My End of the Bargain: Exploring the Influence of Spatial Relationships on the Provisions in Collective Bargaining Agreements

Dan Goldhaber Lesley Lavery Roddy Theobald

Center for Education Data & Research

University of Washington Bothell

March 1, 2012

* *This research was made possible in part by grants from the Smith Richardson Foundation and the Bill and Melinda Gates Foundation, and has benefited from helpful comments from Katharine Strunk and Sean Reardon. We thank Yangru Fang and Dylan D'Entremont for excellent research assistance, and we thank the CBA coders without whom this project would not be possible: Rahn Amitai, Shijia Bian, Scott Bohmke, Stephanie Burns, Jonathan Humphrey, Angela Kim, Gregory Johnsen, Eric Lei, Hanqiao Li, Yi Li, Wanyu Liu, Xijia Lu, Alex McKay, Courtney Polich, Leah Staub, Annie Saurwein, Bifei Xie, Nancy Xu, Youngzee Yi, and Wenjun Yu. The statements made and views expressed in this paper do not necessarily reflect those of the University of Washington Bothell, the Smith Richardson Foundation, or the Bill and Melinda Gates Foundation. Any and all errors are solely the responsibility of the authors.*

The suggested citation for this working paper is:

Goldhaber, D., Lavery, L. & Theobald, R. (2012). My end of the bargain: Exploring the influence of spatial relationships on the provisions in collective bargaining agreements. CEDR Working Paper 2012-2. University of Washington, Seattle, WA.

ABSTRACT

A large literature on teacher collective bargaining describes the potential influence of provisions in teacher union contracts, but little is known about the determinants of these provisions. Using a unique dataset of every active teacher collective bargaining agreement in Washington state, we find that the proximity of two districts is a significant predictor of the similarity between the districts' CBAs. This suggests that spatial relationships play a major role in determining the provisions that appear in a district's CBA. These spatial relationships are more pronounced for different subsets of provisions, including a subset of high-profile provisions and provisions dealing with Association rights, benefits and leave procedures, hiring and transfer policies, and teacher workload agreements. Our findings suggest that policy makers seeking to understand the determinants of teacher collective bargaining provisions must consider the role that spatial relationships play in the process.

I. The Geography of Collective Bargaining Agreements

Collective bargaining agreements cover a wide array of school district rules and regulations and govern everything from hiring and compensation, to the policies that determine transfers between schools, evaluation, professional development, the promotion processes, grievance and termination. Recently, policymakers and pundits alike have pointed to CBAs, and particular CBA provisions (e.g. seniority-based job protections), as key inhibitors to effective school district operation and student achievement. For instance the federal government's Race to the Top grant competition incents states to make dramatic changes in teacher evaluation policies, changes that often must be negotiated as part of the collective bargaining process. And in response to recent budget woes the governors of two Midwestern states illegalized several public employee contract demands.

A large literature on teacher collective bargaining describes the potential influence of union contracts on school organization, finance, staffing and operations (Hannaway and Rotherham 2006). But we have little quantitative evidence about how the contract provisions in CBAs are actually determined. In this paper we focus on one potential determinant of the provisions in a district's CBA: the provisions in the CBAs of nearby districts. Specifically, we assess the influence of geographic and institutional proximity on bargaining outcomes using a unique dataset of every collective bargaining agreement governing policies during the 2010-11 school year in the state of Washington.

Our findings suggest that spatial relationships play a major role in determining the provisions that appear in a district's CBA. When we consider the full range of provisions that appear in teacher collective bargaining agreements (633 in all), we find that the proximity of two districts is a significant predictor of the similarity between the districts' CBAs. For example, holding other observable measures of similarity constant, CBAs of the average pair of districts within 25 miles of each other share 1.5% more provisions than other pairs of districts. These results are robust to various model specifications, and become more pronounced when we consider two subsets of provisions: a subset of provisions objectively derived through methods from Strunk and Reardon (forthcoming), and a subset of provisions subjectively derived by "cherry picking" high-profile provisions. Following Strunk and Reardon, we also calculate a measure of the underlying "restrictiveness" of each contract that utilizes all provisions in the contract. We find that district proximity is also a significant predictor of the similarity in the restrictiveness of two districts' contracts. Finally, we divide all provisions into eight categories and find that the influence of spatial relationships is particularly strong for provisions dealing with Association rights, benefits and leave procedures, hiring and transfer policies, and teacher workload agreements. These findings suggest that policy makers seeking to understand the determinants of teacher collective bargaining provisions must consider the role that spatial relationships play in the process.

II. Background

There are two fundamentally different reasons to expect a spatial relationship among CBA provisions or overall CBA restrictiveness. First, closely clustered districts are likely to share similar demographic, social, political and economic environments. Environmental similarities may explain noted contract similarities. But proximate districts may also have similar bargaining agreements if the bargain struck in one district has a causal impact on the negotiations and subsequent contracts in nearby districts. We might imagine that a "bargaining shock"—a departure from standard contract design which may arise from external factors or changes in personnel influential in the bargaining process—in one district changes the focus and scope of local bargaining and negotiations in neighboring districts. For example, a strike in Tacoma is likely to play up union power in that district and in neighboring locations if the strike is met with broad community support. Teacher Incentive Fund funding inspiring a new teacher evaluation system in Seattle may incite evaluation reform throughout the Puget Sound. Changes in union or district contract negotiation teams or key personnel (superintendent, union president, etc.) may also spur intra- and inter-district reform. And, if there is a spatial component to bargaining then statewide influence is perhaps more likely when districts are centrally located.[1]

Researchers from a variety of disciplines have described models that predict the nature of what is included in CBAs. Kochan and Wheeler (1975) argue that bargaining agreements are a function of environmental characteristics and the organizational characteristics of union and management. Environmental characteristics such as previously existing policies, voter characteristics and preferences, and a district's financial health may directly and indirectly constrain or facilitate bargaining in proximate districts. Social learning theory also offers an environmental explanation for contract similarity. Theorists suggest that a unit (in this case a district) will examine peers with similar ideological, political, demographic, cultural, racial or ethnic profiles and needs to introduce reforms best suited to specific contexts (Bennett 1991).

Several social and economic theories imply a more active role for leader districts. Pattern bargaining theory posits that an agreement reached by a firm (district) sets the pattern for subsequent or nearby firms (districts) (Marshall and Merlo 2004). And networks and conferences designed to facilitate communication between district managers may also facilitate information sharing on bargaining by leaders in geographically proximate districts.

Sociologists DiMaggio and Powell (1983) provide another oft-cited theory. They define "institutional isomorphism" as the result of organizations competing for resources and legitimacy. In their framework, proximate districts could have similar collective bargaining agreements because of coercive isomorphism (due to districts and unions exerting pressure to adopt similar provisions) or mimetic isomorphism (due to districts and unions purposely

[1] There have been relatively few teacher strikes in Washington state (four since 2003), and Seattle is the only district in the state receiving TIF funding, so we do not explicitly model these bargaining shocks in our analysis.

modeling their contract on those of nearby districts).[2] Either way, the theory suggests that firms (district) actively respond to the actions of their neighbors. Finally, policy diffusion – a descendent of social learning theory – suggests that policy spread results from political decision-makers' efforts to solve public policy problems. Leaders actively search for successful solutions and engage in a form of "satisficing" to simplify the process, emulating solutions posed by other municipalities, states or countries rather than incurring the cost of independently evaluating all possible alternatives and hedging political risks (Boemke 2004, Bennett 1991).[3]

Empirical work on teacher contracts has tended to focus on specific contract provisions. For example, recent studies of the teacher labor force find "wage spillovers" between districts in states that allow teachers salaries to be collectively bargained (Babcock et al 2004, Winters 2010). Thus far, however, research has overlooked the influence of geography on other aspects of teacher labor contracts – association rights, hiring and transfers, workload, evaluation, grievance, benefits and leave, layoff and recall. This is surprising given 1) concerns that CBAs may have a significant impact on teacher quality, teacher distribution, and ultimately student achievement (Hoxby 1996; Leven and Quinn 2003; Moe 2009; Strunk forthcoming) and 2) existing conceptual frameworks that suggest districts and unions alike may learn from and contribute to their neighbors' bargaining decisions.

