A Spatial Study of Teachers' Salaries in Pennsylvania School Districts*

ROBERT T. GREENBAUM
The Ohio State University, Columbus, OH 43210

I explore the relationship among teacher salaries across Pennsylvania school districts. Using techniques developed in spatial econometrics, I find that the error terms in a salary regression are spatially correlated, suggesting evidence of omitted labor market factors. I also find evidence that salaries in nearby, financially similar districts directly influence teacher salaries in a particular district, which is evidence of pattern bargaining or more informal social comparisons across districts. Econometric specifications that ignore these factors overstate the influence of own-district variables, such as economic indicators, on salary.

I. *Introduction*

Economists expect that "comparable" firms — firms that share a common labor market and are similar in other ways — will generally pay comparable salaries. There are two distinct reasons for this prediction. The first follows from standard market analysis. If the labor market is reasonably competitive, primarily market forces determine salaries; shocks that affect salaries in one firm will similarly affect other firms.

Second, when salaries are not determined in a perfectly competitive market (e.g., when salaries are set via collective bargaining), salaries of one firm might directly affect salaries in other "comparable" firms through a formal process such as pattern bargaining (Kochan, 1980; Budd, 1992) or more informally through a process of social comparisons (Festinger, 1954; Frank, 1985; Babcock et al., 1996).

These influences cause salaries to be positively correlated across firms, even after accounting for observable characteristics that affect salaries. As a result, in a standard firm-level salary regression, the residuals of firms that share a common labor market may be correlated because of omitted features of the market or because of a direct causal link between salaries in firms. Failure to account for a direct causal link can lead to inconsistent parameter estimates (Anselin, 1988; Case, 1991).

Herein, I explore the relationship among teachers' salaries in Pennsylvania school districts using techniques developed in spatial econometrics to identify whether salaries are related due to omitted labor market features or due to direct causal relationships in salaries across districts. I find empirical evidence of both types of relationships. The estimates suggest that a district's average salary increases by 6.6 percent when salaries in nearby, financially similar districts increase by 10 percent.

Section II discusses in detail why there might be direct casual relationships between salaries across firms. Section III develops the econometric model that will account for the two determinants of correlation of salaries across firms. Section IV describes the data, and section V presents the empirical results.

II. *Covet Thy Neighbor*

Social Comparisons. Psychologists contend that "people evaluate themselves by means of comparisons with others" (Messick and Sentis, 1983, p. 61). Research in sociology and economics suggests that these social comparisons are an important part of explaining both individuals' and firms' behavior (Festinger, 1954; Frank, 1985). Because workers believe they should be paid salaries commensurate with their cohorts, the relationship between salaries and the value of marginal product is often not perfect (Frank, 1985).

Social comparisons play an especially important role in the public sector because of the reduced role of competitive market forces. The comparison of salaries in public sector wage negotiations is common, as "unions and employers frequently allude to salaries paid to workers in other municipalities and to other types of workers in the same municipality" (Babcock et al., 1996, p. 2). In fact, in most levels of government, workers are entitled by law to wages comparable to those paid in the private sector for similar work, i.e., the "prevailing wage" principle (Fogel and Lewin, 1988). Indeed, when arbitrators rule in public sector contract disputes, they are often required to consider the wages paid for similar public and private sector work in comparable communities (Brown and Medoff, 1988; Kochan, 1980).

Furthermore, public sector wage and output decisions are made by elected officials who must answer to the voters and the workers rather than to the market (Valletta, 1993; Fogel and Lewin, 1988). Thus, decisions tend to be more political than in the private sector, and there is greater pressure for politically expedient solutions. It may be important politically to pay wages comparable to what others are paying. For instance, in describing a new contract agreement with the Pittsburgh Public Schools a few years ago, Pennsylvania and Pittsburgh Federation of Teachers president Albert Fondy tactically noted that the new contract was comparable to others reached both in districts locally and to settlements in urban districts of similar size across the country (Lee, 1995). Such a settlement sends a message to the teachers that they are being paid fairly, to the parents that their children will be taught by reasonably compensated teachers, and to the taxpayers (voters) that they will not be shouldering a disproportionate burden.

Pattern Bargaining. One way in which the social comparisons of salaries has been formalized is through pattern bargaining, that is, the sequentially structured relationships among settlements. With pattern bargaining, wage settlements are based on earlier settlements at other places. Typically, an agreement reached first with one employer serves as the basis for subsequent bargaining at other firms, both within and across industries (Cappelli, 1990).

Pattern bargaining in the U.S. dates back to the period just after World War II, when the War Labor Board encouraged pattern settlements as part of its efforts to devise a national wage policy (Kochan, 1980). United Auto Workers (UAW) settlements are a prominent and successful example. Rather than negotiating independently with each of the Big Three automakers and their many suppliers and subcontractors, the UAW instead targets one of the Big Three. The contract negotiated with the target company is a pattern for subsequent negotiations with other employers (Budd, 1992). This allows the union to take wages out of competition. Note that unions can only use this tactic successfully in fairly concentrated industries when there is little nonunion or international competition (Kochan, 1980).

There has been some debate in the bargaining literature as to whether or not pattern bargaining was just a passing fad. For example, Budd (1992) finds that the spillover effects from UAW target settlements were much larger in the period from 1955–1979 than during 1987–1990. Freedman (1982) and Freedman and Fulmer (1982) argue that the economic stagnation of the late 1970s and early 1980s caused a breakdown in formal pattern bargaining. They argue that during wage bargaining in the early 1980s, management became more concerned with factors internal to the firm than with external influences (Freedman, 1982; Cappelli, 1990). Ready (1990), however, argues that the variance in wages paid fell between 1977 and 1983, supposedly evidence of more patterning rather than less.

