ECONOMIC RESEARCH
FEDERAL RESERVE BANK OF ST. LOUIS

WORKING PAPER SERIES

Spatial Probit and the Geographic Patterns of State Lotteries

Authors	Cletus C. Coughlin, Thomas A. Garrett, and Rubén Hernández-Murillo
Working Paper Number	2003-042B
Revision Date	December 2003
Citable Link	https://doi.org/10.20955/wp.2003.042
Suggested Citation	Coughlin, C.C., Garrett, T.A., Hernandez-Murillo, R., 2003; Spatial Probit and the Geographic Patterns of State Lotteries, Federal Reserve Bank of St. Louis Working Paper 2003-042. URL https://doi.org/10.20955/wp.2003.042

Federal Reserve Bank of St. Louis, Research Division, P.O. Box 442, St. Louis, MO 63166

Spatial Probit and the Geographic Patterns of State Lotteries *

Cletus C. Coughlin †
Thomas A. Garrett ‡
and
Rubén Hernández-Murillo §

Research Division
Federal Reserve Bank of St. Louis
411 Locust St.
St. Louis, MO 63102

*Molly Castelazo and Deborah Roisman provided research assistance. The views expressed in this paper do not reflect official positions of the Federal Reserve Bank of St. Louis nor the Federal Reserve System.
†E-mail:coughlin@stls.frb.org, Phone: (314) 444-8585.
‡E-mail:garrett@stls.frb.org, Phone: (314) 444-8601.
§E-mail:hernandez@stls.frb.org,Phone: (314) 444-8588.

Spatial Probit and the Geographic Patterns of State Lotteries

Abstract

We implement a spatial probit model to differentiate states with a lottery from those without a lottery. Our analysis extends the basic spatial probit model by allowing spatial dependence to vary across geographic regions. We also separate the spatial effects of neighbors versus non-neighbors. The methodology provides consistent and efficient coefficient estimation in light of the simultaneity in spatial dependence. We find evidence of spatial dependence and spatial heterogeneity in lottery usage, and we find that spatial patterns differ significantly by geographic region. The importance of spatial dependence in state lottery usage suggests the need to consider spatial effects in empirical models examining the use of any policy tool by subnational governmental units.

Keywords: spatial probit, state lotteries.
JEL Classification: H71, H72, H73, C21, C25, R12.

1 Introduction

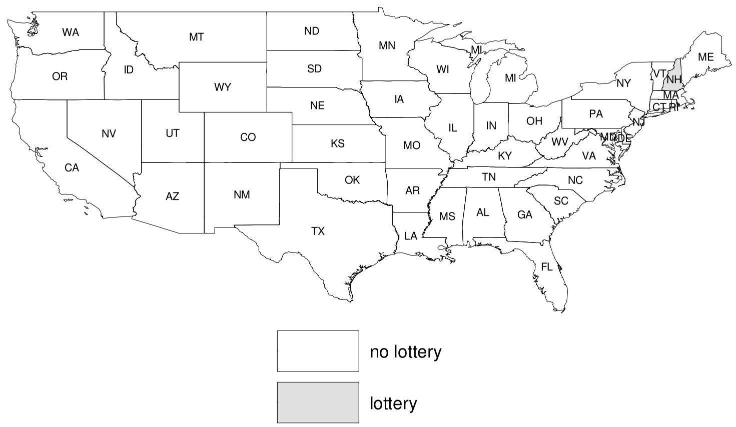
Spatial econometric techniques have been developed to effectively capture spatial processes in natural or experimental data (Anselin, 1988, 1995; Haining, 1990). In this paper we introduce an extension of the spatial probit model and apply this model to the issue of state lottery usage. Our spatial probit model is a less restrictive form of traditional spatial models that assume the degree of correlation is the same for all cross-sectional units. We allow spatial correlation to vary by geographic region and demonstrate that reducing restrictions on spatial correlation coefficients can result in significant gains in model power. In addition, we allow spatial effects to differ between neighboring and non-neighboring states. This analysis suggests significant differences and reveals that estimates from spatial models can be sensitive to the spatial structure imposed.

The spatial probit model not only allows direct estimation of spatial dependence in states' decisions to have a lottery, it also provides for consistent and efficient estimation not afforded by earlier studies of lottery usage. Our methodology provides consistent and efficient coefficient estimation in light of the simultaneity in spatial dependence. Previous studies on lottery usage attempted to capture the influence of neighboring states' lottery status by including a variable that reflects whether a neighboring state has a lottery or a variable reflecting the number of neighboring states having a lottery. However, the very nature of spatial correlation suggests these variables are endogenous. The results of previous studies may be inconsistent because they fail to address the simultaneity problem.

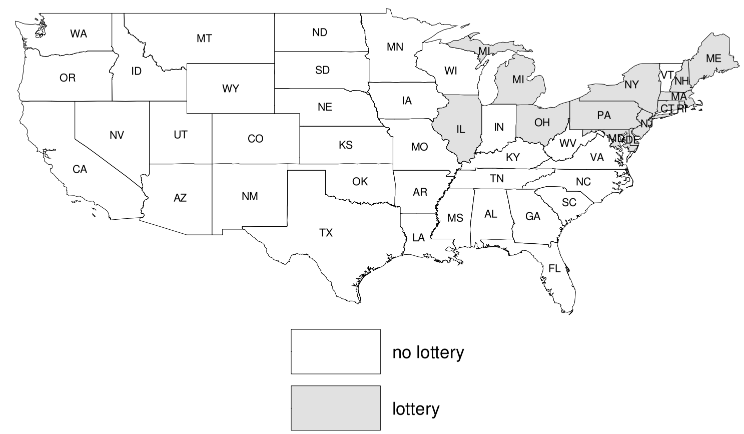
Spatial econometric models have been applied to numerous economic issues. Among the issues examined are: 1) the decision by farmers to adopt new technologies by Case (1992); 2) the location of foreign direct investment in China by Coughlin and Segev (2000); 3) cross-border shopping by Garrett and Marsh (2002); 4) the adoption of environmental treaties by European countries by Murdoch et al. (2003); and 5) state fiscal decisions by Case et al. (1993), Brueckner and Saavedra (2001), and Hernández-Murillo (2003). The majority of these studies have applied spatial econometric techniques to models with continuous dependent variables. Spatial models with discrete dependent variables have received little application in the literature. Case (1992), Marsh et al. (2000), and Murdoch et al. (2003) are noteworthy exceptions.

