A Dynamic Model of Demand for Houses and Neighborhoods

Bayer, McMillan, Murphy, Timmins
forthcoming Econometrica

Florian Oswald

Graduate Labor, SciencesPo 2017

April 13, 2017

Intro

- Residential Sorting is an important policy concern.
- People want to live where education is good, crime is low, and air quaility is high.
    - Rich people want to be close to rich people.
    - White people close to white.
- All of this affects welfare.
- How should the government provide public goods, when there are different jurisdictions of rich/poor neighborhoods?
- How will such a public goods provision affect the equilibrium of the region/city?
- For example: The Paris House Price Cliff. Who pays for the Metro?
- This is related to the Tiebout Model.

Tiebout Model

- Charles Tiebout (1956)
- Theory of local taxation (and tax competition between localities)
- Most important assumptions
  - Zero Moving Costs
  - Complete Information
  - No Commuting costs
  - No spillover of public goods
- Communities will try to attract people by offering an attractive mix of local public goods and tax rates.

Need for a Dynamic Model

- Pretty much all models in this literature are static:
- Household location decisions are inherently dynamic.
    1. Large transaction costs make moving rare
    2. Household circumstances change over time
    3. Local amenities and house prices change over time
- There is reason to be concerned about bias in static estimates.

Why has this not been done?

- Data: require large sample of households including their characteristics, location features and housing choices.
  - i.e. household data with a high resultion geographical identifier.
- Computational intensity: Given many locations, with characteristics, the state space of such models becomes very large.
  - Heterogeneous consumers
  - Heterogeneous Locations
  - Kennan and Walker (2001): 10m points per age
  - Oswald (2015): 25m savings problems

What this paper does

- Develop a model of dynamic neighborhood choice.
  - Devises a computationally light estimator.
  - Builds upon durable demands literature.
  - Compares results to a static setting and finds wildly different estimates.
- Households decide whether and where to move.
  - This decision depends on how they think neighborhoods characteristics (mainly: price) will evolve
  - This evolution affects the expected value of living in that location.

What this paper does NOT do

- General Equilibrium.
- The evolution of amenities (air quality etc), and more importantly, house prices, is exogenous.
- It is not specified what makes house price move around, supply shocks, etc.

What the model could be used for

- The machinery set up here could be useful in several other applications:

1. microdynamics of residential segregation
2. microdynamics of gentrification
    1. there are some theoretical papers (Guerrieri, Hartley and Hurst)
    2. very little on the empirical (certainly dynamic) side.

Data

- 6 counties of the San Francisco metropolitan area (Bay Area)
- Two data sources: dataquick (proprietary) and HMDA
- DataQuick:
  - each housing unit sold 1994–2004
  - buyer's and seller's name, transaction price, exact address, square footage, year built etc..
- Home Mortgage Disclosure Act (HMDA):
  - characteristics of ALL mortgage applicants
- Merge both based on census tract id, loan amount, date and mortgage lender name.
  - Census tract: ca 4000 people
  - unique match for 70% of sales

Summary Stats

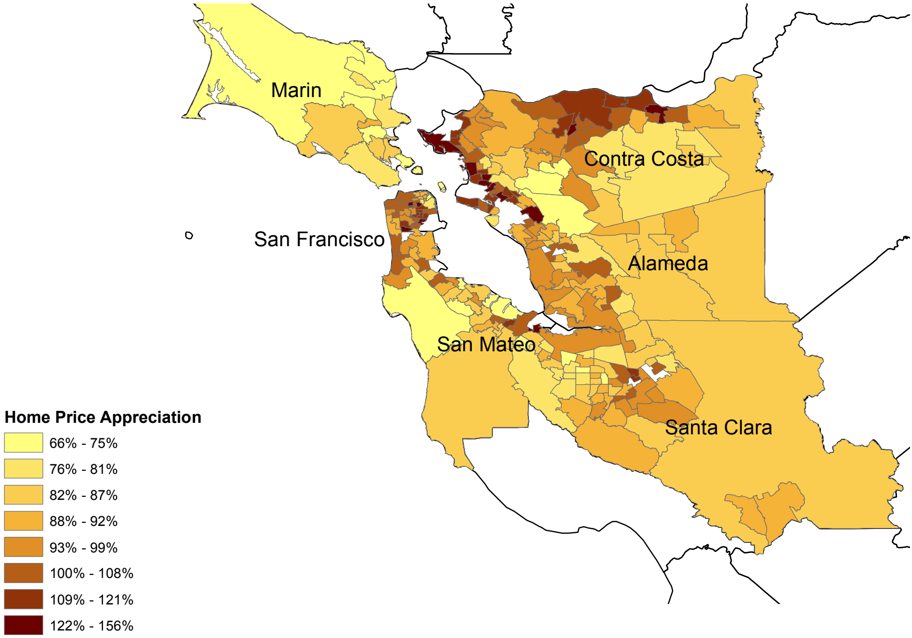
Household Characteristics

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
Income	220403	106.87	45.44	0.89	240.00
Down-payment	220403	82.46	51.92	0.00	240.00
Sales Price	220403	382.86	163.70	98.53	1536.71
White	220403	1	0	1	1
Year	220403	1999.04	3.17	1994	2004

