following sections.

Current-period income, if employed in firm f and job j, is given by the sum of the wage w_{fjc} and potential amenities ψ_{fjc}. At the end of the period, there is a probability δ of losing the job and becoming unemployed. This probability depends on both the death rate of firms and the death rate of matches; i.e. $\delta = \delta^e + (1 - \delta^e)\delta^m$. Hence, the discounted expected utility of employment is represented as:

$$V^E_{fjc}(w_{fjc}) = w_{fjc} + \psi_{fjc} + \rho\left[\delta V^U_{fjc} + (1-\delta)V^E_{fjc}(w'_{fjc})\right], \tag{23}$$

where the value of the amenity, ψ_{fjc}, is zero if job j is associated with a non-union firm, and where ρ is the (effective) discount rate of the worker. V^U_{fjc} is the expected value of unemployment (the value of search) for a worker previously employed in firm f, job j and city c. In this specification, w'_{fjc} corresponds to the wage that will be paid in the next period if the job is not terminated. Following TD, we assume that workers and firms believe that the next period's wage will be set optimally and that they cannot affect it through actions during this period.

Note that workers will assume the same expected wage holds for all future periods. As a result,

$$V^E_{fjc}(w'_{fjc}) = w'_{fjc} + \psi_{fjc} + \rho\left[\delta V^U_{fjc} + (1-\delta)V^E_{fjc}(w'_{fjc})\right], \tag{24}$$

Rearranging and plugging back into equation (23), yields:

$$V^E_{fjc}(w_{fjc}) = w_{fjc} + \mathbb{A}\psi_{fjc} + \rho(1-\delta)\mathbb{A}w'_{fjc} + \rho\delta\mathbb{A}V^U_{fjc} \tag{25}$$

where $\mathbb{A} = \frac{1}{1-\rho(1-\delta)}$, and $\mathbb{A} > 1$.

In the current period, an unemployed worker previously employed in job j earns unemployment benefits b. In the subsequent period, they search for work opportunities across all available jobs and find employment in job k with probability $q^u_c T_{kc|j}$, the steady state transition probability derived from the matching function, as described in equation (5). With probability $(1 - q^u_c)$ they remain unemployed. Hence, the discounted expected unemployment value is given by:

$$V^U_{fjc} = b + \rho\left[q^u_c \sum_k T_{kc|j} V^E_{kc}(w'_{kc}) + (1 - q^u_c)V^U_{fjc}\right], \tag{26}$$

where the relevant value of employment $V^E_{kc}(w'_{kc})$ will depend on the expected wage next period, w'_{kc}, which is an average of next period's wage across all firms in the kc cell.

Note that none of the parameters or driving forces in equation (26) depend on the firm. Hence $V^U_{fjc} = V^U_{jc}$, and solving for V^U_{jc} :

$$V^U_{jc} = b\mathbb{A}_c + \rho q^u_c \mathbb{A}_c \sum_k T_{kc|j} V^E_{kc}(w'_{kc}), \tag{27}$$

where $\mathbb{A}_c = \frac{1}{1-\rho(1-q^u_c)} > 1$.

To further simplify equation (27), we now derive an expression for $\sum_k T_{kc|j} V^E_{kc}(w'_{kc})$, the expected value of employment, conditional on having been employed in job j prior to unemployment. In what follows, for notational simplicity, we drop the dependence on w'_{kc} in $V^E_{kc}(w'_{kc})$.

First, averaging equation (24) across firms, we obtain the following expected value of employment in job k and city c:

$$V^E_{kc} = w'_{kc} + \psi_{kc} + \rho\left[\delta V^U_{kc} + (1-\delta)V^E_{kc}\right],$$

where ψ_{kc} is the expected amenity for job k in city c. Solving for V^E_{kc}, we have:

$$V^E_{kc} = \mathbb{A}w'_{kc} + \mathbb{A}\psi_{kc} + \rho\delta\mathbb{A}V^U_{kc} \tag{28}$$

Then, substituting (27) into (28) we obtain:

$$V^E_{kc} = \mathbb{A}w'_{kc} + \mathbb{A}\psi_{kc} + \rho\delta\mathbb{A}b\mathbb{A}_c + \rho^2\delta\mathbb{A}q^u_c\mathbb{A}_c\sum_{k'}T_{k'c|k}V^E_{k'c}, \tag{29}$$

where, for clarity, we have used k' subscripts to denote potential destination jobs.

Finally, pre-multiplying by the relevant transition probabilities (where $\sum_k T_{kc|j} = 1$) and summing across all jobs, we get:

$$\sum_k T_{kc|j}V^E_{kc} \;=\; \mathbb{A}\sum_k T_{kc|j}w'_{kc} + \mathbb{A}\sum_k T_{kc|j}\psi_{kc} + \rho\delta\mathbb{A}b\mathbb{A}_c + \rho^2\delta\mathbb{A}q^u_c\mathbb{A}_c\sum_k T_{kc|j}\sum_{k'}T_{k'c|k}V^E_{k'c}$$

To simplify the last term of this equation, we assume that the mobility terms $\varphi_{k|j}$s are path-independent, i.e. $\varphi_{k'|k}\varphi_{k|j} = \varphi_{k'|j}$, as in Tschopp (2017). This assumption implies that $\chi_{k'c|k}\chi_{kc|j} = \chi_{k'c|j}$ and, in consequence, $T_{k'c|k}T_{kc|j} = T_{k'c|j}\eta_{kc}$.[4] It follows that:

$$\sum_k T_{kc|j}V^E_{kc} \;=\; \mathbb{A}\sum_k T_{kc|j}w'_{kc} + \mathbb{A}\sum_k T_{kc|j}\psi_{kc} + \rho\delta\mathbb{A}b\mathbb{A}_c + \rho^2\delta\mathbb{A}q^u_c\mathbb{A}_c\sum_{k'}T_{k'c|j}V^E_{k'c}\underbrace{\sum_k \eta_{kc}}_{=1}$$

Solving for $\sum_k T_{kc|j}V^E_{kc}$, we obtain:

$$\sum_k T_{kc|j}V^E_{kc} = \Gamma_c + \mathbb{B}_c\sum_k T_{kc|j}w_{kc}, \tag{30}$$

where $\Gamma_c = \psi\mathbb{B}_c + \rho\delta b\mathbb{A}_c\mathbb{B}_c > 0$ and $\mathbb{B}_c = \frac{\mathbb{A}}{1-\rho^2\delta\mathbb{A}q^u_c\mathbb{A}_c} > 0$. Note that ψ_{kc} can be decomposed into two components, a mean zero job-city component $\widetilde{\psi}_{kc}$ independent of factors that determine $T_{kc|j}$, and a constant component ψ. Given this, $\sum_k T_{kc|j}\widetilde{\psi}_{kc}$ is, on average, equal to zero and we are left with the constant term ψ.

Equation (30) shows that the expected value of employment conditional on being previously employed in job j depends on the weighted average of wages in different jobs available in the

[4] Recall from B that $T_{kc|j} = \chi_{kc|j}\eta_{kc}$, where $\chi_{kc|j} = \frac{\varphi_{k|j}}{\sum_{k'}\varphi_{k'|j}\eta_{k'c}}$ is a relative mobility term capturing the ease with which a worker in job j can move to job k relative to moving to any other job in city c (including the current job), where $\varphi_{k|j}$ is a nation-wise mobility friction from job j to job k, and η_{kc} is the employment share of job k in city c.

local economy, with weights capturing the probability that a worker can actually transition to these jobs. As shown in B, these transition probabilities depend on both the proportions of these jobs in the economy and mobility frictions across the different jobs. We will refer to this average wage as the worker's 'outside option'.

We can now substitute equation (30) into equation (27) to obtain an expression for the expected unemployment value:

$$V_{jc}^{U} = \mathbb{D}_c + \rho q_c^u \mathbb{A}_c \mathbb{B}_c \sum_k T_{kc|j} w'_{kc}, \tag{31}$$

where $\mathbb{D}_c = \mathbb{A}_c \left[\rho\psi q_c^u \mathbb{B}_c + b\left(1 + \rho^2 \delta q_c^u \mathbb{A}_c \mathbb{B}_c\right)\right] > 0$.

C.3.2 Worker Surplus in Union and Simple Non-union Firms

For workers bargaining in either union or simple non-union arrangements, the relevant outside option is losing their jobs and ending up unemployed. The surplus in those cases is obtained by subtracting equation (31) from (25):

$$\begin{aligned}
V_{fjc}^{E} - V_{fjc}^{U} &= w_{fjc} + \mathbb{A}\psi_{fjc} + \rho(1-\delta)\mathbb{A}w'_{fjc} - (1-\rho)\mathbb{A}V_{jc}^{U} \\
&= w_{fjc} + \mathbb{A}\psi_{fjc} + \rho(1-\delta)\mathbb{A}w'_{fjc} - (1-\rho)\mathbb{A}\left[\mathbb{D}_c + \rho q_c^u \mathbb{A}_c \mathbb{B}_c \sum_k T_{kc|j} w'_{kc}\right] \\
&\forall \quad j = \{1, i\} \text{ or } \{2, i\}
\end{aligned} \tag{32}$$

where $\psi_{fjc} = 0$ if the worker is currently employed in a non-union firm, i.e. $\forall j = \{2, i\}$.

C.3.3 Worker Surplus in Emulating Non-union Firms

For workers who are in emulating type non-union arrangements, the relevant outside option to a breakdown in bargaining with the firm is to revert to their threat and unionise the firm, after paying a one-shot unionisation cost, λ_c^*. So, for these workers, using equation (25), the relevant surplus is:

$$\begin{aligned}
V_{fjc}^{E} - \left(V_{f1ic}^{E} - \lambda_c^*\right) &= (w_{fjc} - w_{f1ic}) - \mathbb{A}\psi_{f1ic} + \rho(1-\delta)\mathbb{A}(w'_{fjc} - w'_{f1ic}) \\
&+ \rho\delta\mathbb{A}(V_{jc}^{U} - V_{1ic}^{U}) + \lambda_c^* \quad \forall j = \{3, i\}
\end{aligned}$$

Substituting in equation (31), this becomes:

$$\begin{aligned}
V_{fjc}^{E} - (V_{f1ic}^{E} - \lambda_c^*) &= (w_{fjc} - w_{f1ic}) + \rho(1-\delta)\mathbb{A}(w'_{fjc} - w'_{f1ic}) + (\lambda_c^* - \mathbb{A}\psi_{f1ic}) + \Delta_{jc,1ic} \\
&\forall \quad j = \{3, i\},
\end{aligned} \tag{33}$$

where $\Delta_{jc,1ic} = \rho^2 \delta q_c^u \mathbb{A}\mathbb{A}_c \mathbb{B}_c \sum_k \left(T_{kc|j} - T_{kc|1i}\right) w'_{kc}$.

C.4 Bargaining

C.4.1 Collective Bargaining: Union Firms

Collective bargaining satisfies the following bargaining rule:

$$\beta S_{fjc} = (1-\beta)n\left(V^E_{fjc} - V^U_{jc}\right) \quad \forall j = \{1, i\} \tag{34}$$

Substituting equations (13) and (32) into the surplus splitting rule, and solving for the wage, one obtains:

$$\begin{aligned} w_{fjc} &= \beta\left[p_i\frac{y_{fjc}(n)}{n} + \rho^e(1-\delta^m)\frac{\kappa}{q^v_c}\right] - (1-\beta)\mathbb{A}\left[\psi_{fjc} + \rho^e(1-\delta^m)w'_{fjc} - (1-\rho)V^U_{jc}\right] \\ \forall &\quad j = \{1, i\} \end{aligned}$$

In steady state, $w_{fjc} = w'_{fjc}$. Therefore, the wage equation can be rewritten as:

$$\begin{aligned} w_{fjc} &= \beta\mathbb{A}^{-1}\mathbb{B}\left[p_i\frac{y_{fjc}(n)}{n} + \rho^e(1-\delta^m)\frac{\kappa}{q^v_c}\right] - (1-\beta)\mathbb{B}\left[\psi_{fjc} - (1-\rho)V^U_{jc}\right] \\ \forall &\quad j = \{1, i\} \end{aligned} \tag{35}$$

where $\mathbb{B} = \frac{1}{[1-\beta\rho^e(1-\delta^m)]} > 1$ and where, in steady state, $V^U_{jc} = \mathbb{D}_c + \rho q^u_c \mathbb{A}_c\mathbb{B}_c \sum_k T_{kc|j}w_{kc}$.

Finally, substituting in optimal firm size (as given by equation (43) below in Section C.5.1), we get:

$$w_{fjc} = \tilde{\xi}_0\mathbb{D}_c + \tilde{\xi}_{1c}\sum_k T_{kc|j}w_{kc} + \tilde{\xi}_2\frac{\kappa}{q^v_c} + \tilde{\xi}_3 p_i\epsilon_{fic} - \tilde{\xi}_4\psi_{fjc} \quad \forall j = \{1, i\} \tag{36}$$

where $\tilde{\xi}_0 = (1-\rho)\mathbb{B}\left[\frac{1}{2}\beta\mathbb{A}^{-1} + (1-\beta)\right] > 0$, $\tilde{\xi}_{1c} = \tilde{\xi}_0\rho q^u_c\mathbb{A}_c\mathbb{B}_c > 0$, $\tilde{\xi}_2 = \beta\mathbb{A}^{-1}\mathbb{B}\left[\frac{\mathbb{A}^{-1}}{2(1-\beta)} + \rho^e(1-\delta^m)\right] > 0$, $\tilde{\xi}_3 = \frac{1}{2}\beta\mathbb{A}^{-1}\mathbb{B} > 0$, and $\tilde{\xi}_4 = \mathbb{B}\left[\frac{1}{2}\beta\mathbb{A}^{-1} + (1-\beta)\right] > 0$.

Hence, the wage of a worker currently employed in a union firm depends on the tightness of the labour market (via the city-specific ξs, $\mathbb{D}_c$ and q^v_c), the union-(industry-city)-specific outside option, as captured by $\sum_k T_{kc|j}w_{kc}$, the recruitment costs (through $\frac{\kappa}{q^v_c}$), the productivity term $p_i\epsilon_{fic}$, and the amenity created by the union.

C.4.2 Individual Bargaining: Simple Non-union Firms

Individual bargaining satisfies the following bargaining rule:

$$\beta S_{fjc} = (1-\beta)\left(V^E_{fjc} - V^U_{jc}\right) \quad \forall j = \{2, i\} \tag{37}$$

Substituting equations (18) and (32) into the surplus splitting rule yields the following differential equation:

$$w_{fjc} = \beta p_i \frac{\partial y_{fjc}(n)}{\partial n} - \beta \frac{\partial w_{fjc}}{\partial n} n + \beta \rho^e (1-\delta^m) \frac{\kappa}{q_c^v} - (1-\beta) \left[\rho^e (1-\delta^m) \mathbb{A} w'_{fjc} - (1-\rho) \mathbb{A} V_{jc}^U \right] \quad \forall \quad j = \{2, i\}$$

The solution to this differential equation is:

$$w_{fjc} = \beta \left[p_i \left(\epsilon_{fic} - \frac{1}{1+\beta} \sigma_i n \right) + \rho^e (1-\delta^m) \frac{\kappa}{q_c^v} \right] - (1-\beta) \left[\rho^e (1-\delta^m) \mathbb{A} w'_{fjc} - (1-\rho) \mathbb{A} V_{jc}^U \right] \quad \forall \quad j = \{2, i\}$$

Hence, in steady state where $w_{fjc} = w'_{fjc}$, the wage equation for simple non-union firms is given by:

$$w_{fjc} = \beta \mathbb{A}^{-1} \mathbb{B} \left[p_i \left(\epsilon_{fic} - \frac{1}{1+\beta} \sigma_i n \right) + \rho^e (1-\delta^m) \frac{\kappa}{q_c^v} \right] + (1-\beta)(1-\rho) \mathbb{B} V_{jc}^U \quad \forall \quad j = \{2, i\} \tag{38}$$

where, in steady state, $V_{jc}^U = \mathbb{D}_c + \rho q_c^u \mathbb{A}_c \mathbb{B}_c \sum_k T_{kc|j} w_{kc}$.