Regardless of the question of whether or not unions today are more or less likely to attempt to establish a key settlement and copy it at other locations, "it is not uncommon for one party or the other (and sometimes both) to refer to the pattern of settlements in the community during the negotiations of a wage adjustment" (Horowitz, 1994, p. 38). In some concentrated industries such as steel, companies adjusted wages in unison even before centralized bargaining or unions (Kochan, 1980). With the added presence of unions, labor unions often come to the bargaining table with data showing intercompany comparisons to back up their positions (Reynolds, 1982).

The Impact of Salary Correlation on Empirical Results. Although the importance of relationships in salaries across firms has been discussed in the literature cited above, its effect on empirical analysis has yet to be fully explored. If unobservable labor market forces are pushing salaries in the same direction, the error term of an OLS salary regression will be spatially correlated across districts that share a common labor market. As with temporal serial correlation, OLS estimation in the presence of a spatially autocorrelated error term yields unbiased coefficients, but the estimates of their standard errors will be incorrect.

When salaries negotiated in one firm directly affect the negotiations in another firm, settlements cannot be considered in isolation. Whether the relationship comes from pattern bargaining or social comparisons, the actual salaries paid by other firms do matter. Failure to consider these causal relationships causes an omitted variable bias and can lead to inconsistent parameter estimates.

III. *The Empirical Model*

Equation 1 can be considered the typical approach to estimating how a set of explanatory variables affects teacher salary. The assumption is that the error terms are independent across districts:

$$Y = X\beta + \varepsilon, \text{ with } E[\varepsilon] = 0, E[\varepsilon\varepsilon'] = \sigma^2 I, \tag{1}$$

where Y is an $(N \times 1)$ vector measuring teacher salaries; X is an $(N \times K)$ matrix of exogenous variables; β is the $(K \times 1)$ vector of parameters; and ε is the error term.

My focus is the spatial correlation in salaries that remains after controlling for the observable characteristics. Failure to account for the remaining spatial dependence has been shown to lead to inconsistent or inefficient parameter estimates, depending on the nature of the dependence (Anselin, 1988; Case, 1991). I therefore propose three modifications to the OLS model. First, I account for the possibility that the error term is no longer distributed independently but is perhaps spatially correlated. Second, I account for the direct influence that salary in some districts may have on other districts. Third, I estimate a full model that accounts for both effects.

It is difficult to capture all aspects of the labor market that could affect salary. These labor market factors not modeled will be captured by the error term, ε. Consequently, random labor shocks that affect districts near each other in similar ways will cause the error terms to be correlated across these districts.

This dependence among spatial units is often referred to as spatial autocorrelation. Analogous to the problem in time series data, in which autocorrelation refers to the correlation of an event in recent time periods with the same event in the present time period, spatial autocorrelation refers to the correlated impacts of a nearby event. As with serial correlation, OLS estimation in the presence of a spatially autocorrelated error term still yields unbiased coefficients, but the estimates of their standard errors will be incorrect. Thus, inference based on t and F statistics will be misleading and R^2 measures of fit will be incorrect (Anselin, 1992).

Unlike the problem with serial correlation in time series data, spatial autocorrelation is multidirectional in nature. Therefore, feasible generalized least squares (FGLS) procedures such as the Cochrane-Orcutt estimator are not appropriate (Anselin, 1988). However, one can account for the spatial autocorrelation by estimating a maximum likelihood regression that includes a spatially lagged error term. Equation 2 is referred to in the literature as the spatial disturbances or *spatial error model*:

$$Y = X\beta + \varepsilon;\ \varepsilon = \lambda\varepsilon_c + u, \text{ with } E[u] = 0, E[uu'] = \sigma^2 I, \tag{2}$$

where $\varepsilon_c = W_c\varepsilon$, and W_c is an $(N \times N)$ matrix that contains information about which districts share a common labor market. λ is the spatial correlation of the error term. If there is no correlation among neighbors' error terms, λ equals zero and Equation 2 is equivalent to Equation 1. The specification of W_c is discussed at the end of this section.

The spatial dependence among districts' salaries may also be the result of a more direct process. If social comparisons or pattern bargaining produce causal relationships in salaries across districts, this should be explicitly modeled by including the salary of "comparable" districts as an explanatory variable. The model that includes a lagged dependent variable as an explanatory variable is often referred to as the spatial spillover or *spatial lag model*:

$$Y = \rho Y_s + X\beta + \varepsilon, \text{ with } \mathrm{E}[\varepsilon] = 0, \mathrm{E}[\varepsilon\varepsilon'] = \sigma^2 I, \tag{3}$$

where $Y_s = W_s Y$, and W_s is the $(N \times N)$ weighting matrix that contains information about which districts compare themselves with one another. ρ is the spatial coefficient on the lagged dependent variable. The spatial lag coefficient will equal zero only if pattern bargaining and social comparisons do not affect salary.

Failure to include a measure of the salary in other school districts when it does significantly influence a district's salary is a more serious problem than failure to correct for a spatially correlated error term. The included explanatory variables that are correlated with the omitted lagged dependent variable will be inconsistently estimated by OLS and will lead to incorrect inference (Anselin, 1992).

This omitted variable bias has been demonstrated in previous research. In an examination of states' expenditures, Case et al. (1993) found that a state's spending depended on the spending of similarly situated states: ceteris paribus, a $1 increase in a state's neighbors' expenditures increased its own expenditures by more than 70 cents. Indeed, they also found that the failure to include neighbors' expenditures resulted in erroneous inferences regarding some of the other explanatory variables. The magnitudes of the coefficient estimates of the influences of population density, federal grants, elderly population, and race were different in the spatial and nonspatial models. Likewise, Doreian (1980) found biased coefficients in the nonspatial specifications of a Philippine Huk insurgency and in American presidential electoral support models. His conclusion was that "The nonspatial model estimated by conventional regression procedures is not a reliable representation and should be avoided when there is a spatial phenomenon to be analyzed" (Doreian, 1980, p. 51).

