Environmental and Urban Economics

Residential Sorting

Sophie Mathes
FGV EPGE
2nd Quarter 2023

Residential Sorting

Samuelson ReStat 1954 "A pure theory of public expenditures"

What is the optimal level of expenditures on public goods?

Establish the standard solution to optimal expenditures on public goods:

- Provision is socially optimal when marginal cost equal marginal social benefits

Market outcomes do not achieve optimal levels.

Tiebout JPE 1956

A pure theory of local expenditures

One of JPE's greatest hits - 24,000 citations as of May 2023

Practical problem of the "optimal provision of public goods" solution: The government does not know the preferences of the individuals

In the U.S. many public goods are determined at local level, not federal (police, fire fighters, schools, hospitals)

1954: Total amount spent by local governments exceeded federal expenditures on goods+services (without military spending)

Contribution: Model to obtain optimal public expenditures on the local (versus federal) level.

Tiebout JPE 1956

Are local expenditure levels optimal?

When an individual chooses a location - what factors affect location choice?

People with children might value high level of expenditures on schools, other people might prefer expenditures on parks

Availability+quality of beaches, parks, public safety/police, roads will enter decision-making process

On the local level, individuals are able to choose locations according to their preferences

On the federal level: individual preferences are given

$\Rightarrow$ Recognize that individual location choice reveals preferences for public good provision

Model local government

m communities, n public goods to be provided in each community. Assume

- No moving cost, individuals choose location with preferred bundle of public goods
- Full information
- Large number of locations to choose from
- Fixed income, independent of location
- Locations are "isolated" / public good levels in one location do not affect public good levels in another location
- Define optimal community size
  - Given bundle of public goods can be provided at the lowest average cost per unit
  - Necessary to have communities of finite sizes in equilibrium
- Communities with population size < optimal want to attract new residents

Tiebout JPE 1956

Equilibrium

An allocation of individuals across communities is an equilibrium if no individual wishes to move

Individuals reveal their preferences for public goods by choosing the location that most closely matches their preferences compared to all available alternatives

Tiebout JPE 1956

Papers inspired by Tiebout have built increasingly comprehensive models of residential sorting, political economy/voting, taxation, and provision of public goods.

Integrating all these components is a big challenge. Usually you need to impose very strong assumptions on some parts to be able to analyze a few variations in other parts.

Typical problems in finding equilibria, and in ensuring existence of equilibria

- Endogenous size of communities,
- endogenous number of communities,
- defining political process for voting,
- prediction of voting outcomes,
- optimal allocation of residents across communities.

Epple Filimon Romer RSUE 1993

Existence of voting and housing equilibrium in a system of communities with property taxes

Continuum of heterogenous individuals, choose between n different communities

In each community, residents are taxed and the tax revenue is used to provide the local public good.

Tax rates and amount of public good provision are determined by voting $\longrightarrow$ depend on the set of individuals who chose the community!

Equilibrium: allocation of individuals across communities s.t. no individual wishes to move

When does an equilibrium exist? Which conditions ensure existence of an equilibrium?

Contribution: present structure of model and a formal proof of existence of equilibrium. need to restrict preferences and public good provision technology to ensure existence of equilibrium

Epple Filimon Romer RSUE 1993

Model

- continuum of individuals
- J communities
- individuals consume: numeraire composite, housing, local public goods
- level of public good provision is determined by majority-rule voting
- public good is financed by taxes on housing

An equilibrium is an allocation of individuals across J communities, such that

- no individual wants to move
- each community has a nonzero population
- each individual inhabits only one community
- demand for housing equals supply of housing in each community
- community budget balances,
- there is political equilibrium
- no individual can improve utility by changing decision on housing or numeraire good consumption

Epple Filimon Romer RSUE 1993

Model

Individuals have preferences $u(x, h, b)$ and income $y \sim unif[y_{min}, y_{max}]$.

- Utility functions are increasing, quasi concave, twice continuously differentiable
- x public good consumption
- h housing consumption
- b numeraire consumption

Individuals maximize

$$v(x, p, y) = \max_{h,b} u(x, h, b) \quad s.t.\ y \geq ph + b$$

Assume

- u is increasing in x, p, y, separable in x and (h, b)
- u strictly quasi-concave and twice continuously differentiable
- on $\mathbb{R}^3_+$: $u(x, h, b) > u(0, \bar{h}, \bar{b})$ for any arbitrarily large $\bar{h}$, $\bar{b}$
- analogous for h and b. "A total loss of any of the 3 is catastrophic"

Epple Filimon Romer RSUE 1993

Model

Marginal rate of substitution between public good x and housing price p

$$m(x,p,y) := -\frac{v_x}{v_p}, \quad m > 0 \text{ for finite } p$$

- m is continuously differentiable and increasing in y whenever $x > 0$ and $p > 0$
- Slope of (x, p) indifference curve becomes steeper in y
  - individuals care less about price differences and more about public good quality differences when y is higher
- J communities with housing supply function $h_s^j(p_h^j)$
- p_h^j is the net price, individuals pay for housing $p_h^j + \tau^j$ per unit
- set of individuals C, $C \subset [y_{min}, y_{max}]$, residing in community j
- population in the community is $n^j = \int_C dy$
- cost of public good provision is $c(x^j, n^j)$
  - cost of public good provision is identical conditional on the size of the population

Epple Filimon Romer RSUE 1993

Model: Internal equilibrium in community j

Net housing prices p_h^j and tax rates τ^j are an equilibrium within community j iff

