Simple, Consistent Estimation of the Marginal Willingness to Pay Function: Recovering Rosen's Second Stage Without Instrumental Variables

Kelly C. Bishop
Olin Business School

Christopher Timmins
Duke University and NBER

Camp Resources XV, Wilmington, NC

August 7, 2008

Motivation

- Property value hedonics allows us to recover the value of a marginal change in Z:

$$P_{j,k} = \beta_0 + \beta_1 Z_{j,k} + \beta_2 Z_{j,k}^2 + X_{j,k}'\alpha + \epsilon_{j,k}$$

$$\rho_{j^*(i),k} = \left.\frac{\partial P_{j,k}}{\partial Z_{j,k}}\right|_{j=j^*(i)}$$

- Need to recover MWTP function to value non-marginal changes.
- May want to consider preference heterogeneity.
- Potential for omitted variable bias.

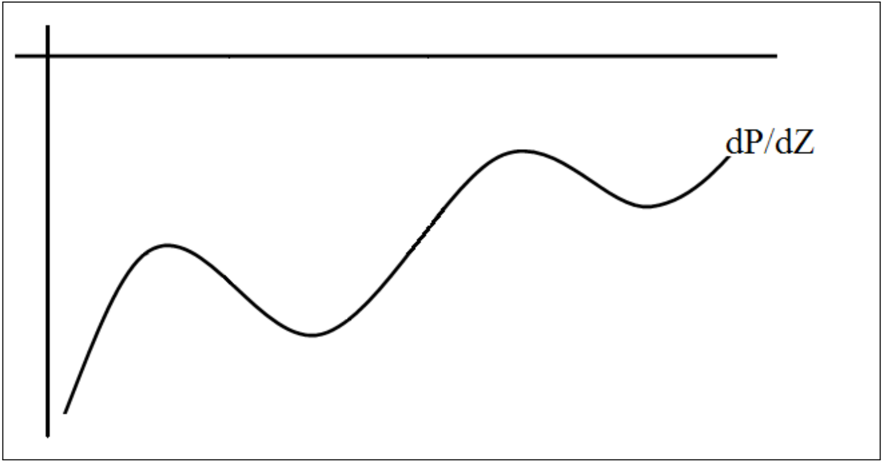
Motivation - Rosen (1974)

- Second-stage regression of $\rho_{j^*(i),k}$ on Z and individual attributes.
- Identification problem [Brown and Rosen (1982), Mendelsohn (1985)].
  - Use multi-market data
  - Estimate a non-parametric first-stage [Heckman, Ekeland, and Nesheim (2004)]
- Endogeneity problem [Epple (1987), Bartik (1987)].
  - Instrumental variables [Palmquist (1984), Chattopadhyay (1999), Boyle, Poor, and Taylor (1999)]

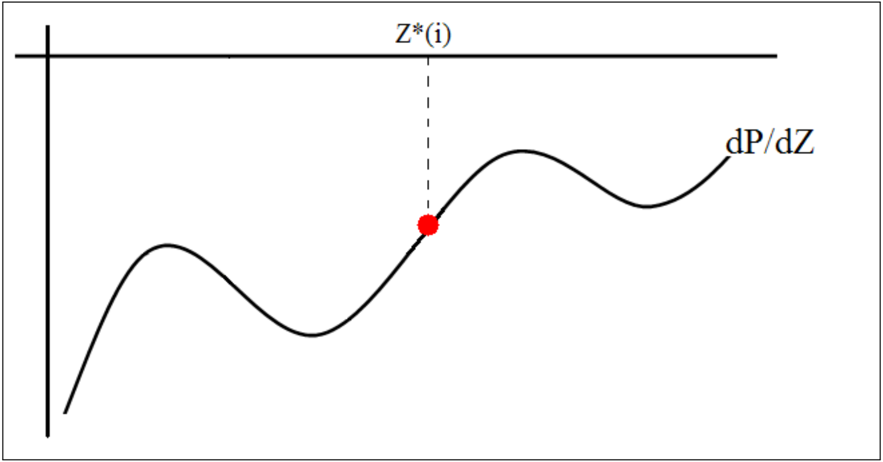
Motivation - Bajari and Benkard (2005)

- Begins by estimating a non-parametric first-stage gradient.
- Second-stage preference inversion method that avoids endogeneity problems of Rosen (1974).
- Recovers individual-specific utility-function parameters from FOCs.
- Requires strong functional form assumptions on utility.
- MWTP is neither a function of Z nor a function of expenditure.

