

A New Approach to Estimating the Production Function for Housing

Dennis Epple, Brett Gordon, and Holger Sieg

Carnegie Mellon University

April 16, 2007

Motivation

- Dating at least to the classic works of Alonso, Mills, and Muth, the production function for housing has played a central role in urban economics and local public finance.
- It entails a powerful abstraction. Houses are viewed as differing only in the quantity of services they provide, with housing services being homogeneous and divisible.
- Estimating the housing production function is then an important undertaking to facilitate the calibration of computational models and to provide evidence about the validity of the strong assumptions entailed in the concept of a housing production function.

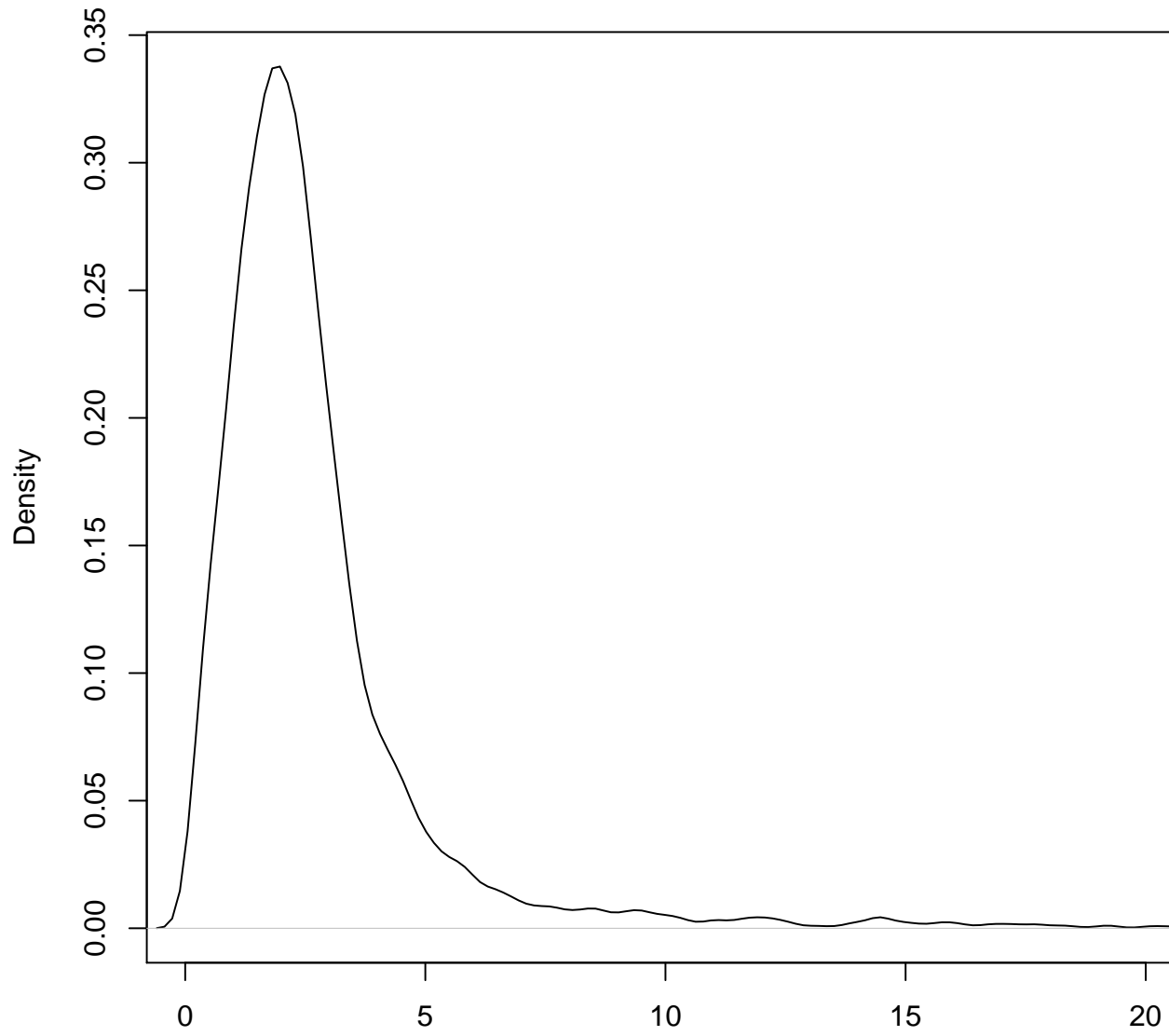
Challenges

- The key problem encountered in estimation is that that the quantity and the price of housing services per dwelling is not observed by the econometrician.
- Instead, we only observe the value of a housing unit which is the product of the price per unit of housing services and the quantity of housing services per dwelling.
- The main objective of this paper is to develop and apply new techniques for estimating production functions which properly treat the quantity and the price of housing services as latent variables unobserved by the econometrician.

