Environment and Planning C: Government and Policy 2005, volume 23, pages 657 – 677

DOI:10.1068/c16r

Spatial diversity in local government revenue effort under decentralization: a neural-network approach

Mildred E Warner
Department of City and Regional Planning, 215 W Sibley Hall, Cornell University, Ithaca, NY 14853-6701, USA; e-mail: mew15@cornell.edu

James E Pratt
Department of Applied Economics and Management, Warren Hall, Cornell University, Ithaca, NY 14853, USA; e-mail: jep3@cornell.edu

Received 26 May 2004; in revised form 28 April 2005

Abstract. Decentralization reflects a global trend to increase the responsiveness of state and local governments to economic forces, but it raises the challenge of how to secure redistributive goals. Theoretically, as the equalizing impact of federal aid declines under devolution, we expect subnational state-level government policy to become more important, and geographic diversity in local governments' efforts to raise revenue to increase. In this paper we explore the impact of state fiscal centralization and intergovernmental aid on local revenue effort with the aid of Census of Governments data for county areas from 1987 for the Mid-Atlantic and East North Central region of the United States, with particular attention paid to rural counties. The 1987 period was chosen because it is the first year in which state policy trends diverged from federal decentralization trends and both state aid and state centralization *increased* while federal aid to localities continued to decline. Using a neural-network approach, we explore the spatially differentiated impact of state policy and find complementary responses in effort among some localities and substitution responses among others. Classification-tree analysis of this diversity suggests that decentralization and the competitive government it promotes are likely to exacerbate inequality among local governments.

Introduction

Theoretically, local government is considered the developmental or competitive state (O'Connor, 1973; Peterson, 1981). The promise of decentralization is that competition between levels of government will promote greater responsiveness to local needs and fiscal discipline (to prevent oversupply of public goods), thereby enhancing efficiency and democracy (Bennett, 1990). Limitations on sources of local revenue require that local governments pay careful attention to the economic development impact of their expenditures (Peterson, 1981). As local responsibility for services rises under decentralization, local governments' efforts to raise revenue (effort) and intergovernmental aid from higher levels of government become important in determining the potential for redistribution. One critical challenge of decentralization is that it reduces the capacity for redistribution, especially for areas with higher need (Prud'homme, 1995). As the nation shifts power downward under decentralization, more attention needs to be focused on the subnational state and its role in redistribution (Jessop, 1994; Lefevbre, 1974).

In this paper we examine the role of state policy on local revenue effort in the Mid-Atlantic and East North Central region of the United States in the late 1980s. This period was chosen because 1987 was the first year that state policy trends diverged from federal trends and both state aid and state centralization *increased* while federal aid to localities continued to decline. Specifically, we model the spatially differentiated impact of state fiscal centralization and intergovernmental aid on effort, paying particular attention to rural counties. Theoretically, we expect these policies to play out differently across localities—potentially increasing diversity and inequality. We demonstrate the use of

neural-network and classification techniques to capture the variation in local government response. Results show both complementary and substitution responses among localities. These results reinforce notions that decentralization and the competitive government it promotes can lead both to vicious and to virtuous cycles, each likely to exacerbate inequality over time—especially for higher need rural areas.

Decentralization, local fiscal effort, and spatial inequality

Tensions between national and state-level authority have led both to decentralizing and to recentralizing tendencies in US public finance (Conlan, 1998; Donahue, 1997; Gold, 1995; Katz, 2001; Powers, 1999). The period since 1970 has been characterized by three different decentralization approaches (Nathan and Lago, 1990). Under President Nixon in the 1970s, decentralization took the form of real fund transfers to local governments through a new program of general revenue sharing, and federal aid to localities increased significantly—peaking in 1977. In the 1980s, under President Reagan, decentralization took the form of a transfer of responsibility to state and local government, with a reduction in federal aid and the elimination of general revenue sharing. In the 1990s federal aid continued to fall and President Clinton responded with 'mandate relief' by decentralizing authority for some programs—most notably, welfare. Although such policies may reduce local government fiscal burden, they lead to a varied landscape of entitlement and social rights which Katz (2001) has termed 'the price of citizenship'. The first period of decentralization, under Nixon, represented an explicit effort to reduce inequality across localities. But in the following two decentralization periods, under Reagan and Clinton, little regard was given to spatial inequality.

Real declines in federal aid and the devolution of responsibilities to lower levels of government put increased pressure on local governments to raise their own funds. Decentralization encourages state and local governments to be more directly involved in promoting their own economic competitiveness and, as local government investment becomes more tied to economic competitiveness, we also expect increasing divergence in the fiscal capacity of local governments. Keenly aware of the mobility of higher income citizens and capital, local governments will increase investments in services which have a developmental benefit (Peterson, 1981; Tiebout, 1956). Redistributive expenditures, which have never been a high priority for local governments, will be even less attractive in an environment focused on economic competitiveness and this will likely result in greater inequality across space (Brenner, 1999; Katz, 2001; Lobao et al, 1999). Thus, it is imperative that studies of decentralization pay special attention to differences across space.

Neoclassical (Peterson, 1981) and Marxist (O'Connor, 1973) scholars both describe local government as the developmental state, unlikely to invest in redistributive expenditures because these will not promote economic growth and the local tax base. They argue that the redistributive role is best handled at the federal level, where the base across which taxes and expenditures may be distributed is wider. Under devolution, we expect an intensification of uneven geographic development as national policy focuses on global competitiveness at the expense of redistribution. Jessop (1994) argues that, as the nation shifts power downward through decentralization, this "creates space for a subnational resurgence" (page 264). The regional-level state plays a key role in helping to manage space on a larger scale than local government (Lefevbre, 1974). Although local fiscal capacity is primarily determined by economic forces which affect local employment and income (Ladd and Yinger, 1994), structural factors such as service responsibilities, tax limits, and intergovernmental aid (Bradbury et al, 1984; Ladd and Yinger, 1989) also determine local fiscal capacity. These structural factors are primarily determined at the state (subnational) level in the USA. In this paper we focus on the interaction between state-level fiscal policy and local effort.

State policy affects local effort primarily through state aid to localities and the level of centralization of fiscal responsibility. If decentralization is to promote efficiency without increasing inequality, then the nature of the response of local governments to state policy is critical. The key question explored in this paper is whether state aid and state centralization have a complementary or a substitution effect on local effort. State expenditure can substitute for effort through higher aid or more fiscal centralization. If state aid were targeted to poorer places, or if state centralization were targeted to services which cost more in areas of higher relative need, then a local substitution (negative) response would have a redistributive effect—reducing effort in poorer places and increasing it in richer places. A complementary (positive) local revenue response might increase spatial differences in effort. The 1987 period was chosen because it was the first year in which state policy trends diverged from federal decentralization trends, and both state aid and state centralization *increased* while federal aid to localities continued to decline. Studies of the symmetry of local revenue responses to increases or decreases in intergovernmental aid (Stine, 1994), or increases in aid as compared with increases in local income (Whyckoff, 1985; Zampelli, 1986) for this period have found some support for a complementary response, but these studies were focused primarily on urban areas. Among the studies which include rural counties for the 1987 period, only one found a complementary effect: state aid grew less in counties with lower government expenditure (Reeder and Jansen, 1995). Other rural-focused studies found state aid to have a substitution effect (Johnson et al, 1995), or no effect (Warner, 2001), on effort. All three of the rural-focused studies of the 1987 period found that state centralization had a substitution effect on effort (Johnson et al, 1995; Reeder and Jansen, 1995; Warner, 2001). Studies which include rural counties are restricted to data from the quinquenial Census of Government finance, so looking at changes in response over time is limited to a five-year time lag. The time lag in local government response to changes in state policy is quite short and has to occur within the same budget year because local governments are not allowed to deficit spend.

We are interested in the spatial variation in local effort responses to state fiscal policy. Quantitative models of cross-sectional data face the challenge that units may be more heterogeneous across space than through time. Of the three rural-focused studies cited above, only one controlled for heteroscedasticity. Warner (2001) used state weights to acknowledge the importance of differences in state policy on effort. Another, increasingly popular, method of explicitly introducing space is to correct for spatial autocorrelation (Anselin, 1988; Mencken, 2000).[1] One problem with such statistical approaches is that they require the researcher to stipulate, a priori, the nature of the spatial relation. In the following analysis, we demonstrate an alternative approach to understanding spatial differences which is based on neural networks.

