Policy Innovation in Local Jurisdictions: Testing for Neighborhood Influence in School Choice Policies

Johannes Rincke*

June 8, 2006

Abstract

Before making difficult decisions, individuals tend to collect information on decision makers in reference groups. With respect to policy innovations in a decentralized public sector, this may give rise to positive neighborhood influence on adoption decisions. In this paper, U.S. school district data are used to show that decision makers indeed are heavily affected by decision makers in reference groups. The policy innovation under consideration is inter-district public school choice. The results suggest that if a given district's neighbors' expected benefits from adopting school choice policies increase, this substantially increases the original district's probability of adoption. The paper thus supports the view that the diffusion of policy innovations is stimulated by horizontal interaction among local governments.

Keywords: policy innovation, school choice, spatial interaction, neighborhood influence

JEL Classification: D78, H77

---

* Department of Economics, University of Munich.
I thank Thiess Büttner and Robert Schwager for their advice and support and an anonymous referee for helpful comments on an earlier draft. I am also indebted to Kerry J. Gruber at the NCES for technical advice and Andreas Schaich for valuable research assistance.

Correspondence:

Johannes Rincke
Department of Economics
Seminar for Economic Policy
University of Munich
Akademiestr. 1/II
D-80799 Munich
phone: +49/89/21806753
email: johannes.rincke@lrz.uni-muenchen.de

1 Introduction

Recent research has put forward the idea that when making decisions, individuals are affected by the attitude or actual behavior of other individuals in certain reference groups. Collecting information on benchmark agents may be rational for a number of reasons. First of all, since it is costly to evaluate alternatives and to find out which one is to be pursued, individuals may seek to benefit from information gathered by others. A straightforward example is what Hirshleifer and Teoh (2003) in their taxonomy of social learning and behavioral convergence call rational observational learning: individuals learn by rational Bayesian inference on information conveyed in the behavior of others. Scharfstein and Stein (1990), for instance, discuss herd behavior at financial markets. They show that it may be in the best interest of managers concerned about their reputation to ignore their private information and to mimic investment decisions of other managers. Katz and Shapiro (1986) argue that with respect to technology adoption, network externalities may motivate decision makers to choose similar actions. Brock and Durlauf (2001) provide a model of discrete choice with social interactions, where individual utility directly depends on the choices of others in a reference group. In the context of political decisions, Besley and Case (1995) argue that it may be rational for office-motivated governments to choose policies similar to those in benchmark jurisdictions if voters use relative rather than absolute performance for their inference on the quality of locally provided services.

A particularly interesting and practically relevant application for theories of behavioral convergence is the horizontal diffusion of policy innovations in federal systems. It has often been claimed that a decentralized political system with a large number of independent local jurisdictions offers favorable conditions for policy experiments and the implementation of policy innovations. Oates (1999), for instance, gives an optimistic view of 'laboratory federalism', where many jurisdictions simultaneously engage in policy experiments and where jurisdictions learn from the experience made by others. A less optimistic perspective is taken by Strumpf (2002). The paper points to the fact that innovative political activities often exhibit learning externalities, creating a standard sort of incentive for free-riding on other jurisdictions' experimentation efforts. In a related contribution, Kotsogiannis and Schwager (2004) deal with career concerns of political actors in federal systems. They show how the incentive to signal above-average ability to the electorate can motivate politicians at the local level to implement new policies with uncertain outcomes. The role of yardstick competition in the diffusion of policy innovations is discussed in Rincke (2005). The paper shows how comparative performance evaluation by voters can force office-motivated local governments to implement political best-practice technologies.

Using data on a large sample of U.S. school districts, this paper provides evidence on behavioral convergence among local jurisdictions with respect to policy experimentation and policy innovation. Specifically, I investigate the school districts' decisions on the adoption of inter-district public school choice. A simple measure for geographical proximity is used to describe reference districts. More specifically, for any given district, a composite reference district is defined from the set of all local school districts belonging to the same county. Using an approach proposed by Case (1992), a spatial probit is estimated in the cross-section of districts. The model allows a district's predisposition towards a new policy to depend on the predisposition of the composite reference district. Controlling for a large number of district characteristics describing local preferences and for spatial correlation in errors, the results indicate that the school districts' predispositions towards policies of open enrollment are strongly interdependent. Districts which are exposed to a composite neighbor with a strong predisposition towards adoption are significantly more likely to participate. Thus, similarly situated school districts indeed tend to affect each other in the decision whether to experiment with new political technologies. The results support the view that policy diffusion is stimulated by horizontal interaction among local governments.