Previous research examining state fiscal decisions has shown spatial dependence in state and local government policies such as tax rates and budget and expenditure levels. Figure 1 provides visual evidence suggesting that state lottery decisions are also interdependent. New Hampshire adopted the first modern-day state lottery in 1964 (panel a). Between 1964 and 1976, states in New England, the Mid-Atlantic, and the Great Lakes approved lotteries (panel b). In the subsequent 12 years, lotteries spread throughout states in the Midwest and on the Pacific Coast and began appearing in states in the Plains and Rocky Mountains. By 2000, the only groups of states without lotteries were in the South and the Rocky Mountains (panel d). While state lottery usage is certainly a function of a state's economic, political, and demographic characteristics, our analysis allows a determination of whether the pattern of state lottery usage is also due to a state's proximity to other states having a lottery.

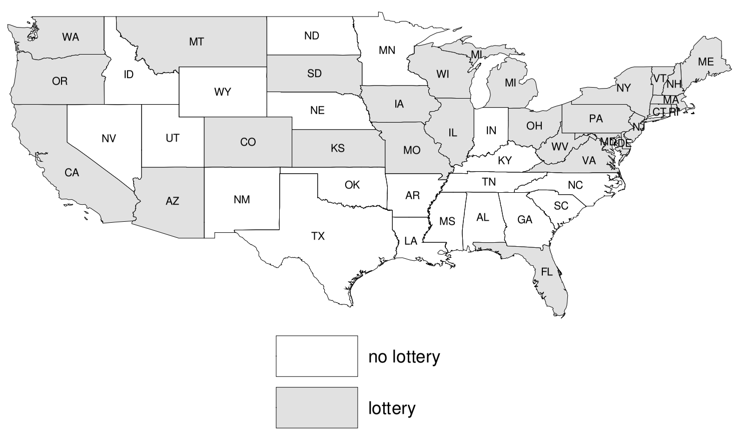
The conceptual frameworks that have been used to explain state lottery usage rely on the rational behavior by legislators. What differentiates the frameworks is the objective function



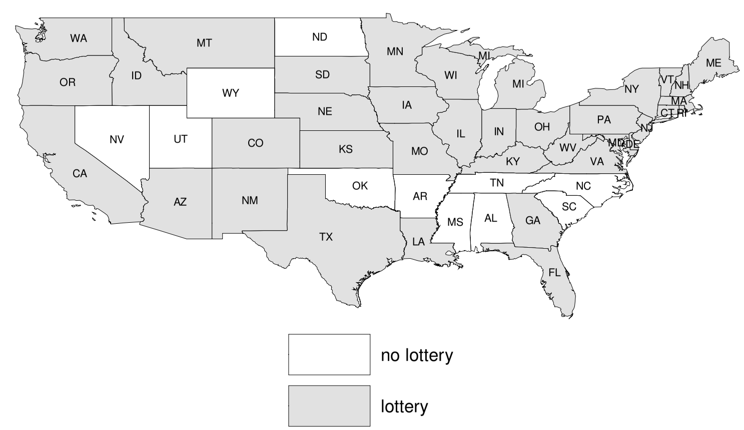
(a) State lottery usage as of 1964



(b) State lottery usage as of 1976



(c) State lottery usage as of 1988



(d) State lottery usage as of 2000

Figure 1: Patterns of State Lottery usage

of the legislators. The most frequently used framework is the legislator-support maximization approach (Filer at al., 1988). Legislators recognize that increased state spending can increase their political support by increasing the welfare of their constituents. Martin and Yandle's (1990) duopoly transfer mechanism approach views the state as a rent-seeker and a redistributive agent. Lotteries compete with both legal and illegal gambling operations, and thus provide a mechanism that allows the state to generate some portion of the revenues that they miss by their inability to tax illegal gambling. A third framework presented by Erekson et al. (1999) views the representative legislator as maximizing utility subject to a constraint.

The empirical implementation of these frameworks has proceeded in two ways, each shedding light on different issues. Filer et al. (1988), Martin and Yandle (1990), and Davis et al. (1992) address the question of whether a state has a lottery as of a specific year. Filer et al. (1988) and Davis et al. (1992) estimate cross section binary choice models using probit models, while Martin and Yandle (1990) present results using ordinary least squares. Alternatively, the estimation of duration models provides evidence on which variables increase or decrease the expected time until a state adopts a lottery. Berry and Berry (1990), Alm et al. (1993), Caudill et al. (1995), Mixon et al. (1997), and Erekson et al. (1999) estimate hazard functions in their lottery adoption studies. Here the legislator receives utility from improving the state's fiscal wellbeing, but the legislator is constrained by his reelection desires that hinge on the satisfaction of his constituents.

We explore the questions of what differentiates the states with a lottery from those without a lottery, and to what extent the existence of lotteries in other states affects the lottery choice in a specific state. Our empirical approach, which is explained in detail later, differs from the previously-used approaches.

2 Data

We use data on the 48 contiguous states and the District of Columbia over the period 1964 to 2000. Table 1 presents summary statistics of the variables we use in our analysis.

