Population Changes and the Political Economy of Local Public Finance*

Ryan Booth[†] Emanuel Vespa[‡]

September 1, 2009

Abstract

We estimate the frictions that city governments face when adjusting public good expenditures in response to population changes and investigate the role that public workers play in this process. Unlike private sector workers, who can unionize and bargain collectively with their employers, public sector workers can also influence the reelection of the policy makers that determine the city's terms of employment. We develop a theoretical model that includes both of these channels and show that population frictions are primarily caused by the political economy channel, with collective bargaining acting to intensify the effect. Using a panel of 909 US cities, we test the implications of the model for five types of local public goods. We find that services whose workers are traditionally organized, such as police, exhibit population frictions, whereas other services, such as administrators, do not. In addition, we estimate the effect of the collective bargaining channel and show that the data verifies several of the model's predictions.

1 Introduction

It is a well documented fact that many cities in the United States are experiencing dramatic population changes, and that growth rates exhibit considerable heterogeneity by geographic region. For instance, between 1970 and 2000, the median growth rate of

*We thank Alessandro Lizzeri for insightful guidance during this project. We also thank Saki Bigio, Oeindrila Dube, Raquel Fernández, Guillaume Fréchette, Alessandro Gavazza, Sanford Gordon, Catherine Hafer, Debraj Ray, Michael Richter, Nicola Persico, Shing-Yi Wang and participants of the NYU Third Year Paper Seminar.

[†]Department of Economics, New York University, ryan.booth@nyu.edu

[‡]Department of Economics, New York University, emanuel.vespa@nyu.edu

"Rust Belt" cities in the states of Michigan, Ohio and Pennsylvania was -14.8%, while cities in the Southwest states of Arizona, Nevada and Texas grew at a rate of 85.2%. Nevertheless, in the same Rust Belt states, per capita expenditures on locally provided public goods, such as fire, increased 24%, while simultaneously increasing only 7.4% in the Southwest.

Given these stylized facts, a natural question is, how do such population dynamics influence a city's expenditures on local public goods? Both the political economy literature and common sense say that public policies tend to persist once implemented, potentially explaining the dynamic within declining cities. Growing cities, on the other hand, may spend less per capita as such frictions are not as prevalent when a city's population is increasing. While there are many potential stories as to why public goods may exhibit such a relationship with population dynamics, one classic explanation is that public sector workers attempt to influence policy through legal or political means.

In this paper, we estimate the frictions that city governments face when adjusting public good expenditures as population changes, and then investigate the role that public workers play in this process. Unlike private sector workers, who can unionize and bargain collectively with their employers, public sector workers can also influence the reelection of the policy makers that determine the city's terms of employment, as discussed by Moe (2006). We develop a theoretical model in which unionized public workers can influence policy via these two channels. It is shown that population frictions are primarily caused by the political economy channel, and that collective bargaining acts to intensify these effects. Using a panel of 909 US cities, we test the implications of the theoretical model for five types of local public goods. We find that services that are traditionally highly organized and, historically more politically active, such as police and fire departments, exhibit population frictions, whereas other services, such as administrators, do not. In addition, we provide an estimate of the influence of the latter channel and show that the data verifies several of the model's predictions.

Consider two cities that are identical with respect to preferences, population and technology, but differ in the rate of population change. In a frictionless world, where a social planner is unconstrained in hiring the optimal amount of workers that produce the public good, these two cities should offer the same level of the service. In reality however, policy makers may be faced with constraints that are related to population dynamics. For instance, restrictions on capital, uncertainty of future population dynamics and contracts may constrain the policy maker from selecting the optimal level of public good provision. Such constraints will be reflected in a relationship between expenditures in public goods and the rate of population change. Persistence is a special case of this phenomenon such that when population shrinks, expenditures remain higher than optimal.

There is a literature in political economy that studies policy persistence. One particularly relevant paper on special interests is Coate and Morris (1999). The environment includes voters, politicians and a special interest group. Each period, the politician in power decides whether or not to initiate a policy that is costly to the voters but benefits the special interest group. The interest group has the ability to influence the politician in a principal-agent environment, by offering a monetary transfer to the politician in exchange for policy concessions.

This paper will focus on how public sector workers that produce public goods contribute to the persistence observed in city public finances. Since Niskanen (1975) bureaucrats have been modeled as self-interested agents having preferences that differ from those of the constituency they serve. In Romer and Rosenthal (1979), for instance, a model is presented in which bureaucrats strategically attempt to increase their budget allocation through political means. Another modeling approach when thinking about the interaction between bureaucracy and government is through the lens of collective bargaining, as presented in McDonald and Solow (1981). In this paper the authors model a labor union that collectively bargains with a firm and that has preferences over the wages that workers receive and the fraction of current union members that will be reemployed. Empirical studies, such as those presented in Freeman (1986) take this perspective when studying the influence of unionization on public sector wages and employment.

In this paper, we merge elements from both of these approaches to examine how bargaining and political influence interact with population dynamics. In the spirit of Coate and Morris (1999) we model public workers as having the ability to support particular candidates in an upcoming election. Moe (2006) points out that the unions representing bureaucracies also have significant political power with which they can influence election outcomes and, consequently, those controlling the public finances. In the US, 43% of city-employed bureaucrats are dues-paying members of their respective union (teachers, health care workers, police, etc). Empirical evidence suggests that bureaucratic unions do indeed wield considerable political influence in local elections, superior even to political parties: Using data from school board elections in California, Moe (2006) shows that candidates supported by the teachers union win 50% more often than candidates that fail to gain such support.

Our model includes citizens, a politician and a labor union in a setup where population is stochastic. Each period the incumbent politician sets public employment and wages, subject to satisfying an employment contract collectively negotiated with the union. At the end of the period the politician is subject to reelection and the probability of remaining in office depends on both the utility that citizen's derive from the policy, as well as campaign support provided by the union. The union has preferences over wages and reemployment and can offer campaign support to the politician at a cost. The model generates population frictions in the sense that, when population declines, per-capita expenditures on the public good increase, so that the public welfare for the remaining citizens falls as the per-capita tax revenue increases. Moreover, the result shows that bargaining power exacerbates this friction when the population declines. On the other hand, when the population growth rate is sufficiently high, public utility welfare increases as employment and wages approach the planners optimal.

In order to test for population frictions and the main comparative statics generated by the model, we create a panel of 909 cities by merging information on city employment, finance and demographics for years between 1970 and 2000. Several authors have used portions of this data for a variety of studies,[1] some of which have looked at the effects of unionization on public finance.[2] Unlike the previous literature, however, we study the interaction between population change and public finance, with a particular focus on the dynamic influence of public workers.

We use the dataset to estimate the effect of population change on local public good expenditures for five different services. Our estimates show a negative and significant effect of population change on the wage expenditures on police and fire services, and a statistically non-significant effect on highway, parks and administration wage expenditures. We argue that these results are consistent with the model's prediction that frictions caused by unions are primarily driven by political means, as police and fire are highly organized and historically more politically active. Moreover, we study the interaction between population change and bargaining power. Unionization has a level effect on per-capita expenditures consistent with some of the previous literature, as well as an influence on the marginal effect of population change in the case of fire services.

One natural concern with our approach is related to the literature on Tiebout-type sorting, according to which expenditures on public goods influence population dynamics. In other words, citizens would make their decisions to move by comparing the level of public goods provided in neighboring cities. We discuss such concerns at length in the

[1]See for instance, Alesina, Baqir and Easterly (1999), Baqir (2002) and **?**.

[2]Hoxby (1996) studies the effect of teacher unions on wages and employment, as well as school quality. She uses a panel from 1970 to 1990 and a differences in differences approach. Her findings suggest that unionization causes an increase in school expenditures, but nevertheless a decrease in school quality. Using a different measure of unionization for a set of three US States, Lovenheim (2009) also studies teacher unions, but finds results that are contradictory to Hoxby (1996). His findings state that unionization does not influence teachers' wages, but increases the number of teachers in a district. Nevertheless, there exists some evidence that school quality improves when a district's teachers are unionized.

empirical section of the paper, where we provide estimates of a spatial model that controls for the vector of public good expenditures in other cities. The main results are robust to this modification.

The rest of the paper is organized as follows: The following section introduces the theoretical model. Section 3 discusses the main empirical implications of the model. Section 4 describes the data, and then Section 5 discusses the empirical strategy and presents the results. Finally, section 6 concludes.

2 Theoretical Model

A brief overview of the game is as follows. There are three types of players: Citizens, an incumbent politician and a labor union representing workers. At the beginning of period t, the measure of citizens, z_t, and the status quo number of workers represented by the union, L_{t-1}, are publicly observed. Each period, the politician has discretion over selecting the time t number of public employees, $L_t \in \mathcal{L} \subseteq \Re_+$, as well as the wage the workers receive, $w_t \in \mathcal{W} := [\underline{w}, \infty)$, where $\underline{w}$ represents the workers' reservation wage. The L_t workers produce a public good according to a linear production technology, so that time t output equals L_t.[3] As is detailed below, the union can attempt to influence (w_t, L_t) via two channels: First, the union can negotiate a collectively bargained contract with the politician. Second, if the union is politically organized, it can offer the politician campaign support s_t in exchange for policy concessions, as the probability that the politician wins reelection is increasing in s_t.

2.1 Primitives

2.1.1 Citizens

The measure of citizens z_t is stochastic and drawn each period from a continuously differentiable cdf G with support $\mathcal{Z} := [\underline{z}, \infty) \subset \Re_{++}$ and pdf g.[4] A representative citizen has preferences over (w_t, l_t), where $l_t = L_t/z_t$ is the per capita production of the

[3] The results of the model hold if a general Cobb-Douglas production function is assumed instead. Here, the additional inputs can be, say, capital or a non-unionized labor input.

[4] For technical reasons, we assume that $\sup \mathcal{L} < \inf \mathcal{Z}$, so that $L/z < 1$ for all $(z, L) \in \mathcal{Z} \times \mathcal{L}$. In equilibrium, this constraint will never bind for $\underline{w}$ large enough. This is natural if we interpret L and z literally as the number of workers and population size.

public good at time t. Preferences for the citizen are represented by the utility function

$$\overline{u}(w_t, l_t) = \frac{l_t^\sigma}{\sigma} - w_t l_t$$

where $\sigma \in (0,1)$ implies concavity in l_t. It is implicitly assumed that the city maintains a balanced budget, as the representative citizen pays $w_t l_t$ in taxes.

2.1.2 The Politician

The incumbent politician is office motivated and faces reelection at the end of each period. As in Grossman and Helpman (1996), elections are not modeled explicitly, but it is assumed that the politician's utility is increasing and additively separable in the representative citizen's time t welfare, $\overline{u}(w_t, l_t)$, and a function of the union's per worker campaign support, s_t. The politician's preferences are assumed to be represented by

$$p(w_t, L_t; z_t, L_{t-1}, s_t) = \overline{u}(w_t, L_t/z_t) + \left(\frac{L_{t-1}s_t}{z_t}\right)^\gamma$$

where $\gamma \in (0,1)$. Since s_t represents campaign support per worker, then $(L_{t-1}s_t)/z_t$ is the politician's per capita support. Consequently, the politician's benefit of such support is increasing and concave.

Each period, the politician's choice of (w_t, L_t) is constrained by a collectively bargained contract $(\tilde{w}, \tilde{\eta}) \in [\underline{w}, \infty) \times [0,1]$, which stipulates that $(w_t, L_t) \geq (\tilde{w}, \tilde{\eta}L_{t-1})$. The intuition behind $(\tilde{w}, \tilde{\eta})$ is that of a wage and rehiring requirement that arises from cooperative bargaining between the politician and the union. We do not model the bargaining procedure, as our qualitative results would remain unaffected. For algebraic simplicity, we parameterize $(\tilde{w}, \tilde{\eta})$ by the union's bargaining power, $B \in \mathcal{B} \subset \Re$, and simply impose that $\partial\tilde{w}/\partial B > 0$ and $\partial\tilde{\eta}/\partial B > 0$. A detailed defense of this modeling assumption is given in the discussion below.

2.1.3 The Union

At time t, the union has preferences over w_t, s_t and the fraction of L_{t-1} that are rehired. The politician is assumed to first rehire the current union members, so that the fraction of L_{t-1} that are rehired is

$$\eta(L_t; L_{t-1}) = \min\left\{\frac{L_t}{L_{t-1}}, 1\right\} \tag{1}$$

It is assumed that union members incur a cost to supply s_t, so that a representative union member's preferences are

$$v(w_t, L_t, s_t; L_{t-1}) = \left[\eta(L_t; L_{t-1})w_t + \Big(1 - \eta(L_t; L_{t-1})\Big)\underline{w}\right] - \frac{\gamma}{\kappa}s_t^{\kappa}$$

where $\kappa > 1$, so that cost is convex, and $\eta : \mathcal{L}^2 \to [0,1]$ is as defined in (1). The interpretation is that the representative union member faces an $\eta(\cdot)$ probability of being rehired and receiving wage w_t, and a $1 - \eta(\cdot)$ probability of being fired and receiving the reservation, $\underline{w}$. Note that $v(\cdot)$ depends on the wages received by the current L_{t-1} union members, and *not new employees hired during period* t. This assumption is made to correspond to anecdotal evidence that unions cater to their current membership, thus ignoring potential benefits enjoyed by future union members.

In addition to the contract $(\tilde{w}, \tilde{\eta})$ discussed above, the union can influence the politician by offering s_t prior to the time t election. The union's strategy is a function $(w_t^*, L_t^*, s_t^*) : \mathcal{Z} \times \mathcal{L} \times \mathcal{B} \to \mathcal{W} \times \mathcal{L} \times \Re_+$ that maps the current state, (z_t, L_{t-1}), and bargaining power, B, into a quid pro quo campaign offer s_t^* in exchange for a "renegotiated" constraint on the politician's decision, $(w_t, L_t) \geq (w_t^*, L_t^*)$. The politician has the ability to accept or reject this offer.

2.1.4 Notational Convention

The following notation will be followed throughout the paper:

- Variables and functions that are covered by tildes, such as $(\tilde{w}, \tilde{\eta})$, will represent objects related to the "bargained" contract
- Variables and functions denoted by stars, such as (w_t^*, L_t^*, s_t^*), will represent the union's time t equilibrium strategy
- Variables and functions without tildes or stars represent general objects

2.1.5 Timing and Equilibrium Concept

For clarification, the timing in period t is:

1. The state (z_t, L_{t-1}) and contract $(\tilde{w}, \tilde{\eta})$ are publicly observed
2. The union plays (w_t^*, L_t^*, s_t^*)

3. The politician chooses (w_t, L_t) and period payoffs are realized

Here, we assume that all players are myopic, and derive several results that we test using data. Given this, our solution concept is Subgame Perfect Nash.

In Appendix B, we develop a dynamic version of the two period game in which the union is forward-looking and internalizes the effect that L_0 has on the path of play. Several additional insights resulting from the dynamic model are discussed, and proofs are given.

2.2 Discussion of Modeling Assumptions

2.2.1 Voting

Empirical evidence suggests that public worker unions influence local elections by offering campaign support for candidates. For instance, Moe (2006) studies school board elections in California and shows that, in general, all candidates receive a non-trivial share of votes. Moreover, the median vote difference separating the lowest vote-getting winner and highest vote-getting loser is only 3%. As candidates running for local office typically invest limited amounts of resources in their campaigns, this suggests that voters have relatively little information when heading to the polls. Nevertheless, candidates supported by the teacher union win 50% more often than candidates that fail to gain such support, suggesting that voters may interpret the support of teachers union as a signal of candidate quality. In city council and mayoral elections, some groups of public workers, such as police and fire fighters, are often active in supporting political campaigns.

The reduced form voting assumption is introduced to capture the trade-off incumbent politicians face between enacting the public's optimal policy and offering concessions to unions in order to gain reelection support. Similar modeling strategies have been employed by several papers, such as Grossman and Helpman (1996), Baron (1994) and Prat (2002). These papers offer micro-foundations to justify this approach. Grossman and Helpman (1996), for instance, assume that there are two types of voters: Those who are informed about policy platforms of two candidates and vote based on this information, and those who are uninformed and are influenced by campaign spending. In Prat (2002), on the other hand, all voters are fully rational but have imperfect information about candidate quality. Special interest groups observe these dimensions, however, and offer campaign funds that, in equilibrium, serve as a quality signal for the voters.

2.2.2 Bargaining Power

The contract $(\tilde{w}, \tilde{\eta})$ is introduced to allow for analysis of how a union's bargaining power and ability to influence electoral outcomes interact over time. Such bargaining power varies widely across cities: State laws in the Northeast and Midwest typically grant unions strong rights in the collective bargaining process, whereas other states greatly restrict public sector bargaining, or even ban collective bargaining, such as in North Carolina.

The connection between bargaining power and $\tilde{w}$ is clear: When negotiating a contract, the union and government explicitly agree on a wage schedule. In theory, higher bargaining power should translate into higher wages, thus justifying the assumption that $\partial\tilde{w}/\partial B > 0$. The interpretation for $\tilde{\eta}$, however, is more subtle. Union-negotiated contracts typically stipulate the order in which workers can be laid off, with workers of longer tenure enjoying greater job security. Moreover, unions with stronger bargaining power may use tactics such as striking or lawsuits to restrict the number of layoffs, with the effectiveness of such measures increasing in union strength. To introduce this employment friction, we take $\tilde{\eta}$ to be the fraction of current workers that the government cannot fire, with $\partial\tilde{\eta}/\partial B > 0$.

3 Model Analysis

3.1 Social Planner

As a baseline, we begin by studying the policy chosen by a planner that maximizes citizen welfare. First, we assume that the planner is fully unconstrained, thus implicitly imposing that $(\tilde{w}, \tilde{\eta}) = (\underline{w}, 0)$, then each period it solves

$$\max_{(w,L)\in\mathcal{W}\times\mathcal{L}} \frac{1}{\sigma}\left(\frac{L}{z}\right)^{\sigma} - \frac{wL}{z}$$

Comment 1. *For any state (z, L_{-1}), the unconstrained planner sets $w = \underline{w}$ and $l = \underline{w}^{-\frac{1}{1-\sigma}}$. Per capita expenditures are $wl = \underline{w}^{-\frac{\sigma}{1-\sigma}}$.*

When the planner is fully unconstrained, the per capita number of workers l is constant across population realizations, and consequently, per capita expenditures are constant as well. Next, we consider a planner that is constrained to respect the union's contract, so that $(w, L) \geq (\tilde{w}, \tilde{\eta}L_{-1})$.

