A Flexible Approach to Estimating Production Functions When Output Prices are Unobserved

Dennis Epple, Brett Gordon, and Holger Sieg

Carnegie Mellon University

May 18, 2006

Motivation

Prices for homogeneous, non-tradeable goods often vary by location:

- A house in Beverly Hills or Santa Monica costs up to five times as much as the same house in Riverside or San Bernadino county.
- The cheapest room in a Hilton Hotel in New York City starts at $439. The same room in a Hilton Hotel in Cleveland, Ohio, costs $109.
- Renting a Ford Taurus from Hertz at LAX costs $74 a day. Renting the same car from Hertz at the Eureka Airport costs $48 a day.
- Tickets for “The Producers” range between $31.25 and $51.25 in Tallahassee and $36.25 to $111.25 in New York.

What Explains These Price Differences?

- Differences in output prices are often due to differences in local factor 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.

- Wages for experienced mechanics range from \$25 to \$100 in our sample. In the absence perfect labor mobility, wages can thus differ substantially among a set of labor markets. Kennan and Walker (2005) document large differences in local wage rates in the U.S.

Challenges

- One potential problem encountered in estimation is that the quantity and the price of output are not separately observed by the econometrician.

- Instead, we observed the value of the output. If prices differ for the same good, the value of output is not necessarily a good measure for the quantity of output.

- 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 output as latent variables unobserved by the econometrician.

Why Are Prices And Quantities Not Observed Separately?

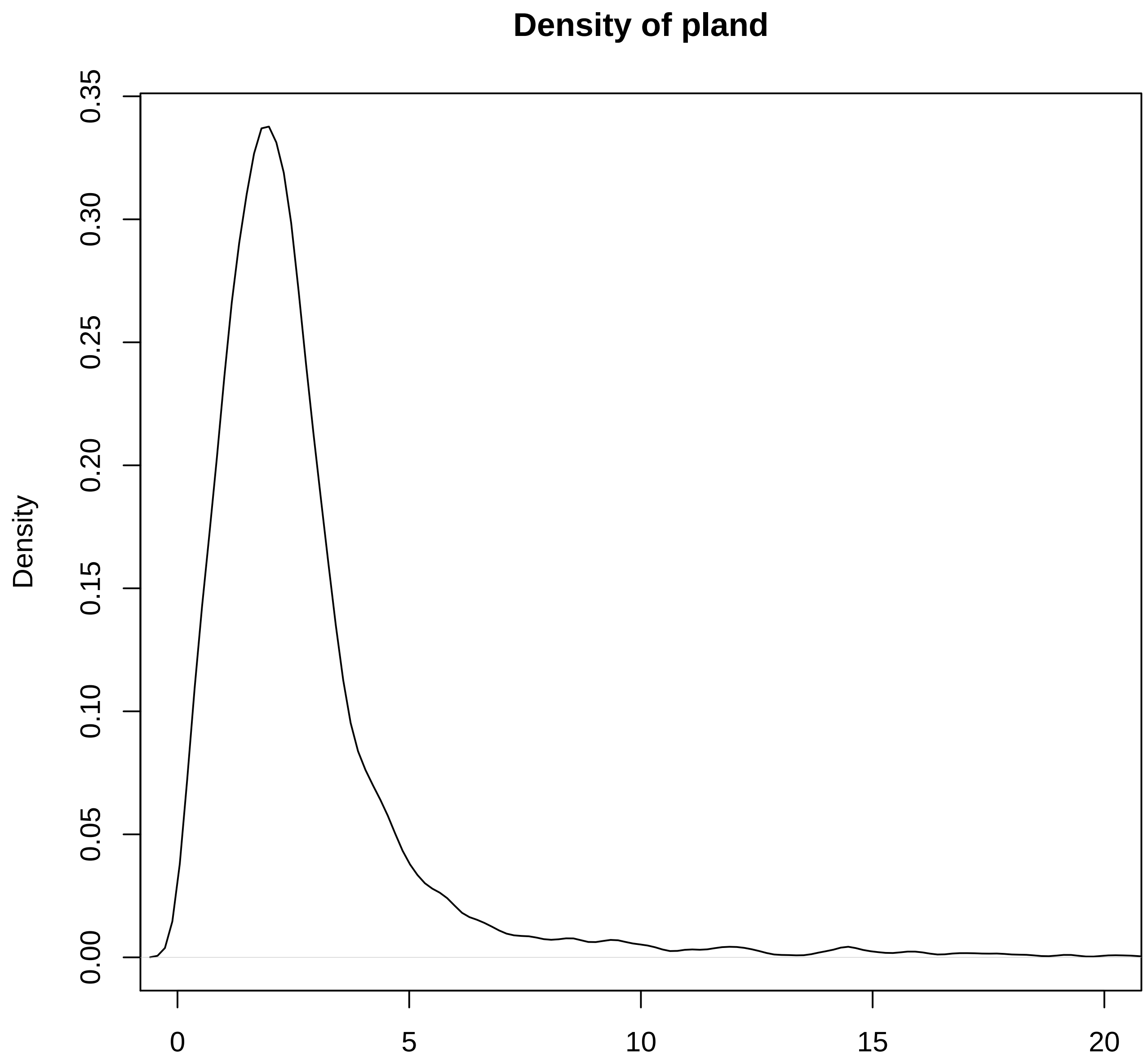
- It is often convenient to assume that the amount of a good can be measured in terms of a homogeneous unit.
- This assumption is value in theoretical modeling, for tractability and convenience.
- It is rare outside of agricultural commodities to observe goods that are easily measured in homogeneous units.
- The existence of a production function itself often entails a powerful abstraction. There may not be an easily measurable price associated with this theoretical construct.