Finally, substituting in optimal firm size (as given by equation (45) below in Section C.5.2), we get:

$$w_{fjc} = \xi_0 \mathbb{D}_c + \xi_{1c} \sum_k T_{kc|j} w_{kc} + \xi_2 \frac{\kappa}{q_c^v} + \xi_3 p_i \epsilon_{fic} \quad \forall j = \{2, i\} \tag{39}$$

where, defining $\mathbb{C} = \frac{1}{1+\beta\rho^e(1-\delta^m)} > 0$, we have $\xi_0 = (1-\rho)\mathbb{B}\left[\beta\mathbb{A}^{-1}\mathbb{C} + (1-\beta)\right] > 0$, $\xi_{1c} = \xi_0 \rho q_c^u \mathbb{A}_c \mathbb{B}_c > 0$, $\xi_2 = \beta \mathbb{A}^{-1} \mathbb{B} \left[\frac{1}{(1-\beta)} \mathbb{A}^{-1} \mathbb{C} + \rho^e (1-\delta^m) \right] > 0$, and $\xi_3 = \beta^2 \rho^e (1-\delta^m) \mathbb{A}^{-1} \mathbb{B} \mathbb{C} > 0$.

The structure of the wage equation for workers in non-union firms is similar to that of unionised workers, except that the latter has different parameters, a union-specific outside option, and a term capturing amenities created in the union firm.

C.4.3 Collective Bargaining: Emulating Non-union Firms

In non-union emulating firms, the wage solves the following bargaining game:

$$\beta S_{fjc} = (1-\beta) n \left[V_{fjc}^E - (V_{f1ic}^E - \lambda_c^*) \right] \quad \forall j = \{3, i\} \tag{40}$$

Substituting equations (22) and (33) into the bargaining rule, we get:

$$w_{fjc} - w_{f1ic} = \frac{(1-\beta)(1-\rho^e)}{\beta + (1-\beta)(1-\rho^e)} \left[(\mathbb{A}\psi_{f1ic} - \lambda_c^*) - \rho^e (1-\delta^m) \mathbb{A} (w'_{fjc} - w'_{f1ic}) - \Delta_{jc,1ic} \right] \quad \forall \quad j = \{3, i\}$$

In steady state where $(w_{fjc} - w_{f1ic}) = (w'_{fjc} - w'_{f1ic})$, the wage in non-union emulating firms is then:

$$w_{fjc} = w_{f1ic} + \bar{\xi}\left[(\mathbb{A}\psi_{f1ic} - \lambda_c^*) - \Delta_{jc,1ic}\right] \quad \forall j = \{3, i\} \tag{41}$$

where $\bar{\xi} = \frac{(1-\beta)(1-\rho^e)}{\beta + (1-\beta)(1-\rho^e)\mathbb{A}} > 0$, and, in steady state, $\Delta_{jc,1ic} = \rho^2 \delta q_c^u \mathbb{A}\mathbb{A}_c\mathbb{B}_c \sum_k \left(T_{kc|j} - T_{kc|1i}\right) w_{kc}$.

Hence, the wage differential between an emulating non-union firm and a union firm captures both the difference between the amenity offered by the union and the unionisation cost, and the difference between outside options in union vs emulating non-union firms.

C.5 Optimal Firm Size

Optimal firm size satisfies the condition given by equation (9), which states that marginal profits have to equate the marginal cost of creating a new vacancy, adjusted for the proportion $(1-\delta)$ of the workforce that remains with the firm from one period to another one. In what follows, we derive the optimal size from the profit function, noting that V_{jc}^U is taken as given by the firm.

C.5.1 Union Firms

Using equation (35), the profits of a union firm are given by:

$$\begin{aligned}\pi_{fjc}(n) \;\;&=\;\; \mathbb{B}\left[(1-\beta)p_i y_{fjc}(n) - \frac{\beta\rho^e(1-\delta^m)}{\mathbb{A}}\frac{\kappa}{q_c^v}n + (1-\beta)\psi_{fjc}n - (1-\beta)(1-\rho)V_{jc}^U n\right] \\ &\forall \quad j = \{1, i\}\end{aligned} \tag{42}$$

Taking the derivative of equation (42) with respect to employment, using equation (9) and solving for n, optimal firm size is given by:

$$n_{fjc}^* = \frac{1}{p_i\sigma_i}\left[p_i\epsilon_{fic} - \frac{1}{(1-\beta)\mathbb{A}}\frac{\kappa}{q_c^v} + \psi_{fjc} - (1-\rho)V_{jc}^U\right] \quad \forall j = \{1, i\}, \tag{43}$$

where the firm-job-city subscript has been added back to firm size.

C.5.2 Simple Non-union Firms

Using equation (38), the profits of a simple non-union firm are given by:

$$\begin{aligned}\pi_{fjc}(n) \;\;&=\;\; \mathbb{B}\left\{(1-\beta)p_i\epsilon_{fic}n - \frac{(1-\beta)}{2(1+\beta)}\left[1+\beta\rho^e(1-\delta^m)\right]p_i\sigma_i n^2 - \frac{\beta\rho^e(1-\delta^m)}{\mathbb{A}}\frac{\kappa}{q_c^v}n\right\} \\ &-\;\; \mathbb{B}(1-\beta)(1-\rho)V_{jc}^U n \quad \forall j = \{2, i\}\end{aligned} \tag{44}$$

Using equation (44) to derive marginal profits, combining with (9) and solving for n, we obtain:

$$n_{fjc}^* = \frac{1}{p_i\sigma_i}\frac{(1+\beta)}{\left[1+\beta\rho^e(1-\delta^m)\right]}\left[p_i\epsilon_{fic} - \frac{1}{(1-\beta)\mathbb{A}}\frac{\kappa}{q_c^v} - (1-\rho)V_{jc}^U\right] \quad \forall j = \{2, i\} \tag{45}$$

C.5.3 Emulating Non-union Firms

From equation (41), we can see that the wage in emulating non-union firms is independent of firm size. Therefore, the optimality condition implies the following optimal firm size:

$$n^*_{fjc} = \frac{1}{p_i \sigma_i} \left[p_i \epsilon_{fic} - \frac{1}{\mathbb{A}} \frac{\kappa}{q_c^v} - w_{fjc} \right] \quad \forall j = \{3, i\}, \tag{46}$$

where w_{fjc} is given by equation (41).

D Linear Approximation

We next move from our derived wage expressions, equations (36) and (39), to our estimating specifications using a linear approximation. To do this, we expand with respect to the vector

$$\mathbf{x} = \left[\psi, b, p_i, \sum_k T_{kc|j} w_{kc}, ER_c, \epsilon_{fic}, \psi_{fjc} \right] \quad \forall j = \{1, i\} \text{ or } j = \{2, i\}$$

where, recall, ER_c denotes the employment rate in city c. We take the linear approximation around a point $\mathbf{x_0}$ where the employment is equally spread across industries (which occurs when, inter alia, the nation-wise mobility frictions are constant, i.e. when $\varphi_{k|j} = \varphi$) and the employment rate takes the same value in all cities. In particular, we expand around $\mathbf{x_0} = [0, 0, p, w, ER, \epsilon, \psi]$. For equation (41) we use the linear approximation of the union wage equation and further expand with respect to the vector $\left[\{T_{kc|1i}\}_k, \{T_{kc|2i}\}_k, \{w_{kc}\}_k, \psi_{f1ic}, \lambda^*_c \right]$, which is $[\{1/I\}_k, \{1/I\}_k, \{w\}_k, \psi, \lambda]$ around $\mathbf{x_0}$.

D.1 Union Wage Equation

The result for the union wage equation is:

$$w_{fjc} = \tilde{\gamma}_{0i} + \tilde{\gamma}_1 \sum_k T_{kc|j} w_{kc} + \tilde{\gamma}_2 ER_c + \tilde{\gamma}_3 \epsilon_{ic} + \tilde{\gamma}_3 u_{fic} - \tilde{\gamma}_4 \psi_{fjc} \quad \forall j = \{1, i\} \tag{47}$$

where $\tilde{\gamma}_{0i}$ is a function of the price p_i and constant terms stemming from the expansion point values. $\tilde{\gamma}_1 = \tilde{\xi}_{1c}\big|_{\mathbf{x_0}} > 0$, and is, at $\mathbf{x_0}$, independent of the city subscript. $\tilde{\gamma}_2$ is a complicated positive function of the underlying parameters evaluated at the common ER value, and $\tilde{\gamma}_3 = \tilde{\xi}_3 p > 0$, and $\tilde{\gamma}_4 = \tilde{\xi}_4 > 0$.[5] Also note that we have been using the fact that the productivity term can be decomposed as $\epsilon_{fic} = \epsilon_{ic} + u_{fic}$.

[5] In particular, $\tilde{\gamma}_2 = \tilde{\xi}_0 \frac{\partial \mathbb{D}_c}{\partial ER_c}\big|_{\mathbf{x_0}} + w \frac{\partial \tilde{\xi}_{1c}}{\partial ER_c}\big|_{\mathbf{x_0}} + \tilde{\xi}_2 \frac{\partial}{\partial ER_c} \left(\frac{k}{q_c^v} \right)\big|_{\mathbf{x_0}}$, where $\frac{\partial \mathbb{D}_c}{\partial ER_c}\big|_{\mathbf{x_0}} = 0$ (since $b = 0$ and $\psi = 0$ around $\mathbf{x_0}$), $\frac{\partial \tilde{\xi}_{1c}}{\partial ER_c}\big|_{\mathbf{x_0}} = \tilde{\xi}_0 \rho (1 - \rho) \left[\mathbb{A}_c^2 \mathbb{B}_c \left(1 + \rho^2 \delta q_c^u \mathbb{A}_c \mathbb{B}_c \right) \frac{\partial q_c^u}{\partial ER_c} \right]\big|_{\mathbf{x_0}} > 0$ (since $\frac{\partial q_c^u}{\partial ER_c} > 0$), and $\frac{\partial}{\partial ER_c} \left(\frac{k}{q_c^v} \right)\big|_{\mathbf{x_0}} > 0$ (since $\frac{\partial q_c^v}{\partial ER_c} < 0$).

D.2 Simple Non-union Wage Equation

For the simple non-union firms, the wage equation is given by:

$$w_{fjc} = \gamma_{0i} + \gamma_1 \sum_k T_{kc|j} w_{kc} + \gamma_2 ER_c + \gamma_3 \epsilon_{ic} + \gamma_3 u_{fic} \quad \forall j = \{2, i\} \tag{48}$$

As with the union equation, γ_{0i} is a function of the price p_i and constant terms stemming from the expansion point values. $\gamma_1 = \xi_{1c}\big|_{\mathbf{x_0}} > 0$, and is, at $\mathbf{x_0}$, independent of the city subscript. γ_2 is, again, a complicated positive function of the underlying parameters evaluated at the common ER value, and $\gamma_3 = \xi_3 p > 0$.[6]

D.3 Emulating Non-union Wage Equation

Finally, for the emulating non-union wage equation, we have:

$$\begin{aligned} w_{fjc} &= w_{f1ic} + \bar{\xi}\left(\mathbb{A}\psi_{f1ic} - \lambda_c^*\right) \quad \forall j = \{3, i\} \\ &= \tilde{\gamma}_{0i} + \tilde{\gamma}_1 \sum_k T_{kc|1i} w_{kc} + \tilde{\gamma}_2 ER_c + \tilde{\gamma}_3 \epsilon_{ic} + \tilde{\gamma}_3 u_{fic} - (\tilde{\gamma}_4 - \bar{\xi}\mathbb{A})\psi_{f1ic} - \bar{\xi}\lambda_c^* \end{aligned} \tag{49}$$

where $\tilde{\gamma}_4 - \bar{\xi}\mathbb{A} > 0$.[7]

Note that the equation includes the outside option if the worker in this firm were union rather than non-union. It has a positive sign both because better outside options for union workers mean the union wage is higher and, so, the emulating firm has to pay higher wages, and because it will imply better search possibilities when non-employed if union workers have better access to other union jobs.

E Union Arrangement Determination and the Mean Non-union Wage Equation

E.1 Union Determination Without Firm Responses

We begin by discussing union determination in the scenario where workers choose to unionise, and there is no response from the firm to the threat of unionisation. In this context, the probability of meeting an emulating firm is zero, hence there are only two possible union status: $j = \{1, i\}$ and $j = \{2, i\}$. A firm becomes unionised if

$$V^E_{f1ic} - \lambda^*_c > V^E_{f2ic}, \tag{50}$$

or, defining the index function $I_{fic} = V^E_{f1ic} - \lambda^*_c - V^E_{f2ic}$, if $I_{fic} > 0$. Using equation (25) and the steady state condition that $(w_{f1ic} - w_{f2ic}) = (w'_{f1ic} - w'_{f2ic})$, we can express I_{fic} as:

[6] Note that $\gamma_2 = \xi_0 \frac{\partial \mathbb{D}_c}{\partial ER_c}\Big|_{\mathbf{x_0}} + w \frac{\partial \xi_{1c}}{\partial ER_c}\Big|_{\mathbf{x_0}} + \xi_2 \frac{\partial}{\partial ER_c}\left(\frac{k}{q^v_c}\right)\Big|_{\mathbf{x_0}}$, where $\frac{\partial \xi_{1c}}{\partial ER_c}\Big|_{\mathbf{x_0}} = \frac{\xi_0}{\tilde{\xi}_0} \frac{\partial \tilde{\xi}_{1c}}{\partial ER_c}\Big|_{\mathbf{x_0}} > 0$

[7] Specifically, $\tilde{\gamma}_4 - \bar{\xi}\mathbb{A} = \frac{\beta\mathbb{A}^{-1} + (1-\beta)(1+\rho^e)}{\beta + (1-\beta)(1-\rho^e)\mathbb{A}} \frac{1}{2}\beta\mathbb{B} > 0.$

$$I_{fic} = \mathbb{A}(w_{f1ic} - w_{f2ic}) + (\mathbb{A}\psi_{f1ic} - \lambda_c^*) + \rho\delta\mathbb{A}(V_{f1ic}^U - V_{f2ic}^U), \quad (51)$$

where, using equation (31), $(V_{f1ic}^U - V_{f2ic}^U) = \rho q_c^u \mathbb{A}_c \mathbb{B}_c \sum_k (T_{kc|1i} - T_{kc|2i}) w_{kc}$, in steady state. Hence, whether firms unionise positively depends on the wage differential between union and non-union firms, the difference between the outside options of their respective workers, and the difference between union amenities and the cost to unionise. The larger these differences, the more likely a firm will unionise.