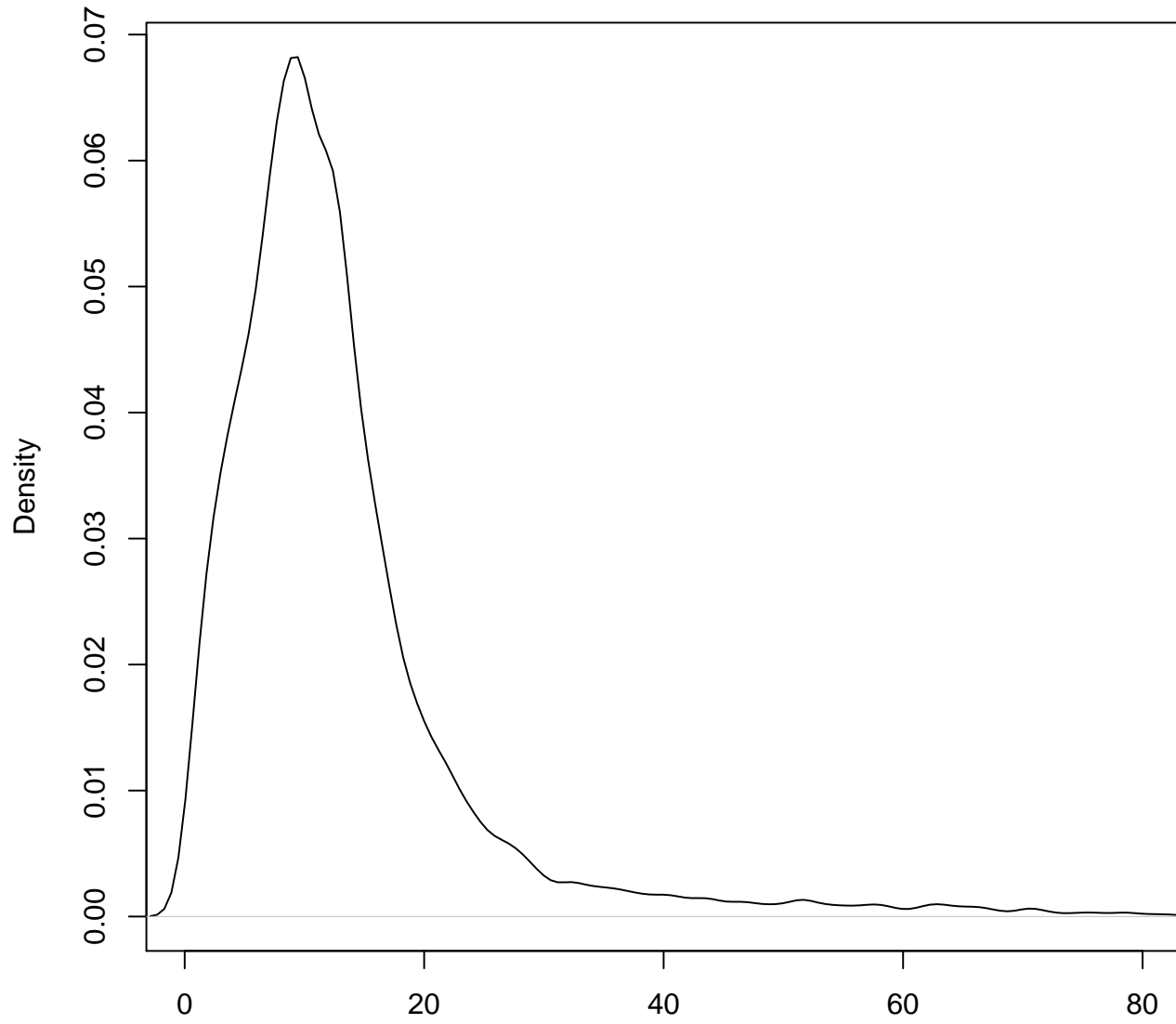
Variation in Land Prices and Housing Values

- Differences in housing values are often due to differences in local land prices.
- The 5th and 95th percentiles of land prices differ in the Pittsburgh metropolitan area by a factor of five; the 1st and 99th percentiles by a factor of fifty.
- Variation in land prices induces variation in the relative proportions of land and non-land factors used in housing production.
- As a consequence we observe large differences in the value of housing per unit of land.

Density of pland



Density of v



Intuition

- Under Constant Returns to Scale, we can normalize output in terms of one input factor (land).
- While we do not observe the price of housing services (p_q), we observe the value of the house per unit of land (v).
- The price of housing is a monotonically increasing function of the value of housing per unit of land.
- We observe the equilibrium relationship between v and the price of land (p_l). If firms make zero profits, we can then characterize the indirect profit function as a function of p_l and v .
- The rest is duality theory.

Some Notation

- We assume that housing is homogeneous and divisible good, Q that can be produced from two factors, M and L via a production function $Q(L, M)$.
- We can think of L as land and M as all other non-land factors.
- The price of mobile factors, p_m , is constant throughout the (metropolitan) area.
- The price of land (the non-mobile factor), p_l , depends of the location.
- As a consequence the price of housing services, p_q , also depends on the location.

Regularity Conditions

We consider an industry with constant returns to scale:

Assumption 1: The production function $Q(L, M)$

a. exhibits constant returns to scale, implying that

$$Q(L, M) = L \cdot Q(1, M/L);$$

b. is strictly increasing, strictly concave, and twice differentiable.

We also assume that the housing industry is competitive and that there are no barriers to entry into the industry:

Assumption 2: There is free entry and firms are price takers.

The Indirect Profit Function

The firm's profit per unit of land can then be written:

$$\pi = \frac{\Pi}{L} = p_q q(m) - m - p_l$$

Solving this optimization problem, we obtain a supply function per unit of land denoted by $s(p_q)$. Let $m(p_q)$ denote the normalized factor demand function. The indirect profit function per unit of land is defined:

$$\pi(p_q, p_l) = p_q s(p_q) - m(p_q) - p_l$$

By the envelope theorem, we have:

$$\frac{\partial \pi(p_q, p_l)}{\partial p_q} = s(p_q)$$

The Relationship between v and p_q

The value of housing per unit of land is defined as:

$$v = p_q s(p_q) = v(p_q)$$

Since $s(p_q)$ is monotonically increasing and differentiable, it follows that $v(p_q)$ is a monotonically increasing, differentiable function of p_q . Hence, this function can be inverted to obtain:

$$p_q = p_q(v)$$

An Example

Consider a Cobb-Douglas production function $Q = M^\alpha L^{1-\alpha}$ which implies that $q = m^\alpha$. Solving the firms optimization problem yields:

$$\begin{aligned}m(p_q) &= (\alpha p_q)^{\frac{1}{1-\alpha}} \\s(p_q) &= (\alpha p_q)^{\frac{\alpha}{1-\alpha}}\end{aligned}$$

as a consequence we have:

$$v(p_q) = p_q s(p_q) = \alpha^{\frac{\alpha}{1-\alpha}} p_q^{\frac{1}{1-\alpha}}$$

