Donald J. Lacombe

Does Econometric Methodology Matter? An Analysis of Public Policy Using Spatial Econometric Techniques

A popular approach to examining the effects of public policy has been to rely on a spatial data sample of border counties as in Holmes (1998)—border counties from a sample of states that are used in conjunction with least-squares estimation techniques in an attempt to isolate the policy impact while controlling for spatial dependence that often arises from latent or unobserved variables. This technique is in the spirit of control-group methodologies from the laboratory sciences. This paper contrasts border-county estimation results from Holmes' (1998) approach and those from a related methodology set forth in Holcombe and Lacombe (2003), with estimates from a spatial autoregressive model explicitly accounting for within-state and between-state public policy effects. As an illustration, the paper examines the effects of Aid to Families with Dependent Children (AFDC) and Food Stamp payments on female-headed households and female labor force participation using the three different methods.

1. INTRODUCTION

Economists have long been interested in applied policy questions, and the answers to these questions may vary depending upon the means used to control for certain unobserved factors, some of which may vary geographically in a systematic manner. Various econometric techniques have been developed to address problems that can bias estimates and associated inferences. Early work tended to ignore the fact that latent or unobserved variables can vary systematically over geographical regions, creating spatial dependence among cross-sectional observations on states or counties. More recently, a general recognition that policy effects may exhibit spatial correlation between nearby geographical regions has led to the adoption of quasi-experimental control-group methodologies (Holcombe and Lacombe 2003; Holmes 1998; Isserman and Rephann 1995). These approaches attempt to control for spatial dependence using border counties from neighboring states as the sample data. If this

The author would like to thank the Urban Affairs Center at the University of Toledo for generously providing the county-level time-series data used in this study. James P. LeSage provided assistance in constructing maximum-likelihood estimation functions used in the study. Any remaining errors are the sole responsibility of the author.

Donald J. Lacombe is a professor at the Ohio University, Department of Economics (lacombe@ohio.edu).

Geographical Analysis, Vol. 36, No. 2 (April 2004) The Ohio State University
Submitted: November 1, 2002. Revised version accepted: September 22, 2003.

approach is successful in eliminating spatial autocorrelation, least-squares estimates should possess desirable properties on which to draw inferences regarding the significance and magnitude of policy effects. Section 2 of the paper provides a literature review of these approaches.

This paper examines the question of whether the border-county methodology is successful in dealing with spatial dependence, using the effect of Aid to Families with Dependent Children (AFDC) and Food Stamp programs on female headed households as well as female labor force nonparticipation as an illustrative example. The empirical results indicate strong evidence of spatial dependence between neighboring counties in the same state as well as neighboring counties from adjacent states. This suggests that least-squares estimates based on the border-county approach to controlling for latent geographical factors may be biased and inconsistent. Section 3 presents estimation results for the border-county methodologies of Holmes (1998) and Holcombe and Lacombe (2003), using an illustrative dataset on AFDC and Food Stamp Programs for a sample of 1129 border counties.[1] Section 4 of the paper sets forth an alternative specification and estimation methodology that explicitly models spatial dependence in the dependent variable. Two spatial autoregressive lagged dependent variables are used, one involving nearby counties from within the state (i.e., intrastate), and a second based on bordering counties in neighboring states (i.e., interstate). It is argued that this approach holds several theoretical as well as econometric advantages over the border-county matching methodology. Estimates based on the AFDC and combined AFDC and Food Stamp policy effects model are used to illustrate these advantages. Conclusions can be found in Section 5, where theoretical and applied advantages of the methodology proposed here are summarized. In terms of econometric advantages, the example suggested here allows for an increase in the sample size, which can improve predictive fit as well as precision of the estimates. Since these aspects of modeling policy effects are often important, this represents an applied advantage. The increased sample size comes from inclusion of neighboring counties within the states in the spatial autoregressive model, observations that are ignored in the border-county approaches. Although these additional observations are included in the spatial autoregressive model, a separate parameter is assigned to these influences, allowing within-state impacts to be separated from those associated with between-state effects. This provides a theoretical advantage, in that the relative influence of latent or unobservable factors arising from cultural or geographical proximity of neighboring observations within-states are explicitly accounted for by the model. Estimates and inferences regarding the relative importance of these within-state effects are available directly from the proposed model as are estimates and inferences concerning between-state policy effects.