The remainder of the paper proceeds as follows. In the next section, the estimation approach is discussed. Section 3 describes inter-district public school choice as a policy innovation and discusses potential factors affecting political preferences of school districts. Section 4 presents the data and estimation results, and section 5 concludes.

2 Estimation approach

Given the mere number of almost 15,000 school districts in the U.S., it seems reasonable to assume that decision makers at the district level tend to perceive the situation in nearby districts as particularly informative with regard to the prospects of new policies. Thus, the estimation approach of Hautsch and Klotz (2003), where neighbors are defined in an abstract social space, does not seem to be appropriate for the current analysis. Instead, the analysis will rely on a spatial probit specification introduced by Case (1992). The model has been developed for the cross-sectional analysis of spatial dependence in discrete choice decisions. The approach is particularly well suited in the context of this study because it can be applied in a situation where, due to sampling, the information on the spatial distribution of adoption decisions is incomplete. Using a straightforward definition of composite neighbors based on a simple measure of spatial proximity, the model allows for the identification of spatial interactions in policy choices of the districts' local governments.

The model is a slightly modified version of the original approach proposed by Case (1992). Methodologically, it also draws on a large body of literature dealing with strategic interaction in tax setting, public goods provision, and expenditures.[1] In the following, the model is briefly discussed.

Consider a standard latent variable model where a latent variable y_i^*, describing the predisposition of some district $i = 1, \ldots, N$ to experiment with a new policy, determines actual adoptions y_i,

$$y_i = 1[y_i^* > 0]. \tag{1}$$

A structural spatial auto-regressive model for the predisposition towards adoption is specified as

$$y_i^* = \phi\, y_{-i}^* + x_i \beta + u_i, \tag{2}$$

where y_{-i}^* is the predisposition of a composite neighbor of i, x_i is a vector of control variables, and u_i is an i.i.d. error with zero mean. Suppose that each district i belongs to some county $m(i)$ and that $n_{m(i)}$ is the number of districts in that county. The composite neighbor's predisposition is then conveniently defined as the average predisposition of districts belonging to the same county as district i,

$$y_{-i}^* = \frac{1}{n_{m(i)} - 1} \sum_{j \neq i}^{N} d_{ij}\, y_j^*, \tag{3}$$

where d_{ij} is an indicator taking value 1 if j belongs to $m(i)$ and zero otherwise. In this model, a positive ϕ would mean that the districts' predispositions towards adopting the new policy are positively interdependent. Thus, if a given district's local government had a positive attitude towards adoption, this would positively affect the attitude in neighboring districts, and, hence, increase the probability of adoption among neighbors.

[1]See Besley and Case (1995), Brett and Pinkse (1997), Hayashi and Boadway (2000), Büttner (2001), Brueckner and Saavedra (2001), and Bordignon, Cerniglia, and Revelli (2003) for the choice of tax rates. Bivand and Szymanski (1997) deal with public goods provision, and Case, Hines, and Rosen (1993) and Baicker (2005) with expenditures. Brueckner (2003) gives a survey.

Of course, y^*_{-i} is endogenous in eq. 2. Stacking the equations for all N districts and solving for the N-dimensional column vector of latent variables gives the reduced form. In matrix notation, it is

$$y = (I - \phi W)^{-1} (x\beta + u), \tag{4}$$

where $W = \{w_{ij}\}_{i,j=1}^{N}$ is a square matrix of weights with $w_{ij} = d_{ij}/(n_{m(i)} - 1)$ if $i \neq j$ and $w_{ij} = 0$ if $i = j$. As Case (1992) has shown, from eq. 4 the individual equations can be rewritten as