[1] 'Spatial autocorrelation' refers to spatial dependence in which an observation at one point in space is functionally related to nearby observations. When values are correlated geographically, the statistical assumption of independence is violated. This can be caused by measurement error, when data are aggregated and miscalculations in one spatial unit 'spill over' to neighboring units, or to interdependence in space when high values in one unit are associated with low values in another (negative autocorrelation which creates a checkerboard pattern), or like values cluster together creating a lattice effect (positive autocorrelation), or values follow a gradient of diffusion (for example, high values in metropolitan areas which fall as one moves toward the suburbs).

Using neural networks to model spatial diversity

Neural networks have been used in mathematics and computer science for some time. They first began appearing in the social science literature in the early 1990s. *Social Science Computer Review* devoted a special issue to the use of neural network applications in social science topics in 1991 (Garson, 1991). In many of these papers neural-network results were compared with regression, discriminant, or path analysis and superior forecasting abilities of neural networks were found (Garson, 1991; Kastens et al, 1995; Meraviglia, 1996; Wier and Phoha, 2002). Policy often involves poorly defined problems and poorly measured variables. In situations of incomplete or inaccurate data, and incomplete theoretical understanding, performance of traditional multivariate statistical procedures deteriorates badly, but neural networks excel (Woelfel, 1993). Although use of neural networks for policy analysis is still limited, these techniques are proving especially useful for classification (Garson, 1998). A particular advantage of neural-network models for the study at hand is their ability to address and reveal individual differences in observational units (county areas). We expected a wide variation of county responses to changes in state aid or state centralization within the study area.

Brief background on neural-network modeling

A neural network is a collection of many simple and highly interconnected processors or 'neurons' that process information in parallel.[(2)] Most neural networks have an input layer, an output layer, and an unspecified number of 'hidden layers'. Hidden layers allow for interactive associations among the inputs. Activation functions for the weights in the direct connections and the hidden layers are used to introduce complex nonlinearities in the direct effects and in the interactions. The ability to represent nonlinearity makes neural-network models with hidden layers extremely powerful predictors. Almost any nonlinear function can be used for the activation function, but it must be monotonic. Sigmoidal functions (that is, logistic and Gaussian functions) are the most common choices and are used in this analysis. Much of the recent interest in neural-network models can be explained by the inclusion of hidden layers to model interactions and the flexibility possible in functional forms.

Neural networks may be classified into two broad categories: feedforward and feedback. In feedforward networks, signals flow in only one direction, and outputs are dependent only on the signal incoming from the neurons in the previous layer. Feedback networks, by contrast, have looping features built into the system. A feedforward approach is used in this analysis.

As with regression, neural-network models are based on a process of minimizing errors. As has been well documented (Wang et al, 2004), a regression problem can be restated as a mathematical-program problem with the minimization of the sum of squared errors serving as the objective. The normal regression procedure is equivalent to an optimization problem that maximizes R^2. By means of mathematical optimization to search for a set of weights on activation functions that minimize the predicted errors, a neural network is 'trained'—much like a regression can be estimated by means of quadratic optimization. Through this system of determining observation-dependent weights, differing responses to variables for individual observations (counties, in this case) are computed. Flexibility in variable type (interval, binary, ordinal, and nominal) permits the inclusion of geographic and sociological concerns that do not lend themselves well to numerical expression (such as state, and the urban–rural continuum code in our model).

[(2)] Neural-network methods encompass a broad class of flexible nonlinear regression and discriminant models, data-reduction models, and nonlinear dynamic systems (Bishop, 1995; Lawrence, 1994). To distinguish the use of neural network models from true biological neural networks, models such as ours are often called 'artificial' neural networks.

The disadvantages of neural-network methods include the current lack of a generally accepted procedure for conducting tests of significance and building confidence intervals (Hwang and Ding, 1997; Veaux et al, 1998). The estimated weights that result from the neural-network analysis often require that additional analysis be performed to generate meaningful interpretations. Additionally, unlike the somewhat precise computational recipe for regression, the various numerical algorithms used for optimizing the complex neural network error functions may produce varying results. The addition of SAS Inc.'s Enterprise Miner (which we used for this analysis) to the suite of neural-network tools available to academic researchers should improve overall access to neural-network tools and replicability of neural-network results.

Study region, time period, and variables

To understand the differential spatial impact of state policy on local effort better we analyzed 587 county areas in the eight-state Mid-Atlantic and East North Central region of the USA—New York, New Jersey, Pennsylvania, Ohio, Indiana, Illinois, Wisconsin, and Michigan) (see figure 1). These states share a similar history as long-time industrial/manufacturing states which experienced deindustrialization prior to 1987. The region also has a long history of civic engagement and popular support for local and state government investment. National studies that include states experiencing later industrialization (the South) and states with minimalist government (South and West), often include the use of regional variables to distinguish the differing effects between the 'rust belt' and 'sun belt' (Grant and Wallace, 1994; Mollenkopf, 1983). By selecting a region with a similar economic and political history, we are able to focus on intraregional differences at the county level.

Our data are drawn from the US Census of Government for 1987 and the US Census of Population and Housing for 1990. The US Census of Government collects data on local government revenue and expenditure every five years, and is the only source of comprehensive data on rural governments. Our unit of analysis is county areas, which are aggregations of all governments within a county. Special districts which cross county boundaries are counted within the county where the administrative headquarters is located. Studies addressing rural areas typically use county areas rather than municipalities as the unit of analysis because county areas include all local

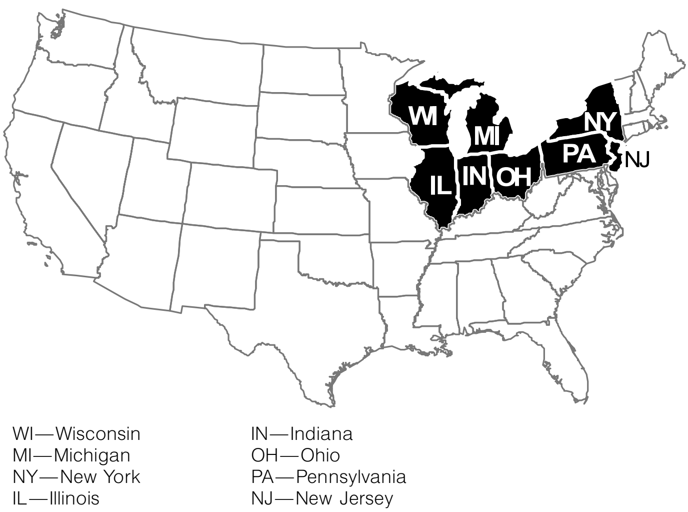


Figure 1. The eight-state, Mid-Atlantic–East North Central Region of the USA.

governmental units (school districts, towns, villages, and counties) (Dewees et al, 2003; Lobao et al, 1999).