2. AN INTRODUCTION TO THE LITERATURE REGARDING BORDER-COUNTY TECHNIQUES

As stated earlier, researchers have come to recognize that geographical factors can play an important role in determining the effects of public policy. If unobserved or latent variables vary systematically over geographic regions, spatial dependence between cross-sectional observations on the dependent variable can arise. As in the case of simultaneity, least-squares estimates will be biased and inconsistent (Kelejian and Prucha,1998). Various sampling techniques have been devised to account for vari-

1. The approach used by Holcombe and Lacombe (2003) relies on a "matched" border-county technique and as a result increases the sample size to 1319 border counties due to the fact that some counties in one state touch more than a single county in another state.

ables that can vary geographically. For example, Holmes (1998) looks at border counties to examine the effect of right-to-work laws on manufacturing employment and finds that right-to-work laws increase manufacturing employment relative to states that do not have such a law. In another frequently cited study, Card and Krueger (1994) empirically investigated the impact of minimum-wage laws by comparing similar restaurants in New Jersey and eastern Pennsylvania, restricting their sample to these two locations. Fox (1986) used matched metropolitan areas to conclude that increases in state and local sales-tax rates reduce the level of retail activity for two of the three metropolitan areas studied. Isserman and Rephann (1995) compared counties within Appalachia with a "twin" county outside of Appalachia and determined that between 1969 and 1991, Appalachian counties grew faster than their control group "twins." Bronars and Lott (1998) examine concealed-weapons laws using differences along state borders, concluding that concealed handguns deter criminals.

The common thread linking all of these studies is their attempt to control for unobserved spatial variation using strategic spatial selection of the sample observations. The rationale for this approach is that geographic differences should be minimized across state borders, while variation in policy impacts are more easily detected, producing more precise estimates of the public policy effects.

3. LITERATURE ON AFDC PAYMENTS AND AN ANALYSIS OF AFDC AND FOOD STAMP PAYMENTS USING TWO BORDER-COUNTY METHODOLOGIES

The literature on the effects of Aid to Families with Dependent Children (AFDC) and Food Stamp Payments on female-headed households and female labor-force participation is voluminous. To set the stage for estimation results presented here, a brief literature review of existing studies on the subject is provided. One of the earliest studies, Garfinkel and Orr (1974), finds that the AFDC program reduced labor supply of mothers eligible for the program. In a widely cited study, Levy (1979) argues that analysts should look at the impact on all female households, rather than just those who participate in the program, and finds that more generous AFDC benefits decrease hours worked. Bieker (1981), looking at the effect of AFDC in Delaware, found that more generous benefits increased participation in the program and reduced labor-force participation. Danziger, Haveman, and Plotnick (1981) also found that AFDC payments had a negative effect on labor supply, but note several reasons why there may be ambiguities that make it difficult to estimate the magnitude of these effects. Moffitt (1986) examined the effect of 1981 changes in the structure of the program, and consistent with other studies, found that reducing benefits to offset earned income led to a reduction in earnings of women eligible to participate. Cloutier and Loviscek (1989) find that local conditions, and especially the unemployment rate, can affect program participation rates. Consistent with other studies, Robins (1990a, 1990b) and Nord and Sheets (1990) both find that reduced benefit levels result in reduced participation rates in AFDC, although each study uses different methodologies and data. Walters (1990) finds that because transfers can blunt the incentive to work, more generous AFDC payments can increase the measured poverty rate. Smith (1991) concludes that AFDC recipients tend to migrate from lower benefit states to higher benefit states; in contrast, Levine and Zimmerman (1999) find little evidence that there is widespread migration in response to welfare benefits. Moffitt (1992), in an extensive review of the literature, notes that AFDC reduces labor supply, that more generous benefits increase program participation, and that these programs increase the likelihood of single-parent households. Yelowitz (1995) found that more generous Medicaid eligibility reduced AFDC participation, because more families could participate in Medicaid without being on AFDC. Chapman, Duncan, and Gray (1998), in contrast with most literature, find that for people who earn income while on AFDC, welfare payments do not appear to make individuals less likely to pursue work.

The main conclusions one draws from previous empirical research is that women tend to behave in the expected manner and reduce their labor supply in the face of increased public assistance payments, move from states with less generous benefits to states with more benefits, and are more likely to be single parents. The obvious differences between these studies are: (1) the data used, and (2) the fact that only some of the methodologies attempt to control for the influence of latent geographic factors. For example, as Levine and Zimmerman point out, "...states differ in their welfare generosity and a myriad of other characteristics" (1999, p. 397). The fact that states differ in many unobservable ways that will not be captured by explanatory variables can lead to spatial dependence between the dependent variable observations. As already noted, this results in a situation akin to simultaneity, where the dependence is across equations rather than between observations in the spatial cross-section. Ordinary least-squares estimates may exhibit bias in the face of spatial dependence, providing misleading inferences regarding the effects of welfare programs (or other public policy impacts). In terms of geographic differences, Levine and Zimmerman note, "...if low benefit states have a nicer climate, then the flow of treatments out of the state would be reduced relative to the flow of treatments out of higher benefit states with a worse climate" (1999, p. 397).

While very few studies of the effects of AFDC explicitly mention geography as a factor that can exert an important influence on the results, this paper compares three different methodologies that explicitly take spatial location into account. One methodology will be the border-county technique similar to the one employed by Holmes (1998). Another methodology will be the "matched" border-county technique used by Holcombe and Lacombe (2003) in their analysis of the effects of AFDC and Food Stamp payments on female-headed households and female labor force nonparticipation. The final methodology is a spatial econometric technique that takes explicit account of within-state as well as between-state effects. In this paper, a spatial autoregressive model that utilizes two spatial weight matrices to separate out the within-state and between-state effects is developed and utilized.