Comment 2. *Let* (z, L_{-1}) *denote the state. The constrained planner sets* $w = \tilde{w}$ *and*

$$l = \begin{cases} \tilde{w}^{-\frac{1}{1-\sigma}} & \text{if } \tilde{\eta} L_{-1} \leq z\tilde{w}^{-\frac{1}{1-\sigma}} \\ \frac{\tilde{\eta} L_{-1}}{z} & \text{otherwise} \end{cases}$$

Per capita expenditures are

$$wl = \begin{cases} \tilde{w}^{-\frac{\sigma}{1-\sigma}} & \text{if } \tilde{\eta} L_{-1} \leq z\tilde{w}^{-\frac{1}{1-\sigma}} \\ \frac{\tilde{w}\tilde{\eta} L_{-1}}{z} & \text{otherwise} \end{cases}$$

There are several points to mention here. First, note that when $\tilde{\eta} L_{-1} > z\tilde{w}^{-\frac{1}{1-\sigma}}$, the rehiring constraint $\tilde{\eta}$ binds for the planner. This occurs when z is sufficiently small relative to L_{-1}, causing employment and per capita expenditures to be larger than in the absence of $\tilde{\eta}$.

Second, note that per capita expenditures can be smaller under the constrained planner, as $\tilde{w}^{-\frac{\sigma}{1-\sigma}} < \underline{w}^{-\frac{\sigma}{1-\sigma}}$. Given the representative citizen's preferences and the linear production technology, citizen demand for the public good is elastic. Consequently, if w increases, then the planner's choice of l should result in smaller per capita expenditures (so long as the $\tilde{\eta}$ rehiring constraint does not bind).

Finally, note that the constrained planner's solution corresponds to the politician's optimal strategy if the union is unable or unwilling to provide campaign support. In the absence of a union offer (w^*, L^*, s^*), the politician simply maximizes the public's utility in order to maximize the probability of reelection.

3.2 Equilibrium

We solve the static version of the game via backward induction. Given the game is static, we drop the t subscript and define the status quo union size as L_{-1}. If the politician chooses to satisfy the union's campaign offer (w^*, L^*, s^*), then its problem is

$$u(w^*, L^*; z) := \max_{(w,L) \geq (w^*, L^*)} \overline{u}(w, L/z)$$

where $u(w^*, L^*; z)$ is the representative citizen's indirect utility, given population z and constraint $(w, L) \geq (w^*, L^*)$. The politician thus maximizes public welfare, subject to satisfying the union's campaign support demand. At the optimum, $w = w^*$ and

$$l = \begin{cases} (w^*)^{-\frac{1}{1-\sigma}} & \text{if } L^* \leq z(w^*)^{-\frac{1}{1-\sigma}} \\ \frac{L^*}{z} & \text{otherwise} \end{cases}$$

Substituting the optimal policy into $u(\cdot)$, we have

$$u(w^*, L^*; z) = \begin{cases} \dfrac{1-\sigma}{\sigma}(w^*)^{-\frac{\sigma}{1-\sigma}} & \text{if } L^* \leq z(w^*)^{-\frac{1}{1-\sigma}} \quad (2a) \\ \dfrac{1}{\sigma}\left(\dfrac{L^*}{z}\right)^{\sigma} - \dfrac{w^* L^*}{z} & \text{otherwise} \quad (2b) \end{cases}$$

For brevity, we define $\tilde{u}^B(z, L_{-1}) := u(\tilde{w}, \tilde{\eta}L_{-1}; z)$, the representative citizen's utility when the state is (z, L_{-1}), the union's bargaining power is B and the politician adheres to the contract $(\tilde{w}, \tilde{\eta})$. Note that this corresponds to the solution to the constrained planner's problem discussed above.

Regardless of (w^*, L^*, s^*), the politician can always achieve a payoff of $\tilde{u}^B(z, L)$. Consequently, the politician's participation constraint is

$$\tilde{u}^B(z, L) \quad \leq \quad u(w^*, L^*; z) + \left(\frac{L_{-1}s^*}{z}\right)^{\gamma} \tag{3}$$

for a union offer of (w^*, L^*, s^*), and thus the union solves

$$\max_{w,L,s} v(w, L, s; L_{-1})$$

subject to (3). One subtlety of (3) is the piecewise functional form of $\tilde{u}^B(\cdot)$ and $u(\cdot)$, which can take the "unconstrained" [equation (2a)] or "constrained" form [equation (2b)] depending on z, L_{-1} and L^* relative to the other variables.

Proposition 1 defines the equilibrium of the stage game and is instrumental in defining the comparative static results that we map to the data. The Proposition partitions the state space $\mathcal{Z} \times \mathcal{L}$ into three sets, $\mathcal{S}^{\mathcal{H}}, \mathcal{S}^{\mathcal{M}}$ and $\mathcal{S}^{\mathcal{L}}$ (see Figure 1). In the "high" region, $\mathcal{S}^{\mathcal{H}}$, status quo employment L_{-1} is high relative to the population realization, z. When $(z, L_{-1}) \in \mathcal{S}^{\mathcal{H}}$, employees are fired in equilibrium so that $L < L_{-1}$. On the other hand, the "low" region, $\mathcal{S}^{\mathcal{L}}$, corresponds to cases in which L_{-1} is small relative to the z. When $(z, L_{-1}) \in \mathcal{S}^{\mathcal{L}}$, new employees are hired so that $L > L_{-1}$ in equilibrium. Finally, the "medium" region, $\mathcal{S}^{\mathcal{M}}$, lies between $\mathcal{S}^{\mathcal{H}}$ and $\mathcal{S}^{\mathcal{L}}$ and is characterized by a maintenance of union size, $L = L_{-1}$.

Proposition 1. *There exist three nonempty sets $\mathcal{S}^{\mathcal{H}}, \mathcal{S}^{\mathcal{M}}, \mathcal{S}^{\mathcal{L}} \subset \mathcal{L} \times \mathcal{Z}$ that constitute a partition of $\mathcal{L} \times \mathcal{Z}$, such that*

1. *if $(z, L_{-1}) \in \mathcal{S}^{\mathcal{H}}$, then $L < L_{-1}$ and $\partial w / \partial L_{-1} > 0$ in equilibrium*
2. *if $(z, L_{-1}) \in \mathcal{S}^{\mathcal{M}}$, then $L = L_{-1}$ and $\partial w / \partial L_{-1}$ can either be positive or negative in equilibrium*

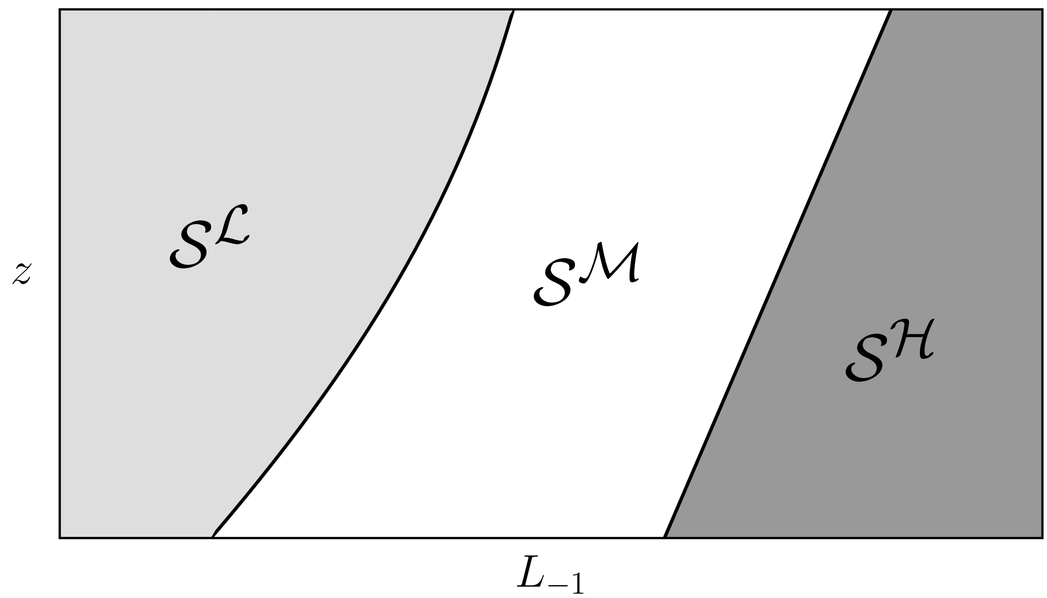

Figure 1: The Partition of $\mathcal{Z} \times \mathcal{L}$

3. *if* $(z, L_{-1}) \in \mathcal{S}^{\mathcal{L}}$*, then* $L > L_{-1}$ *and* $\partial w/\partial L_{-1} > 0$ *in equilibrium*

Proof. All proofs are located in the appendix. □

In addition to describing employment dynamic, Proposition 1 also provides details on the relationship between w and L_{-1}. When $(z, L_{-1}) \in \mathcal{S}^{\mathcal{H}} \cup \mathcal{S}^{\mathcal{L}}$, then $\eta(\cdot)$ is independent of L_{-1}: When L_{-1} is "low," the entire union is reemployed (and additional workers are hired), whereas in the "high" case, L is determined solely by z due to the linearity of w in both the union and citizen objective function. Consequently, w is increasing in L_{-1} via two effects: First, campaign support s is more valuable to the politician when the L_{-1} is larger, thus allowing bigger unions to exert more influence over the politician. And second, the public's default utility under the contract, $\tilde{u}^B(z, L_{-1})$, is (weakly) decreasing in L_{-1}, in effect lowering the politician's outside option from ignoring the union's campaign support offer.

When $(z, L_{-1}) \in \mathcal{S}^{\mathcal{M}}$, on the other hand, w is increasing in L_{-1} in the two ways described above. Nevertheless, since $L = L_{-1}$ in equilibrium, a larger L_{-1} also implies that the citizenry must pay w to more workers, which influences w to decrease in L_{-1}. In general, it is indeterminate which effect will dominate: There exist states $(z, L_{-1}) \in \mathcal{S}^{\mathcal{M}}$ for which $\partial w/\partial L_{-1} > 0$ and others for which $\partial w/\partial L_{-1} < 0$. A full characterization of these strategies is detailed in the appendix.

3.3 Comparative Statics

We now derives three comparative static results that we will test in the empirical section. These results extend Proposition 1 to examine how per capita expenditures on the public good are related to the current population level, z_t, and how these expenditures change with regards to the previous period's population level, z_{t-1}. We thus analyze the two period, repeated version of the static game so that $t \in \{0, 1\}$.

The first result investigates $\partial(w_1 l_1)/\partial z_1$ and explains how this comparative static depends on the size of z_1 relative to the previous population level, z_0.

Result 1. *There exists a population cutoff $z^S \geq z_0$ such that*

$$\frac{\partial(w_1 l_1)}{\partial z_1} \begin{cases} < 0 & \text{if } z_1 < z^S \\ > 0 & \text{if } z_1 > z^S \end{cases}$$

Moreover, $\lim_{z_1 \to \infty} w_1 l_1 = \tilde{w}^{-\frac{\sigma}{1-\sigma}}$.

Result 1 says several things of interest. First, if $z_1 < z_0$, then in equilibrium, $\partial(w_1 l_1)/\partial z_1 < 0$. In words, this implies that as a city declines in population, its per capita expenditures increase. The intuition is that as z_1 falls further below z_0, the union becomes relatively more influential, thus allowing the union to secure a higher wage. Given that citizen preferences are elastic in the public good, an increase in per capita expenditures implies that in equilibrium, the representative citizen's utility *declines as the city's population falls.*

When $z_1 > z_0$, on the other hand, then $\partial(w_1 l_1)/\partial z_1$ can either be positive or negative, depending on the magnitude of growth. The population level z^S (which depends on L_0) determines the cutoff: When $z_1 > z^S$, then $(z_1, L_0) \in \mathcal{S}^{\mathcal{L}}$, and thus from Proposition 1, we know that $L_1 > L_0$. The reasoning is that when z_1 is sufficiently high, the union becomes relatively less powerful and the public demands a higher level of L_1. In equilibrium, the union's wage demand, w_1^*, is sufficiently low relative to z_1 so that the politician is unconstrained in setting $l_1 = (w_1^*)^{-\frac{1}{1-\sigma}}$. Moreover, w_1^* is *decreasing* in z_1. Given that the public's demand for the good is elastic, this implies that $\partial(w_1 l_1)/\partial z_1 > 0$ when $z_1 > z^S$. The union's wage demand, however, is bounded below by $\tilde{w}$, and in the limit, as $z_1 \to \infty$, we have $w_1^* = \tilde{w}$.

While the comparative static can take either sign depending on the relative decline or growth of z_1, note that $\partial(w_1 l_1)/\partial z_1 > 0$ only when the growth rate z_1/z_0 is sufficiently larger than 1. Consequently, $\partial(w_1 l_1)/\partial z_1 < 0$ for the majority of population realizations z_1. Moreover, we see that as $z_1 \to \infty$, then $\partial(w_1 l_1)/\partial z_1 \to 0$, whereas when z_1 decreases, then $\partial(w_1 l_1)/\partial z_1$ decreases without bound.

We now turn to the comparative static between per capita expenditures and bargaining power.

Result 2. *There exists a population cutoff $z^S \geq z_0$ such that*

$$\frac{\partial(w_1 l_1)}{\partial B} \begin{cases} > 0 & \text{if } z_1 < z^S \\ < 0 & \text{if } z_1 > z^S \end{cases}$$

Not surprisingly, Result 2 shows that per capita expenditures are increasing in B when $z_1 < z_0$, which occurs because the union's wage demand is increasing in B. On the other hand, when $z_1 > z^S$, then $\partial(w_1 l_1)/\partial B < 0$. As in Result 1, when $z_1 > z^S$, then $(z_1, L_0) \in \mathcal{S}^{\mathcal{L}}$, which again implies that $w_1 l_1 = (w_1^*)^{-\frac{\sigma}{1-\sigma}}$. Given that w_1^* is *increasing* in B, then the equilibrium $(w_1 l_1)$ is *decreasing* in B.

Finally, Result 3 examines the cross partial of $(w_1 l_1)$ with respect to z_1 and B.

Result 3. *For any z_0 and z_l,*

$$\frac{\partial^2(w_1 l_1)}{\partial B \partial z_1} \leq 0$$

Moreover, there exists a population cutoff $z^S \geq z_0$ such that the inequality is strict when $z_1 > z^S$.

This result shows that for any initial population z_0, the friction outlined in Result 1 is marginally more detrimental to the public, the higher the union's bargaining power. The weak inequality occurs because $\tilde{u}^B(z, L)$ enters linearly into $(w_1 l_1)$ when $z_1 < z^S$. When z_1 is small enough relative to L_0 so that the $\tilde{\eta}$ rehiring constraint constrains the politician's outside option, then higher B will translate into higher wages for the union. On the other hand, if z_1 is high enough so that $\tilde{u}^B(z, L)$ takes form (2a), then $\partial\tilde{u}^B/\partial z_1 = 0$, which implies that $\partial^2(w_1 l_1)/(\partial B \partial z_1) = 0$.

On the other hand, when $z_1 > z^S$, the union's wage demand is increasing in B. Since the politician chooses the optimal public employment, conditional on wage demand w_1^*, and the public demand for the good is elastic, this implies that the marginal per capita expenditure is smaller when B is higher.

4 Empirics

4.1 Data Overview

Our data is compiled from several sources, with the majority of the information coming from the US Census of Governments (CoG) and the County and City Data Book (CCDB). Since 1957, the CoG has been conducted every five years by the US Census Bureau. This data reports detailed finance and employment information for all government bodies within the US, such as states, cities and counties. For cities, information is disaggregated by service, so that the number of employees, wage outlay and total expenditures are reported separately for the police, fire and administrative departments, among others.[5] We extract information on the number of full time employees, the wage bill paid to full time employees and operational expenditures for five services. Both the wage bill and operational expenditures are adjusted to reflect annual expenditures, reported in 2005 dollars. A brief description of each service, as summarized by the Census Bureau, is as follows:

Administrative: Handling of city finance, government-wide planning and legal matters.

Fire: Fire protection and prevention, as well as ambulance and rescue services.

Highway: Maintenance of roads, bridges and street lighting.

Parks and Recreation: Maintenance of public parks, museums, swimming pools and other recreational services.

Police: Enforcement of law and order.

These five services were chosen because the vast majority of cities in our sample directly produce these goods by hiring the capital and labor.[6] A small fraction of the cities in our sample, particularly in the West, outsource these services to either another governmental body or the private sector. In several medium-size cities in California, for instance, fire protection is provided by county fire departments affiliated with the California Department of Forestry and Fire Protection.[7]

[5] In general, data for public education is not classified under cities. Public schools are typically governed by an elected school board that operates independently from the municipal government.

[6] In our sample, all cities provide administrative services; both highway and police services are provided by 98% of cities; 95% of cities provide parks service; and 94% provide fire protection.

[7] Levin, Tadelis and Building (2007) study the outsourcing of public goods by municipal governments and they provide evidence from city managers that several of the services we are studying are difficult to outsource due to monitoring problems. Our sample contains only a few cases in which a city once provided a services in-house, and then decided to outsource the production (and vice versa).

From 1972 until 1987, the CoG also contains data on labor relations between the government and public workers. The information includes union membership in each service, the number of labor contracts negotiated between the city and its departments, and whether the city engages in collective bargaining. We also include this information in our dataset and utilize it to produce an indicator of whether a particular service, say police, was unionized in city i at time t. For the five services we're analyzing, union membership information is only available for the fire, park and police departments. Consequently, analysis on bargaining power is limited to these three services.

Our second main data source, the CCDB, is also produced by the Census Bureau since includes demographic information for cities and counties with population $25,000$ and more. The demographic measures, such as the city's population, racial profile, education and median household income, are derived from the decennial US Population Census, and consequently, demographic information is only observed every 10 years.[8] In addition to demographic measures, we also extract information on the city's land size and the number of serious crimes committed per capita, where "serious" entails both violent crimes and theft.