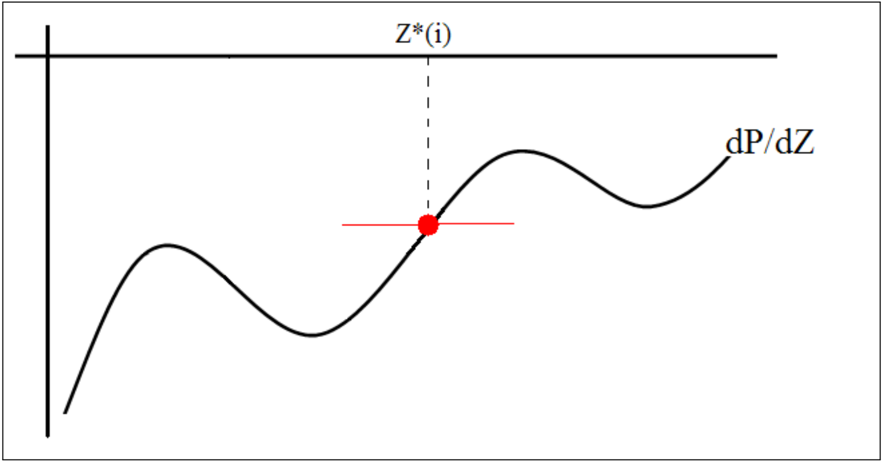
Motivation - Bajari and Benkard (2005)



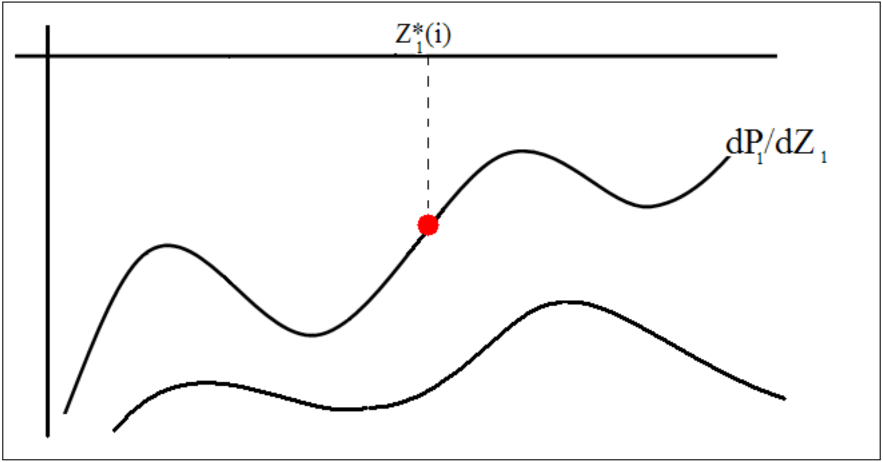
Motivation - Bajari and Benkard (2005)



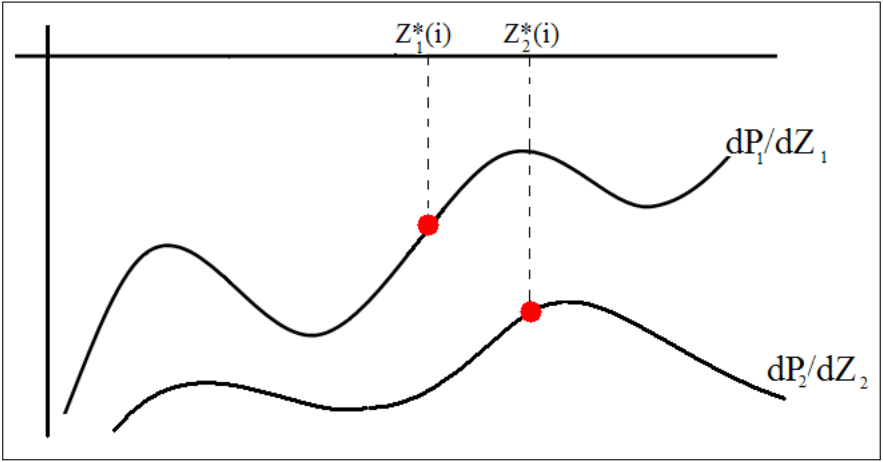
Motivation - Bajari and Benkard (2005)



