Hedonic Prices and Implicit Markets: Estimating Marginal Willingness to Pay for Differentiated Products Without Instrumental Variables

by

Kelly Bishop
Olin Business School
Washington University in St. Louis

Christopher Timmins
Duke University and NBER

November 2011

Since Rosen (1974), property value hedonic models have become the workhorse model for the non-market valuation of local public goods and amenities.


Suggests two-stage procedure for recovering marginal willingness to pay (MWTP) function.

---

1 4734 cites on Google Scholar.

(1) Use hedonic price gradient to recover implicit amenity prices.

\[ P_{i,k} = \beta_{0,k} + \beta_{1,k}Z_{i,k} + \beta_{2,k}Z_{i,k}^2 + \varepsilon_{i,k} \]

\[ P_{i,k}^Z = \beta_{1,k} + 2\beta_{2,k}Z_{i,k} \]

(2) Use implicit prices as dependent variables in MWTP estimation.

\[ P_{i,k}^Z = \alpha_{0,k} + \alpha_{1,k}Z_{i,k} + \alpha_{2,k}X_i^d + \nu_{i,k} \]
Motivation: *Rosen (1974)*

(5) Use area under MWTP (demand) curve to calculate welfare effect of an increase or decrease in $Z$. 

![Diagram showing area under MWTP curve](image)
Motivation: *Rosen (1974)*

Problems:

- omitted variables bias
- MWTP identification
- MWTP endogeneity
- stickyness/initial conditions
- re-sorting in response to a non-marginal change
- endogenous attributes (determined by sorting equilibrium)
- choice set definition
- dynamics

We will be primarily concerned with MWTP endogeneity (Epple, 1987; Bartik, 1987).
Motivation: Bartik-Epple

Hedonic price of product attribute varies systematically with the quantity consumed (unless hedonic price function is linear).

\[ P_{i,k}^Z = \alpha_{0,k} + \alpha_1 Z_{i,k} + \alpha_2 X_i^d + \nu_{i,k}^d \]

Those with strong idiosyncratic tastes for Z will both consume a lot of it and pay a high hedonic price for it.

\[ E[Z_{i,k} \nu_{i,k}^d] \neq 0 \]
Bartik-Epple: *Proposed Solutions*

(1) Instrumental Variables Strategies

Epple (1987): Supplier attributes do not constitute valid instruments, since those with strong idiosyncratic tastes will sort to suppliers who can supply Z more cheaply.

Instead, instrumental variables strategies were based on…

- non-linearity or difficult-to-justify exclusion restrictions
- cross-market preference homogeneity assumptions

(Bartik, 1987)
“To date no hedonic model with site specific environmental amenities has successfully estimated the second stage marginal willingness to pay function.” (Deacon et al., 1998)
Bartik-Epple: *Proposed Solutions*

(2) Bajari-Benkard (2005)

- convert second-stage estimation into “preference inversion” under strong parametric restrictions on utility
- assumes value of MWTP elasticity


- additive separability assumption
- exclusion restrictions based on non-linearity
Motivation: *Bartik-Epple*

Difficulty in dealing with this problem has led researchers to forego estimating MWTP curves altogether. Restrict attention to first-stage of Rosen’s hedonic procedure:

Only appropriate for valuing **marginal** changes in amenities. Very few of the policies being considered are actually marginal.
Motivation: Bartik-Epple

- Black (1999)
- Figlio & Lucas (2004)
- Davis (2004)
- Chay & Greenstone (2005)
- Gayer, Hamilton & Viscusi (2000)
- Greenstone & Gallagher (2008)
- Linden & Rockoff (2008)
- Pope (2008)
- Davis (2010)
- Gamper-Rabindran & Timmins (2011)
- Bajari, Cooley, Kim & Timmins (2011)
Motivation: *Bartik-Epple*

Propose a new estimation procedure based on the notion of “stratification” found in Ellickson (1971) – systematic relationship between quantity of amenity being consumed and attributes of those doing the consumption.

