Lecture 6: Modelo Hedónico

Urban Economics

Ignacio Sarmiento-Barbieri

Universidad de los Andes

September 10, 2025

Modelos

- Where do you want to live?
  - Spatial equilibrium
  - Within cities: Alonso-Muth-Mills (Monocentric/Polycentric Model)
  - Hedonic pricing of amenities and local public goods (Rosen)
  - Across locations: Rosen-Roback

Rosen's Hedonic Model

- Goods are valued for their utility-bearing attributes
- Heterogeneous or differentiated goods are products whose characteristics vary in such a way that there are distinct product varieties even though the product is sold in one market (e.g. houses, cars, computers, etc).
- The variation in product variety gives rise to variations in product prices within each market.
- The hedonic method relies on market transactions for these differentiated goods to determine the implied value or implicit price of characteristics.

The Consumer's Problem

- House: $z = (z_1, \ldots, z_n)$
- Price: $p(z) = p(z_1, \ldots, z_n)$
- Consumer utility is $U(x, z)$ where x is non-housing consumption
- The consumer buys one house and has budget $y = x + p(z)$
  - y denotes exogenous income
  - x denotes consumption of non-housing goods

The Consumer's Problem

The Consumer's Problem

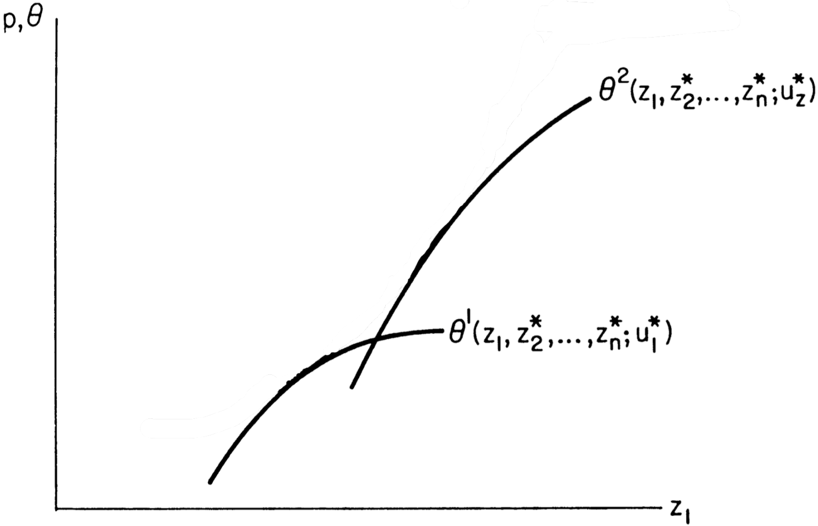
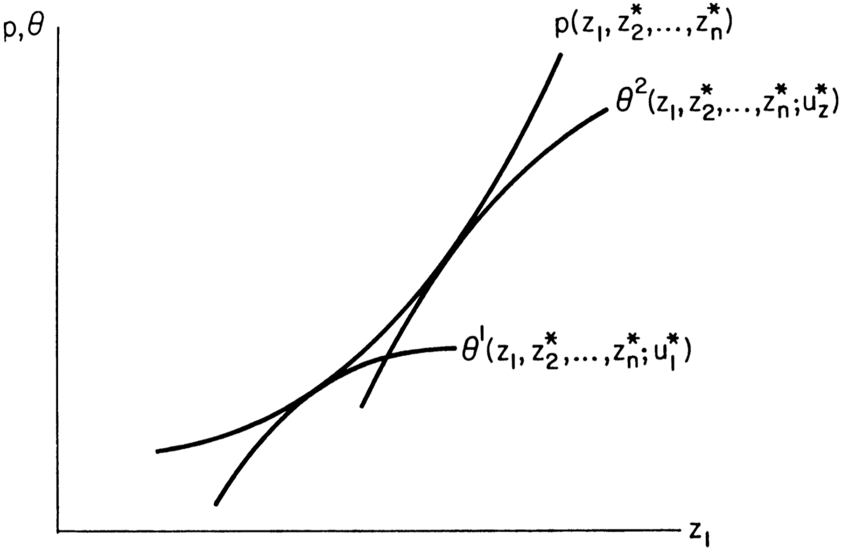