Finally, a model may contain both spatially correlated error terms (Equation 2) and spatial lags (Equation 3). If I only allow one of the effects to occur, my model may be mis-specified. For example, in the spatial error model (Equation 2), I set ρ equal to zero. As a consequence, if there is correlation among the neighbors' salaries due to pattern bargaining or other social comparisons, λ will pick it up. Similarly, if I then estimate the spatial lag model (Equation 3), ρ may be picking up some of the effect of correlated error terms. In other words, if I find that both λ and ρ are significant in the two different models, I cannot isolate the nature of the spatial dependence in order to distinguish between pattern bargaining and social comparisons or unobserved labor market conditions. Only by estimating a full model that includes both a spatial error and spatial lag term will we be able to estimate the effect of pattern bargaining and social comparisons independently of the unobserved labor market conditions. Equation 4 is the full model:

$$Y = \rho Y_s + X\beta + \varepsilon;\ \varepsilon = \lambda\varepsilon_c + u, \text{ with } \mathrm{E}[u] = 0, \mathrm{E}[uu'] = \sigma^2 I, \tag{4}$$

where $Y_s = W_s Y$, and $\varepsilon_c = W_c \varepsilon$. Below, I argue that W_s and W_c should not be the same matrix. First, different types of weighting matrices will appropriately capture the phenomenon of dependence in a spatial lag and spatial error model. Second, different weighting matrices may be required to correctly identify Equation 4 (Anselin, 1988).[1]

The weighting matrix specifies which districts are interrelated. The matrix is likely to differ, theoretically, depending on whether one wants to model a spatial lag model or a spatial error model. Therefore, I make use of two different kinds of weighting matrices: a contiguity matrix and a similarity matrix.

Contiguity Matrix for a Spatial Error Model. In a spatial error model, I attempt to capture econometric specification failure due to omission of important labor market factors that affect regional salaries. In the case of school districts, districts that are physically close are more likely to share a labor pool than are distant districts. Therefore, my weighting matrix for the spatial error model should allow spatial correlation among districts that are physically proximate.

I use a dichotomous contiguity variable as the building block for the elements of the weighting matrix. The matrix, C, is defined so that an element $C_{ij} = 1$ if districts i and j are contiguous and $C_{ij} = 0$ if not. I use "queen's case" contiguity, which means that district borders need only touch to be considered contiguous neighbors. Taking the row sum for the i^{th} row of C, C_i, produces the number of districts that border district i. The elements of an $(N \times N)$ geographic contiguity matrix, W_c, are then just:

$$W_{ij} = C_{ij}/C_i, \text{ where } C_i = \Sigma_j C_{ij}\,. \tag{5}$$

The resulting contiguity matrix causes each of a district's contiguous neighbors to have an equal influence on it. Each row of W, W_i, creates a summary of information about observation i's neighbors. For example, if i has N_i neighbors, a value of $1/N_i$ in the column corresponding to each neighbor will create a simple average when the i^{th} row is pre-multiplied by a variable.

I initially use only first-order contiguity, i.e., districts are neighbors only if they share a border.[2] Alternative schemes have included weighting based on the percentage of the border shared or the distance between centroid points — the shorter the distance, the greater the weight given to a neighbor. However, in a similar application, Case et al. (1993) found that results were rather insensitive to alternative distance measures. I estimated my models using higher-order contiguity weighting matrices, but the parameter estimates and standard errors were not markedly different than those using first-order contiguity. Therefore, the empirical results presented here use the first-order contiguity matrix.

Similarity Matrix for a Spatial Lag Model. For the spatial lag model, I create a similarity matrix. Negotiators likely look beyond their most immediate geographic neighbors when they compare salaries or attempt to set wage patterns. Another district may be considered a reference district because it is a geographic neighbor or because it is similar in terms of size, or financial or demographic characteristics. For example, Babcock and Olson (1992) found that arbitrators in Wisconsin school districts rely heavily on comparisons to other districts in the same athletic conference in cases where final arbitration is used to resolve contract disputes.

Survey results support the contention that negotiators consider settlements in districts that are both nearby and similar. Using a survey of negotiators in Pennsylvania school districts, Babcock et al. (1996) found that geographic proximity was the most important determinant of "comparability" and that financial similarity was the second most important determinant. Gerwin (1973) found that negotiators are likely to discount the salaries paid to neighboring districts that have very different characteristics.

To capture the fact that negotiators look to other districts that are similar to their own, I create a continuous measure of similarity. The more similar districts i and j are in terms of per-student income in the community,[3] the more they can be considered as referents. I start with a matrix that contains the inverse of the difference in the log personal income between each pair of districts.[4] If two districts are very close in income (X), then $(|X_i - X_j|)^{1/q}$ is small and $(|X_i - X_j|)^{-1/q}$ is large. The effect of taking the qth root of $|X_i - X_j|$ is to reduce the weighting of districts that have very similar levels of personal income. As with the contiguity matrix, I also want to normalize the inverse distance function by the row sum.

Because negotiators place a greater emphasis on the salary of nearby and financially similar districts, I need to modify the similarity weighting matrix to include only similar districts that are within a certain distance of their district. Thus, the similarity matrix is a function of the inverse difference in personal income if that district is nearby. An element of this matrix, W_s, takes on the following values:

$$W_{ij} = |X_i - X_j|^{-1/q} / [\Sigma_j |X_i - X_j|^{-1/q}] \text{ if } D[j-i] < n \text{ for } i \neq j; \quad (6)$$

$$W_{ij} = 0 \text{ otherwise,}$$

where n can be thought of as "number of districts away," and $D[j-i]$ is the distance measure between districts i and j — the number of districts physically between two districts. For example, if n is three, then comparable districts used as referents are those districts that can be reached by passing through two or fewer other districts.[5] By increasing n only slightly, I greatly increase the number of districts that are included as potential reference districts. I can get a good idea of what the cut-off should be by varying the value of n and choosing the model that maximizes the log likelihood.