- Housing demand matches supply $\int_C h(p^j, y)dy = h_s^j(p_h^j)$
    - notice that $p^j = p_h^j(1+\tau^j)$
    - h is optimal housing demand given p^j, y. $h = \arg\max u(x^j, h, y - p^j h)$
- Community government budget balances $\tau^j p_h^j \int_C h(p^j, y)dy = c(x^j, n^j)$
- Tax rates τ^j emerge from majority voting in the community
- The following needs to hold $\forall (x, \tau)$ for at least half of the mass of voters

$$u(x^j, h(p^j, y), y - p^j h(p^j, y)) \geq u(x, h(p, y), y - p_h(1+\tau)h(p, y))$$

Epple Filimon Romer RSUE 1993

Model: Equilibrium across communities

A partition C of $[y_{min}, y_{max}]$ with J elements and a set of pairs $(x_1, p_1), \ldots, (x_J, p_J)$ is an equilibrium iff

- $\int_{C^j} dy \neq 0$
- (x^j, p^j) is an internal equilibrium in community j
- Every individual $y \in C^j$ prefers (x^j, p^j) to (x^k, p^k)

Epple Filimon Romer RSUE 1993

Existence proof.

Required assumptions

- housing demand increasing in income y
- housing supply function "nice", continuous, etc.
- cost of public good provision is $c(x, n) = c_0 + c_1 x n$ for some $c_0 > 0$, $c_1 > 0$
- individuals do not anticipate equilibrium effects when voting / do not anticipate changes in n^j and p_h^j from their own voting behavior

Find a suitable partition that satisfies boundary indifference / stratification.

With some more restrictions on the MRS between p and x, show there exist internal equilibria on this partition.

Show that nobody strictly preferes a different community.

Use Brouwer's fixed point theorem a few times.

Epple Platt JUE 1998

Introduce additional source of individual heterogeneity: Preferences α

Simplify "public good provision" g to plain "redistribution" g

Epple Sieg JPE 1999

Estimating equilibrium models of local jurisdictions

Want to test sorting model predictions empirically

Build structural general equilibrium (GE) integrated approach

Test equilibrium conditions

- distribution of household income across communities
- relationship between housing prices and demographic characteristics of communities

Match observed quantiles of the income distributions in communities with model-predicted quantiles

Epple Sieg JPE 1999: Model

Set of communities J, continuum of households C

A competitive housing market determines housing prices p and quantities within each community. Moving is free+households behave as price takers

Each community has a local congestable public good g (endogenous public good)

The public good g is financed by local property tax on housing

Voting process determines tax rates+public expenditures in each community. The gross price p of housing is $p = (1+t)p^h$ (voting process not part of the empirical framework)

Households are characterized by income y and taste parameter α. Income and taste are distributed across households along a continuous density $f(a, y)$

Households have preferences over local public good g, housing consumption h, numeraire consumption b

$$\max_{h,b} u(\alpha, g, h, b) \quad \text{s.t. } ph = y - b$$

Epple Sieg JPE 1999: Model

Solve the household utility maximization problem and find indirect utility function

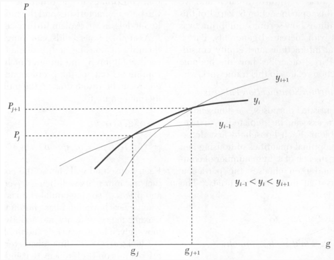
$$v(\alpha, g, p, y) = u(\alpha, g, h(p, y, \alpha), y - ph(p, y, \alpha))$$

where housing is "optimized out".

Consider the indifference curve in the (g, p) plane (increasing slope) and assume that the slope is increasing in y (higher income households are less sensitive to prices and more sensitive to public good provision)

Assumptions

1. The slope of said indifference curve is monotonically increasing in $y \Rightarrow$ "single crossing property"
2. Slope monotonically increasing in $\alpha \Rightarrow$ "single crossing property"



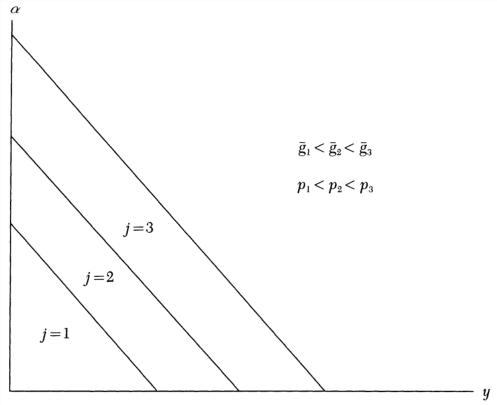
Epple Sieg JPE 1999: Model

Necessary equilibrium properties

1. Boundary indifference. The set of households that is indifferent between community i and j is characterized by

$$\{(a, y) : v(\alpha, g_j, p_j, y) = v(\alpha, g_i, p_i, y)\}$$

If the slope of the indifference curve is monotonic in α and y, then there is a well-defined monotonic function $\alpha \to y$ such that $(\alpha, y(\alpha))$ is indifferent between i and j



Epple Sieg JPE 1999: Model

Testable implications

1. Boundary indifference (see above)
2. Stratification: if $y_j(\alpha)$ defines the set of indifferent households between community j and $j+1$, then for a given α the set of residents in community j are

$$y : y_{j-1}(\alpha) < y < y_j(\alpha)$$

3. Increasing bundles: Assume $p_i > p_j$, then $g_i > g_j \Leftrightarrow y_i(\alpha) > y_j(\alpha)\ \forall \alpha$

Epple Sieg JPE 1999

Towards Estimation

Assume the following functional form for the indirect utility function

$$v(g,p,y,\alpha) = \left[\alpha g^{\rho} + \left(\exp(\frac{y^{1-\nu}-1}{1-\nu})\exp(-\frac{Bp^{\eta+1}-1}{1+\eta})\right)^{\rho}\right]^{\frac{1}{\rho}}$$

CES utility function of public good g and housing consumption/numeraire consumption

- $\rho < 0$, $\alpha > 0$, $\eta < 0$, $\nu > 0$, $B > 0$
- all parameters except α are constant across all households
- assume joint distribution of $\log(\alpha)$, $\log(y)$ is bivariate normal with correlation λ