III. Data

This paper utilizes a unique new dataset derived from the collective bargaining agreements (CBAs) of all 270 Washington state public school districts that are governed by

[2] Coercive isomorphism resembles economic theories of "union threat" whereby increases in union strength cause increases in nonunion wages and decreases in nonunion employment (Leicht 1989).

[3] Note that both the theory of institutional isomorphism and social learning theory contain a time element; that is, they predict that collective bargaining agreements should look more similar over time, and specific provisions should spread geographically over time. Our data is a snapshot in time of every collective bargaining agreement operating in Washington state, so we do not address the time element of this process. Rather, we measure the extent to which the provisions in CBAs are spatially dependent at a specific moment in time.

collective bargaining.[4] We obtained the CBA for all 270 of these districts and coded each CBA across 633 different provisions using a rubric developed by Katharine Strunk and described in greater detail in Goldhaber et al (2012).[5] Following Strunk and Reardon (forthcoming), we also transformed all data to run Partial Independence Item Response (PIIR) models that treat each provision in a CBA as a binary "response" to a survey that includes all contract provisions covered in collective bargaining agreements. For more details on this process see Goldhaber et al (2012).

To our knowledge, there are two other datasets that contain data on collective bargaining agreements across a large range of provisions: the National Council on Teacher Quality TR3 database (NCTQ 2009) contains data on the CBAs of over 100 of the largest districts in the country, while Katharine Strunk maintains a dataset–coded with the same rubric used in this analysis—of the CBAs of a large representative sample of California school districts with four or more schools.[6] Our dataset is the first to include *every* CBA in a state, which is important for exploring the spatial relationships between these agreements.

We link our coding data to district-level data from three other sources. District enrollment, demographic, teacher, and test score data come from the website of Washington's Office of the Superintendent of Public Instruction.[7] District funding data—both the percent of each district's funding that comes from local, state, and federal sources and per pupil spending figures—come from the National Center for Education Statistics' "Local Education Agency Finance Survey", part of the Common Core of Data.[8] Finally, we derive an indicator of the

[4] 25 districts in the state do not have a collective bargaining agreement and are therefore not included in this analyses. We are indebted to Dylan D'Entremont for his tireless work collecting these CBAs.

[5] We originally coded across 766 provisions but the "full data" for this paper include the 633 provisions described in Goldhaber et al. (2012).

[6] David Lipsky and colleagues at the Scheinman Institute on Conflict Resolution at Cornell University are also currently coding approximately 700 collective bargaining agreements from New York state.

[7] We use data from the 2010-11 school year, the same year every collective bargaining agreement in our sample was active, accessed from http://reportcard.ospi.k12.wa.us/DataDownload.aspx?schoolId=1&OrgTypeId=1&reportLevel=State&orgLinkId=

[8] We use data from the most recent version of this survey (2009), accessed from http://nces.ed.gov/ccd/f33agency.asp

partisanship of each district from the voting record of the congressperson representing each district in the U.S. House of Representatives, reported as the "Partisan Voting Index" in the Cook Political Report[9].

Table 1 gives summary statistics for these variables over the 270 districts in our sample, and for comparison, also gives summary statistics for the 25 districts not in our sample because they do not have collective bargaining agreements. The districts not in our sample are significantly smaller on average and have far lower poverty levels than the districts in our sample[10], but since these districts don't have CBAs and therefore do not speak to our research question, we do not have any concerns with selection bias.

IV. Analytic Approach

Our primary goal is to quantify the role of spatial relationships in the adoption of collective bargaining provisions. Prior research on spatial relationships that focus on a single variable, like teacher wages (Babcock et al 2004, Winters 2010), have used a spatial econometric approach (Anselin 1988) that directly models the correlation between wages in a district and the wages in nearby districts. However, our interest is in spatial relationships across potentially hundreds of individual provisions, so we adopt a pairwise approach in which each observation is a pair of school districts and our dependent variable is the similarity between the CBAs in those districts *measured at the individual provision level*. Specifically, we estimate variants of a pairwise model in which the similarity between the CBAs of each pair of districts is a function of the observable differences between the districts and an indicator for the proximity of the

[9] Data was accessed from http://cookpolitical.com/sites/default/files/pvistate.pdf. The Partisan Voting Index is reported as a party affiliation and a magnitude of partisanship (e.g., D10 or R14), so we recode this as a continuous variable that increases with increasing Democrat partisanship. We report this variable as "Congressional Democrat Index" in our results. For school districts represented by more than one congressperson, the index for that district is a weighted average by the number of schools in the district represented by each congressperson.
[10] The average district not in our sample has fewer than 100 students.

districts.[11] If spatial relationships play a role in the adoption of CBA provisions, we would expect the proximity of two districts to be a significant predictor of the similarity of those districts' contracts, holding other variables constant. Before giving the details of this model, we outline the measures of CBA similarity and district proximity we use in our analysis.

Calculating provision-level similarity

There are several challenges in calculating the similarity between two CBAs (or performing any quantitative analysis with CBAs). First, provisions of interest may be qualitative ("Is seniority the sole determinant of teacher transfers?") or quantitative ("What is the length of the school day?"). Second, language in different contracts may place different levels of restriction on the same provision; for example, if we are interested in the role of seniority in teacher transfer policies, seniority could play no role, be a tiebreaker, be one of several determinants, or be the sole determinant of which teachers are able to transfer between jobs.[12] Finally, and perhaps more importantly, not all provisions apply to all CBAs. That is, some provisions (such as whether seniority is the deciding factor in transfer decisions) can only appear if other provisions (such as whether seniority plays *any* role in transfer decisions) also appear in the contract.

All of these challenges can be overcome by using a clever coding strategy introduced by Strunk and Reardon (forthcoming) that builds on methodology developed by Reardon and Raudenbush (2006). Strunk and Reardon treat each provision in a CBA as the "response" to a conditionally-structured survey that includes binary questions about all provisions covered in collective bargaining agreements. This conditional structure allows them to address quantitative provisions (e.g., "Is the length of the school day longer than 360 minutes?"; "If so, is it longer

[11] This model is similar to the pairwise models in Koedel et al (2011). Our choice to include the absolute differences of observable district characteristics is informed by theory (cited in the previous section) that contracts in neighboring districts could be similar because of shared environmental characteristics.
[12] For an in-depth analysis of similar provisions in California, see Koski and Horng (2007)

than 390 minutes?", etc.) and provisions that can only appear if other provisions appear as well ("Is seniority play a role in transfer provisions?"; "If so, is seniority the sole determinant of which teachers can transfer jobs?", etc.) We follow Strunk and Reardon's method to code all 270 CBAs in our sample, and Goldhaber et al (2012) gives the details of this coding process. This process results in a final binary dataset of 633 different provisions that appear in at least 10 but no more than 260 of the 270 collective bargaining agreements in the state.

The structure of our data is central to our calculations of the similarity of each CBA. First, since the data are binary (i.e., does this provision appear in the CBA?), we employ three measures of similarity that are commonly used for binary data:[13] the simple matching coefficient, the Jaccard coefficient, and the Sorensen coefficient. To define these coefficients in the context of CBAs, let $a_{11}^{(i,j)}$ be the number of provisions that appear in both CBA *i* and CBA *j*, let $a_{01}^{(i,j)}$ and $a_{10}^{(i,j)}$ be the number of provisions that appear in one contract and not the other, and let $a_{00}^{(i,j)}$ be the number of provisions that appear in neither contract. The simple matching coefficient is simply the number of provisions that appear or don't appear in both contracts over the total number of provisions:

$$S_{ij}^{(simple)} = \frac{a_{11}^{(i,j)} + a_{00}^{(i,j)}}{a_{11}^{(i,j)} + a_{01}^{(i,j)} + a_{10}^{(i,j)} + a_{00}^{(i,j)}} \qquad (1)$$

While the interpretation of this coefficient is straightforward, for our purposes (as in many applications), it may be more important when provisions *do appear* in two different contracts than when they *do not appear* in either contract. CBAs are of finite length, so the limited number of provisions included in each provide signals about what districts find important, signals that we don't want drowned out by the larger number of provisions that will rarely occur in any pair of CBAs. The Jaccard coefficient ignores all "double zeroes" and calculates the number of provisions that appear in *both* CBAs over the number of provisions that appear in *either* CBA:

[13] See, for instance, Levandowsky and Winter (1971) and McGarigal et al (2000).

$$S_{ij}^{(Jaccard)} = \frac{a_{11}^{(i,j)}}{a_{11}^{(i,j)} + a_{01}^{(i,j)} + a_{10}^{(i,j)}} \qquad (2)$$

A slight modification of the Jaccard coefficient is the Sorensen coefficient, which gives double weight to provisions that appear in both CBAs:

$$S_{ij}^{(Sorensen)} = \frac{2a_{11}^{(i,j)}}{2a_{11}^{(i,j)} + a_{01}^{(i,j)} + a_{10}^{(i,j)}} \qquad (3)$$

We use the Jaccard coefficient in our base models, but check the robustness of our findings with the other two similarity coefficients.