Understanding local revenue effort

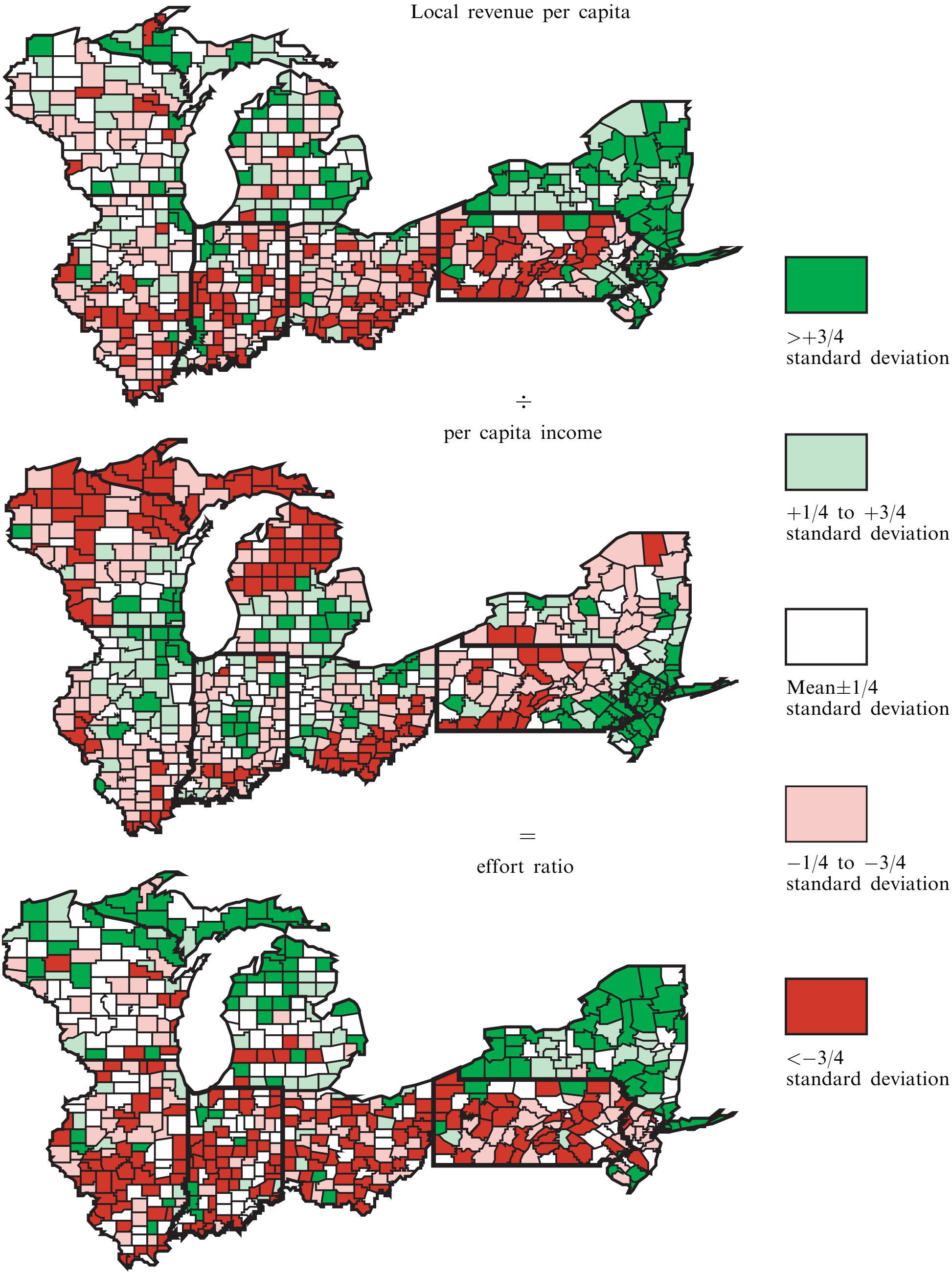
This analysis is focused on local 'effort', which is measured as the ratio of per capita local own-source revenues to per capita income. Given the difficulty of comparing the assessed value of real property across jurisdictions (because of the lack of equalization of assessments), per capita income is used as our standard measure of the capacity of local governments to raise revenues. Local government revenues depend primarily on property taxes, although dependence on sales tax and user charges is increasing. However, ultimately, taxes are paid with income. The maps in color plate 1 demonstrate visually how the components of the effort ratio are spatially related at the county level across the study area.

Data in each map are divided into five categories. White is the center of the distribution for each variable and includes one quarter of a standard deviation above and below the 587-county mean. The next two categories in each direction are also one half of a standard deviation in size. The lighter shade of red represents the distribution from one quarter to three quarters of a standard deviation below the mean, and the lighter shade of green represents a similar area above the mean. The two extreme categories, dark green and dark red, include all values either above or below three quarters of a standard deviation from the mean.

In the map of locally raised revenue per capita, New York and New Jersey counties stand out as predominantly above the 587-county mean, whereas Illinois, Indiana, Ohio, and Pennsylvania are predominantly below average. However, all states except New Jersey and New York have county units that span the entire range of locally raised revenue per capita. The level of locally raised revenue varies across counties according to need, costs of service delivery, and the capacity to raise revenue. Core metropolitan counties raise 50% more local revenue per capita than their rural counterparts.

The second map represents the distribution of per capita income. Again, we see visual evidence of broad contiguous groups of counties within states that can be characterized as either higher or lower than the 587-county average. Yet all states except New Jersey have counties that span the entire range of per capita income categories. High incomes are primarily clustered around metropolitan areas—the Philadelphia to New York corridor, the western shore of Lake Michigan, the Detroit metropolitan region, and the other major metropolitan regions: Chicago, Indianapolis, Columbus, Cleveland, Buffalo, and Pittsburgh. The lowest incomes are found in rural counties.

The third map shows the distribution of 'effort', which is defined as locally raised revenue per capita divided by per capita income. Visually, effort appears to be dispersed similarly to locally raised revenue. Closer inspection of the two components of this simple measure, however, reveals some interesting relationships. In New Jersey the relatively high locally raised revenue, combined with the very high per capita income, results in generally low levels of effort. In New York locally raised revenue is high enough across a wide spectrum of high and low per capita income counties to produce an almost universally high level of effort. That is, low-income New York counties raise local revenues at a rate that places them relatively high in their effort and, unlike New Jersey, high-income New York counties raise even higher levels of local revenue so that they, too, are placed high on effort. In northern Michigan and Wisconsin, the low levels of per capita income were sufficiently low that, when combined with the mixed levels of locally raised revenue, generally high levels of effort resulted. Across Illinois, Indiana, and Ohio, relatively low levels of locally raised revenue, even in many low per capita



Effort = Per capita locally raised revenues (taxes, user fees, and miscellaneous income) 1987/per capita income 1989 (deflated 1987—100)

N = 587 county areas, Mid-Atlantic and East North Central States

Color plate 1. Effort ratio as a function of locally raised revenues and per capita income (source: US Census of Governments Finance Files, 1987 for County Areas).

income counties, result in generally low levels of effort. Yet, even in these areas, individual exemplars of high effort in low-income counties can be found. High effort in the urban counties may be a reflection of higher costs of service provision despite the higher incomes. For the rural counties, high effort may reflect extremely low incomes in the face of minimum service-provision levels. Although there are discernable differences by state and metropolitan status, we see wide spatial variation among counties in the study area.

The geographical variation in effort does not lend itself to the standard spatial degradation functions commonly used to control for spatial autocorrelation (decay over distance or consistent near-neighbor effects). Additionally, statistical procedures, such as regression, that produce a single coefficient relating an independent variable to the conditional mean of the dependent variable would not be able to accommodate the constellation of conditions we suspect are instrumental in determining individual county responses to conditions. Effort is affected by more than per capita income differences: the level and redistributive nature of state aid, and the impact of state centralization are expected to play important roles. Although easy to interpret and explain to policymakers, a single regression-determined relationship between effort and per capita income, even one conditioned on one or more dummy variables such as 'state', or 'rural' versus 'urban', would be inadequate to reveal the true variation and complexity of the underlying relationships.

The model

To address the impact of these broader state policy variables specifically we constructed a neural-network model, with effort as the target variable and the following input variables.

Local effort = f {population, costs (U-shaped cost curve), spatial effects (urban-rural continuum, percentage urban, state), need (percentage poverty), local capacity (per capita income), federal policy (federal aid), state policy (state aid, level of state centralization, average overall state and local government expenditure), and local income inequality}.

Figure 2 represents the neural-network model used in our analysis. Direct effects of inputs on outputs are represented by the direct lines in figure 2, and interactions are modeled by the lines connecting the input layer to the hidden layer.

Input variables

Population (POP) *and density* (DENSITY)

Local effort levels reflect differences in costs, need, and demand. Although costs of service delivery and level of need and demand vary across localities, with the exception of Reeder and Jansen (1995), most researchers attempting to quantify these differences in the 1980s decade focused primarily on urban areas (Bradbury et al, 1984; Ladd and Yinger, 1989). Population is a common measure of need, but it does not reflect the higher costs at the two ends of the density spectrum. For urban areas, congestion requires more services be provided publicly: for example, water and sewers to protect public health. In rural areas sparsity of population may reduce the need for provision of services, but when services are publicly provided, sparsity increases the per unit costs. We account for these higher costs with a DENSITY and DENSITY^2 variable which allows for a U-shaped cost curve. The places with the lowest values for these variables are the smallest rural counties and Indian reservations in Michigan, Wisconsin, and New York; the highest values are found in the largest urban centers: Cook County (Chicago) and Manhattan (NY). Table 1 (over) provides descriptive statistics for all variables.