Alm et al. (1993) hypothesize that a state is more likely to adopt a lottery when fiscal and economic conditions in the state are relatively weak. To account for these factors, we include in our model specifications the level of real per capita personal income, the percent change in real state and local tax revenues, and the levels of real per capita transfers from the federal government to state and local governments.[1]

A state's decision to have a lottery may also be influenced by the amount of debt held by state and local governments. As debt increases, a state lottery may become a more attractive alternative for raising revenues. To capture this scenario, we include real per capita short and long-term state and local debt in our empirical models.[2] To date, the evidence concerning the relationship between debt and lottery usage is mixed. Martin and Yandle (1990) and Alm et al. (1993) have found statistically significant evidence that higher levels of per capita debt are associated positively with lottery usage. On the other hand, Caudill et al. (1995) and Mixon et al. (1997) did not find a significant relationship.

Political and demographic factors are also considered to be important in a state's decision to have a lottery (Alm et al., 1993; Clotfelter and Cook, 1989). We include a dummy variable that takes the value of 1 if the state has a referendum or initiative process, and 0 otherwise.[3] The feasibility of a lottery is also correlated with the potential number of players in a state and the ability of the state to run a more cost-effective operation (DeBoer, 1985). Thus, we also include population density in our empirical models.

Two groups likely to have strong views about gambling are the elderly and those belong-

[1]Per capita income is from the Bureau of Economic Analysis; state and local tax revenue and intergovernmental transfer data were obtained from the U.S. Census' *Governmental Finances*, various years. All nominal variables were deflated using the GDP deflator (1996=100).

[2]Debt data were obtained from the U.S. Census' *Compendium of State Government Finances.* Because 1987 data were not available, total state debt was multiplied by the average short-term percent of total and the average long-term percent of total to arrive at long-term and short-term state debt in 1987.

[3]Obtained from Alm et al. (1993) and the Initiative and Referendum Institute.

ing to religious denominations (Clotfelter and Cook, 1989). We include the percentage of the state's population that is older than age 65 and the percent of the population that is Catholic.[4] Those older than 65 can be characterized as typical lottery players. Moreover, Alm et al. (1993) noted that this group is frequently opposed to other forms of tax increases. In their study, however, this variable was not statistically significant.

On the other hand, the views of religious groups toward gambling have been shown to matter empirically. Martin and Yandle (1990), Berry and Berry (1990), Caudill et al. (1995), Mixon et al. (1997), and Erekson et al. (1999) suggest fundamentalist religious groups (e.g., Baptists) are more likely to oppose a lottery than Catholics. The authors find a negative relationship between religious groups generally thought to oppose gambling and the existence of a state lottery. Meanwhile, Alm et al. (1993) found a significant positive relationship between the percentage of a state's population that was Catholic and the adoption of a lottery.

Our empirical model allows us to control for the impact of whether a neighbor has a lottery. As described later, we incorporate the impact of whether a neighboring state has a lottery on a specific state via spatial econometric techniques. To date, four measures have been used: 1) a dummy variable distinguishing whether an adjacent state has a lottery (Filer et al. 1988; Alm et al. 1993); 2) the number of adjacent states with a lottery (Berry and Berry 1990); 3) the percentage of adjacent states with a lottery (Erekson et al. 1999); and 4) the percentage of a state's border shared with states having a lottery (Davis et al. 1992; Caudill et al. 1995; Mixon et al. 1997). Excluding Filer et al. (1988), the preceding studies found a positive relationship between their proxy and whether a state had a lottery.

[4]Percent of the population exceeding age 65 is from the U.S. Census' *Statistical Abstract of the United States*, various years. The percent of the population that is Catholic is from *Religious Congregations and Membership in US-2000* CD-ROM, published by Glenmary Research Center. Religious affiliation data is only collected in decennial census years. Religious affiliation data is unavailable for Alaska and Hawaii until 1971; therefore 1971 data is used for those states for the period from 1964 to 1970.

Table 1: Summary Statistics

Variable	Mean	Min	Max	St. Dev.
Lottery Usage Status (yes 1; no 0)	0.389	0	1	0.394
Population Density (inhabitants per sq. mi.)	376.6	3.3	13,082.0	1,531.4
Real Personal Income (dollars per capita)	18,169.7	6,750.7	38,595.0	2,760.2
Real Tax Revenues (3-year moving average growth rate, %)	4.2	-14.0	32.0	2.5
Real Short Term Debt (dollars per capita)	92.9	0.0	1,271.9	118.2
Real Long Term Debt (dollars per capita)	2,996.8	132.9	9,073.2	1,191.4
Real Federal Grants (dollars per capita)	0.66	0.08	4.54	0.33
Referendum (yes 1; no 0)	0.520	0	1	0.504
Population Age 65 or Older (%)	11.3	1.0	18.6	1.7
Population Catholic (%)	18.8	0.7	63.5	13.6

The summary statistics use yearly cross-section data for 1964-2000.
Tax revenues, short and long term debt, and federal grants refer to state and local governments.
All nominal variables were deflated using the GDP deflator (1996=100).

3 Empirical Model: The Spatial Probit

The basic model of spatial dependence innovated by Cliff and Ord (1981) and Anselin (1988) allows for spatial dependence in the dependent variable or in the error component. The model with spatial dependence in the dependent variable is often referred to as a *spatial lag model* or as a *spatial autoregressive model*. The model with spatial dependence in the error term is referred to as a *spatial error lag model*, or as a *spatial autocorrelation model*. Somewhat different than autocorrelation in time series models, spatial dependence in cross-sectional data is multi-dimensional in that it depends upon all contiguous or influential units of observations.

The framework we use is similar to the standard spatial econometric model. The discrete nature of our dependent variable and the panel structure of the data necessitate some modifications of the standard model. Maximum likelihood estimation traditionally produces consistent estimators for spatial models with continuous dependent variables. However, unless corrected for, spatial dependence in models with binary dependent variables introduces heteroskedasticity that renders coefficient estimates inefficient.