Inverting this function yields

$$p_q(v) = \alpha^{-\alpha} v^{1-\alpha}$$

The Equilibrium Locus $p_l = r(v)$

Substituting the function $p_q(v)$ into the indirect profit function yields:

$$\pi(p_q, p_l) = p_q(v)s(p_q(v)) - m(p_q(v)) - p_l$$

Since firms make zero profits in equilibrium, we have

$$p_l = p_q(v)s(p_q(v)) - m(p_q(v)) = r(v)$$

Note that this locus only depends on observables. In the Cobb-Douglas

Example, we have

$$p_l = r(v) = (1 - \alpha) v$$

An Implicit Characterization of $s(p_q)$

In equilibrium, we have the following alternative representation of the indirect profit function:

$$\pi(p_q, p_l) = r(p_q s(p_q)) - p_l = 0$$

Differentiating this expression, we obtain:

$$\frac{\partial \pi(p_q, p_l)}{\partial p_q} = r'(p_q s(p_q)) [s(p_q) + p_q s'(p_q)]$$

The supply function per unit of land is implicitly characterized by the solution to the following differential equation:

$$r'(p_q s(p_q)) \cdot [s(p_q) + p_q s'(p_q)] = s(p_q)$$

An Example (continued)

Suppose the relationship between p_l and $r(v)$ is linear:

$$p_l = r(v) = (1 - \alpha) v$$

The differential equation for the supply function is then:

$$(1 - \alpha) [s + p_q s'] = s$$

Integrating and rearranging, we obtain the following supply function:

$$s = c p_q^{\frac{\alpha}{1-\alpha}}$$

where c is the constant of integration.

Three Representations of the CD function

The direct production function:

$$q = m^\alpha$$

The dual in housing price and land price:

$$p_q^{\frac{1}{1-\alpha}} \alpha^{\frac{\alpha}{1-\alpha}} (1-\alpha) - p_\ell = 0$$

The dual in housing value and input price:

$$(1-\alpha)v - p_\ell = 0$$

A General Solution

Consider the differential equation:

$$(r'(ps) - 1) s dp + r'(ps) p ds = 0$$

The integrating factor ps converts the equation above into an exact differential equation. As a consequence the solution to the differential equation is:

$$\int \frac{r'(ps)}{p} dp + \int \left[\frac{r'(ps)}{s} - \frac{\partial}{\partial s} \int \frac{r'(ps)}{p} dp \right] ds = c + \ln(p)$$

The Production Function

Once we have derived the supply function, it is straight forward to derive the underlying production function. We have:

$$\pi(p_q, p_\ell) = p_q s(p_q) - m(p_q) - r(p_q s(p_q)) = 0$$

We can solve this equation for the factor demand function:

$$m(p_q) = p_q s(p_q) - r(p_q s(p_q))$$

Let the inverse of $m(p_q)$ be $p_q(m)$. The production function per unit of land is then:

$$q(m) = s(p_q(m))$$

The General Case

Consider now a model with three inputs, M , N , and L . We assume that the prices of N and L (p_n and p_l) vary across space.

As before let $\pi(p_q, p_n)$ denote the indirect profit function. The envelope theorem now implies that

$$\begin{aligned}\frac{\partial \pi}{\partial p_q} &= s(p_q, p_n) \\ \frac{\partial \pi}{\partial p_n} &= -n(p_q, p_n)\end{aligned}$$

The Equilibrium Loci

Using a similar logic as in the previous section, we can still show that there exists an equilibrium locus that relates p_l to v given p_n . Thus we have

$$p_l = r(v, p_n)$$

Similarly, there exist a function $n(v, p_n)$ such that:

$$n = n(v, p_n)$$

An Alternative Representation of the Indirect Profit Function

The zero profit condition implies the following alternative representation of the indirect profit function:

$$\pi(p_q, p_n) = r(p_q, s(p_q, p_n), p_n) - p_l$$

and note that

$$\frac{\partial \pi}{\partial p_q} = \frac{\partial r}{\partial v} \left[s(p_q, p_n) + p_q \frac{\partial s(p_q, p_n)}{\partial p_q} \right]$$

and

$$\frac{\partial \pi}{\partial p_n} = \frac{\partial r}{\partial v} p_q \frac{\partial s(p_q, p_n)}{\partial p_n} + \frac{\partial r}{\partial p_n}$$

The Characterization of $s(p_q, p_n)$

The supply function is then the solution to the following fundamental system of partial differential equations:

$$\begin{aligned} s(p_q, p_n) &= \frac{\partial r}{\partial v} \left[s(p_q, p_n) + p_q \frac{\partial s(p_q, p_n)}{\partial p_q} \right] \\ -n(p_q s(p_q, p_n), p_n) &= \frac{\partial r}{\partial v} p_q \frac{\partial s(p_q, p_n)}{\partial p_n} + \frac{\partial r}{\partial p_n} \end{aligned}$$

Note that the functions r and n only depend on observables.

An Example

In the three factor Cobb-Douglas Case, we have

$$q = m^\alpha n^\beta$$

The equilibrium loci $r(\cdot)$ and $n(\cdot)$ are given by:

$$r(v, p_n) = (1 - \alpha - \beta)v$$

$$n(v, p_n) = \beta \frac{v}{p_n}$$

An Example (continued)

Hence we can write the fundamental system of partial differential equations as:

$$\begin{aligned} s(p_q, p_n) &= (1 - \alpha - \beta) \left[s(p_q, p_n) + p_q \frac{\partial s(p_q, p_n)}{\partial p_q} \right] \\ -\beta \frac{p_q s(p_q, p_n)}{p_n} &= (1 - \alpha - \beta) \left[p_q \frac{\partial s(p_q, p_n)}{\partial p_n} \right] \end{aligned}$$

and it is straight forward to verify that the solution to this problem is

$$s(p_q, p_n) = c p_n^{\frac{-\beta}{1-\alpha-\beta}} p_q^{\frac{\alpha+\beta}{1-\alpha-\beta}}$$

where c is a constant of integration.