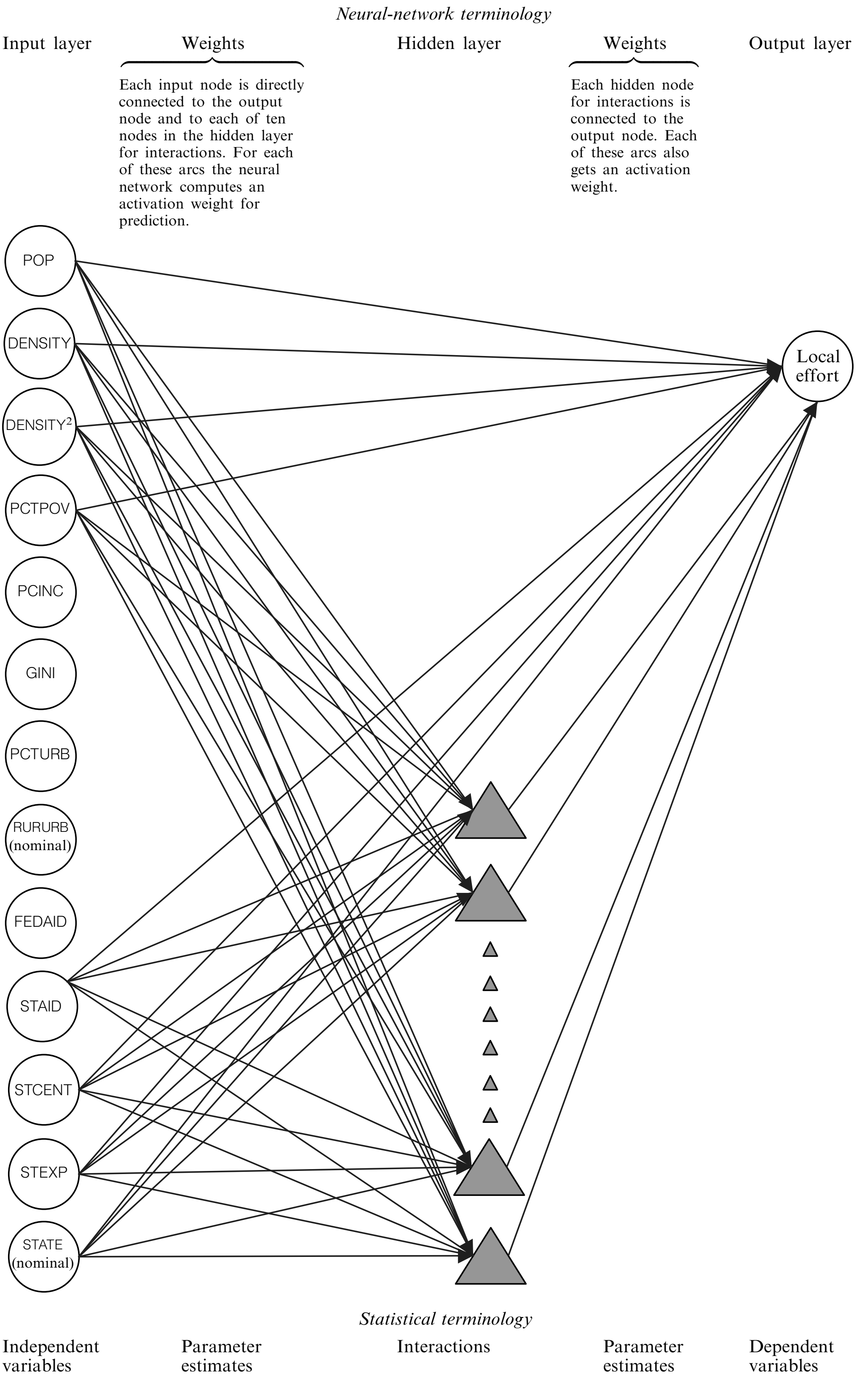


Figure 2. Conceptual framework of a three-layer neural network for local effort.

Table 1. Model variables—descriptive statistics.

Variable[a]	Mean	Standard deviation	Minimum	Maximum
Effort, 1987[b]	8.3	3.06	2.80	28.18
Population (POP), 1987	135 247	331 282	2 000	5 291 100
DENSITY, 1990	527	3 025	3.1	52 415
$DENSITY^2$, 1990	9 411 495	126 967 787	9.6	2.75×10^9
Percentage poverty (PCTPOV), 1990	12.09	5.00	2.2	48.7
Per capita income (PCINC), 1989 (deflated, consumer price index 1987 = 100)	10 516	2 334	4 991	23 911
GINI, 1990	0.40	0.027	0.335	0.574
Percentage urban (PCTURB), 1990	41.7	27.8	0	100
Rural–urban continuum codes (RURURB), 1993	na	na	0	9
Federal aid (FEDAID), 1987	50	38	6	261
State aid (STAID), 1987	538	177	123	1 231
State centralization (STCENT), 1987[c]	40.5	2.5	34.2	42.6
STATEAVREXP, 1987[d]	2 694	492	2 171	3 917
STATE	na	na	na	na

[a] All finance variables are given in $ per capita.
[b] Effort is defined as 100 (locally raised revenue/per capita income).
[c] State centralization (STCENT) is defined as 100 (direct general expenditures by state/direct general expenditure by state and local government). This is a state-level variable: it does not vary by county.
[d] STATEAVREXP—average state and local expenditures per capita, 1987 (taken from the denominator of state centralization). This is a state-level variable: it does not vary by county.
Sources: 1990 figures—US Census of Population and Housing, 1990; 1987 figures—US Census of Government Finance Files, 1987; 1993 Beale—USDA, 1993. Rural-Urban Continuum Codes, based on 1990 Census Data.
na—not applicable.

Rural–urban continuum

We also expect higher costs for more urban places—especially for core metropolitan counties that suffer from aging infrastructure compared with their fringe suburban counterparts. For rural areas, costs are higher for nonadjacent places which cannot benefit from tax exporting[(3)] or service spillovers from neighboring counties. We include the rural–urban continuum codes (RURURB) based on size of central place, and adjacency to a metropolitan county.[(4)]

(3) Tax exporting is the ability to shift tax burden to nonresidents through commuting, sales, and income taxes (Ladd and Yinger, 1989).

(4) Rural–urban continuum codes are developed by the US Department of Agriculture, on the basis of data collected with each decennial census. Counties are grouped into ten categories based on size of central place and adjacency to metropolitan counties as follows: '0 Large Metro Core'—central counties of metropolitan areas of 1 million population or more: '1 Large Metro Fringe'—fringe counties of metropolitan areas of 1 million population or more; '2 Medium Metro'—counties in metro areas of 250 000 to 1 million population; '3 Small Metro'—counties in metro areas of less than 250 000 population; '4 Large Rural Adjacent'—urban population greater than 20 000 adjacent to a metropolitan area; '5 Large Rural Non-Adjacent'—urban population greater than 20 000, not adjacent to a metropolitan area; '6 Medium Rural Adjacent'—urban population 2500 to 19 999, adjacent to a metropolitan area; '7 Medium Rural Non-Adjacent'—urban population 2500

Poverty and inequality
Localities with high poverty often need more services, but theory suggests that the effectiveness of low-income residents' demand for services is weak (Peterson, 1981). The capacity to pay is also weak—at least for rural areas. Reeder and Jansen (1995) found that effort levels were lowest in the poorest Southern US rural counties, where need is highest. Rural counties with high levels of poverty tend to have low per capita income. Urban counties with high levels of poverty tend to have high per capita income. Higher income communities can provide a higher level of services with lower effort because they have higher capacity. All else being equal, we would expect lower effort in places with higher incomes. Schneider (1989) suggests that, where demand is more homogeneous, effort levels will be higher. Thus places with lower GINI coefficients are expected to have higher effort. There is a wide range in poverty, income, and inequality in the sample. The minimum and maximum values for poverty rates are found in the rural counties; but Manhattan (NY) shows the highest income and inequality.

Federal and state aid (FEDAID *and* STAID)
The traditional justification for intergovernmental aid has been to equalize service provision, given unequal need and unequal capacity (Ladd and Yinger, 1994). Government investment can play an equalizing role across space and time (Johnson et al, 1995). With increasing emphasis on the developmental role of the state brought about by globalization, the potential for redistributive impacts may decline. In 1987 federal aid was less than one tenth the level of state aid, on average. Interestingly, the counties with the minimum and maximum values for federal aid are both found in rural Indiana. State aid is larger and more variable than federal aid across counties in our study region. We expect state aid may have differential effects across localities—causing a reduction in effort in some places and encouraging increased effort in others. New York and Wisconsin have the highest levels of state aid in our study region, but these states also have more of the counties with high effort. By contrast, the other states in the study area have lower levels of state aid, but their counties have lower effort (except for Michigan, which has lower state aid and higher effort). This suggests that state aid alone is not sufficient to cause a substitution effect on effort.

State centralization (STCENT)
Decentralization implies shifts in the responsibilities for service provision. Even while service delivery responsibility is being shifted downward to lower levels of government, fiscal responsibility can remain centralized at the state level. We measure state centralization as the state share of total direct state and local expenditures, including capital investment. State aid to localities is counted in the local share. This measures the degree of centralization in fiscal responsibility for governmental services. Centralization varies considerably across the states in our study region, with the lowest levels in New York and Wisconsin and higher levels in the other states.