Housing demand is then (Roy's identity v_p/v_y)

$$h(p,y) = Bp^{\eta}y^{\nu}$$

Towards Estimation

SO, we can actually find the slope of the indifference curve $(-v_g/v_p)$ (call it M) as

$$M(g,p,y,\alpha) = \frac{\alpha g^{\rho-1}\left(\exp(\frac{y^{1-\nu}-1}{1-\nu})\right)^{-\rho}\left(\exp(-\frac{Bp^{\eta+1}-1}{1+\eta})\right)^{-\rho}}{Bp^{\eta}} > 0$$

$\rho < 0$ has to hold to ensure single crossing property (slope strictly increasing in y and α)

With this functional form, can find the boundary indifference condition $(v(p,g,y(\alpha),\alpha) = v(p',g',y(\alpha),\alpha))$

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

$Q_j := \exp(-\frac{\rho}{1+\eta}(Bp_j^{\eta+1}-1))$; $K_j := \log\left(\frac{Q_{j+1}-Q_j}{g_j^{\rho}-g_{j+1}^{\rho}}\right)$

Towards Estimation

Notice $\log(\alpha) = K_j + \rho\left(\frac{y^{1-\nu}-1}{1-\nu}\right)$

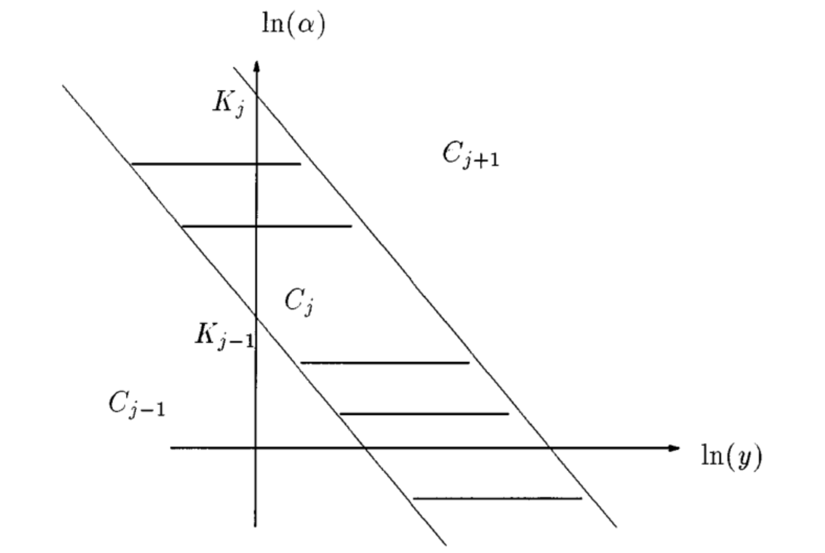


FIG. 1.—Distribution of households across communities, given the parameterization of the model.

$log(y) = 0$ when $y = 1$. The lowest-income community contains all households lying below the K_1 boundary. Obtain population in community j by integrating between the lines through K_{j-1} and K_j.

Epple Sieg JPE 1999

Towards Estimation

- boundary indifference functions in the $(\log(y), \log(\alpha))$ space are determined by parameters ρ and ν. Integrate between lines that intersect $\log(\alpha)$ at K_j, K_{j-1}
- income distribution per community is specified by parameters $(\mu_y, \mu_\alpha, \lambda, \sigma_y, \sigma_\alpha)$, preference parameters ρ, ν, and the community specific intercepts K_j

Mass of people in community j

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

- f are the marginal density functions of $\log(y)$ and $\log(\alpha)$

Can solve boundary indifference conditions recursively (start at lowest or highest) to estimate community-specific intercepts as a function of parameters, and then community sizes $P(C_1), \ldots, P(C_J)$

Epple Sieg JPE 1999

Towards Estimation

With these population mass per community, we can get at the income distribution.

Calculate the qth quantile of the income distribution in community j, $\zeta_j(q)$

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

Finally, let $\theta \in \Theta$ be an element of the underlying parameters space Θ / one of the many parameters in the model.

A differentiable function maps $(0, 1) \times \Theta$ on the income quantiles of community j: $\log(\zeta_j(p, \theta))$ (a.k.a. the resulting quantile of the income distribution function changes "smoothly" in all parameters / the quantiles are differentiable functions of the parameters)

Epple Sieg JPE 1999: Estimation

Two steps

1. Match quantiles of community-specific income distribution predicted by model with empirical counterparts (estimate some parameters)
2. Use empirical housing prices and information on public good provision. Use boundary indifference conditions, orthogonality conditions. (estimate remaining parameters)

Let $q \in (0,1)$ and $\zeta_j(q)$ qth quantile of the income distribution in community j. Observe empirical quantiles of income distro and

$$\zeta_j^N(q) = F_{j,N}^{-1}(q)$$

1. Moment conditions: Match empirical 25th, 50th, 75th percentile of income distribution to model predicted percentiles; for each community. Yields $3 \times J$ moment conditions.
2. Use information on community population sizes. Express community-specific intercepts K_j recursively as a function of model parameters and observed community sizes

Epple Sieg JPE 1999: Estimation

Identify in first step $\theta_1 = (\nu, \frac{\rho}{\sigma_\alpha}, \mu_y, \sigma_y, \lambda)'$ by estimating

$$\theta_1^N = \arg\min_{\theta_1}\{e_1^N(\theta_1)' A_1^N e_1^N\}$$

A_1 positive semi-definite matrix. $e_1^N(\theta)$ vector of moment conditions $(\log(\zeta_1(0.25,\theta)) - \log(\zeta_1^N(0.25,\theta)), \ldots, \log(\zeta_J(0.75,\theta)) - \log(\zeta_J^N(0.75,\theta)))$

So far: need only the quantiles of income distribution by community to estimate this first set of parameters!