Neighborhood Characteristics

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
Percent White	2398	69.63	16.21	26.69	96.79
Violent Crime	2398	453.67	247.02	46.03	2011.05
Ozone	2398	2.17	2.57	0.002	18.25
Sales Price	2398	429.13	206.27	122.75	1792.01

Note: Income, Down-payment, and Sale Price are measured in $1000's.


Marin
Contra Costa
San Francisco
Alameda
San Mateo
Santa Clara
Home Price Appreciation
66% - 75%
76% - 81%
82% - 87%
88% - 92%
93% - 99%
100% - 108%
109% - 121%
122% - 156%

Dynamic Considerations?

$$\text{share}_{jt} = \beta_1 \%\text{white}_{jt} + \beta_2 \text{violence}_{jt} + \beta_3 \text{ozone}_{jt} + \beta_4 \text{price}_{jt} + u_{jt}$$

	Share	Share
Percent White	0.02479	0.02709
	(0.00026)	(0.00329)
Violent Crime	-0.00092	-0.00047
	(0.00002)	(0.00003)
Ozone	0.07284	0.04831
	(0.00183)	(0.00184)
Price	-0.01331	-0.00734
	(0.00017)	(0.00073)
Lagged Percent White		-0.00316
		(0.00328)
Lagged Violent Crime		-0.00034
		(0.00003)
Lagged Ozone		0.07092
		(0.00160)
Lagged Price		-0.00577
		(0.00078)

Dynamic Model

- There are 2 main considerations:
    - wealth accumulation: how are prices going to evolve?
    - moving costs: how costly in terms of utility and money is moving?
    - monetary moving costs are 6% of house value in the US.
- Model for Homeowners who decide to stay or move in the Bay Area.
    - Renting and moving away from the Bay area is the outside option.

Discrete Choice Setup

- Decide whether to move or not.
- $d_{i,t}$ encodes for HH i in period t choice to
  - move: $d_{i,t} = j \in \{0, 1, \ldots, J\}$
  - stay: $d_{i,t} = J + 1$
- Observed State Variables:
  - $X_{j,t}$: price of housing, local crime, racial composition etc
  - $Z_{i,t}$: household characteristics, income, wealth and race.
  - $h_{i,t} \in \{0, 1, \ldots, J\}$ is neighborhood choice in $t - 1$. history.
- Unobserved States:
  - g_i: unobserved houshold type. Love area 1, e.g.
  - $\xi_{j,t}$: unobserved neighborhood quality
  - $\varepsilon_{i,j,t}$: idiosyncratic shock

Model Primitives

- Write the model primitives as (u, MC, q, β)
- $u_{i,j,t} = u\left(X_{j,t}, \xi_{j,t}, g_i, \varepsilon_{i,j,t}\right)$ per period utility of living in j, net of movign costs.
- $MC_{i,t} = MC(Z_{i,t}, X_{h_{i,t}})$: function of where you move *away from.*
- full utility is then

$$u_{i,j,t}^{MC} = u_{i,j,t} - MC_{i,t}\mathbf{1}\left[j \neq J+1\right]$$

- $q(s_{i,t+1}, h_{i,t+1}, \varepsilon_{i,t+1} | s_{i,t}, h_{i,t}, \varepsilon_{i,t}, d_{i,t})$: Markovian law of motion of state space.

Choice Problem

The objective is to

$$\max_{\{d_{i,r}\}_{r=t}^{T}} \left[\sum_{t=r}^{T} \beta^{r-t} \left(u^{MC} \left(X_{j,r}, \xi_{j,r}, Z_{i,r}, g_i, \varepsilon_{i,j,r}, X_{h_{i,r}} \right) \right) | s_{i,t}, h_{i,t}, \varepsilon_{i,t}, d_{i,t} \right]$$

which admits a recursive formulation:

$$\begin{aligned} V(s_{it}, h_{i,t}, \varepsilon_{i,t}) &= \max_{j} \Big\{ u_{i,j,t}^{MC} \\ &+ \beta E \left[V(s_{i,t+1}, h_{i,t+1}, \varepsilon_{i,t+1}) | s_{i,t}, h_{i,t}, \varepsilon_{i,t}, d_{i,t} = j \right] \end{aligned}$$