Calculating these coefficients over all the provisions in our sample produces misleading results, though, because a provision can be coded as a "0" either because the negotiators chose not to include the provision or because the negotiators *could not* include that provision given the other provisions they did include in the contract. We want to draw conclusions about the similarities and differences due to negotiating decisions, so we only want to consider provisions that *could appear* in a CBA given the other provisions in the CBA. Fortunately, the conditional structure of our data allows us to calculate the "risk set" of each CBA (Reardon and Raudenbush, 2006), which is the set of provisions that could be addressed in the CBA given the other provisions in the contract.[14] In calculating the similarity between CBAs *i* and *j*, then, we only consider provisions that are *in the risk set of both CBAs*. This allows us to calculate a measure of similarity between each pair of CBAs that only considers provisions that could have been considered by negotiators in both districts.

Measures of proximity

As outlined in the literature review, the role of proximity in collective bargaining could be geographical or institutional. As a measure of linear geographical proximity, we calculate indicators of whether the centers of each pair of districts are within 25 miles, between 25 and 50

[14] See Goldhaber et al (2012) for more details.

miles, or between 50 and 100 miles of each other. However, treating distance as linear is problematic, because districts are significantly smaller in western Washington than in the eastern part of the state; for example, there are 24 districts within 25 miles of Seattle School District, while there are no other districts within 25 miles of the center of Methow Valley School District. So, we also calculate two non-linear measures of proximity: indicators for whether each pair of districts shares a common district boundary (adjacent districts)[15], and indicators for whether each pair of districts falls into the same metropolitan statistical area (MSA) or micropolitan statistical area (MiSA), as defined by the U.S. census.

In the context of collective bargaining, institutional proximity may be even more important than geographic proximity. Specifically, school districts in Washington are organized into nine Education Service Districts (ESDs) and 21 local UniServ councils. Each type of organization is affiliated with one side of the bargaining table—ESDs with district administration, and UniServ councils with local Education Associations—and thus it is plausible that districts in the same ESD or UniServ council would have similar CBAs. As institutional measures of proximity, then, we calculate indicators for whether each pair of districts is in the same ESD or UniServ council.

One consequence of calculating seven measures of proximity is that it is difficult to include all measures in the same model, because not all measures are sufficiently identifiable across the pairs of districts in our sample (for example, there is no pair of districts to identify the effect of sharing a district boundary if we include all the linear indicators of proximity in the model.) Therefore, we include the three types of measures—linear, non-linear, and institutional—in separate models, and run a final model with both non-linear and institutional measures, since those effects are sufficiently identifiable (at least 84 pairs of districts identify each coefficient.)

[15] Census shape files in ArcGIS provide district boundaries that extend into water, meaning that we can code for neighbors of island school districts as well.

Pairwise model

Our empirical results are all estimated using variants of the following pairwise model, which models the similarity between each pair of districts as a function of the observable differences between the districts and indicators for the proximity of the districts:

$$S_{ij} = B_{ij}\beta_1 + \left|X_i - X_j\right|\beta_2 + \sum_k \omega_k W_{ijk} + \varepsilon_{ij} \qquad (4)$$

The dependent variable is the similarity between the CBAs in districts *i* and *j;* as mentioned above, we use the Jaccard coefficient in our base models. B_{ij} is a vector of indicators for whether districts *i* and *j* both border Oregon, both border Idaho, or both border Canada. We include these controls in all models because we only have data for districts in the state of Washington, but the impact of proximity on the adoption of CBA provisions could cross state or national boundaries.[16] Because we cannot quantify these effects with our data, we include these indicators to control for the confounding influence of potential interstate or international relationships.[17]

$\left|X_i - X_j\right|$ is a vector of the absolute differences between the observable characteristics of districts *i* and *j*. In various model specifications, the covariates in X_i may include logged student enrollment[18], student racial composition, student poverty-level, teacher characteristics, a district partisan index, student performance, district funding sources, and per pupil spending[19]. We need to be cautious about including all of these variables in our models, however, because some could plausibly *result from* bargaining outcomes. Thus, we estimate some models with only the

---

[16] For example, districts on opposite sides of state borders may compete for teachers who are credentialed in both states.

[17] Specifically, if two nearby districts X and Y are both on the same state border, we cannot distinguish whether the similarity between the CBAs of the two districts is due to their proximity or because of the *shared influence* of a district Z on the other side of the border. The border indicators control for this potentially confounding effect.

[18] We take the log of district enrollment as a variance stabilizing transformation, as recommended in Box and Cox (1964).

[19] District funding and per pupil spending data are missing for three districts, so for those districts we impute the mean of the other 267 districts.

control variables we consider to be clearly exogenous to bargaining outcomes (enrollment, student covariates, and district partisanship) and additional models that add control variables that could plausibly be endogenous to bargaining outcomes (teacher characteristics, student performance, district funding, and per pupil spending) to test the robustness of our findings.

Each parameter of interest, ω_k, is the coefficient on W_{ijk}, which is an indicator for the proximity between districts *i* and *j* measured by proximity measure *k*. As discussed above, we include different combinations of proximity measures in different parameterizations of the model, so in each model, ω_k can be interpreted as the marginal effect of proximity measure *k* on the similarity between two districts' CBAs, holding all other variables constant.

Subsets of Data

Our base models use as the dependent variable the similarity between the CBAs in districts *i* and *j* measured over *all* provisions in the risk set of each CBA. However, it is also of empirical interest whether spatial relationships can be detected for subsets of provisions. Thus, we create one objectively-derived subset of provisions and one subjectively-derived subset of provisions to test the robustness of our results. The objectively-derived subset of provisions (which we will call our "restricted data") is created following a procedure described in Strunk and Reardon (forthcoming). The idea is to remove provisions from the dataset until all provisions contribute sufficiently to an underlying latent factor (details of this procedure are in Goldhaber et al (2012).) The resulting restricted dataset contains 218 of the 633 provisions in the full data. The subjectively-derived subset of provisions (which we call our "cherry-picked data") is a set of 40 provisions that have received a lot attention in the press and teacher labor literature; they include provisions addressing seniority-based transfer and layoff policies, evaluation procedures, class size restrictions, and teacher workday agreements, among others (a full list is available in Goldhaber et al (2012).) We run variants of model (4) that use as the dependent variable the similarity between the CBAs in districts *i* and *j* calculated only considering the provisions either in the restricted data or in the cherry-picked data.

Categories of Provisions

Another question of interest is whether spatial relationships are more important for some types of provisions than others. To investigate this, we divide all contract provisions into one of seven categories: Association rights, evaluation, grievance, layoff and recall, benefits and leave, hiring and transfer, and teacher workload. We also consider an additional category of "provisions" that concern the accessibility of the contract to teachers and administrators (such as the length of the contract, the number of provisions, and the number of contacts necessary to acquire the CBA.) As with the restricted and cherry-picked subsets of provisions, we run variants of model (4) that use as the dependent variable the similarity between the CBAs in districts *i* and *j* calculated only considering the provisions in each of the eight categories above.