Fig. 1

The Consumer's Problem



The Producer's Problem

- Each firm produces a specific bundle of attributes $z = (z_1, \ldots, z_n)$
- Production costs are $C(M, z, \beta)$ where
  - $M(z)$ denotes number of units produced of designs offering specification z
  - Producers have different technologies parametrized by β
- The firm is a price taker $p(z)$ and maximizes profits

$$\pi = Mp(z) - C(M, z) \tag{1}$$

The Producer's Problem

The Producer's Problem

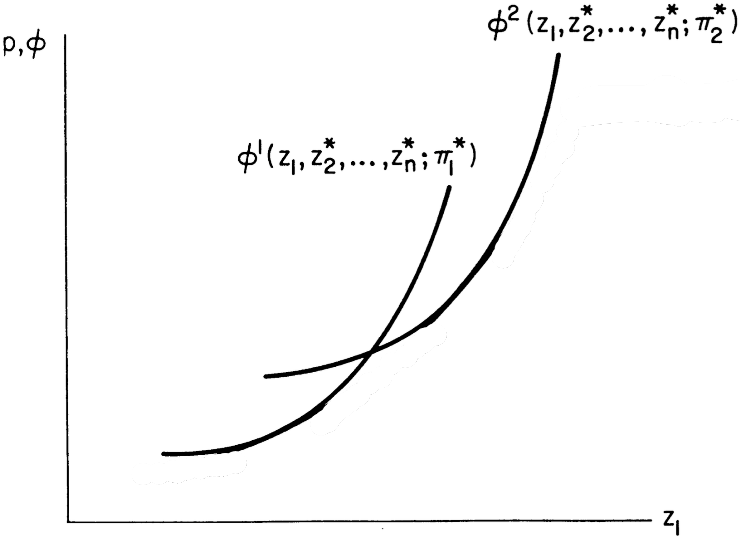
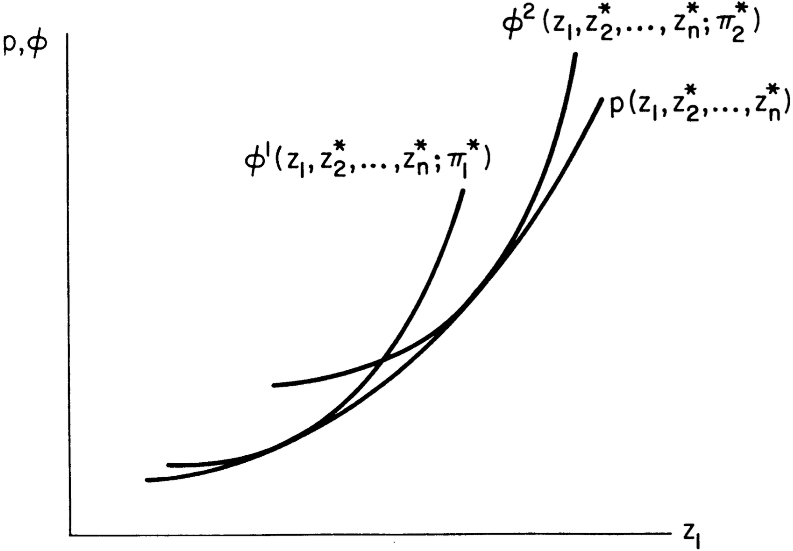


FIG. 2

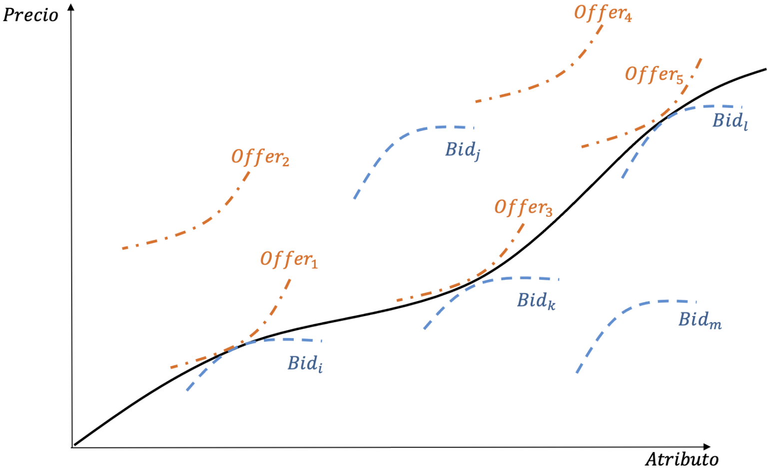
The Producer's Problem



Market Equilibrium

- The market hedonic function $p(z)$ is a joint envelope
  - Upper envelope of consumer's bid functions
  - Lower envelope of producer's offer functions
- Quantities demanded and supplied at each z depend on all of $p(z)$