To create the panel, we merge the CCDB population and demographic information from 1960, 1970, 1980, 1990 and 2000 with the CoG finance and employment data from 1972, 1982, 1992 and 2002 by city. We choose 1972 as the first year for the finance and employment data because the CoG altered its reporting format in this year. The CoG information from years ending in "7" is omitted because the corresponding city demographic information is not available. While $1,302$ cities appear at least once in both the CCDB and CoG, demographic information for many cities appear only once or twice, as city demographic information is only observed if population is greater than $25,000$.[9] Consequently, we restrict the sample to cities for which the demographic information appears in at least three consecutive decades (1970 to 1990, or 1980 to 2000), which reduces the sample to 909 cities.[10]

One point worth mentioning concerns the redistricting of cities. Over the course of the panel, several cities merge with neighboring municipalities to form a larger government. For instance, Athens, Georgia merged with the entire Clarke County in 1991 to form one jurisdiction. As a result, the reported population of Athens jumped from $45,734$ in 1990

[8]The 2000 CCDB does not include education, income or poverty information. We added this information directly from the 2000 US Population Census via the website `http://factfinder.census.gov/`.

[9]For instance, the population of Key West, Florida, was $29,312$ in 1970, $24,382$ in 1980, $24,832$ in 1990 and $25,478$ in 2000.

[10]This should not introduce sample selection bias, as the reduction is based on population, which is an observable that is included in the model. Dropped cities are distributed across the US, suggesting that the reason for exclusion should not be caused by unobservable variables correlated with the covariates.

to $100,266$ in 2000. To control for jurisdiction changes, we construct an indicator variable $boundary_{it}$ that equals 1 if the land area under the city's control more than doubled in the previous decade. This methodology identifies 109 cities that experienced boundary changes.

In addition to the two sources described above, we also used data from the National Oceanic and Atmospheric Administration to compute distances between cities.[11] Using the haversine formula for computing distances, we use the latitude and longitude information in the NOAA dataset to compute a matrix of distances between the 909 cities in our panel.

Finally, we explain how we use create a measure of collective bargaining and a categorical measure of population change, that will be used in the next section. We have built a categorical measure of collective bargaining (B) based on three conditions by combining information from two variables following Hoxby (1996). We observe the fraction of employees that are members of the union at each point in time and we also know the number of contracts that are outcomes of collective bargaining between the city and workers, but we cannot match signed contracts to specific unions. We consider that a group of workers is bargaining collectively with the city if, first, the fraction of employees that are members of the union corresponds to at least 50% of the total size of the full time workforce. Our second requirement is that the city has bargained collectively at least once. Using a measure that satisfies these two conditions results in some instances where services transition from having to not having bargaining power. Lovenheim (2009), focussing on teacher's unions, discusses possible sources of measurement error in the variable capturing the number of union members and he presents evidence that typically once a union starts bargaining collectively it remains doing so in the future. Thus, in an attempt to reduce measurement error, we impose the condition that once a service bargains collectively it continues to do so in the future.

We also create a categorical measure of population change (C), that classifies cities as experiencing decline, a normal expansion or significant growth. A city i at time t will be considered to be in decline if its net population change rate is below the 25th-percentile of the distribution of growth rates for the whole sample. Using the same distribution, the city will be consider to be in growth if its rate of population change is at or above the 75-th percentile. In any other case, a city is considered to have a normal rate of growth.

Table 1 lists summary statistics for the main variables of interest. From the CCDB, we include the following measures for each city: population, the percentage of population that is white and black, the percent of adults that completed college and high school,

[11]The data can be downloaded at `http://www.nws.noaa.gov/noaaport/html/geoex.exe`.

the unemployment and poverty rates, median household income and the percentage of population that is school-age (between the ages of 5 and 17) and elderly (65 years and older).

4.2 Methodology

The theory suggests that in equilibrium, per capita expenditures are a function of preference parameters (σ, γ, κ), the current population level, the population growth rate and union bargaining power. We thus use the following reduced form equation to capture these effects:

$$\log\left(\frac{pay^j_{i,t}}{z_{i,t}}\right) = \alpha_i + \theta_1 \log(z_{i,t}) + \theta_2 \log\left(\frac{z_{i,t}}{z_{i,t-1}}\right) + \boldsymbol{\beta}\mathbf{X_{i,t}} + \boldsymbol{\tau}\mathbf{Z_i} + \epsilon_{i,t} \tag{4}$$

where $pay^j_{i,t}$ represents the payroll for fulltime employees of service j in city i at time t; $z_{i,t}$ stands for population; $\mathbf{X_{i,t}}$ and $\mathbf{Z_i}$ are vectors of covariates to be described; and $\epsilon_{i,t}$ is a normally distributed iid random variable with mean zero and variance σ^2_e.[12] Following Bergstrom and Goodman (1973) and Alesina, Baqir and Easterly (1999), we implicitly assume that expenditures are independent across services.

The level of population may affect expenditures via two different channels: Nonlinearities in the production function and preferences. Though the theoretical model assumes that production is linear in labor, including population allows for the possibilities of economies or diseconomies of scale. Additionally, citizen preferences may depend on population size, if for example citizens demand more of a public good when the city is big.

City-specific variables that change through time and are indicative of preferences are included in the vector $\mathbf{X_{i,t}}$. In particular, we control for economic variables such as (log) median household income, unemployment and poverty; and demographic characteristics including race (percentage of black, whites and other races), the percentage of citizens with college education, and variables related to population age (the share of people above 65 years of age and the share between the ages of 5 and 17). Covariates that do not change in time, such as state fixed effects or climate, are captured by the variables in $\mathbf{Z_i}$. Since our estimation strategy of (4) is fixed effects, it will not be necessary to specify variables in $\mathbf{Z_i}$.

As is shown in Comment 1, in a world without frictions and perfect certainty, a planner is free to select the first best optimum at each point in time. Consequently, in a frictionless

[12]We have also used total operational expenditures in service j and found results that are robust to this change.

world, the provision of public goods should not depend on population change. As exemplified by the theoretical model, however, provision may depend on population change when workers have the ability to influence or constrain the policy maker. From Result 1 and the accompanying discussion, we know that if the service is politically organized, then for most population realizations, per capita expenditures should be decreasing in population change. From Comment 2, however, if the union is not politically organized, then the magnitude of the comparative static should be smaller and, for most of the population space, equal to zero. Consequently, we expect θ_2 to be negative if a service is politically active and, in the absence of other motives for frictions, zero or very small otherwise. Indeed, we will present estimates of (4) that confirm these expectations.

As discussed in the Introduction, there may be other reasons why $\theta_2 < 0$, such as other population-related frictions that are unrelated to the strategic decisions of public workers. The theory describes two ways in which public workers can influence policy: Collective bargaining and campaign support. Because our data provides information on collective bargaining, we can engage in capturing the effect of the former. We thus modify (4) to control for collective bargaining, $B_{i,t}$, using the measure we defined in the Data subsection:

$$\begin{aligned}\log\left(\frac{pay^j_{i,t}}{z_{i,t}}\right) &= \alpha_i + \theta_1 \log(z_{i,t}) + \theta_2 \log\left(\frac{z_{i,t}}{z_{i,t-1}}\right) + \phi_1 B_{i,t} + \phi_2 B_{i,t} \times \log\left(\frac{z_{i,t}}{z_{i,t-1}}\right) \\ &+ \boldsymbol{\beta}\mathbf{X_{i,t}} + \boldsymbol{\tau}\mathbf{Z_i} + \epsilon_{i,t}\end{aligned} \tag{5}$$

From Result 2, we know that for a wide range of population outcomes, per capita expenditures are increasing in bargaining power. Consequently, we expect ϕ_1 to be positive. Moreover, Result 3 shows that an increase in bargaining power weakly increases the effect of the population friction, and thus ϕ_2 should be negative.

Because the theory highlights that the effect of collective bargain differs based on the rate of population change, we try a second specification of (5). We classify each city in one of six categories according to two variables: Whether the union bargains collectively ($B_{i,t}$) and a categorical variable that identifies whether the city has experienced severe decline, relatively small population changes or a large growth rate ($C_{i,t}$, as presented in the previous section). Splitting cities into these categories allows us to assess how the effect of collective bargaining changes in different regions of population growth.

$$\log\left(\frac{pay^j_{i,t}}{z_{i,t}}\right) = \alpha_i + \theta_1 \log(z_{i,t}) + \theta_2 \log\left(\frac{z_{i,t}}{z_{i,t-1}}\right) + \boldsymbol{\phi}\mathbf{Q_{i,t}} + \boldsymbol{\beta}\mathbf{X_{i,t}} + \boldsymbol{\tau}\mathbf{Z_i} + \epsilon_{i,t} \tag{6}$$

where $Q_{i,t}$ is a 5×1 vector of dummy variables $q^{k,m}$ where $k \in \{decline, normal, grow\}$,

$m \in \{0, 1\}$ and the dummy for cities with steady growth and no collective bargaining ($q^{normal,B_{i,t}=0}$) is omitted. As suggested by the theory, we should observe a positive effect of collective bargain for cities in decline or with a normal level of growth, which would translate into positive estimates for $\phi^{decline,B=1}$ and $\phi^{normal,B=1}$. Finally, because at high levels of population growth, collective bargaining reduces per capita expenditures, we expect a negative sign on $\phi^{growth,B=1}$.

4.3 Endogeneity

One potential concern with the empirical approach is endogeneity arising from residential sorting (Tiebout (1956), Fernandez and Rogerson (1996)). The concern is that households choose to move to (or away from) a certain city because of the quality and cost of public goods. To explain how we treat this problem, we rearrange (4) in the following manner:

$$\log(pay^j_{i,t}) = \alpha_i + \delta_1 \log(z_{i,t}) + \delta_2 \log(z_{i,t-1}) + \boldsymbol{\beta}\mathbf{X_{i,t}} + \boldsymbol{\tau}\mathbf{Z_i} + \epsilon_{i,t} \tag{7}$$

where $\delta_1 = \theta_1 + \theta_2 - 1$ and $\delta_2 = -\theta_2$.

First, we consider migration between metropolitan areas, referred to as "long-distance internal migration." A voluminous literature surveyed in Greenwood (1997) considers employment as one of the main driving forces behind the decision to move to another metropolitan area. Herzog and Schlottmann (1986) and Fox, Herzog and Schlottman (1989) explicitly examine the effect of public goods on long-distance internal migration and conclude that there is a negative impact of spending per pupil on out-migration from US metropolitan areas.[13] Moreover, **?**, conclude that local public good spending is not significantly correlated with long-term population dynamics, whereas other economic variables such as employment prospects and education are strong predictors of population change. In short, the evidence suggests that the main determinants of internal migration are not related to the type of local goods that we will examine empirically. Consequently, if employment considerations are the main determinants of long distance migration, endogeneity has little or no role in metropolitan areas that are experiencing an exodus of citizens.

Endogeneity may occur, however, when households move within a labor market, or when

[13] More recently, Hunt and Mueller (2004) conclude that individuals locate to labor markets where the returns to their skills are highest. A series of papers that start with Day (1992) (see Day and Winer (2006) for a survey) studied the impact of public policies on migration between Canadian provinces. The authors conclude that the prime determinants of inter-provincial migration are differentials in earnings, employment prospects and moving costs. The impact of public policies, though, is small regardless of the specification employed.

households choose which municipality to locate in when moving to a new metropolitan area. In other words, such households face a problem of community choice that involves comparing the basket of attributes of the different cities within a metropolitan area, as described in McFadden et al. (1978).

If there is an endogeneity problem with the model as presented in (4), then there must be an omitted variable that is correlated with $z_{i,t}$ through its effect on individual residential decisions. The empirical literature on residential location (see, for instance, Nechyba and Strauss (1998)) specifies that consumers take characteristics of the communities, local public good levels, tax rates and equilibrium housing prices into account. In short, when deciding where to reside, an individual compares the vector of attributes of a particular municipality with other municipalities in the metropolitan area. This implies that that the omitted variables in (5) are a the vector of attributes of other neighboring communities. In particular, if local public goods, such as police services, affect location decisions, then a key omitted variable is the provision of the public goods in neighboring communities.

In order to account for the effects of public good provision in neighboring cities, we modify (4) to include spatial interactions, implying that a vector of expenditures in other cities will constitute an explanatory variable. As is standard in the spatial econometrics literature, the vector will be weighted by spatial closeness, with the particular weights being imposed prior to the estimation. In our setup of I cities, we define $w_{i,i'}$ as the weight given to the spatial relationship between cities i and i'. For each city i, the vector $\mathbf{w_i}$ groups all i weights where, naturally, the weight given to city i is zero. If $\mathbf{log(pay^j/pop)}$ represents the $Ix1$ vector of city expenditures on service j, then estimating an equation that controls for the level of public goods provided in neighboring regions involves the model in (8).

$$\begin{aligned} \log\left(\frac{pay^j_{i,t}}{pop_{i,t}}\right) &= \rho \mathbf{w_i'} \mathbf{log}\left(\frac{\mathbf{pay^j}}{\mathbf{pop}}\right) + \alpha_i + \theta_1 \log(pop_{i,t}) + \theta_2 \log\left(\frac{pop_{i,t}}{pop_{i,t-1}}\right) \\ &+ \boldsymbol{\beta} \mathbf{X_{i,t}} + \boldsymbol{\tau} \mathbf{Z_i} + \epsilon_{i,t} \end{aligned} \tag{8}$$

If we define $\mathbf{W}$ as a matrix with columns given by $\mathbf{w_i}$; and α, $\mathbf{log(pop)}$, $\mathbf{X}$, $\mathbf{Z}$, and ϵ as the corresponding matrices for the variables in the right hand side of (4), then (8) in matrix form is:

$$\mathbf{log}\left(\frac{\mathbf{pay^j}}{\mathbf{pop}}\right) = \rho \mathbf{W} \mathbf{log}\left(\frac{\mathbf{pay^j}}{\mathbf{pop}}\right) + \boldsymbol{\alpha} + \theta_1 \mathrm{log}(\mathbf{pop}) + \theta_2 \mathbf{R} + \boldsymbol{\beta}\mathbf{X} + \boldsymbol{\tau}\mathbf{Z} + \epsilon \tag{9}$$

where $\mathbf{R}$ is the vector of the log of the gross rate of population growth. As presented in (8), this corresponds to the Spatial Autorregresive Model. The coefficient ρ captures the correlation between the level of public good provision in city i and the weighted vector of public good provision in all other cities. Notice that the specification imposes this correlation to be the same for all cities in the sample. At the same time, in order for estimation to be possible by standard techniques, a set of weights must be imposed. In this paper we present results where the non-diagonal elements of the matrix have been computed as the inverse of the spherical distance between the two cities, thus giving a higher weight to cities that are closer in distance. Results are robust to other specifications, such as the standard procedure of giving equal weight to the 6 closest cities (see LeSage and Pace (2009)). We estimate all models discussed above also using the Spatial Autorregresive Model (SAR).

Finally, notice that although one important motivation for including a spatial lag is related to the potential endogeneity of intra-city migration, there are may be other reasons for ρ to be positive. For instance, it can directly capture the fact that public policy decisions in one city may directly affect the decisions of the leaders of other communities close by.

4.4 Results

We now report the estimation results of the models defined above. Our estimates are divided into two parts: First, we report the results regarding population frictions. Then, we then provide different measures to parse out the effect of organized bargaining.

Table 2 presents estimates for models (4) and (9). For each one of five services, the first column presents the coefficients for the within estimator, while the second column shows the fixed effects spatial model estimated by maximum likelihood (see Elhorst (2003)).[14]

There is a negative and significant effect of population change on fire and police expenditures. In fact, notice that fire and police are two of the three services for which we have unionization information, presumably because the unions of workers in these services are relatively more organized. Moreover, on average, police officers constitute approximately a quarter of the total number of full time employees in our sample, while fire workers constitute the second largest group, with 16%. In other words, we find a significant population friction effect precisely for services that are more likely to be valuable for the policymakers in terms of campaign support. Although it is not possible to directly compare the elasticities across models, both are centered around -0.2 and -0.1 and are

[14] The Hausman test leads to rejection of the random effects model. At the same time likelihood ratio tests side in favor of the presence of heterogeneities.

actually higher when controlling for spacial effects.[15]

Contrarily, we observe no statistical evidence of population frictions in the case of administrative, highway and parks services. Insofar as it can be ascertained, this evidence is consistent with these groups having little or no political power. In a few words, the estimates are consistent with Result 1: We observe a significant negative effect of population changes for services that traditionally are more politically active.[16]

Finally, notice that the general pattern in the fixed effects regression is not significantly altered when we control for spatial effects, even though the coefficient on spatial correlation is always positive and significant. The rest of the variables show estimates that are consistent across services. In particular, cities with relatively higher income, less poverty and a higher percentage of blacks spend more on these local public goods. Additionally, time dummies are significant for some of these services, perhaps capturing a general pattern of decentralization in the period under study.[17]

What is driving these population frictions? In the model, unions can influence policy through two distinct channels: First, by influencing elections, and second, via formal bargaining. By estimating (5), we provide a lower bound for the latter effect, and thus a lower bound for the total effect that these workers contribute to this friction.

The relevant time for which bargaining power matters is when the union actively negotiates the contract with the government. As discussed in Freeman (1986) it typically takes several years for the union to become effective at negotiating with the government, as contract tends to improve in the union's favor over time. Given these requirements we classify workers in a particular department as unionized if they have had collective bargaining power for ten years.[18]

We present the estimations of (5) for a 10 year lag in the collective union measure in table 3. Because data collected on unionization started in 1970 and was discontinued

[15]The comparison is partial because in the spatial model, the elasticities only measure the direct effect of a change in (log) gross population growth. There is a chain of indirect effects as a consequence of the spatial correlation. Notice, however, that this chain of indirect events can only increase the effect of the friction.

[16]One may be concerned that these expenditures are higher because there is a higher demand for these services when population declines. For example, expenditures in police may be related to the prevalence of crime and arson. We have run auxiliary regressions and observed that per capita number of serious crimes is not related to the city's population growth rate.

[17]The full sample of cities provides administrative services, and thus no observation is dropped from the sample in that case. In some cities in California fire services is provided at the county level, as is the case for park services in some cities in Illinois. In the case of the spatial regressions we require a balanced panel and thus the size of the sample considerably drops. Results are robust to the estimation of the non-spatial regressions constrained to a balanced panel

[18]We've also used a five year lag, but results are stronger for the 10 year lag (which corresponds to our previous population measure)

from 1990 onwards, we must restrict our sample to the 1970-1990 period. Moreover, because we are using a 10-year lag on the unionization measure, we can only conduct the analysis using wage bill data from 1980 and 1990.