$$y^*_i = \varrho_{m(i)} x_i \beta + \vartheta_{m(i)} \bar{x}_{m(i)} \beta + \varrho_{m(i)} u_i + \vartheta_{m(i)} \bar{u}_{m(i)}, \tag{5}$$

where $\bar{x}_{m(i)}$ is the vector of mean characteristics for districts in $m(i)$, $\bar{u}_{m(i)}$ is the mean of errors in $m(i)$, $\varrho_{m(i)} = (n_{m(i)} - 1)/(\phi + n_{m(i)} - 1)$, and $\vartheta_{m(i)} = n_{m(i)}\phi/[(1-\phi)(\phi + n_{m(i)} - 1)]$.

As in any model aiming at the identification of spatial dependence, a particular problem is to separate spatial correlation in the variables of interest from spatial error correlation. The reason is that spatial correlation in adoption decisions could be driven by spatially correlated shocks. Not accounting for spatial error dependence could then lead to false conclusions. Therefore, the model is extended to incorporate spatially correlated shocks

$$u_i = \rho u_{-i} + \epsilon_i, \tag{6}$$

where now ϵ_i is an i.i.d. error and u_{-i} is the average error across neighbors in $m(i)$. In matrix notation, the composite error of the reduced form is

$$v = (I - \phi W)^{-1}(I - \rho W)^{-1}\epsilon. \tag{7}$$

Denote the variance covariance matrix of the system as $\Omega = E(vv')$ and note that v is heteroskedastic by definition. Therefore, parameter estimates obtained from eq. 5 would be inconsistent. To restore homoskedasticity, the model is pre-multiplied by the inverse of the square root of diag(Ω) before estimation. Using standard maximum likelihood techniques for limited dependent variable models, the parameter of interest, ϕ, can then be estimated from the transformed system together with β and ρ.

The reduced form eq. 5 reveals that the identification of the parameters relies on information on district characteristics being available for all districts. Therefore, variation in exogenous variables provided by sampling procedures will not suffice. Note, however, that it is not necessary to use all N equations to estimate the model. Thus, given that information on district characteristics is complete, nothing precludes estimation of the model based on a sufficiently large sample of districts. I exploit this situation by combining data on district characteristics readily available for all school districts with data on district policies available only for a large sample of districts.

Before the data and the estimation of the model is described, it is useful to briefly discuss inter-district open enrollment as a local policy innovation in the U.S.

3 Inter-district open enrollment as a local policy innovation

As mentioned above, the policy innovation under investigation in this paper is inter-district public school choice, sometimes also called inter-district open enrollment. Basically, it allows students to attend a public school in a school district other than the district of residence. The data used in the analysis are a sample of school districts in Arkansas, California, Idaho, Massachusetts and Ohio. All five selected states share the common feature that they established inter-district choice programs between 1989 and 1993 and that districts were given discretionary power to decide whether they would admit nonresident students at local schools.[2] Thus, with regard to the selected states, open enrollment can be considered a policy innovation implemented at the school district level.

In the U.S., school choice policies have been a much discussed topic of educational reform in recent years. The significance of the inter-district version of school choice comes from the fact that it will tend to increase competition for students between districts. Of course, school boards as local authorities in individual districts will rather be interested in the effects of school choice on the number and the composition of students in local schools than in potential overall effects of increased competition on school productivity.[3] A valid model of the diffusion of open enrollment policies among districts must therefore take into account the crucial factors affecting the districts' willingness to participate in statewide inter-district choice programs.

In 1993, 40.5% of all local school districts sampled in the Schools and

[2] Choice programs started in Arkansas and Ohio in school year 1989/90, in Massachusetts and Idaho in 1991/92, and in California in 1993.

[3] Hoxby (2000, 2003) provides evidence suggesting that competition among public schools in the U.S. has substantial effects on educational achievement and school productivity.

Staffing Survey in the selected states reported to admit non-resident students at local schools. In all five states, fiscal incentives for participation were set by rewarding receiving districts by additional funds. Participating districts could thus hope to raise additional revenues by attracting transfer students.