IV. *Data*

Teacher salaries in Pennsylvania's public schools are bargained for locally at each of the 500 school districts, all of which are unionized. Each district is represented by either the Pennsylvania Federation of Teachers (affiliated with the American Federation of Teachers) or the Pennsylvania State Education Association (affiliated with the National Education Association). Often, the same union representative will negotiate salaries for many school districts in a region. Thus, although salaries are bargained at the local level in each district, settlements in one district are not independent of negotiated settlements in neighboring districts.

The data are from school years 1982–1983 to 1987–1988. The variables are averaged across the six school years to avoid the complications of including observations across time. Because of missing data, I conduct the analysis on 483 of the 500 districts.

The dependent variable in all models is the log of the district salary. Teachers are paid based on a salary matrix that is a function of the teacher's education level and experience. Using the average of the salaries in the district would cause unwanted variation across districts due to differences in such factors as average education and experience across districts. Instead, the salary that a full-time teacher with a bachelor's degree and 15 years of experience would receive in each district was estimated from the personnel files of the Pennsylvania Department of Education (Babcock and Engberg, 1993). The average annual salary for a teacher with a bachelor's degree and 15 years of experience over the school years 1982–1983 through 1987–1988 is $24,009.

To account for observable differences in district, community, and labor market characteristics, I include a number of control variables. The appendix describes all of the variables and their sources. The school district variables include the size of the district (enrollment) and an indicator of which teachers' union represents the district. Characteristics of the community include the average level of education for adults in the district, the percentage of registered voters who are Democrats, the percentage of children who belong to families receiving AFDC, the percentage of black students, personal income per student, and the per-student market value of property. Characteristics of the labor market include the total number of nonagricultural employees in the region, the mean annual salary in the county, the county's unemployment rate, and the region's unionization rate. Additional control variables include the violent crime rate in the county and the state senator's labor rating based on pro-labor votes. Note that some of the variables are measured at the county rather than the district level, but regression results do not appear to be sensitive to the inclusion of county-measured variables.

V. *Results*

The first step is to examine whether salaries are indeed correlated spatially. Two rough indicators suggest that they are. The first is visual. Figure 1 shows the mean salary for a teacher with a master's degree and 15 years of experience in Pennsylvania (averaged across 1982–1983 through 1987–1988). As can be seen on the map, teachers' salaries appear to be clustered into distinct areas in which similar salaries are paid. The highest paid teachers are in the eastern and western parts of the state, primarily near the cities of Philadelphia and Pittsburgh. The lower paying districts are clustered predominately in central Pennsylvania.

The second indicator is a statistical measure of spatial autocorrelation, Moran's *I* (Anselin, 1992), which measures whether a particular variable in one district moves together with that same variable in neighboring districts. If the first-order contiguity weighting matrix is used, the test is whether the mean salary in a district moves in the same direction as the salary in contiguous districts. Indeed, I find that salary is highly correlated with the salary in contiguous districts. The Moran's *I* statistic for salary is

Figure 1

Pennsylvania School District Salary
1986–87 School Year

.8180 and the corresponding Z-value is 29.932 (P-value < 0.000001). Inference is based on the standard normal, so I find that salary is very significantly correlated across neighboring districts.

Undoubtedly, much of the correlation in the salary of neighbors is due to similarities in many of the measurable characteristics that I include as explanatory variables. To control for these characteristics, I estimate regressions.[6] The first column of Table 1 presents the OLS regression results (Equation 2, in which $\lambda=\rho=0$). Even when I control for the independent variables by estimating the OLS model, there remains correlation in salary across neighboring districts. Moran's I was reduced to 0.3447, and the Z value is 13.5384, which is still clearly significant (P-value < 0.000001). The problem can also be seen on the map of the OLS residuals (Figure 2), which shows that the residuals are not randomly distributed across space. For districts in which the model under-predicted salary, the neighbors' salary was also likely to be under-predicted. This is most evident in the west-central, northeast, and southeast parts of the state. Likewise, when the model over-predicted, it often over-predicted for the neighboring districts as well. This can especially be seen in the southwest, central, south-central, and three different places in the eastern part of the state.

When comparing the different spatial models, R^2 will not be useful since the spatial models contain nonspherical errors (Anselin, 1992). To directly compare the alternative models to the OLS specification, two goodness-of-fit measures are reported for all models: the maximized log likelihood and the Akaike Information Criterion (AIC) (Akaike, 1981; Buck and Hakim, 1981). The fit is better the higher the log likelihood or the lower the AIC. Furthermore, results of a likelihood ratio test are provided.

In a first attempt to take account of the spatial dependence, I estimate the spatial error (Equation 2) maximum likelihood model with a first-order geographic contiguity weighting matrix. These results are reported in Column (2) of Table 1. Based on the likelihood ratio tests, the spatial error model fits significantly better than the OLS model. The coefficient on the spatial error term, λ, is 0.759, which is large and significant at the 0.001 level. The finding that neighbors' errors are significantly spatially correlated supports the contention that some unobserved labor market conditions may be causing neighbors to pay salaries similar to one another.

Next, I estimate the spatial lag model (Equation 3) using a weighting matrix based on both similarity of personal income and proximity (Equation 6). The matrix used is the row-standardized inverse of the cube root of the difference in personal income of districts that are within two districts of one another. Results are presented in Column (3) of Table 1. In this specification, ρ is large (0.660) and significant at the .001 level: For every 10 percent increase in contiguous neighbors' salaries, a district's salary will increase 6.6 percent. This is evidence that either pattern bargaining or social comparisons (or both) may exist across financially similar districts near each other.