The lottery status for a state $i = 1, ..., N$ at time period $t = 1, ..., T$ is derived, as in the usual binary choice model, from a latent variable, y^*_{it}, and the rule $y_{it} = \mathbf{1}\,[y^*_{it} \geq 0]$. The first-order spatial lag model for the latent variable is given in vector form, stacking cross-sectional observations over all time periods, by:

$$(1) \quad y^* = \rho \mathbf{W} y^* + X\beta + \varepsilon,$$

where X is a $(TN \times K)$ matrix of exogenous variables and ε is a $(TN \times 1)$ zero-mean error term. The scalar ρ is the spatial lag coefficient and reflects positive spatial correlation in the dependent variable if $\rho > 0$, negative spatial correlation if $\rho < 0$, and no spatial correlation if $\rho = 0$. [5] $\mathbf{W}$ is a $(TN \times TN)$ block diagonal matrix having T copies of an $(N \times N)$ spatial weights matrix W along the diagonal. The individual elements of $W = \{\omega_{ij}\}$ are specified later.

Spatial dependence can also occur in the error term, ε. Spatially correlated errors may occur due to spatial correlation among the independent variables, spatial heterogeneity in functional form, omitted variables, and spatial correlation in the dependent variable when a spatially lagged dependent variable is not included in the model (Anselin, 1988; chapter 8). The first-order spatial error lag model is given as:

$$(2) \quad \varepsilon = \lambda \mathbf{W} \varepsilon + \upsilon,$$

where υ is a $(TN \times 1)$ vector of i.i.d. random variables with zero mean and covariance matrix $\sigma_\upsilon^2 I$.[6] $\mathbf{W}$ is the $(TN \times TN)$ matrix described earlier, and λ is a scalar that measures spatial error correlation. The errors are positively correlated if $\lambda > 0$, negatively correlated if $\lambda < 0$, and not correlated if $\lambda = 0$.

Many alternative weighting schemes for W have been used in the literature. One of the most common is the binary joins matrix (Cliff and Ord, 1981; Anselin, 1988; Case, 1992)

[5] Unlike the standard first-order autoregressive model in time series, spatial correlation coefficients do not necessarily have to lie between −1 and 1 in the first-order spatial model. Generally, coefficient values are between the inverse of the largest and smallest eigenvalues of the weights matrix. See Anselin (1995).

[6] Our model makes the assumption that the off-diagonal elements of the covariance matrix are zero. Relaxing this assumption, while potentially increasing efficiency, greatly complicates the estimation procedure. Researchers have recently explored several alternative methods for estimating the spatial probit models that use information in the off-diagonal elements (see Anselin (2002) and Fleming (forthcoming)).

where $\omega_{ij} = 1$ if observations i and j ($i \neq j$) share a common border, and $\omega_{ij} = 0$ otherwise. In this specification, the elements of matrix W are row-standardized by dividing each ω_{ij} by the sum of each row i.

A limitation of the binary joins matrix is that it assumes equal weights across all bordering spatial neighbors and does not allow the effective capture of spatial distances across all cross-sectional units. Thus, we also considered various measures of spatial distance (d) that have been discussed in the literature (Bodson and Peeters (1975), Dubin (1988), Garrett and Marsh (2002), Hernández-Murillo (2003)). Measures of spatial contiguity include the inverse distance between states, where $\omega_{ij} = 1/d_{ij}$, the inverse distance squared, where $\omega_{ij} = 1/d_{ij}^2$ and exponential distance decay, where $\omega_{ij} = exp(-d_{ij})$. As the distance between states i and j increases (decreases), ω_{ij} decreases (increases), thus giving less (more) spatial weight to the state pair when $i \neq j$. In all cases, $\omega_{ij} = 0$ for $i = j$. While there is no consensus on how distance between cross-sectional units should be measured, we follow Hernández-Murillo (2003) and consider the distance between state population centers.[7]

We use the inverse distance spatial weights matrix and the binary joins matrix in our empirical models. We chose to use both forms of weights matrices to highlight any differences in spatial patterns of lottery usage in neighboring states only (captured by the binary joins matrix) and between all states (captured by the inverse distance matrix).

Our basic specification assumes that the influence of spatial dependence is the same for all states. To reveal differences in spatial correlation for geographic regions, we also specify distinct spatial correlation coefficients for states in each of the eight BEA regions. Allowing for regional spatial correlation coefficients gives the following structure:

$$(3) \qquad y^* = \sum_{k=1}^{R} \rho_k \mathbf{W}_k y^* + X\beta + \varepsilon,$$

[7]The distance was computed using the geographic coordinates for the population centroids computed by the Bureau of the Census for the year 2000. Population centroids did not differ significantly in earlier decades.

where, potentially,

$$(4) \qquad \varepsilon = \sum_{k=1}^{R} \lambda_k \mathbf{W}_k \varepsilon + \upsilon.$$

Here R denotes the total number of regions, and ρ_k and λ_k denote the spatial lag and spatial error lag coefficients, respectively, for region k. $\mathbf{W}_k$ remains an $(TN \times TN)$ block diagonal matrix having T copies of an $(N \times N)$ spatial weights matrix W_k along the diagonal.[8] Now, however, the elements of each matrix W_k are constructed to capture spatial correlation between each state in region k and the remaining 47 states and the District of Columbia. Thus, for each state i in region k, row i of W_k contains our measure of distance between state i and all remaining 47 states and the District of Columbia. If state i is not in region k, then row i of W_k contains all zeros. In essence, each matrix W_k is constructed by multiplying each row of the matrix W_k by a dummy variable that has a value of 1 if state i is in region k, and 0 otherwise.