Incorporating contract restrictiveness

It is intuitive to consider individual contracts as being either more or less "restrictive" to district administration, and it is interesting to consider whether proximity can predict the relative restrictiveness of the contracts in two districts. Thus we calculate the restrictiveness of all 270 contracts using a Partial Independence Item Response (PIIR) model developed by Reardon and Raudenbush (2006) and applied to CBAs by Strunk and Reardon (forthcoming). An analysis of these restrictiveness estimates is given in Goldhaber et al (2012). The PIIR model exploits the conditional structure of the data by only considering provision *k* in CBA *i* if it is in the risk set of that CBA. Specifically, if Y_{ik} represents the outcome of provision *k* in contract *i*, and h_{ik} is an indicator for whether provision *k* is in the risk set of CBA *i*, we can let $\varphi_{ik} = \Pr(Y_{ik} = 1 \mid h_{ik} = 1)$. The model is then:

$$\log\left(\frac{\varphi_{ik}}{1-\varphi_{ik}}\right) = \theta_i + \sum_{j=1}^{K} \gamma_j D_{ij} \qquad (5)$$

Here, θ_i represents the latent level of restrictiveness for CBA i, γ_j represents the conditional "severity" of each provision in the CBA, and D_{ij} is a dummy variable indicating which provision is considered.[20] Although the model allows simultaneous calculation of the restrictiveness of each CBA as a whole and the severity of each individual provision, we only use the estimates of the restrictiveness of each CBA (θ_i).[21] To assess whether proximity can predict the relative restrictiveness of the contracts in two districts, we run a variant of model (4) where we use the absolute difference between the restrictiveness of CBAs *i* and *j* $\left|\theta_i - \theta_j\right|$--multiplied by -1 and then standardized so greater values indicate more similar districts—as the dependent variable S_{ij}. The results of this model indicate whether geographically proximate districts are more likely (all else equal) to have similarly restrictive contracts.

V. Results

To ease interpretation of our results, we standardize the similarity measures across the 36,135 pairs of districts in our sample. We also standardize all district variables before taking the absolute differences. **Table 2** reports the coefficients from our base specifications of model (4), which uses the Jaccard measure of CBA similarity calculated across all 633 provisions in our full dataset. It is not clear which district variables we should include in the absolute differences $\left|X_i - X_j\right|$, so we run specifications of model (4) using three different sets of district variables: specification 1 (columns 1-4) controls only for the border indicators and the absolute difference in district enrollment (as similarity in district size is likely to be highly correlated with the

---

[20] As detailed in Reardon and Raudenbush (2006), this model can be fit either as a fixed effect model or a random effect model. We choose to run a fixed-effect specification, but the correlation between estimates from the two specifications is over .99.

[21] Restrictiveness is measured relative to other CBAs in our sample, so provisions adopted by the vast majority of districts contribute less to this measure than provisions adopted by fewer districts.

similarity of two contracts); specification 2 (columns 5-8) adds variables that are clearly exogenous to bargaining outcomes, such as differences in student-level variables and the difference in the Congressional Democrat Index; and specification 3 (columns 9-12) adds variables that are plausibly endogenous to bargaining outcomes, such as funding sources, teacher characteristics, student performance, and per pupil spending.[22] Within each specification, our variables of interest are the coefficients on the distance indicators: columns 1, 5, and 9 include linear distance indicators, columns 2, 6, and 10 include non-linear distance indicators, columns 3, 7, and 11 include institutional distance indicators, and columns 4, 8, and 12 include both non-linear and institutional distance indicators.

The results in Table 2 give strong evidence of spatial relationships. Since the results are quite robust to model specification, we will restrict attention to our preferred specification (specification 2) that includes all clearly exogenous variables but excludes variables that are plausibly endogenous. In column 5 of Table 2, we see that linear proximity is strongly correlated with the similarity between two contracts; on average and all else equal[23], districts within 25 miles have CBAs that are .135 standard deviations more similar than the average pair of districts, districts between 25 and 50 miles apart have CBAs that are .091 standard deviations more similar than the average pair of districts, and districts between 50 and 100 miles apart have CBAs that are .092 standard deviations more similar than the average pair of districts. These results are highly significant for all the distance measures. and the magnitudes of these coefficients suggest that the average pair of districts within 25 miles of each other share 1.5% more provisions than districts over 100 miles apart, and the average pair of districts between 25 and 100 miles of each other share 1.0% more provisions than districts over 100 miles apart.[24]

[22] We explain our motivation for these different models in the analytic approach section.
[23] For clarity of exposition, we will not repeat these caveats for the rest of our results, but all results are regression-adjusted and should thus be interpreted as holding all control variables constant.
[24] We illustrate this calculation for the coefficient on the indicator for districts within 25 miles of each other (.135). Using the Jaccard measure of similarity, the standard deviation of the percent of provisions shared in the shared risk set of each pair of CBAs is 6.07%. Thus the coefficient of .135 corresponds to a .135*.0607 = 0.82 percentage point

These spatial relationships also persist in models that consider non-linear and institutional measures of distance. Column 6 of Table 2 shows that districts in that share a border have CBAs that are .084 standard deviations more similar than the average pair of districts ($p < .05$), while districts in the same metropolitan or micropolitan area have CBAs that are .069 standard deviations more similar than the average pair of districts ($p < .05$).

The spatial findings are somewhat more modest in column 7 of Table 2, which shows that districts in the same Education Service District have CBAs that are .038 standard deviations more similar than the average pair of districts, while the marginal effect of being in the same UniServ council is not statistically significant. Column 8 of Table 2 demonstrates the relative importance of geographic distance over institutional distance, as sharing a district border and being in the same MSA or MiSA is significantly positively correlated with CBA similarity, while being in the same ESD or UniServ is not. This is a somewhat surprising result, as it suggests that institutions that contain districts or unions may not contribute to contract similarities when geographic proximity is taken into account.

It is interesting to note the sizable coefficients on many of the control variables in specification 2. For example, in the distance controls in Column 5 of Table 2, we see that districts that are both on the Oregon border have CBAs that are .135 standard deviations more similar than the average pair of districts, while districts that are both on the Idaho border have CBAs that are .378 standard deviations more similar than the average pair of districts. As we discussed in the previous section, it may be that districts along the border in Washington face different institutional incentives because they compete with neighboring states for teachers.

Observable differences are also highly predictive of CBA similarity; in Column 5 of Table 2, we see that a one standard deviation difference in log enrollment is correlated with

---

increase in the percent of shared provisions. The average pair of districts shares 53.00% of the provisions in its shared risk set, so the expected percent increase corresponding to this coefficient is [(53% + 0.82%) / 53%] – 100% = 1.5%.

having CBAs that are .276 standard deviations less similar, while a one standard deviation difference in the percent of students eligible for free or reduced price lunch is correlated with having CBAs that are .108 standard deviations less similar. These results emphasize the importance of shared environmental characteristics in similarities in CBAs, as predicted by the reviewed literature.

Models Using Different Subsets of Provisions and Similarity Measures

Because the results in Table 2 are robust to model specification, the remainder of our results control for our preferred set of distance difference covariates (specification 2 in Table 2). **Tables 3A** and **3B** present the coefficients from variants of equation (4) that are run on different subsets of data and using different similarity measures. The models in Table 3A only include linear distance indicators (so Column 1 of Table 3A is identical to Column 5 of Table 2), while the models in table 3B include non-linear and institutional distance indicators (so Column 1 of Table 3B is identical to Column 8 of Table 2.) Focusing on Columns 1-3 of Table 3A, we see that the proximity results are similar whether we use the Jaccard or Sorensen similarity measure, but are considerably less striking when we use the Simple Matching similarity measure. This is not surprising since the noise from provisions not included in either contract—which are included in the simple matching measure but not in the Jaccard or Sorensen measures—may dilute the signal from provisions that are included in both contacts. We see similar patterns across Columns 5-7 and 9-11 of Table 3A, and Columns 1-3, 5-7, and 9-11 of Table 3B, although the dilution of results using the simple matching coefficient is more pronounced in some specifications than others. From these results, we conclude that our results are relatively robust to the measure of similarity, and that the Jaccard measure is our preferred measure of provision-level similarity.