3.1. Sample and Variables

The sample used in this study is comprised of all of the border counties in the contiguous United States.[2] A total of 1129 county-level observations are used, with the exception of the AFDC and combined AFDC and Food Stamp variables, which are state-level programs and are therefore state level observations. The year 1990 is used primarily because the data for the dependent and independent variables are accessible for this year.[3] The dependent variables used in this study will be the percentage of households headed by a female with no husband and children under the age of 18 as well as females aged 16 and over not in the labor force as a percentage of the total population of females 16 years old and older, as a measure of female labor-force participation[4].

2. The border counties that border both Canada and Mexico are excluded from the sample as well as observations from Alaska and Hawaii.

3. The data used in this study was obtained from the CensusCD (Geolytics, Inc.) and the Regional Economic Information System, 1969-1999 which is accessible via the internet at http://fisher.lib.virginia.edu/reis/. Data on AFDC and Food Stamp payments are available in Overview of Entitlement Programs, 1990 Green Book: Background Material and Data on Programs Within the Jurisdiction of the Committee on Ways and Means, Committee on Ways and Means, U.S. House of Representatives, Washington, D.C., 1990, 553–55.

4. A referee pointed out that using percentages in this manner may introduce a limited-dependent variable problem in that the dependent variables will be bound between zero and one. Two items are worth mentioning in regards to this comment. First, the female, no-husband-dependent variable has a minimum

The independent variables are also county-level observations, with the noted exception of the AFDC and Food Stamp Variables. The AFDC and Food Stamp variables are the maximum AFDC grant in a state and the combined maximum AFDC and Food Stamp payment in a state. Other independent variables used in this study are variables used in similar studies and include such items as: population density, median age of the population, average wage per job per week, average tax rate, percentage of the population with a high school diploma, percentage of the population employed in manufacturing, local per-capita government expenditures, the unemployment rate, and percentage of the population living in urban areas. Additionally, all of the models estimated have had the variables standardized, i.e. the variables have had their means subtracted and have been divided by their standard deviations.

The first empirical technique to be used will be a border-county approach that tries to control for unobserved geographic factors that may be inherent in the data. This technique is similar to that employed by Holmes (1998) in his examination of the effects of right-to-work laws on manufacturing employment. The rationale for using the border counties in each individual state is that geographic differences across the state line should be minimized given their close proximity. If one were to examine the effects of right-to-work laws on manufacturing employment at the state level, incorrect inferences may occur because other factors relating to geography may not be taken into account. For example, Florida is a right-to-work state while Maine is not. Empirical investigation may reveal that manufacturing employment is higher in Florida relative to Maine, and this effect might be attributed to the fact that Florida is a right-to-work state while Maine is not. However, there are many geographic differences between Maine and Florida, one of which is weather. Unless this geographic difference is taken into account, the higher relative manufacturing employment in Florida may be attributed to its right-to-work status, rather than some uncontrolled for geographic characteristic. By using border counties on either side of a state line where policy differs, it is hoped that this sampling technique will control for unobserved geographic factors.

The first empirical model will examine the effects of AFDC payments and combined AFDC and Food Stamp payments on both female-headed households and female labor-force nonparticipation. A total of four models are presented, two for each dependent variable with each of the measures of public assistance. All of these models use the border-county sample and ordinary least-squares (OLS) estimation.

Estimates from the first model presented in second column of Table 1 use female headed households as the dependent variable and the list of independent variables described earlier. This model explains approximately 50% of the variation in the dependent variable and the AFDC measure of public assistance is insignificant, a result that is inconsistent with previous studies. Other independent variables have the correct sign and are significant or have the correct sign and are insignificant. For example, the median-age variable has a negative-coefficient estimate and is significant, indicating that as the median age rises, we should see fewer female-headed households. However, other independent variables, such as the wage rate, have the expected sign, but are insignificant. The third column of Table 1 contains the same

value of 0 and a maximum value of 0.2155 with only one of those values exactly equal to zero. The female labor-force-participation dependent variable has a minimum value of 0.1669 and a maximum value of 0.5705, so there does not appear to be any limited-dependent variable problem in the dependent variables. Second, the spatial autoregressive models were reestimated with log(dependent variable) on the right-hand side and log(denominator of dependent variable) on the left hand side in addition to all of the other independent variables. The results indicated that the signs, magnitudes, and significance of the coefficient estimates were virtually identical to those reported in Table 3. I wish to express my gratitude to the referee for pointing this potential problem out.

