A Dynamic Model of Demand for Houses and Neighborhoods

Patrick Bayer *Duke University*

Robert McMillan *University of Toronto*

Alvin Murphy *Arizona State University*

Christopher Timmins *Duke University*

Introduction

- Previous urban literature has been static
- Theory would suggest a dynamic approach is necessary:
    - Large transactions costs
    - As an asset, housing choice affects wealth accumulation
        - Predictable changes in neighborhood characteristics (including prices)
- Static models will bias estimates of marginal willingness to pay for amenities
- Dynamics are key to understanding many features of the housing market:

Introduction

1. Introduce a new dataset
2. Outline a new approach to estimating dynamic demand
3. Provide an insight into housing market dynamics

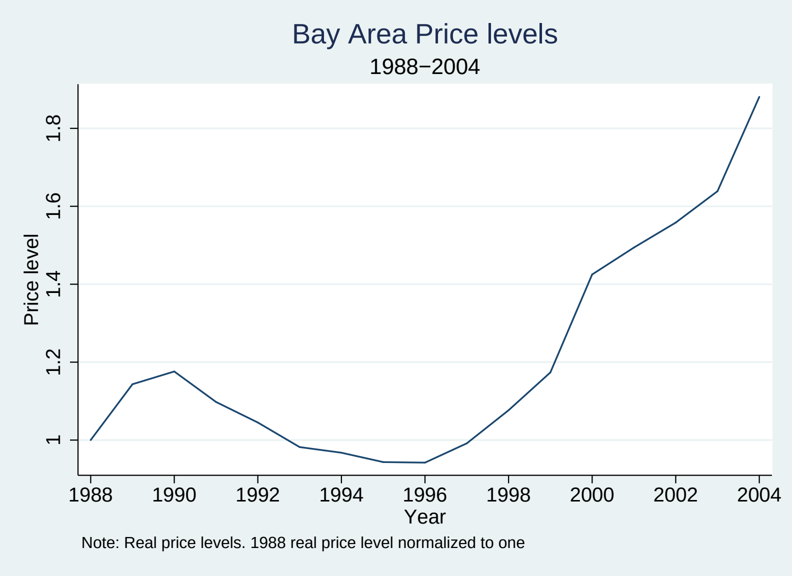
Data

- The unique dataset uses data from multiple sources:
  - housing transactions data
  - mortgage application data
  - the Decennial Census, Rand, California Air Resources Board
- In the merged dataset, both the personal characteristics of buyers and the house characteristics (including location) are observed for house sales in the six counties of the Bay Area of California between 1994 and 2004.
- Also observed is the date an individual resells their house and sales price.
- We impute neighborhood race, crime, and air quality for each neighborhood in every year.

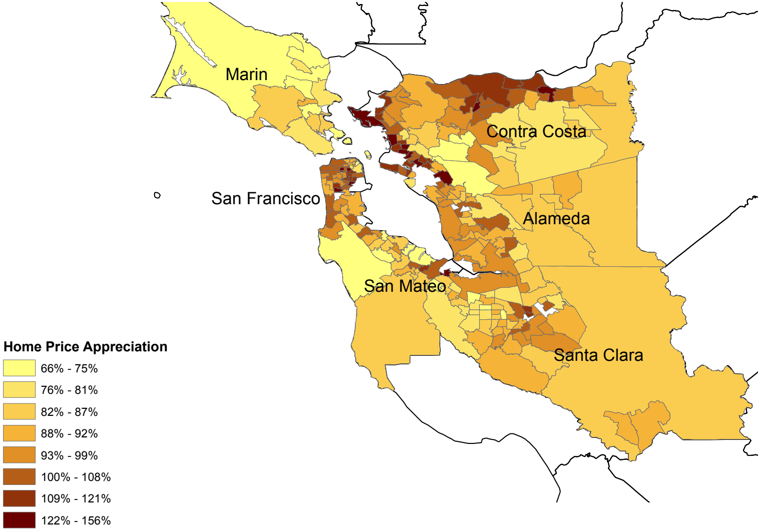
Motivation for Dynamic Approach – Descriptive Analysis

- Agents take into account their expectations about future prices and costs
- In the Bay Area there were significant changes in the housing prices during sample period.
- Large cross section variation in price appreciation.