We report results for the estimation of (5) (using the within estimator), as well as the "spatial" counterpart of 5, for the three services for which we have unionization information. In this specification, $B_{i,t}$ indicates if the union engaged in collective bargaining for at least the previous 10 years. According to Result 2, we expect bargaining power to increase the level of expenditures per capita for a wide range of population levels. In the specification presented in table 3 we capture a positive effect of unionization, significant at the 5% level for all specifications and all services.

The coefficient on the effect of population change is still negative and significant at the 5% level for fire and police, even when we control for bargaining power. In the case of highway services, θ_2 was not significant when estimating (4), but here is negative and significant in both specifications.[19] In short, population frictions are still present when we account for bargaining power, suggesting that there are additional frictions coming from other sources.

Result 3 indicates that ϕ_2 should be negative and significant for all three services. We inquire about the presence of this effect by including the interaction between the dummy variable and the change in population and find a negative effect, significant at the 6% level, for only fire services.

Finally, we estimate (6) in order to capture that, as suggested by the theory, the effects of collective bargain on per capita expenditures differ with population change. Results are presented in table 4.[20]

As expected, union bargaining power has a positive effect on expenditures per capita for cities in decline, as revealed by the estimation for fire services (at the 5% level) and police services (at the 10% level). In the case of growing cities, Result 2 states that the effect of bargaining on per capita expenditures should be negative. We find no significant effects in the case of fire services for growing cities, find weak evidence in favor of bargaining power for cities that are growing in the case of highway services and a positive significant effect for police.

Finally, notice first that the coefficient on population change remains significant at the

[19]Note that θ_2 is not significant in the highway regression when we estimate (4) using the 1980-1990 sample.

[20]A city is classified as *decline* (*grow*) if net population change is in the bottom 25th (top 75th) percentile of the distribution of growth rates. We have verified the results using several cutoffs. Results are robust to specifications where *decline* is the bottom 33rd, 20th or 15th percentile; and *grow* is the top 67th, 80th or 85th percentile, respectively.

5% level, regardless of specification, a result that is consistent with the fact that the political channel is still present as part of the frictions.

5 Conclusion

In this paper, we presented a theoretical model that examines the interaction between public sector workers' ability to bargain collectively and influence policy outcomes by campaigning for elected officials. We use this model to analyze the interaction between population dynamics and policy frictions and show frictions are increasingly strong as the city's population declines. The political economy channel is shown to be sufficient to generate frictions related to population change, and collective bargaining acts to exacerbate these effects. We proceed to test the model's implications using a panel of 909 US cities, and show that per capita spending on wages in traditionally organized services, such as police and fire, increases (decreases) as population falls (grows). In the less organized departments of administration, parks and road-related maintenance, we do not find evidence of a relationship between population change and wage expenditures. Moreover, there exists some evidence that collective bargaining intensifies this effect for fire departments.

Potential future research may involve studying the interaction between different services in a city. In this paper, we've viewed different services in isolation, and the question remains as to whether these departments collude or compete for resources. Moreover, this paper has focused exclusively on labor expenditures for public goods; a simultaneous examination of capital investment may yield further insights into how governments and public workers respond to population changes. Another important public good which we have omitted is education. Extending our methodology to this service would provide another take on the relationship between the provision local public goods and population dynamics. Education is a local public good that is unique because its governed by its own entity, the board of education. Finally, one additional avenue for extending this research could involve studying the interaction between population frictions and the relationship with various forms of local governments.

References

Alesina, A., R. Baqir, and W. Easterly, "Public Goods and Ethnic Divisions*," *Quarterly Journal of Economics*, 1999, *114* (4), 1243–1284.

Baqir, R., "Districting and government overspending," *Journal of Political Economy*, 2002, *110* (6), 1318–1354.

Baron, D.P., "Electoral competition with informed and uniformed voters," *American Political Science Review*, 1994, pp. 33–47.

Bergstrom, T.C. and R.P. Goodman, "Private demands for public goods," *The American Economic Review*, 1973, pp. 280–296.

Coate, Stephen and Stephen Morris, "Policy Persistence," *American Economic Review*, December 1999, *89* (5), 1327–1336.

Day, K.M., "Interprovincial migration and local public goods," *Canadian Journal of Economics*, 1992, pp. 123–144.

_ **and S.L. Winer**, "Policy-induced internal migration: An empirical investigation of the Canadian case," *International Tax and Public Finance*, 2006, *13* (5), 535–564.

Elhorst, J.P., "Specification and estimation of spatial panel data models," *International Regional Science Review*, 2003, *26* (3), 244.

Fernandez, R. and R. Rogerson, "Income distribution, communities, and the quality of public education," *The Quarterly Journal of Economics*, 1996, pp. 135–164.

Fox, W.F., H.W. Herzog, and A.M. Schlottman, "Metropolitan Fiscal Structure and Migration," *Journal of Regional Science*, 1989, *29* (4), 523–536.

Freeman, R.B., "Unionism comes to the public sector," *Journal of Economic Literature*, 1986, pp. 41–86.

Glaeser, E.L., J.A. Scheinkman, and A. Shleifer, "Economic growth in a cross-section of cities," *Journal of Monetary Economics*, 1995, *36* (1), 117–143.

Grossman, Gene M and Elhanan Helpman, "Electoral Competition and Special Interest Politics," *Review of Economic Studies*, April 1996, *63* (2), 265–86.

Herzog, H. and A. Schlottmann, "State and local tax deductibility and metropolitan migration," *National Tax Journal*, 1986, *39*, 189–200.

Hoxby, C.M., "How teachers' unions affect education production," *The Quarterly Journal of Economics*, 1996, pp. 671–718.

Hunt, G.L. and R.E. Mueller, "North American migration: returns to skill, border effects, and mobility costs," *Review of Economics and Statistics*, 2004, *86* (4), 988–1007.

LeSage, J.P. and R.K. Pace, *Introduction to spatial econometrics*, Chapman & Hall/CRC, 2009.

Levin, J., S. Tadelis, and L.E. Building, "Contracting for government services: theory and evidence from US cities," *NBER Working Paper*, 2007.

Lovenheim, M.F., "The Effect of Teachers Unions on Education Production: Evidence from Union Election Certifications in Three Midwestern States," *Journal of Labor Economics*, 2009, *27* (3).

McDonald, I.M. and R.M. Solow, "Wage bargaining and employment," *The American Economic Review*, 1981, pp. 896–908.

McFadden, D. et al., "Modelling the choice of residential location," *Spatial interaction theory and planning models*, 1978, *25*, 75–96.

Moe, Terry M., "Political Control and the Power of the Agent," *Journal of Law, Economics, and Organization*, 2006, *22* (1), 1–29.

Nechyba, TJ and RP Strauss, "Community choice and local public services: A discrete choice approach," *Regional Science and Urban Economics*, 1998, *28* (1), 51–73.

Niskanen, W.A., "Bureaucrats and politicians," *Journal of Law & Economics*, 1975, *18*, 617.

Prat, Andrea, "Campaign Spending with Office-Seeking Politicians, Rational Voters, and Multiple Lobbies," *Journal of Economic Theory*, March 2002, *103* (1), 162–189.

Romer, T. and H. Rosenthal, "Bureaucrats versus voters: On the political economy of resource allocation by direct democracy," *The Quarterly Journal of Economics*, 1979, pp. 563–587.

Tiebout, C.M., "A pure theory of local expenditures," *The journal of political economy*, 1956, *64* (5), 416.

A Proofs

A.1 Proof of Proposition 1

In this proof, we refer to (z_1, L_0) as the state of the game, and solve for the equilibrium (w_1, L_1). We begin by defining two functions that will aid in the proof. First, recall that the union's objective at $t = 1$ is $\max_{w,L,s} v(w, L, s; L_0)$ subject to (3). Clearly, (3) is satisfied with equality, and thus we can define the optimal s as a function of (w, L, z_1, L_0):[21]

$$s(w, L; z_1, L_0) \quad := \quad \left(\tilde{u}^B(z_1, L_0) - u(w, L; z_1)\right)^{\frac{1}{\gamma}} \left(\frac{z_1}{L_0}\right)$$

Second, we define is the politician's optimal L_1, conditional on (w, z_1):

$$\phi(w; z_1) \quad := \quad z_1 w^{-\frac{1}{1-\sigma}}$$

Note that if $L_1^* < \phi(w_1^*; z_1)$, then the politician will play $(w_1, L_1) = (w_1^*, \phi(w_1^*; z_1))$. Consequently, without loss of generality we can restrict $L_1^* \geq \min\{\phi(w_1^*; z_1), L_0\}$.

Next, note that any strategy (w', L') where $L' > L_0$ is weakly dominated by (w', L_0), because $\eta(L'; L_0)[w' - \underline{w}] = \eta(L_0; L_0)[w' - \underline{w}]$ but $s(w', L'; z_1, L_0) \geq s(w', L_0; z_1, L_0)$. Consequently, without loss of generality we restrict $L_1^* \leq L_0$. The union thus solves[22]

$$\max_{(w,L)\in\mathcal{W}\times[\min\{\phi(w;z_1),L_0\},L_0]} \frac{L}{L_0}\Big[w - \underline{w}\Big] - \frac{\gamma}{\kappa} s(w, L; z_1, L_0)^\kappa$$

Letting λ_L, λ_ϕ denote the multipliers for L, we have

$$\begin{aligned}
&\text{FOC } w: && \frac{L}{L_0} + u_w\left(\frac{z_1}{L_0}\right)^\gamma s^{\kappa-\gamma} - \lambda_\phi\phi_w \\
&\text{FOC } L: && \frac{[w - \underline{w}]}{L_0} + u_L\left(\frac{z_1}{L_0}\right)^\gamma s^{\kappa-\gamma} + \lambda_\phi - \lambda_L \\
&\text{SOC } w: && u_{ww}\left(\frac{z_1}{L_0}\right)^\gamma s^{\kappa-\gamma} - u_w^2\left(\frac{z_1}{L_0}\right)^{2\gamma}\left(\frac{\kappa-\gamma}{\gamma}\right)s^{\kappa-2\gamma} + \lambda_\phi\phi_{ww} \qquad (10) \\
&\text{SOC } L: && u_{LL}\left(\frac{z_1}{L_0}\right)^\gamma s^{\kappa-\gamma} - u_L^2\left(\frac{z_1}{L_0}\right)^{2\gamma}\left(\frac{\kappa-\gamma}{\gamma}\right)s^{\kappa-2\gamma} \\
&\text{Cross:} && \left[\frac{1}{L_0} + u_{wL}\left(\frac{z_1}{L_0}\right)^\gamma s^{\kappa-\gamma}\right] - \left[u_L u_w\left(\frac{z_1}{L_0}\right)^{2\gamma}\left(\frac{\kappa-\gamma}{\gamma}\right)s^{\kappa-2\gamma}\right]
\end{aligned}$$

[21] In the proof, we take s_1 as defined here and refer to the union's strategy as a pair (w_1, L_1).

[22] Since $L \leq L_0$, then $\eta(L; L_0) = L/L_0$. Moreover, the preferences represented by $\big[\eta(L; L_0)w + \big(1 - \eta(L; L_0)\big)\underline{w}\big]$ are identical to those represented by $\eta(L; L_0)\big[w - \underline{w}\big]$.

The proof defines the partition $\{\mathcal{S}^{\mathcal{H}}{}_1, \mathcal{S}^{\mathcal{M}}{}_1, \mathcal{S}^{\mathcal{L}}{}_1\}$ of the state space $\mathcal{Z} \times \mathcal{L}$ and defines the union's best response function, $(w_1^*, L_1^*) : \mathcal{Z} \times \mathcal{L} \times \mathcal{B} \to \mathcal{W} \times \mathcal{L}$, that maps state/bargaining power triples (z_1, L_0, B) into an action, (w_1, L_1). The first subsection deals with the region $\mathcal{S}^{\mathcal{H}}{}_1$, while the second subsection deals with $\mathcal{S}^{\mathcal{M}}{}_1 \cup \mathcal{S}^{\mathcal{L}}{}_1$.

A.1.1 $(z_1, L_0) \in \mathcal{S}^{\mathcal{H}}{}_1$

Define $z^D(L) := L\underline{w}^{\frac{1}{1-\sigma}}$ and $\mathcal{S}^{\mathcal{H}}{}_1 := \{(z, L) \in \mathcal{Z} \times \mathcal{L} \mid z < z^D(L)\}$. The characterization requires two steps: First, we formulate the function $(w_1^D, L_1^D) : \mathcal{S}^{\mathcal{H}}{}_1 \times \mathcal{B} \to \operatorname{int}\{\mathcal{W} \times [\phi(w; z_1), L_0]\}$ that maps to the unique local maximum on $\mathcal{W} \times [\phi(w; z_1), L_0]$.[23] Second, we verify that $(w_1^*, L_1^*) = (w_1^D, L_1^D) \iff (z_1, L_0) \in \mathcal{S}^{\mathcal{H}}{}_1$.

Step 1: Formulating (w_1^D, L_1^D): By assumption, (w_1^D, L_1^D) maps to the interior of $\mathcal{W} \times [\phi(w; z_1), L_0]$, and thus $L_1^D > \phi(w_1^D; z_1)$ for all $(z_1, L_0) \in \mathcal{S}^{\mathcal{H}}{}_1$. Consequently, $u(\cdot)$ takes the form of (2b). Dividing the two FOCs in (10),

$$\frac{L_1^D}{w_1^D - \underline{w}} = \frac{-u_w}{-u_L} = \frac{L_1^D / z_1}{w_1^D / z_1 - (L_1^D)^{\sigma-1} z_1^{-\sigma}}$$

There exists a unique solution, with $L_1^D(z_1, L_0, B) = z_1 \underline{w}^{-\frac{1}{1-\sigma}} < L_0$. Plugging L_1^D into the w FOC in (10), algebra yields

$$w_1^D(z_1, L_0, B) = \frac{\underline{w}}{\sigma} + \underline{w}^{\frac{1}{1-\sigma}} \left[\left(\frac{L_0}{z_1} \right)^{\frac{\gamma(\kappa-1)}{\kappa-\gamma}} - \tilde{u}^B(z_1, L_0) \right]$$

and $s(\cdot) = (z_1/L_0)^{\frac{1-\gamma}{\kappa-\gamma}}$. Note that $\partial w_1^D / \partial L_0 > 0$. To conclude Step 1, we verify that (w_1^D, L_1^D) is a local maximum. Calculus yields $u_{ww}, u_{LL} \leq 0$, which implies the

[23] $\operatorname{int}\{\mathcal{W} \times [\phi(w; z_1), L_0]\}$ is nonempty: $\forall (z_1, L_0) \in \mathcal{S}^{\mathcal{H}}{}_1$ and $\forall w \in \mathcal{W}$, $z_1 < L_0 \underline{w}^{\frac{1}{1-\sigma}} \leq L_0 w^{\frac{1}{1-\sigma}}$, and thus $\phi(w; z_1) < L_0$.

w, L SOCs in (10) are strictly negative. Moreover,

$$\begin{aligned}
\textit{Discriminant} &= u_{ww}u_{LL}\left(\frac{z_1}{L_0}s^{\kappa-\gamma}\right)^2 - u_{ww}u_L^2\left(\frac{z_1}{L_0}\right)^{3\gamma}s^{2\kappa-3\gamma} - u_{LL}u_w^2\left(\frac{z_1}{L_0}\right)^{3\gamma}s^{2\kappa-3\gamma} \\
&\quad + \left[u_Lu_w\left(\frac{z_1}{L_0}\right)^{2\gamma}\left(\frac{\kappa-\gamma}{\gamma}\right)s^{\kappa-2\gamma}\right]^2 - \left[u_Lu_w\left(\frac{z_1}{L_0}\right)^{2\gamma}\left(\frac{\kappa-\gamma}{\gamma}\right)s^{\kappa-2\gamma}\right]^2 \\
&\quad - \left[\frac{1}{L_0} + u_{wL}\left(\frac{z_1}{L_0}\right)^{\gamma}s^{\kappa-\gamma}\right]^2 + 2\left[\frac{1}{L_0} + u_{wL}\left(\frac{z_1}{L_0}\right)^{\gamma}s^{\kappa-\gamma}\right]\left[u_Lu_w\left(\frac{z_1}{L_0}\right)^{2\gamma}\left(\frac{\kappa-\gamma}{\gamma}\right)s^{\kappa-2\gamma}\right] \\
&= u_{ww}u_{LL}\left(\frac{z_1}{L_0}s^{\kappa-\gamma}\right)^2 - u_{ww}u_L^2\left(\frac{z_1}{L_0}\right)^{3\gamma}s^{2\kappa-3\gamma} - u_{LL}u_w^2\left(\frac{z_1}{L_0}\right)^{3\gamma}s^{2\kappa-3\gamma} \\
&\quad - \left[\frac{1}{L_0} + u_{wL}\left(\frac{z_1}{L_0}\right)^{\gamma}s^{\kappa-\gamma}\right]\left[\frac{1}{L_0} + u_{wL}\left(\frac{z_1}{L_0}\right)^{\gamma}s^{\kappa-\gamma} - 2u_Lu_w\left(\frac{z_1}{L_0}\right)^{2\gamma}\left(\frac{\kappa-\gamma}{\gamma}\right)s^{\kappa-2\gamma}\right] \\
&= u_{ww}u_{LL}\left(\frac{z_1}{L_0}s^{\kappa-\gamma}\right)^2 - u_{ww}u_L^2\left(\frac{z_1}{L_0}\right)^{3\gamma}s^{2\kappa-3\gamma} - u_{LL}u_w^2\left(\frac{z_1}{L_0}\right)^{3\gamma}s^{2\kappa-3\gamma} \\
&> 0
\end{aligned}$$

The second equality follows from factoring; the third equality follows from $u_{wL}(z_1/L_0)^{\gamma}s^{\kappa-\gamma} = -1/L_0$ at (w_1^D, L_1^D); and the strict inequality follows from $u_{LL} < 0$.