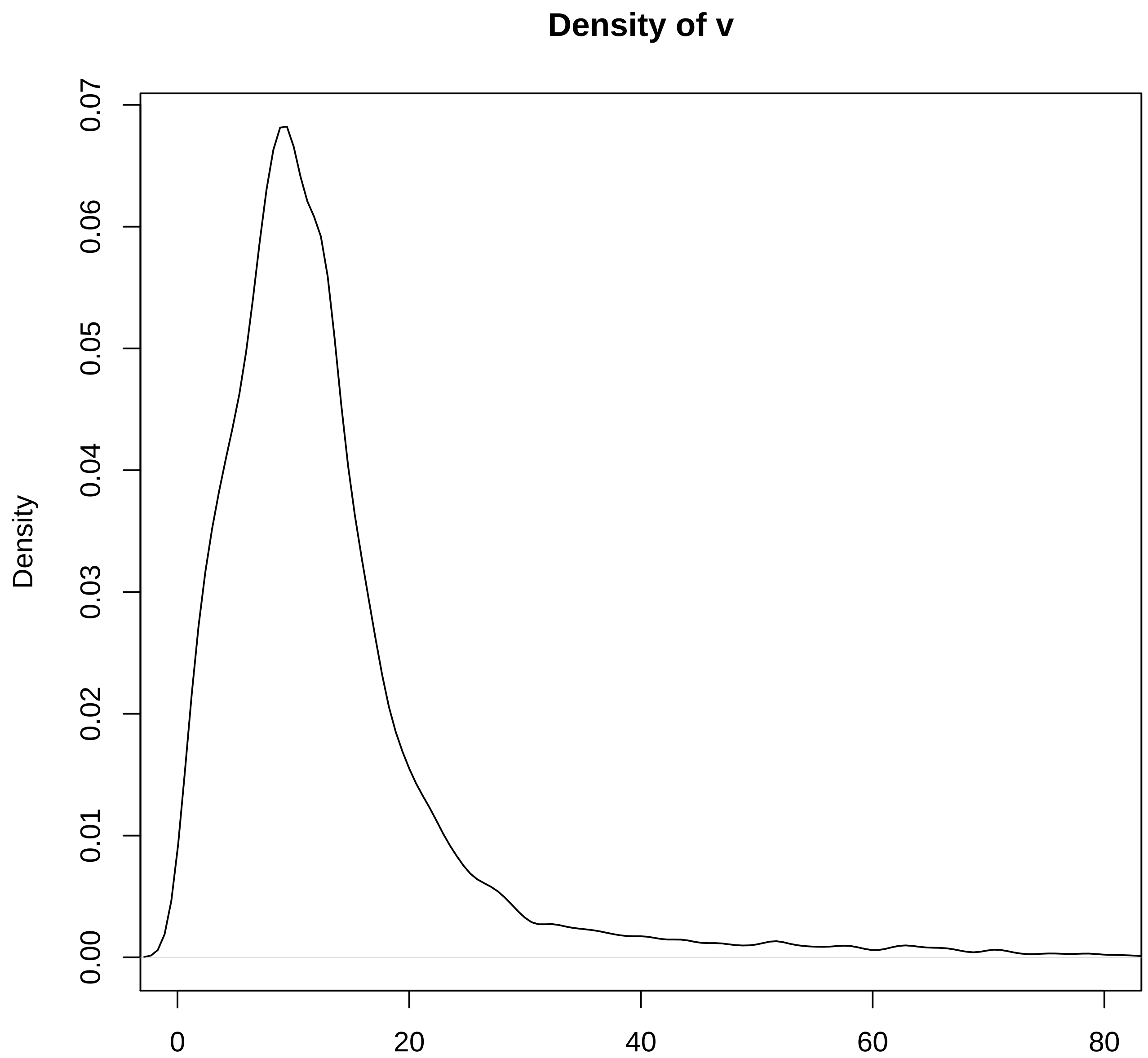
Applications: Housing and Car Repair Services