TABLE 1

Effect of Combined AFDC and Food Stamp Payments and the Effect of Maximum AFDC Grant on Female Headed Households and Female Labor Force Non-Participation OLS Results (t-statistics in parentheses)

	Dependent Variables			
Independent Variable	Female Heads, No Husband, Children 0–17 as a % of Total Households	Female Heads, No Husband, Children 0–17 as a % of Total Households	Females Age 16+ Not in Labor Force as a % of Female Population	Females Age 16+ Not in Labor Force as a % of Female Population
Combined Maximum AFDC Grant and Food Stamp Payment 1990		−0.021159 (−0.77)		−0.074808 (−3.20)***
Maximum AFDC Grant 1990	−0.016184 (−0.60)		−0.072037 (−3.12)***	
Average Tax Rate	−0.065057 (−2.12)**	−0.064945 (−2.12)**	−0.041439 (−1.57)	−0.041239 (−1.57)
Percentage of Population that are High School Graduates 1990	−0.360506 (−10.99)***	−0.357200 (−10.78)***	−0.449515 (−15.95)***	−0.446580 (−15.68)***
Manufacturing Employment % 1990	−0.067638 (−2.65)***	−0.067639 (−2.65)***	−0.335781 (−15.29)***	−0.336575 (−15.34)***
Average Wage per Job per Week	−0.015670 (−0.65)	−0.015754 (−0.65)	0.132399 (6.37)***	0.132581 (6.38)***
Median Age 1990	−0.377009 (−15.77)***	−0.376096 (−15.69)***	0.530562 (25.82)***	0.531836 (25.82)***
Population Density 1990	0.016439 (0.71)	0.016922 (0.74)	−0.048663 (−2.46)**	−0.048154 (−2.43)**
Local Per Capita Government Expenditures	0.106816 (3.41)***	0.107400 (3.44)***	−0.105727 (−3.93)***	−0.106602 (−3.98)***
Unemployment Rate % 1990	0.253552 (10.08)***	0.254136 (10.10)***	0.246185 (11.39)***	0.246583 (11.40)***
Urban Population % 1990	0.276023 (10.13)***	0.275405 (10.11)***	−0.070902 (−3.03)***	−0.071548 (−3.05)***
Number of Observations	1129	1129	1129	1129
Adjusted R^2	.4989	.4990	.6297	.6299

**** significant at the 1% level. ** significant at the 5% level. * significant at the 10% level* specification as the second with the exception of the measure of entitlement payments, which is now the combined AFDC and Food Stamp payment. The results from this specification are similar to the results when only the maximum AFDC payment was utilized.

The next set of results contained in Table 1 will examine the same specifications as in columns two and three of Table 1 but replace the dependent variable with females aged 16 and over not in the labor force as a percentage of the total population of females 16 years old and older. The results in the third column indicate that a good portion of the variation in the dependent variable is explained by the independent variables. Here the results indicate that the AFDC variable is significant, but of opposite sign from previous studies where public assistance payments are inversely related to female labor-force participation.

The final specification in Table 1 is similar to the one appearing in the third column of Table 1, but replaces the independent variable with the combined AFDC and Food Stamp variable. The results indicate a similar pattern to the results in the third column of Table 1. The combined independent variable is significant but has the incorrect sign, contradicting previous studies.

Summarizing these results based on border counties as our unit of observation, we see that much of that variation in the dependent variable is explained but the independent variables of interest are either insignificant, have the wrong sign, or both. These results cast doubt on the technique of simply using border-county observations to control for latent, unobserved geographic factors.

A second method of controlling for latent geographic factors that also relies on least-squares estimation was proposed by Holcombe and Lacombe (2003)[5]. They use a "matched" border-county technique that links observations geographically by transforming variables in each county to reflect a percentage of the same variable in the adjacent county or counties across the state border. For example, the first dependent variable is the percentage of households headed by females, with children under 18, and no husband. For this analysis, the percentage of female-headed households in county i is matched with the percentage of female-headed households in counties, $j = 1,\ldots,m$, which represent m counties in the adjacent state that touch county i. This produces the matched value for female-headed households, FHH_i, following the formula

$$FHH_i = \frac{county_i}{\frac{1}{m}\sum_{j=1}^{m} county_j}$$

where $county_i$ is the percentage of female-headed households on one side of the state border and $county_j$ is the percentage of female-headed households in each of the m counties of the bordering state that touch county i. Observations calculated in this way are referred to as matched values, and all of the dependent and independent variables in the regressions that follow are transformed in this way.

As an illustrative example, in Columbus County, North Carolina, female-headed households with children under 18 make up 7.26% of all households. In Horry County, South Carolina, which is adjacent to Columbus County and across the state border, female-headed households with children account for 6.34% of households. Substituting into the above formula, the matched value for female-headed house-

5. The discussion that follows borrows heavily from Holcombe and Lacombe (2003).

holds with children in Columbus County is 7.26/6.34 = 1.15. Horry County borders not only Columbus County, but also Brunswick County, North Carolina, where female-headed households with children are 5.33% of total households, so the matched value for Horry County is 6.34/((7.26+5.33)/2) = 1.01. The only South Carolina county Brunswick County borders is Horry, so its matched value for female-headed households with children is 5.33/6.34 = 0.84. These matched values are interpreted as the value of the county's observation as a percentage of those it matches, so, for example, female-headed households are 115% as common in Columbus County as the counties it borders across the state line, while female-headed households are only 84% as common in Brunswick County as its matched county across the state border. This example shows the difference between the matched values and the raw county level data used in the previous specifications contained in Table 1.