In both the spatial error and lag models, the coefficient estimates on the control variables are generally smaller than in the OLS estimation. This is an indication that

Table 1

Salary Regressions, 483 Pennsylvania School Districts

(Dependent Variable: Average district salary[a]; standard errors in parentheses)

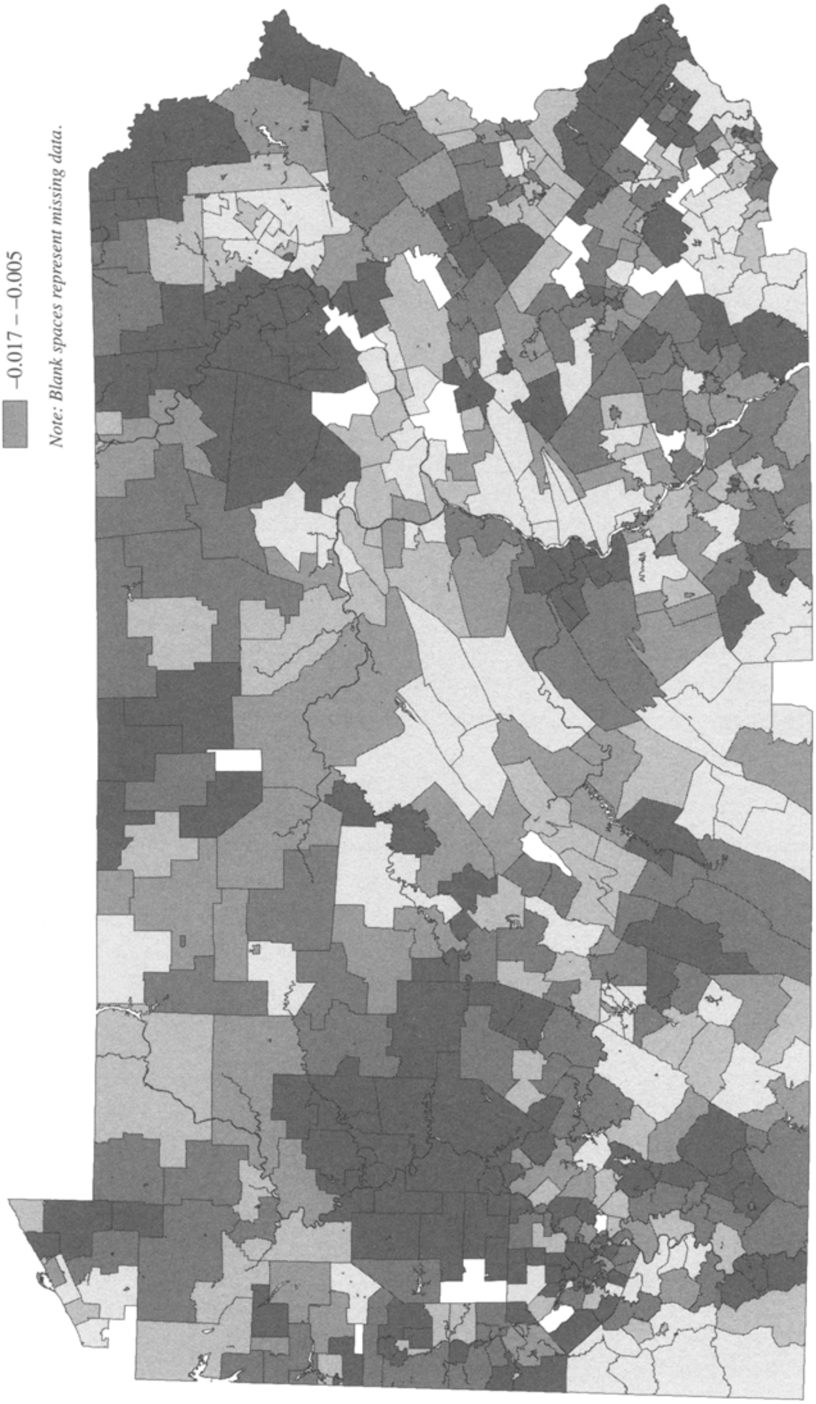
Coefficient	OLS	Spatial Error	Spatial Lag	Full Model
Spatial Lag (ρ)			0.660****	0.656****
			(0.043)	(0.051)
Spatial error (λ)		0.759****		0.260****
		(0.036)		(0.067)
AFDC	0.119**	0.024	0.049	0.039
	(0.053)	(0.043)	(0.042)	(0.042)
Affiliation	0.041***	0.022**	0.029***	0.025**
	(0.013)	(0.011)	(0.011)	(0.011)
Alternative wage	0.181****	0.100***	0.067***	0.068***
	(0.026)	(0.031)	(0.022)	(0.025)
Black	–0.123*	–0.078	–0.108*	–0.096*
	(0.069)	(0.056)	(0.055)	(0.055)
Crime	0.015	–0.017	–0.002	–0.006
	(0.023)	(0.025)	(0.018)	(0.020)
Democrat	0.059**	0.042	0.051**	0.052**
	(0.027)	(0.036)	(0.021)	(0.025)
Education	0.040****	0.029****	0.027****	0.028****
	(0.006)	(0.006)	(0.005)	(0.006)
Enrollment	0.025****	0.023****	0.027****	0.027****
	(0.004)	(0.003)	(0.003)	(0.003)
Local unionization	0.407****	0.315**	0.205****	0.216***
	(0.072)	(0.132)	(0.059)	(0.072)
Market size	0.015****	0.011***	0.004**	0.004*
	(0.002)	(0.003)	(0.002)	(0.002)
Personal income	–0.029**	–0.030**	–0.030***	–0.032***
	(0.014)	(0.014)	(0.011)	(0.012)
Property values	0.061****	0.039***	0.031****	0.030***
	(0.011)	(0.011)	(0.009)	(0.010)
Senate	0.079***	0.037	0.007	0.010
	(0.026)	(0.030)	(0.021)	(0.024)
Unemployment	–0.002	–0.004	–0.002	–0.002
	(0.002)	(0.002)	(0.001)	(0.002)
Constant	0.410	1.658	–0.008	0.036
	(0.269)	(0.315)	(0.218)	(0.245)
Likelihood	750.803	831.306	844.811	860.440
LR test		161.007****	188.016****	219.247****
R squared	0.729			
AIC	–1,471.61	–1,632.61	–1,657.62	–1,688.88

Notes: [a] Log of annual salary for a teacher with bachelor's degree and 15 years of experience.