Rewriting the full spatial autoregressive model by incorporating the structure in (3) and (4) gives

$$(5) \qquad y^* = (I - \sum_{k=1}^{R} \rho_k \mathbf{W}_k)^{-1} X\beta + (I - \sum_{k=1}^{R} \rho_k \mathbf{W}_k)^{-1} (I - \sum_{k=1}^{R} \lambda_k \mathbf{W}_k)^{-1} \upsilon.$$

This structure induces heteroskedasticity. Correcting for heteroskedasticity is done using the method of Case (1992) and Marsh et al. (2000). Renaming the last term in equation (5)

$$(6) \qquad u = (I - \sum_{k=1}^{R} \rho_k \mathbf{W}_k)^{-1} (I - \sum_{k=1}^{R} \lambda_k \mathbf{W}_k)^{-1} \upsilon,$$

[8]Notice, however, that $\sum_{k=1}^{R} \mathbf{W}_k = \mathbf{W}$, and therefore, the original model with constant spatial lag coefficient is a restricted version of the model with regional spatial lags.

we obtain its covariance matrix as follows

$$(7) \quad E[uu'] = \sigma_\upsilon^2[(I - \sum_{k=1}^{R} \rho_k \mathbf{W}_k)'(I - \sum_{k=1}^{R} \lambda_k \mathbf{W}_k)'(I - \sum_{k=1}^{R} \lambda_k \mathbf{W}_k)(I - \sum_{k=1}^{R} \rho_k \mathbf{W}_k)]^{-1},$$

where σ_υ^2 is the common variance of the υ_{it}'s. As in the standard probit specification, for identification purposes, we assume that $\sigma_\upsilon^2 = 1$.

The full spatial model in (5) is then premultiplied by the variance-normalizing transformation $Z = (\mathrm{diag}(E[uu']))^{-1/2}$, a $(TN \times TN)$ diagonal matrix. The transformed model is:

$$(8) \quad Zy^* = Z(I - \sum_{k=1}^{R} \rho_k \mathbf{W}_k)^{-1} X\beta + Z(I - \sum_{k=1}^{R} \rho_k \mathbf{W}_k)^{-1}(I - \sum_{k=1}^{R} \lambda_k \mathbf{W}_k)^{-1}\upsilon.$$

Clearly the event $y_{it}^* > 0$ is equivalent to the event $Zy_{it}^* > 0$; therefore $y_{it} = \mathbf{1}\ [Zy_{it}^* \geq 0]$ and the log-likelihood can be stated in terms of the transformed model as follows:

$$(9) \quad \ln L = \sum_{i=1}^{N} \sum_{t=1}^{T} \{y_{it} \ln \Phi[X_{it}^* \beta] + (1 - y_{it}) \ln(1 - \Phi[X_{it}^* \beta])\},$$

where $X^* = Z(I - \rho\mathbf{W})^{-1}X$ and Φ is the c.d.f. of a standard normal distribution. Setting $\rho = 0$ or $\lambda = 0$ allows estimation of the spatial lag or spatial error lag model, respectively, and setting $\rho = \lambda = 0$ gives the log-likelihood for the standard probit model.

4 Estimation Results and Discussion

Tables 2 and 3 present the probit estimation results for alternative specifications of the determinants of a state's lottery usage.

Table 2: Standard Pooled Probit and Probit with Binary Spatial Weights

	Pooled Probit		Binary Spatial Weights	
	No Spatial Effects [1]	NLP in RHS [2]	Spatial Lag [3]	Spatial Error [4]
Population Density (log)	*** 0.0311	*** 0.0499	*** 0.0945	*** 0.0414
	[4.7367]	[7.0646]	[9.9670]	[5.6539]
Per Capita Real Personal Income	*** 4.9305	*** 3.5128	*** 3.1514	*** 5.3120
	[21.6798]	[11.1929]	[8.7396]	[18.8166]
Real Tax Revenues (growth rate)	***-1.0235	***-0.8761	***-1.0879	***-1.1274
	[4.0927]	[3.4869]	[4.4849]	[3.8947]
Per Capita Real Short Term Debt	-0.0100	0.0162	*-0.1368	-0.0138
	[0.1617]	[0.2634]	[1.7108]	[0.1896]
Per Capita Real Long Term Debt	0.7338	** 1.4607	** 2.3401	* 1.3369
	[1.0247]	[2.1023]	[2.4317]	[1.7823]
Per Capita Real Federal Grants	***-0.7988	***-1.2132	***-1.1785	***-0.8927
	[3.4161]	[5.2924]	[3.4769]	[3.4823]
Referendum	** 0.0330	** 0.0390	*** 0.2382	** 0.0428
	[2.0073]	[2.4172]	[6.9647]	[2.3931]
Population Age 65 or Older (share)	*** 2.8696	*** 2.6585	*** 2.0160	*** 3.2785
	[7.2885]	[6.8300]	[3.2464]	[6.8139]
Population Catholic (share)	*** 0.8731	*** 0.8752	*** 0.8709	*** 0.8245
	[13.6201]	[13.9817]	[9.8439]	[11.3221]
Share of Neighbors with Lottery (NLP)		*** 2.2364		
		[6.7913]		
RHO			*** 0.7451	
			[25.1612]	
LAMBDA				***-1.1678
				[-24.7462]
σ^2	0.104	0.099	0.126	0.111
Log-Likelihood	-570.712	-547.165	-443.78	-553.721
LR stat.	1298.754	1345.849	1552.617	1332.736
MF's R^2	0.532	0.552	0.636	0.546
No. Obs.	1813	1813	1813	1813

Top panel: mean marginal effects; middle panel: probit coefficients; z-statistics in square brackets.
Asterisks *, **, and ***, indicate statistical significance at the 10%, 5%, and 1% levels, respectively.
LR is the likelihood ratio test of the joint significance of all coefficients other than the constant term, $2[l_u - l_r]$, where l_u is the log-likelihood of the unrestricted model, and l_r is the log-likelihood of the restricted model.
McFadden's pseudo R^2 is defined by $R^2 = 1 - l_u/l_r$.