Other state differences
Historical patterns of government investment are also important. Important differences in the average level of expenditures across states reflect citizen preference for more or less government, or the inertia of past policies which get embedded in current budgets. 'State average expenditure' (STEXP) is the average per capita expenditure of state and local governments averaged across all local governments in the state. Indiana

(4) continued.
to 19 999, not adjacent to a metropolitan area; '8 Small Rural Adjacent'—places with population of less than 2500, adjacent to a metro area; '9 Small Rural Non-Adjacent'—places with population less than 2500, not adjacent to a metro area.

Table 2. State policy variables (source: Census of Government Finance Files 1987).

State	State centralization[a,b]	Average state expenditure (US $ per capita)[a,c]	State aid (US $ per capita)			
			mean	min.	max.	no. of counties
Illinois	41.1	2505	402	214	777	102
Indiana	42.4	2171	497	333	680	92
Michigan	42.6	2896	491	234	889	83
New Jersey	41.3	3040	610	296	1012	21
New York	34.2	3917	830	516	1231	62
Ohio	40.8	2452	508	360	754	88
Pennsylvania	42.1	2364	446	123	710	67
Wisconsin	38.4	2847	685	423	1038	72

min., minimum; max. maximum.
[a] State-level variable—does not vary by county.
[b] State centralization (STCENT) is defined as 100 (direct general expenditures by state/direct general expenditure by state and local government). This is a state-level variable: it does not vary by county.
[c] Average state expenditure (STATEAVREXP) is average state and local expenditure per capita, 1987 (taken from the denominator of state centralization). This is a state-level variable: it does not vary by county.

has the lowest and New York the highest state average expenditure per capita. A nominal variable for state (STATE) is included to capture attitudes about government (such as state-level differences in powers delegated to localities, limits on revenue raising, etc) not reflected in financial variables. Table 2 provides state-level detail on these variables.

Neural-network results

The weights that result from the training of a neural network do not possess a readily interpretable meaning—unlike regression coefficients. Because our primary interest was to understand the impact of state policy on local effort better, we had to develop a way of measuring the impacts of model variables on the target variable—local effort. Given the differences in units across the various input variables, we decided to compute elasticities for each of the 587 county units of observation with the aid of the neural-network results. An impact elasticity measures the percentage change in the target variable given a 1% change in the input variable. In ordinary least squares models, elasticities are usually reported at the mean values for the variables and, for most a priori functional form specifications, vary for different levels of the input variables. To obtain elasticities from the neural network, we created a dataset in which the input variable of interest (for example, state centralization—STCENT) was increased by 1%. We then applied this new dataset to the previously trained network (trained on the original data) to generate predictions of effort corresponding to the changed values of the input variable and calculated the new elasticities. This provides a unique elasticity measure for every county. For example:

Impact elasticity of state centralization (STCENT) on effort:

$$\frac{\widehat{\mathrm{EFFORT}}_{t_1} - \mathrm{EFFORT}_{t_0}}{\widehat{\mathrm{EFFORT}}_{t_0}} \bigg/ \frac{\mathrm{STCENT}_{101\%} - \mathrm{STCENT}_{100\%}}{\mathrm{STCENT}_{100\%}} = \varepsilon_{\frac{\mathrm{EFFORT}}{\mathrm{STCENT}}} .$$

This process was repeated for each independent variable of interest to generate the distribution of elasticities shown in figure 3. State centralization has the largest relative impact on effort of any of the variables: on average, a 1% increase in state

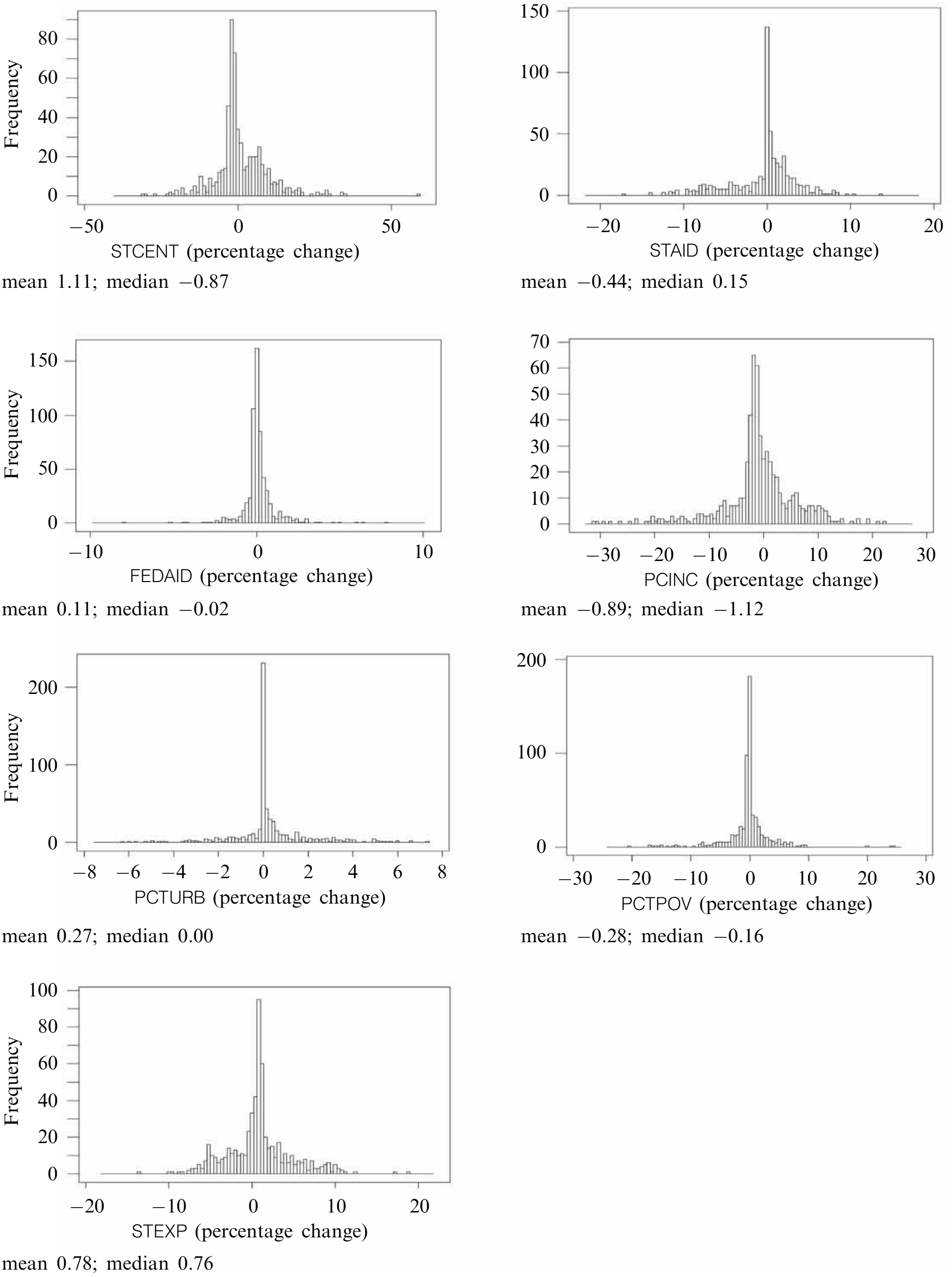


Figure 3. Distributions of impact elasticities on effort, given a 1% change in input variables ($N = 587$ county areas).

centralization (STCENT) results in a 1.1% increase in effort (see figure 3). Average state and local expenditure (STEXP) is the policy variable with the next largest relative impact: on average, a 1% increase in average expenditure is associated with a 0.78% increase in effort. These two state policy variables—centralization and average government expenditure level—have greater relative impacts on effort than does state aid. The next largest impact is found in per capita income (PCINC) where, on average, a 1%

increase in income results in a 0.89% decrease in effort. The importance of economic capacity of a locality is reflected in this per capita income impact.

The average effects, however, are not the focus of this neural-network analysis. Neural networks do not create a single impact value for all observations—unlike regression. The flexibility in functional form and the complexity of interactions permitted by the hidden layer produce a unique impact value for each county area in the study region. Histograms of the distribution of elasticity values for each variable show a strong central tendency, but wide dispersion.(5)

Our key interest is determining whether state aid and state centralization have a positive (complementary) or negative (substitution) effect on effort. The neural-network analysis reveals, however, that the signs of the average (mean value) and median impact elasticities for state aid and state centralization are different. Although the average impact of state centralization is positive, more than half of the study-area counties have a negative response to state centralization. Prior research (Johnson et al, 1995; Reeder and Jansen, 1995; Warner, 2001) showed state centralization to have a negative impact on effort. This substitution effect may be true primarily for counties that have high effort or high need. Effort in these places may already be at a political maximum and an increase in state fiscal centralization may substitute for local effort. Positive impacts occur where state centralization complements effort. This may be stronger in counties with low effort levels that use the funds freed up by state centralization to invest in other preferred local projects. However, positive impacts also may be found in counties with high need.