Estimation

To obtain a well behaved econometric model we introduce an additively separable error term to the indirect profit function:

$$p_l = r(v) + \epsilon$$

Assuming $E(\epsilon|v) = 0$, the model can be estimated using nonparametric techniques using a sample of size N . If we assume that

$$r(v) = r(v, \theta_k)$$

and treat k as a function of the sample size N , $k = k(N)$, we can reinterpret the model above as a semi-nonparametric model (Chen, 2006).

In our application it is convenient to approximate the function with a polynomial of arbitrary order k , $r_k(v) = \sum_{i=1}^k \frac{r_i}{i} v^i$.

The supply function in the general polynomial case can be expressed solely in terms of v and $\{r_i\}$, is

$$s = \frac{v^{1-r_1}}{\exp \left\{ \sum_{i=2}^k \frac{r_i}{i-1} (v^{i-1} - 1) \right\}}$$

and

$$p = v^{r_1} \exp \left\{ \sum_{i=2}^k \frac{r_i}{i-1} (v^{i-1} - 1) \right\}$$

Data: New Housing Construction

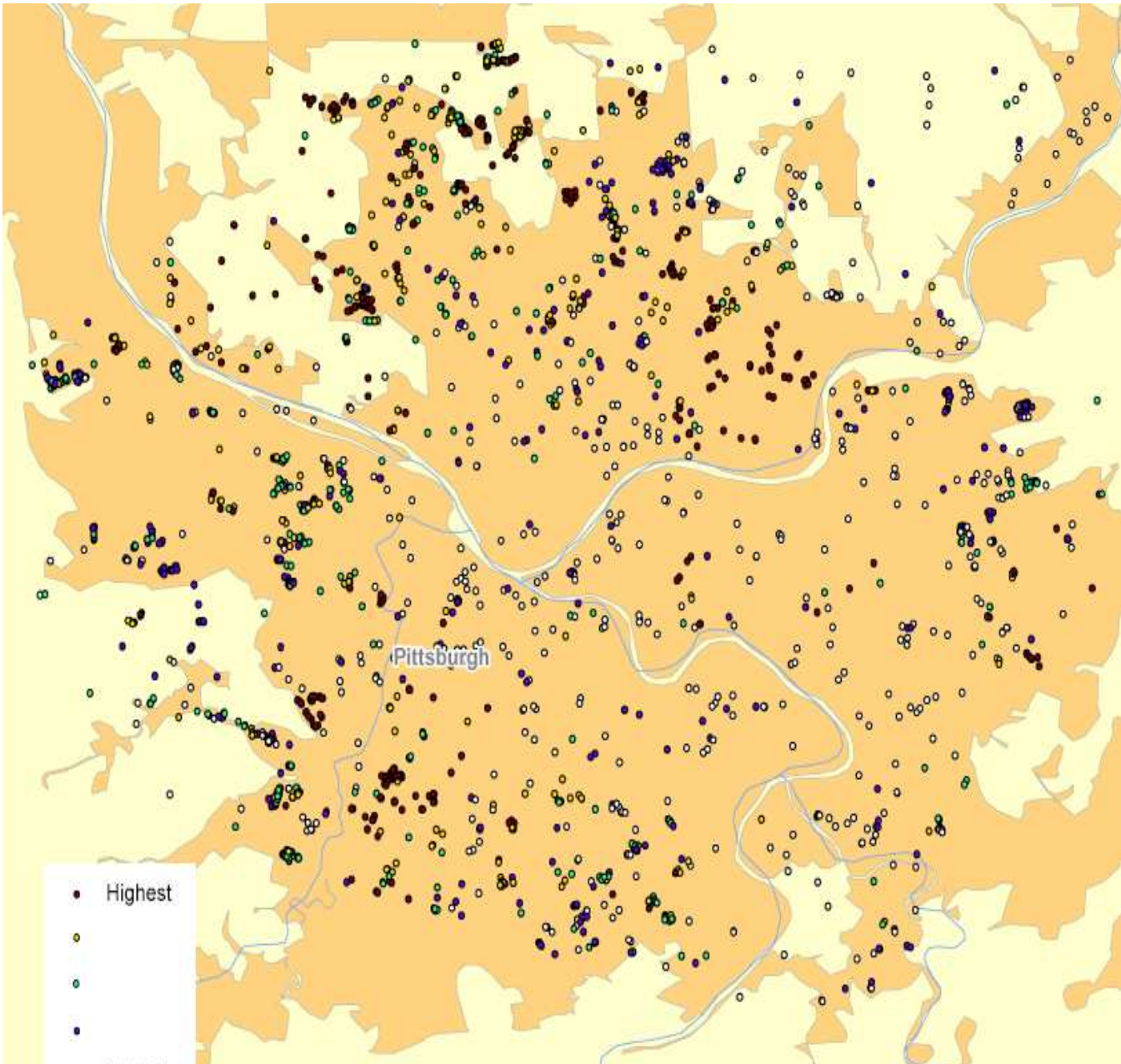
We use data from Allegheny County in Pennsylvania, which contains the greater metropolitan area of Pittsburgh.

Most of the analysis uses residences that were built in or after 1995, which yields our final sample size as 6,362.

Table 1: Descriptive Statistics

Variable	Mean	Median	Stdev	Min	Max
value per unit of land	21.44	14.29	26.91	0.15	366.62
price of land	3.32	2.28	3.86	0.05	41.75
lot area	26756	15507	52197	540	1207483
travel time	29.12	30	9.47	1	59

The sample size is 6,362.