- Housing is assumed to be a homogeneous good, despite the observed differences in quality of housing units.
- Different houses are viewed as differing only in the quantity of services they provide. Thus, a grand house and a modest house differ only in the number of homogeneous service units they contain.
- We observe the value of a house, but we never observe prices and homogeneous service units.
- Similarly, it may be useful to ignore differences among routine car repairs.

Some Stylized Facts

- There is an enormous variation in land prices within a metropolitan area.
- This variation arises from differences in proximity to places of employment and commerce, access and availability of public goods and amenities among locations.
- Variation in land prices induces variation in the relative proportions of land and non-land factors used in housing production. We observe a large variation in the value of housing per unit of land.


Density of pland
Density
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0
5
10
15
20


Density of v
Density
0.00
0.01
0.02
0.03
0.04
0.05
0.06
0.07
0
20
40
60
80

Intuition

- Variation in land prices induces variation in the relative proportions of land and non-land factors used in housing production.
- We can trace out the equilibrium relationship between land prices and housing values per unit of land. Moreover, this equilibrium locus implicitly characterizes the supply of housing per unit of land.
- Estimating the supply function per unit of land allows us to decompose the observed house value per unit of land into a price and a quantity component.
- The production function of housing can be recovered from the supply of housing per unit of land.

Literature:

- Omitted Price Bias: Marschak and Andrews (1944) and Klette and Griliches (1996).
- Transmission Bias due to the endogeneity of input factors: Marschak and Andrews (1944), Mundlack and Hoch (1965), Zellner, Kmenta, and Dreze (1966), ...
- Specification Bias: Diewert (1971), Christensen, Jorgenson, and Lau (1973), Gallant (1981), Vinod and Ullah (1988), ...
- Dynamic Selection Bias: Olley and Pakes (1996), Levinsohn and Petrin (2003).

Some Notation

- We assume that a homogeneous (non-tradeable) good, Q can be produced from two factors, M and L via a production function $Q(L, M)$.
- The price of mobile factors, p_m, is constant throughout the area.
- The price of the non-mobile factor, p_l, depends of the location.
- As a consequence the price of output, 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 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 Supply per Unit of the Non-Mobile Factor

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

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

Normalize $p_M = 1$ and let $\pi^*(\cdot)$ denote the corresponding indirect profit function. By the envelope theorem:

$$\frac{\partial \pi^*(p_q, p_l)}{\partial p_q} = s(p_q) \tag{1}$$

Moreover the supply function has the following properties:

Proposition 1 *$s(p_q)$ is strictly increasing in p_q, $s(p_q) > 0$ for $p_q > 0$, and $s(p_q)$ approaches zero as p_q approaches zero.*