By contrast, although the average impact of state aid is negative (suggesting that state aid is a substitute for effort), for more than half the counties state aid has a complementary effect—increasing effort. This complementary effect may reflect matching grants, which require increased effort to secure additional state aid. The average and median values for federal aid elasticities also switch signs, illustrating that federal aid is a complement to effort in some counties and a substitute in others. The neural network has allowed us to see that the elasticity of response to changes in an input variable may vary considerably from county to county.

Both the mean value and the median share the same sign for the impact elasticities for average state and local expenditure, percent poverty, per capita income, and percent urban. The impact elasticity of state average expenditure on effort is generally positive. Counties in states with higher average expenditure will exhibit higher effort. The impact elasticities for per capita income show that, on average, a 1% increase in income will result in a reduction in effort—a result consistent with theoretical and empirical predictions (Peterson, 1981; Reeder and Jansen, 1995). Although the majority of places exhibit a negative impact elasticity to income, some counties show a positive response. Positive elasticities may reflect increased demand for services with rising wealth. Newly urbanizing areas often show increased effort to meet new infrastructure needs. For low-income rural counties, positive elasticities reflect low effort despite high need—a finding confirmed by Reeder and Jansen (1995).

The negative elasticity of poverty with respect to effort supports theoretical predictions that redistributive expenditures are unlikely among poorer places. However, for a significant number of places, an increase in poverty has a positive impact on effort,

(5) One of the reviewers expressed concern about the 'variability of the estimates'. In our judgment, the explicit computation of observation-dependent estimates and the resulting opportunity to study the distribution of these estimates (figure 3) is an important benefit of the use of a neural-network model for this analysis. In a case such as this, where we fully expect the county-level responses to exhibit variation across the various data dimensions, including space, allowing observation-dependent estimates provides useful policy information.

showing that redistribution is possible. This is more likely in urban areas which exhibit both high poverty and high income than in rural counties where high poverty is associated with lower income.

Although the average impact elasticity for percentage urban (PCTURB) is positive, in half the counties an increase in urbanization is associated with a drop in effort. In some cases, increased urbanization may reduce the costs of service provision as the density of hookups to preexisting infrastructure is increased—thus reducing per unit costs. This is the logic behind much urban-planning literature promoting infill development (Orfield, 1997). In other cases, increased urbanization may increase the costs of service provision by requiring extension of infrastructure to sparsely developed strip malls and subdivisions. This is reflected in the literature opposing sprawl (Altschuler et al, 1999). Neural-network analysis allows that each of these descriptions may be appropriate simultaneously for different counties.[6]

Classification of local effort responses to state centralization

The model underlying the interaction of state policy with effort is complex. The neural-network analysis has shown that this complexity results in large variation in the elasticity of effort to changes in state policy. Can we determine why some counties have a positive response and others a negative response to state centralization? To understand the nature of this complexity better, we used a classification tree. Decision, or classification, trees are a widely used form of quantitative logic (Brieman et al, 1984). This method classifies observations on the basis of their attributes. The output includes a decision tree that provides rule-based paths to the final classifications.

We took the predicted impact elasticities for state centralization from the neural-network model and developed a decision tree based on the same attribute data as was used for the original neural-network model. The classification algorithm then created decision rules to determine whether impact elasticities would be positive or negative. Beginning with the root node, the model develops rules, based on the attribute data, to split the data into two or more segments for classification. The number of splits at each level, the minimum size of each terminal node, and the number of levels in the tree can be specified by the user. We limited the number of splits to four, the minimum node

[6] As a result of concern with the 'variability' of the estimates produced by the neural-network analysis, a reviewer correctly questioned the interpretation of the results without the usual statistical tests of significance typically found in regression analyses. To our knowledge, generally accepted versions of these tests for neural-network analyses have not yet been developed. Also, although a sample size of 587 is not large, especially in this age of terrabyte databases, as the number of observations increases, the small-sample tests of significance we typically use with regression analysis will produce statistically significant results, even when there is little or no practical significance associated with the finding. There is a growing body of literature that documents the misuse of statistical significance testing (Ziliak and McCloskey, 2004) and calls for a more 'analytical significance' approach. This same reviewer also correctly suggested that part (one half) of the observations be used for the analysis and part (the other half) be used for validation. Splitting of samples for this purpose has long been suggested as a way to avoid pretest bias resulting from sequential modeling performed on a single dataset in search of "results that validate preconceived notions" (Tomek, 1985, page 911). Using a subset of the available observations (normally much smaller than one half) to validate the neural network results is highly recommended, especially for models that will be used for predictive, and not descriptive, purposes. Out-of-sample validation avoids the development of estimates that are highly 'trained' to reproduce the observations while not generalizing well to other data. It was our judgment that previous studies (Johnson et al, 1995; Reeder and Jansen, 1995; Warner, 2001) had provided sufficient guidelines with respect to the variables to be included in the analysis, so that no pretesting was needed. Additionally, for reasons noted above, we did not intend to generalize the results to counties outside the geographic study area, nor to time periods prior to 1987.

size to ten, and the maximum number of levels to five. The output of this classification method can best be described visually as a set of nodes and connecting arcs (see figure 4).

The first classification rule determined by the decision-tree algorithm was STATE. At each node after the root, a classification is assigned and either the path is terminated or another split is performed. At the second level, the attributes for the next split are different for each group of states. Each node is represented by a box in figure 4. The top line in each box shows the decision rule, expressed as a greater-than or less-than argument for some value of the splitting attribute. The bottom line shows the number of observations and, for terminal nodes, the percentage correctly classified in that node. At the bottom of each path in the tree is the terminal node, shown white for negative elasticity values, and dark for positive; undetermined nodes are shown in grey. The size of each terminal node varies relative to the number of observations it covers. Given that the splitting rules are mutually exclusive, the sum of all terminal nodes completely covers all the cases. From this tree, 77% of the predicted observations were correctly classified as either positive or negative.

The nominal variable STATE was the most important attribute in the classification—not state centralization. The policy structure, history, and attitudes toward government are better captured in the nominal STATE variable. This provides empirical support for the importance of state-level policy under decentralization (Jessop, 1994). The first level of the classification tree groups states according to the percentage of negative effort-response values. The highest is Indiana, at 88% negative elasticities, followed by New Jersey, New York, and Wisconsin at 66% negative, and then by Illinois and Ohio at 47% negative. The states with the lowest percentage of negative values (37%) are Michigan and Pennsylvania.

Our primary theoretical expectation was that the response of effort to increased state centralization would be negative—a substitution effect. Local government, as the developmental state, faces fiscal constraints which should discourage increased effort (Peterson, 1981; Schneider, 1989). The classification tree may be most interesting in helping us understand why some counties exhibit a positive response.

The paths in the tree show three primary explanations for positive responses.

Excess capacity which allows increased effort

This is illustrated by mid-sized suburban counties ($100\,550 < \text{population} > 218\,950$) in New York, New Jersey, and Wisconsin, which are experiencing growth, and hence need for infrastructure development, and are capable of a complementary response. Similarly, counties with above-average federal aid (between \$47 and \$65 per capita) in Illinois and Ohio show positive responses.

Lower income inequality which encourages increased effort

Theoretically, we expect counties with low effort, relative homogeneity, and higher capacity (such as more federal or state aid) to be more likely to exhibit an increase in effort in response to an increase in state centralization. Terminal nodes exhibiting these characteristics are found in Indiana ($\text{FEDAID} > \$44$), among rural counties with average state aid and lower inequality ($\text{GINI} < 0.38$) in New Jersey, New York, and Wisconsin, and among counties with average poverty ($\text{PCTPOV} < 0.17\%$) and average inequality ($\text{GINI} < 0.42$) in Michigan and Pennsylvania. These nodes provide partial support for Schneider's (1989) notion that lower income inequality could promote the mobilization of local revenue for local needs.