Use new estimation procedure to recover welfare effects of large reductions in violent crime in San Francisco Bay Area observed in 1990’s.
Identification: *Linear Gradients, Two Markets*

Begin by showing intuition in a simple case (linear hedonic gradient, two markets):

(1) Hedonic Price Function:

\[
P_{i,k} = \beta_{0,k} + \beta_{1,k}Z_{i,k} + \beta_{2,k}Z_{i,k}^2 + \varepsilon_{i,k}
\]

\[
P_{i,k}^Z = \beta_{1,k} + 2\beta_{2,k}Z_{i,k}
\]

(2) MWTP Function:

\[
P_{i,k}^Z = \alpha_{0,k} + \alpha_{1}Z_{i,k} + \alpha_{2,k}X_{i,d} + \nu_{i,k}
\]
Identification: *Linear Gradients, Two Markets*

(3) Hedonic Equilibrium (Stratification Equation):

\[
Z_{i,k} = \frac{\alpha_{0,k} - \beta_{1,k}}{2\beta_{2,k} - \alpha_1} \pi_{0,k} + \frac{\alpha_{2,k}}{2\beta_{2,k} - \alpha_1} \pi_{1,k} X_i^d + \frac{1}{2\beta_{2,k} - \alpha_1} u_{i,k}^d
\]
Identification: *Linear Gradients, Two Markets*

(4) With two markets, a system of six equations in six unknowns:

\[
\hat{\pi}_{0,1} = \frac{\alpha_{0,1} - \beta_{1,1}}{2\beta_{2,1} - \alpha_1} \quad \hat{\pi}_{0,2} = \frac{\alpha_{0,2} - \beta_{1,2}}{2\beta_{2,2} - \alpha_1}
\]

\[
\hat{\pi}_{1,1} = \frac{\alpha_{2,1}}{2\beta_{2,1} - \alpha_1} \quad \hat{\pi}_{1,2} = \frac{\alpha_{2,2}}{2\beta_{2,2} - \alpha_1}
\]

\[
\hat{\sigma}_{u_1} = \frac{\sigma_v}{2\beta_{2,1} - \alpha_1} \quad \hat{\sigma}_{u_2} = \frac{\sigma_v}{2\beta_{2,2} - \alpha_1}
\]
Identification: *Linear Gradients, Two Markets*

(5) Solve for structural parameters:

\[
\alpha_1 = \frac{2(\hat{\sigma}_{u2}\hat{\beta}_{2,2} - \hat{\sigma}_{u1}\hat{\beta}_{2,1})}{\hat{\sigma}_{u2} - \hat{\sigma}_{u1}}
\]

\[
\alpha_{0,1} = \hat{\beta}_{1,1} + (2\hat{\beta}_{2,1} - \hat{\alpha}_1)\hat{\pi}_{0,1} \quad \alpha_{0,2} = \hat{\beta}_{1,2} - (2\hat{\beta}_{2,2} - \hat{\alpha}_1)\hat{\pi}_{0,2}
\]

\[
\alpha_{2,1} = (2\hat{\beta}_{2,1} - \hat{\alpha}_1)\hat{\pi}_{1,1} \quad \alpha_{2,2} = (2\hat{\beta}_{2,2} - \hat{\alpha}_1)\hat{\pi}_{1,2}
\]
Identification:  *Non-Linear Gradient, N Markets*

(1) Hedonic Gradient:  \[ P_i^Z = f(Z; \varphi) \]

(2) MWTP Function:  \[ P_i^Z = \alpha_0 + \alpha_1 Z_i + \alpha_2 X_i^d + v_i^d \]

(3) Hedonic Equilibrium:

\[ v_i^d = f(Z; \varphi) - (\alpha_0 + \alpha_1 Z_i + \alpha_2 X_i^d) \]
Identification: *Non-Linear Gradient, N Markets*

(4) Maximum Likelihood:

\[
\max \{\bar{\alpha}, \sigma\} \prod_{i=1}^{n} l(Z_i, X_i^{d}; \bar{\alpha}, \sigma)
\]

\[
l(Z_i, X_i^{d}; \bar{\alpha}, \sigma) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left\{ -\frac{1}{2\sigma^2} v_i^{d2} \right\} \left| \frac{\partial v_i^{d}}{\partial Z_i} \right|
\]

\[
\frac{dv_i^{d}}{dz_i} = f'(Z; \varphi) - \alpha_1
\]
Identification: Non-Linear Gradient, N Markets
Data: *Crime*

- RAND California Data Base
- Incidents per 100,000 residents
- Imputed for each house using \( \left( \frac{1}{distance^2} \right) \) weights
- 80 cities in the SF Bay Area

- Violent Crime: crimes against people, including homicide, forcible rape, robbery, and aggravated assault

- Property Crime: crimes against property, including burglary and motor vehicle theft
Map of California Counties
Figure 1: Locations of Crime-reporting Cities within the San Francisco Metro Area
Distribution of Violent Crime Rate (Incidents per 100,000)  
San Francisco Bay Area
Time Series of Violent Crime Rate (Incidents per 100,000)  
San Francisco and Los Angeles Metropolitan Areas
Data: Crime