A reason for hesitation in switching to a policy of open enrollment may be limited capacity in local schools. In general, districts with crowded schools will be less willing to allow for the enrollment of transfer students. Furthermore, crowded schools are perceived as less attractive by potential transfer students and, from an ex-ante perspective, decrease the probability that the district will be successful in attracting non-resident students.

Another factor influencing participation of districts in public school choice may be the districts' location relative to large central cities. Traditionally, suburban school districts have been opposing the idea of inter-district open enrollment (Ryan and Heise 2002). Given their social and economic characteristics, suburban schools are, on average, better than urban schools, and residents in suburban districts tend to perceive inter-district transfers as a threat to the superior quality of local public schools. Furthermore, apart from a districts location relative to central urban areas, the racial composition of local public schools alone as well as the income of an average resident household may affect districts' predisposition towards open enrollment. Finally, the analysis shall account for the fact that in the sample there are three different types of districts with regard to the grades served. In Arkansas, Idaho and Ohio, all districts are unified school districts and (at least potentially) serve all grades. In California and Massachusetts, unified school districts serve students in elementary as well as secondary schools, while elementary and high school districts are more specialized.

Based on the preceding discussion, I include as control variables in the empirical specification the student-teacher ratio as a variable measuring the capacity for enrollment of transfer students; the district's revenue per student as a measure for fiscal stress; the share of minority students in local public schools[4]; the median household income; and four dummy variables, one for districts in large or mid-size central cities, one for suburban school districts, one for elementary school districts and one for high school districts.

[4]The share of minority students is defined as one minus the share of white non-Hispanic students.

4 Data, estimation and results

4.1 The data

The information on school districts' open enrollment policies is from the Schools and Staffing Survey (SASS) 1993-94, providing data on a large sample of local school districts.[5] The survey asked districts whether they had 'a choice program in which students can enroll in another school or district outside their attendance area without justification based on individual special needs'. Districts which affirmed were then asked whether the program allowed for enrollment of students from other districts. In the empirical analysis, the answer to this last question is used to determine which districts did participate in inter-district open enrollment in the 1993-94 school year.

As mentioned above, I use the Schools and Staffing Survey data on school districts in Arkansas, California, Idaho, Massachusetts and Ohio. For the five selected states, the SASS provides information on open enrollment policies in 649 local school districts. For the empirical analysis, the sample was reduced to 511 districts.[6] Table 1 displays the number of school districts and observations by state. In the first column, the total number of local school districts in 1993/94 is provided. The second column shows the number of districts sampled in the Schools and Staffing Survey, and the last column displays the number of districts used in estimations. Data on the control variables are from the Public Education Finance Data of the Bureau of the Census (revenues and district type indicator), the School District Demographic System of the National Center for Education Statistics (NCES) (median household income[7]) and the Common Core of Data of the NCES (enrollment, number of minority students, number of teachers, and urbanicity indicator). Apart from a small

[5] To access the data, refer to National Center for Education Statistics (1998). For technical information, see National Center for Education Statistics (1996).

[6] 17 districts had to be excluded from the sample since they represent a whole county, i.e. they have no neighbors. Another 14 districts had missing values for explanatory variables. In a next step, districts with less than 800 students were removed from the sample. The reason for doing so is the presumption that the political behavior of a very small district will resemble that of an average school more closely than that of a larger district. The threshold of 800 students was determined by increasing the minimum number of students by increments of 100 (starting from zero) until each of the remaining districts had at least two schools. Finally, in order to identify influential observations, a linear probability model was estimated using the remaining 539 observations. Based on the approach proposed by Krasker, Kuh, and Welsch (1983), 28 observations were removed. This left 511 school districts for the analysis.

[7] Data on median household income is from 1999.

Table 1: Number of school districts and observations by state

	# school districts	Nob SASS	Nob sample
Arkansas	315	119	84
California	1000	192	151
Idaho	113	74	40
Massachusetts	259	115	101
Ohio	611	149	135
Sum	2298	649	511

number of districts with missing values for some control variables[8], the data described are available for all school districts. Table 2 provides descriptive statistics for the sample used in estimation.