The independent variables are matched in a similar manner as are all county-level observations, with the exception of the state-level AFDC and Food Stamp variables. For example, Alabama in 1990 had a maximum AFDC benefit of $118, while Georgia's maximum benefit was $273 and Florida's was $294. Houston County, Alabama, is in the southeast corner of the state, and borders Early and Seminole Counties in Georgia, and Jackson County in Florida. Using the formula above, Houston County's matched AFDC variable for the Georgia counties is $118/$273 = 0.43, and is $118/294 = 0.40 when matched with Jackson County in Florida. The Georgia counties have a matched value of 2.31 with Houston County, Alabama, using the same formula, and the Jackson County, Florida, has a value of 2.49.

The following regression models use the "matched" dependent and independent variables and ordinary least-squares (OLS) regression techniques. Note that that data are not simply the raw values, but are matched according to the above formula in order to tie geographically each observation and hopefully control for unobserved geographic variation. Also note that some counties touch counties in more than one state. These observations were included twice to account for the specific geography of the particular counties which raises the sample size to 1319 "matched" border counties.

Columns two and three of Table 2 lists the results when the dependent variable is the percentage of female headed households with the set of independent variables from the previous unmatched regression models contained in Table 1.

Examining the results from the regression models contained in columns two and three of Table 2, we see that now the public assistance variables are both statistically significant and have positive coefficient estimates, consistent with previous research. However, these two models do have some shortcomings. For one, the variation of the dependent variable explained by the model is quite low, hovering around 30%. This may be due to the fact that by matching the dependent variable, some variation is removed. Another problem is that some independent variables are significant, but they have the wrong sign. For example, the average wage per job per week variable is significant but has a positive coefficient estimate, indicating that an increase in the wage increases the number of female-headed households.

The next regression models provide information on the relationship between public assistance payments and females labor force participation, again using the "matched" border-county sample.

The specifications in columns four and five of Table 2 indicate that AFDC and combined AFDC and Food Stamp payments increase the incidence of females age 16 and over who are not in the labor force, consistent with the findings of other researchers. However, only 42% of the variation in the dependent variable is explained by this model, and some of the independent variables either have the wrong sign, are insignificant, or both (e.g., manufacturing employment).

TABLE 2

Effect of Combined AFDC and Food Stamp Payments and the Effect of Maximum AFDC Grant on Female Headed Households and Female Labor Force Non-Participation Matched Sample (t-statistics in parentheses)

	Dependent Variables			
Independent Variable	Female Heads, No Husband, Children 0–17 as a % of Total Households	Female Heads, No Husband, Children 0–17 as a % of Total Households	Females Age 16+ Not in Labor Force as a % of Female Population	Females Age 16+ Not in Labor Force as a % of Female Population
Combined Maximum AFDC Grant and Food Stamp Payment 1990	0.115922 (4.82)***		0.113728 (5.15)***	
Maximum AFDC Grant 1990		0.113930 (4.79)***	0.111326 (5.09)***	
Average Tax Rate	−0.028164 (−1.06)	−0.029367 (−1.11)	−0.076524 (−3.14)***	−0.075372 (−3.09)***
Percentage of Population that are High School Graduates 1990	−0.137582 (−5.29)***	−0.136267 (−5.24)***	−0.344648 (−14.43)***	−0.345986 (−14.47)***
Manufacturing Employment % 1990	−0.051222 (−2.18)**	−0.050346 (−2.15)**	0.081325 (3.77)***	0.080488 (3.73)***
Average Wage per Job per Week	0.075294 (2.83)**	0.072749 (2.74)**	−0.055166 (−2.26)**	−0.052624 (−2.15)**
Median Age 1990	−0.347421 (−14.57)***	−0.347245 (−14.56)***	0.450340 (20.56)***	0.450187 (20.56)***
Population Density 1990	0.127950 (5.31)***	0.127469 (5.29)***	0.008350 (0.38)	0.008822 (0.40)
Local Per Capita Government Expenditures	0.056418 (2.02)**	0.060953 (2.19)**	0.017659 (0.69)	0.013112 (0.51)
Unemployment Rate % 1990	0.269329 (11.03)***	0.271816 (11.15)***	0.132200 (5.90)***	0.129725 (5.78)***
Urban Population % 1990	0.132115 (5.48)***	0.132214 (5.48)***	−0.018510 (−0.83)	−0.018580 (−0.84)
Number of Observations	1319	1319	1319	1319
Adjusted R^2	.3150	.3148	.4216	.4218

*** significant at the 1% level. ** significant at the 5% level. * significant at the 10% level

In summary, the "matched" border-county technique does a better job of explaining the relationship between public assistance payments and the dependent variables in a manner consistent with the many studies on this topic. The matching technique, by directly linking observations geographically, can be argued to provide a better method of controlling for geographic factors. On the downside, however, the variation in the dependent variables that these models explain is lower than in the non-matched sample, although a direct comparison is not possible because the dependent variables in the two techniques are different.