Market Equilibrium



Empirics

- Rosen (1974) proposed a two-step empirical strategy
  1. Estimate hedonic prices $p(z)$ with the best fitting functional form
  2. Take partial derivatives of the estimate $\hat{p}(z)$ at the sample values and estimate the simultaneous demand and supply equations

$$\frac{\partial p}{\partial z_i} = F_i(z, x^d, y - p(z)) \tag{2}$$

$$\frac{\partial p}{\partial z_i} = G_i(z, x^s, p(z)) \tag{3}$$

Empirics

- Rosen (1974) proposed a two-step empirical strategy
  1. Estimate hedonic prices $p(z)$ with the best fitting functional form
  2. Take partial derivatives of the estimate $\hat{p}(z)$ at the sample values and estimate the simultaneous demand and supply equations

$$\frac{\partial p}{\partial z_i} = F_i(z, x^d, y - p(z)) \tag{2}$$

$$\frac{\partial p}{\partial z_i} = G_i(z, x^s, p(z)) \tag{3}$$

- Problems?

Empirics
Solutions

- Bartik (1987): exogenous shifts in the consumer's budget constraint
  - Exogenous income changes if you can find them (field experiments)
- Urban economists have mostly shied away from structural estimation
  - Stop at the first-step hedonic regression
  - Focus on omitted-variable bias

Hedonic Price Function

$$P_i = \beta_0 + \sum_{j=1}^{h} \beta_j H_{ij} + \sum_{k=1}^{n} \beta_k N_{ik} + u_i \tag{4}$$

- P the sales price of a house;
- H represents structural and property characteristics of the house (e.g. square footage of the living area, lot size, etc)
- N represents location characteristics, (e.g. quality of schools,distance to CBD, parks, etc.)

Empirics

Example: Linden and Rockoff (2008)

American Economic Review 2008, 98:3, 1103–1127
http://www.aeaweb.org/articles.php?doi=10.1257/aer.98.3.1103

Estimates of the Impact of Crime Risk on Property Values from Megan's Laws

By Leigh Linden and Jonah E. Rockoff*

Empirics

Example: Linden and Rockoff (2008)

1112 THE AMERICAN ECONOMIC REVIEW JUNE 2008

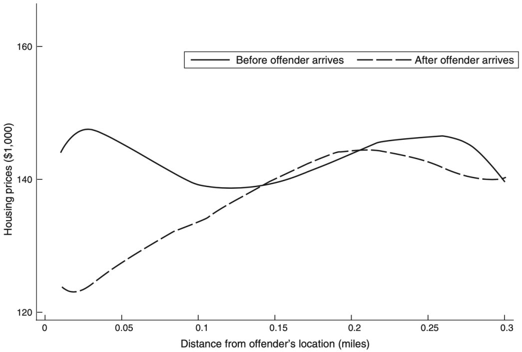


FIGURE 2B. PRICE GRADIENT OF DISTANCE FROM OFFENDER
(Sales during year before and after arrival)

Note: Results from local polynomial regressions (bandwidth = 0.075 miles) of sale price on distance from offender's future/current location.

Source: Estimates of the impact of crime risk on property values from Megan's laws Linden and Rockoff (2008)

Empirics

Example: Linden and Rockof (2008)

1116 THE AMERICAN ECONOMIC REVIEW JUNE 2008

TABLE 3—IMPACT OF SEX OFFENDERS' LOCATIONS ON PROPERTY VALUE AND SALE PROBABILITY

	Log (sale price) pre-arrival		Log (sale price), pre- and post-arrival				Probability of sale†
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Within 0.1 miles of offender	−0.340 (0.052)*	−0.007 (0.013)	−0.007 (0.012)	<0.001 (0.013)	−0.006 (0.012)	−0.006 (0.012)	−0.029 (0.035)
Within 0.1 miles × post-arrival			−0.033 (0.019)+	−0.041 (0.020)*	−0.036 (0.021)+	−0.116 (0.059)+	0.126 (0.059)*
Dist*≤ 0.1 miles × post-arrival (0.1 Miles = 1)						0.107 (0.064)+	
Within 1/3 miles of offender				−0.010 (0.007)			
Within 1/3 miles × post-arrival				0.010 (0.010)	0.003 (0.016)	0.004 (0.016)	−0.055 (0.040)
H_0: within 0.1 miles × post-arrival = 0			*p-value* = 0.079	*p-value* = 0.0443	*p-value* = 0.0828	*p-value* = 0.0502	*p-value* = 0.0361
Housing characteristics		✓	✓	✓	✓	✓	✓
Year fixed effects	✓						
Neighborhood-year fixed effects		✓	✓	✓			
Offender area-year fixed effects					✓	✓	✓
Restricted to offender areas 2 years pre- and post-arrival					✓	✓	✓
Standard errors clustered by…	*Neighbor-hood*	*Neighbor-hood*	*Neighbor-hood*	*Neighbor-hood*	*Offender area*	*Offender area*	*Offender area*
Sample size	164,993	164,968	169,557	169,557	9,086	9,086	1,519,364
R^2	0.01	0.84	0.83	0.83	0.75	0.75	0.01