Epple Sieg JPE 1999: Estimation

Public good provision

Assume the level of public-good provision can be expressed as an index $g_j = x_j'\gamma + \epsilon_j$,

- need to estimate γ
- x_j are the observable components
- ϵ_j are the unobservable components of public goods

(Problem: ϵ_j will enter estimation non-linearly)

Community-specific intercepts K_j are a function housing prices p_j and public good provision g_j. Re-arrange the definition of K_j

$$g_{j+1}^{\rho} = g_j^{\rho} - (Q_{j+1} - Q_j)\exp(-K_j)$$

this is recursive! Q_j monotonic function of p_j; take K_j from previous section, rewrite

$$\epsilon_j = x_j'\gamma - \left[g_1^{\rho} - \sum_{i=1}^{j}(Q_i - Q_{i-1}exp(-K_i)\right]^{\frac{1}{\rho}}$$

Epple Sieg JPE 1999: Estimation

$$\epsilon_j = x_j'\gamma - \left[g_1^\rho - \sum_{i=1}^{j} (Q_i - Q_{i-1} \exp(-K_i) \right]^{\frac{1}{\rho}}$$

Now assume $E(\epsilon_j | x_j, p_j) = 0 \rightarrow$ Non-linear regression

$\Rightarrow$ identify remaining parameters $\theta_2 = (\rho, \eta, \gamma, \sigma_\alpha, \mu_\alpha)$

Identifying assumption: ϵ_j unrelated with x_j and p_j (problematic). Of course housing prices will be affected by unobserved amenities.

Need instrument. Here: use community income rank as instrument

$\Rightarrow$ estimate GMM

Epple Sieg JPE 1999: Data

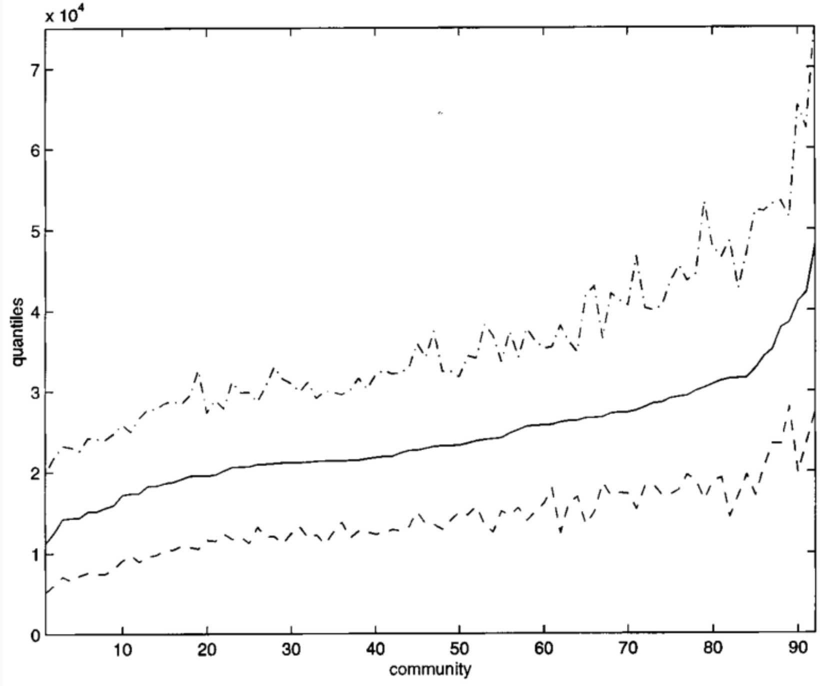
U.S. Census 1980, focus on one metro area: Boston

Focus on one metro area because

- households can choose to re-locate within a metro area without giving up their jobs
- Boston metro contains 17 cities, 75 towns, 92 municipalities $\longrightarrow$ can test theory
- median income across communities ranges from \$11,200 to \$47,646 (1980 USD)
  Community sizes were used yet not targeted specifically. Aim to match pop sizes

Epple Sieg JPE 1999: Results

Income distribution quantiles



Notation: — median, - - 25% quantile, -. 75% quantile.

FIG. 2.—Estimated quantiles

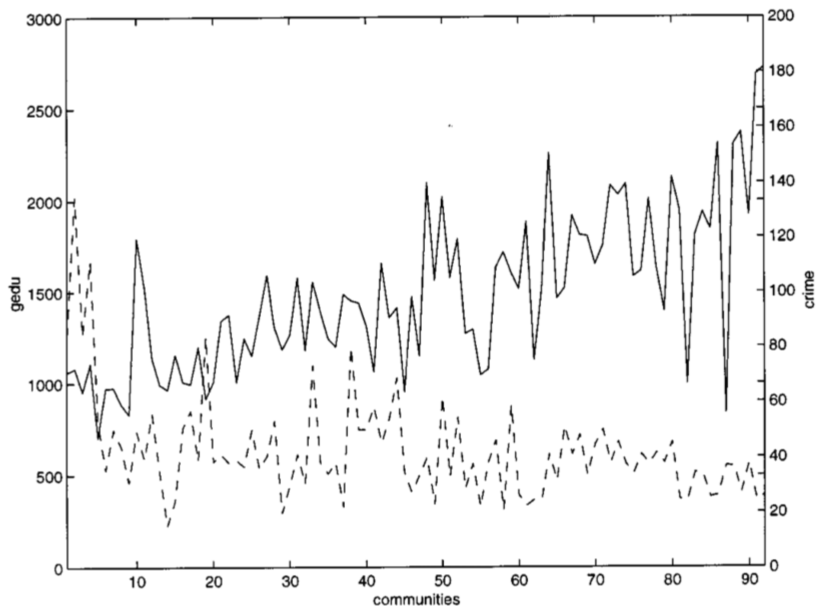
Model prediction: Ranking by 25th percentile vs 50th percentile etc should yield the same ranking. Looks good.