Step 2: $(w_1^*, L_1^*) = (w_1^D, L_1^D) \iff (z_1, L_0) \in \mathcal{S}^{\mathcal{H}}{}_1$**:** By contradiction, assume that $(w_1^*, L_1^*) = (w_1^D, L_1^D)$ but $(z_1, L_0) \notin \mathcal{S}^{\mathcal{H}}{}_1$. Since $(z_1, L_0) \notin \mathcal{S}^{\mathcal{H}}{}_1$, then $z_1 > L_0\underline{w}^{1/(1-\sigma)}$. But then $L_1^*(\cdot) = z_1\underline{w}^{-\frac{1}{1-\sigma}} > L_0$, which contradicts $L_1^* \leq L_0$.

Conversely, let $(z_1, L_0) \in \mathcal{S}^{\mathcal{H}}{}_1$. From Step 1, (w_1^D, L_1^D) maps to the unique local maximum on $\mathcal{W} \times [\phi(w; z_1), L_0]$. Thus, we must check other (w_1', L_1') on the boundary. First, consider $L_1' = L_0$. As is shown in the $\mathcal{S}^{\mathcal{M}}{}_1 \times \mathcal{S}^{\mathcal{L}}{}_1$ subsection, when $L_1' = L_0$, the optimal wage is

$$w_1' = \frac{1}{\sigma}\left(\frac{z_1}{L_0}\right)^{1-\sigma} + \left(\frac{z_1}{L_0}\right)^{\frac{\kappa(1-\gamma)}{\kappa-\gamma}} - \left(\frac{z_1}{L_0}\right)\tilde{u}^B(z_1, L_0)$$

and $s_1' = (z_1/L_0)^{\frac{1-\gamma}{\kappa-\gamma}}$. Plugging into $v(\cdot)$, we have

$$\begin{aligned}
v(w_1', L_1', s_1'; L_0) &= \left[\frac{1}{\sigma}\left(\frac{z_1}{L_0}\right)^{1-\sigma} - \underline{w}\right] - \left(\frac{z_1}{L_0}\right)\tilde{u}^B(z_1, L_0) + \frac{k-\gamma}{k}\left(\frac{z_1}{L_0}\right)^{\frac{\kappa(1-\gamma)}{\kappa-\gamma}} \\
&< \left[\frac{1-\sigma}{\sigma}\frac{z_1\underline{w}^{-\frac{\sigma}{1-\sigma}}}{L_0}\right] - \left(\frac{z_1}{L_0}\right)\tilde{u}^B(z_1, L_0) + \frac{k-\gamma}{k}\left(\frac{z_1}{L_0}\right)^{\frac{\kappa(1-\gamma)}{\kappa-\gamma}} \\
&= v(w_1^D, L_1^D, s_1^D; L_0)
\end{aligned}$$

Thus (w_1^D, L_1^D) is preferred.[24] Next, consider (w_1', L_1') where $L_1' = \phi(w_1'; z_1)$. Then

$$\begin{aligned}\frac{\partial}{\partial L_1}\left[\frac{L_1}{L_0}[w-\underline{w}] - \frac{1}{\kappa}s(\cdot)^\kappa\right](w', L') &= \frac{[w'-\underline{w}]}{L_0} + \frac{1}{z_0}\left[\left(\frac{z_0}{L'}\right)^{1-\sigma} - w'\right]\left(\frac{z_1}{L_0}\right)^\gamma s^{\kappa-\gamma} + \lambda_\phi \\ &= \frac{1}{L_0}[w'-\underline{w}] + \lambda_\phi \\ &> 0\end{aligned}$$

where the second equality follows from $L' = \phi(w'; z_1) = z_1 (w')^{-\frac{1}{1-\sigma}}$. Consequently, $L' = \phi(w'; z_1)$ cannot be optimal when $(z_1, L_0) \in \mathcal{S^H}_1$.

Finally, when $w_1' = \underline{w}$, then $v(\cdot) \leq 0$, which is suboptimal. On the other hand, when $w_1' \to \infty$, then $s(\cdot)^\kappa \to \infty$ at a faster rate then $[w_1' - \underline{w}] \to \infty$ because $\kappa/\gamma > 1$, and thus $\lim_{w_1' \to \infty} v(\cdot) = -\infty$. Consequently, $(w_1^*, L_1^*) = (w_1^D, L_1^D)$.

A.1.2 $(z_1, L_0) \in \mathcal{S^M}_1 \cup \mathcal{S^L}_1$

When $(z_1, L_0) \notin \mathcal{S^H}_1$, then $L_1^* = L_0$,[25] and thus the union's objective is $\max_w [w - \underline{w}] - (\gamma/\kappa) s(w, L_0; z_1, L_0)^\kappa$. The union's optimal strategy depends on the functional form of $s(\cdot)$: If $\phi(w; z_1) < L_0$, then $u(\cdot)$ is of form (2b), whereas if $\phi(w; z_1) \geq L_0$, then $u(\cdot)$ is of form (2a). Define[26]

$$w_1^C(z_1, L_0, B) = \arg\max_w [w - \underline{w}] - \left(\frac{\gamma}{\kappa}\right)\left(\frac{z_1}{L_0}\right)^\kappa \left[\tilde{u}^B(z_1, L_0) - \left(\frac{1}{\sigma}\left(\frac{L_0}{z_1}\right)^\sigma - \frac{L_0 w}{z_1}\right)\right]^{\frac{\kappa}{\gamma}} \tag{11a}$$

$$w_1^G(z_1, L_0, B) = \arg\max_w [w - \underline{w}] - \left(\frac{\gamma}{\kappa}\right)\left(\frac{z_1}{L_0}\right)^\kappa \left[\tilde{u}^B(z_1, L_0) - \left(\frac{1-\sigma}{\sigma}\right) w^{-\frac{\sigma}{1-\sigma}}\right]^{\frac{\kappa}{\gamma}} \quad \text{s.t.} \quad \phi(w; z_1) \geq L_0 \tag{11b}$$

---

[24] The strict inequality occurs because of the following: When $z_1 = z^D(L_0) = L_0 \underline{w}^{\frac{1}{1-\sigma}}$, then

$$\left[\frac{1}{\sigma}\left(\frac{z^D(L_0)}{L_0}\right)^{1-\sigma} - \underline{w}\right] = \left(\frac{1-\sigma}{\sigma}\right)\underline{w} = \left[\frac{1-\sigma}{\sigma}\frac{z^D(L_0)\underline{w}^{-\frac{\sigma}{1-\sigma}}}{L_0}\right]$$

But since $(L_0/z_1)^\sigma > \underline{w}^{-\frac{\sigma}{1-\sigma}}$ for any $(z_1, L_0) \in \mathcal{S^H}_1$, then

$$\begin{aligned}\frac{\partial}{\partial z}\left[\frac{1}{\sigma}\left(\frac{z}{L_0}\right)^{1-\sigma} - \underline{w}\right] &= \left(\frac{1-\sigma}{\sigma}\right)\frac{1}{L_0}\left(\frac{L_0}{z}\right)^\sigma \\ &> \left(\frac{1-\sigma}{\sigma}\right)\frac{1}{L_0}\underline{w}^{-\frac{\sigma}{1-\sigma}} \\ &= \frac{\partial}{\partial z}\left[\frac{1-\sigma}{\sigma}\frac{z\underline{w}^{-\frac{\sigma}{1-\sigma}}}{L_0}\right]\end{aligned}$$

[25] From the $\mathcal{S^H}_1$ subsection, if $z_1 \geq z^D(L_0)$, then there are no critical points on $\mathcal{W} \times [\phi(w; z_1), L_0]$. Moreover, $L' = \phi(w'; z_1)$, $w' = \underline{w}$ and $w' \to \infty$ are suboptimal strategies by identical arguments.

[26] To reduce notation, we allow $w_1^C, w_1^G : \Re_+ \times \mathcal{L} \times \mathcal{B} \to \mathcal{W}$, thus allowing the first argument $z \in \Re_+$ as opposed to $z \in \mathcal{Z}$.

and define $z^G(L;B)$ by the implicit function

$$v\Big(w_1^C(z^G(L;B),L,B),L,s(\cdot);L\Big) = v\Big(w_1^G(z^G(L;B),L,B),L,s(\cdot);L\Big) \tag{12}$$

For now, we assert that $z^G : \mathcal{L} \times \mathcal{B} \to \Re$ is a well-defined function. This assertion will be verified below. Given this assertion, we can define $\mathcal{S}^{\mathcal{M}}{}_1 := \{(z_1, L_0) \in \mathcal{Z} \times \mathcal{L} \mid z_1 \in [z^D(L_0), z^G(L_0;B)]\}$ and $\mathcal{S}^{\mathcal{L}}{}_1 := \{(z_1, L_0) \in \mathcal{Z} \times \mathcal{L} \mid z_1 > z^G(L_0;B)\}$.

The proof consists of five steps: Steps 1 and 2 formulate $w_1^C(\cdot)$ and $w_1^G(\cdot)$, respectively. Step 3 shows $\forall L \in \mathcal{L}$, $\exists z', z'' \in \Re_+$ where $w_1^* = w_1^C$ when $z_1 = z'$ and $w_1^* = w_1^G$ when $z_1 = z''$. Step 4 shows $\forall L \in \mathcal{L}$, there exists, at most, one $z' \in \Re_+$ such that $w_1^* = (z'/L)^{1-\sigma}$, which is used to show z^G is well-defined. Finally, Step 5 shows that $\forall L \in \mathcal{L}$, $\exists! z' \in \Re_+$ satisfying $v\Big(w_1^G(z',L,B),L,s(\cdot);L\Big) = v\Big(w_1^C(z',L,B),L,s(\cdot);L\Big)$.

Step 1: Formulating $w_1^C(\cdot)$: The union's objective function in (11a) is strictly concave. The FOC yields

$$w_1^C(z_1,L_0,B) = \frac{1}{\sigma}\left(\frac{z_1}{L_0}\right)^{1-\sigma} + \left(\frac{z_1}{L_0}\right)^{\frac{\kappa(1-\gamma)}{\kappa-\gamma}} - \left(\frac{z_1}{L_0}\right)\tilde{u}^B(z_1,L_0)$$

and $s_1^C = (z_1/L_0)^{\frac{1-\gamma}{\kappa-\gamma}}$. Note that $\partial w_1^C/\partial L_0$ can be either positive or negative, and $\phi(w_1^C; z_1) < L_0$ iff $(z_1/L_0)^{1-\sigma} < w_1^C$.

Step 2: Formulating $w_1^G(\cdot)$: It is straightforward to show that the objective in (11b) is strictly increasing at $w = \underline{w}$, strictly concave when $w \in \Big[\underline{w}, \tilde{w}[1+(\kappa-1)\sigma/\gamma]^{\frac{1-\sigma}{\sigma}}\Big)$ and strictly convex when $w \in \Big(\tilde{w}[1+(\kappa-1)\sigma/\gamma]^{\frac{1-\sigma}{\sigma}}, \infty\Big)$. Consequently, w_1^G will equal either (1) the boundary, $(z_1/L_0)^{1-\sigma}$ or (2) the local maximum on $w \in \Big[\underline{w}, \tilde{w}[1+(\kappa-1)\sigma/\gamma]^{\frac{1-\sigma}{\sigma}}\Big)$ (if it exists). The local maximum (if it exists) can be implicitly defined as

$$w_1^G = \left\{\left(\frac{z_1}{L_0}\right)^{\kappa}\left[\tilde{u}^B(z_1,L_0) - \left(\frac{1-\sigma}{\sigma}\right)(w_1^G)^{-\frac{\sigma}{1-\sigma}}\right]^{\frac{\kappa-\gamma}{\gamma}}\right\}^{1-\sigma} \tag{13}$$

Taking the total derivative of (13) with respect to L_0, we find

$$\frac{\partial w_1^G}{\partial L_0} = \frac{1}{L_0}\frac{\kappa(1-\sigma)w_1^G\left[\tilde{w}^{-\frac{\sigma}{1-\sigma}} - (w_1^G)^{-\frac{\sigma}{1-\sigma}}\right]}{(w_1^G)^{-\frac{\sigma}{1-\sigma}}\left[\frac{\gamma+(\kappa-\gamma)\sigma}{\gamma}\right] - \tilde{w}^{-\frac{\sigma}{1-\sigma}}} > 0$$

Step 3: $\forall L \in \mathcal{L}$, $\exists z', z'' \in \Re_+$ such that $w_1^*(z',L,B) = w_1^C(z',L,B)$ and $w_1^*(z'',L,B) = w_1^G(z'',L,B)$: Since $w^D(\cdot) = w^C(\cdot)$ at $(z_1^D(L),L,B)$, then from the $\mathcal{S}^{\mathcal{H}}{}_1$ subsection we know that

$w_1^*(z^D(L), L, B) = w_1^C(z^D(L), L, B)$.

On the other hand, if $w_1^C \leq (z/L)^{1-\sigma}$, then $L_1 \leq \phi(w_1^C; z)$ by the definition of ϕ. Rearranging $w_1^C \leq (z/L)^{1-\sigma}$, we have

$$\frac{1-\sigma}{\sigma}\left(\frac{L}{z}\right)^{\sigma} + \left(\frac{L}{z}\right)^{\frac{\gamma(\kappa-1)}{\kappa-\gamma}} \leq \tilde{u}^B(z, L)$$

As $z \to \infty$, the $LHS \to 0$ while the $RHS \to \tilde{w}^{-\frac{\sigma}{1-\sigma}}(1-\sigma)/\sigma > 0$. Consequently, $\exists z'$ large such that $L < \phi(w^C(z', L, B); z')$. But then $w_1^*(z', L, B) = w^G(z', L, B)$ by definition.

Step 4: $\forall L \in \mathcal{L}$. $\exists$ at most one $z' \in \mathcal{Z}$ such that $w_1^*(z', L, B) = (z'/L)^{1-\sigma}$: By contradiction, assume $\exists z', z'' \in \mathcal{Z}$ such that $w_1^*(z', L, B) = (z'/L)^{1-\sigma}$ and $w_1^*(z'', L, B) = (z''/L)^{1-\sigma}$. Note that $\forall (z, L) \in \mathcal{Z} \times L$, the two objectives in (11a) and (11b) are equal and have the same derivative at $w = (z/L)^{1-\sigma}$. Consequently, by the strict concavity of the objective in (11a), it must be that $w_1^C(z', L, B) \leq (z'/L)^{1-\sigma}$ and $w_1^C(z'', L, B) \leq (z''/L)^{1-\sigma}$, as otherwise w_1^C would yield a higher payoff.

From Step 3, $\exists! z$ satisfying $w_1^C(z, L, B) = (z/L)^{1-\sigma}$. Thus, without loss of generality, we assume that $w_1^C(z'', L, B) < (z''/L)^{1-\sigma}$. By strict concavity of the objective in (11a), utility is decreasing in w at $(z''/L)^{1-\sigma}$. Since the derivatives of the objectives in (11a) and (11b) are equal at $(z''/L)^{1-\sigma}$, then $v_w((z''/L)^{1-\sigma}, L, s(\cdot); L) < 0$, and thus $\exists w'' < (z''/L)^{1-\sigma}$ such that $v(w'', L, s(\cdot); L) > v((z''/L)^{1-\sigma}, L, s(\cdot); L)$, a contradiction.

Step 5: $\forall L \in \mathcal{L}$, $\exists! z'$ such that $v\Big(w_1^G(z', L, B), L, s(\cdot); L\Big) = v\Big(w_1^C(z', L, B), L, s(\cdot); L\Big)$: Step 4 implies that when $w_1^* = w_1^G$, then w_1^G satisfies (13). In other words, when w_1^G is played, the constraint in (11b) does not bind. Note that the objective in (11b) is strictly convex in z: The SOC is[27]

$$\frac{1}{z}\left(\frac{z}{L}\right)^{\kappa}\left[\tilde{u}^B(z, L) - \left(\frac{1-\sigma}{\sigma}\right)(w_1^G)^{-\frac{\sigma}{1-\sigma}}\right]^{\frac{\kappa-\gamma}{\gamma}}\left\{-\frac{\gamma(\kappa-1)}{z}\left(\frac{1-\sigma}{\sigma}\right)\left[\tilde{w}^{-\frac{\sigma}{1-\sigma}} - (w_1^G)^{-\frac{\sigma}{1-\sigma}}\right] - \kappa (w_1^G)^{-\frac{1}{1-\sigma}}\frac{\partial w_1^G}{\partial z}\right\}$$

Taking the total derivative of (13) with respect to z, we find

$$\frac{\partial w_1^G}{\partial z} = \frac{1}{z}\frac{\kappa(1-\sigma)w_1^G\left[\tilde{w}^{-\frac{\sigma}{1-\sigma}} - (w_1^G)^{-\frac{\sigma}{1-\sigma}}\right]}{\tilde{w}^{-\frac{\sigma}{1-\sigma}} - (w_1^G)^{-\frac{\sigma}{1-\sigma}}\left[\frac{\gamma+(\kappa-\gamma)\sigma}{\gamma}\right]} < 0$$

[27] From Step 4, when w_1^G is played, then $L \leq z(w_1^G)^{-\frac{1}{1-\sigma}} < z\tilde{w}^{-\frac{1}{1-\sigma}}$ since $w_1^G > \tilde{w}$. Consequently, $\tilde{u}^B(z, L)$ takes the form (2a) when w_1^G is played, and thus $\tilde{u}^B(z, L)$ doesn't depend on z.

Substituting in, we see that the sign of the SOC depends on

$$\begin{aligned}\frac{\kappa^2}{1-\left(\frac{w_1^G}{\tilde{w}}\right)^{\frac{\sigma}{1-\sigma}}+\left(\frac{(\kappa-\gamma)\sigma}{\gamma}\right)}-\frac{\gamma(\kappa-1)}{\sigma} &> \frac{\kappa^2}{\kappa-\gamma}-(\kappa-1)\\ &> 0\end{aligned}$$

where the first inequality follows from $w_1^G > \tilde{w}$ and the second inequality follows from $\kappa > 1 > \gamma$. On the other hand, the objective in (11a) is strictly concave in z; the SOC is

$$\frac{1}{z_1 L_0}\left[-(1-\sigma)\left(\frac{L_0}{z_1}\right)^{\sigma}-\frac{\gamma(1-\gamma)(\kappa-1)}{\kappa-\gamma}\left(\frac{L_0}{z_1}\right)^{\frac{\gamma(\kappa-1}{\kappa-\gamma}}-2z_1\tilde{u}_z-z_1^2\tilde{u}_{zz}\right]<0$$

Since $v(\cdot)$ is continuous, then $v(w_1^*, L_0, s(\cdot); L_0)$ is continuous in z by the Theorem of the Maximum. Thus from Step 3, $\exists z \in \Re_+$ such that (12) is satisfied: Let $z' := \min\{z \geq z^D(L) | (12) \text{ is satisfied}\}$. From Step 3, the derivative of (11a) at $z' <$ the derivative of (11b) at z'. Since (11a) is concave and (11b) is convex, z' is the unique z satisfying (12). ■

A.2 Proof of Result 1

Let (z_0, L_{-1}) denote the initial state of the game. The following lemma shows that the equilibrium L_0 is such that if $z_1 = z_0$ at $t = 1$, so that the same population realization occurs twice, then $(z_1, L_0) \in \mathcal{S}^{\mathcal{M}}$, which is the "medium" region of the state space.