4.2 Estimation and results

Before estimation results are presented and discussed it should be stressed that the districts covered in the Schools and Staffing Survey are selected on the basis of a complex survey design. Thus, the analysis is not based on a random sample. The effect of the survey design on the composition of the sample is accounted for by including the inverse of the sampling probability as a weight for the contribution of each district in the likelihood function.[9] Since the sample design may also induce unknown correlation in errors, robust standard errors for parameter estimates are computed using a Huber-White formula for probit models.[10]

The first step in the empirical analysis is a simple baseline regression where I completely ignore the potential impact the predispositions of neighbors may have on the attitude towards open enrollment in any given district. The baseline regression is meant as a first, albeit crude test whether the approach of estimating a discrete choice model for the adoption of open enrollment policies with the given set of control variables is meaningful at all. Table 3 reports the results of a weighted maximum likelihood estimation of a standard probit

[8]See footnote on page 8.

[9]The variation of sampling probabilities across districts is also reflected in the means displayed in Table 2. For instance, as mentioned above, the unweighted share of participating districts is 0.41, while the weighted share is 0.46.

[10]See Wooldridge (2002, p. 496) for details.

Table 2: Descriptive statistics (Nob=511)

	Mean[a]	Std. Dev.[a]	Min	Max
Indicator for school choice	0.461	0.499	0.00	1.00
Number of neighbors[b]	17.6	18.4	0.00	80.0
Central city	0.104	0.305	0.00	1.00
Suburb	0.303	0.460	0.00	1.00
Elementary school district	0.084	0.279	0.00	1.00
High school district	0.030	0.172	0.00	1.00
Student-teacher ratio	19.6	3.52	11.9	29.5
% minority students	0.205	0.250	0.00	0.960
Revenues per student[c]	4.98	1.16	2.97	11.5
Median household income[c]	39.5	13.1	17.1	98.2

[a] Weighted by inverse of sampling probabilities; [b] Number of districts belonging to the same county; [c] In thousands of dollars.

framework. The latent variable model is specified as

$$y_i^* = x_i\beta + u_i, \tag{8}$$

where x_i includes a constant, the control variables discussed above, and a series of state dummies. The state dummies account for all kinds of state-specific influences on the predisposition towards open enrollment, such as differences in school choice laws, state-specific fiscal incentives for districts promoting participation in open enrollment programs or the length of time the program was in place at the time of data collection. In Table 3, the first column displays the parameter estimates for the specification in eq. (8), the second column gives the robust standard errors, and the third column provides the average partial effects, i.e. the sample averages of estimated changes in the probability of adoption associated with a change in the explanatory variable. A quick inspection of the results shows that districts with crowded schools, as we presumed, seem to be less willing to open up local schools for non-resident students. Furthermore, districts with a higher share of minority students are less likely to participate. In addition, districts with lower revenues per student are more inclined towards open enrollment than high revenue districts.

The significance of a number of district characteristics together with the fact that the model correctly predicts almost two thirds of all adoption decisions suggests that all explanatory variables together provide a strong signal for the predisposition of school districts to participate in open enrollment pro-

Table 3: Baseline Probit

Explanatory variables	Estimates	S.E.	Slope[a]
Central city	0.052	0.214	0.019
Suburb	-0.270	0.194	-0.098
Elementary school district	0.337	0.297	0.122
High school district	-0.515	0.407	-0.187
Student-teacher ratio	-0.094**	0.047	-0.034
% minority students	-0.959**	0.407	-0.348
Revenues per student[b]	-0.222**	0.102	-0.081
Median household income[b]	-0.003	0.007	-0.001
Log-likelihood		-323.73	
Percent correctly predicted		65.2	

		Actual adoptions	
		Yes	No
Predicted adoptions	Yes	104	74
	No	104	229

[a] Average of estimated individual changes in probabilities, weighted by inverse of sampling probabilities; [b] In thousands of dollars; Significance of coefficients: ** 5% level; Standard errors (S.E.) are Huber-White accounting for sampling weights; Estimation including state dummies.

grams. Thus we can hope that the spatial probit, where we rely on neighbors' mean characteristics in order to identify the impact of neighbors predispositions towards adoption, is capable to provide significant results on potential interdependencies among districts.

