Local Public Goods and Household Sorting in Metropolitan Areas

Holger Sieg

Carnegie Mellon University and NBER

May 29, 2007

Motivation

- Economic mobility coupled with the choice provided by a diverse set of local communities makes the analysis of local public economics interesting and different from that of the federal government.
- State and local governments comprise about 40 percent of the total public sector in the United States and provide some of the most important public goods such as education, protection from crime, income maintenance, and environmental quality.
- Local governments can be viewed as a laboratory for evaluating new public policies such as welfare reform, vouchers, and school reform.
- Local governments are fascinating objects to study!

A Historical Perspective

- Samuelson (1954, 1955) provided conditions for allocative efficiency for public good provision. In general, we have no reason to believe that public goods will be provided efficiently.
- Tiebout (1956) conjectured that competition among publicly elected governments for mobile households may yield an efficient provision of local public goods.
- Barr and Davis (1966) provided some of the foundations for a political theory of local expenditures.
- Much of the research in local public economics has attempted to formalize and test these ideas.

We will focus on answering the following questions in this lecture:

- Are simple general equilibrium models of interjurisdictional competition consistent with the main stylized facts observed in the data?
- How can we estimate household preferences for local public goods and amenities? What do we learn about household sorting?
- What can we say about the mechanism of local public good provision? How well do simple voting models explain the observed policy outcomes?
- Can we use these models to estimate the willingness to pay for improvements in local public goods and amenities?
- How can we estimate housing production functions?

Empirical Approach

The empirical approach can be best characterized as theory-based micro-econometric modeling. That means we are interested in evaluating the empirical implications of microeconomic theory. As a consequence, the empirical analysis and the underlying economic model must be internally consistent.

Some Stylized Facts

- Community boundaries rarely change. School district boundaries change more often.
- There is a large amount of housing price variation across communities.
- There is a large variation in observed expenditures and other community specific amenities across communities.
- There is a large amount of turn-over in housing markets.
- There is a significant amount of income and housing expenditure heterogeneity within communities and across communities.

A Locational Equilibrium Model

- Metropolitan Area which consists of J communities each of them has fixed boundaries. Households behave as price takers.
- Continuum of households which differ by income, y, and tastes for public goods, α.
- Households have preferences defined over a local public good, g, a local housing good, h, and a composite private good, b.
- Levels of public good provision are determined by majority rule.
- The budgets of the communities must be balanced.
- Mobility between communities is costless.

Household Preferences

Households maximize utility with respect to their budget constraint:

$$\max_{(h,b)} U(\alpha, g, h, b)$$
$$s.t. \ \ (1+t)\, p^h \, h \ = \ y \ - \ b$$

Alternatively, we can represent the preferences by the corresponding indirect utility function.

$$V(\alpha, g, p, y) \ = \ U(\alpha, g, h(p, y, \alpha), y - ph(p, y, \alpha))$$

where $p = (1+t)p^h$.

The Single-Crossing Properties

Consider the slope of an "indirect indifference curve" in the (g, p)-plane:

$$\begin{aligned} M(\alpha, g, p, y) &= \frac{dp}{dg}\Big|_{V=\bar{V}} \\ &= -\frac{\partial V(\alpha, g, p, y)/\partial g}{\partial V(\alpha, g, p, y)/\partial p} \end{aligned}$$

If $M(\cdot)$ is monotonic in y for given α, then indifference curves in the (g, p)-plane satisfy the "single-crossing" property. Likewise, monotonicity of $M(\cdot)$ in α provides single-crossing for given y.

The Decision Problem of a Household

There are no mobility costs, and hence households maximize:

$$\max_j \; V(\alpha, g_j, p_j, y)$$

Define the set C_j to be the set of households living in community j:

$$C_j \;=\; \{(\alpha, y) | V(\alpha, g_j, p_j, y) \geq \max_{i \neq j} \; V(\alpha, g_i, p_i, y)\}$$

Moreover, define the boundary indifference loci $\alpha_j(y)$ as follows:

$$V(\alpha_j(y), g_j, p_j, y) \;=\; V(\alpha_j(y), g_{j+1}, p_{j+1}, y)$$

Community Size and Housing Demand

A measure of the community size is given by:

$$P(C_j) = \int_{C_j} f(\alpha, y)\, dy\, d\alpha$$

Similarly, aggregate housing demand is defined as:

$$H_j^d = \int_{C_j} h(p_j, \alpha, y) \; f(\alpha, y)\, dy\, d\alpha$$

Community Budget Constraint

We assume that the budget of community j must be balanced. This implies that:

$$t_j \, p_j^h \int_{C_j} h(p_j, \alpha, y) \;\; f(\alpha, y) \, dy \, d\alpha \; / \; P(C_j) \; = \; c(g_j)$$

where $c(g)$ is the cost per household of providing g.

Housing Supply

We assume that housing is owned by absentee landlords and that aggregate housing supply in community j depends on the net price of housing p_j^h and a measure of the land area of community j denoted l_j. Hence, we have:

$$H_j^s \;=\; H(l_j, p_j^h)$$

The Government-Services Possibility Frontier

The GPF is defined as the locus of (g_j, p_j) such that housing markets are in equilibrium:

$$F_j(g_j, p_j, t_j) \;=\; H_j^d(g_j, p_j, t_j) \;-\; H_j^s(p_j, t_j) \;=\; 0$$

and the community budget is balanced:

$$G_j(g_j, p_j, t_j) \;=\; c(g_j) \;-\; p_j \, \frac{t_j}{1+t_j} \, H_j^d(g_j, p_j, t_j) / \, P(C_j) \;=\; 0$$

given the perceived migration effects.

Majority Rule

Consider a point (g^*, p^*) on the GPF. We say (g^*, p^*) is a majority rule equilibrium, if there does not exist another point on the GPF $(\hat{g}, \hat{p})$ which would beat (g^*, p^*) in a majority vote:

$$
\begin{aligned}
& P\{(\alpha, y) \in C_j | V(\alpha, g^*, p^*, y) \geq V(\alpha, \hat{g}, \hat{p}, y)\} \\
\geq\ & P\{(\alpha, y) \in C_j | V(\alpha, g^*, p^*, y) < V(\alpha, \hat{g}, \hat{p}, y)\}
\end{aligned}
$$

We assume that voters do not behave strategically, i.e. we assume sincere voting.

Preferences for Local Public Goods and Taxes

A voter's preferred level of g is then obtained by maximizing the indirect utility function subject to the feasibility constraint given by the GPF:

$$\max_{p,g} \quad V(\alpha, g, p, y)$$
$$\text{s.t.} \quad p = p(g)$$

where $p(g)$ characterizes th GPF. The FOC is given by:

$$\frac{dp}{dg}|_{GPF} = -\frac{V_g}{V_p} = M(\alpha, g, p, y)$$

Decisive or Pivotal Voters

Hence there will be a set of voters (α, y) which will have the same preferred level for (p, g).

Consider a point (p_j^*, g_j^*) which is on the GPF of community j and the preferred point for a set of households denoted by $\tilde{\alpha}_j(y)$.