Substituting the linearized wage equations (47) and (48), and further expanding with respect to the vector $\left[\{T_{kc|1i}\}_k, \{T_{kc|2i}\}_k, \{w_{kc}\}_k, \psi_{f1ic}, \lambda_c^*\right]$ (as we do in D), we obtain the following expression:

$$I_{fic} = \alpha_{0i} + \alpha_1 \sum_k T_{kc|1i} w_{kc} - \alpha_2 \sum_k T_{kc|2i} w_{kc} + \alpha_3 ER_c + \alpha_4 \epsilon_{ic} + \alpha_4 u_{fic} + \alpha_5 \psi_{f1ic} - \lambda_c^*, \quad (52)$$

where the αs are constant terms obtained from the linear approximation, In particular, $\alpha_1 > 0$, $\alpha_2 > 0$, $\alpha_3 < 0$, $\alpha_4 > 0$ and $\alpha_5 > 0$.[8]

Using equation (48), this standard selection set-up implies that the conditional mean of the non-union wage for workers at firms that are actually observed to be non-union is given by:

$$E\left(w_{f2ic} \,\middle|\, I_{fic} \leq 0\right) = \gamma_{0i} + \gamma_1 \sum_k T_{kc|jc} w_{kc} + \gamma_2 ER_c + \gamma_3 \epsilon_{ic} + \gamma_3 E\left(u_{fic} \,\middle|\, I_{fic} \leq 0\right) \quad (53)$$

E.2 Incorporating Firm Responses

We now turn to the scenario where firms respond to the threat of unionisation. Of course, firms do not need to consider employing any response if their workers are content to remain non-union. That is, if the costs of unionising, the amenities available to the workers at this firm if they unionise, the union wage they would receive if they organise, and the non-union wage they receive, if they don't, are such that $I_{fic} \leq 0$, then there is no reason for the firm to bear costs to incentivize its workers not to form a union. However, if $I_{fic} > 0$, workers will want to unionise, and firms will consider whether to respond or to let unionisation happen.

The possibility of a firm response implies that there are two amenity thresholds determining a firm's union status. Specifically, there is a first amenity threshold, ψ_{fic}^*, below which workers are content to remain non-union and, hence, below which a firm remains non-union. At ψ_{fic}^*, $I_{fic} = 0$, and workers are indifferent between being union or non-union.

There is also a second amenity threshold, ψ_{fic}^{**}, above which the firm will be unionised and at which the firm is indifferent between responding (and thus becoming an emulating non-union firm) and being unionised. In our derivations of worker and firm surpluses in the non-union emulation state, the outside option is being unionised. Thus, a firm-worker arrangement will be unionised if the total surplus in the emulating non-union state (the

[8] Specifically, we have that $\alpha_{0i} = \mathbb{A}(\tilde{\gamma}_{0i} - \gamma_{0i})$, $\alpha_1 = \mathbb{A}\tilde{\gamma}_1 > 0$, $\alpha_2 = \mathbb{A}\gamma_1 > 0$, $\alpha_3 = \mathbb{A}(\tilde{\gamma}_2 - \gamma_2) < 0$, $\alpha_4 = \mathbb{A}(\tilde{\gamma}_3 - \gamma_3) > 0$, and $\alpha_5 = \mathbb{A}(1 - \tilde{\gamma}_4) > 0$.

worker plus the firm surplus) is negative. At ψ^{**}_{fic}, the total surplus is zero, while it is strictly positive for amenities that fall between the two thresholds, making it optimal for the firm to respond to the threat of unionisation.

In what follows, we derive these two amenity thresholds that determine union status, as well as the mean non-union wage equation in this context.

E.2.1 Union Statuses Thresholds

Amenity Threshold $\boldsymbol{\psi^*_{fic}}$. As discussed above, workers do not initiate unionisation if $I_{fic} \leq 0$, i.e., if the surplus the median worker would receive if the firm remained non-union exceeds that if it became unionised, net of unionisation costs. At the threshold ψ^*_{fic}, $I_{fic} = 0$, and workers are indifferent between being union and non-union.

Using equation (51) and linearizing as we do above, we obtain:

$$\psi^*_{fic} = \lambda_c - \left[w_{f1ic}(\psi^*_{fic}) - w_{f2ic}\right], \tag{54}$$

where $w_{f1ic}(\psi^*_{fic})$ corresponds to equation (47) evaluated at ψ^*_{fic}, w_{f2ic} is given by (48), and $\lambda_c = \frac{1}{\mathbb{A}}\lambda^*_c$.

Substituting in the wage equations and solving for ψ^*_{fic}, we obtain:

$$\psi^*_{fic} = \frac{1}{1-\tilde{\gamma}_4}\lambda_c - \Delta^*_{1ic,2ic} - \frac{\alpha_4}{\alpha_5}\epsilon_{fic}, \tag{55}$$

where $\Delta^*_{1ic,2ic} = \frac{\alpha_{0i}}{\alpha_5} + \frac{\alpha_1}{\alpha_5}\sum_k T_{kc|1i}w_{kc} - \frac{\alpha_2}{\alpha_5}\sum_k T_{kc|2i}w_{kc} + \frac{\alpha_3}{\alpha_5}ER_c$, and where $0 < \frac{1}{1-\tilde{\gamma}_4} < 1$, $\frac{\alpha_4}{\alpha_5} > 0$, $\frac{\alpha_{0i}}{\alpha_5} > 0$, $\frac{\alpha_1}{\alpha_5} > 0$, $\frac{\alpha_2}{\alpha_5} > 0$, $\frac{\alpha_3}{\alpha_5} < 0$. Notice that the threshold is specific to each firm-industry-city combination; i.e., within an industry-city, firms face different thresholds depending on their productivity ϵ_{fic}.

Amenity Threshold $\boldsymbol{\psi^{**}_{fic}}$. Firms remain non-union $\forall \psi_{f1ic} \leq \psi^*_{fic}$. However, for values of $\psi_{f1ic} > \psi^*_{fic}$, workers will want to unionise, and firms will consider whether to respond to forestall unionisation or to let workers unionise. A firm response is optimal as long as the total surplus obtained in the emulating non-union state is positive. At the threshold ψ^{**}_{fic}, the total surplus is zero and becomes negative above this point.

Combining equations (22) and (33), along with the steady state condition that $w_{fjc} = w'_{fjc}$, the total surplus in the non-union emulation state is:

$$TS_{f3ic} = \left\{-\frac{\rho^e\delta^m}{1-\rho^e}\mathbb{A}(w_{f3ic} - w_{f1ic}) - \mathbb{A}\psi_{f1ic} + \lambda^*_c + \Delta_{3ic,1ic}\right\} n \tag{56}$$

Substituting equation (41) into (56), and linearizing (as done earlier on), we obtain:

$$\begin{aligned} TS_{f3ic} &= \left\{-\frac{\rho^e\delta^m}{1-\rho^e}\mathbb{A}\overline{\xi}\left[\mathbb{A}\psi_{f1ic} - \lambda^*_c - \Delta_{3ic,1ic}\right] - \mathbb{A}\psi_{f1ic} + \lambda^*_c + \Delta_{3ic,1ic}\right\} n \\ &= \left[1 + \frac{\rho^e\delta^m}{1-\rho^e}\mathbb{A}\overline{\xi}\right]\left(\lambda^*_c - \mathbb{A}\psi_{f1ic}\right) n, \end{aligned} \tag{57}$$

where $\Delta_{3ic,1ic} = 0$ due to the linearization.

Since $\left[1+\frac{\rho^e\delta m}{1-\rho^e}\mathbb{A}\bar{\xi}\right]>0$ and $n>0$, equation (57) implies that the total surplus in the non-union emulation state is zero if $\lambda_c^*-\mathbb{A}\psi_{f1ic}=0$, i.e., given equation (49), if the wage difference between the emulation and union statuses is zero.

Therefore,

$$\psi_{fic}^{**}=\lambda_c, \tag{58}$$

and the threshold is city-specific only.

Productivity Threshold ϵ_{ic}^*. Another determinant of the response is the productivity term ϵ_{fic}. As we will see below, for small values of the productivity term that fall below a threshold ϵ_{ic}^*, there is no set of amenity values where firms would be emulators. This implies that emulating type non-union firms are only found in industry-city pairs characterized by $\epsilon_{fic}>\epsilon_{ic}^*$.

From equations (54) and (58) it is clear that $\psi_{fic}^*<\psi_{fic}^{**}$ as long as $w_{f1ic}(\psi_{fic}^*)>w_{f2ic}$, which occurs when the productivity term ϵ_{fic} is sufficiently large. If $w_{f1ic}(\psi_{fic}^*)\leq w_{f2ic}$, then $\psi_{fic}^{**}\leq\psi_{fic}^*$, and there is no set of ψ_{f1ic} values where firms would be emulators. Therefore, the lower bound productivity threshold ϵ_{ic}^* solves $w_{f1ic}(\psi_{fic}^*,\epsilon_{ic}^*)=w_{f2ic}(\epsilon_{ic}^*)$ (or, equivalently, $\psi_{fic}^*(\epsilon_{ic}^*)=\psi_{fic}^{**}$). Using equations (47) and (48) together with (55), the productivity threshold that solves $w_{f1ic}(\psi_{fic}^*,\epsilon_{ic}^*)=w_{f2ic}(\epsilon_{ic}^*)$ is given by:

$$\epsilon_{ic}^*=\underline{\gamma}_{0i}+\underline{\gamma}_1\sum_k T_{kc|2i}w_{kc}-\underline{\gamma}_2\sum_k T_{kc|1i}w_{kc}+\underline{\gamma}_3 ER_c+\underline{\gamma}_4\lambda_c,$$

where the $\underline{\gamma}$s are positive terms obtained from the linearization.[9]

To summarize, for values of $\epsilon_{fic}\leq\epsilon_{fic}^*$, firms will be non-union $\forall\ \psi_{f1ic}\leq\psi_{fic}^*$ and unionised if $\psi_{f1ic}>\psi_{fic}^*$. The emulation state is only observed for productivity levels above ϵ_{ic}^*: If $\epsilon_{fic}>\epsilon_{ic}^*$, firms will be non-union if $\psi_{f1ic}\leq\psi_{fic}^*$, emulators if $\psi_{fic}^*<\psi_{f1ic}\leq\psi_{fic}^{**}$ and unionised for amenity values above ψ_{fic}^{**}. The various statuses described above are summarized in Table 1.

E.2.2 Observed Mean Non-union Wage

The expected, observed non-union wage, conditional on being non-union of either type, w_{ic}^n, is given by:

$$w_{ic}^n=\frac{1}{P_{ic}^n}\int_0^\infty\int_0^{\psi_{fic}^*}w_{f2ic}f(\psi)g(\epsilon)d\psi d\epsilon+\frac{1}{P_{ic}^n}\int_{\epsilon_{ic}^*}^\infty\int_{\psi_{fic}^*}^{\psi_{fic}^{**}}w_{f3ic}f(\psi)g(\epsilon)d\psi d\epsilon, \tag{59}$$

where, for notational clarity, we have omitted the subscripts on amenities (ψ_{f1ic}) and productivity (ϵ_{fic}) in the integrals, and where P_{ic}^n, the probability of being observed non-union of either type, is:

$$P_{ic}^n=\int_0^\infty\int_0^{\psi_{fic}^*}f(\psi)g(\epsilon)d\psi d\epsilon+\int_{\epsilon_{ic}^*}^\infty\int_{\psi_{fic}^*}^{\psi_{fic}^{**}}f(\psi)g(\epsilon)d\psi d\epsilon, \tag{60}$$

[9] In particular, $\underline{\gamma}_{0i}=\frac{\gamma_{0i}-\tilde{\gamma}_{0i}}{\tilde{\gamma}_3-\gamma_3}>0$, $\underline{\gamma}_1=\frac{\gamma_1}{\tilde{\gamma}_3-\gamma_3}>0$, $\underline{\gamma}_2=\frac{\tilde{\gamma}_1}{\tilde{\gamma}_3-\gamma_3}>0$, $\underline{\gamma}_3=\frac{\gamma_2-\tilde{\gamma}_2}{\tilde{\gamma}_3-\gamma_3}>0$, and $\underline{\gamma}_4=\frac{\tilde{\gamma}_4}{\tilde{\gamma}_3-\gamma_3}>0$.

Table 1: Summary of Union Statuses

		$\epsilon_{fic} \leq \epsilon^*_{ic}$	$\epsilon_{fic} > \epsilon^*_{ic}$
$\psi^{**}_{fic} \leq \psi^*_{fic}$	$\psi_{f1ic} \leq \psi^*_{fic}$	Simple non-union	
	$\psi_{f1ic} > \psi^*_{fic}$	Union	
$\psi^*_{fic} < \psi^{**}_{fic}$	$\psi_{f1ic} \leq \psi^*_{fic}$		Simple non-union
	$\psi^*_{fic} < \psi_{f1ic} \leq \psi^{**}_{fic}$		Emulating non-union
	$\psi_{f1ic} > \psi^{**}_{fic}$		Union

and where the thresholds are derived in the previous section. Substituting in equations (48) and (49), we have:

$$\begin{aligned} w^n_{ic} &= \gamma_{0i} + \tilde{\gamma}_1 P^{ne}_{ic} \sum_k T_{kc|1i} w_{kc} + \gamma_1 (1 - P^{ne}_{ic}) \sum_k T_{kc|2i} w_{kc} + \gamma_2 ER_c + \gamma^*_{0i} P^{ne}_{ic} \\ &- \gamma_6 P^{ne}_{ic} \lambda_c + \mu_{ic} + \tilde{\gamma}_3 P^{ne}_{ic} \epsilon_{ic} + \gamma_3 (1 - P^{ne}_{ic}) \epsilon_{ic}, \end{aligned} \tag{61}$$

where $\gamma^*_{0i} = (\tilde{\gamma}_{0i} - \gamma_{0i})$, $\gamma_6 = \bar{\xi}\mathbb{A} > 0$ and where we have used the fact that at reasonable structural parameter values, $\tilde{\gamma}_2 \approx \gamma_2$.[10] P^{ne}_{ic} is the probability of being an emulating firm conditional on being a non-union firm; i.e.:

$$P^{ne}_{ic} = \frac{1}{P^n_{ic}} \int_{\epsilon^*_{ic}}^{\infty} \int_{\psi^*_{fic}}^{\psi^{**}_{fic}} f(\psi) g(\epsilon) d\psi d\epsilon. \tag{62}$$

Finally, μ_{ic} is an error mean term capturing the selection of firms into simple non-union or emulating non-union status. Specifically:

$$\mu_{ic} = \gamma_3 E(u_{fic}|\tau = 2) + \tilde{\gamma}_3 E(u_{fic}|\tau = 3) - (\tilde{\gamma}_4 - \bar{\xi}\mathbb{A}) E(\psi_{f1ic}|\tau = 3).$$

where

$$E(u_{fic}|\tau = 2) = \frac{1}{P^n_{ic}} \int_0^{\infty} \int_0^{\psi^*_{fic}} u_{fic}(\epsilon) f(\psi) g(\epsilon) d\psi d\epsilon,$$

$$E(u_{fic}|\tau = 3) = \frac{1}{P^n_{ic}} \int_{\epsilon^*_{ic}}^{\infty} \int_{\psi^*_{fic}}^{\psi^{**}_{fic}} u_{fic}(\epsilon) f(\psi) g(\epsilon) d\psi d\epsilon,$$

[10] Using the definitions of $\tilde{\gamma}_2$ and γ_2 in footnotes 5 and 6, note that the $\tilde{\xi}_0$ and ξ_0 play a key role. Referring to their definitions in Section C.4, they differ because of a multiplicative factor $\mathbb{C} = \frac{1}{1+\beta\rho^e(1-\delta^m)}$. Using reasonable values for β, ρ^e, and δ^m, $\mathbb{C}$ takes a value near 1 (e.g. with $\beta = 0.5$, $\rho^e = 0.03$, and $\delta^m = 0.05$, then $\mathbb{C} = 0.986$). Therefore, $\tilde{\gamma}_2 \approx \gamma_2$ and $[P^{ne}_{ic}\tilde{\gamma}_2 + (1 - P^{ne}_{ic})\gamma_2] ER_c = \gamma_2 ER_c$.

and

$$E(\psi_{f1ic}|\tau=3)=\frac{1}{P^n_{ic}}\int_{\epsilon^*_{ic}}^{\infty}\int_{\psi^*_{fic}}^{\psi^{**}_{fic}}\psi f(\psi)g(\epsilon)d\psi d\epsilon.$$

E.3 Intimidation response

Instead of raising wages, firms might react to the threat of unionisation through intimidation tactics, i.e., by increasing the fixed cost of unionising. In this section, we explore that possibility, assuming that firms can increase λ^*_c at a cost to themselves. For example, they could lock out workers, ceasing production or hiring less productive replacement workers, known as scabs. The firm may also take legal actions to delay the union vote, imposing more costs on the workers.