Results for the probit with spatial correlation in the latent variable are presented in Table 4. They suggest that predispositions towards adoption of inter-district open enrollment are positively interdependent among school districts. The positive and significant coefficient for the spatial lag, ϕ, measuring the impact of composite neighbors' attitudes on the predisposition towards adoption, indicates that open enrollment policies of local school districts are significantly affected by the anticipated behavior of neighboring districts. Apart from the predisposition of neighbors, a number of district characteristics affect the discrete choice decision whether to participate in open enrollment. High school districts are about 30% less likely to allow for inter-district transfers relative to unified school districts. As expected, districts with crowded schools are less in-

Table 4: Spatial probit

Explanatory variables	Spatial correlation in latent variable only: Estimates	Slope[a]	Spatial correlation in latent variable and errors: Estimates	Slope[a]
Neighbors' predisposition, ϕ	0.491 **	-	0.534 ***	-
	(0.192)		(0.205)	
Central city	-0.045	-0.017	-0.042	-0.016
	(0.219)		(0.199)	
Suburb	-0.165	-0.062	-0.145	-0.056
	(0.175)		(0.170)	
Elementary school district	0.289	0.109	0.267	0.103
	(0.259)		(0.250)	
High school district	-0.789 *	-0.297	-0.811 *	-0.312
	(0.433)		(0.420)	
Student-teacher ratio	-0.088 **	-0.033	-0.080 *	-0.031
	(0.040)		(0.041)	
% minority students	-0.639 *	-0.240	-0.581 *	-0.223
	(0.344)		(0.334)	
Revenues per student[b]	-0.200 **	-0.075	-0.177 **	-0.068
	(0.078)		(0.085)	
Median household income[b]	-0.002	-0.001	-0.002	-0.001
	(0.005)		(0.005)	
Spatial correlation in errors, ρ	-	-	-0.447 *	-
			(0.243)	
Log-likelihood	-319.98		-319.53	
Percent correctly predicted	65.9		66.7	

		Actual adoptions: Yes	No	Actual adoptions: Yes	No
Predicted adoptions	Yes	107	73	109	71
	No	101	230	99	232

[a] Average of estimated individual changes in probabilities, weighted by inverse of sampling probabilities; [b] In thousands of dollars; Significance of coefficients: *** 1% level, ** 5% level, * 10% level; Huber-White standard errors accounting for sampling weights in parentheses; Estimations including state dummies.

clined towards admitting transfer students. An additional student per teacher lowers the probability that open enrollment policies are adopted by 3.3%. At the same time, a one percentage point increase in a district's share of minority students, with all other things being equal, makes the district 0.24% less likely to adopt open enrollment. This may reflect the fact that districts with a higher share of minority students will, on average, expect to be less successful in attracting students from elsewhere. Thus, it may not be worthwhile for these districts to adjust their policies towards open enrollment regulations. Furthermore, higher revenues per student make districts less willing to participate in inter-district school choice. $ 1,000 of additional revenues per student make the average district 7.5% less likely to admit non-resident students.

As mentioned in section 2, it is important to test for spatial error dependence as a potential source for spatial correlation in the districts' predispositions towards the adoption of school choice. I do this by estimating a weighted probit with interaction in the latent variable while allowing the errors to be spatially correlated according to eq. 6. The output for his regression is displayed as the second set of results in Table 4. The first and most important thing to note is that allowing for spatial error dependence does not break the link between neighbors' predispositions. On the contrary, the link becomes even stronger: The estimate for ϕ increases, and it is now significant at the 1% level. At the same time, weak evidence is found for the presence of negative spatial error correlation. The remaining parameter estimates are of similar size as before and need not be discussed again. Note that accounting for spatial error dependence does only marginally increase the log-likelihood of the model. The ability of the model to correctly predict the school districts' decisions is moderately strengthened. We can conclude that the positive spatial correlation in adoption decisions is not driven by spatial error correlation.