Levitt (2004): Six factors not responsible for decline in crime rates

- economic growth and reduced unemployment
- shifting age and racial demographics
- changes in policing strategies
- changes in gun control/concealed weapon laws
- changes in capital punishment

He argues instead for role of:

- increasing size of the police force
- increased incarceration rates
- declines in the crack epidemic
- legalization of abortion twenty years prior
Data: *DataQuick Inc. Housing Transactions*

Housing transaction information:

- transaction dates
- latitude and longitude
- year built
- number of bedrooms
- number of stories
- loan amount loan date
- transaction prices
- square footage
- number of bathrooms
- number of total rooms
- mortgage lender

Cuts for:

- land sales
- excessive or depreciation
- rebuilds & major renovations
- missing values
Table 4: Property Transactions Summary Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price (constant 2000 dollars)</td>
<td>345,823.00</td>
<td>192,210.00</td>
</tr>
<tr>
<td>Year Built</td>
<td>1966.64</td>
<td>23.16</td>
</tr>
<tr>
<td>Lot Size (sq. ft)</td>
<td>6,453.55</td>
<td>7,927.96</td>
</tr>
<tr>
<td>Square Footage</td>
<td>1662.76</td>
<td>671.69</td>
</tr>
<tr>
<td>Number Bathrooms</td>
<td>2.07</td>
<td>0.73</td>
</tr>
<tr>
<td>Number Bedrooms</td>
<td>3.01</td>
<td>1.09</td>
</tr>
<tr>
<td>Property Crimes (per 100,000 residents)</td>
<td>1,803.43</td>
<td>771.48</td>
</tr>
<tr>
<td>Violent Crimes (per 100,000 residents)</td>
<td>445.69</td>
<td>241.19</td>
</tr>
</tbody>
</table>
Data: *Home Mortgage Disclosure Act*

HMDA (1975) collects information from mortgage applications on race, income, and sex of buyers. Determine whether financial institutions are serving the housing needs of their communities and look for evidence of discriminatory lending practices.

Provides data on the following variables, which can be matched to DataQuick:

- loan amount
- loan date
- lender name
- census tract

Match quality described in Bayer, MacMillan, Murphy and Timmins (2010).
Table 5: Buyer Summary Statistics (Full- and 1999- Samples)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Full Sample</th>
<th>1999 Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Dev.</td>
</tr>
<tr>
<td>Price</td>
<td>354,002.90</td>
<td>193,464.20</td>
</tr>
<tr>
<td>Violent Crime</td>
<td>445.46</td>
<td>242.24</td>
</tr>
<tr>
<td>Income</td>
<td>111,312.90</td>
<td>111,607.30</td>
</tr>
<tr>
<td>White</td>
<td>0.63</td>
<td>0.48</td>
</tr>
<tr>
<td>Asian</td>
<td>0.24</td>
<td>0.42</td>
</tr>
<tr>
<td>Black</td>
<td>0.04</td>
<td>0.19</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.10</td>
<td>0.30</td>
</tr>
</tbody>
</table>

Prices and incomes are expressed in constant 2000 dollars. The violent crime rate is per 100,000 residents.
Application: *Hedonic Price Functions*

\[ \ln p_{j,k,t} = \beta_{0,k,t} + \beta_{1,k,t} BUILT_{j,k,t} + \beta_{2,k,t} LOT_{j,k,t} + \beta_{3,k,t} SQFT_{j,k,t} + \beta_{4,k,t} BATH_{j,k,t} + \beta_{5,k,t} BED_{j,k,t} + \beta_{6,k,t} PC_{j,k,t} + \beta_{7,k,t} VC_{j,k,t} + TRACT'_{j,k,t} \psi_{k,t} + \varepsilon_{j,k,t} \]

*VC* = violent crime rate (incidents per 100,000 population)

*PC* = property crime rate (incidents per 100,000 population)