We now turn our attention to the results for the restricted data (Columns 5-8 of Tables 3A and 3B) and the cherry-picked data (Columns 9-12 of Tables 3A and 3B.) Interestingly, the estimates of spatial dependence become *more* pronounced as we restrict the dataset. For example, districts within 25 miles have CBAs that are .135 standard deviations more similar than the average pair of districts when we use the full set of provisions (Column 1 of Table 3A), but this difference becomes .183 standard deviations when we use the restricted dataset, and .232 standard deviations when we use the cherry-picked dataset. So, while the average pair of districts within 25 miles of each other share 1.5% more provisions over the full data than districts over 100 miles apart, this percent increase becomes 4.7% over the restricted dataset and 5.8% over the cherry-picked dataset.[25] We see similar results in Columns 1, 5, and 9 of Table 3B, as the marginal effect of non-linear and institutional measures of proximity also increase as we restrict the dataset. The improved results using the restricted dataset are not surprising, as removing noise from the data should improve our ability to detect signal in the spatial relationships. The results from the cherry-picked dataset are more interesting, as they suggest the spatial dependence between CBAs is stronger for high-profile provisions.[26]

Models Using CBA Restrictiveness Estimates

Columns 4, 8, and 12 of Tables 3A and 3B present coefficients from models that measure similarity as the relative restrictiveness of the CBAs in each pair of districts.[27] Interestingly, for the full data and restricted data (Columns 4 and 8 of Table 3A), the effects of linear distance

---

[25] We derive each of these percent increases using the standard deviation of the similarity measures calculated using each data subset (see footnote 24).

[26] ANOVA results reinforce these findings. We run a Type I and Type III ANOVA on the models in Tables 3A and 3B, which provides a lower and upper bound for the percent of explained variability due to proximity indicators. When we consider linear measures of distance (as in Table 3A), for example, we find that the proximity indicators account for 3.4%-4.2% of the explained variability in CBA similarity in the full data, 4.9%-5.4% of the explained variability in CBA similarity in the restricted data, and 4.8%-7.8% of the explained variability in CBA similarity in the cherry-picked data. Again, this suggests that proximity has particularly strong explanatory power in the adoption of high-profile provisions.

[27] To make the results comparable to the other measures of similarity, we multiply the absolute difference between the estimated restrictiveness of each pair of districts by -1 and then standardize over all pairs of districts. This results in a measure of relative restrictiveness that increases as the contracts become more similar.

indicators are robust to this new measure of similarity. In other words, districts within 25 miles, between 25 and 50 miles apart, and between 50 and 100 miles apart have more similarly restrictive contracts than the average pair of districts. These results do not persist to the cherry-picked data, suggesting either that restrictiveness resulting from high-profile provisions is not subject to spatial dependence or that it is not appropriate to apply the PIIR model to a subset of only 40 provisions.

In Columns 4, 8, and 12 of Table 3B, we see some interesting differences between models that use relative restrictiveness as the measure of similarity and models that use a provision-level measure of similarity. Specifically, while the provision-level models indicate that spatial dependence is driven by districts that share borders or are in the same MSA or MiSA, the models using relative restrictiveness suggest that districts in the same ESD have contracts that are considerably *more* similarly restrictive than the average pair of districts, while districts in the same UniServ council have contracts that are considerably *less* similarly restrictive than the average pair of districts. The results for ESDs are not surprising, as we hypothesized *a priori* that districts may work together within these institutions to negotiate more restrictive contracts. However, we can only speculate about the negative association between being in the same UniServ council and the relative restrictiveness of CBAs. Perhaps this negative association reflects choices made within UniServ councils to focus on negotiations in some member districts at the expense of others.[28]

Models Using Different Categories of Provisions

[28] This may be due to the fact that each year, negotiators within a UniServ council must choose how to allocate their time across all the districts with expiring contracts. To test this hypothesis, we add a term to the model that interacts the UniServ indicator with an indicator for whether the two CBAs were bargained in the same year. Under this model, the coefficient on the interaction term is significantly negative, but the coefficient on the UniServ indicator is also still significantly negative. This suggests that this result may be partially (but not completely) driven by allocation of UniServ resources within a bargaining year.

Tables 4A and **4B** present the coefficients from models that use as the dependent variable the Jaccard measure of CBA similarity calculated over different categories of provisions. Considering the geographic measures of proximity in Table 4A, we see that marginal effects of proximity are particularly striking for two categories: the provisions describing the "accessibility" of each contract, and the provisions governing hiring and transfer policies. Specifically, districts within 25 miles have CBAs that are .342 standard deviations more similar in accessibility and .247 standard deviations more similar in hiring and transfer policies than the average pair of districts. This gives evidence that districts take cues from proximate districts about how accessible to make their CBAs, and districts are particularly likely to bargain similar provisions regarding hiring and transfer provisions (often a bitterly negotiated set of provisions[29]) as geographically proximate districts. Geographic proximity is also a significant predictor of the similarity of provisions regarding Association rights, benefits and leave, and workload agreements. The same categories of provisions appear to be driving the marginal effects of non-linear and institutional measures of proximity in Table 4B.

VI. Conclusions

Our findings demonstrate that the similarity between the CBAs is a function of both district characteristics and the proximity of the districts to one another. This suggests that spatial relationships play a major role in determining the provisions that appear in any particular district's CBA. These results are robust across models that use various control variables, data subsets, and measures of similarity, and are most striking for two subsets of provisions: an objectively derived subset of provisions (see Goldhaber et al., 2012) and a subjectively chosen set of high-profile provisions.

[29] For example, a major sticking point in contract negotiations that led to a major teacher strike in Tacoma School District at the beginning of the 2011-12 school year was over seniority transfer policies (Baker 2011).

This analysis is a first step in exploring the spatial relationships between the provisions in collective bargaining agreements. Our findings and the theory we review suggest two natural extensions. First, while provocative in the sense that the findings strongly suggest what happens in one district bargain likely affects the nature of the bargain in other districts, we do not know that the correlation between districts' CBAs is in fact causal. Spatial lag models (Anselin 1988) provide a natural setting to tease out the potential causal impact of specific provisions—or the underlying restrictiveness—of CBAs in geographically proximate districts on the provisions or restrictiveness of a given contract. Bargaining shocks discussed earlier, such as teacher strikes, external grants, and changes in district or union leadership may allow an instrumental variable approach.

Second, our findings are consistent with contagion theories that predict the spread of policies over time as districts respond to the provisions bargained in nearby districts. But cross-section data represent a snapshot in time and do not allow us to model the temporal aspect of this process. An analysis of the adoption of specific CBA provisions across districts over time could highlight both the spatial and temporal components of this process and give policy makers an even clearer picture of the role of spatial relationships in the adoption of collective bargaining provisions.

References

Anselin, Luc (1988). *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer Academic Publishers.

Babcock, Linda, John Engberg, and Robert Greenbaum (2004). “Wage spillovers in public sector contract negotiations: the importance of social comparisons.” *Regional Science and Urban Economics,* 35 (2005), 395–416.

Baker, Mike. (2011). “Tentative Pact in Tacoma Teachers Strike.” *The Seattle Times*, September 21, 2011.

Bennett, Colin J. (1991). "What is Policy Convergence and What Causes It?" *British Journal of Political Science* 21; 215-33.

Berry, Frances Stokes and William D. Berry. (1990). "State Lottery Adoptions as Policy Innovations: An Event History Analysis." *American Political Science Review* 84: 395-415.

Boemke, Frederick J. and Richard Witmer. (2004). "Disentangling Diffusion: The Effects of Social Learning and Economic Competition on State Policy Innovation and Expansion." *Political Research Quarterly*, 57(1): 39-51.

Box, George and David Cox (1964). “An analysis of transformations.” *Journal of the Royal Statistical Society. Series B (Methodological),* pp. 211-252.

Boyd, Donald, Hamilton Lankford, Susanna Loeb, and James Wyckoff. (2003). "The Draw of Home: How Teachers' Preferences for Proximity Disadvantage Urban Schools." NBER Working Paper 9953.

Glick, Henry R., and Scott P Hays (1991). "Innovation and Reinvention in State Policymaking: Theory and the Evolution of Living Will Laws." *Journal of Politics* 53:835-50.

Goldhaber, Dan, Lesley Lavery, Roddy Theobald, Dylan D’Entremont, and Yangru Fang (2012). “Teacher Collective Bargaining in Washington State: Assessing the Internal Validity of Partial Independence Item Response Measures of Contract Restrictiveness.” CEDR Working Paper.

Hannaway, Jane and Andrew J. Rotherham (2006). *Collective Bargaining in Education: Negotiating Change in Today's Schools*. Cambridge, MA: Harvard Education Press.

Hoxby, Caroline. (1996). "How Teachers' Unions Affect Education Production." *The Quarterly Journal of Economics*, 111(3): 671-718.

Koedel, Cory, Jason Grissom, Shawn Ni, and Michael Podgursky (2011). “Pension-Induced Rigidities in the Labor Market for School Leaders.” Koedel Working Paper 11-15.