Lemma 1. $\forall (z_0, L_{-1}) \in \mathcal{Z} \times \mathcal{L}$, *the equilibrium* L_0 *is such that* $(z_0, L_0) \in \mathcal{S}^{\mathcal{M}}$.

Proof. We consider three cases: $(z_0, L_{-1}) \in \mathcal{S}^{\mathcal{H}}$, $(z_0, L_{-1}) \in \mathcal{S}^{\mathcal{M}}$ and $(z_0, L_{-1}) \in \mathcal{S}^{\mathcal{L}}$.

Case 1: If $(z_0, L_{-1}) \in \mathcal{S}^{\mathcal{H}}$, then $L_0 = z_0 \underline{w}^{-\frac{1}{1-\sigma}}$. But $(z_0, z_0 \underline{w}^{-\frac{1}{1-\sigma}}) \in \mathcal{S}^{\mathcal{M}}$ by definition.

Case 2: If $(z_0, L_{-1}) \in \mathcal{S}^{\mathcal{M}}$, then $L_0 = L_{-1}$, and thus $(z_0, L_0) \in \mathcal{S}^{\mathcal{M}}$.

Case 3: If $(z_0, L_{-1}) \in \mathcal{S}^{\mathcal{L}}$, then $L_0 > L_{-1}$. By contradiction, assume that $(z_0, L_0) \notin \mathcal{S}^{\mathcal{M}}$.

First, note that $L_0 = z_0 w_0^{-\frac{1}{1-\sigma}} < z_0 \underline{w}^{-\frac{1}{1-\sigma}}$ since $w_0 > \underline{w}$. Consequently, $L_0 \underline{w}^{\frac{1}{1-\sigma}} = z^D(L_0) < z_0$, which implies that $(z_0, L_0) \notin \mathcal{S}^{\mathcal{H}}$.

Consequently, $(z_0, L_0) \in \mathcal{S}^{\mathcal{L}}$, which implies that $z_0 w_1^{-\frac{1}{1-\sigma}} = L_1 > L_0 = z_0 w_0^{-\frac{1}{1-\sigma}}$, and thus $w_1 < w_0$. By revealed preference at $t = 1$, we know that $v(w_1, L_1, s(\cdot); L_0) > v(w_0, L_0, s(\cdot); L_0)$. Rearranging this inequality, we have[28]

[28]This equation uses the fact that $\tilde{u}^B(z, L) = \tilde{w}^{-\frac{\sigma}{1-\sigma}}(1-\sigma)/\sigma$ when $(z, L) \in \mathcal{S}^{\mathcal{L}}$.

$$
\begin{aligned}
\frac{\gamma}{\kappa}\left[\frac{z_0}{L_0}\right]^{\kappa}\left[\frac{1-\sigma}{\sigma}\right]^{\frac{\kappa}{\gamma}}\left\{\left[\tilde{w}^{-\frac{\sigma}{1-\sigma}}-w_0^{-\frac{\sigma}{1-\sigma}}\right]^{\frac{\kappa}{\gamma}}-\left[\tilde{w}^{-\frac{\sigma}{1-\sigma}}-w_1^{-\frac{\sigma}{1-\sigma}}\right]^{\frac{\kappa}{\gamma}}\right\} \quad &> \quad [w_0-\underline{w}]-[w_1-\underline{w}] \\
&> \quad 0
\end{aligned}
$$

where the second inequality follows from $w_1 < w_0$. But since $L_0 > L_{-1}$, then

$$
\begin{aligned}
&\frac{\gamma}{\kappa}\left[\frac{z_0}{L_{-1}}\right]^{\kappa}\left[\frac{1-\sigma}{\sigma}\right]^{\frac{\kappa}{\gamma}}\left\{\left[\tilde{w}^{-\frac{\sigma}{1-\sigma}}-w_0^{-\frac{\sigma}{1-\sigma}}\right]^{\frac{\kappa}{\gamma}}-\left[\tilde{w}^{-\frac{\sigma}{1-\sigma}}-w_1^{-\frac{\sigma}{1-\sigma}}\right]^{\frac{\kappa}{\gamma}}\right\} \\
> \quad &\frac{\gamma}{\kappa}\left[\frac{z_0}{L_0}\right]^{\kappa}\left[\frac{1-\sigma}{\sigma}\right]^{\frac{\kappa}{\gamma}}\left\{\left[\tilde{w}^{-\frac{\sigma}{1-\sigma}}-w_0^{-\frac{\sigma}{1-\sigma}}\right]^{\frac{\kappa}{\gamma}}-\left[\tilde{w}^{-\frac{\sigma}{1-\sigma}}-w_1^{-\frac{\sigma}{1-\sigma}}\right]^{\frac{\kappa}{\gamma}}\right\}
\end{aligned}
$$

Rearranging these two inequalities, $v(w_1, L_1, s(\cdot); L_{-1}) > v(w_0, L_0, s(\cdot); L_{-1})$, which implies that (w_1, L_1) is optimal for the union at $t = 0$, a contradiction.

$\square$

Returning to the main proof, in equilibrium,

$$
w_1 l_1 = \begin{cases}
\frac{1}{\sigma}\underline{w}^{-\frac{\sigma}{1-\sigma}} + \left(\frac{L_0}{z_1}\right)^{\frac{\gamma(\kappa-1)}{\kappa-\gamma}} - \tilde{u}^B(z_1, L_0) & \text{if } (z_1, L_0) \in \mathcal{S}^{\mathcal{H}} \\
\frac{1}{\sigma}\left(\frac{L_0}{z_1}\right)^{\sigma} + \left(\frac{L_0}{z_1}\right)^{\frac{\gamma(\kappa-1)}{\kappa-\gamma}} - \tilde{u}^B(z_1, L_0) & \text{if } (z_1, L_0) \in \mathcal{S}^{\mathcal{M}} \\
w^G(z_1, L_0, B)^{-\frac{\sigma}{1-\sigma}} & \text{if } (z_1, L_0) \in \mathcal{S}^{\mathcal{L}}
\end{cases} \tag{14}
$$

and

$$
\frac{\partial(w_1 l_1)}{\partial z_1} = \begin{cases}
-\frac{1}{z_1}\frac{\gamma(\kappa-1)}{\kappa-\gamma}\left(\frac{L_0}{z_1}\right)^{\frac{\gamma(\kappa-1)}{\kappa-\gamma}} - \tilde{u}_z^B(z_1, L_0) & \text{if } (z_1, L_0) \in \mathcal{S}^{\mathcal{H}} \\
-\frac{1}{z_1}\left\{\left(\frac{L_0}{z_1}\right)^{\sigma} + \frac{\gamma(\kappa-1)}{\kappa-\gamma}\left(\frac{L_0}{z_1}\right)^{\frac{\gamma(\kappa-1)}{\kappa-\gamma}}\right\} - \tilde{u}_z^B(z_1, L_0) & \text{if } (z_1, L_0) \in \mathcal{S}^{\mathcal{M}} \\
-\frac{\sigma}{1-\sigma} w^G(z_1, L_0, B)^{-\frac{1}{1-\sigma}}\frac{\partial w^G}{\partial z_1} & \text{if } (z_1, L_0) \in \mathcal{S}^{\mathcal{L}}
\end{cases}
$$

Since $\tilde{u}_z^B(z_1, L_0) > 0$, then $\partial(w_1 l_1)/\partial z_1 < 0$ when $(z_1, L_0) \in \mathcal{S}^{\mathcal{H}} \cup \mathcal{S}^{\mathcal{M}}$. Since $\partial w^G/\partial z_1 < 0$ (shown in the proof of Proposition 1), then $\partial(w_1 l_1)/\partial z_1 > 0$ when $(z_1, L_0) \in \mathcal{S}^{\mathcal{L}}$. Consequently, $z^S = z^G(L_0, B)$ satisfies the statement in the Result. From Lemma 1, it follows that $z_0 \leq z^G(L_0, B)$. To conclude, note that $\lim_{z_1 \to \infty} w^G(z_1, L_0, B) = \tilde{w}$, and thus $\lim_{z_1 \to \infty} w_1 l_1 = w^G(z_1, L_0, B)^{-\frac{\sigma}{1-\sigma}}$. $\blacksquare$

A.3 Proof of Result 2

From (15), it follows that

$$\frac{\partial(w_1 l_1)}{\partial B} = \begin{cases} -\frac{\partial \tilde{u}^B(z_1,L_0)}{\partial B} & \text{if } (z_1, L_0) \in \mathcal{S}^{\mathcal{H}} \cup \mathcal{S}^{\mathcal{M}} \\ -\frac{\sigma}{1-\sigma} w^G(z_1, L_0, B)^{-\frac{1}{1-\sigma}} \frac{\partial w^G}{\partial B} & \text{if } (z_1, L_0) \in \mathcal{S}^{\mathcal{L}} \end{cases}$$

Note that

$$\frac{\partial \tilde{u}^B(z_1, L_0)}{\partial B} = \begin{cases} -\frac{\sigma}{1-\sigma} \tilde{w}^{-\frac{1}{1-\sigma}} \frac{\partial \tilde{w}}{\partial B} & \text{if } \tilde{\eta} L_{-1} \leq z \tilde{w}^{-\frac{1}{1-\sigma}} \\ \frac{L_0}{z_1} \left[\left(\frac{z_0}{\tilde{\eta} L_0} \right)^{1-\sigma} - \tilde{w} \right] \frac{\partial \tilde{\eta}}{\partial B} - \frac{\tilde{\eta} L_0}{z_1} \frac{\partial \tilde{w}}{\partial B} & \text{otherwise} \end{cases} \tag{15}$$

Since $\partial \tilde{\eta}/\partial B > 0$ and $\partial \tilde{w}/\partial B > 0$, then $\partial \tilde{u}^B(z_1, L_0)/\partial B < 0$. Consequently, if $(z_1, L_0) \in \mathcal{S}^{\mathcal{H}} \cup \mathcal{S}^{\mathcal{M}}$, then $\partial(w_1 l_1)/\partial B > 0$.

Taking the total derivative of (13) with respect to B, it follows that

$$\frac{\partial w^G}{\partial B} = \frac{\frac{\partial \tilde{w}}{\partial B} \sigma \left(\frac{\kappa - \gamma}{\gamma} \right) \left(\frac{w^G}{\tilde{w}} \right)^{\frac{1}{1-\sigma}}}{\sigma \left(\frac{\kappa-\gamma}{\gamma} \right) + 1 - \left(\frac{w^G}{\tilde{w}} \right)^{\frac{\sigma}{1-\sigma}}} > 0 \tag{16}$$

Consequently, if $(z_1, L_0) \in \mathcal{S}^{\mathcal{L}}$, then $\partial(w_1 l_1)/\partial B < 0$. Define $z^S = z^G(z_1, L_0)$. From Lemma 1, we know that $z_0 \leq z^S$. ■

A.4 Proof of Result 3

From (15), it follows that

$$\frac{\partial^2(w_1 l_1)}{\partial B \partial z_1} = \begin{cases} -\frac{\partial^2 \tilde{u}^B(z_1,L_0)}{\partial B \partial z_1} & \text{if } (z_1, L_0) \in \mathcal{S}^{\mathcal{H}} \cup \mathcal{S}^{\mathcal{M}} \\ \frac{\sigma}{(1-\sigma)^2} w^G(z_1, L_0, B)^{-\frac{2-\sigma}{1-\sigma}} \frac{\partial w^G}{\partial B} \frac{\partial w^G}{\partial z_1} - \frac{\sigma}{1-\sigma} w^G(z_1, L_0, B)^{-\frac{1}{1-\sigma}} \frac{\partial^2 w^G}{\partial B \partial z_1} & \text{if } (z_1, L_0) \in \mathcal{S}^{\mathcal{L}} \end{cases}$$

From (15), we have

$$\frac{\partial^2 \tilde{u}^B(z_1, L_0)}{\partial B \partial z_1} = \begin{cases} 0 & \text{if } \tilde{\eta} L_{-1} \leq z \tilde{w}^{-\frac{1}{1-\sigma}} \\ \frac{L_0}{z_1^2} \left\{ \left[\tilde{w} - \left(\frac{z_0}{\tilde{\eta} L_0} \right)^{1-\sigma} \right] \frac{\partial \tilde{\eta}}{\partial B} + (1-\sigma) \left(\frac{z_0}{\tilde{\eta} L_0} \right)^{1-\sigma} \frac{\partial \tilde{\eta}}{\partial B} + \tilde{\eta} \frac{\partial \tilde{w}}{\partial B} \right\} & \text{otherwise} \end{cases}$$

Consequently, $\partial^2 \tilde{u}^B(z_1, L_0)/(\partial B \partial z_1) \geq 0$, and thus $\partial^2(w_1 l_1)/(\partial B \partial z_1) \leq 0$ when $(z_1, L_0) \in \mathcal{S}^{\mathcal{H}} \cup \mathcal{S}^{\mathcal{M}}$.

One can show that $\partial^2 w^G/(\partial B \partial z_1) > 0$. From (16), we know that $\partial w^G/\partial B > 0$, and from

Proposition 1, we know that $\partial w^G/\partial z_0 < 0$. Combining these three inequalities, it follows that $\partial^2(w_1 l_1)/(\partial B \partial z_1) < 0$. Consequently, $z^S = z^G(z_1, L_0, B)$ satisfies the requirement on z^S. ■

B Dynamic Version of the Theoretical Model

In this appendix, we present a dynamic version of the model defined in the text. We consider a two period version of the game in which the union is forward-looking, and thus internalizes the $t = 0$ equilibrium labor outcome, L_0, on the equilibrium outcome at $t = 1$.[29] Given the two period nature of the model, the equilibrium defined in Proposition 1 corresponds to the outcome at the terminal period. We now subscript the sets defined in Proposition 1 with 1, so that we have ${\mathcal{S}^{\mathcal{H}}}_1, {\mathcal{S}^{\mathcal{M}}}_1$ and ${\mathcal{S}^{\mathcal{L}}}_1$ correspond to the state space partition at $t = 1$. We first present a Corollary to Proposition 1 that will aid in understanding the dynamic case. This Corollary examines the sign of the cross-partial $\partial^2 w_1/(\partial L_0 \partial B)$, which captures how the marginal effect of an additional union member on the wage, $\partial w_1/\partial L_0$, changes with bargaining power, B.

Corollary 1. *If $(z_1, L_0) \in {\mathcal{S}^{\mathcal{H}}}_1 \cup {\mathcal{S}^{\mathcal{L}}}_1$, then $\partial^2 w_1/(\partial L_0 \partial B) \geq 0$ in equilibrium. If $(z_1, L_0) \in {\mathcal{S}^{\mathcal{M}}}_1$, then $\partial^2 w_1/(\partial L_0 \partial B)$ can be either positive or negative.*

Proof. It is straightforward to show that

$$\frac{\partial^2 w_1}{\partial L_0 \partial B} = \begin{cases} -\underline{w}^{\frac{1}{1-\sigma}} \frac{\partial^2 \tilde{u}}{\partial L_0 \partial B} & \text{if } (z_1, L_0) \in {\mathcal{S}^{\mathcal{H}}}_1 \\ \frac{z_1}{L_0^2} \frac{\partial \tilde{u}}{\partial B} - \frac{z_1}{L_0} \frac{\partial^2 \tilde{u}}{\partial L_0 \partial B} & \text{if } (z_1, L_0) \in {\mathcal{S}^{\mathcal{M}}}_1 \\ \dfrac{\frac{\partial \tilde{w}}{\partial B} \frac{\sigma(\kappa-\gamma)}{\gamma} \left(\frac{w_1^G}{\tilde{w}} \right)^{\frac{1}{1-\sigma}}}{\left[\frac{\gamma + \sigma(\kappa-\gamma)}{\gamma} \right] - \left(\frac{w_1^G}{\tilde{w}} \right)^{\frac{\sigma}{1-\sigma}}} & \text{if } (z_1, L_0) \in {\mathcal{S}^{\mathcal{L}}}_1 \end{cases} \tag{17}$$

Moreover, $\partial^2 \tilde{u}/\partial L_0 \partial B < 0$ if $\tilde{u}^B(z_1, L_0)$ takes functional form (2b), and $\partial^2 \tilde{u}/\partial L_0 \partial B = 0$ if $\tilde{u}^B(z_1, L_0)$ takes functional form (2a). Consequently, when $(z_1, L_0) \in {\mathcal{S}^{\mathcal{H}}}_1$, then $\partial^2 w_1/(\partial L_0 \partial B) \geq 0$, with strict inequality when $\tilde{u}^B(z_1, L_0)$ takes functional form (2b).

Note that $\partial \tilde{u}/\partial B < 0$, regardless of whether $\tilde{u}^B(z_1, L_0)$ takes functional form (2b) or (2a). Consequently, when $(z_1, L_0) \in {\mathcal{S}^{\mathcal{M}}}_1$ and $\tilde{u}^B(z_1, L_0)$ takes functional form (2a), then $\partial^2 w_1/(\partial L_0 \partial B) < 0$.

Finally, when $(z_1, L_0) \in {\mathcal{S}^{\mathcal{L}}}_1$, then $\partial^2 w_1/(\partial L_0 \partial B) > 0$. □

[29]Deriving general results for this case has been difficult because of the friction imposed in the union's utility function, $\eta(L_t; L_{t-1})$.

When the city is in the “high” or “low” region, then the marginal benefit of an additional union member is *increasing in* B. Consequently, at $t = 0$, unions with higher B may have an incentive to push for larger membership in order to procure a higher wage at $t = 1$. Cities in the constant region, however, can exhibit either positive or negative $\partial^2 w_1/(\partial L_0 \partial B)$.