Significance levels: (*, **, ***, ****) = (0.1, 0.05, 0.01, 0.001).

Figure 2

OLS Residuals



in the nonspatial, OLS model the other coefficients were picking up part of the influence of the omitted neighbors' salary.

Although both spatial models are an improvement over the nonspatial specification, it is difficult to say whether the spatial error or spatial lag is the better model. Aside from the different implications about the cause of the spatial relationships, both models have similar coefficient estimates. Doreian (1980, p. 43) warns that the choice between the two models "is not one that is readily settled by examining the estimation outcomes from the two specifications. Rather, the choice concerning the way in which spatial autocorrelation is to be dealt with is made on substantive grounds."

On substantive grounds, I would like to control for both spatial effects. Indeed, it may be misleading to include one spatial term without controlling for the other — the coefficient on the included term is likely to pick up some of the influence of the excluded term. In order to isolate the direct causal influences, I estimate a full model that controls for both a spatial error and a spatial lag (Equation 4).

In the full model (Column (4), Table 1), the weighting matrix for the spatial lag component is the similarity matrix, W_s, which allows the influence of districts that have similar levels of personal income and are within two districts away. The weighting matrix for the spatial error component is the first-order geographic contiguity matrix, W_c, which allows spatially correlated error terms for districts that share a border. Both spatial coefficients are positive and significant. The spatial error coefficient, λ, is 0.260 and is significant at the .001 level, suggesting that unobserved shared labor market forces push the salaries of neighboring school districts together. Controlling for the spatially correlated error term, ρ is 0.656 and significant at the .001 level. This coefficient suggests that a district's salary increases by 6.6 percent when salaries of financially similar and proximate districts increase by 10 percent, which supports the hypothesis that salary comparisons, via pattern bargaining or social comparisons, influence the salary determination process.

Holding constant the effect of neighbors' salaries on a district's salary, I find that a number of the district and community characteristics have an influence on salary. Districts with higher property values per student pay higher salaries. This is not surprising, since much of the funding for salaries comes from property taxes. What is somewhat surprising is the negative coefficient on the personal income variable: Controlling for property values, as personal income increases, mean salary falls. One possible explanation may be that teachers are willing to accept lower pay to teach the children of wealthier parents. Another possibility is that children of wealthier parents may be more likely to attend private schools. If that is the case, one might expect that those parents would have less of a desire to insure that public school teachers are paid well in their own school district.

With regard to labor market conditions, the coefficient on alternative wage is large and significant. For every 10 percent increase in the mean annual salary in the county of the district, the mean teacher salary increases by 0.7 percent. The county unemployment rate does not significantly affect salary.

The estimated coefficients from the full model differ from those from the OLS specification in a number of ways. The size of the coefficient on the level of local unionization was halved from 0.407 to 0.216. Furthermore, the coefficient on the state senator's labor rating dropped substantially from 0.079 to 0.010 and is no longer significant. Although the coefficient on enrollment remained virtually unchanged, the market size coefficient is now much smaller, dropping from 0.015 to 0.004. In the full model, the percentage of children in AFDC families no longer significantly affects salary. The coefficients on both alternative wage and property values fell by at least half. Taken as a whole, the district-specific factors in the model that controls for spatial correlation are much less important than the estimates from the OLS model would indicate.

Inspection of the map of residuals from the full spatial model (Figure 3) leads to the observation that the residuals look much less clustered than the OLS residuals. Thus, I can conclude that much, but not all of the spatial autocorrelation has been removed. Clustered residuals, such as the group of under-predictions in the west-central part of the state, remain.

VI. *Summary*

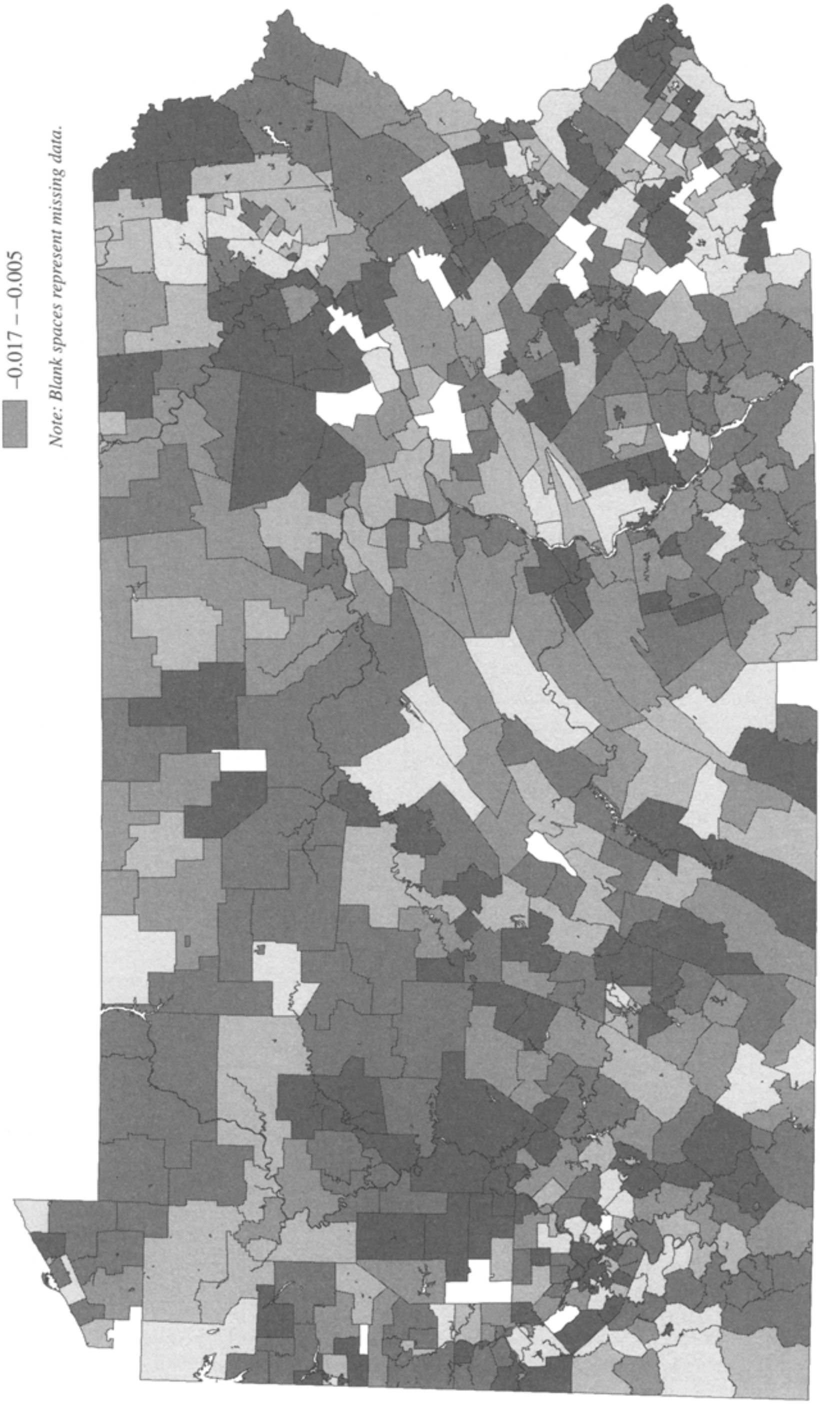
The empirical results indicate that controlling for measurable characteristics, teacher salaries in Pennsylvania are spatially correlated. I estimate a model that allows for two types of spatial correlation. Using a spatial error term, I find that the residuals of districts near one another are significantly positively correlated. Thus, in my regressions, there may be unobserved labor market conditions that force salaries of neighboring districts together. These could include unmeasured alternative opportunities, cost of living differences, or regional differences in the supply of available teachers.