Koski, William & Eilene Horng (2007). “Curbing or Facilitating Inequality? Law, Collective Bargaining, and Teacher Assignment Among Schools in California.” *Education Finance and Policy, 2(3),* 262-300.

Leicht, Kevin T. (1989). On the Estimation of Union Threat Effects. *American Sociological Review* 54(6): 1035-1047.

Levin, Jessica and Meredith Quinn, (2003). *Missed Opportunities: How We Keep High-Quality Teachers Out of Urban Classrooms*, The New Teacher Project.

Levandowsky, Michael and David Winter (1971). "Distance Between Sets". *Nature* 234:5323, pp. 34-35.

Lewin, David and Mary McCormick, (1981). "Coalition Bargaining in Municipal Government: The New York City Experience*, Industrial and Labor Relations Review* 34, 175-184.

McGarigal, Kevin, Sam Cushman and Susan Stafford (2000). *Multivariate Statistics for Wildlife and Ecology Research.* New York: Springer.

Moe, Terry M. (2009). "Collective Bargaining and the Performance of the Public Schools." *American Journal of Political Science*, 53(1): 156-174.

National Council on Teacher Quality (2009). "Teacher Rules, Roles, and Rights." www.nctq.org/tr3

Peterson, Paul, and Mark C. Rom. (1990). *Welfare Magnets:A New Case for a National Standard*. Washington, DC: The Brookings Institute

Reardon, Sean & Stephen Raudenbush (2006). "A Partial Independence Item Response Model for Surveys With Filter Questions." *Journal of Sociological Methods 36(1)*, 257-300.

Schneider, Anne and Helen Ingram (1988). Systematically Pinching Ideas: A Comparative Approach to Policy Design. *Journal of Public Policy* 8(1): 61-80.

Strunk, Katharine (Forthcoming). "Are Teachers' Unions Really to Blame? The Relationships Between Teachers' Union Contract Restrictiveness, District Resource Allocation and Student Test Scores in California." *Education Finance and Policy.*

Strunk, Katharine and Jason Grissom (2010). "Do Strong Unions Shape District Policies? Collective Bargaining, Teacher Contract Restrictiveness, and the Political Power of Teachers' Unions." *Educational Evaluation and Policy Analysis, 32(3),* 389-406.

Strunk, Katharine and Sean Reardon (Forthcoming). "Measuring Union Strength: A Partial Independence Item Response Approach to Measuring the Restrictiveness of Teachers' Union Contracts." *Journal of Educational and Behavioral Statistics*.

Tiebout, Charles M. (1956). "A Pure Theory of Local Expenditures." *The Journal of Political Economy* 64(5): 416-424.

Winters, John V, (2010). "Teacher Salaries and Teacher Unions: a Spatial Econometric Approach." MPRA Paper 21202, University Library of Munich, Germany

Tables

Table 1: Summary Statistics for Districts With and Without CBAs in Washington State

	Districts with CBAs	**Districts without CBAs**
	N = 270	N = 25
	Mean (SD)	**Mean (SD)**
% on Oregon Border	10.4%	20.0%
% on Idaho Border	5.6%	4.0%
% on Canada Border	3.7%	4.0%
Average Enrollment	3827 (6205)	97 (200)
Average % American Indian Students	5.9% (14.1%)	6.3% (20.0%)
Average % Asian/Pacific Islander Students	3.4% (4.9%)	1.2% (2.6%)
Average % Black Students	2.1% (3.5%)	0.2% (0.6%)
Average % Hispanic Students	15.4% (21.0%)	12.2% (25.8%)
Average % Migrant Students	1.9% (5.4%)	4.3% (12.0%)
Average % Bilingual Students	6.0% (9.8%)	5.5% (15.9%)
Average % Special Education Students	13.0% (3.1%)	9.4% (7.6%)
Average % Students with Free/Reduced Price Meals	48.3% (19.7%)	27.5% (32.1%)
Average Congressional Democrat Index	-2.6 (7.8)	-6.2 (5.8)
Average % Local Funding	23.8% (9.2%)	12.8% (10.1%)
Average % State Funding Difference	66.9% (8.2%)	75.3% (12.0%)
Average Mean Teacher Experience	13.0% (2.1%)	13.4% (6.9%)
Average % Teachers with Masters Difference	65.2% (10.5%)	69.9% (22.9%)
Average % Passing State Exam in Reading	66.5% (11.2%)	64.1% (14.4%)
Average % Passing State Exam in Math	47.2% (12.5%)	53.3% (19.6%)
Average Per Pupil Spending on Instruction	$6295 ($2313)	$9731 ($5642)

Table 2: Predictors of Jaccard Measure of CBA Similarity Across Entire Contract

	Specification 1				Specification 2				Specification 3			
	1	2	3	4	5	6	7	8	9	10	11	12
Distance Indicators												
Within 0-25 mi.	0.132 ***				0.135 ***				0.098 ***			
Within 25-50 mi.	0.089 ***				0.091 ***				0.067 ***			
Within 50-100 mi.	0.087 ***				0.092 ***				0.07 ***			
Share Border		0.083 *		0.095 *		0.084 *		0.092 *		0.070		0.078 *
Same MSA/MiSA		0.094 **		0.1 **		0.069 *		0.072 *		0.047 +		0.052
Same Uniserv			-0.042 +	-0.052 *			-0.031	-0.040			-0.030	-0.037
Same ESD			0.048 *	0.021			0.038 +	0.018			0.029	0.014
Distance Controls												
Both on OR Border	0.152 **	0.158 **	0.157 **	0.156 **	0.135 **	0.14 **	0.139 **	0.138 **	0.172 ***	0.176 ***	0.176 ***	0.174 ***
Both on ID Border	0.337 ***	0.362 ***	0.383 ***	0.381 ***	0.378 ***	0.389 ***	0.402 ***	0.401 ***	0.429 ***	0.438 ***	0.451 ***	0.45 ***
Both on Can. Border	-0.066	-0.072	-0.050	-0.068	-0.096	-0.102	-0.084	-0.099	-0.113	-0.118	-0.102	-0.115
Absolute Differences												
Log Enrollment	-0.279 ***	-0.279 ***	-0.28 ***	-0.279 ***	-0.276 ***	-0.278 ***	-0.279 ***	-0.278 ***	-0.267 ***	-0.267 ***	-0.268 ***	-0.268 ***
% Amer Indian Stu.					-0.03 ***	-0.031 ***	-0.031 ***	-0.031 ***	-0.035 ***	-0.036 ***	-0.036 ***	-0.036 ***
% Asian Stu.					-0.035 ***	-0.031 ***	-0.03 ***	-0.031 ***	-0.051 ***	-0.047 ***	-0.047 ***	-0.048 ***
% Black Stu.					0.081 ***	0.086 ***	0.086 ***	0.086 ***	0.087 ***	0.09 ***	0.090 ***	0.09 ***
% Hispanic Stu.					0.071 ***	0.066 ***	0.065 ***	0.065 ***	0.068 ***	0.063 ***	0.063 ***	0.063 ***
% Migrant Stu.					0.017 **	0.019 **	0.020 **	0.020 **	0.012 *	0.014 *	0.014 *	0.014 *
% Bilingual Stu.					0.012	0.014	0.014	0.013	0.016 *	0.017 *	0.017 *	0.017 *
% Special Ed. Stu.					0.062 ***	0.061 ***	0.060 ***	0.061 ***	0.065 ***	0.065 ***	0.064 ***	0.065 ***
% Reduced Meals Stu.					-0.108 ***	-0.106 ***	-0.106 ***	-0.106 ***	-0.133 ***	-0.131 ***	-0.131 ***	-0.131 ***
Democrat Index					-0.009	-0.029 ***	-0.032 ***	-0.031 ***	-0.015 *	-0.030 ***	-0.033 ***	-0.032 ***
% Local Funding									0.024 **	0.023 **	0.023 **	0.023 **
% State Funding									-0.06 ***	-0.061 ***	-0.061 ***	-0.061 ***
Avg. Tch. Experience									-0.017 **	-0.017 **	-0.017 **	-0.017 **
% Tch. w/ Masters									-0.13 ***	-0.131 ***	-0.132 ***	-0.131 ***
% Passing in Reading									0.047 ***	0.05 ***	0.051 ***	0.05 ***
% Passing in Math .									0.04 ***	0.039 ***	0.039 ***	0.039 ***
Per Pupil Spending									0.033 ***	0.034 ***	0.034 ***	0.034 ***

Significance levels based on p-values from Wald test: + p<0.1, * p<.05, ** p < .01, *** p < .001

NOTE: Each observation is a pair of districts in the sample. The dependent variable is the Jaccard measure of similarity between the two CBAs, standardized over all pairs of districts. Distance variables are binary indicators, and controls are absolute differences of standardized covariates.