For any level of income y, the single crossing properties imply that households with higher (lower) values of α will have higher (lower) demands for local public goods.

As a consequence $\tilde{\alpha}_j(y)$ characterizes the set of pivotal voters, if the following condition holds:

$$\int_0^\infty \int_{\alpha_{j-1}(y)}^{\tilde{\alpha}_j(y)} f(\alpha, y) \, d\alpha \, dy = \frac{1}{2} P(C_j)$$

The Myopic Voting Model

According to this hypothesis, voters treat the population boundaries of the communities as fixed. If voters also treat the housing demand as fixed when voting, then we obtain the simple myopic voting model:

$$\frac{dp_j}{dg_j}|_{MV} = \frac{c'(g_j)}{H_j/P(C_j)}$$

This is equivalent to the assumption that when voting, each resident of the community takes the net-of-tax price of housing, community population, and the aggregate housing demand as fixed

Existence of Equilibrium

Existence of equilibrium can be shown under a number of regularity conditions. The basic idea is to break down the proof into two steps:

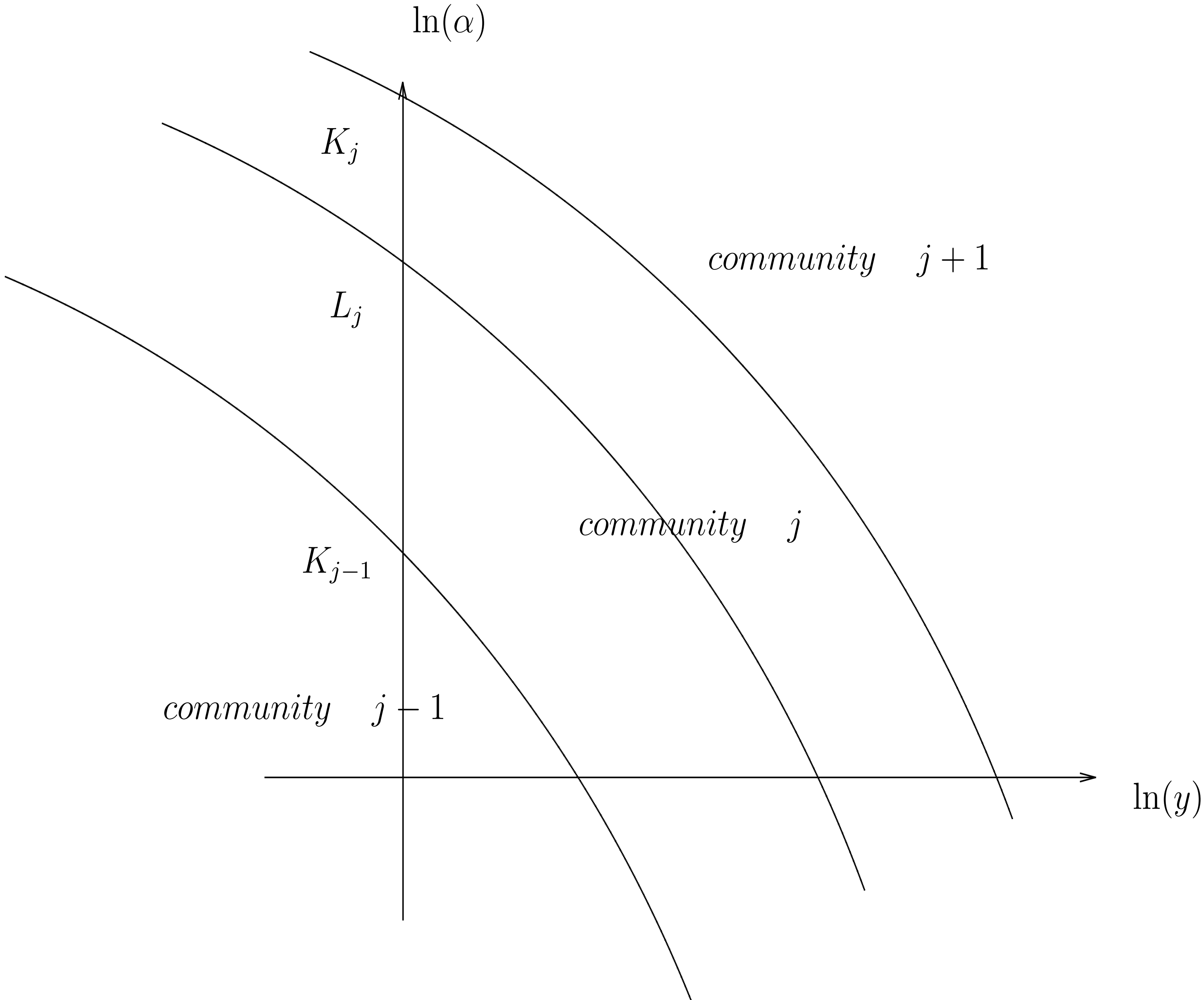
1. We impose assumptions which guarantee that an internal equilibrium exists for a community with given population and show that this internal equilibrium is unique.
2. One sets up a mapping on the space characterizing community boundaries and shows that a fixed point of that algorithm is an equilibrium.

For details, see Epple, Filimon, Romer (1984,93) and Calabrese, Epple, Romer and Sieg (2006).

Properties of the Equilibrium

For such an allocation to be a locational equilibrium , there must be an ordering of community pairs, $\{(g_1, p_1), ..., (g_J, p_J)\}$, such that:

1. **Boundary Indifference:** The set of "border" individuals are indifferent between the two communities: $I_j = \{(\alpha, y) \mid V(\alpha, g_j, p_j, y) = V(\alpha, g_{j+1}, p_{j+1}, y)\}$
2. **Stratification:** Let $y_j(\alpha)$ be the implicit function defined by the equation above. Then, for each α, the residents of community j consist of those with income, y, given by: $y_{j-1}(\alpha) < y < y_j(\alpha)$.
3. **Increasing Bundles:** Consider two communities i and j such that $p_i > p_j$. Then $g_i > g_j$ if and only if $y_i(\alpha) > y_j(\alpha)$.



$\ln(\alpha)$
K_j
L_j
K_{j-1}
community $j+1$
community j
community $j-1$
$\ln(y)$

Computation of Equilibrium

- There are no analytical solutions for equilibrium. But we can compute equilibria numerically.
- You need to parametrize the model and pick numerical values for each parameter.
- An equilibrium is characterized by a vector $(t_j, p_j, g_j)_{j=1}^{J}$.
- To computing an equilibrium we need to solve a system of $J \times 3$ nonlinear equations: budget constraints, housing market equilibria, voting conditions.
- Check SOC after you find a solution to the system of equations.

A Parameterization of the Model

- Let the joint density of $\ln(\alpha)$ and $\ln(y)$ be bivariate normal with correlation λ.

- Assume that the indirect utility function is given by:

$$V(g, p, y, \alpha) \; = \; \{\alpha g^{\rho} \; + \; [e^{\frac{y^{1-\nu}-1}{1-\nu}} \; e^{-\frac{Bp^{\eta+1}-1}{1+\eta}}]^{\rho}\}^{\frac{1}{\rho}}$$

- Housing Supply: $H_j^s \; = \; l_j \; [p^h]^{\tau}$.