Price per Unit Land vs. House Price per Unit Land

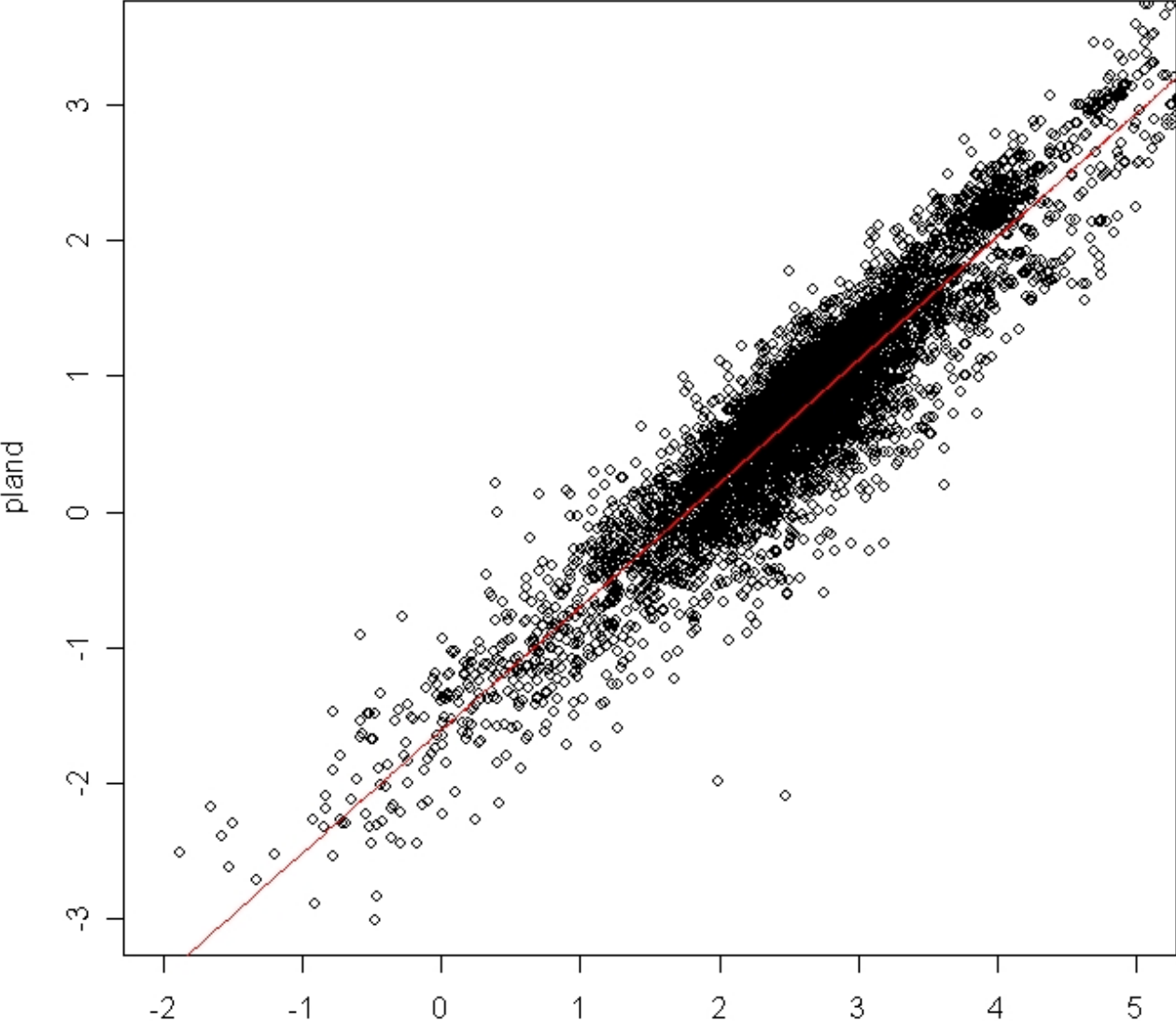


Table 2: OLS estimates of $r(v)$

OLS Estimates				
	Log-linear	Linear	Quadratic	Cubic
v		0.1394***	0.1685***	0.1622***
v^2			-0.0002***	-0.0001
v^3				3.9e-7*
Constant	-1.6051***			
$\log(v)$	0.9090***			
R^2	0.8649	0.8014	0.8382	0.8391
N	6,362	6,362	6,362	6362

* indicates significance at the 90% level, ** at the 95% level, and *** at the 99% level.

Table 3: IV estimates of $r(v)$

2SLS Estimates				
	Log-linear	Linear	Quadratic	Cubic
v		0.1440 ***	0.1631***	0.1732***
v^2			-0.0002***	-0.0005***
v^3				1.1e-6*
Constant	-1.6129***			
$\log(v)$	0.9119***			
R^2	0.8649	0.7992	0.8360	0.8135
N	6,362	6,362	6362	6,362

* indicates significance at the 90% level, ** at the 95% level, and *** at the 99% level.