Note: Pre-arrival (post-arrival) refers to the two-year period before (after) the date upon which offenders registered their current address. Standard errors in parentheses.

* Significant at 5 percent level.

+Significant at 10 percent level.

† Probability sale is measured as percentage points, e.g., probability of sale + 1 would be 100 percentage points.

Source: Estimates of the impact of crime risk on property values from Megan's laws Linden and Rockof (2008)

Empirics

Example: Currie et al (2015) AER

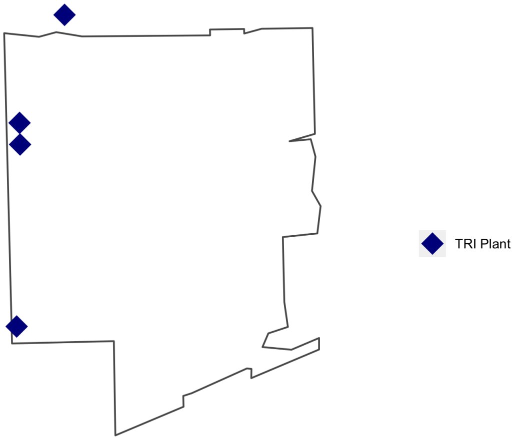
American Economic Review 2015, 105(2): 678–709
http://dx.doi.org/10.1257/aer.20121656

Environmental Health Risks and Housing Values: Evidence from 1,600 Toxic Plant Openings and Closings†

By Janet Currie, Lucas Davis, Michael Greenstone, and Reed Walker*

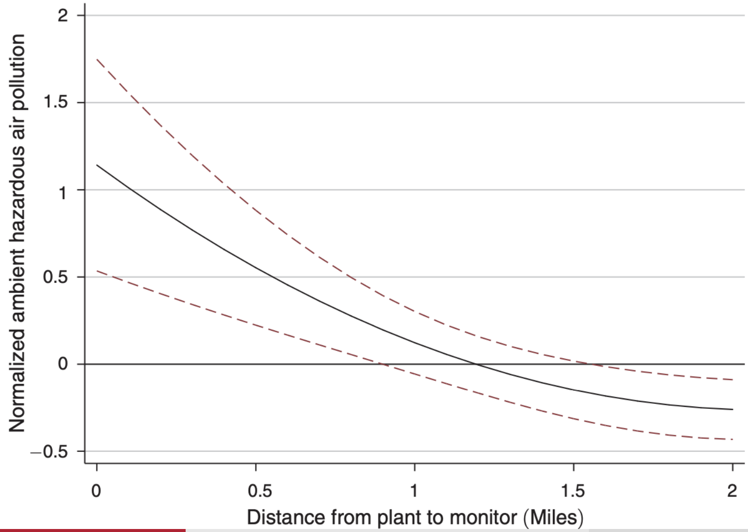
Empirics

Zip Code with TRI Toxic Plants within one mile



Empirics

Example: Currie et al (2015) AER



Empirics

Example: Currie et al (2015) AER

TABLE 2—THE EFFECT OF TOXIC PLANTS ON LOCAL HOUSING VALUES

	0–0.5 Miles		0.5–1 Miles		0–1 Miles		0–1 Miles (+/– 2 years)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel C. First difference: Estimated effect of plant openings and closings								
1(Plant Opening) × Near	−0.096***	−0.107***	−0.007	−0.008	−0.020	−0.022	−0.030	−0.038
	(0.036)	(0.034)	(0.023)	(0.020)	(0.022)	(0.019)	(0.028)	(0.025)
1(Plant Closing) × Near	0.017	0.010	0.008	0.003	0.010*	0.005	0.005	0.001
	(0.011)	(0.009)	(0.005)	(0.004)	(0.006)	(0.005)	(0.007)	(0.005)
H_0: Opening = −Closing (p-value)	0.051	0.013	0.968	0.827	0.688	0.438	0.402	0.164
Observations	1,114,248	1,114,248	1,305,780	1,305,780	1,375,751	1,375,751	1,196,000	1,196,000
State × year fixed FE	X		X		X		X	
County × year FE		X		X		X		X