Assuming it costs the firm 1 dollar to raise unionisation costs by 1 dollar, to thwart unionisation, the firm would need to increase the per-worker cost of unionisation by cu_{fic}, an amount that renders a worker indifferent between union and non-union status. We assume that a firm employing intimidation tactics will adjust employment levels, considering the impact of its decision on total intimidation costs, but that it continues to pay the simple non-union wage. Denoting the intimidation status as $\tau=4$, in this scenario, the firm solves the following optimization problem:[11]

$$\begin{aligned}\Pi_{fjc}(n_{-1})=\max_{v}\quad & [p_i y_{fjc}(n)-[w_{f2ic}+cu_{fic}]\,n-\kappa v+\rho^e\Pi_{fjc}(n)]\\ \text{s.t.}\quad & n=n_{-1}(1-\delta^m)+q^v_c v\\ & cu_{fic}=(V^E_{f1ic}-\lambda^*_c)-V^E_{f2ic}\qquad\qquad \forall j=\{4,i\}\end{aligned}$$

where, for clarity, we have dropped the subscript fjc on employment and vacancies. Solving this problem, optimal firm size is given by:

$$n^*_{f4ic}=\frac{1}{p_i\sigma_i}\Big\{p_i\epsilon_{fic}-(w_{f2ic}+cu_{fic})-[1-\rho^e(1-\delta^m)]\,\frac{\kappa}{q^v_c}\Big\},\tag{63}$$

where $cu_{fic}=\mathbb{A}\left[(w_{f1ic}-w_{f2ic})+(\psi_{f1ic}-\lambda_c)+\rho\delta(V^U_{f1ic}-V^U_{f2ic})\right]$.

To determine the type of firm's response along the range of possible amenity values, it is useful to examine, for each union status, (i) how the firm's optimal value function changes with ψ_{f1ic} and (ii) the optimal value function under the intimidating status compared to the emulating status at the thresholds ψ^*_{fic} and ψ^{**}_{fic}. To do so, note that, given the first-order condition (equation (9)), the optimal value function for each union status is given by $\Pi_{fjc}(n^*_{fjc})=(1-\delta^m)\frac{\kappa}{q^v_c}n^*_{fjc}\ \forall\ j$, where n^*_{fjc} denotes optimal employment.

[11] If the firm were to adjust both employment and the wage paid, then it can be shown that $w_{f4ic}<w_{f2ic}$, implying $V_{f4ic}<V_{f2ic}$ (since $V_{f4ic}-V_{f2ic}\approx\mathbb{A}(w_{f4ic}-w_{f2ic})$ from equation (25)). Thus, workers would be worse off if the firm responds compared to the status quo where the firm remains a simple non-union firm. Anticipating this, workers would have no incentive to threaten to unionise in the first place, and the union status would never be observed, which is inconsistent with the fact that we observe unions in practice. For this reason, we consider the more realistic case where $V_{f4ic}=V_{f2ic}$, which occurs when the intimidator pays the simple non-union wage.

First, using equations (43), (45), (46) and (63), we obtain:

$$
\begin{aligned}
\frac{\partial \Pi_{f1ic}(n^*_{f1ic})}{\partial \psi_{f1ic}} &= (1-\delta^m)\frac{\kappa}{q^v_c}\frac{1}{p_i\sigma_i} > 0 \\
\frac{\partial \Pi_{f2ic}(n^*_{f2ic})}{\partial \psi_{f1ic}} &= 0 \\
\frac{\partial \Pi_{f3ic}(n^*_{f3ic})}{\partial \psi_{f1ic}} &= (1-\delta^m)\frac{\kappa}{q^v_c}\frac{1}{p_i\sigma_i}(\tilde{\xi}_4 - \bar{\xi}\mathbb{A}) > 0 \\
\frac{\partial \Pi_{f4ic}(n_{f4ic})}{\partial \psi_{f1ic}} &= -(1-\delta^m)\frac{\kappa}{q^v_c}\frac{1}{p_i\sigma_i}\mathbb{A}(1-\tilde{\xi}_4) < 0 \qquad (64)
\end{aligned}
$$

and where $0 < (\tilde{\xi}_4 - \bar{\xi}\mathbb{A}) < 1$. Therefore, the optimal values in both the emulation and the union statuses increase with ψ_{f1ic}, but the optimal value increases faster with amenities in union firms compared to emulating firms. The optimal value of non-union firms is independent of amenities, while that of intimidating firms is a decreasing function of ψ_{f1ic}.

Next, we can compare optimal firm sizes to evaluate whether intimidation tactics dominate an emulation response at the thresholds. Combining equations (63) and (46), we have that:

$$
n^*_{f4ic} - n^*_{f3ic} = \frac{1}{p_i\sigma_i}\left[w_{f3ic} - (w_{f2ic} + cu_{fic})\right], \qquad (65)
$$

when δ^e is small.

At ψ^*_{fic}, the per-person intimidation cost is zero. Combining equations (49) and (54), we obtain:

$$
\begin{aligned}
n^*_{f4ic}(\psi^*_{fic}) - n^*_{f3ic}(\psi^*_{fic}) &= \frac{1}{p_i\sigma_i}\left[w_{f3ic}(\psi^*_{fic}) - w_{f2ic}\right] \\
&= \frac{1}{p_i\sigma_i}\left(1-\bar{\xi}\right)\left[w_{f1ic}(\psi^*_{fic}) - w_{f2ic}\right] > 0, \qquad (66)
\end{aligned}
$$

where $w_{f1ic}(\psi^*_{fic}) - w_{f2ic} > 0$ because $\psi^*_{fic} < \lambda_c$. Therefore, at ψ^*_{fic}, $\Pi_{f4ic}(n^*_{f4ic}) > \Pi_{f3ic}(n^*_{f3ic})$, and intimidation dominates the emulation response.

At $\psi^{**}_{fic} = \lambda_c$, the per-person intimidation cost simplifies to $cu_{fic} = \mathbb{A}\left[w_{f1ic}(\psi^{**}_{fic}) - w_{f2ic}\right]$, and, as can be seen from equation (49), $w_{f1ic}(\psi^{**}_{fic}) = w_{f3ic}(\psi^{**}_{fic})$. Hence,

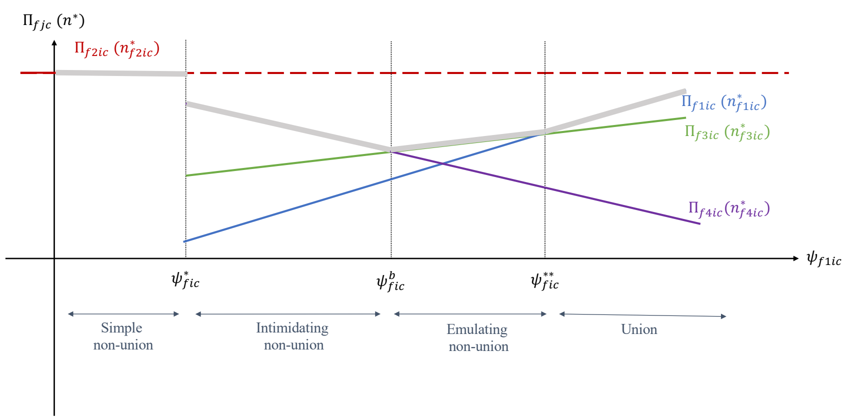
$$
\begin{aligned}
n^*_{f4ic}(\psi^{**}_{fic}) - n^*_{f3ic}(\psi^{**}_{fic}) &= -\frac{1}{p_i\sigma_i}\left(\mathbb{A}-1\right)\left[w_{f1ic}(\psi^{**}_{fic}) - w_{f2ic}\right] \\
&= -\frac{1}{p_i\sigma_i}\frac{\mathbb{A}-1}{\mathbb{A}}cu_{fic} < 0 \qquad (67)
\end{aligned}
$$

In consequence, at ψ^{**}_{fic}, $\Pi_{f3ic}(n^*_{f3ic}) > \Pi_{f4ic}(n^*_{f4ic})$ and the emulation response dominates intimidation at the upper threshold. Therefore, there must exist a threshold $\psi^b_{fic} \in (\psi^*_{fic}, \psi^{**}_{fic})$ at which both $\Pi_{f4ic}(n^*_{f4ic})$ and $\Pi_{f3ic}(n^*_{f3ic})$ are equalized and above which firm's optimal response switches from an intimidation to an emulation one.

Figure 1 plots the optimal value functions, by firm status.[12] The figure illustrates that introducing the possibility of intimidation introduces an additional amenity threshold, ψ^b_{fic}, that determines the intimidation status. Specifically, for $\psi_{f1ic} \in [\psi^*_{fic}, \psi^b_{fic}]$ the firm prefers to respond by intimidation rather than allowing workers to unionise or responding through emulation tactics. However, since $\frac{\partial \Pi_{f4ic}(n^*_{f4ic})}{\partial \psi_{f1ic}} < 0$ and $\frac{\partial \Pi_{f3ic}(n^*_{f3ic})}{\partial \psi_{f1ic}} > 0$, as amenities increase, it becomes increasingly costly for the firm to respond by intimidation, while becoming an emulator becomes more beneficial. Therefore, beyond the threshold ψ^b_{fic}, acting as an emulator is the dominating strategy. Finally, for $\psi_{f1ic} > \psi^{**}_{fic}$, the surplus of the firm in the emulation status becomes negative, and allowing workers to unionise dominates any type of firm response.

While the model suggests the existence of one additional type of firm, it is not clear that there is any advantage in separating simple non-union workers and intimidated non-union workers in practice since we are just studying the overall non-union wage and both of these subgroups receive the same wage. Moreover, there is nothing in our data to allow us to identify the simple non-union workers from the workers who are remaining non-union only because of threats. Given this, our approach is to estimate a specification that allows for emulation, testing to see if it exists but making no attempt to separate non-union workers into simple non-union, intimidated non-union and co-opted (emulating) non-union.

Figure 1: Optimal Value Functions by Firm Union Status



[12]Note that $\Pi_{f4ic}(n^*_{f4ic}) < \Pi_{f2ic}(n^*_{f2ic}) \ \forall \ \psi_{f1ic} \geq \psi^*_{fic}$ since

$$p_i\sigma_i\left(n^*_{f4ic} - n^*_{f2ic}\right) = -\frac{\beta}{1+\beta}\mathbb{A}^{-1}\mathbb{B}p_i\sigma_i n^*_{f2ic} - cu_{fic} < 0,$$

as can be seen by combining equations (63) and (38).

F Allowing the Job Destruction Rate to Vary by Union Status

In this section, we allow the job destruction rate to vary by union status, assuming that the death rate of matches can differ across jobs. Consequently, the job destruction rate becomes $\delta_k = \delta^e + (1-\delta^e)\delta_k^m$.

F.1 Matching Process

The matching process remains similar to that described in our baseline model, with slight adjustments to the equations that ensure a constant matching rate and sectoral composition in steady state. Importantly, none of these adjustments impact q_{kc}^v and $q_{kc|jc}^u$.

As before, the following conditions must hold in steady state. First, the number of effective matches in job k must equal the number of type k jobs that are destroyed, i.e. $\delta_k N_{kc} = M_{kc}$. Therefore,

$$\frac{\delta_k \eta_{kc}}{\sum_{k'} \delta_{k'} \eta_{k'c}} = \frac{M_{kc}}{M_c} \tag{68}$$

Second, effective matches must aggregate properly and determine both ϕ_{kc} and θ_{jc}. In particular, it must be that $\sum_k M_{kc|jc} = M_{c|jc}$, which, as before, implies that $\sum_k \phi_{kc}\chi_{kc|j}(\varphi_{k|j}) = 1 \ \forall j$. Given our assumed form for $\chi_{kc|j}(\varphi_{k|j})$, this condition is satisfied if $\phi_{kc} = \eta_{kc}$.

Finally, it must be that $\sum_j M_{kc|jc} = M_{kc}$, which implies that the proportion of city-specific matches that will be in job k are given by:

$$\frac{M_{kc}}{M_c} = \eta_{kc} \sum_j \theta_{jc}\chi_{kc|j}(\varphi_{k|j}) \tag{69}$$

Combining equation (69) together with equation (68) we obtain the following condition:

$$\frac{\delta_k}{\sum_{k'} \delta_{k'} \eta_{k'c}} = \sum_j \theta_{jc}\chi_{kc|j}(\varphi_{k|j}) \quad \forall k \tag{70}$$

Therefore, equation (70) yields a set of K equations, which, given job destruction rates, frictions, and steady state job shares, jointly determines the set of $\{\theta_{jc}\}_j$, within city c. Comparing equations (3) and (70), one can see that when the job destruction rate varies across jobs, $\sum_j \theta_{jc}\chi_{kc|j}(\varphi_{k|j})$ no longer equals one but becomes a function of both employment shares and the job destruction rates.

Importantly, combining these conditions, it remains the case that $q_{kc}^v = q_c^v$ and $q_{kc|jc}^u = q_c^u T_{kc|j}$, where $T_{kc|j} = \eta_{kc}\chi_{kc|j}(\varphi_{k|j})$ and $\chi_{kc|j}(\varphi_{k|j}) = \frac{\varphi_{k|j}}{\sum_{k'} \eta_{k'c}\varphi_{k'|j}}$. Therefore, in terms of the matching process, accommodating job-specific job destruction rates only affects how both θ_{jc} and $\frac{M_{kc}}{M_c}$ adjust.

F.2 Non-union Wage Equation

In what follows, we derive the wage equation for non-union workers. We first rewrite equations assuming that the job destruction rate is job-specific, and then consider the case where it varies by union status.