*BUILT* = year built

*SQFT* = square footage

*BATH* = number of bathrooms

*BED* = number of bedrooms

*TRACT* = vector of census tract dummy variables
<table>
<thead>
<tr>
<th>Year Built</th>
<th>Lot Size</th>
<th>Sq. Footage</th>
<th>Bathrooms</th>
<th>Bedrooms</th>
<th>Prop. Crime</th>
<th>Violent Crime</th>
</tr>
</thead>
<tbody>
<tr>
<td>1994</td>
<td>8.26E-04***</td>
<td>6.47061E-06***</td>
<td>3.45E-04***</td>
<td>6.00E-03**</td>
<td>0.0348***</td>
<td>-4.77E-05***</td>
</tr>
<tr>
<td></td>
<td>(9.36E-05)</td>
<td>(2.57963E-07)</td>
<td>(3.24E-06)</td>
<td>(2.77E-03)</td>
<td>(1.69E-03)</td>
<td>(5.81E-06)</td>
</tr>
<tr>
<td>1995</td>
<td>8.19E-04***</td>
<td>6.63532E-06***</td>
<td>3.49E-04***</td>
<td>9.51E-03***</td>
<td>0.0370***</td>
<td>-1.04E-05</td>
</tr>
<tr>
<td></td>
<td>(1.09E-04)</td>
<td>(2.95071E-07)</td>
<td>(4.76E-06)</td>
<td>(3.53E-03)</td>
<td>(1.93E-03)</td>
<td>(6.92E-06)</td>
</tr>
<tr>
<td>1996</td>
<td>7.41E-04***</td>
<td>6.53443E-06***</td>
<td>3.58E-04***</td>
<td>0.0136***</td>
<td>0.0432***</td>
<td>-1.64E-05**</td>
</tr>
<tr>
<td></td>
<td>(8.85E-05)</td>
<td>(3.3423E-07)</td>
<td>(4.29E-06)</td>
<td>(3.24E-03)</td>
<td>(1.69E-03)</td>
<td>(7.50E-06)</td>
</tr>
<tr>
<td>1997</td>
<td>4.40E-04***</td>
<td>7.12735E-06***</td>
<td>3.57E-04***</td>
<td>0.0186***</td>
<td>0.0430***</td>
<td>8.05E-07</td>
</tr>
<tr>
<td></td>
<td>(7.61E-05)</td>
<td>(2.64862E-07)</td>
<td>(4.62E-06)</td>
<td>(3.03E-03)</td>
<td>(1.67E-03)</td>
<td>(8.48E-06)</td>
</tr>
<tr>
<td>1998</td>
<td>2.34E-04***</td>
<td>7.05921E-06***</td>
<td>3.56E-04***</td>
<td>0.0251***</td>
<td>0.0430***</td>
<td>1.13E-04***</td>
</tr>
<tr>
<td></td>
<td>(8.55E-05)</td>
<td>(3.33583E-07)</td>
<td>(6.01E-06)</td>
<td>(3.59E-03)</td>
<td>(2.03E-03)</td>
<td>(9.16E-06)</td>
</tr>
<tr>
<td>1999</td>
<td>9.28E-05</td>
<td>7.38838E-06***</td>
<td>3.55E-04***</td>
<td>0.0211***</td>
<td>0.0451***</td>
<td>6.18E-05***</td>
</tr>
<tr>
<td></td>
<td>(7.99E-05)</td>
<td>(2.52938E-07)</td>
<td>(3.30E-06)</td>
<td>(2.51E-03)</td>
<td>(1.37E-03)</td>
<td>(9.68E-06)</td>
</tr>
<tr>
<td>2000</td>
<td>3.86E-04***</td>
<td>8.016E-06***</td>
<td>3.35E-04***</td>
<td>0.0246***</td>
<td>0.0431***</td>
<td>1.32E-04***</td>
</tr>
<tr>
<td></td>
<td>(8.55E-05)</td>
<td>(3.6165E-07)</td>
<td>(4.24E-06)</td>
<td>(2.97E-03)</td>
<td>(1.71E-03)</td>
<td>(1.03E-05)</td>
</tr>
</tbody>
</table>

Data are mean-differenced to remove 830 tract-level fixed effects. Significance is indicated by: 
*** (0.01), ** (0.05), and * (0.10).
Table 7: Simple Estimates of MWTP to Avoid Violent Crime

<table>
<thead>
<tr>
<th>Year</th>
<th>MWTP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1994</td>
<td>-4.59</td>
</tr>
<tr>
<td>1995</td>
<td>-8.71</td>
</tr>
<tr>
<td>1996</td>
<td>-8.02</td>
</tr>
<tr>
<td>1997</td>
<td>-10.44</td>
</tr>
<tr>
<td>1998</td>
<td>-16.82</td>
</tr>
<tr>
<td>1999</td>
<td>-17.48</td>
</tr>
<tr>
<td>2000</td>
<td>-21.07</td>
</tr>
</tbody>
</table>
Application: MWTP Function

\[ p_{j,k,t}^{VC} = \alpha_{0,k} + YEAR_{j,k,t}^{'} \Omega_k + \alpha_{1,k} VC_{j,k,t} + \alpha_{2,k} INC_{j,k,t} + \alpha_{3,k} API_{j,k,t} + \alpha_{4,k} BLACK_{j,k,t} + \alpha_{5,k} HISP_{j,k,t} + \nu_{j,k,t} \]

\[ p^{VC} = \text{hedonic price of violent crime at house purchased by individual } j \]

\[ INC_j = \text{income (/1,000) of individual } j \]

\[ API_j = 1 \text{ if individual } j \text{ is Asian-Pacific Islander, } = 0 \text{ otherwise} \]

\[ BLACK_j = 1 \text{ if individual } j \text{ is Black, } = 0 \text{ otherwise} \]

\[ HISP_j = 1 \text{ if individual } j \text{ is Hispanic, } = 0 \text{ otherwise} \]
Table 8: MWTP Function Estimates