Empirics

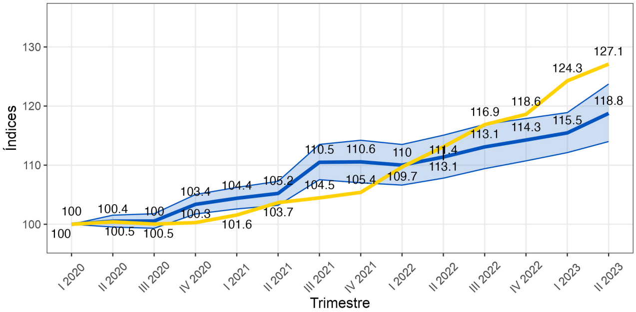
Example: Currie et al (2015) AER

TABLE 4—THE EFFECT OF TOXIC PLANTS ON LOW BIRTHWEIGHT

	0–0.5 Miles		0.5–1 Miles		0–1 Miles		0–1 Miles (+/− 2 years)	
Panel B. Estimated effect of plant openings and closings								
1(Plant Opened)	0.0025	0.0022	0.0024***	0.0027***	0.0024**	0.0024***	0.0031*	0.0037**
× Near	(0.0019)	(0.0018)	(0.0009)	(0.0010)	(0.0009)	(0.0008)	(0.0017)	(0.0017)
1(Plant Closed)	−0.0002	−0.0007	−0.0009	−0.0009	−0.0007	−0.0009	−0.0016	−0.0021*
× Near	(0.0016)	(0.0016)	(0.0009)	(0.0010)	(0.0009)	(0.0009)	(0.0012)	(0.0013)
H_0: Opening = −Closing (p-value)	0.44	0.56	0.32	0.28	0.22	0.24	0.51	0.48
Observations	88,958	88,958	88,958	88,958	88,958	88,958	63,324	63,324
Plant count	3,438	3,438	3,438	3,438	3,438	3,438	3,438	3,438
Plant × Distance-bin FE	X	X	X	X	X	X	X	X
State × Year FE	X		X		X		X	
Plant × Year FE		X		X		X		X

House Price Indices

- The most common approaches for constructing house price indices are hedonic price functions and the repeat sales estimator



House Price Indices

- Both estimators
  - Focus on changes over time in mean prices.
  - Are designed to measure the expected sale price of a home after controlling for features of the structure and location.
- Hedonic price functions control directly for structural and locational characteristics by including them as explanatory variables in the estimated regression.
- The repeat sales approach controls for these variables more indirectly by restricting the analysis to properties that sold at least twice during the sample period.

Hedonic price functions

$$log(P)_{it} = \sum_{t=1}^{T} \delta_t D_{it} + \sum_{j=1}^{h} \beta_j H_{ij} + \sum_{k=1}^{n} \beta_k N_{ik} + u_{it} \tag{5}$$

- $log(P)_{it}$ represents the natural logarithm of the sale price of house i at time t $(t = 1, \dots, T)$,
- D_{it} is a variable indicating the house i sold at time t,
- H represents structural and property characteristics of the house (e.g. square footage of the living area, lot size, etc)
- N represents location characteristics, (e.g. quality of schools,distance to CBD, parks, etc.)

Hedonic price functions

- Advantages
  - Produces a quality-adjusted index by including controls for observable characteristics of the house and location
  - Leverages large sample sizes available for home sales over time.