Table 3A: Robustness Checks for Different Subsets of Provisions and Similarity Measures (Linear Proximity)

	Full Data (633 Provisions)				Restricted Data (218 Provisions)				Cherry-Picked Data (40 Provisions)			
	1	2	3	4	5	6	7	8	9	10	11	12
Similarity Measure	Jaccard	Simple	Sorensen	PIIR	Jaccard	Simple	Sorensen	PIIR	Jaccard	Simple	Sorensen	PIIR
Distance Indicators												
Within 0-25 mi.	0.135 ***	0.016	0.129 ***	0.107 ***	0.183 ***	0.087 **	0.161 ***	0.131 ***	0.232 ***	0.220 ***	0.207 ***	-0.022
Within 25-50 mi.	0.091 ***	0.008	0.096 ***	0.097 ***	0.120 ***	0.047 *	0.121 ***	0.079 ***	0.114 ***	0.105 ***	0.108 ***	0.014
Within 50-100 mi.	0.092 ***	0.031 *	0.092 ***	0.062 ***	0.111 ***	0.059 ***	0.103 ***	0.067 ***	0.058 ***	0.04 **	0.054 ***	-0.005
Distance Controls												
Both on OR Border	0.135 **	0.337 ***	0.133 *	0.099 *	0.043	0.111 *	0.062	0.236 ***	0.141 **	0.296 ***	0.136 **	-0.139 **
Both on ID Border	0.378 ***	0.430 ***	0.326 ***	-0.018	-0.103	0.080	-0.183 +	-0.042	0.033	-0.164 +	0.046	0.292 **
Both on Can. Border	-0.096	-0.361 *	-0.089	-0.010	0.300 *	0.079	0.287 *	-0.128	-0.084	-0.32 *	-0.074	0.100
Absolute Differences												
Log Enrollment	-0.276 ***	-0.211 ***	-0.276 ***	-0.266 ***	-0.290 ***	-0.336 ***	-0.277 ***	-0.318 ***	-0.273 ***	-0.284 ***	-0.271 ***	-0.173 ***
% Amer Indian Stu.	-0.03 ***	0.032 ***	-0.032 ***	-0.032 ***	-0.067 ***	-0.020 ***	-0.059 ***	-0.040 ***	-0.045 ***	-0.008 *	-0.043 ***	-0.005
% Asian Stu.	-0.035 ***	-0.160 ***	-0.034 ***	0.031 ***	0.063 ***	-0.028 ***	0.067 ***	0.069 ***	0.055 ***	0.006	0.056 ***	-0.016 *
% Black Stu.	0.081 ***	0.056 ***	0.080 ***	-0.062 ***	0.014 *	-0.012 ***	0.008	-0.030 ***	0.052 ***	0.009	0.048 ***	-0.045 ***
% Hispanic Stu.	0.071 ***	0.084 ***	0.072 ***	0.101 ***	0.048 ***	0.063 ***	0.045 ***	0.130 ***	-0.025 *	0.028 **	-0.026 *	0.057 ***
% Migrant Stu.	0.017 **	-0.007	0.018 *	0.019 **	0.013 *	-0.030 ***	0.019 **	0.003	0.079 ***	0.030 ***	0.079 ***	0.041 ***
% Bilingual Stu.	0.012	-0.032 ***	0.011	-0.076 ***	0.028 **	0.013	0.028 **	-0.048 ***	-0.003	-0.018 +	-0.002	-0.081 ***
% Special Ed. Stu.	0.062 ***	0.083 ***	0.063 ***	0.034 ***	0.072 ***	0.091 ***	0.076 ***	0.064 ***	-0.014 *	-0.002	-0.010 +	0.030 ***
% Reduced Meals Stu.	-0.108 ***	-0.026 ***	-0.107 ***	-0.073 ***	-0.107 ***	-0.037 ***	-0.103 ***	-0.062 ***	-0.071 ***	-0.040 ***	-0.070 ***	-0.011
Democrat Index	-0.009	-0.06 ***	-0.009	-0.012 +	0.010	-0.029 ***	0.004	-0.042 ***	0.017 *	-0.005	0.014 *	-0.037 ***

Significance levels based on p-values from Wald test: + $p<0.1$, * $p<.05$, ** $p < .01$, *** $p < .001$

NOTE: Each observation is a pair of districts in the sample. The dependent variable is the measure of similarity between the two CBAs indicated for that column, standardized over all pairs of districts. The PIIR difference measure is standardized so that greater values indicate more similarity. Distance variables are binary indicators, and controls are absolute differences of standardized covariates.

Table 3B: Robustness Checks for Different Subsets of Provisions and Similarity Measures (Non-linear and Institutional Proximity)

	Full Data (633 Provisions)				Restricted Data (218 Provisions)				Cherry-Picked Data (40 Provisions)			
	1	2	3	4	5	6	7	8	9	10	11	12
Similarity Measure	Jaccard	Simple	Sorensen	PIIR	Jaccard	Simple	Sorensen	PIIR	Jaccard	Simple	Sorensen	PIIR
Distance Indicators												
Share Border	0.092 *	0.052	0.081 *	0.017	0.093 *	0.035	0.079 *	0.015	0.103 *	0.107 **	0.093 *	-0.015
Same MSA/MiSA	0.072 *	-0.030	0.073 *	0.075 *	0.121 ***	0.051	0.107 ***	0.073 *	0.241 ***	0.174 ***	0.218 ***	0.006
Same Uniserv	-0.040	0.054 *	-0.046	-0.136 ***	-0.097 ***	-0.002	-0.107 ***	-0.176 ***	-0.174 ***	-0.14 ***	-0.17 ***	-0.083 **
Same ESD	0.018	-0.004	0.020	0.064 ***	0.017	0.024	0.020	0.100 ***	0.108 ***	0.12 ***	0.109 ***	0.082 ***
Distance Controls												
Both on OR Border	0.138 **	0.333 ***	0.136 **	0.1 *	0.054	0.107 *	0.073	0.23 ***	0.126 *	0.276 ***	0.121 *	-0.157 **
Both on ID Border	0.401 ***	0.407 ***	0.352 ***	0.037	-0.046	0.076	-0.121	0.019	0.074	-0.141	0.085	0.295 **
Both on Can. Border	-0.099	-0.374 **	-0.091	0.006	0.307 *	0.074	0.297 *	-0.108	-0.087	-0.326 *	-0.076	0.104
Absolute Differences												
Log Enrollment	-0.278 ***	-0.211 ***	-0.277 ***	-0.268 ***	-0.293 ***	-0.336 ***	-0.28 ***	-0.319 ***	-0.273 ***	-0.284 ***	-0.272 ***	-0.172 ***
% Amer Indian Stu.	-0.031 ***	0.031 ***	-0.033 ***	-0.032 ***	-0.067 ***	-0.021 ***	-0.06 ***	-0.039 ***	-0.045 ***	-0.008 ***	-0.043 ***	-0.005
% Asian Stu.	-0.031 ***	-0.158 ***	-0.03 ***	0.034 ***	0.068 ***	-0.025 ***	0.071 ***	0.071 ***	0.057 ***	0.007	0.058 ***	-0.016 ***
% Black Stu.	0.086 ***	0.057 ***	0.084 ***	-0.059 ***	0.020 ***	-0.010	0.013 *	-0.027 ***	0.055 ***	0.012 *	0.050 ***	-0.046 ***
% Hispanic Stu.	0.065 ***	0.084 ***	0.066 ***	0.095 ***	0.039 ***	0.061 ***	0.036 ***	0.124 ***	-0.029 **	0.025 *	-0.03 **	0.059 ***
% Migrant Stu.	0.02 **	-0.007	0.021 **	0.022 ***	0.017 **	-0.029 ***	0.023 ***	0.007	0.082 ***	0.032 ***	0.082 ***	0.042 ***
% Bilingual Stu.	0.013	-0.032 ***	0.012	-0.076 ***	0.03 **	0.013	0.030 **	-0.047 ***	-0.003	- 0.018*	-0.003	-0.083 ***
% Special Ed. Stu.	0.061 ***	0.082 ***	0.063 ***	0.034 ***	0.071 ***	0.090 ***	0.075 ***	0.063 ***	-0.013 *	-0.002	-0.010 +	0.031 ***
% Reduced Meals Stu.	-0.106 ***	-0.025 ***	-0.105 ***	-0.071 ***	-0.104 ***	-0.036 ***	-0.100 ***	-0.060 ***	-0.069 ***	-0.038 ***	-0.068 ***	-0.011 +
Democrat Index	-0.031 ***	-0.061 ***	-0.032 ***	-0.035 ***	-0.023 ***	-0.039 ***	-0.028 ***	-0.066 ***	-0.002	-0.019 **	-0.004	-0.035 ***