One could question the significance of the spatial effect in modelling the school districts' behavior towards school choice with reference to the share of correctly predicted decisions as a measure for the goodness of fit. Although the specifications accounting for interdependencies among districts outperform the baseline model in terms of predicting the districts' choices, the differences are not large. Therefore, as an additional check for the significance of the spatial effect, I computed the corresponding likelihood ratio test statistics. For the null hypothesis that the log-likelihood values are equal for the baseline model and the probit with spatial correlation only in predispositions, the statistic is 7.52. With one degree of freedom, this is significantly different from zero at the 1% level. For the test comparing the baseline model and the model with spatial correlation in predispositions and errors, the statistic is 8.42. With two degrees of freedom, this is significant at the 2% level. Hence, not accounting for the neighborhood influence removes a significant amount of information

Table 5: Non-linear least squares

Explanatory variables	Coef.	S.E.
Neighbors' predisposition, ϕ	0.377 ***	.125
Central city	-0.057	.083
Suburb	-0.088 *	.050
Elementary school district	0.204 **	.085
High school district	-0.213	.190
Student-teacher ratio	-0.045 ***	.012
% minority students	-0.409 ***	.135
Revenues per student^b	-0.079 ***	.026
Median household income^b	-0.003	.002

Significance of coefficients: *** 1% level, ** 5% level, * 10% level; Estimation including state dummies.

from the system.

Table 4 does not provide average partial effects for neighbors' predispositions. The reason is that ϕ is the coefficient of a latent variable. Thus, one cannot say anything about the quantitative effect of neighbors' predispositions on adoption probabilities. To get an impression of the strength of the neighborhood influence, it is useful to estimate a simple spatial linear probability model. This model looks like

$$y_i = \phi\, y_{-i} + x_i\beta + u_i. \tag{9}$$

Now, a district's policy depends on the actual adoption decision of the composite neighbor. The results displayed in Table 5 are derived from estimating a reduced form of eq. 9 by non-linear least squares. The reduced form is constructed along the same lines as described for the spatial probit in Section 2. The results suggest that in their open enrollment policies, the school districts are significantly and substantially influenced by their neighbors' actual policies. The estimate for the coefficient of the spatial lag indicates that a one percentage point increase in the share of neighboring districts allowing for inter-district transfers makes participation 0.38% more likely. Of course, given that the school districts' policies are discrete responses, the model from eq. 9 cannot be fully appropriate. Therefore, the numbers in Table 5 should be interpreted with caution. However, they suggest that the diffusion of school choice among districts is affected by a remarkably strong neighborhood influence.

5 Conclusion

It is in the nature of the political process that policy makers often face difficult discrete choice decisions. Particularly interesting and practically relevant examples are decisions to experiment with new political concepts. This paper provides evidence on the behavior of local governments in the adoption of a significant policy innovation in a large number of local jurisdictions. More specifically, the paper deals with the question how U.S. school districts have interacted with each other in the adoption of inter-district public school choice. The results from estimations allowing for spatial dependence of adoption decisions suggest that local governments are positively affected by anticipated decisions in benchmark jurisdictions. Moreover, the impact of reference jurisdictions on local policies seems to be substantial. This supports the view that in federal systems and, more generally in systems with a decentralized public sector, the diffusion of policy innovations is stimulated by horizontal interaction among jurisdictions.

Still, there are many open questions with respect to decentralized decision making and the diffusion of policy innovations. For instance, in the identification of spatial interactions between jurisdictions, this paper relies on the spatial distribution of adoption decisions in a cross-section of local jurisdictions. Clearly, for future empirical research, it should be worthwhile to also take into account dynamic aspects of the diffusion of new political technologies.

References

Baicker, K. (2005): "The Spillover Effects of State Spending," *Journal of Public Economics*, 89, 529–544.

Besley, T., and A. Case (1995): "Incumbent Behavior: Vote-Seeking, Tax-Setting, and Yardstick Competition," *The Amercian Economic Review*, 85(1), 25–45.

Bivand, R., and S. Szymanski (1997): "Spatial Dependence Through Local Yardstick Competition: Theory and Testing," *Economics Letters*, 55, 257–265.

Bordignon, M., F. Cerniglia, and F. Revelli (2003): "In Search of Yardstick Competition: A Spatial Analysis of Italian Municipality Property Tax Setting," *Journal of Urban Economics*, 54, 199–217.