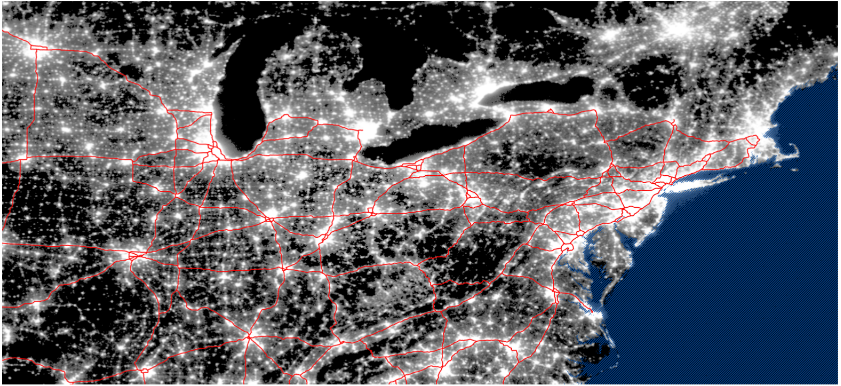
Hedonic price functions
Problems

- Market definition

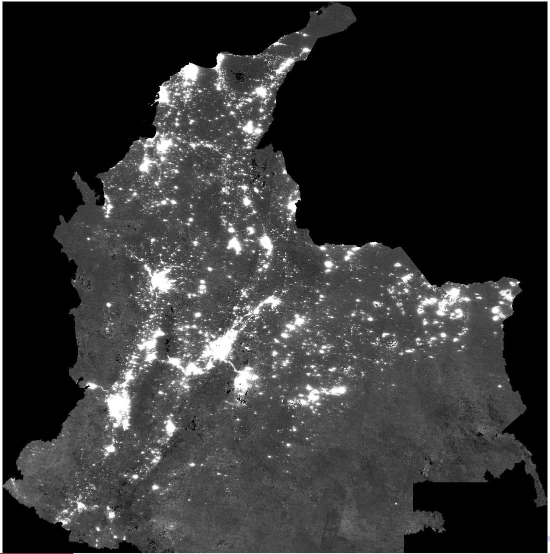
Hedonic price functions

Problems

Figure 1: Lights at night and Interstate Highways in the Northeast US, 2007.



Hedonic price functions
Problems



Hedonic price functions

Problems

- Market definition
- The model's conceptual logic implies that the market should be chosen to satisfy the

  "law of one price function"

  - That is, when a house can be fully defined by a unique bundle of physical characteristics and location-specific amenities, then equivalent bundles sell for the same price throughout that market.
  - A common approach is to define the market as a single metropolitan area over a few years (e.g. Pope 2008b, Abbott and Klaiber 2013).

Hedonic price functions

Problems

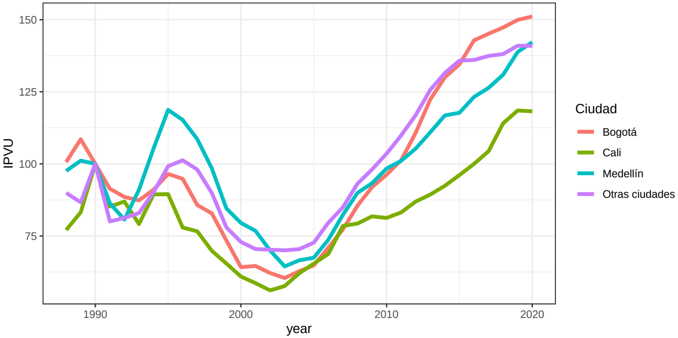
- Functional form
  - Need to specify a functional form and the possibility that omitted variables may affect the results. (Functional form misspecification can also be viewed as a form of omitted variables.)
    - Theoretical and simulation evidence suggest that the hedonic price function should be assumed to be nonlinear (Ekeland, Heckman, and Nesheim, 2004).
    - Kuminoff, Parmeter, and Pope (2010) found that flexible Box-Cox models continue to outperform linear models

Hedonic price functions

- Sample selection
  - A key question in price indices is whether D_{it} are correlated with u_{it}
    - If the sample of sales in a time period tends to be composed of unusually high- or low-quality homes, i.e., homes with high or low average values for X, then the effect of the missing variables will be to bias the estimated value of δ_t upward or downward.

Repeat sales estimator

- Repeat sales approach attempts to control for the effects of missing variables by restricting the analysis to the set of homes that sold at least twice during the sample period (Bailey, Muth and Nourse 1963, Case and Shiller 1987, 1989).



Repeat sales estimator

- The Indice de Precios de la Vivienda Usada (IPVU) estimated by BanRep uses the repeat sales approach proposed by Case and Shiller (1989)
- Built in 3 stages using properties sold at least twice
- Case and Shiller (1989) model the price behavior as

$$\ln(P_{it}) = \beta_t + H_{it} + N_{it} \tag{6}$$

  - β_t is the log of the citywide level of housing prices at time t
  - H_{it} is a gausian random walk (where ΔH_{it} has zero mean and variance σ_h^2) and uncorrelated with C_t. Represents the deviation of a particular house from the market price
  - $N_{it} \sim_{iid} N(0, C)$ are idiosincratic differences