Empirical Results

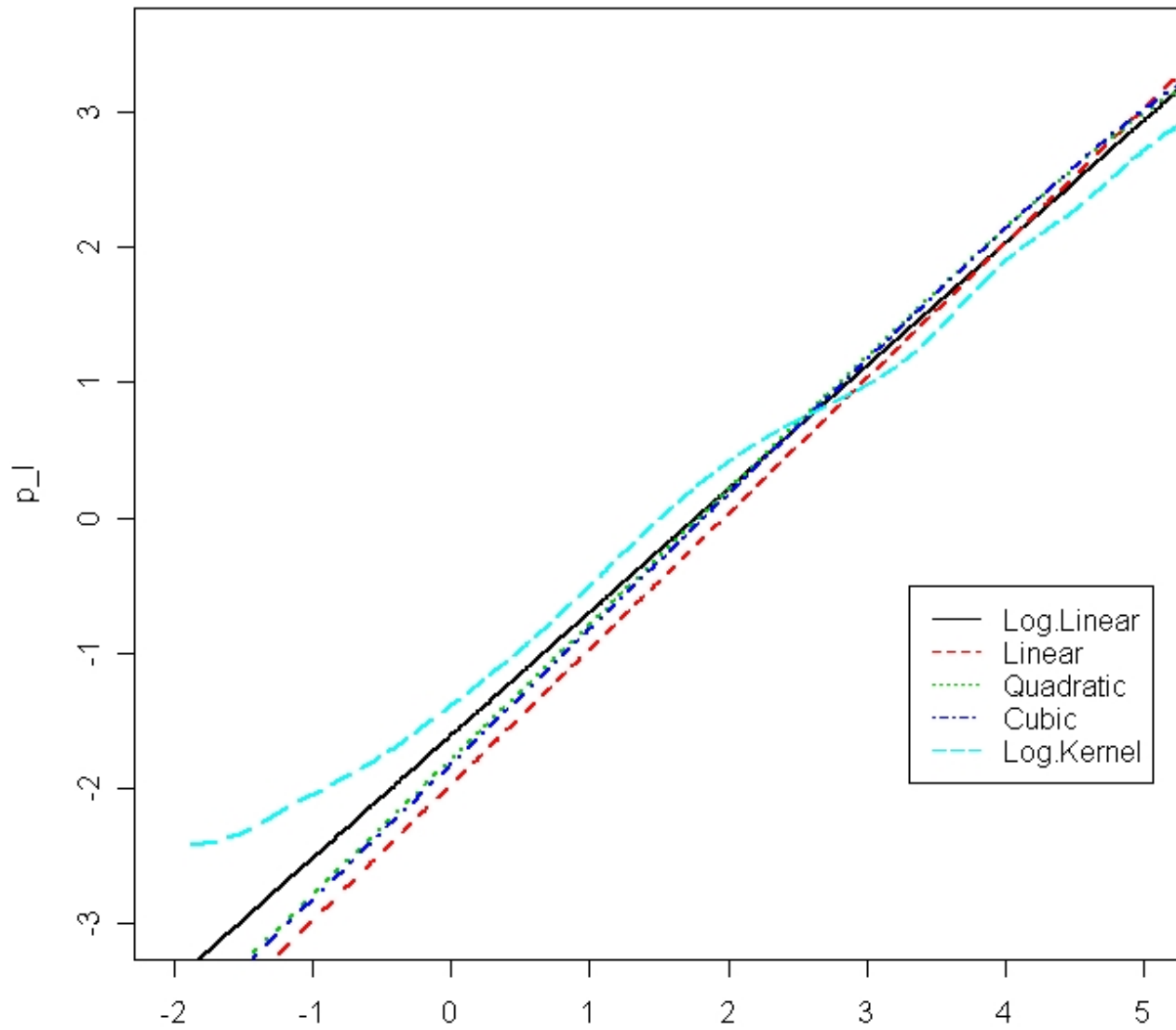
We find that the housing supply function per unit of land is price elastic with average price elasticities ranging between 4.3 and 6.6.

In the linear case, $r(v) = kv$, the estimated slope coefficient is 0.1394. This implies that Cobb-Douglas production function is given by $Q(L, M) = 1.14 * L^{0.14} * M^{0.86}$.

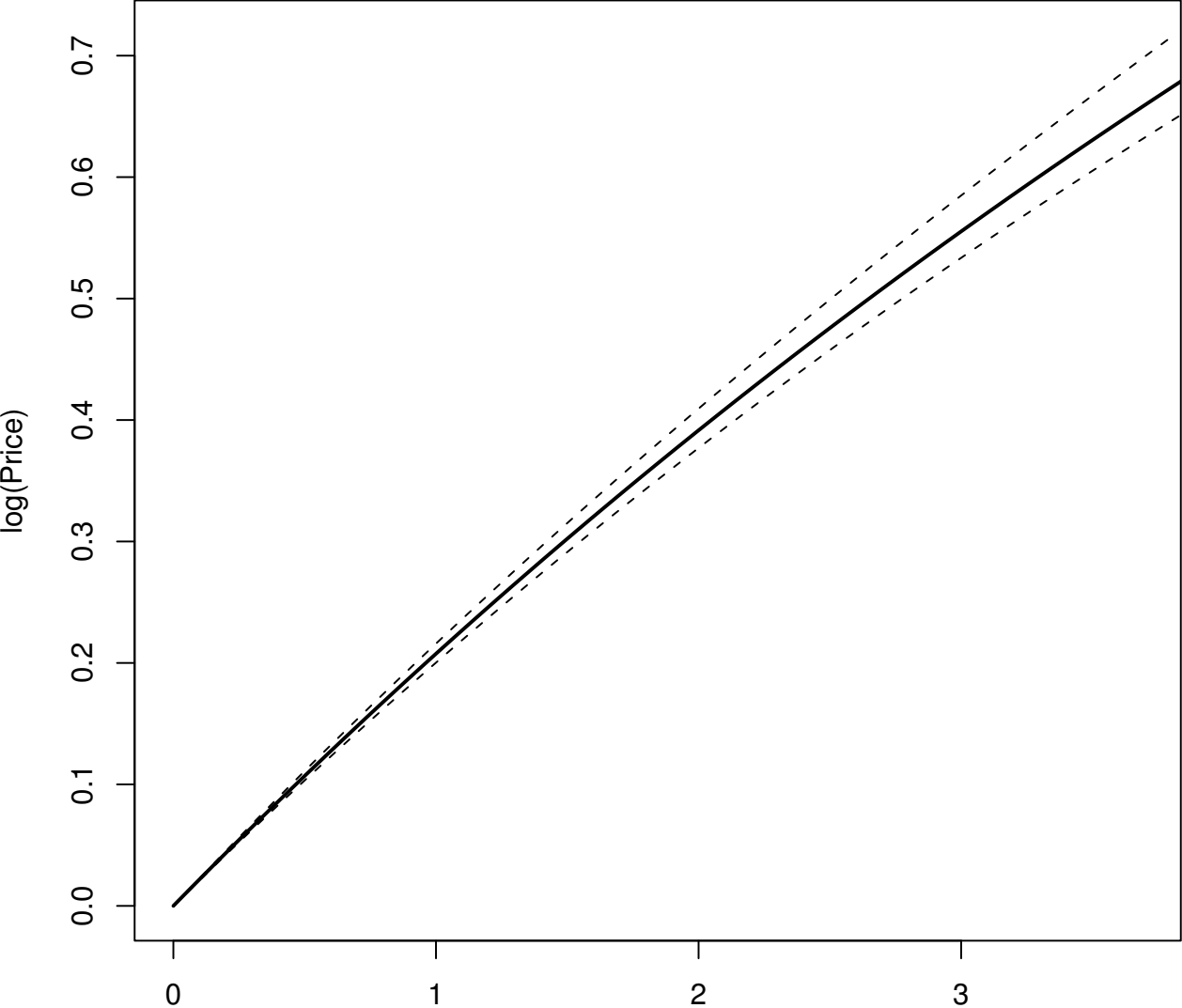
We find that the elasticity of substitution between land and non-land factors ranges between 1 in the linear case and 0.84 in the quadratic case.