4.1 Pooled Probit and Binary Spatial Weights

Column [1] in table 2 corresponds to a standard probit model where no spatial effects among neighboring states are taken into account. The coefficients represent the mean marginal effects of the explanatory variables on the probability of lottery usage.[9] Column [2] corresponds to a specification where the spatial interaction among states is accounted for by including, for each state, the percentage of neighboring states that have already enacted a lottery. This is analogous to the specifications used in previous hazard function approaches to the adoption of state lotteries (Alm et al., 1993). As noted previously, this specification fails to account for the endogeneity of the lottery usage decision among states. This model suggests that the probability of usage depends positively on the percentage of neighboring states that have already adopted a lottery. It is important, therefore, to model the interaction among neighboring states appropriately since the decision to adopt and maintain the policy is an endogenous variable.

Columns [3] and [4] present the results from two alternative specifications in which we account for spatial effects using binary spatial weights to determine contiguity among states. As we discussed before, this scheme considers as neighbors only those states that are adjacent to each other. Column [3] corresponds to the model with a spatial lag in the dependent variable, and column [4] corresponds to the model with a spatial lag in the error term. As we can see from the table, the spatial dependence coefficients are statistically significant. In particular, the spatial lag coefficient, ρ, indicates the presence of strategic interaction among states in the choice of lottery usage.

Specifications [1] through [4] suggest that the level of real per capita state and local long term debt has a positive impact on the lottery usage probability. The same is true for real

[9]For continuous variables, the coefficients in tables 2 and 3 represent the average (taken over all the values of the explanatory variables in the sample) of the derivative of the predicted probability with respect to the variable in question. For the dummy referendum variable, the coefficients represent the average absolute change in the predicted probability when this variable takes a value of 1 relative to the case where it takes a value of 0. The standard errors for the mean marginal effects where computed from the covariance matrix of the probit coefficients using the Delta method.

per capita personal income and for population density. Additionally, the growth rate of real state and local tax revenues and real per capita levels of federal grants to state and local governments have a statistically significant negative impact. These results are consistent with the theory that a state's level of fiscal distress is an important determinant in the usage of new revenue instruments. Real per capita levels of short term debt do not have a statistically significant impact on the lottery usage probability.

Previous studies have found that political variables, such as whether the state has a referendum mechanism, have a significant effect on the lottery usage probability. We find that the effect is positive and statistically significant.

The share of elderly population (65 years or older) and the share of Catholics in the population also have a positive impact on the lottery usage probability. These results are consistent with studies documenting the characteristics of lottery players.

In the models corresponding to the binary spatial weights, both the spatial lag in the dependent variable and the spatial lag in the error term are statistically significant from zero at the 1 percent level. The log-likelihood of model [3], with only a spatial lag in the dependent variable is, however, substantially higher.[10]

4.2 Inverse-Distance Decay Spatial Weights

An alternative definition of neighborhood effects in the lottery usage decision allows decisions in nearby states that are not necessarily adjacent to affect a specific state. In this case, the use of inverse-distance spatial weights is more appropriate to identify spatial interactions among the states. Columns [1] through [5] in table 3 present the estimation results using spatial weights computed as the inverse distance between states' population centroids.

The results are qualitatively similar to those in table 2. The results from a joint spatial model, in column [3], suggest spatial interaction in both the dependent latent variable and in the error term. The coefficients for the spatial lag in columns [1] and [3] are positive,

[10]The model with both a spatial lag in the dependent variable and in the error term could not be estimated using binary weights because of identification problems.

Table 3: Probit with Inverse-Distance Decay Spatial Weights

	Inverse-Distance Decay Spatial Weights				
	Spatial Lag	Spatial Error	Spatial joint	Regional Spatial lags	
	[1]	[2]	[3]	[4]	[5]
Population Density (log)	*** 0.0677	*** 0.0319	*** 0.0659	*** 0.1015	*** 0.1098
	[8.1108]	[4.6908]	[7.9932]	[15.0023]	[12.5342]
Per Capita Real Personal Income	*** 2.5722	*** 4.9142	*** 2.7876	*** 1.0635	* 0.3791
	[7.4206]	[20.8090]	[8.1658]	[4.2104]	[1.7883]
Real Tax Revenues (growth rate)	***-0.8355	***-1.0022	***-0.8701	***-0.3774	**-0.2695
	[4.0598]	[3.9614]	[3.9745]	[2.9331]	[2.5065]
Per Capita Real Short Term Debt	*** 0.2546	-0.0031	*** 0.3384	0.0263	0.0591
	[3.6127]	[0.0468]	[3.6075]	[0.5590]	[1.2132]
Per Capita Real Long Term Debt	* 1.5817	0.7468	* 1.5746	*** 2.1879	*** 2.6637
	[1.8652]	[1.0327]	[1.9043]	[3.0205]	[4.1296]
Per Capita Real Federal Grants	***-1.0844	***-0.7229	***-0.9332	***-1.3641	***-0.4992
	[3.8699]	[2.8678]	[3.1141]	[6.5739]	[2.6940]
Referendum	*** 0.1403	** 0.0347	*** 0.1472	*** 0.0957	*** 0.0872
	[5.7042]	[2.0907]	[5.8546]	[4.1035]	[3.7230]
Population Age 65 or Older (share)	** 1.2902	*** 2.9006	*** 1.5383	*** 1.5594	** 0.8107
	[2.4609]	[6.7848]	[3.0703]	[3.6922]	[2.0706]
Population Catholic (share)	*** 1.3333	*** 0.8808	*** 1.3262	*** 1.2350	*** 0.8270
	[14.0798]	[13.4928]	[12.7001]	[12.0459]	[7.4606]
Share of Neighbors with Lottery (NLP)					
RHO	*** 7.6691		*** 7.4620		
	[18.0835]		[16.1675]		
LAMBDA		** 5.3260	*** 5.0678		*** 5.0059
		[2.4853]	[4.5738]		[3.0486]
RHO1 (New England)				***16.1965	***14.2287
				[15.1686]	[5.7785]
RHO2 (Mideast)				*** 4.3018	*** 9.0582
				[2.9529]	[3.2495]
RHO3 (Great Lakes)				*** 4.2418	*** 3.3517
				[3.4298]	[2.9223]
RHO4 (Plains)				*** 7.4210	*** 8.0902
				[4.4004]	[3.9550]
RHO5 (Southeast)				*** 3.7385	*** 2.4991
				[4.5737]	[3.6667]
RHO6 (Southwest)				*** 5.7908	*** 6.3611
				[3.6374]	[4.2545]
RHO7 (Rocky Mountain)				***18.0134	***16.5425
				[5.1627]	[4.3866]
RHO8 (Far West)				***15.0908	***14.0484
				[4.0690]	[3.6732]
σ^2	0.132	0.104	0.131	0.176	0.207
Log-Likelihood	-497.783	-570.57	-492.69	-434.788	-430.393
LR stat.	1444.612	1299.039	1454.799	1570.601	1579.391
MF's R^2	0.592	0.532	0.596	0.644	0.647
No. Obs.	1813	1813	1813	1813	1813