An Alternative Representation of π^*

There is a relationship that characterizes the optimal factor use in equilibrium

$$p_l \;=\; r(v)$$

The value of output per unit of L, denoted by v, is defined as:

$$v = p_q s(p_q)$$

Hence we have $r(p_q s(p_q)) - p_\ell = 0$. The zero profit assumption implies that:

$$\pi^*(p_q, p_l) = r(p_q s(p_q)) - p_\ell = 0$$

An Implicit Characterization of the Supply Function

Differentiating this expression, we obtain:

$$\frac{\partial \pi^*(p_q, p_l)}{\partial p_q} = r'(p_q s(p_q))[s(p_q) + p_q s'(p_q)] \tag{2}$$

Combining equations (1) and (2), we have the following key result that provides the basis of our approach to estimating the supply function per unit of L:

Proposition 2 *The supply function per unit of the non-mobile factor is implicitly characterized by the solution to the following differential equation:*

$$r'(p_q s(p_q)) \cdot [s(p_q) + p_q s'(p_q)] = s(p_q) \tag{3}$$

Properties of $r(v)$

The following Proposition establishes conditions that must be satisfied by function $r(\cdot)$:

Proposition 3 *The equilibrium locus* $p_l = r(v)$ *must satisfy the following restrictions:*

$$0 < r'(v) < 1$$

for all $v > 0$*.*

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:

$$m^*(p_q) = p_q s(p_q) - r(p_q s(p_q)) \tag{4}$$

Let the inverse of (4) be $p_q^*(m)$. The production function per unit of L is then:

$$q(m) = s(p_q^*(m)) \tag{5}$$

An Example

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

$$p_\ell = r(v) = kv$$

The following differential equation for the supply function per unit of L:

$$k \cdot [s + p_q s'] = s$$

Integrating and rearranging, we obtain:

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

where c is the constant of integration which can be normalized $(c = 1.)$

To derive the production function we substitute the supply function into (4) and obtain

$$m^*(p_q) = (1-k)p_q^{\frac{1}{k}}$$

Inverting:

$$p_q^* = \left(\frac{m}{1-k}\right)^k$$

Substituting this result into (5), we obtain the production function per unit of L:

$$q(m) = \left(\frac{m}{1-k}\right)^{1-k}$$

Recall that $Q(M, L) = Lq(M/L)$. This and the preceding yield the production function:

$$Q(L, M) \;=\; A\, L^k\, M^{1-k}$$

where $A = 1/(1-k)^{(1-k)}$.

A General Solution

Consider the differential equation:

$$(r'(p\ s) - 1)\ s\ dp\ +\ r'(p\ s)\ p\ ds\ =\ 0 \tag{6}$$

We have the following result:

Proposition 4 *The integrating factor p s converts (6) into an exact differential equation. As a consequence the solution to equation (6) is:*

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

A Semi-Nonparametric Approach

In our application it is convenient to approximate the unknown $r(v)$ function with a polynomial of arbitrary order k:

$$p_l = \sum_{i=1}^{k} \frac{r_i}{i} v^i \ + \ \epsilon \tag{7}$$

Assuming $E(\epsilon|v) = 0$, the equation above can be estimated using least squares based on a cross-section of housing units with sample size N. If we treat k as a function of the sample size N, i.e. assume that $k = k(N)$, we can reinterpret the model above as a semi-nonparametric model (Chen, 2006).

After estimating $r(v)$, the key differential equation could be numerically solved for $s(p)$. The next Proposition provides a simpler approach to calculate the supply function.

Proposition 5 *A simple expression for the supply function in the general polynomial case, expressed solely in terms of* v *and* $\{r_i\}$, *is*

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

and

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

A Nonparametric Approach

Alternatively, we can estimate $r(v)$ using a fully nonparametric estimator, such as a kernel estimator.

After we have obtained unrestricted kernel estimate of $r(v)$, we also need to check whether the derivative restrictions are satisfied everywhere.

With the bandwidth set equal to the standard deviation of v, we typically find that the derivative conditions of Lemma 1 are met.

Once we have obtained a reasonable nonparametric estimate of the function $r(v)$, we solved the ordinary differential equation in Proposition 2 using the boundary-value condition $s(1) = 1$.