Quantiles are not 'estimated'? Are empirical. Side note: the poorer communities are larger

Public good provision

Proxy public safety with crime levels, school quality with expenditure on education.
Underlying assumption in the model: can capture 'public goods' in a single index

- Correlation of median income and crime rate: -.43
- Correlation of median income and education expenditures: .74
- Correlation of median income and population density: -.6



Notation: - - crime rates, — education expenditures.

Epple Sieg JPE 1999: Results

First step: match income quantiles

ESTIMATED PARAMETERS: STAGE 1

Parameters	Estimates
$\mu_{\ln(y)}$	9.790
	(.002)
$\sigma_{\ln(y)}$	.755
	(.004)
λ	−.019
	(.031)
$\rho/\sigma_{\ln(\alpha)}$	−.283
	(.013)
ν	.938
	(.026)
Function value	.0368
Degrees of freedom	271

NOTE.—Estimated standard errors are in parentheses.

Epple Sieg JPE 1999: Results

Second step: recursively use boundary indifference conditions

ESTIMATED PARAMETERS: STAGE 2

	NLLS		GMM		
	(1)	(2)	(3)	(4)	(5)
γ	−1.95	−1.97	−1.97	−2.08	−2.26
	(1.88)	(1.91)	(4.95)	(4.99)	(1.12)
$\mu_{\ln(\alpha)}$	−2.48	−1.91	−3.11	−2.91	−3.36
	(.65)	(2.87)	(1.80)	(1.38)	(.73)
$\sigma_{\ln(\alpha)}$	.60	.64	.81	.83	.87
	(.19)	(.58)	(.34)	(.38)	(.24)
ρ	−.17	−.18	−.23	−.23	−.25
	(.05)	(.16)	(.10)	(.11)	(.07)
η	−.30	−.70	−.30	−.50	−.30
		(2.01)			
Instruments	4	5	4	4	6

NOTE.—Estimated standard errors are in parentheses. The sample size is 92.

Sign of ρ provides an empirical test of the model, needs to be negative+is found to be negative

Epple Sieg JPE 1999

Conclusion

New method for estimating spatial equilibrium models

Estimate using income quantiles and using boundary indifference conditions implied by sorting model

Found that spatial models can replicate many 'empirical regularities' in the data

Sieg Smith Banzhaf Walsh IER 2004

Estimating the GE benefits of large changes in spatially delineated public goods

Kuminoff Smith Timmins JEL 2013

The New Economics of Equilibrium Soring and Policy Evaluation Using Housing Markets

Side note: JEL papers provide summaries of the current state of the literature. When a new field emerges, first there are papers, then there are summary papers - like JEL papers, then there are textbooks.

Literature review of sorting models.

Households sort into locations based on wealth/income, preferences, and other socioeconomic characteristics.

The individual choices add up and generate effects on endogenous amenities, for example: traffic congestion, air pollution, quality of local schools.

New sorting frameworks combine theory/characteristics of spatial equilibria with other information on households to estimate structural parameters. These can be used for welfare analysis/policy evaluation.

Kuminoff Smith Timmins JEL 2013

Sorting process reveals information about individuals. Can use this information for policy design; especially for policies that target externalities+public good provision.

Distinguish: preference heterogeneity, constraints, choices that are observable versus unobservable.

Integrate these different 'sources of heterogeneity' (constraints+preferences) into one model.

Use properties/characterization of equilibria to estimate structural parameters that characterize heterogeneity in individuals.

Sorting models can be used to evaluate non-marginal changes (improvement over reduced form work and Rosen-style frameworks)

Sorting frameworks are "general equilibrium" frameworks, can model "feedback effects", "endogenous amenities", and changes in choices in response to changes in the environment

Practically: combine information from hedonic estimations and discrete choice estimation

Kuminoff Smith Timmins JEL 2013

How are "equilibrium sorting frameworks" different from plain old differentiated goods?

- Allow explicitly for unobservable heterogeneity,
- allow multiple dimensions of observable heterogeneity for individuals
- allow for feedback effects/endogenous amenities (congestion, voting, etc). "general equilibrium" approach
- can inform policy design / be used to simulate policy counterfactuals

Popular application for sorting models: housing markets

Typical challenges in estimating sorting frameworks:

- need micro data on individual or household level choices + additional information on individual or household characteristics
- instruments for endogenous amenities
- multiple equilibria
- how to incorporate moving cost?
- forward looking behavior of individuals? dynamics?

Foundations

Motivation: Estimate demand for public goods that are not traded in markets

First pass: Regress local government expenditures on voter characteristics, unit of observation is a location.

Bergstrom and Goldman 1973: Regress public good expenditures $\boldsymbol{a}$ per community on median voter income $\boldsymbol{y}$, median voter marginal tax rates, and other median demographics $\boldsymbol{\delta}$

$$\log a = \beta_0 + \beta_1 \log y_{median} + \beta_2 \tau_{median} + \sum_k \beta_k \delta_{k,median} + u_{median}$$

Tiebout's insight: Individuals reveal information about their preferences by choosing a location. Taste-based sorting means u_{median} will not be iid

Kuminoff Smith Timmins JEL 2013

Foundations

Rubinfeld, Shapiro, Roberts 1987: Individuals choose location that maximizes utility over the choice set

$$j = \arg\max_{j \in J} v(a_j, \tau_j, y_{ij}, d_i)$$

Then estimate on the individual level

$$\log a_{ij} = \beta_0 + \beta_1 y_{ij} + \beta_2 \log \tau_{ij} + \sum_k \beta_k d_{kij} + u_{ij}$$

Error term is a function of individual unobserved heterogeneity. τ will be endogenous to community composition through voting, income y will be endogenous to location choice. Have: Tiebout bias + OVB

Need instruments for y and τ. How to get good instruments for this?