Firm Surplus Firm surplus is similar to the baseline, except that the destruction rate is now job-specific, i.e.:

$$S_{fjc}(n) \quad = \quad \frac{\partial \pi_{fjc}(n)}{\partial n} + \rho^e(1-\delta_j^m)\frac{\kappa}{q_c^v}, \quad \forall j = \{2, i\} \tag{71}$$

Worker Surplus The discounted value of employment takes the form:

$$V_{fjc}^E(w_{fjc}) = w_{fjc} + \mathbb{A}_j\psi_{fjc} + \rho(1-\delta_j)\mathbb{A}_j w'_{fjc} + \rho\delta_j\mathbb{A}_j V_{fjc}^U, \quad \forall j = \{2, i\} \tag{72}$$

where $\mathbb{A}_j = \frac{1}{1-\rho(1-\delta_j)}$, $0 < \mathbb{A}_j < 1$. Moreover, proceeding as described in Section C.3.1, we have the following expression for the discounted value of unemployment:

$$V_{jc}^U = \mathbb{D}_{jc} + \rho q_c^u \mathbb{A}_c \overline{\mathbb{B}}_c \sum_k \mathbb{A}_k T_{kc|j} w'_{kc}, \quad \forall j = \{2, i\} \tag{73}$$

where $\mathbb{D}_{jc} = \mathbb{A}_c \left[\rho\psi q_c^u \overline{\mathbb{B}}_c \sum_k \mathbb{A}_k T_{kc|j} + b\left(1 + \rho^2 q_c^u \mathbb{A}_c \overline{\mathbb{B}}_c \sum_k \delta_k \mathbb{A}_k T_{kc|j}\right)\right] > 0$, $\overline{\mathbb{B}}_c = \frac{1}{1-\rho^2\overline{\delta}_c q_c^u \mathbb{A}_c} > 0$ and $\overline{\delta}_c = \sum_k \delta_k \mathbb{A}_k \eta_{kc} > 0$.

Combining equations (72) and (73), the worker surplus writes as:

$$\begin{aligned}
V_{fjc}^E - V_{fjc}^U \quad &= \quad w_{fjc} + \mathbb{A}_j\psi_{fjc} + \rho(1-\delta_j)\mathbb{A}_j w'_{fjc} - (1-\rho\delta_j\mathbb{A}_j)\, V_{jc}^U \\
&= \quad w_{fjc} + \rho(1-\delta_j)\mathbb{A}_j w'_{fjc} - (1-\rho\delta_j\mathbb{A}_j)\left(\mathbb{D}_{jc} + \rho q_c^u \mathbb{A}_c \overline{\mathbb{B}}_c \sum_k \mathbb{A}_k T_{kc|j} w'_{kc}\right) \\
&\forall \quad j = \{2, i\}
\end{aligned} \tag{74}$$

where, since $j = \{2, i\}$, $\psi_{fjc} = 0$ in the second line of equation (74).

Individual Bargaining Individual bargaining solves the bargaining rule given by (37). Substituting in equations (71) and (74), solving the differential equation and focusing on the steady state where $w_{fjc} = w'_{fjc}$ yields:

$$\begin{aligned}
w_{fjc} \quad &= \quad \beta\mathbb{B}_j\left[p_i\left(\epsilon_{fic} - \frac{1}{1+\beta}\sigma_i n\right) + \rho^e(1-\delta_j^m)\frac{\kappa}{q_c^v}\right] + (1-\beta)\mathbb{B}_j\,(1-\rho\delta_j\mathbb{A}_j)\, V_{jc}^U, \\
&\forall \quad j = \{2, i\}
\end{aligned} \tag{75}$$

where $\mathbb{B}_j = \frac{1}{1+(1-\beta)\rho^e(1-\delta_j^m)\mathbb{A}_j} > 0$ and where, in steady state, $V_{jc}^U = \mathbb{D}_{jc} + \rho q_c^u \mathbb{A}_c \overline{\mathbb{B}}_c \sum_k \mathbb{A}_k T_{kc|j} w_{kc}$. Equations (38) and (75) have a similar structure but in the latter equation the coefficients on the variables ϵ_{fic}, n, $\frac{\kappa}{q_c^v}$ and V_{jc}^U are job-specific.

We can then use equation (75) along with the optimality condition to obtain an expression for optimal firm size. Proceeding as in Section C.5.2 we obtain:

$$\begin{aligned} n_{fjc} &= \frac{1}{p_i\sigma_i}\frac{(1+\beta)}{[1+\beta(1-2\mathbb{B}_j)]}\left\{(1-\beta\mathbb{B}_j)\,p_i\epsilon_{fic}-\left[\beta\mathbb{B}_j\rho^e(1-\delta_j^m)+\mathbb{A}_j^{-1}\right]\frac{\kappa}{q_c^v}\right\} \\ &- \frac{1}{p_i\sigma_i}\frac{(1+\beta)}{[1+\beta(1-2\mathbb{B}_j)]}\left\{(1-\beta)\mathbb{B}_j\,(1-\rho\delta_j\mathbb{A}_j)\,V_{jc}^U\right\} \quad \forall j=\{2,i\} \end{aligned} \tag{76}$$

Substituting equation (76) for optimal firm size into the wage equation (75), we can rewrite the wage of workers in non-union firms as:

$$w_{fjc} = \xi_{0,j}\mathbb{D}_{jc} + \xi_{1c,j}\sum_k \mathbb{A}_k T_{kc|j} w_{kc} + \xi_{2,j}\frac{\kappa}{q_c^v} + \xi_{3,j}p_i\epsilon_{fic} \quad \forall j=\{2,i\} \tag{77}$$

where the coefficients are given by $\xi_{0,j} = \frac{1+\beta(1-\mathbb{B}_j)}{1+\beta(1-2\mathbb{B}_j)}(1-\beta)\mathbb{B}_j\,(1-\rho\delta_j\mathbb{A}_j) > 0$, $\xi_{1c,j} = \xi_{0,j}\rho q_c^u\mathbb{A}_c\overline{\mathbb{B}}_c > 0$, $\xi_{2,j} = \frac{\beta\mathbb{B}_j}{1+\beta(1-2\mathbb{B}_j)}\left[\mathbb{A}_j^{-1}+\rho(1-\delta_j)(1-\beta\mathbb{B}_j)\right] > 0$, and $\xi_{3,j} = \beta^2\mathbb{B}_j(1-\mathbb{B}_j) > 0$.

Hence, equation (77) is also similar to the baseline wage equation (39) except that the coefficients on each variable are job-specific and that weights in the term capturing average wages are now multiplied by the corresponding job destruction rates.

Linear Approximation of the Wage Equation Finally, linearizing equation (77) as done in Section D, we get:

$$w_{fjc} = \gamma_{0,j} + \gamma_{1,j}\sum_k \mathbb{A}_k T_{kc|j} w_{kc} + \gamma_{2,j}ER_c + \gamma_{3,j}\epsilon_{fic} \quad \forall j=\{2,i\} \tag{78}$$

As before, $\gamma_{0,j}$ is a function of the price p_i, the destruction rate δ_j and constant terms stemming from the expansion point values. $\gamma_{1,j} = \xi_{1c,j}\big|_{\mathbf{x_0}} > 0$, and is, at $\mathbf{x_0}$, independent of the city subscript. $\gamma_{2,j}$ is, as for the baseline case, a complicated positive function of the underlying parameters evaluated at the common ER value, and $\gamma_3 = \xi_{3,j}p > 0$. Equation (78) has also a structure that is identical to the baseline wage equation with job-specific coefficients and weights that incorporate the destruction rates in the weighted average wage term.

Our goal is to examine the wage equation for non-union workers when the job destruction differs between union and non-union firms, i.e. when $\delta_j^m = \delta_\tau^m$ ($\delta_1^m \neq \delta_2^m$ but $\delta_2^m = \delta_3^m$). To this end, we rewrite equation (78) using union-status and industry subscripts, in place of job subscripts:

$$\begin{aligned} w_{f2ic} &= \gamma_{0,2i} + \gamma_{1,2i}\sum_{\tau'}\mathbb{A}_{\tau'}\sum_{i'} T_{\tau'i'c|2i} w_{\tau'i'c} + \gamma_{2,2i}ER_c + \gamma_{3,2i}\epsilon_{fic} \\ &= \gamma_{0,2i} + \gamma_{1,2i}\mathbb{A}_1 \underbrace{\sum_{i'} T_{1i'c|2i} w_{1i'c}}_{\text{Union component}} + \gamma_{1,2i}\mathbb{A}_2 \underbrace{\sum_{i'} T_{2i'c|2i} w_{2i'c}}_{\text{Non-union component}} + \gamma_{2,2i}ER_c + \gamma_{3,2i}\epsilon_{fic} \\ &= \gamma_{0,2i} + \gamma_{1,2i}\mathbb{A}_1 E_{1c|2i} + \gamma_{1,2i}\mathbb{A}_2 E_{2c|2i} + \gamma_{2,2i}ER_c + \gamma_{3,2i}\epsilon_{fic}, \end{aligned} \tag{79}$$

where $E_{1c|2i} = \sum_{i'} T_{1i'c|2i} w_{1i'c}$ captures the union component and $E_{2c|2i} = \sum_{i'} T_{2i'c|2i} w_{2i'c}$ reflects the non-union component, both of non-union workers' outside option.

Equation (79) shows that allowing for job destruction rates to vary between union and non-union firms leads to a non-union wage equation which is similar to the baseline model, with one key difference, namely that the coefficients on each variable are non-union-specific. Importantly, since we estimate the wage equation separately for union and non-union workers, our current setup already allows for this possibility.

G Simulating the Effects of Omitting the Rent Capture Term

In the text, we show that the term $(\tilde{\gamma}_3 - \gamma_3)\epsilon_{ict}\Delta P^{ne}_{ict}$ is present in our derived expression for the observed non-union wage and argue that it represents a spillover onto non-union wage setting of one channel through which unions obtain higher wages. In particular, it corresponds to unions capturing a higher share of firm rents (ϵ_{ict}) as reflected in $\tilde{\gamma}_3 > \gamma_3$. That term ends up in the error because we do not have observations on ϵ_{ict} that would allow us to generate a measure for the term. As part of our response, we are interested in simulating the likely extent of the effect of having this term in the error term on our estimated coefficients. In this section, we present this simulation exercise.

To start, we can represent our IV regression of the vector of coefficients in (80) as $\hat{\xi} = (Z'X)^{-1}Z'y$, where X is the matrix of right hand side variables, Z is the matrix of instruments, and $y = \Delta \ln w^n_{ict}$. Then, under the assumption that the instruments are mean independent of $\Delta\epsilon_{ict}$, the expectation of the estimated coefficients would equal $\xi + (\tilde{\gamma}_3 - \gamma_3) \cdot \alpha^*$, where $\alpha^* = (Z'X)^{-1}Z'(\epsilon_{ict}\Delta P^{ne}_{ict})$ and ξ is the vector of true parameter values. We are interested in simulating the α^* vector.

The α^* vector equals the vector of coefficients from an IV regression on the variables in the X matrix, using Z as instruments. We generate simulated versions of $\epsilon_{ict}\Delta P^{ne}_{ict}$ using the actual values of ΔP^{ne}_{ict} and random draws from a mean zero normal distribution with standard deviation equal to the standard deviation of the residuals from (80) to stand in for ϵ_{ict}. Note that the actual ϵ_{ict} values in (80) are correlated with the elements of X, but we don't need to recreate those correlations in our simulated ϵ_{ict} draws because they only enter the estimator in correlations with the elements of Z and ϵ_{ict} is independent of those elements by assumption.

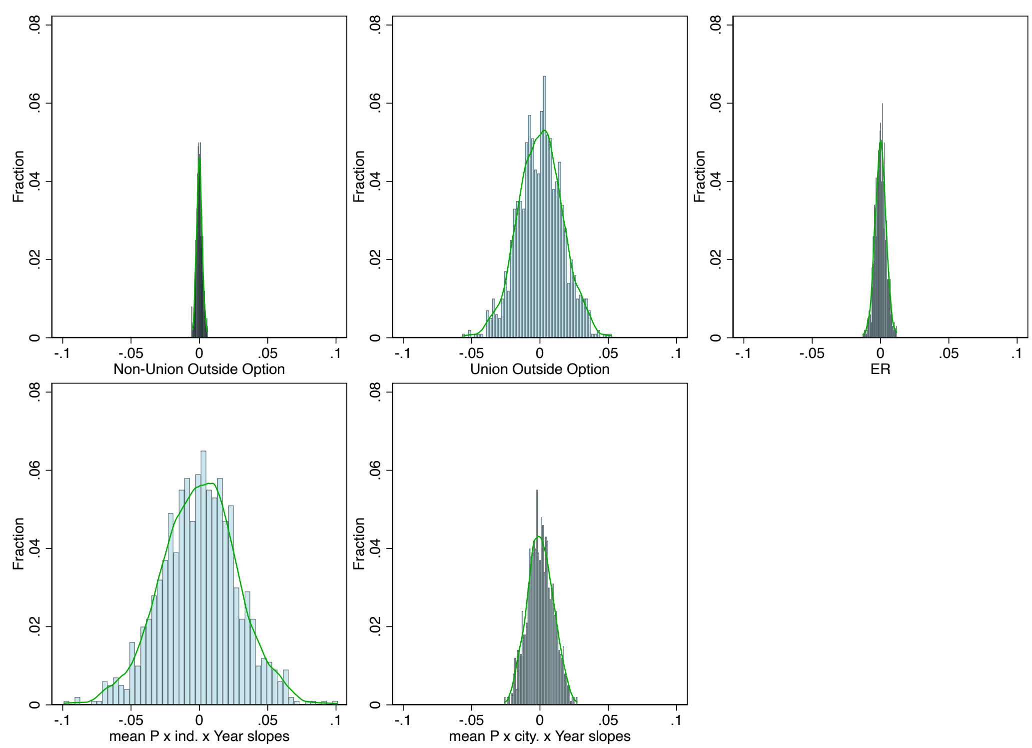
We then run an instrumental variables regression of this simulated $\epsilon_{ict}\Delta P^{ne}_{ict}$ on all the right hand side variables in (80). We re-run this simulated regression 1000 times. In Table 2, we present the mean, standard deviation, median, 5th and 95th percentiles of the distribution of the elements of α^* corresponding to $\Delta((1-P^{ne}_{ict})E_{2ict})$, $\Delta(P^{ne}_{ict}E_{1ict})$, and ΔER_{ct} (in lines 1, 2, and 3 of the table, respectively). We also present the same statistics for the distributions of the mean values of the P^{ne}_{ict-1} interacted with industry$\times$time effects and interacted with city$\times$time effects in lines 4 and 5 of the table, respectively. In Figure 2, we plot the associated histograms of the simulated values for these same coefficients.

The histograms show that all of the simulated distributions are centred very near zero and are roughly symmetric. Their spreads, though, are quite different, with a very narrow spread for the coefficient on $\Delta((1-P^{ne}_{ict})E_{2ict})$ and the widest spread on the interactions of P^{ne}_{ict-1} with industry x time effects. From this, we conclude that the omission of the rent capture term will have very little effect on the estimated $\Delta((1-P^{ne}_{ict})E_{2ict})$ coefficient or on the $\Delta(P^{ne}_{ict}E_{1ict})$, and ΔER_{ct} coefficients and will mostly be loaded onto the estimated $P^{ne}_{ict-1}\times$industry$\times$time effects.