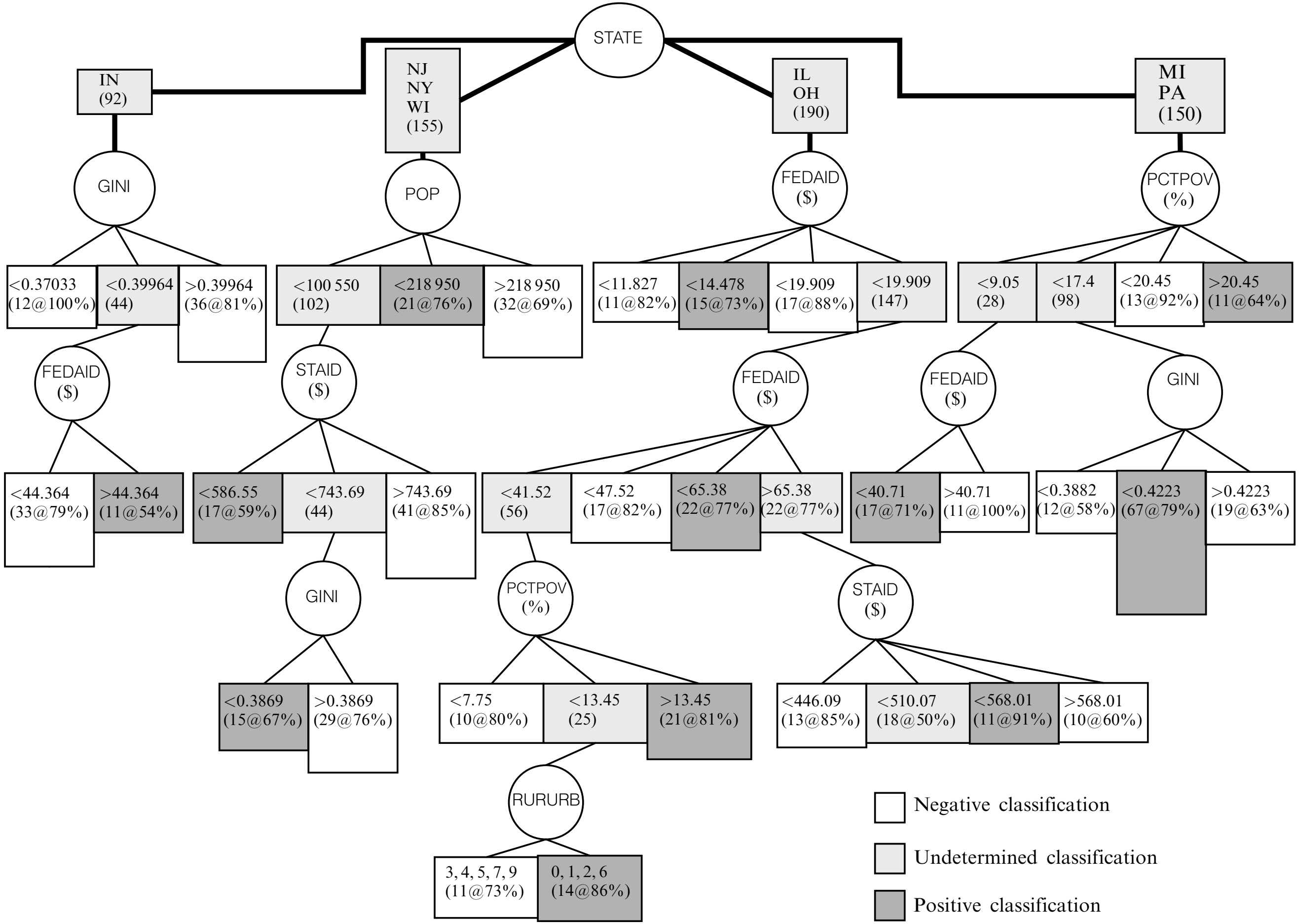


Figure 4. Classification tree for state centralization elasticities.

Higher need which requires increased effort
Rural places which do not have the political power to attract a proportional share of state aid may show higher effort in response to state centralization; these include the smallest counties in New Jersey, New York, and Wisconsin with below-average state aid (STAID < \$586 per capita). In Illinois and Ohio we see positive elasticities among counties with high federal aid but only moderate state aid (STAID between \$510 and \$568), suggesting that these counties are less able to secure state attention either in aid or in centralization. We also see positive responses in counties with lower federal aid (FEDAID < \$42) but higher poverty (PCTPOV > 0.13%) and more urbanization (rural – urban codes 0, 1, 2, 6), suggesting higher need. The counties with lowest federal aid in Michigan and Pennsylvania (FEDAID < \$41) also show a positive effort response. It is also possible that state centralization can be structured in a way that does not meet the needs of the counties with the highest poverty rates as in Michigan and Pennsylvania where the highest poverty-rate (PCTPOV > 20%) counties show a positive response.

The neural-network model presents a mosaic of county-level responses to state centralization. The paths to positive state centralization elasticities are varied; yet for spatially differentiated subsets of counties, meaningful generalizations can be made without masking individual county differences. Some of the paths show counties with greater capacity or lower income inequality, which may permit higher effort in response to higher centralization. Other paths show higher need or lower aid. For these places, state centralization may be structured so that increases do not provide fiscal relief. Clearly, both the nature and the impact of state centralization differs within and across states.

Conclusion
Dencentralization raises the promise of increased efficiency and responsiveness, but also the challenge of redistribution and increasing diversity in local government revenue effort across space. Beginning in the late 1980s in the United States, the regional-level state emerged as a key player by increasing state aid to localities and increasing state centralization of fiscal responsibility. Studying the interaction between state-level policy and local government revenue effort in this early period of federal decentralization may offer lessons for understanding how decentralization may be implemented to exacerbate or ameliorate local government inequality. Our models suggest that state centralization and state aid have a major impact on effort, but the wide dispersion of local government responses suggests that state centralization and state aid may be substitutes for local effort in some counties and complements to local effort in others. It is important to know which process occurs where, and why.

A negative (substitution) response is easier to understand: state aid or state centralization of fiscal responsibility provides relief to local government revenue effort. This substitution effect is the most common response, and has the potential to ameliorate inequality in effort. For a significant proportion of counties, however, increased state centralization has a complementary, procyclical effect which can exacerbate inequality across localities. Using classification-tree techniques, we have begun to identify some of the decision rules which allow us to predict when the elasticity of state centralization on county effort might be positive. This could occur in places with higher capacity, in places with lower income inequality, and in places with higher need. Places with higher capacity can increase effort thereby enhancing local investment and local economic development prospects. However, for a different set of localities, higher effort appears to reflect higher need and suggests that state centralization may provide less fiscal relief.

Decentralization encourages state and local governments to become more developmentally focused. A substitution effect of state policy on effort should serve to narrow the differences between effort and capacity, leading to convergence in economic development prospects across counties. The caution, suggested by our analysis, is that decentralization could lead to vicious and virtuous cycles in the localities with positive, complementary, responses to state policy (see figure 5). For localities with higher capacity, there appears to be a virtuous cycle: state policy promotes greater effort, stimulating even more investment for future development. Medium-sized counties (suburbs) in New York, New Jersey, and Wisconsin are examples where this virtuous cycle may be at work. A more vicious cycle is found in the smallest rural places and higher poverty urban areas, where lower state aid and poorly targeted state centralization force higher effort despite limited capacity. Such expenditure choices may limit other economic development investments in these distressed areas, further frustrating their economic competitiveness.

State policies should not be expected to have a similar effect on every locality. The theoretical literature on geography, locality, and globalization calls for increased attention to the unique, contingent, responses of localities to decentralization (Dewees et al, 2003; LeGales, 1998; Lobao et al, 1999; MacLeod, 2001; Swyngedouw, 1997). Our analysis illustrates just how varied these locality effects might be. If decentralization leads to divergence in local government effort as suggested by the virtuous and vicious cycles illustrated below, then, in contrast to theoretical predictions of increased choice and efficiency, decentralization may lead to increased spatial inequality across local governments.