Turning to the initial period, we now impose two additional requirements on the model:

1. $\overline{z} := \sup \mathcal{Z} < \infty$
2. $L_t^* \in \{L \in \mathcal{L} \mid L \leq L_{-1}\}$

The first restriction places a finite upper bound on the population space. The second restriction imposes that the union's campaign support offer (w_t^*, L_t^*, s_t^*) must be such that $L_t^* \leq L_{t-1}$, so that the union cannot demand that the politician hire more than the current status quo number of employees. This assumption is made to ensure that there exists a nonempty subset of $\mathcal{Z} \times \mathcal{L}$ such that the union size decreases in equilibrium $(L_0 < L_{-1})$. Without such a restriction, employment would always increase at $t = 0$, regardless of the size of L_{-1} relative to z_0.[30] Consequently, if the union wants to induce the politician to hire $L_0 > L_{-1}$, it must offer a low w_0^* so that the politician finds it optimal to select a larger employment size.

In what follows, the phrase “$t = 0$ union” will refer to the initial L_{-1} workers who are unionized at $t = 0$, whereas “$t = 1$ union” will refer to L_0. In the first subsection, we will discuss the union's expected value function, $EV : \mathcal{L}^2 \to \Re$, that maps initial union size, L_{-1}, and the $t = 0$ employment decision, L_0, into expected utility for the status quo union at $t = 0$. As is discussed, EV depends nontrivially on L_{-1}. The following subsection then analyzes the dynamic equilibrium in the case of declining cities, while the final subsection studies growing cities.

B.1 Initial Union's Value Function

The value function of the initial L_{-1} union members can take one of two functional forms. First, if the entire union at $t = 1$ is comprised of members of L_{-1}, so that $L_0 \leq L_{-1}$, then the inter-temporal preferences of the union at $t = 0$ and $t = 1$ are aligned. In this instance, there is no time-inconsistency, as the equilibrium strategy played by the union at $t = 1$ will be identical to the optimal strategy the $t = 0$ union would commit to.

[30] There exist other assumptions that guarantee such a set exists. For instance, placing an upper bound on the union's contract, $(\tilde{\eta}, \tilde{w})$, in conjunction with a restriction on the relationship between κ and γ, are also sufficient.

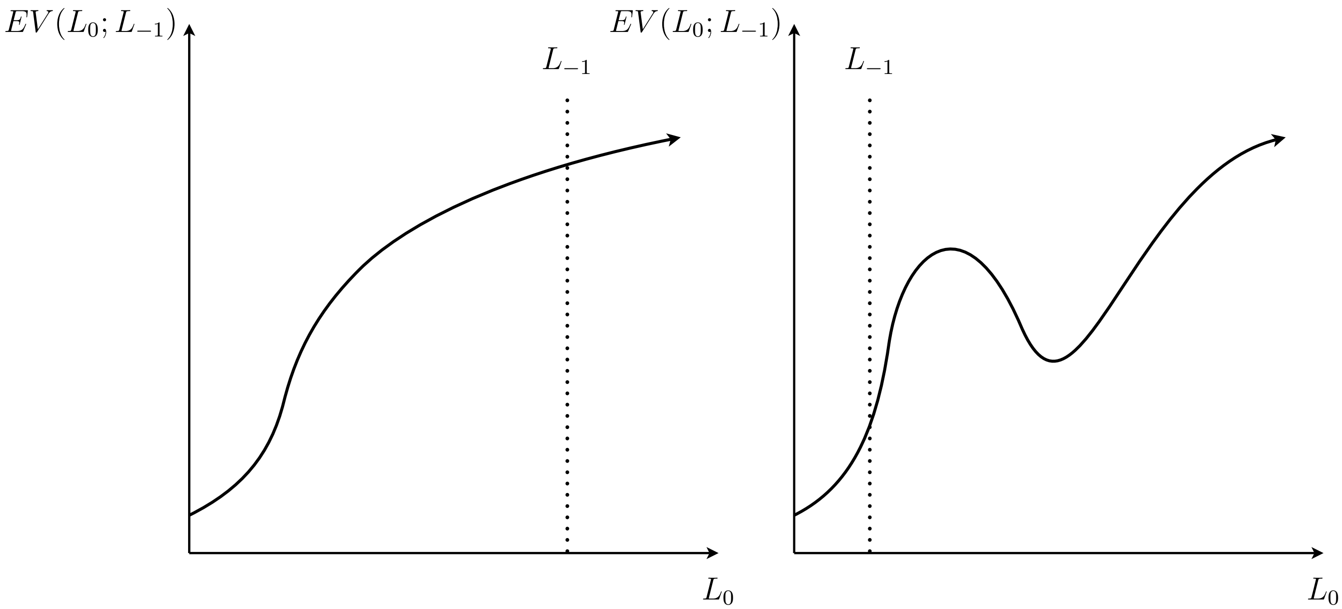

Figure 2: EV, Conditional on Initial Union Size, L_{-1}

If, on the other hand, $L_0 > L_{-1}$ so that the union grows in membership, then there is a discrepancy between the union's preferences at $t = 0$ and $t = 1$: At $t = 0$, the union cares only about maximizing the utility of the initial L_{-1} members, whereas at $t = 1$, the union maximizes the utility of the L_0 members, thus potentially distorting the $t = 0$ union's optimal strategy. Mathematically, EV is defined as

$$EV(L_0; L_{-1}) = \begin{cases} \frac{L_0}{L_{-1}} \int_{\underline{z}}^{\overline{z}} v\Big(w_1^*(z, L_0, B), L_1^*(z, L_0, B), s_1^*(z, L_0, B); L_0\Big) dG(z) & \text{if } L_0 \leq L_{-1} \\ \int_{\underline{z}}^{\overline{z}} \Big\{\eta\Big(L_1^*(z, L_0, B); L_{-1}\Big)\Big[w_1^*(z, L_0, B) - \underline{w}\Big] - \frac{\gamma}{\kappa} s_1^*(z, L_0, B)^\kappa\Big\} dG(z) & \text{if } L_0 > L_{-1} \end{cases}$$

When $L_0 = L_{-1}$, the resulting value function is $\int_{\mathcal{Z}} v(\cdot)dG(z)$. When union size diminishes, so that $L_0 < L_{-1}$, then $\int_{\mathcal{Z}} v(\cdot)dG(z)$ is multiplied by L_0/L_{-1}, in effect just shifting preferences by a scalar. When $L_0 > L_{-1}$, however, preferences change in a nonlinear manner: As was discussed last section, when $t = 1$, the L_0 union maximizes the expected welfare of its membership, and thus $\eta\Big(L_1; L_0\Big)$ appears in its objective. When $t = 0$, on the other hand, the L_{-1} union cares only about the welfare of its L_{-1} members, and thus $\eta\Big(L_1; L_{-1}\Big)$ appears in the value function.

In Figure 2, we plot the typical shape of EV, conditional on different initial values of L_{-1}. The left graph plots EV when L_{-1} is "large," so that the probability that $(z_1, L_{-1}) \in \mathcal{S}^{\mathcal{H}}{}_1$ is high. Note that EV is *increasing* in L_0: When $L_0 \leq L_{-1}$, $EV_L > 0$ because a larger L_0 translates into more initial union members L_{-1} remaining employed. Conversely, the right figure depicts EV for a "small" value of L_{-1}, when the probability that $(z_1, L_{-1}) \in \mathcal{S}^{\mathcal{L}}{}_1$ is high. For low values of L_0, $EV_L > 0$ as more union members translate into greater political influence for the union at $t = 1$. For intermediate values of L_0, however, $EV_L < 0$. The intuition for this is grounded in Proposition 1: When $(z_1, L_0) \in \mathcal{S}^{\mathcal{M}}{}_1$, then $\partial w_1 / \partial L_0 < 0$ for some values of (z_1, L_0). Consequently, the union representing L_{-1} may actually want to *restrict* employment L_0, as greater members may drive down wages in future negotiations with the politician.

Clearly, EV need not be concave. In order to ensure existence of an interior solution, we adopt a technical assumptions on the parameters that ensures the concavity of EV for high values of L_0, when there is a high probability that $(z_1, L_0) \in \mathcal{S}^{\mathcal{H}}{}_1$. Concavity at high L_0 is necessary to ensure existence of a solution, as otherwise $L_0^* \to \sup \mathcal{L}$, which would be unrealistic and uninteresting. This assumption essentially places a restriction between the parameters entering the union's campaign support cost, γ and κ, the public's preference parameter σ and the rehiring requirement from the contract, $\tilde{\eta}$.

A1: $\frac{\gamma(\kappa-1)}{\kappa-\gamma} \in \left[(1-\sigma)\tilde{\eta}^{\sigma}\frac{\kappa}{1-\gamma}, \sigma\right]$

This assumption is somewhat restrictive: When $(\sigma, \kappa, \tilde{\eta}) = (.9, 3, .3)$, for instance, then **A1** requires that $\gamma \in [0.174, 0.877]$, while when $(\sigma, \kappa, \tilde{\eta}) = (.75, 2, .2)$, then **A1** implies $\gamma \in [0.399, 0.751]$.

B.2 Equilibrium in Declining Cities

Similar to Proposition 1, Proposition 2 proves existence of a subset $\mathcal{S}^{\mathcal{H}}{}_0 \subset \mathcal{Z} \times \mathcal{L}$ of initial states (z_0, L_{-1}) that are interpreted as states of "declining" cities. Cities with $(z_0, L_{-1}) \in \mathcal{S}^{\mathcal{H}}{}_0$ exhibit the following characteristics: First, in expectation, such cities will decline in population at $t = 1$, and second, the cities' workers-per-capita are in excess at $t = 0$.

Given that the probability of $(z_1, L_{-1}) \in \mathcal{S}^{\mathcal{H}}{}_1$ is high, EV will look similar to the left graph in figure NUMBER. As is expected, the introduction of EV into the union's $t = 0$ objective will increase the union's incentive to push for a higher employment, L_0. Interestingly, however, Proposition 2 also indicates that *equilibrium* L_0 *is increasing in bargaining power*. Moreover, this effect manifests itself not because it is easier for unions with high B to influence the politician at $t = 0$, but rather because the inter-temporal marginal benefit of additional union members is increasing in B (see Corollary 1).

Proposition 2. *There exists a nonempty set* $\mathcal{S}^{\mathcal{H}}{}_0 \subset \mathcal{Z} \times \mathcal{L}$ *such that if* $(z_0, L_{-1}) \in \mathcal{S}^{\mathcal{H}}{}_0$, *then* $\frac{\partial L_0}{\partial B} > 0$ *in equilibrium.*

Proof. For this proof, $V : \mathcal{Z} \times \mathcal{L}^2 \to \Re$ is simply EV evaluated at a particular point $z \in \mathcal{Z}$. The proof consists of three steps: First, we show several properties of EV.

Second, we define ${\mathcal{S}^{\mathcal{H}}}_0 \subset \mathcal{Z} \times \mathcal{L}$ and formulate a strategy $(w_0^D, L_0^D) : {\mathcal{S}^{\mathcal{H}}}_0 \times \mathcal{B} \to \mathcal{W} \times \mathcal{L}$. Finally, we verify $(w_0^*, L_0^*) = (w_0^D, L_0^D)$ if $(z_0, L_{-1}) \in {\mathcal{S}^{\mathcal{H}}}_0$.

Step 1: Properties of EV

Property 1: $EV_L > 0$ when $L_0 < L_{-1}$:

$$V_L(z_1, L_0; L_{-1}) = \begin{cases} \frac{1}{L_{-1}}\left[\frac{\gamma\kappa-1}{\kappa}\left(\frac{z_1}{L_0}\right)^{\frac{\kappa(1-\gamma)}{(\kappa-\gamma)}} - z_1 \tilde{u}_L^B(z_1, L_0)\right] & \text{if } (z_1, L_0) \in {\mathcal{S}^{\mathcal{H}}}_1 \\ \frac{1}{L_{-1}}\left[\frac{\gamma\kappa-1}{\kappa}\left(\frac{z_1}{L_0}\right)^{\frac{\kappa(1-\gamma)}{(\kappa-\gamma)}} - z_1 \tilde{u}_L^B(z_1, L_0) + \left\{\left(\frac{z_1}{L_0}\right)^{1-\sigma} - \underline{w}\right\}\right] & \text{if } (z_1, L_0) \in {\mathcal{S}^{\mathcal{M}}}_1 \\ \frac{1}{L_{-1}}\left[\left(w_1^*(z_1, L_0) - \underline{w}\right) + \frac{\gamma(\kappa-1)}{\kappa} s_1^*(\cdot)^\kappa\right] & \text{if } (z_1, L_0) \in {\mathcal{S}^{\mathcal{L}}}_1 \end{cases}$$

Since $\tilde{u}_L^B(z_1, L_0) \leq 0$ for all (z_1, L_0, B), then $V_L > 0$ for all (z_1, L_0, L_{-1}). Therefore, $EV_L > 0$ when $L_0 < L_{-1}$.

Property 2: $EV_{LB} \geq 0$ when $L_0 < L_{-1}$:

$$V_{LB}(z_1, L_0; L_{-1}) = \begin{cases} -\frac{z_1}{L_{-1}} \tilde{u}_{LB}^B(z_1, L_0) & \text{if } (z_1, L_0) \in {\mathcal{S}^{\mathcal{H}}}_1 \cup {\mathcal{S}^{\mathcal{M}}}_1 \\ \frac{1}{L_{-1}}\left[\frac{\partial w_1^*}{\partial B} + \left(\frac{1-\sigma}{\sigma}\right)\left(\frac{\frac{\partial w_1^*}{\partial B} - \frac{\partial \tilde{w}}{\partial B}\left(\frac{w_1^*}{\tilde{w}}\right)^{-\frac{1}{1-\sigma}}}{\left(\frac{w_1^*}{\tilde{w}}\right)^{-\frac{\sigma}{1-\sigma}} - 1}\right)\right] & \text{if } (z_1, L_0) \in {\mathcal{S}^{\mathcal{L}}}_1 \end{cases}$$

Since $\tilde{u}_{LB}^B(z_1, L_0) \leq 0$, then $V_{LB} \geq 0$ when $(z_1, L_0) \in {\mathcal{S}^{\mathcal{H}}}_1 \cup {\mathcal{S}^{\mathcal{M}}}_1$, with strict inequality when $\tilde{\eta} L_0 > \phi(\tilde{w}; z_1)$. Moreover, by substituting $\partial w_1^*/\partial B$ from (17) into V_{LB}, we see that

$$\frac{\partial w_1^*}{\partial B} - \frac{\partial \tilde{w}}{\partial B}\left(\frac{w_1^*}{\tilde{w}}\right)^{-\frac{1}{1-\sigma}} > 0$$

which implies that $V_{LB} > 0$ when $(z_1, L_0) \in {\mathcal{S}^{\mathcal{L}}}_1$. Consequently, $EV_{LB} \geq 0$ when $L_0 < L_{-1}$, with strict equality if $(z_1, L_0) \in {\mathcal{S}^{\mathcal{L}}}_1$ or $\tilde{\eta} L_0 > \phi(\tilde{w}; z_1)$ occurs with positive probability.

Property 3: $\exists \hat{L} : \mathcal{L} \to \mathcal{L}$ such that $\forall L > \hat{L}(L_{-1})$, $EV_{LL}(L; L_{-1}) < 0$: First, note that $\forall L_{-1} \in \mathcal{L}$, we have $V_{LL}(z_1, L_0; L_{-1}) < 0$ when $(z_1, L_0) \in {\mathcal{S}^{\mathcal{H}}}_1$ by **A1**. Moreover, if $L_0 < L_{-1}$ and $(z_1, L_0) \in {\mathcal{S}^{\mathcal{M}}}_1$, then

$$V_{LL}(z_1, L_0; L_{-1}) \;=\; \frac{1}{L_{-1} L_0}\left[-\frac{\gamma(1-\gamma)\kappa - 1)}{\kappa(\kappa-\gamma)}\left(\frac{z_1}{L_0}\right)^{\frac{\kappa(1-\gamma)}{\kappa-\gamma}} - z_1 L_0 \tilde{u}_{LL}^B(z_1, L_0) - (1-\sigma)\left(\frac{z_1}{L_0}\right)^{1-\sigma}\right] \;<\; 0$$

by **A1**. When $(z_1, L_0) \in {\mathcal{S}^{\mathcal{L}}}_1$, however, $V_{LL} > 0$.

Second, note that $\exists L' \in \mathcal{L}$ large enough such that $z^D(L') \geq \overline{z}$, which implies that $EV_{LL}(L'; L_{-1}) < 0$, and $\exists L'' \in \mathcal{L}$ small enough such that $z^G(L') \leq \underline{z}$, which implies that $EV_{LL}(L''; L_{-1}) > 0$. Since EV_{LL} is continuous by virtue of V_{LL} and g being continuous, then

$$\hat{L}(L_{-1}) = \max\left\{L \in \mathcal{L} \mid \frac{\partial^2 EV(L; L_{-1})}{\partial L^2} \geq 0\right\}$$

is well defined for all $L_{-1} \in \mathcal{L}$.