- Costs: $c(g) \; = c_0 \; + \; c_1 \; g$.

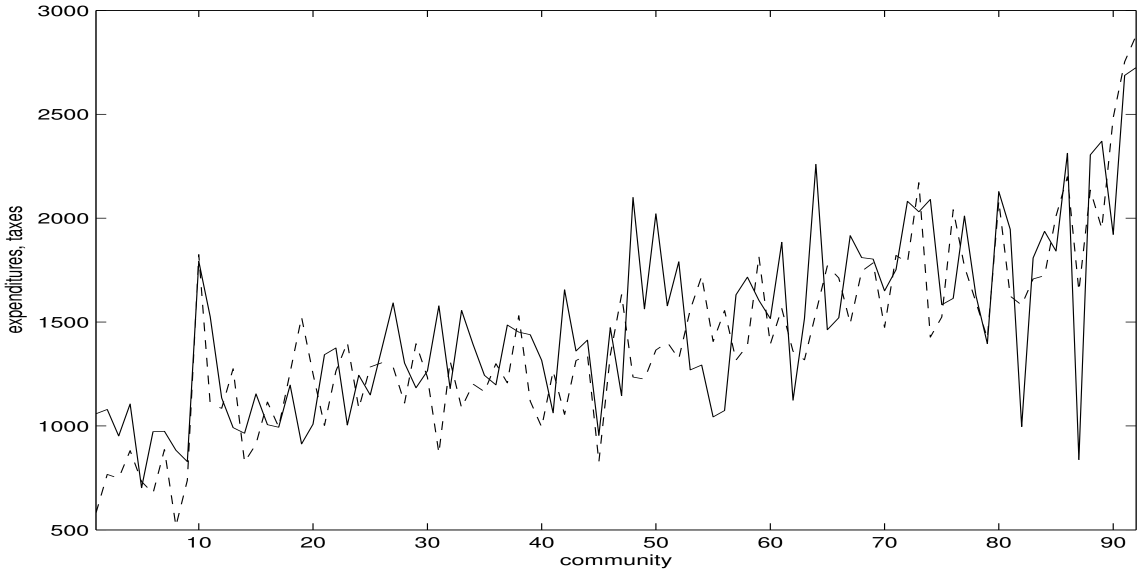
- Parameters: $\mu_{\ln(y)}, \mu_{\ln(\alpha)}, \sigma_{\ln(y)}, \sigma_{\ln(\alpha)}, \lambda, \rho, \nu, \eta, B$ and c_0, c_1, τ.

Table 1: Descriptive Statistics of the Sample

Variable	Mean	Std. Deviation
Population size	30036	59719
Number of households	10769	23335
Mean income[a]	27402	8024
Median income[a]	24108	6481
Education expenditure[a]	1479	435
Property tax rate[b]	0.031	0.009
Median property value[a]	64923	21515
Median gross rent[a]	314.35	58.22
Fraction of renters	0.28	0.16

Table 1 provides summary statistics for the 92 communities in the Boston Metropolitan Area in 1980. The notation [a] indicates that the value is per household. The notation [b] indicates that the variable is measured per dollar of value.

Figure 1: Residential Property Tax Revenue and Educational Expenditure per Household



Notation: — expenditures per household on education, - - property taxes per household

Communities are arrayed in order of increasing median household income.

Testing Predictions of the Model I:

The model predicts *the distribution of households by income* among the set of communities. We can test the predictions of the model by matching the predicted marginal distribution of income in each community, $f_j(y)$, to the distribution reported in the U.S. Census.

The Boundary Indifference Condition

The boundary indifference condition for community j versus community $j+1$ can be written as:

$$\ln(\alpha) \; - \; \rho \left(\frac{y^{1-\nu} - 1}{1 - \nu} \right) \; = \; \ln \left(\frac{Q_{j+1} - Q_j}{g_j^{\rho} - g_{j+1}^{\rho}} \right) \; \equiv \; K_j$$

where

$$Q_j = e^{-\frac{\rho}{1+\eta} \left(B p_j^{\eta+1} - 1 \right)}$$

Community Sizes and Share Inversion

The Size of a Community is given by:

$$P(C_j) \;=\; \int_{-\infty}^{\infty} \int_{K_{j-1}+\rho\frac{y^{1-\nu}-1}{1-\nu}}^{K_j+\rho\frac{y^{1-\nu}-1}{1-\nu}} \; f(\ln(\alpha)\;\ln(y)) \;\; d\ln(\alpha)\, d\ln(y)$$

Given the parameterization of the model, we can (recursively) express the community specific intercepts, $(K_0, ..., K_J)$, as functions of the community sizes, $(P(C_1), ..., P(C_J))$, and the parameters of the model.

Income Quantiles

The Quantiles of the Income Distribution are implicitly defined by: equation:

$$\int_{-\infty}^{\ln(\zeta_j(q))} \int_{K_{j-1}+\rho\frac{y^{1-\nu}-1}{1-\nu}}^{K_j+\rho\frac{y^{1-\nu}-1}{1-\nu}} f(\ln(\alpha)\ , \ln(y)) \ \ d\ln(\alpha)\, d\ln(y) \ = \ q\ P(C_j)$$

Given the parameterization of the model, the income distributions of the J communities are completely specified by the parameters of the distribution function, $(\mu_y, \mu_\alpha, \lambda, \sigma_y, \sigma_\alpha)$, the slope coefficient, ρ, the curvature parameter ν, and the community specific intercepts, $(K_0, ..., K_J)$.

A Minimum Distance Estimator

For every community we have:

$$e_j^N(\theta_1) = \begin{Bmatrix} \ln(\zeta_j(0.25,\ \theta_1)) \ - \ \ln(\zeta_j^N(0.25)) \\ \ln(\zeta_j(0.50,\ \theta_1)) \ - \ \ln(\zeta_j^N(0.50)) \\ \ln(\zeta_j(0.75,\ \theta_1)) \ - \ \ln(\zeta_j^N(0.75)) \end{Bmatrix}$$

Stacking the orthogonality conditions above, yields the following MDE:

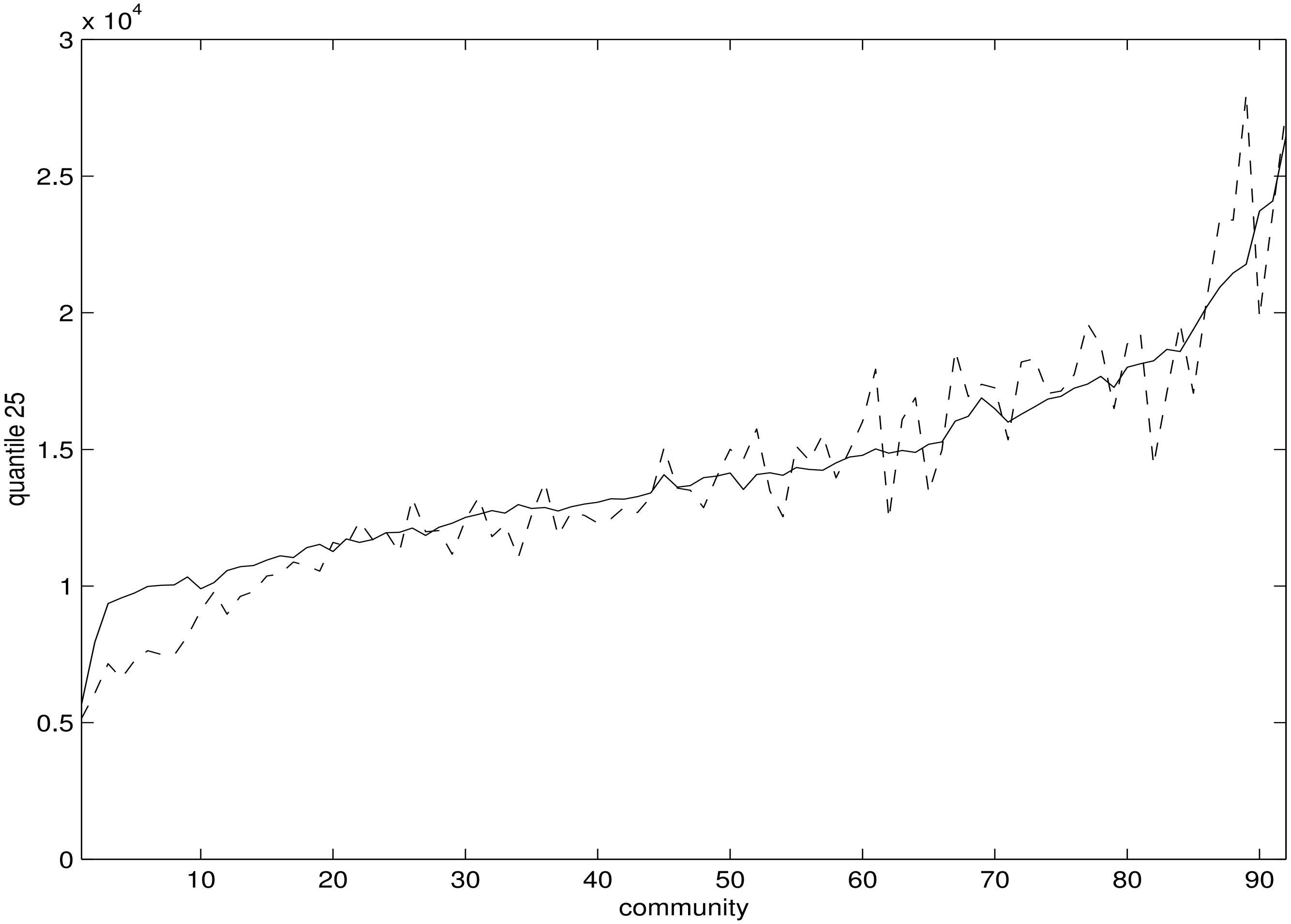
$$\begin{aligned} \theta_1^N \ &= \ \arg\min_{\theta_1 \in \Theta_1} \ \{ \ e_N(\theta_1)' \, A_N \, e_N(\theta_1) \ \} \\ s.t. \ K_j \ &= \ K_j(K_{j-1}, P(C_j) \mid \rho, \mu_y, \sigma_y, \mu_\alpha, \sigma_\alpha, \lambda, \nu) \qquad j = 1, ..., J-1 \end{aligned}$$

where θ_1 is the unknown parameter vector, A_N is the weighting matrix.

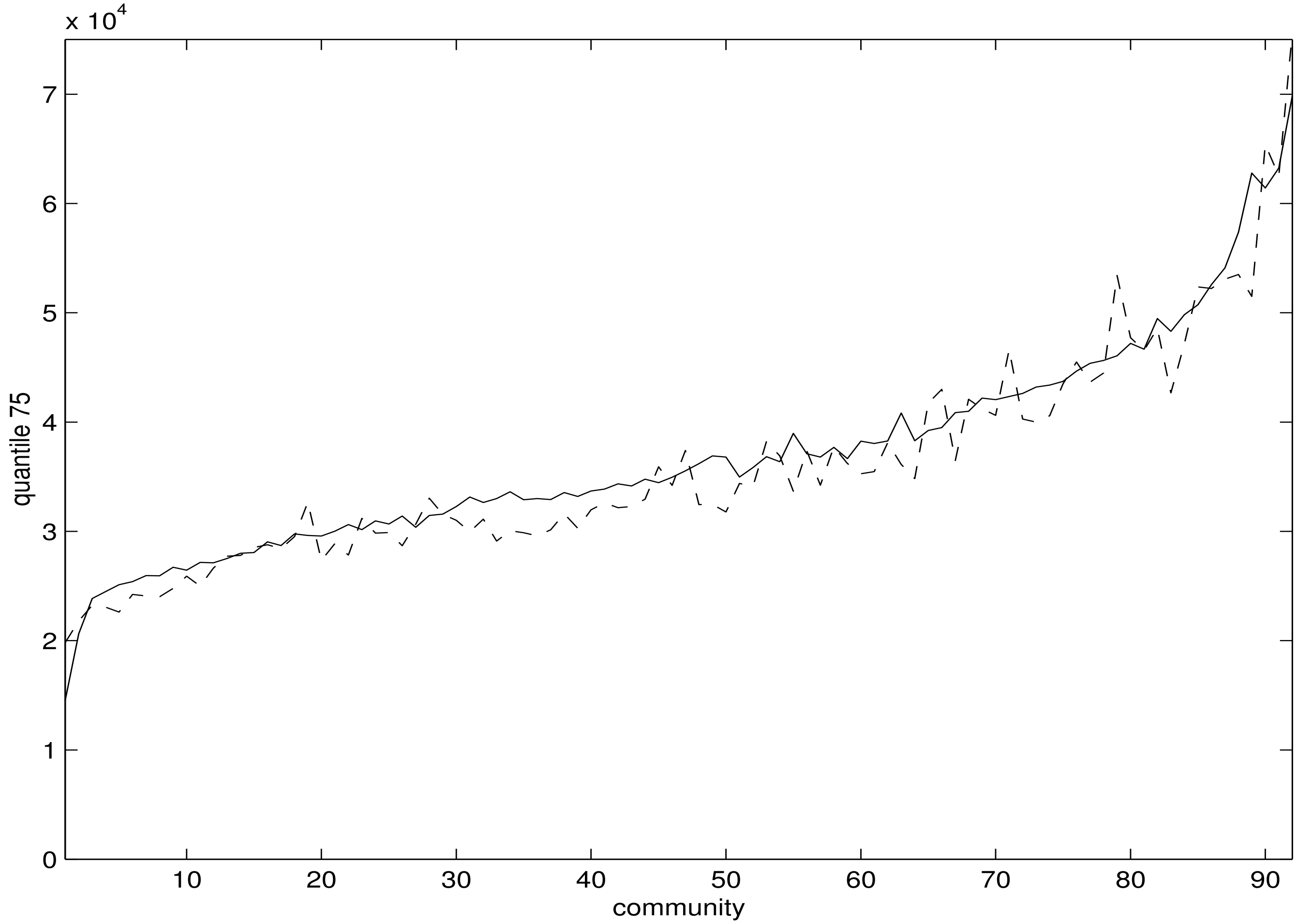
Table 2: Estimated Parameters I

parameters	estimates
$\mu_{\ln(y)}$	9.790
	(0.002)
$\sigma_{\ln(y)}$	0.755
	(0.004)
λ	-0.019
	(0.031)
$\rho/\sigma_{\ln(\alpha)}$	-0.283
	(0.013)
ν	0.938
	(0.026)
function value	0.0368
degrees of freedom	271