4. A SPATIAL AUTOREGRESSIVE MODEL FOR BORDERING MATCHING DATA SAMPLES

In this section, an extension of the traditional spatial autoregressive model described in Anselin (1988) that is ideally suited to the case of border matching spatial data samples is set forth. The model is shown in (1), where W_1 and W_2 represent n by n spatial weight matrices described in the sequel.

$$y = \rho_1 W_1 y + \rho_2 W_2 y + X\beta + \varepsilon \tag{1}$$

In (1), y represents the same dependent variable vector used in the border-matching models whose least-squares estimates were presented in Tables 1 through 4[6]. Identical explanatory variables are in the matrix X, and we assume that ε obeys the traditional least-squares assumptions, $\varepsilon \sim N(0,\sigma^2 I_n)$. Note that this model subsumes the border-matching models as a special case when the scalar parameters ρ_1 and ρ_2 take on values of zero. This provides a simple test of this model specification versus the traditional border-matching specification.

The n by n spatial weight matrix W_1 is constructed using the three nearest within-state neighbors to each county observation and W_2 is based on the three nearest neighbors in the bordering state. Specifically, for county i, the ith row of the weight matrices contain a value of 1/3 in the three columns associated with the three observations representing the nearest within-state neighboring counties. Similarly, the matrix W_2 contains values of 1/3 for the three nearest neighbors in an adjacent state.

Matrix multiplication of these weight matrices with the dependent-variable vector creates one explanatory variable vector consisting of the average of the dependent variable in surrounding counties within the state, (W_1,y), and another representing the average value of y observations from the three nearest counties from a bordering state (W_2,y). These two variables are spatial lags analogous to time-series lagged dependent variables. As in the case of time-series they capture the influence of excluded unobservable latent variables that exhibit spatial autocorrelation. If the scalar parameters ρ_1 and ρ_2 are significantly different from zero, one can show that least-squares estimates of β are biased and inconsistent (Kelejian and Prucha 1998).

Estimation of the model in (1) requires maximum-likelihood methods where the likelihood function is shown in (2), where $e = (I_n - \rho_1 W_1 - \rho_2 W_2)y - X\beta$.

$$L(\beta,\sigma,\rho_1,\rho_2;y,X) = (2\pi\sigma^2)^{-n/2}\,|\,I_n - \rho_1 W_1 - \rho_2 W_2\,|\exp\left(-\frac{1}{2\sigma^2}e'e\right) \tag{2}$$

6. The sample size is different for the spatial autoregressive model which includes county-level observations from bordering counties within the state as well as those from neighboring states, so the vector of dependent and the matrix of independent variables differ in terms of the number of observations.

In practice, the parameters β and σ can be concentrated out of the log-likelihood function, making this an optimization problem involving the two scalar parameters ρ_1 and ρ_2. The most computationally demanding part of solving the optimization problem involves computing the log-determinant expression $|I_n - \rho_1 W_1 - \rho_2 W_2|$ over a grid of values for ρ_1 and ρ_2. This is carried out prior to optimization so that a simple table look-up is needed during optimization calls to evaluate the log-likelihood function. This represents a simple extension of the methods proposed in Pace and Barry (1997). The MATLAB functions used to implement these estimates are available in the spatial econometrics toolbox.[7]

The traditional spatial autoregressive model utilizes a single spatial-weight matrix, taking the form $y = \rho Wy + X\beta + \varepsilon$. In this application, the interest is in separating out the two types of spatial dependence that may exist in the dependent variable vector. Given that the sample data is constructed using bordering counties, each observation will have neighboring counties within the state as well as neighbors in the adjacent state. This provides one motivation for this approach to modeling spatial dependence. A second point is that the magnitude and strength of spatial autocorrelation may be different between observations within the state and observations in adjacent states since cultural practices as well as the economic environment seem likely to differ amongst individual states. For this particular application, it may be the case that there are different spatial dependencies based on the premise that individual's react differently depending on both the county and state in which they live.

Table 3 presents estimation results for the spatial autoregressive model based on border-county observations in their raw form, similar to the models presented in Table 1. Estimates of the variance-covariance for the parameters β presented in the tables were constructed using a numerical Hessian calculated using the maximum-likelihood values of the parameters. In columns two and three of Table 3, the dependent variable is the number of female-headed households and the independent variables are the same as in the previous specifications. There are several interesting observations one can make about the spatial autoregressive model estimates. First, both the within and between-state parameter estimates ρ_1 and ρ_2 are positive and significant at the 99% level, indicating that this specification is superior to least squares. These nonzero parameters provide evidence of statistically significant spatial dependence between neighboring counties within the states as well as between counties in adjacent states. That is, the sample selection strategy of relying on border counties does not appear to eliminate spatial dependence arising from unobserved latent variables that take on a spatial character. A second point is that after controlling for these two sources of spatial dependence, the AFDC and the combined AFDC and Food Stamp payments variables are significant. Entitlement payments seem to have a positive effect on the number of female-headed households, after controlling for intrastate and interstate spatial autocorrelation, which is consistent with many of the studies on this topic. Finally, the models explain approximately 61% of the variation in the dependent variable, which is approximately 11% higher than the least-squares models that do not take into account any spatial variation in the dependent variable.