Table 2: Simulation Results

	(1)	(2)	(3)	(4)	(5)
Coefficient					
	Mean	SD	Median	5th pct.	95th pct.
Non-Union Ouside Option	-0.00001	0.00207	-0.00004	-0.00334	0.00362
Union Ouside Option	0.00019	0.01629	0.00074	-0.02610	0.02711
Employment Rate	0.00003	0.00403	-0.00002	-0.00635	0.00673
Slopes: P x Ind. x Year	0.00006	0.02783	0.00073	-0.04611	0.04709
Slopes: P x City. x Year	0.00027	0.00942	0.00005	-0.01534	0.01629

Figure 2: Simulation Results



H Linearized Full Specification

In this section, we present results from a linearized version of the full specification (equation 30 in the paper). Recall that the full specification is given by

$$\begin{aligned}\Delta \ln w_{ict}^{n} = {} & \Delta\gamma_{0it} + \gamma_1 \Delta((1-P_{ict}^{ne})E_{2ict}) + \gamma_2 \Delta ER_{ct} \\ & + \tilde{\gamma}_1 \Delta(P_{ict}^{ne} E_{1ict}) + \Delta(\gamma_{0it}^{*} P_{ict}^{ne}) - \gamma_6 \Delta(P_{ict}^{ne}\lambda_{ct}) + (\tilde{\gamma}_3 - \gamma_3)\epsilon_{ict}\Delta P_{ict}^{ne} \\ & + \Delta\mu_{ict} + \gamma_3 \Delta\epsilon_{ict} + (\tilde{\gamma}_3 - \gamma_3) P_{ict-1}^{ne} \Delta\epsilon_{ict}.\end{aligned} \tag{80}$$

This specification, includes interactions of P_{ict}^{ne} with the outside options variables, industry×time and city×time effects, and both $\Delta\epsilon_{ict}$ and ϵ_{ict}. The latter interaction, in particular, introduces omitted variables bias issues that we discussed in Appendix Section G. One might be concerned that what we present as the effect of the outside option of union workers is really being identified by movements in P_{ict}^{ne} and, as a result, be interested in the separate effects of the key components in equation (80). For that reason, we estimated what might be viewed as a linearized version of (80) in which $\Delta \ln w_{ict}^{n}$ is regressed on: ΔE_{2ict}, ΔE_{1ict} , ΔER_{ct}, a complete set of P_{ict-1}^{ne} by city×time effects, and a complete set of P_{ict-1}^{ne}×industry×time effects. The latter effects capture effects of ΔP_{ict}^{ne} in a flexible way. We also report on a specification in which we just include ΔP_{ict}^{ne} on its own (i.e., not interacted with city and industry effects).

The results of this exercise are presented in Table 3. They show that the estimated coefficients on both ΔE_{2ict} and ΔE_{1ict} are positive and statistically significant. That is, even when entered on their own both the outside option of non-union workers and, most interestingly, the outside option of union workers affect non-union wage setting. Notice that if we arrive at the linear equation by linearizing (80) around a point where P_{ict}^{ne} takes a common value across cities, time and industries, $\bar{P}^{ne}$ then the coefficient on ΔE_{2ict} in the linearized regression is actually, $\gamma_1 \cdot \bar{P}^{ne}$ and the coefficient on ΔE_{1ict} is given by, $\tilde{\gamma}_1 \cdot (1-\bar{P}^{ne})$. Thus, would expect that the estimated coefficient on ΔE_{2ict} should be smaller than the coefficient on $\Delta(P_{ict}^{ne}E_{2ict})$ in the full, non-linear equation, and the coefficient on ΔE_{1ict} in the linearized equation should be even smaller relative to its counterpart in the non-linear specification. This, in fact, is what we observe.

I Standard Errors

In this Appendix, we consider the implications of recent papers on standard errors in Bartik style estimators: Borusyak et al. (2022) and Adao et al. (2020).

I.1 Borusyak et al. (2022)

Borusyak et al. (2022) examine Bartik-type estimators, arguing that those estimators can be implemented by aggregating to the industry level (job level, in our case) and then running a simple IV at that level. A key equation is the expression of the IV estimator:

$$\hat{\beta} = \frac{\sum_c IV1_{ct} w_{ct}}{\sum_c IV1_{ct} E_{ct}} \tag{81}$$

Table 3: Non-Union Wages and Outside Options: OLS and 2SLS Estimates

	OLS	2SLS				
	(1)	(2)	(3)	(4)	(5)	(6)
ΔE_{2ict}	-0.29***	0.37***	0.38***	0.40***	0.46***	0.69***
	(0.030)	(0.14)	(0.14)	(0.13)	(0.11)	(0.067)
ΔE_{1ict}	1.36***	0.33***	0.32***	0.32***	0.29***	
	(0.026)	(0.10)	(0.10)	(0.10)	(0.11)	
ΔER_c	-0.21***	0.33*	0.30			
	(0.075)	(0.18)	(0.18)			
ΔP_{ict}^{ne}					0.28**	0.35**
					(0.12)	(0.13)
Obs.	6081	6081	6081	6081	5958	5958
R^2	0.84					
Year $\times$ Ind.	Yes	Yes	Yes	Yes	Yes	Yes
$P_{ict-1}^{ne} \times$ Ind. $\times$ Year	Yes	Yes	Yes	Yes		
$P_{ict-1}^{ne} \times$ City $\times$ Year	Yes	Yes	Yes	Yes		
Select controls						
ΔP_{ic} Quadratic			Yes	Yes	Yes	Yes
Instrument set:		$IV1_{2ict}$ $IV1_{1ict}$	$IV1_{2ict}$ $IV1_{1ict}$	$IV1_{2ict}$ $IV1_{1ict}$	$IV1_{2ict}$ $IV1_{1ict}$ $\Delta\hat{P}_{ict}^{ne}$	$IV1_{2ict}$ $IV1_{1ict}$ $\Delta\hat{P}_{ict}^{ne}$
First-Stage p-Stat.:						
ΔE_{2ict}		0.000	0.000	0.000	0.000	0.000
ΔE_{1ict}		0.000	0.000	0.000	0.000	0.000
ΔP_{ict}^{ne}					0.000	
Over-id. p-val		.	.	.	.	.

Notes: This table displays results from the estimation of equation (21) via OLS (column 1) and 2SLS (columns 2 - 5). The dependent variable is the decadal change in the regression adjusted average hourly wage of non-union workers in an industry-city cell, using CPS data from 1980-2019 across 50 industries and 93 cities. Standard errors, in parentheses, are clustered at the city-year level.

$$= \frac{\sum_c (\sum_k \omega_{kct-1} g_{kt}) w_{ct}}{\sum_c (\sum_k \omega_{kct-1} g_{kt}) E_{ct}}$$

where we have substituted in instruments and variables using our terminology instead of theirs. In particular, ω_{kct-1} is a weight that equals the share of employment in city c in period $t-1$ that is in sector k in the simple Bartik instrument case, g_{kt} is the national level growth rate in employment in sector k, w_{ct} is the wage in city c in period t, and E_{ct} is the outside offer value for workers in city c in period t. We have also assumed there is equal weight on all cities.

The numerator in this expression can be re-written:

$$\sum_k g_{kt} \sum_c \omega_{kct-1} w_{ct} = \sum_k g_{kt} \omega_{kt-1} \bar{w}_{kt} \tag{82}$$

where $\omega_{kt-1} = \frac{1}{C} \sum_c \omega_{kct-1}$ and $\bar{w}_{kt} = \frac{\sum_c \omega_{kct-1} w_{ct}}{\sum_c \omega_{kct-1}}$. The same can be done with the denominator, implying that $\hat{\beta}$ can be obtained by first aggregating the left and right-hand side variables in a regression of w on E in to the industry level in a specific way (in our case, to the job level) then running an IV regression with the $g's$ (in our case, the combination of job rents and job employment growth rates at the national level) as instruments.

The question for us is how this applies to our scenario. Note, first, that both our right and left-hand side variables vary at the job and city level, rather than just the city level. We could think of carrying out their approach for each job then aggregating the resulting $\hat{\beta}_i$ estimates. Given our use of the transition probabilities, the ω's would be different for each base job. To get the standard errors right, we could run this as a Seemingly Unrelated Regression specification.

However, there is an extra issue for us which is noticeable when we write out the definition of $IV1_{jct}$:

$$IV1_{jct} = \sum_k \eta_{kct-1} g_{kt} \frac{\varphi_{kt|j}}{\sum_{k'} \eta_{k'ct-1} g_{k't} \varphi_{k't|j}} \nu_{kt} - \sum_k \eta_{kct-1} \frac{\varphi_{kt-1|j}}{\sum_{k'} \eta_{k'ct-1} \varphi_{k't-1|j}} \nu_{kt-1} \tag{83}$$

where ν_{kt} is the job premium in industry k in period t, and g_{kt} is the growth rate of employment in job k, both at the national level.

Now return to the exercise of multiplying this by the dependent variable and summing across cities to get the numerator of the estimator (for the moment, thinking about running this just for one base job, j):

$$\sum_c \Big(\sum_k \eta_{kct-1} g_{kt} \frac{\varphi_{kt|j}}{\sum_{k'} \eta_{k'ct-1} g_{k't} \varphi_{k't|j}} \nu_{kt} - \sum_k \eta_{kct-1} \frac{\varphi_{kt-1|j}}{\sum_{k'} \eta_{k'ct-1} \varphi_{k't-1|j}} \nu_{kt-1}\Big) w_{jct} \tag{84}$$

$$= \sum_c \Big(\sum_k \eta_{kct-1} g_{kt} \frac{\varphi_{kt|j}}{\sum_{k'} \eta_{k'ct-1} g_{k't} \varphi_{k't|j}} \nu_{kt}\Big) w_{jct}$$

$$- \sum_c \Big(\sum_k \eta_{kct-1} \frac{\varphi_{kt-1|j}}{\sum_{k'} \eta_{k'ct-1} \varphi_{k't-1|j}} \nu_{kt-1}\Big) w_{jct}$$

We can, again, do the trick of reversing the summations giving:

$$\sum_k \nu_{kt} g_{kt} \sum_c \eta_{kct-1} \frac{\varphi_{kt|j}}{\sum_{k'} \eta_{k'ct-1} g_{k't} \varphi_{k't|j}} w_{jct}$$

$$
- \sum_k \nu_{kt-1} \sum_c \eta_{kct-1} \frac{\varphi_{kt-1|j}}{\sum_{k'} \eta_{k'ct-1} \varphi_{k't-1|j}} w_{jct}
$$

We could then write this in their aggregation as:

$$
\sum_k \nu_{kt} g_{kt} \sum_c \omega_{kjct} w_{jct} - \sum_k \nu_{kt-1} \sum_c \omega_{kjct-1} w_{jct} \tag{85}
$$

$$
= \sum_k \nu_{kt} g_{kt} \omega_{kjt} \bar{w}_{kt} - \sum_k \nu_{kt-1} \omega_{kjt-1} \bar{w}_{kt-1}
$$

where, the ω_{kjt} and $\bar{w}_{kt}$ are defined as before but because of changes in the transition rates over time, the weights (ω's) and then the weighted wages ($\bar{w}$) will differ for the start and end of the period. That means that no further reduction is possible. Since the same thing is going on in the denominator with the outside option variable, we end up with both a numerator and a denominator which are expressed in terms of differences in the kinds of aggregated variables they create. That means there is not a simple way to implement their approach: we cannot simply aggregate and then run linear IV. This, in turn, means that we can't take the approach of correcting standard errors by taking this aggregation route.

I.2 Adao et al. (2020)

But regardless of whether we can implement the correction approach from Borusyak et al. (2022), we still potentially face a problem with our standard errors. Using a similar sector-focused framework as Borusyak et al. (2022), Adao et al. (2020) argues that in standard Bartik regressor or instrument specifications, we should expect a variant of a clustering problem to affect our standard errors. In particular, they argue that there could be other, unobserved sectoral shifters, μ_{kt} in their notation, that end up in the error term in a Bartik form. That is, the error term contains a term such as $\sum_k \eta_{kct-1}\mu_{kt}$. They assume that the different shifters (the observed ones that appear in the Bartik variable and the unobserved ones that appear in the error term) are independent of each other and, so, the existence of this other shift-share term in the error does not raise consistency concerns. But any two regions with similar weights, η_{kct}'s, will have similar values of the Bartik variable and of the term in the error. That generates standard clustering-type problems, where the clusters here are regions with similar start-of-period sectoral shares.

This raises two questions. First, do their arguments carry over to our specific setting and, second, is their proposed correction implementable in our context? On the first question, we return to the point that the weights used in our Bartik instruments vary not just with sectors (jobs) and the local economy (k and c) – they also vary with the industry (job) the specific worker is in because the weights include the transition rates that depend on the initial job type, j. To have the exact same problem outlined in Adao et al. (2020), we would have to think of unobservable industry shocks that are aggregated in this same specific way with weights that vary by both j and k. Since the weights arise from the specific form of wage bargaining, it is not clear what such other shocks would be. This relates to Adao et al. (2020)'s statement: "standard inference methods lead to over-rejection if the residual contains important shift-share terms that affect the outcome of interest through the same shares ... as those defining the covariate of interest" (p. 1974). In our context, the shares are the $T_{kc|j}$ terms capturing access to other job types and the covariate of interest is $IV1_{jct}$, and we are arguing that in the context of our model there are no other candidate shocks that would be aggregated using these weights. Any sectoral demand shocks, for example, would only affect bargaining through either the wages that are already a part of the outside option variable or through altering labour market tightness, which we control for using the change in the city employment rate. Further, our theory indicates that our error term consists of industry-city level aggregates of firm productivity shocks. Taken together, this suggests that we don't expect the problem outlined in Adao et al. (2020) to plague our standard errors.

The second question is whether we can implement their correction in our estimation context, which we might want to do out of an abundance of caution. We could, in this case, assume that there are unobserved sectoral shifters, μ_{kt}, that enter the error term order through a standard Bartik aggregator: $\sum_k \eta_{kct-1}\mu_{kt}$. In the context of our model, noting that our error term consists of sector-city aggregates of firm specific productivity shocks, this would arise if the productivity shocks in one sector depended on the productivity shocks in all other sectors. While we have no reason to assume this structure for the error term, we present standard errors clustered at the city-year level to address this. Note that one might also be concerned about common sectoral shocks, but we already account for these by including a complete set of industry×time effects in our differenced specifications.

J Data Appendix

Our CPS data is downloaded from the National Bureau of Economic Research (NBER).

We construct potential experience as max(min(age-years of schooling-6, age-16),0), dropping those with negative potential experience. We use the approach in Jaeger & Page (1996) to convert the years of completed schooling recorded in the MORG prior to 1992 to the post-1992 education categories. Because of limitations in the union coverage question, we define union workers as workers reporting belonging to a labour union.

We follow Lemieux (2006) closely in the construction of our wage data, working with weekly wages. Specifically, wages are based on individuals reporting employment in the reference week as wage and salary workers. We drop observations with allocated wages, and for workers paid hourly we use hourly earnings multiplied by usual weekly hours worked. For workers not paid hourly, we use edited weekly earnings, multiplying the weekly earnings topcode by 1.4 for topcoded observations. Wages are converted to 2000 dollars using a CPI deflator. We drop observations with an hourly wage below 1 or greater than 100 in 1979 dollars. All calculations use the earnings weights provided in the data. We aggregate the highest degree obtained into four categories (less than high school, high school graduate, some post-secondary, and university degree). For years before 1992, we use Table 5 from Park (1994) to construct education categories from the number of completed years of education.