Top panel: mean marginal effects; middle panel: probit coefficients; z-statistics in square brackets.
Asterisks *, **, and ***, indicate statistical significance at the 10%, 5%, and 1% levels, respectively.
LR is the likelihood ratio test of the joint significance of all coefficients other than the constant term, $2[l_u - l_r]$, where l_u is the log-likelihood of the unrestricted model, and l_r is the log-likelihood of the restricted model.
McFadden's pseudo R^2 is defined by $R^2 = 1 - l_u/l_r$.

supporting the conclusion from model [3] in table 2 that there is a positive interaction in the choice of lottery usage among states. In a strategic framework, this suggests that the policy choices of states' usage of a lottery behave as strategic complements, in the sense that a state will find it to its advantage to adopt and maintain the policy if its rivals adopt it as well.[11]

The numerical interpretation of the estimated coefficients is as follows.[12] The findings in column [5] of table 3, for example, suggest that a $1,000 increase in real per capita personal income yields, on average, an increase of about 0.38 percentage points in the probability of lottery usage, since real per capita personal income is measured in $100,000s. The same scale is used for the real per capita long term state and local debt. On the other hand, real per capita short term debt is measured in $1,000s, but the coefficient is not significant. An increase of 1 percentage point in the annual growth rate of real tax revenues induces a decrease in the probability of lottery usage of about 0.27 percentage points, as this variable is measured in decimal points (that is, 0.01 equals 1 percent); the same scale is used for the share of elderly population and the share of Catholic population. Finally, real per capita transfers from the federal government are measured in tenths of dollars, therefore, the coefficient indicates that a $1 increase in per capita transfers decreases the probability of lottery usage by about 5 percentage points. The interpretation for the dummy variable is straightforward, as the coefficient indicates that states with a referendum process have a probability of usage that is about 8 percentage points larger, on average, than the probability of usage among states without a referendum.

We have included two models in which we isolate the regional correlation coefficients as we described in section 3. Each coefficient, ρ_k, measures the average correlation between a state in region k and the spatially weighted lottery status of all other states. Differences in

[11]Notice also that although the sign on the spatial error lag is negative in column [4] of table 2 and positive in column [2] of table 3, the models with only a spatial lag in the error term have considerably smaller likelihood than models with only the spatial lag in the dependent variable or with both a spatial lag in the dependent variable and in the error term.

[12]The explanatory variables were transformed to a similar scale for the estimation. This is often necessary to facilitate the convergence of the optimization algorithm in maximum likelihood estimation.

the regional lag coefficients, $\{\rho_k\}$, are attributable to two factors. First, the average distance from a state to all other states varies across regions—that is, the spatial weights matrices W_k differ for different k's. Second, the number of states having a lottery increases over time. Each regional coefficient, ρ_k, captures both factors.

We find considerable evidence that the effects of spatial correlation in the dependent variable vary by region. As reported in columns [4] and [5] of table 3, all the regional correlation coefficients are positive and statistically significant; furthermore, visual inspection of the estimated coefficients suggests differences in the magnitude of spatial correlation between a state in a given region and all other states. A likelihood ratio test for the null hypothesis that all the regional coefficients, $\{\rho_k\}$, are identical (comparing models [4] and [1]), yields a statistic of 125.99. The corresponding *p-value* for a χ^2 distribution with 8 degrees of freedom is essentially zero, so we can safely reject the null hypothesis. Comparing models [5] and [3] with the same null hypothesis, we obtain a statistic of 124.59, so we also reject that the regional spatial lags are identical.

Model [5], where we allow also for spatial dependence in the error term, provides the largest log-likelihood. Comparing this model with model [4], where we assume that there is no spatial dependence in the error term, we obtain a likelihood ratio statistic of 8.79. The corresponding *p-value* for a χ^2 distribution with 1 degree of freedom is 0.003. This implies that considerable explanatory power is gained by relaxing the assumption that spatial effects in lottery usage are identical across regions. Furthermore, the probit specification of lottery usage seems to exhibit spatial dependence in the error term as well.

5 Conclusion

Using spatial econometric techniques, we show that the spatial patterns associated with state lottery usage result from a combination of the characteristics of individual states and the decisions of their neighbors. Five characteristics—population density, real per capita personal income, real per capita state and local debt, the share of a state's population that is elderly, and the share of a state's population that is Catholic—are related positively to whether a state has a lottery.