The model also allows salaries to be correlated for more direct reasons. Pattern bargaining or social comparisons may cause salary in a district to be directly influenced by salaries in other districts. The estimates suggest that a district's salary increases by 6.6 percent when the salary of financially similar and proximate districts increases by 10 percent. I also find that the failure to account for both types of spatial correlation leads to an overstatement of the influence of variables, such as economic indicators, on salary. This finding is consistent with the research of Doreian (1980) and Case et al. (1993).

Finally, it is important to use two different weighting matrices for the spatial lag and spatial error components. While failure to do so leads to identification problems, it is also important to consider the underlying processes leading to spatial correlation in the two models. Because the use of contiguous geographic neighbors is plausible for identifying districts that share a common labor market, I use a contiguity matrix in the spatial error model. To address which districts matter for purposes of social comparison and pattern bargaining, I estimated models with weighting matrices that account for similarity and well as proximity of other districts. I find that districts are most likely to imitate the salary decisions of other districts that are both nearby (within two dis-

Figure 3

Full Model Residuals



tricts away) and have similar levels of personal income. Thus, negotiators weight their closest neighbors differentially based on the similarity of their financial characteristics. This econometric finding supports the survey findings that negotiators look to nearby districts with similar financial conditions for basis of comparison.

NOTES

*I thank colleagues Linda Babcock and John Engberg for very valuable contributions and for use of their data. Thanks also to Wil Gorr, who helped with some of the GIS aspects of this research. A National Science Foundation grant supported data collection.

[1]It can be shown with substitution that if $W_c = W_s = W$, the full model from Equation (4) can be expressed as $Y = (\rho+\lambda)WY - \lambda\rho W^2 Y + X\beta - \lambda WX\beta + u$. In this case, it is possible to obtain two different estimates of ρ, and identification becomes an issue (Anselin, 1988).

[2]Second-order contiguity, for example, would consider two districts that border a common district to be neighbors.

[3]For completeness, I also examined three other measures of similarity: property values, the average level of education in the adult population, and the size of the district (based on enrollment). Based on the results of spatial lag model regressions, none of these measures of similarity was an improvement over per-student income.

[4]The inverse-distance matrix is seen more commonly in the geography literature, where X is the physical distance between two points.

[5]For example, consider states instead of school districts. If $n = 3$, then, moving westward, Pennsylvania would consider Illinois but not Iowa as a neighbor.

[6]All regressions were estimated using SpaceStat, version 1.8 (Anselin, 1992).

REFERENCES

Akaike, Hirotugu. "Likelihood of a Model and Information Criteria." *Journal of Econometrics* 16 (May 1981): 3-14.

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

________. *SpaceStat Tutorial: A Workbook for Using SpaceStat in the Analysis of Spatial Data.* Technical Report S-92-1. National Center for Geographic Information and Analysis. University of California, Santa Barbara, Calif., 1992.

Babcock, Linda and John Engberg. "Wages and Quits in the Unionized Public Sector." Heinz School of Public Policy and Management Working Paper 93-52. Pittsburgh: Carnegie Mellon University, 1993.

Babcock, Linda and Craig Olson. "The Causes of Impasses in Labor Disputes." *Industrial Relations* 31 (April 1992): 348-60.

Babcock, Linda, Xianghong Wang, and George Loewenstein. "Choosing the Wrong Pond: Social Comparisons in Negotiations That Reflect a Self-Serving Bias." *Quarterly Journal of Economics* 111 (February 1996): 1-19.

Brown, Charles C. and James L. Medoff. "Employer Size, Pay, and the Ability to Pay in the Public Sector." In Richard Freeman and Casey Ichniowski, eds. *When Public Sector Workers Unionize.* Chicago: University of Chicago Press, 1988.

Buck, Andrew J. and Simon Hakim. "Appropriate Roles for Statistical Decision Theory and Hypothesis Testing in Model Selection." *Regional Science and Urban Economics* 11 (February 1981): 135-47.

Budd, John W. "The Determinants and Extent of UAW Pattern Bargaining." *Industrial and Labor Relations Review* 45 (April 1992): 523-39.