We define industry using an aggregated grouping of industry codes based on the 1980 industrial classification from the Census Bureau. We obtain a consistent industry classification using crosswalks provided by IPUMS and the Census Bureau that map the 1970, 1990, and 2000 industry codes to the 1980 classification.[13] The result is a consistent classification

[13] Available at https://www.census.gov/topics/employment/industry-occupation/guidance/code-lists.html and https://usa.ipums.org/usa/volii/occ_ind.shtml

system with 50 industries. Table 4 shows the relationship between this detailed industry definition and the 1990 industrial classification system used by the Census Bureau.

We construct a set of cities with as consistent geographic boundaries as possible, given data limitations in the CPS. We are constrained by the number of SMSA's available in the May extract data and end up with 43 cities. Making use of the limited number of counties identified in the CPS, we are able to create a set of cities which are reasonably, though not always perfectly, consistent over time.[14] The final geographic definition we use pools data for these 43 cities and the remaining population. Specifically, we create additional regions made up of the remaining state population absent the population living in these 43 cities. In the end, our core geographic measure is composed of 93 areas that are fairly consistently defined over the course of the sample period. Tables 6 and 5 contain details.

Additionally, we use data on union elections to proxy for the threat probability, P^{ne}_{ict} in our model. The idea is that location x industry cells where the proportion of union certification elections that result in certification is high are more union-friendly and, therefore, will have a higher threat of unionisation for non-union firms. To obtain these proportions we use National Labor Relations Review Board (NLRB) case data for the three year periods for which we use CPS data.[15] We focus on certification elections and cases where a conclusive decision on certification was reached.[16] We use the county of the unit involved in the election to construct our geographic measures, aggregating counties to our city definition discussed above.

More specifically, the procedure to construct our P^{ne}_{ict} proxy is:

1. Using NLRB data, count the number of elections in each ic cell from 1977-2010 that resulted in union certification, NE_{ic}.

2. Using the CBP data, count the number of establishments in each ic cell from 1977-2010, $Estab_{ic}$

3. For each of our main data years where the NLRB data is available (1980, 1990, 2000 and 2010) we compute the ratio of the number of successful certification drives over the previous 4 years in an ic cell divided by the number over non-union establishments in the ic cell at the start of that 4 year period.

Table 6: Changes to SMSA Definitions 1973-2010

	1973-1980		1981-1989	1993-2003	2004-2020
Chicago	Cook	Lake	Kendall Added		
	Du Page	McHenry	Grundy Added	Dekalb Added	
	Kane	Will			
Philadelphia	Burlington	Chester		Salem Added	
	Camden	Delaware			

[14] The metropolitan area definition used by the IPUMS identifies a general pattern of expanding metropolitan area definitions over time that we overcome to some extent, but not perfectly: https://usa.ipums.org/usa/volii/county_comp2b.shtml. Estimation using states as the geographic unit yields very similar results, suggesting that issues related to geographic definitions are not driving our results.

[15] Our thanks to Hank Farber for providing this data.

[16] As opposed to the case being dismissed or withdrawn.

	Gloucester	Montgomery				
	Bucks	Philadelphia				
Detroit	Lapeer	Oakland	Monroe Added		Lenawee Added	
	Livingston	St.Clair			Washtenaw Added	
	Macomb	Wayne				
Washington	District of Columbia	Arlington	Calvert Added		Fauquier Added	King George Dropped
	Montgomery	Fairfax	Charles Added		Clarke & Warren Added	Rappahannock Added
	Prince George's	Fairfax city	Frederick Added		Culpeper Added	
	Alexandria	Falls Church	Loudoun Added		King George Added	
			Prince William Added		Spotsylvania Added	
			Masassas Added		Jefferson Added	
			Masassas Park Added		Fredericksburg Added	
			Stafford Added		Berkeley Added	
Boston	Essex	Plymouth	Bristol Added			Bristol Dropped
	MiddleSex	Suffolk				Essex Dropped
	Norfolk		Worchester Added			
Pittsburgh	Allegheny	Washington	Fayette Added		Butler Added	Armstrong Added
	Beaver	Westmoreland				
St Louis	Clinton	Jefferson	Jersey Added		Lincoln Added	Macoupin Added
	Madison	St. Charles			Warren Added	Bond Added
	Monroe	St. Louis				Calhoun Added
	St. Clair	St. Louis city				
	Franklin					
Baltimore	Anne Arundel	Carroll	Queen Anne's Added			
	Baltimore city	Harford				
	Baltimore	Howard				
Cleveland	Cuyahoga	Lake			Added Ashtabula	
	Geauga	Medina			Added Lorain	
Houston	Brazoria	Liberty			Added Chambers	Added Austin
	Fort Bend	Montgomery				Added Galveston
	Harris	Waller				
Newark	Essex	Sussex				Union Dropped
	Morris	Union				
Minneapolis-	Anoka	Ramsey	Isanti Added		Sherburne Added	
St Paul	Carver	Scott				
	Chisago	Washington				
	Dakota	Wright				
	Hennepin					

Dallas-	Collin	Wise	Wise Dropped		Henderson Added	Wise Added
Fort Worth	Dallas	Hood	Hood Dropped		Hunt Added	Somerwell Added
	Denton	Johnson			Hood Added	
	Ellis	Tarrant				
	Kaufman	Parker				
	Rockwall					
Seattle-Everett	King	Snohomish			Island Added	Pike Added
Atlanta	Cherokee	Gwinnett	Barrow Added		Butts dropped	Butts Added
	Clayton	Henry	Coweta Added		Carroll Added	Dawson Added
	Cobb	Newton	Spalding Added		Bartow Added	Haralson Added
	De Kalb	Paulding				Heard Added
	Douglas	Rockdale				Jasper Added
	Fayette	Walton				Lamar Added
	Forsyth	Butts				Meriwether Added
	Fulton					Morgan Added
Cincinnati	Dearborn	Clermont			Ohio Added	Union Added
	Boone	Hamilton			Gallatin Added	Bracken Added
	Campbell	Warren			Grant & Brown Added	Butler Added
	Kenton				Pendelton Added	
Kansas City	Johnson	Jackson	Lafayette Added		Clinton Added	Linn Added
	Wyandotte	Platte	Leavenworth Added			Bates Added
	Cass	Ray	Miami Added			Caldwell Added
	Clay					
Denver	Adams	Denver				Adams Dropped
	Arapahoe	Douglas				Broomfield Added
	Boulder	Jefferson				Clear Creek Added
						Elbert & Park Added
						Gilpin Added
Indianapolis	Boone	Johnson				Brown Added
	Hamilton	Marion				Putnam Added
	Hancock	Morgan				
	Hendricks	Shelby				
New Orleans	Jefferson	St. Bernard	St Charles Added		St James Added	
	Orleans	St.Tammany	St John the Bap. Added		Plaquemines Added	
Tampa-	Hillsborough	Pinellas	Hernando Added			
St Petersburg	Pasco					
Portland	Clackamas	Washington			Clark Added	
	Multnomah	Yamhill			Columbia Added	
Columbus	Delaware	Madison	Licking Added		Licking Dropped	Licking Added

	Fairfield	Pickaway	Union Added			Hocking Added
	Franklin					Morrow Added
Rochester	Livingston	Orleans			Genesee Added	
	Monroe	Wayne				
	Ontario					
Sacramento	Placer	Yolo			El Dorado Added	
	Sacramento					
Birmingham	Jefferson	Walker	Blount Added		Walker Dropped	Walker Added
	Shelby	St. Clair				Bibb & Chilton Added
Albany-	Albany	Schenectady	Greene Added		Greene Dropped	
Schenectady-Troy	Rensselaer	Montgomery			Schoharie Added	
	Saratoga					
Norfolk-	Currituck	Portsmouth	Currituck Dropped		Currituck Added	
Portsmouth	Chesapeake	Virginia Beach	Gloucester Added		Isle of Wight Added	
	Norfolk		Hampton & Suffolk Added		Mathews Added	Gloucester Added
			James & York Added			
			Newport News Added			
			Poquoson Added			
			Williamsburg Added			
Greensboro-	Forsyth	Yadkin	Davie Added		Alamance Added	Alamance Dropped
Winston-Salem-	Guilford	Stokes				
High point	Randolph	Davidson				
Gary-Hammond	Lake	Porter				Jasper Added
East Chicago						Newton Added
Portland	Clackamas	Multnomah			Columbia Added	
	Washington	Yamhill				

Notes: Changes to the counties/cities/parishes, included in the SMSA definitions over the sample period. There are no county changes for New York, Patterson, Nassau-Suffolk, Los Angeles, San Francisco, Anaheim, Milwaukee, San Diego, Buffalo, Miami, San-Bernadino, San Jose, Akron.

J.1 Job-to-job Transition Rates

We proxy the mobility friction $\varphi_{k|j}$ using the transition rate from the industry-union status cell, j, to the industry-union status cell, k, at the national level, constructed with additional data from IPUMS-CPS. For years 1990, 2000, 2010, and 2016 all transitions are constructed using IPUMS data, which contains a necessary unique identification variable, allowing us to determine which cell a person is in year $t-1$ and year t. We compute $\varphi_{k|j}$ as the proportion of people observed in year k in year t conditional on them being in a cell j in year $t-1$. For 1980, we match IPUMS identification data to the May extracts, as union data is not contained in IPUMS for these years.

We perform the match using household identifiers and personal characteristics. It is not possible to track individuals for most of 1981 and for all of 1982 in the May extracts. To overcome this limitation, we extend the range of years used to calculate transitions. Using the May extracts, we match individuals from 1977 to 1981, and we match individuals from 1983 to 1984 using the MORG data.

J.2 Instrument Construction

In order to construct our instruments, we need (1) estimates of the national industrial premia, and (2) to predict local union and non-union employment composition.

Estimating national wage premia. We estimate separate worker-level log wage regressions for each of our set of sample years at the national level, working with pooled union and non-union workers. The regressions include the same set of skill and demographic variables used when forming our residualized wages for the dependent variable, plus a complete set of industry dummy variables interacted with a union dummy. We interpret the coefficients on the industry dummies as rents that are allowed to differ in the union and non-union sectors. We define the industry dummy variables such that the coefficient values are defined relative to the overall average wage. We then replace the wages, w_{kct}, with the industry-union status cell wage premia, which we call ν_{kt} in the outside option expressions. We do this separately at each decade point.

Predicting local job shares. We construct the predicted local job shares as follows. First, we construct predicted employment levels using start-of-period employment at the job-city level combined with national-level growth rates for the relevant job:

$$\hat{N}_{jct} = N_{jct-1} \cdot \left(\frac{N_{jt}}{N_{jt-1}} \right)$$

We then form predicted city-level employment as $\hat{N}_{ct} = \sum_j \hat{N}_{jct}$ and, from that, we construct predicted job employment shares as $\hat{\eta}_{jct} = \frac{\hat{N}_{jct}}{\hat{N}_{ct}}$.

K Other Results

K.1 Union Wage Estimation

In this section, we report on results using the union wage as the dependent variable. In particular, we work from the wage specification we derive from our model, which we recreate here:

$$w_{fjc} = \tilde{\gamma}_{0i} + \tilde{\gamma}_1 \sum_k T_{kc|j} w_{kc} + \tilde{\gamma}_2 ER_c + \tilde{\gamma}_3 \epsilon_{ic} + \tilde{\gamma}_3 u_{fic} - \tilde{\gamma}_4 \psi_{fjc} \quad \forall j = \{1, i\} \tag{86}$$

As with the non-union wage equation, the specification calls for including a complete set of industry×time effects, so that identification is across cities within industries. Also note that we have been using the fact that the productivity term can be decomposed as $\epsilon_{fic} = \epsilon_{ic} + u_{fic}$.

Estimation of this equation requires addressing the same issues as for the non-union wage equation. In particular, we work with industry×city mean union wages obtained as the coefficients on a complete set of industry×city dummy variables in an individual log wage regression which includes the same flexible set of controls for age, education, sex, and race as in the non-union case. As with the non-union case, we drop cells with fewer than 10 observations. Because of the decline in unionisation, this results in a sample size with about one-third the number of industry×city cells as in the non-union case. We again weight the observations using the square root of the number of observations in the cell.

As with the non-union case, we have to address selectivity (which we again do use the overall proportion of workers in a cell who are union or a union election variable), endogeneity and reflection issues (which we do using the same form of instrument as in the non-union case, differing only in the fact that we use transition rates for union rather than non-union workers). To form the instrument, we use the same predicted, end of decade shares in industry×union status shares ($\hat{\eta}_{kct}$) as before and with those, we form the predicted end-of-decade option value:

$$\hat{E}_{1ict} = \sum_{k \neq 1i} \hat{\eta}_{kct} \frac{\varphi_{kt|1i}}{\sum_{k'} \hat{\eta}_{k'ct} \varphi_{k't|1i}} \nu_{kt}, \tag{87}$$

which differs from $\hat{E}_{2ict}$ only in that the transition rates are formed from union rather than non-union cells (i.e., are $\varphi_{kt|1i}$ rather than $\varphi_{kt|2i}$). We also construct $\tilde{E}_{1ict-1}$, which is a slightly altered version of E_{1ict} where we have replaced w_{kt} with ν_{kt}, the job type wage premium at the national level. Then, using both $\hat{E}_{1ict}$ and $\tilde{E}_{1ict}$, we form our instrument for ΔE_{1ict}:

$$IV1_{1ict} = \hat{E}_{1ict} - \tilde{E}_{1ict-1}, \tag{88}$$

As before, the cross-city variation in this instrument comes from the η_{kct-1}'s since $\hat{\eta}_{kct}$ is a function of η_{kct-1} and all the other terms in (88) do not vary at the city level. Thus, the validity of the instrument requires that the η_{kct-1}'s are independent of the relevant variation in the error term: cross-city variation in productivity growth. That is, we require an assumption that the productivity process follows a random walk (since, as BGS show, the η_{kct}s can be written as functions of the ϵ_{ict}s).

We present the results from estimating (86) in Table 7 below. As with the non-union specification, we present results in which we divide the outside option into a term related to finding a union and one related to finding a non-union job. The first stage remains strong, even with the reduced sample size. The estimated outside option effects are positive, as the theory would predict, but not significant at any standard significance level. In later years in particular, the public sector makes up an important part of the union jobs sample. Because we don't necessarily expect the same model of wage determination to hold for public sector jobs, we present results without the public sector in Table 8. Dropping public sector observations reduces the sample size by 16% and results in estimated outside option effects that are very similar in size to the effect of non-union workers' outside options on the non-union wage. Note that dropping the public sector does not have a substantial effect on the non-union equation estimates because public sector workers are mainly unionised. Interestingly, union jobs as outside options have a larger impact on union wage setting than non-union jobs, which contrasts with the results for non-union workers, where their effects are very similar in size. In Table 9, we present results excluding the public sector but including our selectivity controls. In contrast to the non-union specification, the selection terms are jointly significant at least the 10% significance level, and their inclusion reduces the estimated outside option coefficients to a small degree. In the final column of the table, we include the ΔP^{ne}_{ict} variable as a covariate. There is no reason for this variable to enter according to our theory and, in fact, we cannot reject the null hypothesis that its true effect is zero (the standard error on its coefficient is double the estimated coefficient). In addition, including it does not change the coefficient on the outside option variable. Overall, we view these results as indicating that our specification also fits for union wages in the private sector.