Foundations: Standard model of location choice

- n houses, with housing characteristics h_n;
- Community j has level g_j of public goods (school quality, environmental quality, infrastructure, safety; demographic composition of community, share of college degree holders)

Assume: households are fully mobile, fully informed, can choose location free from discrimination. Household utility is given by

$$\max_{n,b} u(b, h_n, g_j; \alpha_i, d_i) \quad s.t. \; y_{ij} = b + p_n$$

- b numeraire consumption,
- α_i household preference parameter,
- y_{ij} household annual income
- p_n price of house n/annualized user cost of occupying house n for one year

Equilibrium: no household wants to move, given prices, wages, tax rates, public good levels

Kuminoff Smith Timmins JEL 2013

Foundations: Standard model of location choice

Theory papers

Prove existence+uniqueness of equilibria, find sufficient conditions

Empirical papers

Use properties of equilibria to estimate valuation.

Use empirical estimates for welfare analysis and simulation of large changes to amenities

Evolution of sorting theory: Stratification

Tiebout's paper - like may papers of the era - is essentially an essay. Few equations and formal definitions, no standard utility maximization problem.

Earlier papers in the sorting literature: Try to formalize Tiebout's main ideas. Focus on formalizing the demand side of the housing market and the formation of endogenous amenities. Housing supply usually treated as fixed or as CES production function.

There are J communities to choose from. How to characterize the optimal decision for households? First, solve consumption versus housing problem for each community j. Then pick community that offers maximal utility

For each community j, households solve utility max problem

$$\max_{h,b} u(b, h, g_j, \alpha) \quad s.t.\ \ y = b + p_j h$$

- b numeraire consumption, h housing consumption (in continuous units), g_j level of public good g in community j, α preference parameter

Kuminoff Smith Timmins JEL 2013

Evolution of sorting theory: Stratification

Assume there is a continuum of households with income levels $y \in [y_{min}, y_{max}]$

"optimize out" the optimal level of housing consumption and find the indirect utility function

$$v(g, p, \alpha, y) := u(y - h(g, p, \alpha, y), h(g, p, \alpha, y), g_j, \alpha)$$

What can we learn about the composition of communities in equilibrium? Start small:

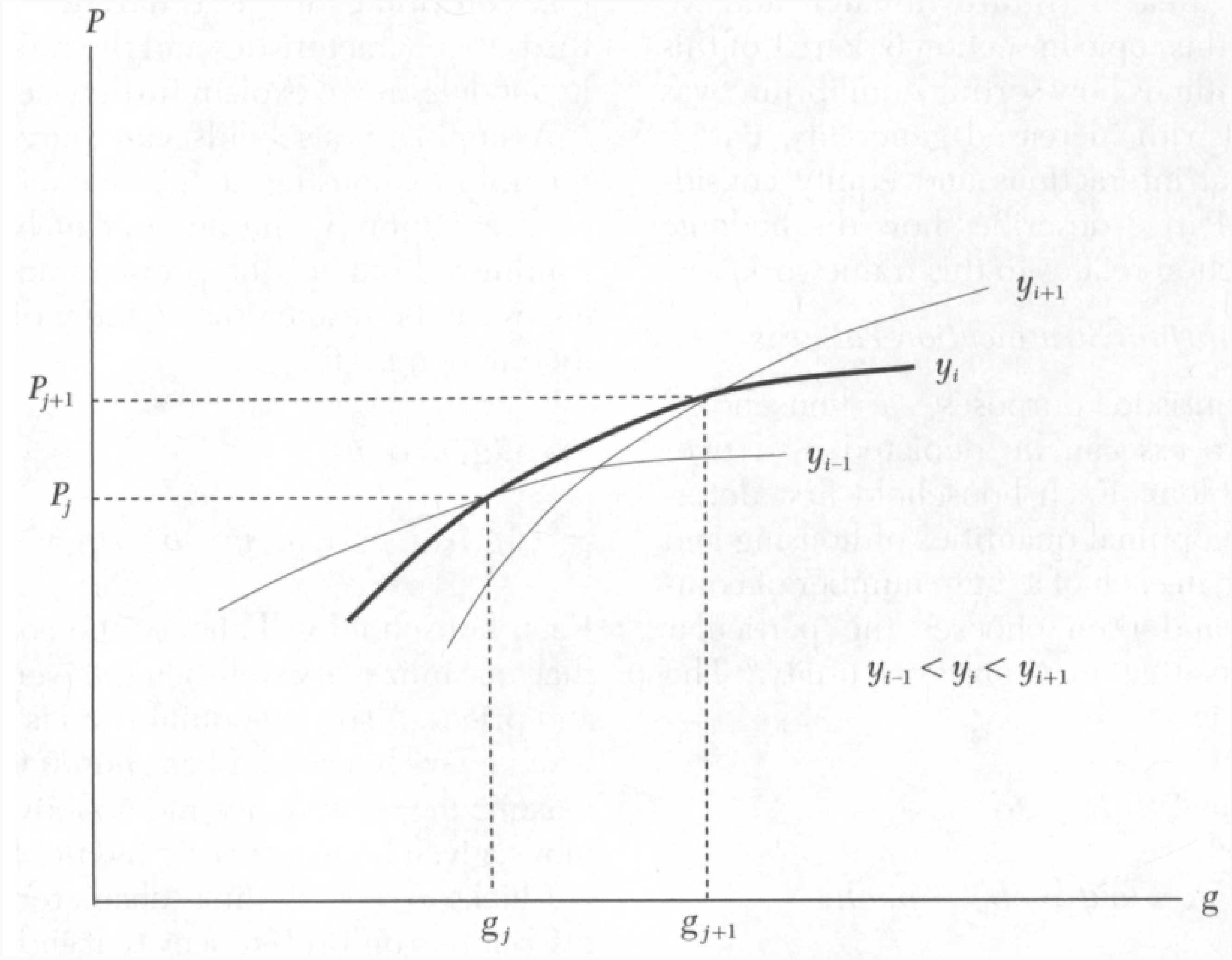
If all households are identical in income y and preference α, there can still be different communities, as long as all communities offer combinations of prices p and public good g that lie on the households' indifference curves!