McDonald (1981) surveys 13 studies and report estimates of the elasticity of substitution ranging between 0.36 and 1.13 with a majority obtaining estimates significantly less than one.

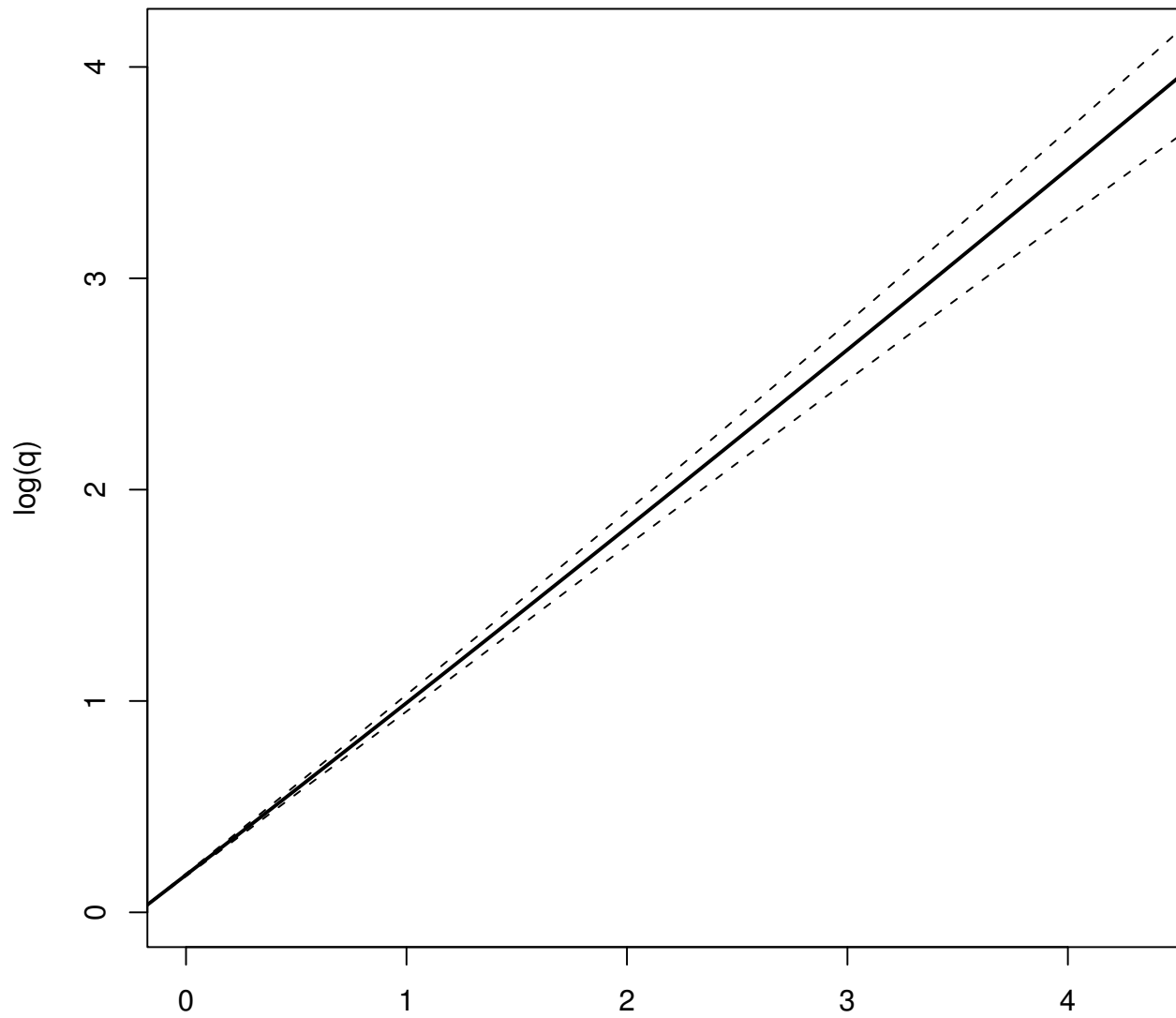
Fitted Regressions for $r(v)$



**Log-linear Supply Function with
95% Confidence Band**



Log-linear Production Function with 95% Confidence Band



Results for Downtown Commercial Properties

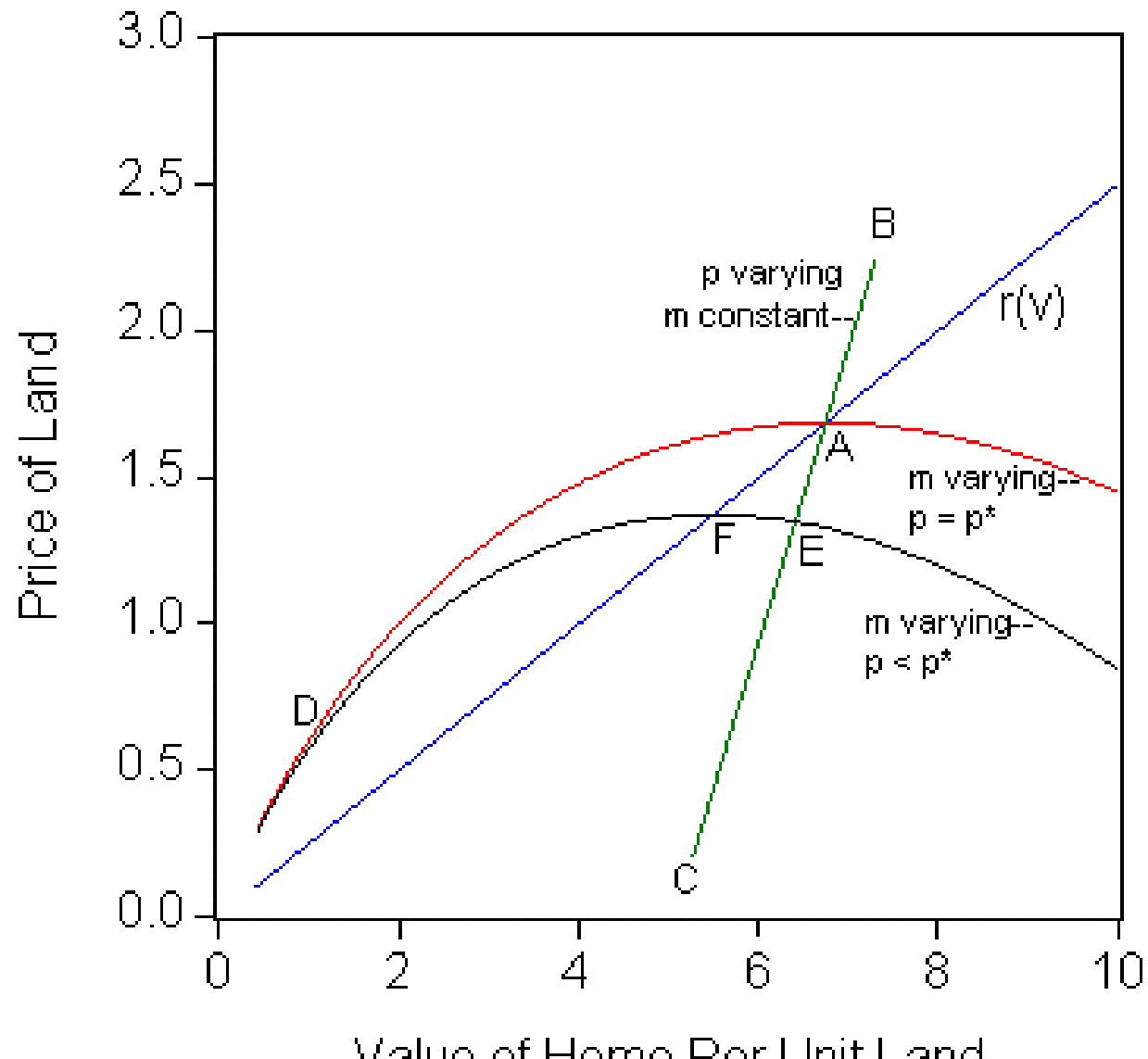
- Estimates are substantially different from residential property case.
- Consider the log-linear case:
 - Commercial: Constant = -0.72 (0.04) Slope = 0.74 (0.02)
 - Residential: Constant = -1.61 (0.00) Slope = 0.91 (0.00)
- Mean supply elasticity is 3.98 (1.43).
- Mean substitution elasticity is 1.39 (0.04).

Zoning

Zoning can have broadly two effects:

1. Zoning may require more housing services per dwelling than would be chosen in the absence of zoning.
2. Zoning may result in a distortion of input ratios, for example, by increasing lot size by a larger proportion than investment in the dwelling is increased.

Potential Effects of Large-Lot Zoning



To test the hypothesis that Fox Chapel residences lie on the $r(v)$ curve, we generated a dummy variable for the Fox Chapel observations, and included the dummy alone and interacted with $\log(v)$ in the regression for $r(v)$.

Table 4: An Empirical Test for Zoning

Fox Chapel		
	Coefficient	p-value
Constant	-1.6068	0.0000
$\log(v)$	0.9095	0.0000
FC	0.1474	0.4431
FC* $\log(v)$	-0.0476	0.4865
R^2	0.8650	
N	6362	

Policy Implications

Examples: School voucher programs, property tax reforms, housing vouchers, welfare reform, urban development policies, or policies aimed at improving access of poor households to economic opportunity.

All of these policies are likely to affect the demand for housing and residential sorting patterns. If the supply of new housing is price elastic, an increase in the demand of housing is largely met by an increase in housing supply. Even large policy changes may only have a small impact on housing prices if the supply is elastic.

Hence welfare effects will largely be driven by household adjustments and changes in housing quantities, and not so much by price changes.

Conclusions

- We have demonstrated how to estimate the housing production functions when housing prices are unobserved.
- We have illustrated the usefulness of the approach using an application that focuses on Pittsburgh.
- We think that this approach can also be helpful in measuring productivity and technological progress in a number of other industries besides housing, where output prices and quantities are not well observed.