Step 2: Defining $\mathcal{S}^{\mathcal{H}}{}_0$ and (w_0^D, L_0^D) First, we define[31]

$$\begin{aligned}\mathcal{S}^{\mathcal{H}}{}_0 \;:=\; & \Big\{(z_0, L_{-1}) \in \mathcal{Z} \times (\hat{L}(L_{-1}), \sup \mathcal{L}) \mid \\ & z_0 \in \Big[\hat{L}(L_{-1})\big[\underline{w} - \hat{L}(L_{-1})EV_L(\hat{L}(L_{-1}); L_{-1})\big]^{\frac{1}{1-\sigma}}, \quad L_{-1}\big[\underline{w} - L_{-1}EV_L(L_{-1}; L_{-1})\big]^{\frac{1}{1-\sigma}}\Big]\Big\}\end{aligned}$$

At $t = 0$, the union solves

$$\max_{w,L} \eta(L; L_{-1})[w - \underline{w}] + EV(L; L_{-1}) - \frac{\gamma}{\kappa} s(w, L; z_0, L_{-1})^{\kappa}$$

First, consider the case where $(z_0, L_{-1}) \in \mathcal{S}^{\mathcal{H}}{}_0$ is such that (w_0^D, L_0^D) maps to $\operatorname{int}\{\mathcal{W} \times [\hat{L}(L_{-1}), L_{-1}]\}$. Dividing the FOCs, we find

$$\frac{L_0^D}{[w_0^D - \underline{w}] + L_{-1}EV_L(L_0^D)} = \frac{s_w}{s_L} = \frac{L_0^D/z_0}{w_0^D/z_0 - (L_0^D)^{\sigma-1} z_0^{-\sigma}}$$

which yields the implicit function

$$L_{-1}EV_L(L_0^D; L_{-1}) + \left(\frac{z_0}{L_0^D}\right)^{1-\sigma} = \underline{w} \tag{18}$$

and

$$w_0^D = \left(\frac{1}{\sigma}\right)\left(\frac{z_0}{L_0^D}\right)^{1-\sigma} + \left(\frac{z_0}{L_0^D}\right)\left[\left(\frac{L_{-1}}{z_0}\right)^{\frac{\gamma(\kappa-1)}{\kappa-\gamma}} - \tilde{u}^B(z_0, L_{-1})\right] \tag{19}$$

By virtue of $(z_0, L_{-1}) \in \mathcal{S}^{\mathcal{H}}{}_0$, we have $L_0^D \in \big(\hat{L}(L_{-1}), L_{-1}\big)$, and thus $V_{LL}(L_0^D; L_{-1}) < 0$. Consequently, an identical argument as the proof of Proposition 1 can be employed to show that the objective is concave at (z_0, L_{-1}). Taking the total derivative of (18) with

[31]When $L_{-1} > \hat{L}(L_{-1})$, then $EV_L(L_{-1}; L_{-1}) < EV_L(\hat{L}(L_{-1}); L_{-1})$ by concavity of EV, and thus

$$\Big[\hat{L}(L_{-1})\big[\underline{w} - EV_L(\hat{L}(L_{-1}); L_{-1})\big]^{\frac{1}{1-\sigma}}, \quad L_{-1}\big[\underline{w} - EV_L(L_{-1}; L_{-1})\big]^{\frac{1}{1-\sigma}}\Big] \neq \emptyset$$

respect to B, we find that

$$\frac{\partial L_0^D}{\partial B}\left(EV_{LL}(L_0^D) - (1-\sigma)z_0^{1-\sigma}(L_0^D)^{\sigma-2}\right) + EV_{LB}(L_0^D; L_{-1}) = 0$$

Since $EV_{LB}(L_0^D; L_{-1}) \geq 0$ and $EV_{LL}(L_0^D) - (1-\sigma)z_0^{1-\sigma}(L_0^D)^{\sigma-2} < 0$, then $\partial L_0^*/\partial B \geq 0$.

Next, consider the case when the $w \geq \underline{w}$ binds, so that w_0^D defined above is less than $\underline{w}$. In this instance, $w_0^D = \underline{w}$ in equilibrium and the union solves

$$\max_L EV(L; L_{-1}) - \frac{\gamma}{\kappa}s(\underline{w}, L; z_0, L_{-1})^\kappa$$

which is strictly concave. Consequently, L_0^D is defined implicitly by

$$EV_L(L_0^D; L_{-1}) = -u_L\left(\frac{z_0}{L_{-1}}\right)^\gamma s^{\kappa-\gamma} \tag{20}$$

Taking the total derivative with respect to B, we see that

$$\frac{\partial L_0^D}{\partial B}\Big[SOC\ L\Big] + \left[EV_{LB}(L_0^D; L_{-1}) + \tilde{u}_B u_L\left(\frac{z_0}{L_{-1}}\right)^\gamma s^{\kappa-\gamma}\right] = 0$$

As state, the SOC is strictly negative, whereas the remaining two terms are positive: $E_{LB} \geq 0$, while $\tilde{u}_B, u_L < 0$. Consequently, $\partial L_0^*/\partial B > 0$.

Step 3: $(w_0^*, L_0^*) = (w_0^D, L_0^D)$ **if** $(z_0, L_{-1}) \in \mathcal{S}^{\mathcal{H}}{}_0$ Let $(z_0, L_{-1}) \in \mathcal{S}^{\mathcal{H}}{}_0$. If $(z_0, L_{-1}) \in \mathcal{S}^{\mathcal{H}}{}_0$ is such that w_0^D is constrained, then $(w_0^*, L_0^*) = (w_0^D, L_0^D)$ from strict concavity.

If $(z_0, L_{-1}) \in \mathcal{S}^{\mathcal{H}}{}_0$ is such that $w_0^D > \underline{w}$, then we must verify (w_0^D, L_0^D) is the global maximum. Since (w_0^D, L_0^D) is the only critical value in $\text{int}\{\mathcal{W} \times [\hat{L}(L_{-1}), L_{-1}]\}$, we must check alternative strategies (w', L') where (1) $w' = \underline{w}$, (2) $L' \leq \phi(w'; z_0)$, (3) $L' = L_{-1}$ or (4) $w' \to \infty$.

First, consider (w', L') such that $w' = \underline{w}$. The optimal L', when $w' = \underline{w}$, is as defined in (20). Since $w' < w_0^D$, then $L' > L_0^D$, and thus

$$L_{-1}EV_L(L'; L_{-1}) + \left(\frac{z_0}{L'}\right)^{1-\sigma} < \underline{w}$$

where the strict inequality comes from the definition of L_0^D and the RHS is decreasing in

L. Equation (20), evaluated at (w', L') and rearranged, yields

$$\begin{aligned} 1 &> \frac{L_{-1} EV_L(L'; L_{-1})}{\underline{w} - \left(\frac{z_0}{L'}\right)^{1-\sigma}} \\ &= \left(\frac{L_{-1}}{z_0}\right)^{1-\gamma} s^{\kappa-\gamma} \end{aligned}$$

The w FOC, evaluated at (w', L'), yields

$$\frac{L'}{L_{-1}} \left[1 - \left(\frac{L_{-1}}{z_0}\right)^{1-\gamma} s^{\kappa-\gamma}\right] > 0$$

and thus (w', L') cannot be the global maximum.

Second, consider (w', L') to be the optimal strategy when $L' = L_{-1}$. Then the w FOC must be satisfied with equality, so that

$$1 = \frac{L'}{z_0} \left(\frac{z_0}{L_{-1}}\right)^{\gamma} s^{\kappa-\gamma}$$

The FOC wrt L, evaluated at (w', L') and substituting in the previous equation, is

$$\begin{aligned} \frac{1}{L_{-1}}[w' - \underline{w}] + EV_L(L'; L_{-1}) - \frac{1}{L'}\left[w' - \left(\frac{z_0}{L'}\right)^{1-\sigma}\right] &= EV_L(L'; L_{-1}) - \frac{1}{L_{-1}}\left[\underline{w} - \left(\frac{z_0}{L_{-1}}\right)^{1-\sigma}\right] \\ &< 0 \end{aligned}$$

where the strict inequality follows from $(z_0, L_{-1}) \in \mathcal{S}^{\mathcal{H}}{}_0$.

Finally, consider either (w', L') where $L' \leq \phi(w'; z_0)$ or $w' \to \infty$. The arguments for these cases are identical to the proof of Proposition 1.

Consequently, $(w_0^*, L_0^*) = (w_0^D, L_0^D)$ if $(z_0, L_{-1}) \in \mathcal{S}^{\mathcal{H}}{}_0$. □

This proposition may help explain two phenomenon regarding public sector unions: First, the wage premium $w^* - \underline{w}$ may be lower than in the private sector because workers trade off the immediate benefit of higher wages for the dynamic insurance provided by a larger union size: In the event of a future decline in the demand for a public good, a larger union will be more adept at extracting rents from future policy-makers via campaign contributions.

Corollary 2. *If $(z_0, L_{-1}) \in \mathcal{S}^{\mathcal{H}}{}_0$, then the equilibrium wage premium $w^* - \underline{w}$ is smaller in the first period of the dynamic game than in the static game.*

Proof. Let (w_0^*, L_0^*) $[(w_s^*, L_s^*)]$ denote the union's $t = 0$ equilibrium wage/labor demand in the dynamic [static] game. If (z_0, L_{-1}) is such that $w_0^* = \underline{w}$, then the result is clear since $w_s^* > \tilde{w} \geq \underline{w}$. Otherwise, since $L_0^* > L_s^*$, then from (18) it follows that $w_0^* < w_s^*$. $\square$

B.3 Equilibrium in Growing Cities

We have been unable to prove a result analogous to Proposition 2 for cities that are expected to grow from $t = 0$ to $t = 1$. The reason is that it has been difficult to define a set $\mathcal{S}_0^L \subset \mathcal{Z} \times \mathcal{L}$ for which (i) in expectation, $z_1 > z_0$, and (ii) the sign of $EV_{LB}(\cdot)$ can be established. Given that we have not established existence of such a set, we have not been able to prove a definitive result. One partial result is the following:

Claim: *If* (z_0, L_{-1}) *is such that* $EV_{LB}(\cdot) < 0$ *in equilibrium, then* $\partial L_0^*/\partial B < 0$ *and* $\partial w_0^*/\partial B > 0$.

This claim states that if the marginal future benefit of L_0 is decreasing in B, then $\partial L_0^*/\partial B < 0$. Numerical approximations, using a variety of parameter values, have shown that, for states for which L_{-1} is "low" relative to z_0, the comparative static $\partial L_0^*/\partial B$ appears to be negative. Nevertheless, a proof establishing existence has been elusive to date.

Table 1: Summary Statistics

Variables	Mean	Median	10th	90th
Population	107813	50756	28254	176572
Gross Population Growth Rate	1.123	1.056	0.914	1.364
Median Household Income	54834	51108	38955	75074
Unemployment Rate	0.055	0.05	0.026	0.091
Poverty Rate	0.117	0.11	0.038	0.207
% College Grad	0.198	0.169	0.081	0.36
% School Age (5-17)	0.197	0.191	0.143	0.259
% Elderly (65+)	0.118	0.116	0.059	0.17
% Black	0.119	0.046	0.004	0.343
% White	0.806	0.862	0.539	0.982
Administration Wage Bill (per capita)	45.84	39.26	19.71	77.91
Fire Wage Bill (per capita)	84.02	78.29	43.10	133.48
Highway Wage Bill (per capita)	31.10	27.58	13.31	53.58
Parks Wage Bill (per capita)	24.23	20.57	7.33	44.96
Police Wage Bill (per capita)	115.12	104.02	62.87	176.52

Table 2: Persistence Regressions

	Administration		Fire		Highway		Parks		Police	
VARIABLES	FE	SAR	FE	SAR	FE	SAR	FE	SAR	FE	SAR
$\log(pop/pop_{-1})$	0.004	-0.038	-0.215***	-0.233***	-0.0478	-0.028	-0.0471	-0.053	-0.114**	-0.218***
	0.935	0.492	0.000	0.000	0.509	0.731	0.485	0.495	0.047	0.000
$\log(pop)$	-0.151***	-0.126***	-0.133***	-0.158***	-0.044	-0.102	0.064	-0.007	-0.227***	-0.125***
	0.003	0.004	0.002	0.000	0.503	0.103	0.346	0.907	0.000	0.000
$\log(income)$	0.300***	0.296***	0.401***	0.306***	0.600***	0.390***	0.480***	0.446***	0.433***	0.302***
	0.006	0.009	0.000	0.000	0.000	0.017	0.002	0.004	0.000	0.000
unemployment	-0.806**	-1.347***	0.190	-0.014	-0.935*	-0.942*	-0.886	-1.583***	-0.299	-0.353*
	0.044	0.000	0.506	0.953	0.100	0.077	0.112	0.002	0.236	0.089
poverty	-0.752**	-0.703***	-0.592**	-0.555***	0.136	-0.200	-0.477	-0.665	-0.641***	-0.635***
	0.017	0.035	0.015	0.010	0.694	0.678	0.214	0.155	0.001	0.001
college	0.197	0.128	0.062	-0.036	0.029	0.112	-0.010	-0.184	-0.257*	-0.192*
	0.291	0.521	0.699	0.782	0.906	0.695	0.969	0.510	0.070	0.094
school_age	-0.644	0.130	-0.613	-0.384	1.987***	2.460***	-1.298*	-0.496	-0.333	-0.212
	0.216	0.786	0.212	0.232	0.007	0.000	0.077	0.456	0.382	0.443
old	0.342	0.329	0.862*	0.021	2.692***	2.577***	1.621***	1.991***	-0.207	-0.584***
	0.409	0.375	0.078	0.931	0.000	0.000	0.003	0.000	0.717	0.006
black	0.554***	0.492***	0.823***	0.355***	0.738**	0.271	0.968***	0.402	0.408***	0.209**
	0.005	0.003	0.000	0.003	0.020	0.280	0.001	0.106	0.002	0.036
white	-0.013	0.023	0.216	-0.070	0.389*	0.286*	0.860***	0.549***	-0.108	-0.068
	0.928	0.846	0.175	0.398	0.064	0.085	0.000	0.001	0.332	0.329
boundary	-0.028	-0.032	0.006	0.037*	-0.042	-0.059	-0.046	-0.020	-0.003	0.011
	0.436	0.331	0.790	0.071	0.410	0.198	0.316	0.651	0.936	0.534
ρ		0.528***		0.760***		0.477***		0.591***		0.782***
		0.000		0.000		0.000		0.000		0.000
Observations	3310	2436	3056	2292	3202	2288	3060	2136	3241	2404
Number of cities	909	609	852	573	890	572	868	534	895	601

Table 3: Interaction Between Union Indicator, Population Change

	Fire		Highway		Police	
VARIABLES	FE	SAR	FE	SAR	FE	SAR
B	0.051***	0.048***	0.105**	0.117***	0.036**	0.031**
	0.017	0.000	0.046	0.001	0.017	0.011
$B \times \log(pop/pop_{-1})$	-0.122	-0.134*	-0.207	-0.140	0.025	0.067
	0.091	0.052	0.203	0.455	0.078	0.309
$\log(pop/pop_{-1})$	-0.198***	-0.157***	-0.230**	-0.358***	-0.146***	-0.216***
	0.053	0.004	0.111	0.008	0.048	0.000
$\log(pop)$	-0.233***	-0.133***	0.146	0.005	-0.314***	-0.247***
	0.072	0.004	0.145	0.966	0.057	0.000
$\log(income)$	0.321***	0.247***	0.619**	0.400	0.294***	0.215***
	0.117	0.003	0.261	0.067	0.111	0.000
unemployment	-0.339	-0.368	-2.364***	-2.076***	-0.090	-0.356
	0.385	0.154	0.908	0.002	0.353	0.147
poverty	-0.225	0.173	0.666	-0.827	-0.464	-0.195
	0.358	0.553	0.955	0.278	0.328	0.481
college	-0.160	0.154	-0.013	0.065	-0.181	0.047
	0.167	0.231	0.385	0.849	0.164	0.659
school_age	-2.197***	-2.410***	1.709	1.422	-1.869***	-1.962***
	0.575	0.000	1.278	0.191	0.454	0.000
old	-0.135	-0.205	1.847	1.056	-1.020**	-0.982***
	0.638	0.546	1.286	0.225	0.423	0.002
black	0.339	-0.052	-0.389	-1.196***	0.265	-0.496***
	0.365	0.735	0.866	0.003	0.280	0.001
white	-0.170	0.239	-0.612	-1.396**	-0.322*	0.092
	0.237	0.343	0.539	0.034	0.167	0.698
boundary	0.053*	0.027	0.020	-0.024	0.033	0.018
	0.030	0.307	0.085	0.725	0.028	0.483
ρ		0.432***		0.035		0.560984**
		0.002		0.794		0.044
Observations	1554	1244	1632	1216	1692	1260
Number of id	777	622	816	608	846	630

Table 4: Interaction Between Union Indicator, Population Growth Indicators

	Fire		Highway		Police	
VARIABLES	FE	SAR	FE	SAR	FE	SAR
$q^{decline,B=1}$	0.0415**	0.074***	0.020	0.080	0.0498***	0.030*
	0.018	0.000	0.059	0.104	0.019	0.092
$q^{decline,B=0}$	-0.011	0.000	-0.110**	-0.045	0.021	-0.008
	0.020	0.996	0.051	0.330	0.022	0.668
$q^{normal,B=1}$	0.002	0.036**	0.033	0.099***	0.003	0.010
	0.016	0.031	0.055	0.018	0.016	0.533
$q^{grow,B=1}$	-0.005	-0.001	0.000	0.136*	0.003	0.050*
	0.030	0.956	0.086	0.078	0.026	0.055
$q^{grow,B=0}$	0.029	0.002	-0.057	0.008	-0.028	-0.049**
	0.023	0.928	0.053	0.879	0.017	0.014
$\log(pop/pop_{-1})$	-0.224***	-0.152**	-0.263**	-0.464***	-0.058	-0.144**
	0.054	0.014	0.121	0.004	0.039	0.015
$\log(pop)$	-0.256***	-0.135***	0.172	-0.015	-0.252***	-0.250***
	0.071	0.005	0.134	0.905	0.043	0.000
$\log(income)$	-0.089	0.251***	0.087	0.378**	-0.153*	0.202***
	0.092	0.003	0.250	0.086	0.083	0.003
unemployment	0.236	-0.316	-1.771*	-2.085***	0.223	-0.290
	0.278	0.224	0.916	0.002	0.279	0.240
poverty	-0.550*	0.159	0.307	-0.840	-0.334	-0.223
	0.286	0.584	0.916	0.271	0.291	0.418
college	0.085	0.167	0.199	0.074	0.074	0.044
	0.139	0.196	0.359	0.828	0.118	0.699
school_age	-1.223**	-2.426***	3.238***	1.404	-1.054***	-2.01***
	0.530	0.000	1.167	0.198	0.346	0.000
old	0.511	-0.254	3.071***	1.085	-0.273	-1.027***
	0.502	0.455	1.041	0.214	0.303	0.001
black	0.707**	0.242	-0.163	-1.390**	0.429**	0.068
	0.297	0.337	0.775	0.035	0.213	0.775
white	0.236	-0.054	-0.317	-1.195***	0.018	-0.519***
	0.169	0.726	0.495	0.003	0.132	0.000
boundary	0.0752***	0.024	0.012	-0.019	0.035	0.003
	0.023	0.399	0.088	0.799	0.023	0.907
ρ		0.434***		0.031		0.532***
		0.002		0.801		0.015
Observations	1556	1244	1636	1216	1692	1260
Number of id	778	622	818	608	846	630