If utility is increasing in $g \Rightarrow$ communities with higher level of g need to have higher housing prices p in any equilibrium (this is a fundamental point. This means that any ranking of communities by public goods needs to be identical to the ranking of communities by price, as long as everyone agrees on what is "good" versus "bad". **increasing bundles**)

What if households are identical in preference α, but vary by income y? Figure...

Kuminoff Smith Timmins JEL 2013

Single crossing condition. Indifferences curves of 3 households



Evolution of sorting theory: Stratification

Have two communities with exogenous public good levels g_j and g_{j+1}, $g_j < g_{j+1}$. If households receive utility from g, any market clearing equilibrium needs to fulfill $p_j < p_{j+1}$

If the slope of the indifference curve $p(g)$ is continuously increasing in income, and there is a continuum of households s.t. for each income level y there exists a household that has this income level

$\rightarrow$ there exists an indifferent household i with income y_i. All households with income $y < y_i$ strictly prefer community j, all households with income $y > y_i$ strictly prefer community $j + 1$ (**stratification**, **boundary indifference**)

Notice this is an inductive argument.

Conditional on preferences, communities will be perfectly stratified by income!

Kuminoff Smith Timmins JEL 2013

Evolution of sorting theory: Stratification

(Epple Platt 1998, Epple Sieg JPE 1999)

In real life, communities are *not* perfectly stratified. Not everyone in community j has lower income than everyone in community $j + 1$

But empirically, ranking communities by median income often reflects pretty well the community ranking by public goods (heuristic). So there is *some* ranking but also *some* overlap

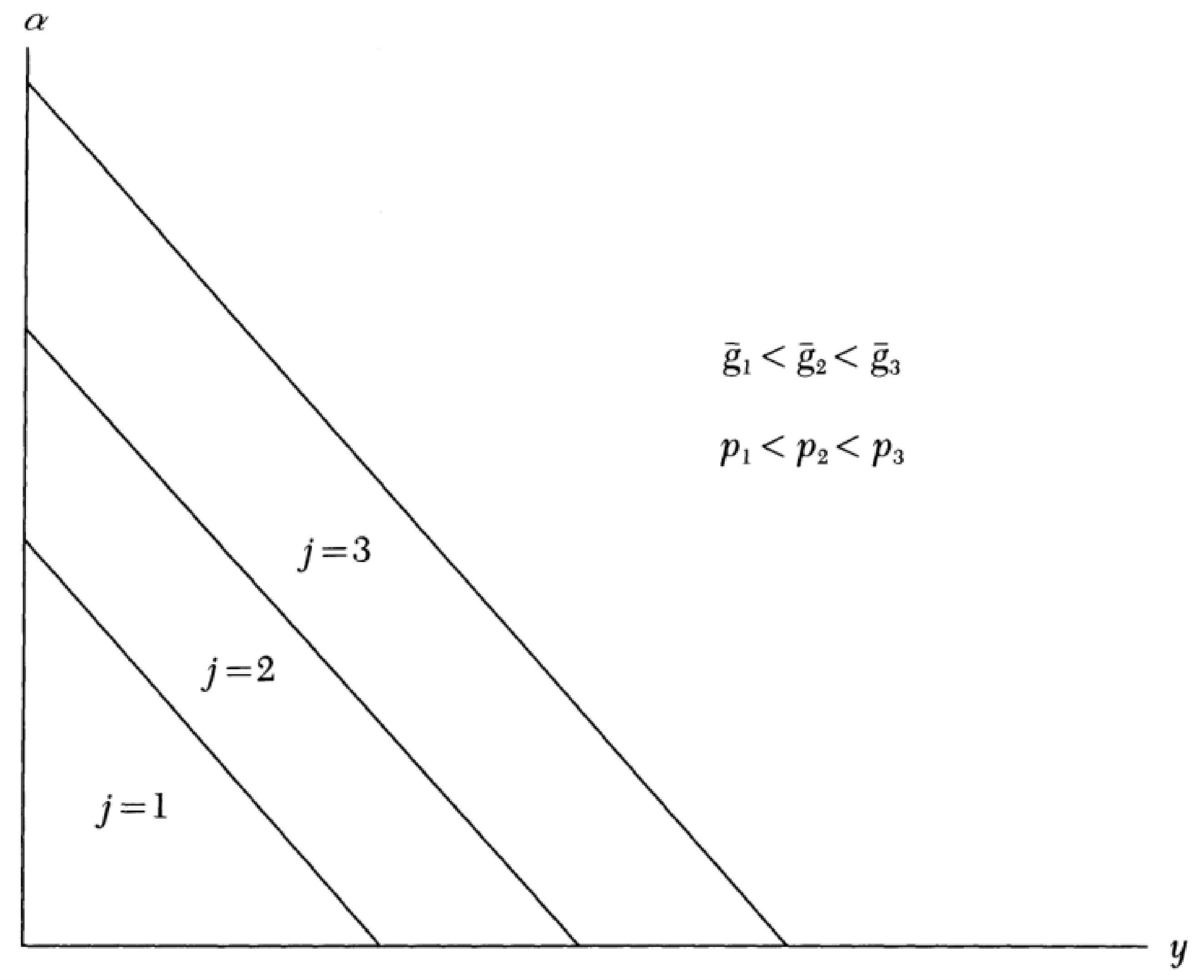
Take a closer look at preferences α

Analogous to the income case! Fix income, and assume that the slopes of indifference curves are either strictly increasing or strictly decreasing in preference parameter $\alpha \Rightarrow$ conditional o income, communities will be perfectly stratified by (one-dimensional) preferences.