Columns four and five of Table 3 contain estimation results for the same independent variable specifications as in columns two and three of Table 3, with the female labor force nonparticipation rate used as the dependent variable. The spatial weight matrices utilized in these models are also the same as in the previous models constructed using the same within-state and between-state spatial connectivity relationships.

7. The spatial econometrics toolbox is a public domain set of functions that can be found at www.spatial-statistics.com.

TABLE 3

Effect of Combined AFDC and Food Stamp Payments and the Effect of Maximum AFDC Grant on Female Headed Households and Female Labor force Non-Participation Spatial Autoregressive Model (Asymptotic t-statistics in parentheses)

Independent Variable	Dependent Variables: Female Heads, No Husband, Children 0–17 as a % of Total Households	Female Heads, No Husband, Children 0–17 as a % of Total Households	Females Age 16+ Not in Labor Force as a % of Female Population	Females Age 16+ Not in Labor Force as a % of Female Population
Combined Maximum AFDC Grant and Food Stamp Payment 1990		0.044016 (1.84)*		−0.002516 (−0.13)
Maximum AFDC Grant 1990	0.044534 (1.89)*		−0.002089 (−0.11)	
Average Tax Rate	−0.043161 (−1.60)	−0.043921 (−1.63)	−0.030258 (−1.36)	−0.030247 (−1.36)
Percentage of Population that are High School Graduates 1990	−0.286680 (−9.94)***	−0.290332 (−9.95)***	−0.346595 (−14.58)***	−0.346299 (−14.42)***
Manufacturing Employment % 1990	−0.063455 (−2.83)***	−0.063115 (−2.81)***	−0.222957 (−12.04)***	−0.222966 (−12.05)***
Average Wage per Job per Week	−0.034018 (−1.60)	−0.033577 (−1.58)	0.091262 (5.20)***	0.091257 (5.20)***
Median Age 1990	−0.280485 (−13.35)***	−0.284006 (−13.46)***	0.408156 (23.55)***	0.408243 (23.50)***
Population Density 1990	0.012585 (0.62)	0.012411 (0.61)	−0.008034 (−0.48)	−0.007990 (−0.48)
Local Per Capita Government Expenditures	0.088106 (3.20)***	0.089239 (3.25)***	−0.055280 (−2.44)***	−0.055244 (−2.44)***
Unemployment Rate % 1990	0.202934 (9.18)***	0.204193 (9.22)***	0.173061 (9.49)***	0.173112 (9.49)***
Urban Population % 1990	0.247034 (10.32)***	0.248151 (10.34)***	−0.076524 (−3.88)***	−0.076580 (−3.88)***
ρ_1 (Intra State)	0.214999 (6.96)***	0.207999 (9.29)***	0.256999 (12.07)***	0.256999 (12.07)***
ρ_2 (Inter State)	0.197999 (3.80)***	0.193000 (8.71)***	0.176000 (5.86)***	0.176000 (5.86)***
Number of Observations	1129	1129	1129	1129
Adjusted R^2	.6039	.6083	.7343	.7343

*** significant at the 1% level. ** significant at the 5% level. * significant at the 10% level

The results in columns four and five of Table 3 are very similar to those in columns two and three Table 3, with the exception that both the AFDC and combined AFDC and Food Stamp payment variables are now insignificant at conventional significance levels. Also of note is that both the within-state (ρ_1) and between-state (ρ_2) parameters on the spatial lags are positive and significant, indicating the presence of spatial autocorrelation for both within-state and between-state observations of this dependent variable. Again the spatial autoregressive model specification proves superior to ordinary least squares. Finally, the models explain approximately 73% of the variation in the dependent variable, which is 10% higher than the least-squares model that does not take into account spatial dependence.

5. CONCLUSIONS

Economists have long been interested in the study of applied policy questions and most disagreement among researchers regards controlling for possible latent unobserved influences. In cross-sectional samples of states, counties, zip-code areas, census tracts, and other geographic delineations, it seems likely that some of the latent unobservable influences may vary systematically with the spatial location of the observations. A comparison of least squares and spatial autoregressive model estimates for a sample of county-level observations selected from state borders was undertaken here. The proposed spatial autoregressive model nests the least-squares specification as a special case. When estimates of the parameters measuring two types of spatial dependence, (1) that between counties within states and (2) between counties in adjacent states are insignificant, the spatial autoregressive model collapses to the least-squares specification. This makes specification testing simple. On the other hand, when these parameters are significantly different from zero, least-squares estimates are biased and inconsistent, similar to the situation that arises with simultaneity. This paper examined one particular policy question, that of whether AFDC and Food Stamp payments have an impact on the number of female-headed households and female labor-force participation. For this illustrative case, strong spatial dependence was found for both within-state and between-state county-level observations, implying that least-squares estimates are biased and inconsistent.