The results from spatial probit models strongly indicate that a state's decision to have a lottery is dependent on the decisions made by other states. Using either a binary or inverse distance weights matrix in the estimation of spatial effects, we find that proximity to states that have a lottery will increase the probability that a given state will itself have a lottery. In addition, we provide strong evidence that spatial dependence in state policies can vary by region. Unlike previous research, this study is the first to directly model and provide estimates on the spatial interdependence of states' decision to have a lottery, and the approach taken here affords the estimation of consistent and efficient coefficients. The importance of spatial effects on the choice of state lottery usage suggests that such effects should be considered in examining the determinants of other state government policies.

References

[1] Alm, James; McKee, Michael and Skidmore, Mark. "Fiscal Pressure, Tax Competition, and the Introduction of State Lotteries." *National Tax Journal*, December 1993, 46(4), pp. 463-476.

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

[3] Anselin, Luc. *SpaceStat, A Software Program for the Analysis of Spatial Data, Version 1.80.* Regional Research Institute, West Virginia University, Morgantown, West Virginia, 1995.

[4] Anselin, Luc. "Under the Hood: Issues in the Specification and Interpretation of Spatial Regression Models." *Agricultural Economics*, November 2002, 27(3), pp. 247-267.

[5] Berry, Frances S. and Berry, William D. "State Lottery Adoptions as Policy Innovations: An Event History Analysis," *American Political Science Review*, June 1990, 84(2), pp. 395-415.

[6] Bodson, P. and Peeters, D. "Estimation of the Coefficients of a Linear Regression in the Presence of Spatial Autocorrelation: An Application to a Belgian Labor Demand Function." *Environment and Planning A*, 1975, 7, pp. 455-472.

[7] Brueckner, Jan K. and Saavedra, Luz A. "Do Local Governments Engage in Strategic Property-Tax Competition?" *National Tax Journal*, June 2001, 54(2), pp. 203-229.

[8] Case, Anne C.; Rosen, Harvey S. and Hines, James R., Jr. "Budget Spillovers and Fiscal Policy Interdependence: Evidence from the States." *Journal of Public Economics*, October 1993, 52(3), pp. 285-307.

[9] Case, Anne. "Neighborhood Influence and Technological Change." *Regional Science and Urban Economics*, September 1992, 22(3), pp. 491-508.

[10] Caudill, Steven B.; Ford, Jon M.; Mixon, Franklin G., Jr. and Peng, Ter Chao. "A Discrete-Time Hazard Model of Lottery Adoption." *Applied Economics*, June 1995, 27(6), pp. 555-561.

[11] Cliff, A. and Ord, J. *Spatial Processes, Models, and Application.* London: Pion, 1981.

[12] Clotfelter, Charles T. and Cook, Philip J. *Selling Hope: State Lotteries in America.* Cambridge: Harvard University Press, 1989.

[13] Coughlin, Cletus C. and Segev, Eran. "Foreign Direct Investment in China: A Spatial Econometric Study." *The World Economy*, January 2000, 23(1), pp. 1-23.

[14] Davis, J. Ronnie; Filer, John E. and Moak, Donald L. "The Lottery as an Alternative Source of State Revenue." *Atlantic Economic Journal*, June 1992, 20(2), pp. 1-10.

[15] DeBoer, Larry. "Administrative Costs of State Lotteries." *National Tax Journal*, December 1985, 38(4), pp. 479-487.

[16] Dubin, Robin A. "Estimation of Regression Coefficients in the Presence of Spatially Autocorrelated Error Terms." *Review of Economics and Statistics*, August 1988, 70(3), pp. 466-474.

[17] Erekson, O. Homer; Platt, Glenn; Whistler, Christopher and Ziegert, Andrea L. "Factors Influencing the Adoption of State Lotteries," *Applied Economics*, July 1999, 31(7), pp. 875-884.

[18] Filer, John E.; Moak, Donald L. and Uze, Barry. "Why Some States Adopt Lotteries and Others Don't," *Public Finance Quarterly*, July 1988, 16(3), pp. 259-283.

[19] Fleming, Mark M. "Techniques for Estimating Spatially Dependent Discrete Choice Models." in L. Anselin, R. Florax, and S. Rey (eds.), *Advances in Spatial Econometrics*. Heidelberg: Springer-Verlag, forthcoming.

[20] Garrett, Thomas A. and Marsh, Thomas L. "The Revenue Impacts of Cross-Border Lottery Shopping in the Presence of Spatial Autocorrelation." *Regional Science and Urban Economics*, July 2002, 32(4), pp. 501-519.

[21] Haining, Robert. *Spatial Data Analysis in the Social and Environmental Sciences*. Cambridge: Cambridge University Press, 1990.

[22] Hernández-Murillo, Rubén. "Strategic Interaction in Tax Policies Among States," Federal Reserve Bank of St. Louis *Review*, May/June 2003, 85(3), pp. 47-56.

[23] Marsh, Thomas L.; Mittelhammer, Ron C. and Huffaker, Ray G. "Probit with Spatial Correlation by Field Plot: Potato Leafroll Virus Net Necrosis in Potatoes." *Journal of Agricultural, Biological, and Environmental Statistics*, 2000, 5(1), pp. 22-38.

[24] Martin, Robert and Yandle, Bruce. "State Lotteries as Duoploy Transfer Mechanisms." *Public Choice*, March 1990, 64(3), pp. 253-264.

[25] Mixon, Franklin G., Jr.; Caudill, Steven B.; Ford, Jon M. and Peng, Ter Chao. "The Rise (or Fall) of Lottery Adoption within the Logic of Collective Action: Some Empirical Evidence." *Journal of Economics and Finance*, Spring 1997, 21(1), pp. 43-49.

[26] Murdoch, James C.; Sandler, Todd and Vijverberg, Wim P. M. "The Participation Decision versus the Level of Participation in an Environmental Treaty: A Spatial Probit Analysis" *Journal of Public Economics*, February 2003, 87(2), pp. 337-362.