<table>
<thead>
<tr>
<th></th>
<th>Bishop-Timmins</th>
<th>Rosen</th>
</tr>
</thead>
<tbody>
<tr>
<td>Violent Crime</td>
<td>-0.1008 (3.66E-03)</td>
<td>3.51E-03 (1.05E-04)</td>
</tr>
<tr>
<td>Income (/1000)</td>
<td>-0.0581 (4.88E-03)</td>
<td>-0.0252 (2.75E-03)</td>
</tr>
<tr>
<td>Asian</td>
<td>2.8352 (0.1868)</td>
<td>0.0949 (0.0337)</td>
</tr>
<tr>
<td>Black</td>
<td>25.0498 (0.9251)</td>
<td>1.9942 (0.0684)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>13.9813 (0.5706)</td>
<td>1.1989 (0.0748)</td>
</tr>
<tr>
<td>Constant</td>
<td>54.1218 (1.6280)</td>
<td>-3.3790 (0.3542)</td>
</tr>
<tr>
<td>1995 dummy</td>
<td>-6.2501 0.5187</td>
<td>-2.2899 0.4489</td>
</tr>
<tr>
<td>1996 dummy</td>
<td>-11.8685 0.5090</td>
<td>-1.5536 0.4167</td>
</tr>
<tr>
<td>1997 dummy</td>
<td>-12.8032 0.5172</td>
<td>-3.1074 0.3957</td>
</tr>
<tr>
<td>1998 dummy</td>
<td>-18.4337 0.6021</td>
<td>-7.2178 0.4639</td>
</tr>
<tr>
<td>1999 dummy</td>
<td>-27.1906 0.7219</td>
<td>-7.5220 0.4835</td>
</tr>
<tr>
<td>2000 dummy</td>
<td>-29.8092 0.8100</td>
<td>-10.7238 0.6091</td>
</tr>
<tr>
<td>$\sigma_\nu$</td>
<td>23.8670</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>0.8092</td>
<td>–</td>
</tr>
</tbody>
</table>

All coefficients are significant at the 1\% level of significance.
Year 2000 Hedonic Gradient and MWTP Function Estimates
Application: Valuing Non-Marginal Change in Crime

Estimating MWTP function has significant effects on valuing non-marginal changes in crime:

Consider all individuals who purchased a house in 1999. What is the value of the change in crime experienced over 1999-2000?
Figure 4: Distribution of One-Year Violent Crime Rate Changes for 1999 Buyers
Table 9: WTP for Non-Marginal Changes in Violent Crime

<table>
<thead>
<tr>
<th></th>
<th>Buyers with Reductions (n = 29,953)</th>
<th>Buyers with Increases (n = 19,485)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average WTP</td>
<td>Std. Dev. WTP</td>
</tr>
<tr>
<td>Bishop-Timmins</td>
<td>269.57</td>
<td>328.26</td>
</tr>
<tr>
<td>Rosen</td>
<td>401.15</td>
<td>543.86</td>
</tr>
<tr>
<td>Horizontal MWTP</td>
<td>395.83</td>
<td>532.83</td>
</tr>
</tbody>
</table>
Conclusions

(1) Researchers regularly ascribe downward sloping demand curves to consumers of all sorts of goods. Because of econometric complications, hedonic analysis generally does not ascribe the same flexibility to the demand for amenities.

(2) Data describing housing transactions and attributes of home buyers at a high level of geographic precision are now available for many areas (e.g., Dataquick/HMDA, restricted access census data). Makes estimation of MWTP function feasible.
Conclusions

(3) Require econometric procedures to recover unbiased estimates of MWTP functions for use in valuing non-marginal policies.

- without strict assumptions on the shape of preferences
- without tenuous exclusion restrictions

(3) Hedonic equilibrium yields sufficient information to identify the MWTP function (stratification equation).

- prove identification in simple linear gradient, two market example
- general non-linear gradient ML estimation algorithm
Conclusions

(5) Application to measuring welfare costs of a non-marginal increase in violent crime:

- practical estimator (computationally light)
- results suggest important implications for valuing non-marginal changes in amenities