These theoretical observations are consistent with the fact that in the illustrative example the least-squares estimates produced results that were inconsistent with most literature on the policy effects of these programs. A second "matched" border-county approach set forth by Holcombe and Lacombe (2003) that relies on least-squares estimation produced results that were more consistent with previous empirical research in this area. However, the statistical fits of the models are quite low, and the signs of parameter estimates for some of the independent variables are not consistent with theoretical expectations. Formal tests for spatial autocorrelation in the least-squares residuals from these two methodologies (not reported here) indicated the presence of significant spatial dependence in the residuals from the Holmes (1998) specification. The residuals from the Holcombe and Lacombe (2003) least-squares estimates exhibited weaker spatial dependence, but still significant in most of the four models examined here. These results are consistent with the finding of strong spatial dependence in the dependent variable when using the spatial autoregressive model on all four specifications.

Economists and other social scientists interested in controlling for the effects of unobserved latent variables that vary systematically over space should find the spatial autoregressive model set forth here to be a parsimonious and reasonably simple alternative specification that effectively models this type of influence on the dependent variable.

LITERATURE CITED

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

Bieker, Richard F. (1998) "Work and Welfare: An Analysis of AFDC Participation Rates in Delaware." *Social Science Quarterly* vol. 62, 1, 169–76.

Bronars, Stephen G. and John R. Lott, Jr. "Criminal Deterrence, Geographic Spillovers, and the Right to Carry Concealed Handguns." *American Economic Review* 88, 2, 475–79.

Card, David and Alan B. Krueger (1994) "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania." *American Economic Review* 84, 4, 772–93.

Chapman, Richard, Kevin Duncan, and Jerry Gray (1998) "Mixing Welfare and Work: Evidence from the PSID, 1980-87." *Eastern Economic Journal* 24, 1, 51–62.

Cloutier, Norman R., and Anthony L. Loviscek (1989) "AFDC Benefits and Inter-Urban Variation in Poverty Among Female-Headed Households." *Southern Economic Journal* 56, 2, 315–22.

Danziger, Sheldon, Robert Haveman, and Robert Plotnick (1981) "How Income Transfer Programs Affect Work, Savings, and the Income Distribution: A Critical Review." *Journal of Economic Literature* 19, 3, 975–1028.

Fox, William F. (1986) "Tax Structure and the Location of Economic Activity Along State Borders" *National Tax Journal* 39, 4, 397–401.

Garfinkel, Irwin, and Larry L. Orr (1974) "Welfare Policy and the Employment Rate of AFDC Mothers." *National Tax Journal* 27, 2, 275–84.

Holcombe, Randall G. and Donald J. Lacombe (Forthcoming) "Using Matched Border Counties for Policy Analysis: The Effect of Entitlement Programs on Female-Headed Households and Female Labor Force Participation" *Eastern Economic Journal*.

Holmes, Thomas J. (1998) "The Effect of State Policies on the Location of Manufacturing: Evidence from State Borders." *Journal of Political Economy* 106, 4, 667–705.

Isserman, Andrew, and Terance Rephann (1995) "The Economic Effects of the Appalachian Regional Commission: An Empirical Assessment of 26 Years of Regional Development Planning." *Journal of the American Planning Association* 61, 3, 345–64.

Kelejian, H., and I.R. Prucha (1998) "A Generalized Spatial Two-Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances." *Journal of Real Estate and Finance Economics* 17, 1, 99–121.

Levine, Phillip B., and David J. Zimmerman (1999) "An Empirical analysis of the Welfare Magnet Debate using the NLSY." *Journal of Population Economics* 12, 3, 391–409.

Levy, Frank (1979) "The Labor Supply of Female Household Heads, or AFDC Work Incentives Don't Work Too Well." *The Journal of Human Resources* 14, 1, 76–97.

Moffitt, Robert (1986) "Work Incentives in the AFDC System: An Analysis of the 1981 Reforms." *American Economic Review* 76, 2, 219–23.

Moffitt, Robert (1992) "Incentive Effects of the U.S. Welfare System: A Review." *Journal of Economic Literature* 30, 1, 1–61.

Nord, Stephen, and Robert G. Sheets (1990) "The Relationships of AFDC Payments and Employment Structure on the Labour Force Participation and Underemployment Rates of Single Mothers." *Applied Economics* 22, 2, 187–99.

Pace, R., Kelley, and Ronald Barry (1997) "Quick Computation of Spatial Autoregressive Estimators." *Geographical Analysis* 29, 232–46.

Robins, Philip K. (1990a) "Explaining Recent Declines in AFDC Participation." *Public Finance Quarterly* 18, 2, 236–55.

Robins, Philip K. (1990b) "A Decade of Declining Welfare Participation: Sorting Out the Causes." *Contemporary Policy Issues* 8, 1, 110–23.

Smith, Patricia K. (1991) "An Empirical Investigation of Interstate AFDC Benefit Competition." *Public Choice* 68, 1-3, 217–33.

Walters, Stephen J.K. (1990) "Business Climate and Measured Poverty: The Evidence Across States." *Atlantic Economic Journal* 18, 1, 20–26.

Yelowitz, Aaron S. (1995) "The Medicaid Notch, Labor Supply, and Welfare Participation: Evidence from Eligibility Expansions." *The Quarterly Journal of Economics* 110, 4, 909–939.