Application #1: Housing

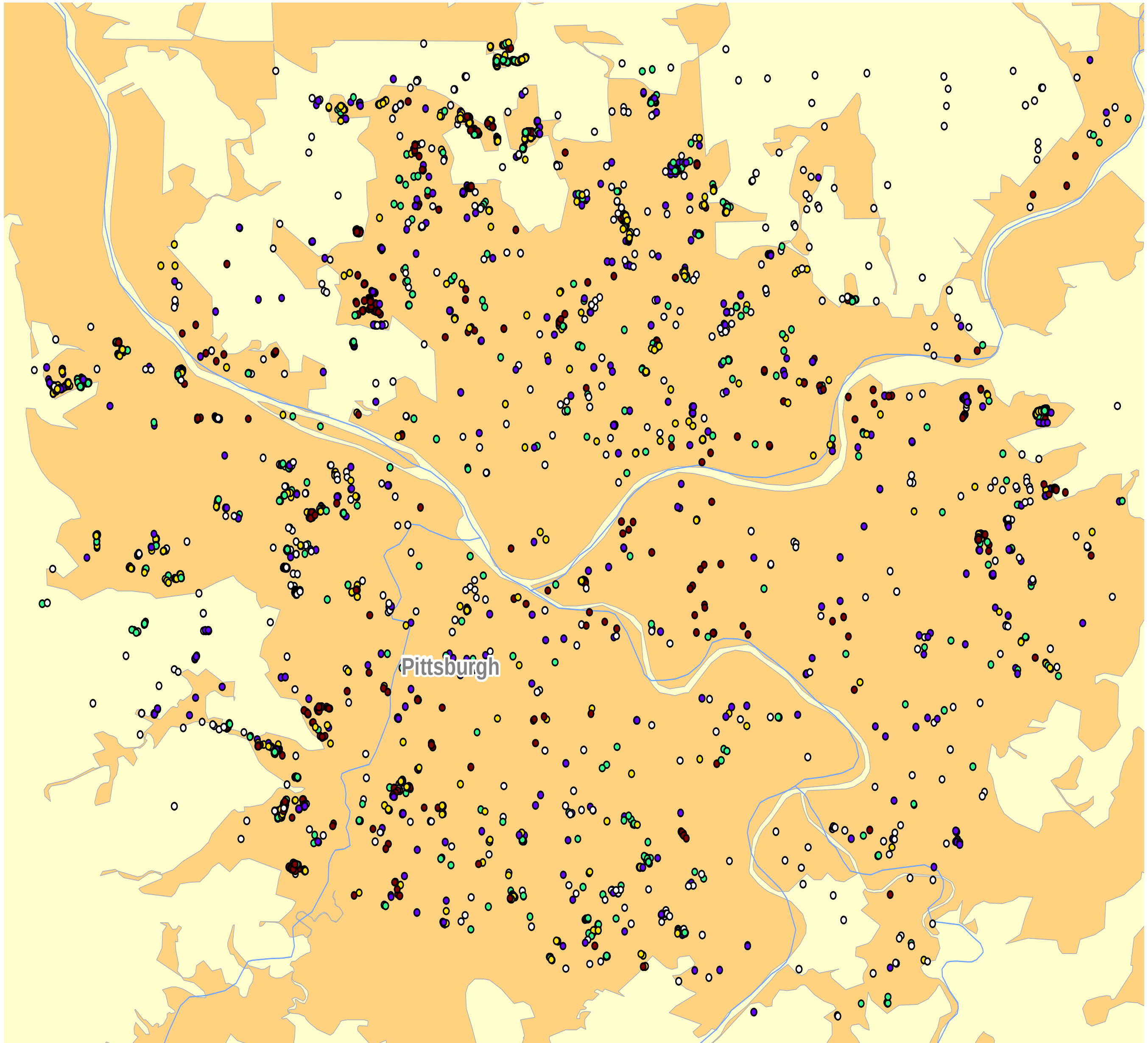
We use data from Allegheny County in Pennsylvania, which contains the greater metropolitan area of Pittsburgh.

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.