Motivation - Bajari and Benkard (2005) with Panel Data

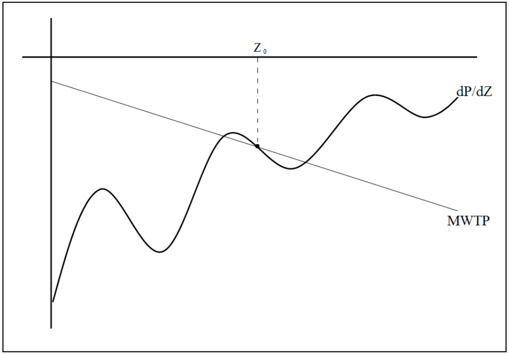


Motivation - Bajari and Benkard (2005) with Panel Data



Motivation - Bajari and Benkard (2005) with Panel Data

- Requires strict assumptions on stability of preferences over time
- HUGE data requirements
- Utility minimization (!)



Data

Property Transactions Panel (Dataquick)

- Bay Area of California 1990 - 2004
- Unique property ID along with housing characteristics
- Buyer name and Seller name

Data

	Full Sample		Repeat-Sale Sample	
	properties = 594,665		*properties = 191,210*	
variable	mean	median	mean	median
Price (2000 $)	377,684	330,000	377,158	336,673
Sq. Ft. House	1,696.95	1,500.00	1,589.71	1,432.00
Sq. Ft. Lot	9,611.62	5,500.00	8,384.23	5,044.00
Year Built	1966.48	1970	1968.60	1972
Total Rooms	6.38	6	6.21	6
Num. Bedrooms	2.94	3	2.86	3
Num. Bath	2.02	2	1.99	2
Days $\geq$ 90 PPB Ozone	2.44	2.09	2.54	2.09

Data

Individual Characteristics

- HMDA data
- Gender, race, income, down payment
- Transaction date, Census tract, lender name, loan amount

Data

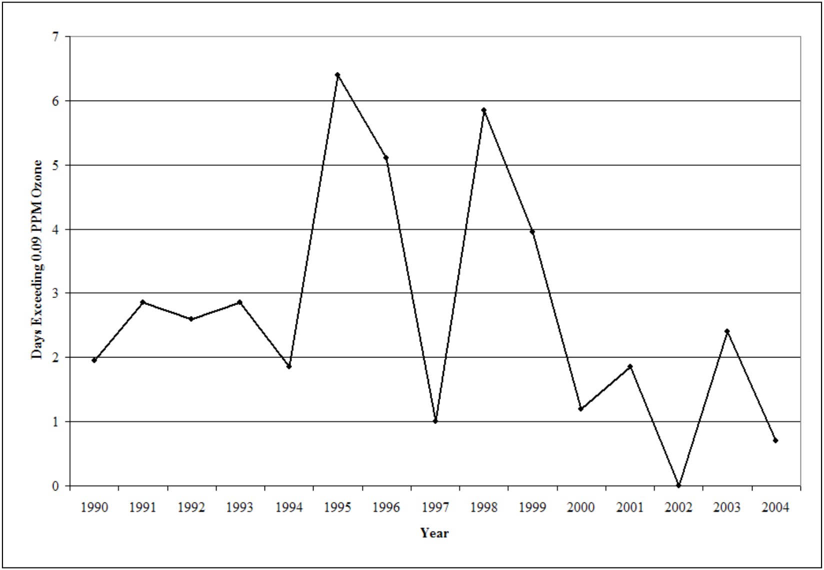
	Full Sample		2-Purchase		3-Purchase	
	ind = 40,092		*ind = 9,805*		*ind = 1,076*	
variable	mean	median	mean	median	mean	median
Asian	0.21	0	0.22	0	0.21	0
Black	0.02	0	0.02	0	0.02	0
Hispanic	0.09	0	0.10	0	0.10	0
White	0.59	1	0.58	1	0.58	1
Other	0.09	0	0.08	0	0.09	0
Income	123,033	98,993	124,805	101,324	131,637	101,826
Price	409.943	344.778	431,786	365,997	437,937	354,653

Data

Ozone Data - CA Air Resources Board

- 37 monitors - days exceeding 90 PPB, monitor coverage
- Geographic coordinates
- Calculate a weighted average for each house ($\frac{1}{dist^2}$)

Data



Days Exceeding 0.09 PPM Ozone
0
1
2
3
4
5
6
7
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
Year