Significance levels based on p-values from Wald test: + p<0.1, * p<.05, ** p < .01, *** p < .001

NOTE: Each observation is a pair of districts in the sample. The dependent variable is the measure of similarity between the two CBAs indicated for that column, standardized over all pairs of districts. The PIIR difference measure is standardized so that greater values indicate more similarity. Distance variables are binary indicators, and controls are absolute differences of standardized covariates.

Table 4A: Predictors of Jaccard Measure of CBA Similarity by Category of Provisions (Linear Proximity)

	1	2	3	4	5	6	7	8	9
Provision Category	**Full Data**	**Accessibility**	**Association**	**Evaluation**	**Grievance**	**Layoffs**	**Leave**	**Transfer**	**Workload**
Distance Indicators									
Within 0-25 mi.	0.135 ***	0.342 ***	0.134 ***	-0.006	-0.017	-0.048	0.099 ***	0.247 ***	0.103 ***
Within 25-50 mi.	0.091 ***	0.179 ***	0.070 ***	-0.014	-0.003	-0.007	0.119 ***	0.202 ***	0.056 **
Within 50-100 mi.	0.092 ***	0.058 ***	0.033 *	0.025 +	0.019	0.005	0.088 ***	0.162 ***	0.080 ***
Distance Controls									
Both on OR Border	0.135 **	-0.206 ***	0.003	-0.013	0.435 ***	0.243 ***	0.032	0.014	0.132 **
Both on ID Border	0.378 ***	0.072	0.616 ***	0.22 *	-0.175 +	0.291 **	0.558 ***	-0.190 *	-0.058
Both on Can. Border	-0.096	-0.401 **	0.079	-0.118	-0.052	0.17	-0.892 ***	0.101	0.455 **
Absolute Differences									
Log Enrollment	-0.276 ***	-0.428 ***	-0.232 ***	-0.073 ***	-0.08 ***	-0.042 ***	-0.075 ***	-0.243 ***	-0.155 ***
% Amer Indian Stu.	-0.030 ***	-0.039 ***	0.006	0.021 ***	-0.032 ***	-0.031 ***	0.062 ***	-0.009 *	-0.092 ***
% Asian Stu.	-0.035 ***	0.133 ***	0.037 ***	0.003	-0.16 ***	-0.001	-0.051 ***	0.112 ***	0.011 +
% Black Stu.	0.081 ***	0.044 ***	0.035 ***	0.033 ***	0.106 ***	0.003	0.07 ***	0.025 ***	0.035 ***
% Hispanic Stu.	0.071 ***	0.019 +	0.022 *	0.012	0.069 ***	-0.045 ***	0.011	-0.147 ***	0.097 ***
% Migrant Stu.	0.017 **	0.061 ***	0.075 ***	0.013 +	-0.011	0.026 ***	0.056 ***	0.133 ***	-0.022 ***
% Bilingual Stu.	0.012	-0.054 ***	-0.030 **	0.04 ***	-0.018 *	-0.012	0.065 ***	0.022 *	-0.012
% Special Ed. Stu.	0.062 ***	-0.039 ***	-0.023 ***	0.018 **	0.056 ***	-0.014 *	0.003	0.025 ***	0.079 ***
% Reduced Meals Stu.	-0.108 ***	-0.041 ***	-0.001	-0.053 ***	-0.012 +	-0.009	-0.034 ***	-0.058 ***	-0.100 ***
Democrat Index	-0.009	0.033 ***	-0.071 ***	-0.001	-0.014 *	0.023 **	0.013 +	0.057 ***	0.020 **

Significance levels based on p-values from Wald test: + p<0.1, * p<.05, ** p < .01, *** p < .001

NOTE: Each observation is a pair of districts in the sample. The dependent variable is the Jaccard measure of similarity between the two CBAs, standardized over all pairs of districts. Distance variables are binary indicators, and controls are absolute differences of standardized covariates.

Table 4B: Predictors of Jaccard Measure of CBA Similarity by Category of Provisions (Non-Linear and Institutional Proximity)

	1	2	3	4	5	6	7	8	9
Provision Category	**Full Data**	**Accessibility**	**Association**	**Evaluation**	**Grievance**	**Layoffs**	**Leave**	**Transfer**	**Workloac**
Distance Indicators									
Share Border	0.092 *	0.077 *	0.009	0.072 +	0.050	0.020	0.111 **	0.08 *	0.031
Same MSA/MiSA	0.072 *	0.331 ****	0.226 ***	-0.011	-0.057 +	-0.023	0.046	0.152 ***	0.1 **
Same Uniserv	-0.040	-0.212 ***	-0.147 ***	0.000	0.022	0.062 *	0.015	-0.125 ***	-0.069 **
Same ESD	0.018	0.138 ***	0.035	-0.012	0.034	-0.05 *	-0.078 ***	0.007	0.009
Distance Controls									
Both on OR Border	0.138 **	-0.22 ***	0.006	-0.014	0.421 ***	0.249 ***	0.060	0.039	0.138 **
Both on ID Border	0.401 ***	0.129	0.672 ***	0.221 *	-0.201 *	0.277 **	0.599 ***	-0.100	-0.021
Both on Can. Border	-0.099	-0.397 **	0.089	-0.126	-0.064	0.162	-0.894 ***	0.120	0.461 **
Absolute Differences									
Log Enrollment	-0.278 ***	-0.430 **	-0.233 ***	-0.073 ***	-0.079 ***	-0.042 ***	-0.079 ***	-0.248 ***	-0.156 ***
% Amer Indian Stu.	-0.031 ***	-0.039 ***	0.007 +	0.021 ***	-0.032 ***	-0.031 ***	0.061 ***	-0.010 ***	-0.092 ***
% Asian Stu.	-0.031 **	0.134 ***	0.037 ***	0.004	-0.159 ***	-0.001	-0.046 ***	0.119 ***	0.014 *
% Black Stu.	0.086 ***	0.048 ***	0.037 ***	0.033 ***	0.106 ***	0.004	0.076 ***	0.034 ***	0.038 ***
% Hispanic Stu.	0.065 ***	0.013	0.018 +	0.011	0.070 ***	-0.045 ***	0.002	-0.162 ***	0.091 ***
% Migrant Stu.	0.020 **	0.065 ***	0.078 ***	0.013 *	-0.011 +	0.026 ***	0.060 ***	0.140 ***	-0.019 **
% Bilingual Stu.	0.013	-0.054 ***	-0.031 ***	0.040 ***	-0.019 *	-0.011	0.067 ***	0.025 **	-0.012
% Special Ed. Stu.	0.061 ***	-0.039 ***	-0.022 ***	0.018 **	0.056 ***	-0.014 *	0.002	0.023 ***	0.078 ***
% Reduced Meals Stu.	-0.106 ***	-0.038 ***	0.000	-0.052 ***	-0.011 +	-0.009	-0.031 ***	-0.053 ***	-0.098 ***
Democrat Index	-0.031 ***	0.006	-0.088 ***	-0.004	-0.011	0.025 ***	-0.016 *	0.006	-0.001

Significance levels based on p-values from Wald test: + p<0.1, * p<.05, ** p < .01, *** p < .001

NOTE: Each observation is a pair of districts in the sample. The dependent variable is the Jaccard measure of similarity between the two CBAs, standardized over all pairs of districts. Distance variables are binary indicators, and controls are absolute differences of standardized covariates.