Brett, C., and J. Pinkse (1997): "Those Taxes are All over the Map: A Test for Spatial Dependence of Municipal Tax Rates in British Columbia," *International Regional Science Review*, 20, 131–151.

Brock, W. A., and S. N. Durlauf (2001): "Discrete Choice with Social Interactions," *Review of Economic Studies*, 68, 235–260.

Brueckner, J. K. (2003): "Strategic Interaction Among Governments: An Overview of Empirical Studies," *International Regional Science Review*, 26, 175–188.

Brueckner, J. K., and L. A. Saavedra (2001): "Do Local Governments Engage in Strategic Property Tax Competition?," *National Tax Journal*, 54(2), 203–229.

Büttner, T. (2001): "Local Business Taxation and Competition for Capital: The Choice of the Tax Rate," *Regional Science and Urban Economics*, 31, 215–245.

Case, A. (1992): "Neighborhood Influence and Technological Change," *Regional Science and Urban Economics*, 22, 491–508.

Case, A., J. R. Hines, and H. S. Rosen (1993): "Budget Spillovers and Fiscal Policy Interdependence: Evidence from the States," *Journal of Public Economics*, 52, 285–307.

Hautsch, N., and S. Klotz (2003): "Estimating the Neighborhood Influence on Decision Makers: Theory and an Application on the Analysis of Innovation Decisions," *Journal of Economic Behavior and Organization*, 52, 97–113.

Hayashi, M., and R. Boadway (2000): "An Empirical Analysis of Intergovernmental Tax Interaction: The Case of Business Income Taxes in Canada," *Canadian Journal of Economics*, 34, 481–503.

Hirshleifer, D., and S. H. Teoh (2003): "Herd Behavior and Cascading in Capital Markets: A Review and Synthesis," *European Financial Management*, 9(1), 25–66.

Hoxby, C. M. (2000): "Does Competition Among Public Schools Benefit Students and Taxpayers?," *The American Economic Review*, 90, 1209–1238.

——— (2003): "School Choice and School Productivity (Or, Could School Choice Be a Rising Tide That Lifts All Boats?)," in *The Economics of School Choice*, ed. by C. M. Hoxby, pp. 287–341. Chicago: University of Chicago Press.

Katz, M. L., and C. Shapiro (1986): "Technology Adaption in the Presence of Network Externalities," *Journal of Political Economy*, 94, 822–841.

Kotsogiannis, C., and R. Schwager (2004): "Policy Innovation in Federal Systems," Mimeo, University of Exeter.

Krasker, W. S., E. Kuh, and R. E. Welsch (1983): "Estimation for Dirty Data and Flawed Models," in *Handbook of Econometrics, Volume I*, ed. by Z. Griliches, and M. D. Intriligator, pp. 651–698. North-Holland Publishing Company, New York.

National Center for Education Statistics (1996): *1993-94 Schools and Staffing Survey: Data File User's Manual.* U.S. Department of Education, Office of Educational Research and Improvement, Washington D.C.

——— (1998): *The Schools and Staffing Survey (SASS) and Teacher Followup Survey (TFS) CD-ROM: Electronic Codebook and Public-Use Data for Three Cycles of SASS and TFS.* U.S. Department of Education, Office of Educational Research and Improvement, Washington D.C.

Oates, W. E. (1999): "An Essay on Fiscal Federalism," *Journal of Economic Literature*, 37, 1120–1149.

Rincke, J. (2005): "Yardstick Competition and Policy Innovation," ZEW Discussion Paper 05-11.

Ryan, J. E., and M. Heise (2002): "The Political Economy of School Choice," *The Yale Law Journal*, 111, 2043–2136.

Scharfstein, D. S., and J. C. Stein (1990): "Herd Behavior and Investment," *American Economic Review*, 80(3), 465–479.

Strumpf, K. S. (2002): "Does Government Decentralization Increase Policy Innovation?," *Journal of Public Economic Theory*, 4(2), 207–241.

Wooldridge, J. M. (2002): *Econometric Analysis of Cross Section and Panel Data*. The MIT Press, Cambridge (Mass.).