First Stage - Estimation of the Hedonic Gradient

Estimate a non-parametric hedonic gradient that controls for house fixed-effects:

$$P_{j,t} = f(Z_{j,t}) + \xi_j + \nu_{j,t}$$

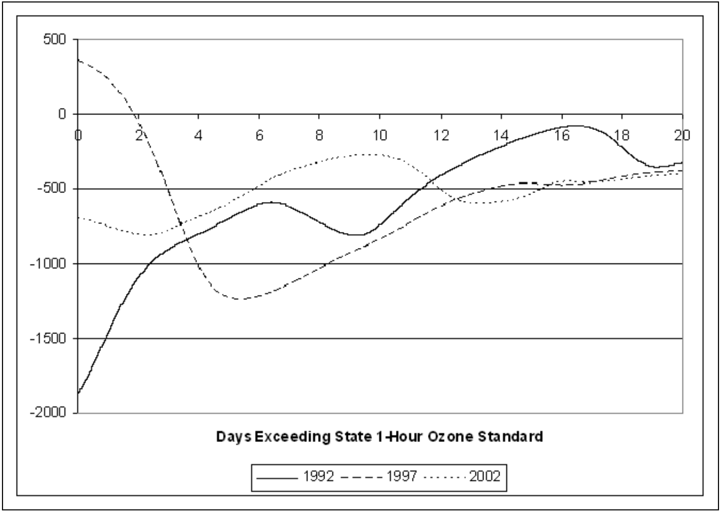
where $f(\cdot)$ is an unspecified, flexible function of $Z_{j,t}$.

Estimation Strategy

- Mean-difference over observations of property j to eliminate ξ_j
- Local linear estimation $\rightarrow$

  OLS at every observed level of $Z_{j,t}$ and smooth results

First Stage - Estimation of the Hedonic Gradient



Second Stage - Estimation of the MWTP Function

Begin with indirect utility:

$$\begin{aligned} V_{i,j,t} &= \alpha_{0,i} + \alpha_{1,i,t} X_j + \alpha_{2,i} X_j^2 + \alpha_{3,i,t} Z_{j,t} + \alpha_{4,i,t} Z_{j,t}^2 \\ &\quad + \alpha_{5,i} Z_{j,t}(I_{i,t} - R_t(X_j, Z_{j,t})) + (I_{i,t} - R_t(X_j, Z_{j,t})) \end{aligned}$$

Specify preference heterogeneity as:

$$\begin{aligned} \alpha_{3,i,t} &= \alpha_{3,0} + \alpha_{3,1} A_i + \epsilon_{i,t} \\ \alpha_{4,i,t} &= \alpha_{4,0} + \alpha_{4,1} A_i + u_{i,t} \\ \alpha_{5,i} &= \alpha_{5,0} + \alpha_{5,1} A_i \end{aligned}$$

Second Stage - Estimation of the MWTP Function

First Order Condition:

$$\epsilon_{i,t} = \frac{\partial R_t}{\partial Z_{j,t}}(1 + \alpha_{5,i} Z_{j,t}) - \Big(\alpha_{3,0} + \alpha_{3,1} A_i + 2(\alpha_{4,0} + \alpha_{4,1} A_i + u_{i,t}) Z_{j,t} +$$

$$\alpha_{5,i}(I_{i,t} - R_t(X_j, Z_{j,t}))\Big)$$

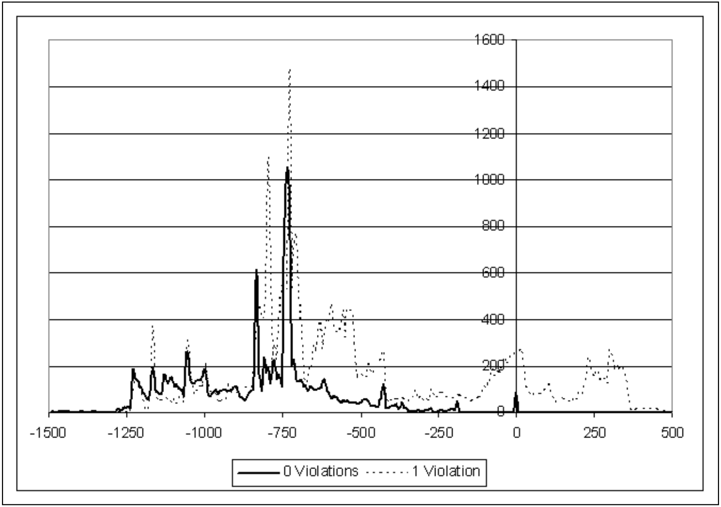
Second Order Condition:

$$u_{i,t} < \alpha_{5,i} \frac{\partial R_t}{\partial z_{j,t}} - \alpha_{4,0} - \alpha_{4,1} A_i + \frac{1}{2} \frac{\partial^2 R_t}{\partial z_{j,t}^2} \left(1 + \alpha_{5,i} z_{j,t}\right) = \Gamma_{i,t}$$