NOTE: Estimated standard errors are given in parentheses.



x 10^4
3
2.5
2
1.5
1
0.5
0
quantile 25
10
20
30
40
50
60
70
80
90
community



x 10^4
7
6
5
4
3
2
1
0
quantile 75
10
20
30
40
50
60
70
80
90
community

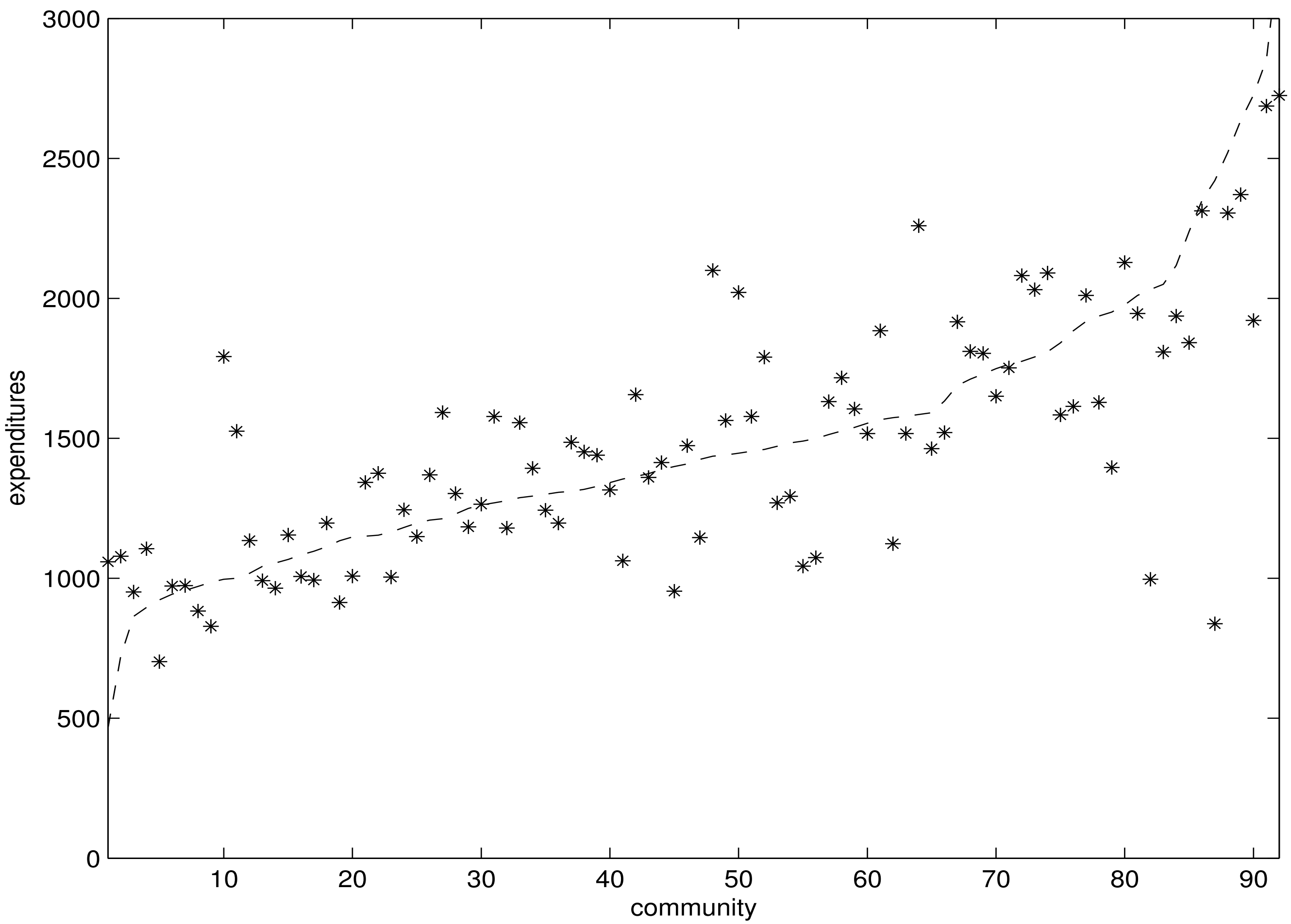
Testing Predictions of the Model II

The model predicts a distribution of *tax rates, expenditures on education, and mean housing expenditures* among the communities observed in the metro area. We test whether the model can fit these observed distributions. We can compute equilibria and match the equilibrium values to the observed ones using a ML estimator which is based on the following three equations:

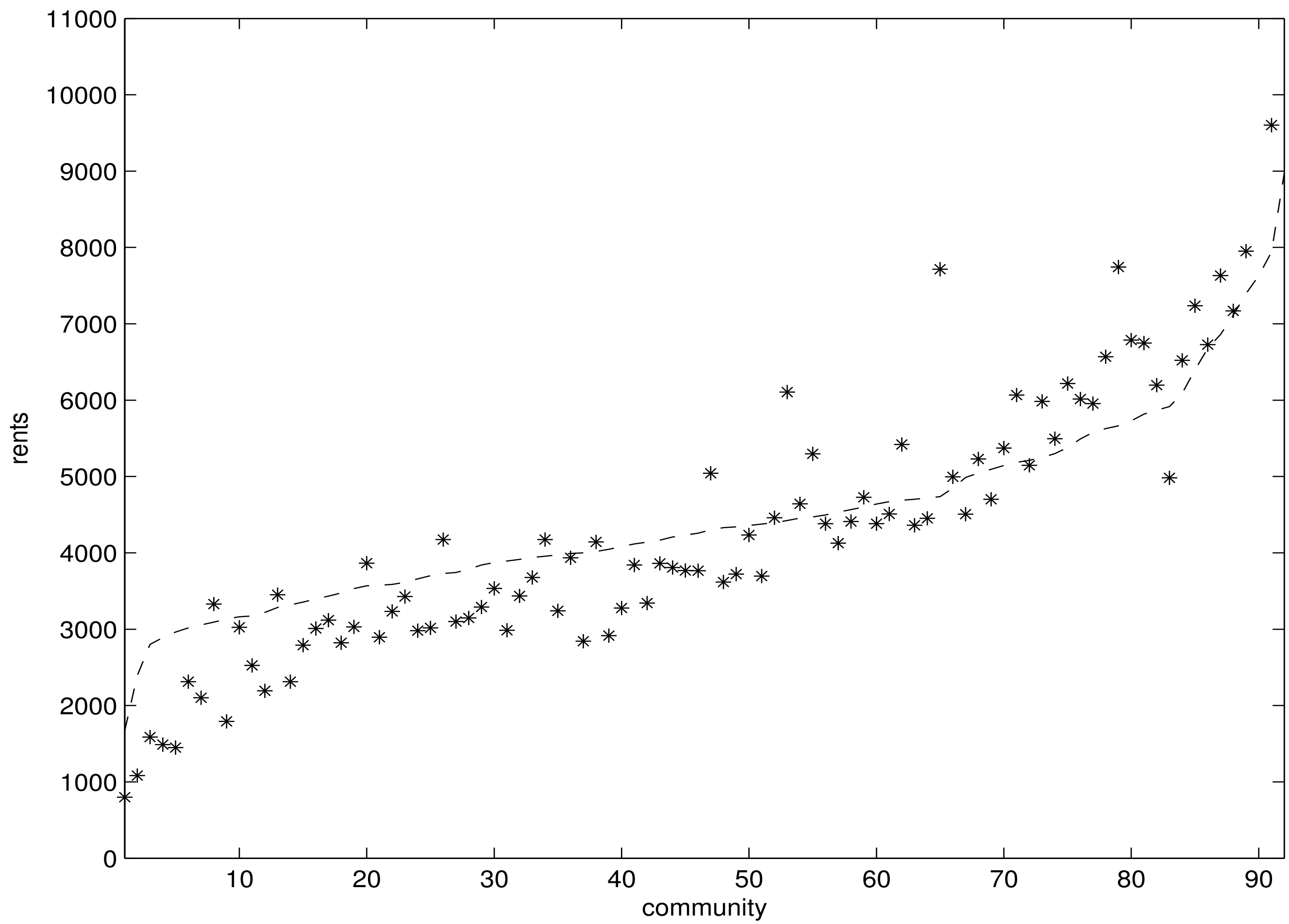
$$\begin{aligned} \tau_j &= \tau_j(\theta_2) + \epsilon_j^\tau \\ g_j &= g_j(\theta_2) + \epsilon_j^g \\ e_j &= e_j(\theta_2) + \epsilon_j^e \end{aligned}$$

Table 3: Estimation Results

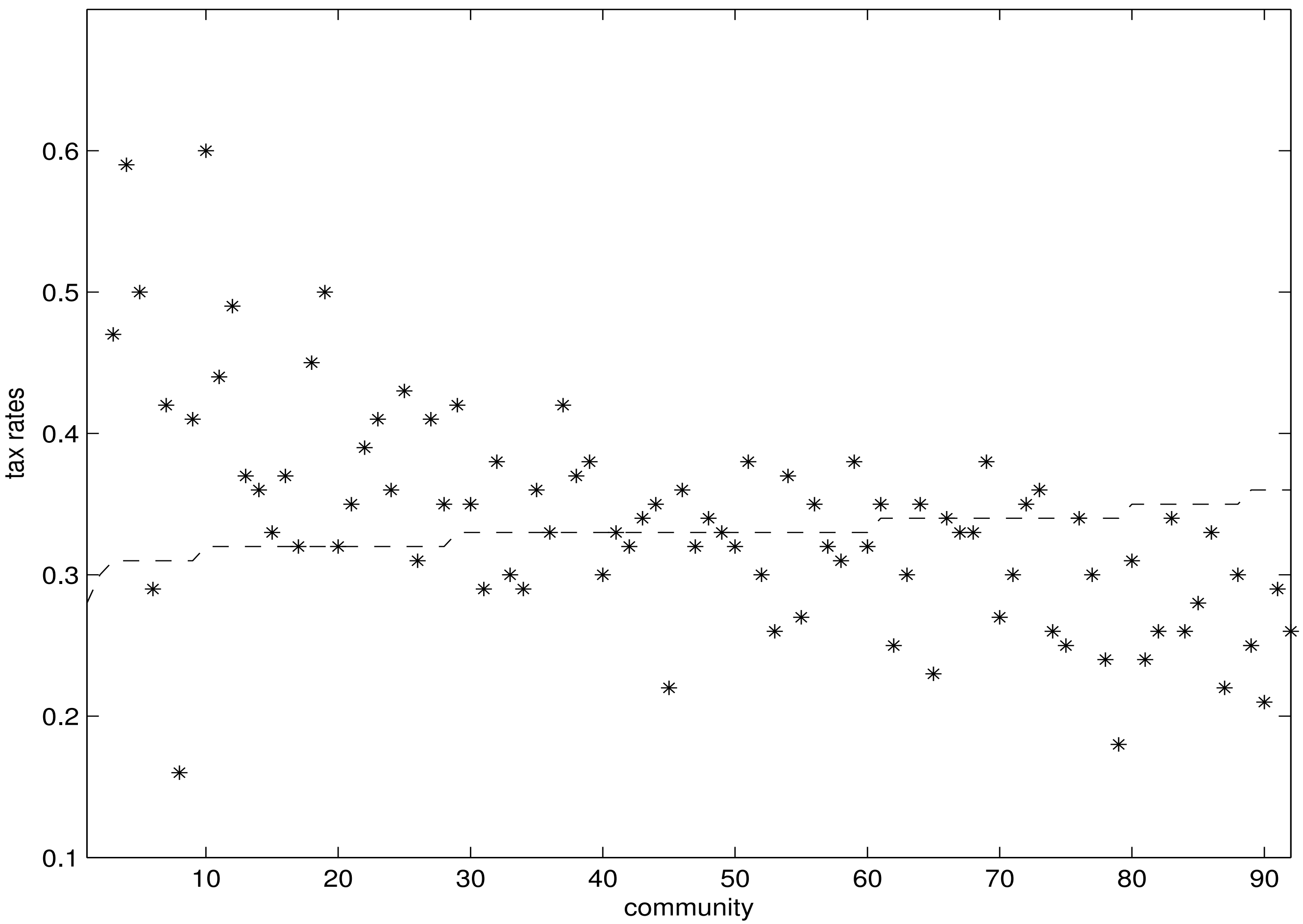
	I	II
parameters	baseline model	extended model
$\mu_{\ln(\alpha)}$	-2.622	-2.643
	(0.021)	(0.017)
$\sigma_{\ln(\alpha)}$	0.1	0.1
	——	——
B	0.325	0.175
	(0.006)	(0.007)
ϕ	0.0	2.623
	——	(0.147)
likelihood function	-1360.92	-996.51
correlation expenditures	0.727	0.727
correlation rents	0.940	0.939
correlation taxes	-0.672	0.747