Bay Area Price Levels



Appreciation Rates by Neighborhood: 1994-2004



Previous Literature

- Static Urban Models:

  Epple & Sieg (1999), Ferreyra (2007), Bayer, Ferreira, McMillan (2007), Kuminoff (2008)

- Demand Models:

  Static: Berry (1994), BLP (1995), ...

  Dynamic: Melnikov (2001), Hendel & Nevo (2006), Aguirregabiria & Nevo (2010)

- Other Dynamic Models:

  Rust (1987), Hotz & Miller (1994), ...

Key Features of Housing Choice

- Households have heterogenous preferences for neighborhood characteristics.
- Large moving costs $\Rightarrow$ prohibitive cost of frequent reoptimization
- Expectations about future housing prices and neighborhood demographics.
- Positive persistence of prices & wealth accumulation from housing asset

Model

- Each period households face two decisions:
  - Stay in current residence or move
  - If move, choose which neighborhood to move to
- Households make both decisions to maximize discounted sum of expected flow utilities
- Households are forward looking
  - The characteristics of the neighborhood may change
  - Neighborhood choice may determine the performance of housing as an asset
  - Neighborhood choice may determine likelihood of future transaction costs

Model : Notation

- The decision variable, d_{it}
  - Move: $d_{it} = j \in \{0, 1, \ldots, J\}$
  - Stay: $d_{it} = J + 1$
- Observed state variables:
  - X_{jt} : vector of characteristics of neighborhood j
  - Z_{it} : vector of characteristics of household i
  - h_{it} : neighborhood chosen in $t - 1$
  - lags of neighborhood/household characteristics
- Unobserved variables:
  - ξ_{jt} : unobserved neighborhood characteristics.
  - ϵ_{ijt} : unobserved idiosyncratic choice specific shock

Model

- Households choose $\{d_{it}\}$ optimally to maximize lifetime utility
- The total flow utility household i receives from neighborhood j at time t is:

$$u_{ijt}^{MC} = u(X_{jt}, \xi_{jt}, Z_{it}) - MC(Z_{it}, X_{h_{it}}) I_{[j \neq J+1]} + \epsilon_{ijt}$$

- The lifetime expected utility is defined as

$$v_j^{MC}(s_{it}, h_{it}) = u_{ijt}^{MC} + \beta E \Big[max_k \{ v_k^{MC}(s_{it+1}, h_{it+1}) + \epsilon_{ikt+1} \} | s_{it}, d_{it} = j]$$

Estimation

- Estimate the model in 4 steps:
  - Estimate flexibly the lifetime expected utilities using observed location choice of movers
  - Estimate moving costs and marginal utility of wealth using timing of moves
  - Estimate transition probabilities and recover the per-period utility function
  - Decompose the per-period utility function

Estimation - Step 1

- Consider the problem faced by a household that has chosen to move.
- The moving cost term, $MC(Z_{it}, X_{h_{it}})I_{[j \neq J+1]}$, is identical for all neighborhoods and so it simply drops out
- Each household therefore chooses j to maximize $v_j(s) + \epsilon_{ij}$:

$$v_j(s_t) = u_{ijt} + \beta E\left[log\big(\sum_{k=0}^{J+1} exp(v_k^{MC}(s_{it+1}, h_{it+1}))\big)|s_{it}, d_{it} = j\right]$$

$$v_{jt}^{\tau} = u_{jt}^{\tau} + \beta E\left[log\left(\sum_{k=0}^{J+1} exp(v_{kt+1}^{\tau_{t+1}} - MC_{jt+1}^{\tau_{t+1}} I_{[k \neq J+1]})\right)\middle| s_{it}, d_{it} = j\right]$$

Estimation - Step 1

- Let P^{τ}_{jt} denote the choice probability.
- Estimate $\tilde{v}^{\tau}_{jt}$ as:

$$\tilde{v}^{\tau}_{jt} = log(\hat{P}^{\tau}_{jt}) - \frac{1}{J+1}\sum_{k=0}^{J} log(\hat{P}^{\tau}_{kt})$$

- $\tilde{v}^{\tau}_{j} = v^{\tau}_{j} - m^{\tau}$ and $m^{\tau} = \frac{1}{J+1}\sum_{j} v^{\tau}_{j}$.
- Large number of types solved by smoothing

Estimation - Step 2

- Households choose to move or stay in each period.
- We estimate moving costs in a dynamic binary choice model.
- Financial component of moving costs is approx. 6% of sales price.
- Incorporate this outside information to recover the marginal value of wealth (convert utils to dollars)
- We recover the baseline utility differences between types.