Putting these two together (income+preferences), we end up with communities that are (imperfectly) stratfied by income, reconciling the ranking of communities by median income with some overlap in the income distribution between communities

Kuminoff Smith Timmins JEL 2013

Distribution of households across communities



Evolution of sorting theory: Stratification

Notice that these 3 properties / testable implications

- boundary indifference
- stratification
- increasing bundles

are necessary conditions in any sorting equilibrium, but not sufficient. You may find allocations of households across J communities that fulfill boundary indifference, stratification, and increasing bundles, but are not equilibria

Evolution of sorting theory: Social interactions

Endogenous amenities! Demographic composition of communities affect local provision of public goods. For example, property crime is a function of income/poverty. An isolated high-income community is likely to have lower property crime.

Social interactions can generate multiple equilibria

Evolution of sorting theory: Social interactions

Nechyba 1999: Case of school quality and vouchers for private schools.

- Assume there are 2 communities and 2 types of households: college-educated and non-college educated households. Otherwise same model as before.
- The public good is "school quality". One community has higher quality than the other
- Assume household demand for "school quality" is increasing in income, and in college status (preference parameter α!)
- Assume that the presence of college educated households in the community increases school quality ("peer group effect").

If the peer group effect is weak $\Rightarrow$ communities will be segregated by college status. If peer effect is strong, non-college households will want to co-locate with college households, may get "integrated" communities with both types in both communities

Kuminoff Smith Timmins JEL 2013

Evolution of sorting theory: Social interactions

Nechyba 1999: Application of "private school vouchers" Different neighborhoods with different levels of public school quality. Find+calibrate "integrated" equilibrium with some overlap of income distribution across communities.

Policy simulation: Add private schools + partial vouchers for private schools.

Result: parents move from expensive districts to cheaper districts with worse public school quality, send their children to private schools. 1. This lowers school quality in more expensive districts. 2. This increases housing prices in cheaper districts.

$\rightarrow$ relatively richer households from cheaper neighborhoods realize capital gains from housing price increases, sell their houses, move to better neighborhoods.

$\rightarrow$ this lowers public school quality in cheaper districts, because the peer effects from the within-cheap district relatively richer households/"best students" are lost

Offsetting effect: higher income households moving into cheaper neighborhoods increases tax revenues in cheaper neighborhoods $\rightarrow$ increase school quality in cheaper neighborhoods

Kuminoff Smith Timmins 2013

Empirical models of sorting: Estimation

Epple and Sieg 1999: Housing is a homogenous good, buy at per-unit price $p_j \rightarrow$ the per-unit price of housing is a function of the local public goods in the community.

What justifies the "constant per unit price"? Show that when utility is separable+homogeneous of degree one in housing characteristics, demand for housing characteristics does not depend on the level of amenities. Expenditure on housing P_n can be expressed as

$$P_n = p(g_j)h(h_n)$$

Take logs, then can estimate using housing-level micro data

$$\log(P_n) = log(h(h_n)) + log(p(g_j))$$

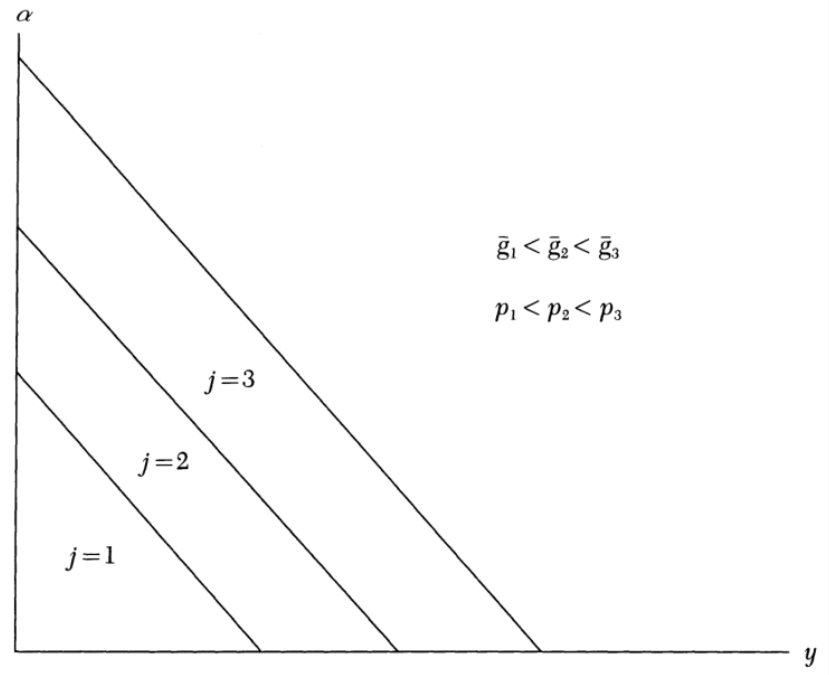
...get community specific per-unit $p(g_j)$ price as fixed effect!

Kuminoff Smith Timmins 2013

Policy evaluation

Kuminoff Smith Timmins JEL 2013

Recap: Distribution of households across communities. Households are heterogeneous in income y and taste ε



Given exogenous amenities and housing-supply curves per community, find a partition of households across communities and per-unit prices such that nobody wants to move and housing markets within communities clear $\Rightarrow$ **Equilibrium**

Thanks!

Slides created via the R package **xaringan**.

The chakra comes from remark.js, **knitr**, and R Markdown.