K.2 Union Wage Premia

In the paper, we raise the possibility that our model can explain the increase in the union wage premium in the 1980s – at a time when union power was declining. Under our model, this could happen because the decline in the threat of unionisation means that fewer of the observed non-union firms are in the emulation category and those that do not have to pay as high an emulating wage. This would have the effect of lowering the observed non-union wage and, so, increasing the union wage differential. To check on whether this mechanism seems to be happening at all, we regress changes in the union wage premium at the industry$\times$city cell level on changes in our union threat probability variable (P_{ict}^{ne}). Constructing the union wage premium, of course, requires the union wage and because unionisation declines to quite low numbers in our period, there are many cells for which we cannot construct the mean union wage. In fact, working with the union premium as the dependent variable results in losing approximately two-thirds of our cells. Because this creates issues with influential outliers, especially in the hospital sector, we winsorize the threat probability variable at the 99th percentile.

The results from this exercise are in Table 10, below. They show that there is, indeed, a negative relationship between the union wage premium and the threat probability that is statistically significant at the 5% level. Thus, periods of declining union threats within city$\times$industry cells are also periods of increasing union wage premia as our model predicts.

K.3 Results Omitting the Public Sector

In this section, we present results from our simple and full specification using data in which we drop observations associated with the public sector. This is done out of a concern that wage setting is substantially different from the private sector and, potentially, substantially different from our model. Note, though, that we keep public sector wages in our outside option values since they continue to be alternative wage options that workers in the private sector could point to during wage bargaining. As it turns out, because the public sector is largely unionised, dropping non-union public sector observations has little impact on our non-union wage estimates. This can be seen in 11.

Table 4: Aggregated Industry Definitions

Category	Code	1990 Industry Codes
Agriculture Service	1	12, 20, 21 , 30
Other Agriculture	2	10 - 11
Mining	3	40 - 50
Construction	4	60
Lumber and Wood Products, except Furniture	5	230 - 241
Furniture and Fixtures	6	242
Stone Clay, Glass, and Concrete Product	7	250 - 262
Primary Metals	8	270 - 280
Fabricated Metal	9	281 - 300
Not Specified Metal Industries	10	301
Machinery, except Electrical	11	310 - 332
Electrical Machinery, Equipment, and Supplies	12	340 - 350
Motor Vehicles and Equipment	13	351
Aircraft and Parts	14	352
Other Transportation Equipment	15	360 - 370
Professional and Photographic Equipment, and Watches	16	371 - 382
Toys, Amusements, and Sporting Goods	17	390
Miscellaneous and Not Specified Manufacturing Industries	18	391 - 392
Food and Kindred Products	19	100 - 122
Tobacco Manufactures	20	130
Textile Mill Products	21	132 - 150
Apparel and Other Finished Textile Products	22	151 - 152
Paper and Allied Products	23	160 - 162
Printing, Publishing and Allied Industries	24	171 - 172
Chemicals and Allied Products	25	180 - 192
Petroleum and Coal Products	26	200 - 201
Rubber and Miscellaneous Plastics Products	27	210 - 212
Leather and Leather Products	28	220 - 222
Transportation	29	400 - 432
Communications	30	440 - 442
Utilities and Sanitary Services	31	450 - 452, 460 - 472
Wholesale Trade	32	500 - 571
Retail Trade	33	580 - 691
Banking and Other Finance	34	700 - 710
Insurance and Real Estate	35	711 - 712
Private Household Services	36	761
Business Services	37	721, 722, 731 - 750, 892
Repair Services	38	751 - 760
Personal Services, except Private Household	39	762 - 791
Entertainment and Recreation Services	40	800 - 802, 810
Hospitals	41	831
Health Services, except Hospitals	42	812 - 830, 832 - 840
Educational Services	43	842 - 860
Social Services	44	861 - 871
Other Professional Services	45	730, 841, 872 - 891, 893
Forestry and Fisheries	46	31 - 32
Justice, Public Order and Safety	47	910
Administration Of Human Resource Programs	48	922
National Security and Internal Affairs	49	932
Other Public Administration	50	900, 901, 921, 930, 931

Notes: List of aggregated industries and corresponding 1990 codes used by the US Census Bureau.

Table 5: SMSA Rankings

1980 Rank	SMSA	1980 Rank	SMSA
1	New York, NY	23	Patterson-Clifton-Passaic, NJ
2	Los Angeles-Long Beach, CA	24	San Diego, CA
3	Chicago, IL	25	Buffalo, NY
4	Philadelphia, PA	26	Miami, FL
5	Detroit, MI	27	Kansas City, MO, KS
6	San Francisco-Oakland, CA	28	Denver, CO
7	Washington, DC, MD, VA	29	San Bernardno-Riverside-Ontario, CA
8	Boston, MA	30	Indianapolis, IN
9	Nassau-Suffolk, NY	31	San Jose, CA
10	Pittsburgh, PA	32	New Orleans, LA
11	St Louis, MO, IL	33	Tampa- St Petersburg, FL
12	Baltimore, MD	34	Portland, OR
13	Cleveland, OH	35	Columbus, OH
14	Houston, TX	36	Rochester, NY
15	Newark, NJ	37	Sacramento, CA
16	Minneapolis-St Paul, MN	38	Birmingham, AL
17	Dallas-Fort Worth, TX	39	Albany-Schenectady-Troy, NY
18	Seattle-Everett, WA	40	Norfolk-Portsmouth, VA
19	Anaheim-Santa Ana-, Garden Grove, CA	41	Akron, OH
		42	Gary-Hammond-East Chicago, IN
20	Milwaukee, WI	43	Greensboro-Winston-Salem- High Point, NC
21	Atlanta, GA		
22	Cincinnati, OH		

Notes: SMSAs consistently available from 1978-2010, ranked by population size in 1980.

Table 7: Union Wages and Outside Options: OLS and 2SLS Estimates

	OLS	2SLS		
	(1)	(2)	(3)	(4)
$\Delta E_{1ct\|1i}$	0.69***	0.27		
	(0.033)	(0.21)		
$\Delta E_{2ct\|1i}$	0.70***	0.16		
	(0.032)	(0.20)		
ΔE_{1cit}			0.13	0.19
			(0.20)	(0.18)
ΔER_{ct}	-0.28**	0.026	0.12	
	(0.14)	(0.23)	(0.25)	
Obs.	2366	2366	2366	2366
R^2	0.45			
Year × Ind.	Yes	Yes	Yes	Yes
Year × City				
Instrument set:		$IV1_{2ct\|1i}$	$IV1_{2ct\|1i}$	$IV1_{2ct\|1i}$
		$IV1_{1ct\|1i}$	$IV1_{1ct\|1i}$	$IV1_{1ct\|1i}$
First-Stage p-Stat.:				
$\Delta E_{2ct\|1i}$		0.000		
$\Delta E_{1ct\|1i}$		0.000		
ΔE_{1cit}			0.000	0.000
Over-id. p-val		.	0.040	0.035

Notes: This table displays results from the estimation of equation (86) via OLS (column 1) and 2SLS (columns 2 - 5). The dependent variable is the decadal change in the regression adjusted average hourly wage of union workers in an industry-city cell, using CPS data from 1980-2019 across 50 industries and 93 cities. Standard errors, in parentheses, are clustered at the city-year level.

Table 8: Union Wages and Outside Options: OLS and 2SLS Estimates No Public Sector

	OLS	2SLS		
	(1)	(2)	(3)	(4)
$\Delta E_{1ct\|1i}$	0.69***	0.57***		
	(0.034)	(0.19)		
$\Delta E_{2ct\|1i}$	0.69***	0.46***		
	(0.033)	(0.17)		
ΔE_{1cit}			0.42***	0.47***
			(0.16)	(0.15)
ΔER_{ct}	-0.20	-0.12	0.0096	
	(0.14)	(0.21)	(0.21)	
Obs.	1977	1977	1977	1977
R^2	0.47			
Year × Ind.	Yes	Yes	Yes	Yes
Year × City				
Instrument set:		$IV1_{2ct\|1i}$	$IV1_{2ct\|1i}$	$IV1_{2ct\|1i}$
		$IV1_{1ct\|1i}$	$IV1_{1ct\|1i}$	$IV1_{1ct\|1i}$
First-Stage p-Stat.:				
$\Delta E_{2ct\|1i}$		0.000		
$\Delta E_{1ct\|1i}$		0.000		
ΔE_{1cit}			0.000	0.000
Over-id. p-val		.	0.021	0.020

Notes: This table displays results from the estimation of equation (86) via OLS (column 1) and 2SLS (columns 2 - 5). The dependent variable is the decadal change in the regression adjusted average hourly wage of union workers in an industry-city cell, using CPS data from 1980-2019 across 50 industries and 93 cities. Standard errors, in parentheses, are clustered at the city-year level.

Table 9: Union Wages and Outside Options: Controlling for Selectivity

	OLS	2SLS		OLS	2SLS		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\Delta E_{1ct\|1i}$	0.69***	0.21		0.66***	0.59***		
	(0.033)	(0.22)		(0.040)	(0.17)		
$\Delta E_{2ct\|1i}$	0.70***	0.090		0.65***	0.47***		
	(0.034)	(0.21)		(0.038)	(0.15)		
$\Delta E_{ct\|1i}$			0.11			0.44***	0.39**
			(0.21)			(0.15)	(0.15)
ΔER_{ct}	-0.27**	0.037	0.12	-0.076	-0.015	0.26	0.25
	(0.14)	(0.25)	(0.26)	(0.18)	(0.33)	(0.31)	(0.32)
ΔP^{ne}_{ict}							0.38
							(0.75)
Obs.	2366	2366	2366	1660	1660	1660	1656
R^2	0.45	0.078	0.068	0.46	0.21	0.21	0.18
Year × Ind.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
P^{ne}_{ict-1}× Ind. × Year							Yes
P^{ne}_{ict-1}× City × Year							Yes
Instrument set:		$IV1_{2ct\|1i}$	$IV1_{2ct\|1i}$		$IV1_{2ct\|1i}$	$IV1_{2ct\|1i}$	$IV1_{2ct\|1i}$
		$IV1_{1ct\|1i}$	$IV1_{1ct\|1i}$		$IV1_{1ct\|1i}$	$IV1_{1ct\|1i}$	$IV1_{1ct\|1i}$
							$\Delta \hat{P}^{ne}_{ict}$
First-Stage p-Stat.:							
$\Delta E_{1ct\|1i}$		0.000			0.000		
$\Delta E_{2ct\|1i}$		0.000			0.000		
$\Delta E_{ct\|1i}$			0.000			0.000	0.000
ΔP^{ne}_{ict}							0.000
Over-id. p-val		.	0.035		.	0.018	
Selection Controls							
P_{ic} Quadratic	Yes	Yes	Yes	No	No	No	Yes
Election Vars.	No	No	No	Yes	Yes	Yes	
Joint Tests:							
p-val	0.69	0.09	0.29	0.02	0.05	0.02	
F-Stat	0.57	2.05	1.24	2.52	2.10	2.46	

Notes: This table displays results from the estimation of equation (86) via OLS (columns 1 and 4) and 2SLS (columns 2,3,5, and 6). The dependent variable is the decadal change in the regression adjusted average hourly wage of union workers in an industry-city cell, using CPS data from 1980-2019 across 50 industries and 93 cities. Standard errors, in parentheses, are clustered at the city-year level.

Table 10: Change in Union Premium and Threat

			Winsorized			
	(1)	(2)	(3)	(4)	(5)	(6)
	All	No Hosp.	All	1980s	1990s	2000s
ΔP^{ne}_{ict}	-0.20	-0.57*	-0.51**	-1.04**	0.39	-1.00***
	(0.13)	(0.31)	(0.23)	(0.47)	(0.28)	(0.31)
Obs.	1546	1451	1546	394	686	466
R^2	0.036	0.041	0.037	0.011	0.002	0.012
Year	Yes	Yes	Yes			

Notes: This table displays the results of regressing the union wage premium on union threat and the *ict* level. Standard errors, in parentheses, are clustered at the city-year level.

Table 11: Non-Union Wages and Outside Options: OLS and 2SLS Estimates No Public Sector

	OLS	2SLS			
	(1)	(2)	(3)	(4)	(5)
$\Delta E_{2ct\|2i}$	1.06***	0.49***			
	(0.022)	(0.12)			
$\Delta E_{1ct\|2i}$	0.99***	0.51***			
	(0.025)	(0.11)			
ΔE_{2ict}			0.55***	0.56***	0.37*
			(0.094)	(0.084)	(0.21)
ΔER_{ct}	-0.37***	0.30*	0.22*		
	(0.077)	(0.16)	(0.13)		
Obs.	8273	8273	8273	8273	8273
R^2	0.68				
Year × Ind.	Yes	Yes	Yes	Yes	Yes
Year × City					Yes
Instrument set:		$IV1_{2ct\|2i}$	$IV1_{2ct\|2i}$	$IV1_{2ct\|2i}$	$IV1_{2c\|2it}$
		$IV1_{1ct\|2i}$	$IV1_{1ct\|2i}$	$IV1_{1ct\|2i}$	$IV1_{1c\|2it}$
First-Stage p-Stat.:					
$\Delta E_{2ct\|2i}$		0.000			
$\Delta E_{1ct\|2i}$		0.000			
ΔE_{2cit}			0.000	0.000	0.000
Over-id. p-val		.	0.287	0.669	0.036

Notes: This table displays results from the estimation of equation (21) via OLS (column 1) and 2SLS (columns 2 - 5). The dependent variable is the decadal change in the regression adjusted average hourly wage of non-union workers in an industry-city cell, using CPS data from 1980-2019 across 50 industries and 93 cities. Standard errors, in parentheses, are clustered at the city-year level.

References

Adao, R., Kolesar, M., & Morales, E. (2020). Shift-share designs: Theory and inference. *The Quarterly Journal of Economics*, 134(4), 1949–2010.

Bassier, I. (2022). *Collective Bargaining and Spillovers in Local Labour Markets*. Discussion Paper 1895, Centre for Economic Performance.

Borusyak, K., Hull, P., & Jaravel, X. (2022). Quasi-experimental, shift-share research designs. *The Review of Economic Studies*, 89, 181–213.

Jaeger, D. A. & Page, M. E. (1996). Degrees matter: New evidence on sheepskin effects in the returns to education. *The Review of Economics and Statistics*, (pp. 733–740).

Lemieux, T. (2006). Increasing residual wage inequality: Composition effects, noisy data, or rising demand for skill? *American Economic Review*, 96(3), 461–498.

Park, J. H. (1994). *Estimation of sheepskin effects and returns to schooling using the old and the new CPS measures of educational attainment*. Number 338. Industrial Relations Section, Princeton University.

Taschereau-Dumouchel, M. (2020). The Union Threat. *The Review of Economic Studies*, 87(6), 2859–2892.

Tschopp, J. (2017). Wage formation: Towards isolating search and bargaining effects from the marginal product. *The Economic Journal*, 127(603), 1693–1729.