3000
2500
2000
1500
1000
500
0
expenditures
10
20
30
40
50
60
70
80
90
community



11000
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
rents
10
20
30
40
50
60
70
80
90
community



tax rates
0.1
0.2
0.3
0.4
0.5
0.6
10
20
30
40
50
60
70
80
90
community

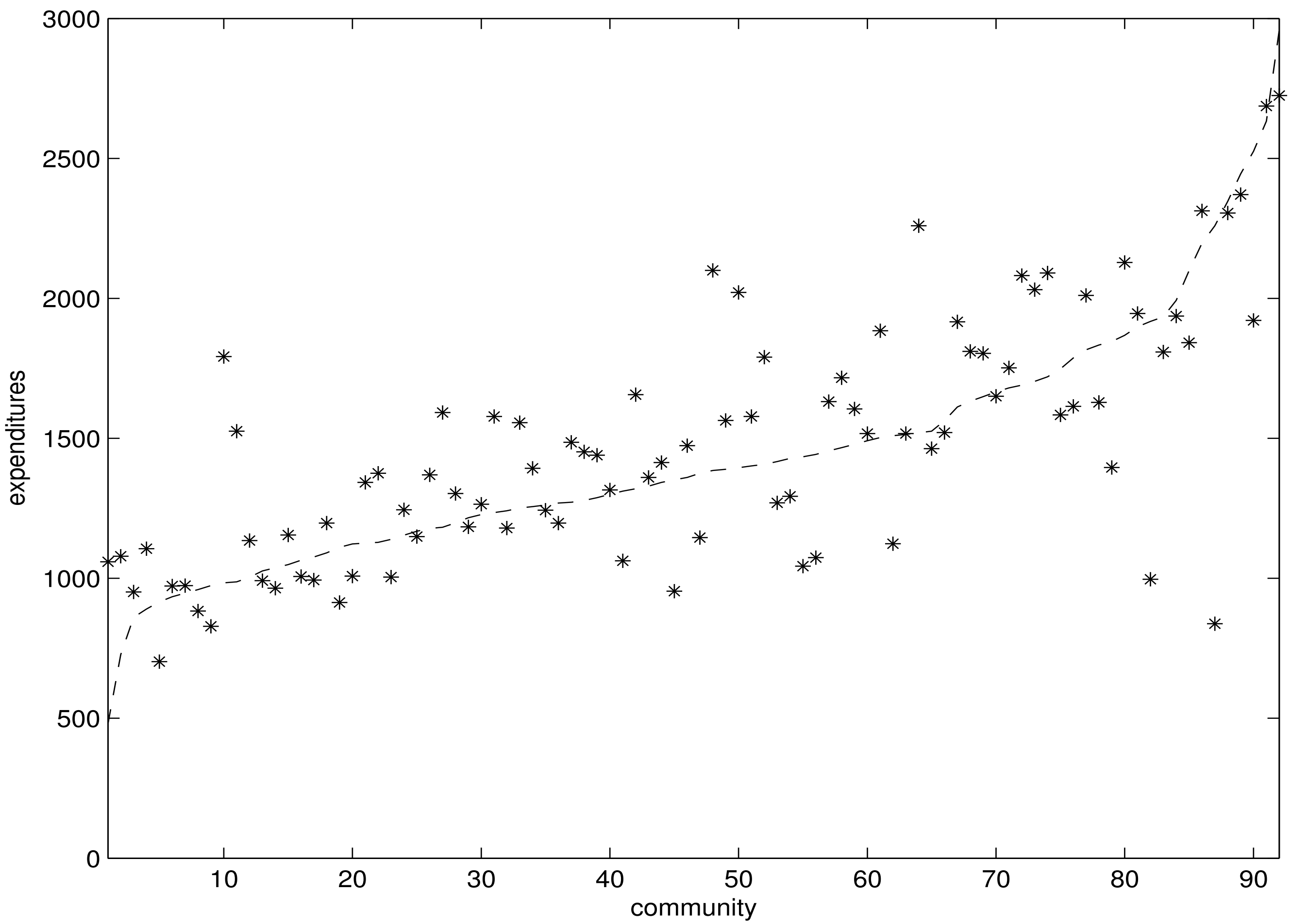
An Extended Model with Peer Effects

The quality of local public good provision denoted by q, depends on expenditures per household, g, and a measure of peer quality, denoted by $\bar{y}$.

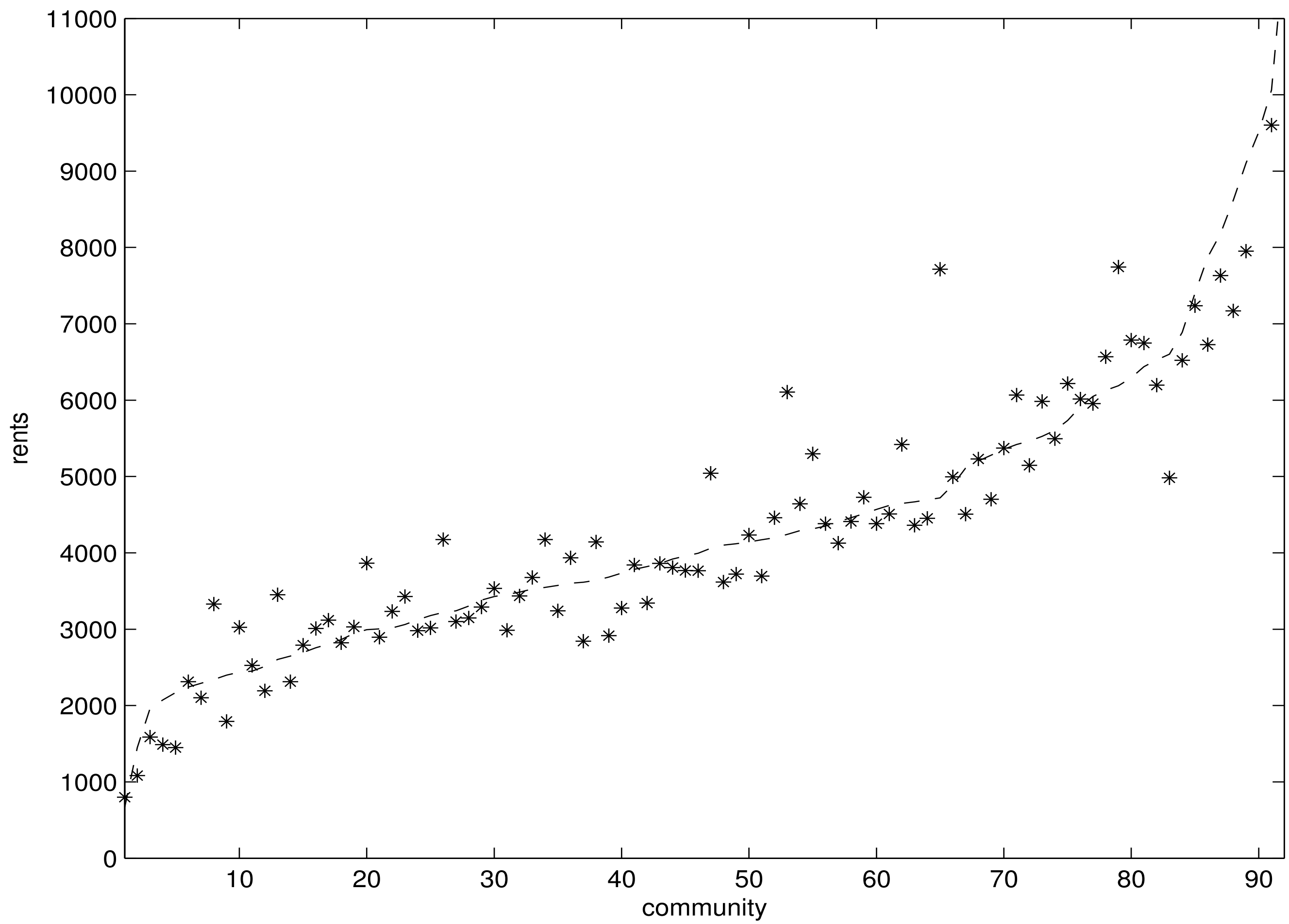
$$q_j = g_j \left(\frac{\bar{y}_j}{\bar{y}}\right)^{\phi}$$

where peer quality can be measured by the mean income in a community:

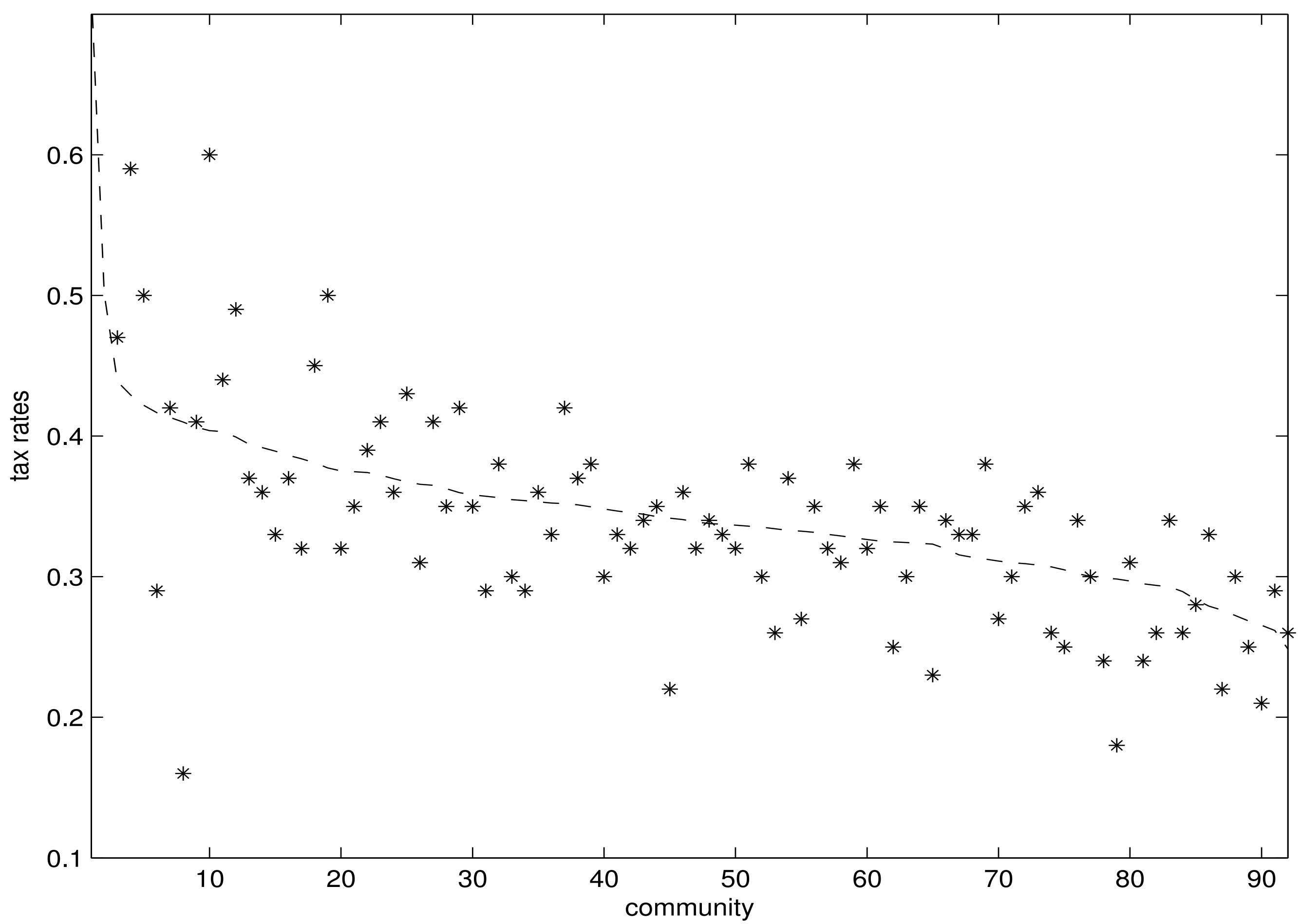
$$\bar{y}_j = \int_{C_j} y\, f(\alpha, y)\, dy\, d\alpha \;/\; n_j$$



3000
2500
2000
1500
1000
500
0
expenditures
10
20
30
40
50
60
70
80
90
community



11000
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
rents
10
20
30
40
50
60
70
80
90
community


0.6
0.5
0.4
0.3
0.2
0.1
tax rates
10
20
30
40
50
60
70
80
90
community

Some Remarks

Identification does not rely on functional form assumptions on the joint distribution of tastes and income (Epple, Peress, and Sieg, 2006).

Once we have estimated the model, we can use the model to conduct some counterfactual policy analysis. In Sieg, Smith, Banzhaf, and Walsh (2004) we estimate the willingness to pay for various scenarios that have been proposed by the EPA for reducing air pollution in Los Angeles.

One can also relax a number of assumptions and allow for more observed and unobserved heterogeneity among households. We are currently working on project that tries to evaluate housing market subsidies such as Section 8 vouchers.

Some Related Empirical Studies

Bajari and Kahn (2004) analyze racial sorting within a structural residential choice model.

Bayer, McMillan, and Reuben (2005) estimate a model of household choice and control for access to employment opportunities.

Bayer, Ferreira, and McMillan (2005) provide a general framework for measuring preferences for local amenities.

Epple, Romano, and Sieg (2006) estimate a model of competition in higher education.

Ferreyra (2005) estimates a model community choice to study the impact of vouchers in local education.

Ferreira (2005) studies the impact of property tax limitations (Proposition 13) in California on household sorting.

Gordon and Knight (2006) estimate a model of school district consolidation.

Hastings, Kane, and Staiger (2005a,2005b) estimate a model of school choice and analyze the impact of school choice on student achievement.

Ioannides and Schmidheiny (2005) are working on a model that tries to capture sorting patterns in Boston.

Nechyba (1997, 2000) studies residential choice, school competition and mobility within a calibrated equilibrium model.

Orthalo-Magne and Rady (2005) analyze household sorting and risk sharing in housing markets.

Timmins (2004) estimates a model of sorting among Brazilian farmers.

Walsh (2005) analyzes open space policies in North Carolina.