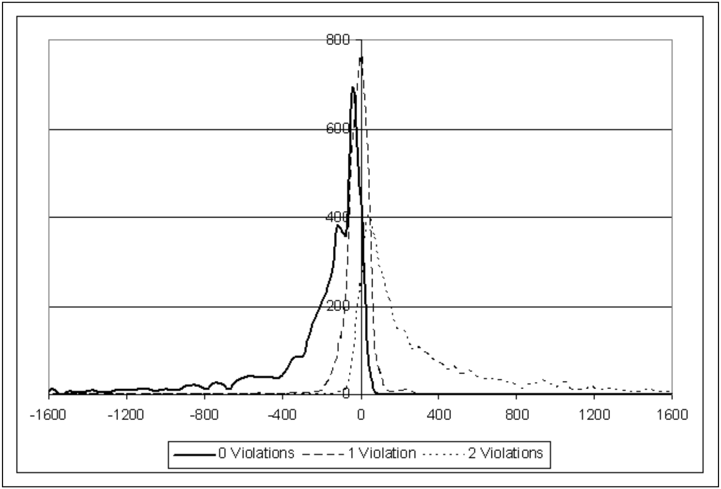
Estimation Strategy

- Integrate out over feasible range of $u_{i,t}$ conditional on $Z_{j^*(i),t}$
- Maximize the likelihood of observing $Z_{j^*(i),t}$

Results - Cross Sectional Inversion



Results - Panel Inversion

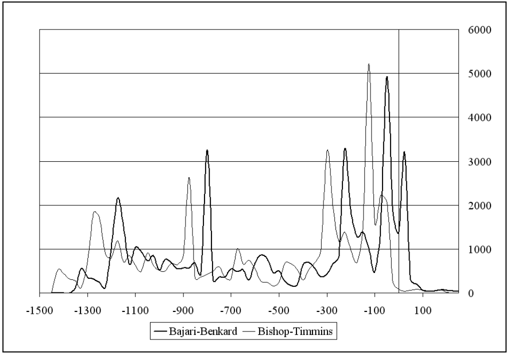


Results - ML Estimator

parameter		value	s.e.	t-stat.
intercept	α_3	200.96	3.45	58.20
ozone	α_4	-46.96	0.81	-57.96
expenditure	α_5	-0.30E-02	0.59E-04	-51.48
stdev(ϵ)	σ_ϵ	466.13	5.39	86.44
stdev(u)	σ_u	36.00	0.75	47.73

Consider a One Day Increase in Ozone...

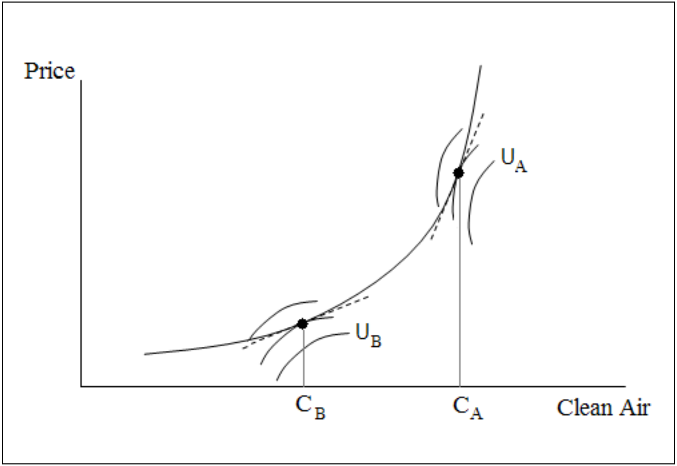
- Bajari and Benkard yields smaller damage estimate (17% less)
- Average MWTP masks considerable heterogeneity ($860.94 at median house)



Summary

- Estimate heterogeneous preferences for clear air.
- Two-stage approach that combines the insights of Rosen and Bajari-Benkard.
- Avoid endogeneity bias of Rosen.
- Doesn't require strict functional form assumptions as Bajari and Benkard.
- Doesn't require strict assumption of time-invariant preferences.
- Simple to estimate.
- Estimate with cross-section of buyers.

Motivation



Motivation