Cappelli, Peter. "Is Pattern Bargaining Dead? A Discussion." *Industrial and Labor Relations Review* 44 (October 1990): 152-55.

Case, Anne C. "Spatial Patterns in Household Demand." *Econometrica* 59 (July 1991): 953-65.

________, Harvey S. Rosen, and James R. Hines, Jr. "Budget Spillovers and Fiscal Policy Interdependence." *Journal of Public Economics* 52 (October 1993): 285-307.

Doreian, Patrick. "Linear Models with Spatially Distributed Data: Spatial Disturbances or Spatial Effects." *Sociological Methods and Research* 9 (August 1980): 29-60.

Festinger, Leon. "A Theory of Social Comparison Processes." *Human Relations* 7 (May 1954): 117-40.

Freedman, Audrey. "A Fundamental Change in Wage Bargaining." *Challenge* 25 (July/August 1982): 14-17.

________ and William E. Fulmer. "Last Rites for Pattern Bargaining." *Harvard Business Review* 60 (March-April 1982): 30-48.

Fogel, Walter and David Lewin. "Wage Determination in the Public Sector." In David Lewin, Peter Feuille, Thomas Kochan, and John Delaney, eds. *Public Sector Labor Relations*. Lexington, Mass.: D.C. Heath, 1988.

Frank, Robert. *Choosing the Right Pond: Human Behavior and the Quest for Status*. New York: Oxford University Press, 1985.

Gerwin, Donald. "An Information Processing Model of Salary Determination in a Contour of Suburban School Districts." *American Educational Research Journal* 10 (Winter 1973): 5-20.

Hirsch, Barry T. and David A. Macpherson. "Union Membership and Contract Coverage Data from the Current Population Survey." Accompanying Data Appendix. Florida State University, May 1992.

Horowitz, Morris A. *Collective Bargaining in the Public Sector*. New York: Lexington Books, 1994.

Kochan, Thomas A. *Collective Bargaining and Industrial Relations*. Homewood, Ill. Richard D. Irwin, 1980.

Lee, Carmen J. "City Teachers OK $6,000 Over 4 Years." *Pittsburgh Post-Gazette*, January 5, 1995.

Messick, David M. and Keith Sentis. "Fairness, Preference, and Fairness Biases." In D. M. Messick and K. S. Cook, eds. *Equity Theory: Psychological and Sociological Perspectives*. New York: Praeger, 1983.

Ready, Kathryn J. "Is Pattern Bargaining Dead?" *Industrial and Labor Relations Review* 43 (January 1990): 272-79.

Reynolds, Lloyd G. *Labor Economics and Labor Relations*, 8th ed. Englewood Cliffs, N.J.: Prentice-Hall, 1982.

Valletta, Robert G. "Union Effects on Municipal Employment and Wages: A Longitudinal Approach." *Journal of Labor Economics* 11 (July 1993): 545-74.

Appendix

Variable Definitions and Descriptive Statistics[a]

Dependent Variable

Salary	Natural log of annual salary for teacher with bachelor's degree and 15 years of experience (in thousands of dollars) Source: Pennsylvania Department of Education Personnel Files mean = 3.173, standard deviation = 0.1, min = 2.95, max = 3.599

Independent Variables

AFDC	Percent of children in district in families receiving AFDC Source: Pennsylvania Department of Education mean = 0.085, standard deviation = 0.09, min = 0, max = 0.566
Affiliation	Dummy variable representing the national affiliation of district's local union Source: Pennsylvania State Education Association mean = 0.038, standard deviation = 0.19, min = 0, max = 1
Alternative wage	Natural log of the mean annual salary in county of district Source: Pennsylvania Department of Labor and Industry mean = 9.738, standard deviation = 0.13, min = 9.35, max = 9.951
Black	Percentage of students in district who are black Source: Pennsylvania Department of Education mean = 0.025, standard deviation = 0.06, min = 0, max = 0.576
Crime	Violent crime rate in county of district (multiplied by 100) Source: Federal Bureau of Investigation mean = 0.2, standard deviation = 0.14, min = 0.04, max = 1.013
Democrat	Percent of registered voters in county who are Democrats Source: Pennsylvania Department of State, Bureau of Elections mean = 0.489, standard deviation = 0.15, min = 0.25, max = 0.785
Education	Average number of years educated beyond 8th grade for people age 25 and older in the district, coded 0 (no high school) through 8 (college graduate) Source: 1980 and 1990 Census STF3a (linear interpolation) mean = 4.048, standard deviation = 0.64, min = 2.68, max = 6.528
Enrollment	Natural log of the number of students in the district Source: Pennsylvania Department of Education mean = 7.782, standard deviation = 0.67, min = 5.65, max = 12.2
Local unionization	Percent of local labor market employees that are unionized Source: Hirsch and Macpherson, 1992 mean = 0.228, standard deviation = 0.04, min = 0.12, max = 0.309
Market size	Natural log number of nonagricultural employees in the local labor market (thousands) Source: Pennsylvania Department of Labor and Industry mean = 4.913, standard deviation = 1.66, min = 0.36, max = 7.62
Personal income	Natural log of personal income per student in the district Source: Pennsylvania Department of Education mean = 10.82, standard deviation = 0.42, min = 9.98, max = 12.67
Property values	Natural log of per student market value of property Source: Pennsylvania Department of Education mean = 11.33, standard deviation = 0.42, min = 10.5, max = 12.82
Senate	State senator's labor rating: percent of "correct" votes on labor issues in Pennsylvania State Senate Source: Pennsylvania State Education Association mean = 0.769, standard deviation = 0.13, min = 0.62, max = 0.978
Unemployment	Unemployment rate in county of district Source: Pennsylvania Department of Labor and Industry mean = 9.263, standard deviation = 2.74, min = 4.68, max = 14.68

Note: [a]The sample size is 483. Data are averaged over 1983–1988.