Estimation - Step 3

- From stages one and two, we know:
  - Moving costs for each type
  - The marginal value of changing type
  - The true mean value terms, v_j^τ.
- Want to recover

$$u_{jt}^{\tau} = v_{jt}^{\tau} - \beta E\left[log\left(\sum_{k=0}^{J+1} exp(v_{kt+1}^{\tau_{t+1}} - MC_{jt+1}^{\tau_{t+1}} I_{[k \neq J+1]})\right)\Big| s_{it}, d_{it} = j\right]$$

Estimation - Step 3

- We assume households use today's states to directly predict future values of the lifetime utilities, v, rather than predicting the values of the variables upon which v depends.
- Households need to predict how their types will change.
  - Need to predict future house values
- Transitions are modeled as:

$$v_{jt}^{\tau} = \rho_{0,j}^{\tau} + \sum_{l=1}^{L} \rho_{1,l}^{\tau} v_{jt-l}^{\tau} + \sum_{l=1}^{L} X'_{jt-l} \rho_{2,l}^{\tau} + \rho_{3,j}^{\tau}\ t + \omega_{jt}^{\tau}$$

$$price_{jt} = \varrho_{0,j} + \sum_{l=1}^{L} X'_{jt-l} \varrho_{2,l} + \varrho_{3,j}\ t + \varpi_{jt}^{\tau}$$

Estimation - Step 3

- Knowing v, MC, and the transition probabilities allows us to calculate mean flow utilities for each type and neighborhood, u^{τ}_{jt}, according to:

$$u^{\tau}_{jt} = v^{\tau}_{jt} - \beta E\left[log\left(\sum_{k=0}^{J+1} exp(v^{\tau_{t+1}}_{kt+1} - MC^{\tau_{t+1}}_{jt+1} I_{[k \neq J+1]}) \right) \middle| s_{it}, d_{it} = j \right]$$

- where in practice, s includes all the variables on the right hand side of the transition equations
- For each type, τ, neighborhood, j, and time, t, we have the necessary information to calculate the expectation on the right hand side

Estimation - Step 4

- Once we recover the mean per-period utilities, we can decompose them into functions of the observable neighborhood characteristics, X_{jt}
- Straightforward if we assume that unobservable characteristics, ξ, are uncorrelated with the other neighborhood characteristics.

$$u_{jt}^{\tau} = \alpha_0^{\tau} + \alpha_c^{\tau} + \alpha_t^{\tau} + X_{jt}' \alpha_x^{\tau} + \xi_{jt}^{\tau}$$

- X includes user cost, % white, crime rate, ozone

Estimation - Step 4: Potential Endogeneity Problems

- We need to control for endogeneity of user cost.
  - We already know marginal value of wealth from Stage 2
  - Assume that the effect of a marginal change in wealth on lifetime utility is the same as the effect of a marginal change in income on one period's utility
  - The negative of the marginal utility of income can be interpreted as the coefficient on rent
- Therefore we estimate the following regression where $\widehat{\gamma_{fmc}^{\tau}}$ is known from Stage 2 and $\tilde{X}$ denotes the non-price components of X.

$$u_{jt}^{\tau} + \widehat{\gamma_{fmc}^{\tau}} usercost_{jt} = \alpha_0^{\tau} + \alpha_c^{\tau} + \alpha_t^{\tau} + \tilde{X}_{jt}' \alpha_x^{\tau} + \xi_{jt}^{\tau}$$

Results

- Estimate the model for 75 types:
  - Whites only
  - 3 Income Types: \$40,000, \$120,000, and \$200,000
  - 25 wealth types: \$10,000 increments – \$0 to \$240,000
- Estimate the marginal willingness to pay for:
  - Increase in a neighborhood's percent white
  - Increase in violent crime rate
  - Increase in number of days ozone exceeds threshold

Results – Stage 2

Table: Marginal Utility of Wealth Results

	Estimate
Psychological Costs	
Constant	9.50612
	(0.04270)
Income	-0.00209
	(0.00035)
t	-0.15111
	(0.00427)
Financial Costs	
Constant*6% House Value	0.03515
	(0.00148)
Income*6% House Value	-0.00008
	(0.00001)

Income and House Value are measured in $1000.