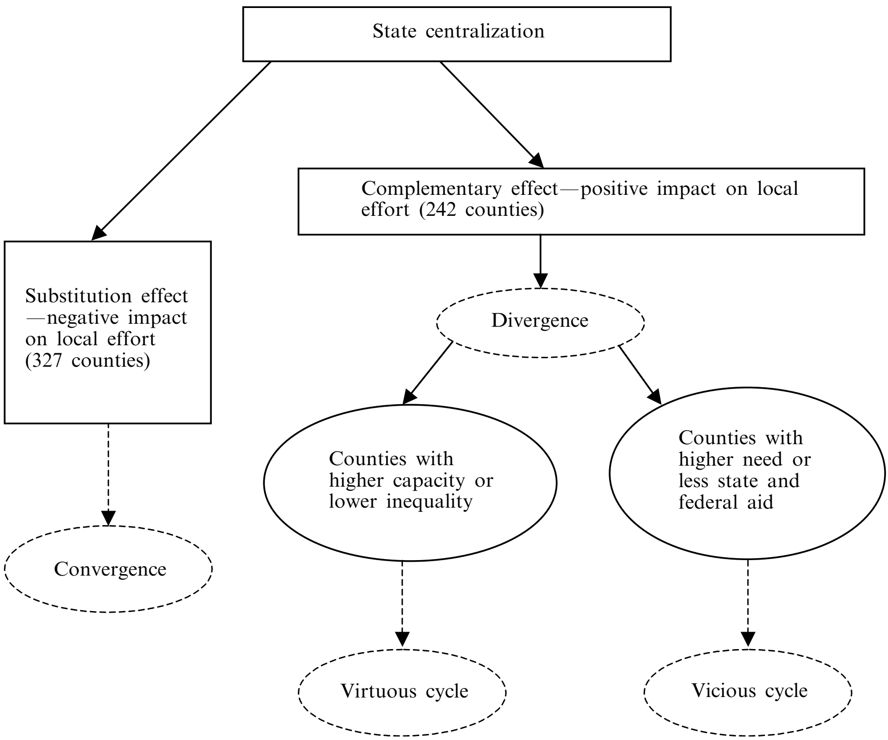


Figure 5. Impact of state centralization on local effort (N = 587 counties, 18 counties not classified). Neural network and classification model results.

Acknowledgements. This research was supported by a USDA National Research Initiative Grant #97-35401-4350. The authors thank Linda Lobao, Ann Tickameyer, and Greg Hooks, and the anonymous reviewers for their helpful comments.

References

Altshuler A, Morrill W, Wolman W, Mitchell F (Eds), 1999 *Governance and Opportunity in Metropolitan America* (National Academy Press, Washington, DC)

Anselin L, 1988 *Spatial Econometrics: Methods and Models* (Kluwer Academic, Boston, MA)

Bennett R, 1990, "Decentralization, intergovernmental relations and markets: towards a post-welfare agenda?", in *Decentralization, Local Governments and Markets: Towards a Post-Welfare Agenda* Ed. R Bennett (Clarendon Press, Oxford) pp 1–26

Bishop C M, 1995 *Neural Networks for Pattern Recognition* (Oxford University Press, Oxford)

Bradbury K, Ladd H, Perrault M, Reschousky A, Yinger J, 1984, "State aid to offset fiscal disparities across communities" *National Tax Journal* **37**(2) 151–170

Brenner N, 1999, "Globalisation as reterritorialisation: the re-scaling of urban governance in the European Union" *Urban Studies* **36** 431–451

Brieman L, Friedman J, Olshen R, Stone C, 1984 *Classification and Regression Trees* (Wadsworth, Belmont, CA)

Conlan T, 1998 *From New Federalism to Devolution: Twenty Five Years of Intergovernmental Reform* (Brookings Institution Press, Washington, DC)

Dewees S, Lobao L, Swanson L, 2003, "Local economic development in an age of devolution: the question of rural localities" *Rural Sociology* **68** 182–206

Donahue J D, 1997 *Disunited States* (Basic Books, New York)

Garson G D, 1991, "A comparison of neural network and expert system algorithms with common multivariate procedures for analysis of social science data" *Social Science Computer Review* **9** 299–434

Garson G D, 1998 *Neural Networks: An Introductory Guide to Social Scientists* (Sage, Thousand Oaks, CA)

Gold S, 1995 *The Fiscal Crisis of the States: Lessons for the Future* (Georgetown University Press, Washington, DC)

Grant D, Wallace M, 1994, "The political economy of manufacturing growth and decline across the American states, 1970–1985" *Social Forces* **73** 33–63

Hwang J T G, Ding A A, 1997, "Prediction intervals for artificial neural networks" *Journal of the American Statistical Association* **92** 748–757

Jessop R, 1994, 'Post-Fordism and the state", in *Post-Fordism: A Reader* Ed. A Amin (Blackwell, Cambridge, MA) pp 251–279

Johnson K M, Pelissero J P, Holien D B, Maly M T, 1995, "Local government fiscal burden in nonmetropolitan America" *Rural Sociology* **60** 381–398

Kastens T L, Featherstone A M, Biere A W, 1995, "A neural networks primer for agricultural economists" *Agricultural Finance Review* **55** 54–73

Katz M, 2001 *The Price of Citizenship: Redefining the American Welfare State* (Metropolitan Books, New York)

Ladd H F, Yinger J, 1989 *America's Ailing Cities: Fiscal Health and the Design of Urban Policy* (Johns Hopkins University Press, Baltimore, MD)

Ladd H F, Yinger J, 1994, "The case for equalizing aid" *National Tax Journal* **77** 211–224

Lawrence J, 1994 *Introduction to Neural Networks: Design, Theory, and Applications* (California Scientific Software, Nevada City, CA)

Lefevbre H, 1974 *La production de l'espace* (Anthropos, Paris); *The Production of Space* translated in 1991 by Donald Nicholson-Smith (Blackwell, Cambridge, MA)

LeGales P, 1998, 'Regulations and governance in European ciies" *International Journal of Urban and Regional Research* **47** 482–506

Lobao L, Rulli J, Brown L A, 1999, 'Macro-level theory and local-level inequality: industrial structure, institutional arrangements, and the political economy of redistribution, 1970 and 1990" *Annals of the Association of American Geographers* **89** 571–601

MacLeod G, 2001, "New regionalism reconsidered: globalization and the remaking of political economic space" *International Journal of Urban and Regional Research* **50** 804–829

Mencken C, 2000, "Federal spending and economic growth in Appalachian counties" *Rural Sociology* **65** 126–147

Meraviglia C, 1996, "Models of representation of social mobility and inequality systems: a neural network approach" *Quality and Quantity* **30** 231–252

Mollenkopf J H, 1983 *The Contested City* (Princeton University Press, Princeton, NJ)
Nathan R A, Lago J R, 1990, 'Intergovernmental fiscal roles and relations" *Annals of the American Academy of Political and Social Science* **509** 36 – 47
O'Connor J, 1973 *The Fiscal Crisis of the State* (St Martin's Press, New York)
Orfield M, 1997 *Metropolitics* (Brookings Institution, Washington, DC)
Peterson P, 1981 *City Limits* (University of Chicago Press, Chicago, IL)
Powers E, 1999, "Block granting welfare: fiscal impact on the states", OP 23, The Urban Institute, Washington, DC
Prud'homme R, 1995, "The dangers of decentralization" *The World Bank Research Observer* **10** 201 – 214
Reeder R, Jansen A, 1995, "Rural government—poor counties, 1962 – 1987", RR 88, Rural Development, US Department of Agriculture, Washington, DC
Schneider M, 1989 *The Competitive City: The Political Economy of Suburbia* (University of Pittsburgh Press, Pittsburgh, PA)
Stine W F, 1994, 'Is local government revenue response to federal aid symmetrical? Evidence from Pennsylvania county governments in an era of entrenchment" *National Tax Journal* **47** 799 – 816
Swyngedouw E, 1997, "Neither global nor local: 'glocalization' and the politics of scale", in *Spaces of Globalization* Ed. K Cox (Guilford Press, New York) pp 137 – 166
Tiebout C M, 1956, "A pure theory of local expenditures" *Journal of Political Economy* **64** 416 – 424
Tomek W G, 1985, "Limits on price analysis" *American Journal of Agricultural Economics* **67** 905 – 915
Veaux R D, Schumi J, Schweinsberg J, Ungar L H, 1998, "Prediction intervals for neural networks via nonlinear regression" *Technometrics* **40** 273 – 282
Wang D Q, Chukova S, Lai C C, 2004, "On the relationship between regression analysis and mathematical programming" *Journal of Applied Mathematics and Decision Sciences* **8**(2) 131 – 140
Warner M E, 2001, "State policy under devolution: redistribution and centralization" *National Tax Journal* **54** 541 – 556
Wier T G, Phoha V V, 2002, "Neural networks and regional science modeling: a survey of techniques for complex spatial analysis" *The Review of Regional Studies* **32** 309 – 324
Woelfel J, 1993, "Artificial neural networks in policy research: a current assessment" *Journal of Communication* **43** 63 – 80
Wyckoff P, 1985, "Revenue sharing and local public expenditures: old questions, new answers" *Economic Review Federal Reserve Bank of Cleveland* second quarter, 13 – 28
Zampelli E M, 1986, "Resource fungibility, the flypaper effect and the expenditure impact of grants-in-aid" *The Review of Economics and Statistics* **68** 33 – 41
Ziliak S T, McCloskey D N, 2004, "Size matters: the standard error of regressions in the *American Economic Review*" *Journal of Socio-Economic* **33** (special issue on Statistical Significance) 527 – 546

© 2005 a Pion publication printed in Great Britain