Assumptions

- Standard Rust (1985) assumptions.
  - Additive Separability of utility from shocks.
  - Conditional Independence Assumption of q from shocks.
  - Shocks are Type 1 extreme value distributed.
- Allows to write the *choice-specific value function*

$$
\begin{aligned}
v_j^{MC}(s_{i,t}, h_{i,t}) &= u_{i,j,t} - MC(Z_{i,t}, X_{h_{i,t}})\mathbf{1}\left[j \neq J+1\right] \\
&+ \beta E\left[\log\left(\sum_{k=0}^{J+1} \exp\left(v_{k,}^{MC}(s_{i,t+1}, h_{i,t+1})\right)\right) | s_{i,t}, d_{i,t} = j\right]
\end{aligned}
$$

Location Choice

- Rewrite this as

$$v_j^{MC}(s_{i,t}, h_{i,t}) \quad = \quad v_j(s_{i,t}) - MC(Z_{i,t}, X_{h_{i,t}})\mathbf{1}\,[j \neq J+1]$$

where

$$v_j(s_{i,t}) \quad = \quad u_{i,j,t} + \beta E\left[\log\left(\sum_{k=0}^{J+1}\exp\left(v_{k,}^{MC}(s_{i,t+1}, h_{i,t+1})\right)\right)|s_{i,t}, d_{i,t} = j\right]$$

- Notice that $v_j(s_{i,t})$ is *independent* of previous neighborhood $h_{i,t}$.
- If move, go to highest utility neighborhood.
- Assume that moving costs are identical across neighborhoods.
- Based on characteristics $Z_{i,t}$, put households into bins index by type $\tau$and get a type specific value

$$v_{j,t}^{\tau} \quad = \quad u_{j,t}^{\tau} + \beta E\left[\log\left(\sum_{k=0}^{J+1}\exp\left(v_{k,t+1}^{\tau'} - MC_{j,t+1}\right)\right)\right]$$

Estimation Outline

1. Use location decisions to estimate v_j^T for each neighborhood
2. Obtain estimates of per period utility (fully flexible)

- The proceedure is very low in computational cost
- This is a result of
  - Type 1 EV assumption: get closed form expressions for choice probabilities and future values
  - A closed form expression for the FOC of the resulting log-likelihood function.
  - They have to do some data smoothing in order to deal with zero size bins
- Proceedure is then extended to account for unobserved types.

Results

- Moving Costs have a psychological and a financial component
- Psychological costs are very large.

Table 3: Moving Cost Estimates

Psychological Costs	
Constant	9.50612
	(0.04344)
Income	-0.00209
	(0.00038)
t	-0.15111
	(0.00392)
Financial Costs	
Constant*6% House Value	0.03515
	(0.00148)
Income*6% House Value	-0.00008
	(0.00001)

WTP for 10% increase in amenities

Table 4: Willingness to Pay for a 10-Percent Increase in Amenities

	I	II	III	IV
Percent White	2256.09	2470.99	2188.18	2349.79
	(88.16)	(116.17)	(85.45)	(112.86)
Violent Crime	-760.33	-620.10	-725.19	-573.48
	(43.16)	(43.96)	(41.02)	(42.50)
Ozone	-359.89	-315.50	-347.14	-299.36
	(22.16)	(23.80)	(21.36)	(23.42)
County Dummies	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes
Type Dummies	Yes	Yes	Yes	Yes
Estimator	LAD	OLS	LAD	OLS
Wealth Outliers	NO	NO	YES	YES

These Models need Home Bias

- Model dynamics: Need to avoid too many movers.
  - High moving cost.
  - high utility from living at home.
  - for a \$120,000 income HH, living in the preferred ("home") neighborhood is equivalent to a on off \$73,372 payment.

Dynamic vs Static

- Caveats for this model:
  - worse data than large static model, relying on confidential census data
  - amenities are exogenous here.
- Sources of differences across the two: Time varying neighborhood characteristics.
  - **Mean-reversion**: If crime is known to mean-revert, seeing a high-crime neighborhood today means that it can only get better. Households will have a *higher* WTP for a house in that neighborhood. In static model, this will be downward biased.
  - **Persistence**: racial composition is likely to be very persistent. so many whites today means more white tomorrow. WTP will again be higher than in a static model, since this future "benefit" is lost.

Dynamic vs Static

Table 6: Willingness to Pay for a 10-Percent Increase in Amenities – Static versus Dynamic Estimates by Income

	Static			Dynamic		
	\$40,000	\$120,000	\$200,000	\$40,000	\$120,000	\$200,000
Percent White	1627.02	1901.43	2221.66	612.14	2428.91	4888.42
	(11.28)	(18.76)	(48.55)	(84.45)	(116.72)	(277.96)
Violent Crime	-291.14	-380.67	-448.88	-350.15	-962.19	-1298.80
	(7.68)	(11.08)	(19.02)	(48.66)	(71.46)	(94.06)
Ozone	-66.24	-80.71	-97.04	-302.06	-380.03	-395.58
	(2.13)	(2.43)	(3.15)	(28.30)	(30.12)	(39.32)