Results – Stage 4

- Per-period MWTP $= \alpha_x^{\tau} / \gamma_{fmc}^{\tau}$
- Willingness to Pay for 10% Increase in Amenities at:
  - Amenity Means: %white: 69.6%, Violent Crime: 453.7, Ozone: 2.2
  - Income: $120,000

Table: Willingness to Pay for 10% Increase in Amenities

Percent White	2256.08
	(95.65)
Violent Crime	-760.31
	(40.24)
Ozone	-359.88
	(21.11)

Results – Stage 4: Heterogeneity

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

	\$40,000	\$120,000	\$200,000
Percent White	612.09	2428.93	4888.46
	(87.09)	(121.89)	(275.08)
Violent Crime	-350.18	-962.20	-1298.81
	(47.68)	(67.29)	(92.79)
Ozone	-302.06	-380.03	-395.58
	(24.82)	(26.98)	(37.12)

Static versus Dynamic – Omitted Variable Problem

$$v_{jt}^{\tau} = u_{jt}^{\tau} + \beta E\left[log\left(\sum_{k=0}^{J+1} exp(v_{kt+1}^{\tau_{t+1}} - MC_{jt+1}^{\tau_{t+1}} I_{[k \neq J+1]}) \right) \Big| s_{it}, d_{it} = j \right]$$

- Current neighborhood characteristics determine the choice specific value functions in two ways:
  1. They affect the flow utility directly
  2. They help predict future neighborhood utility.
- Estimating a static model omits the second effect.
- The omitted variable can be expressed as:

$$\beta E\left[log\left(\sum_{k=0}^{J+1} exp(v_{k,t+1}^{\tau_{t+1}} - MC_{j,t+1}^{\tau_{t+1}} I_{[k \neq J+1]}) \right) \Big| s_{i,t}, d_{i,t} = j \right] - \beta v_{j,t}^{\tau}$$

Results – Static versus Dynamic

Table: 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	$200,00
Percent White	1627.03	1901.43	2221.66	612.09	2428.93	4888.4
	(12.92)	(18.67)	(48.56)	(87.09)	(121.89)	(275.0
Violent Crime	-291.14	-380.66	-448.88	-350.18	-962.20	-1298.8
	(8.13)	(10.98)	(19.40)	(47.68)	(67.29)	(92.7
Ozone	-66.24	-80.71	-97.04	-302.06	-380.03	-395.5
	(2.00)	(2.39)	(3.25)	(24.82)	(26.98)	(37.1

Static versus Dynamic

- Static and Dynamic estimates differ considerably
- Static underestimates MWTP for ozone, crime
- Race differentials vary with income.
- Bias is potentially quite large
- Finding direction of bias is an empirical question
  - Direction of bias is consistent with our hypothesis based on time series properties of attributes

Conclusion

- We develop a tractable model of neighborhood choice in a dynamic setting
  - Allow heterogeneity in valuation of both observed and unobserved neighborhood characteristics
  - No restriction on size of state space or size of choice set
  - Type (wealth) of individuals changes endogenously
  - Novel approach to dealing with endogeneity of prices
- We use information on neighborhood choice and the timing of moves to:
  - estimate preferences for neighborhood attributes
  - illustrate large biases from myopic model

More Results : Amenity Time Series

Table: Time-Series Properties of Amenities

	Δ Percent White$_{t+1}$	Δ Violent Crime$_{t+1}$	Δ Ozone$_{t+1}$
Percent White$_t$	**0.0156**	-0.4466	0.0089
	(0.0014)	(0.1123)	(0.0027)
Violent Crime$_t$	0.0010	**-0.1520**	-0.0016
	(0.0001)	(0.0083)	(0.0002)
Ozone$_t$	-0.0162	-0.5958	**-0.6566**
	(0.0105)	(0.8491)	(0.0201)
County Dummies	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes

Note: For the dependent variable, $\Delta X_{t+1} = X_{t+1} - X_t$.