Pittsburgh

Price per Unit Land vs. House Price per Unit Land

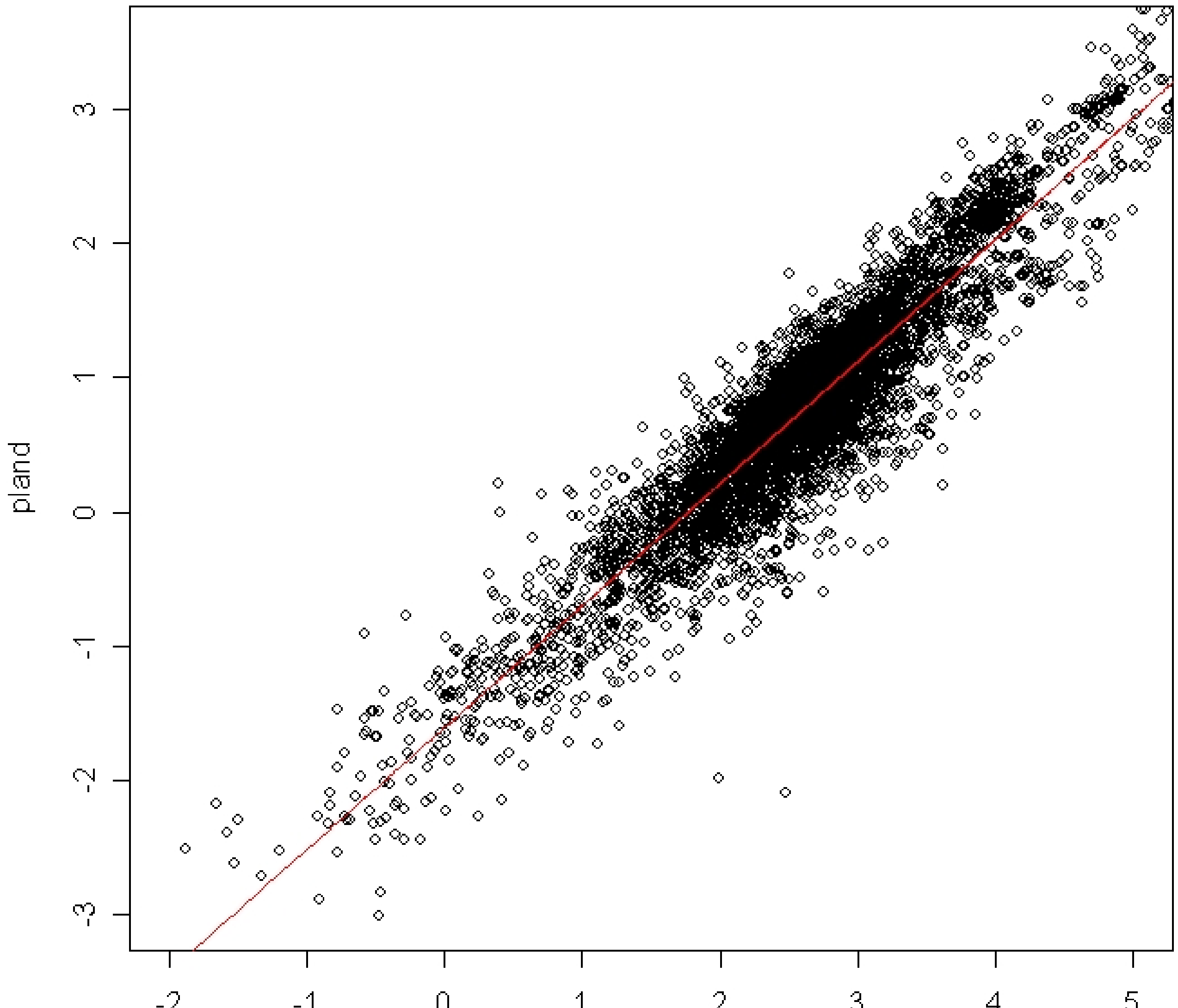


Table 2: OLS estimates of $r(v)$

	OLS Estimates			
	Log-linear	Linear	Quadratic	Cubic
v		0.1394***	0.1685***	0.1622***
v^2			-0.0002***	-0.0001
v^3				$3.9\mathrm{e}{-7}$*
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.

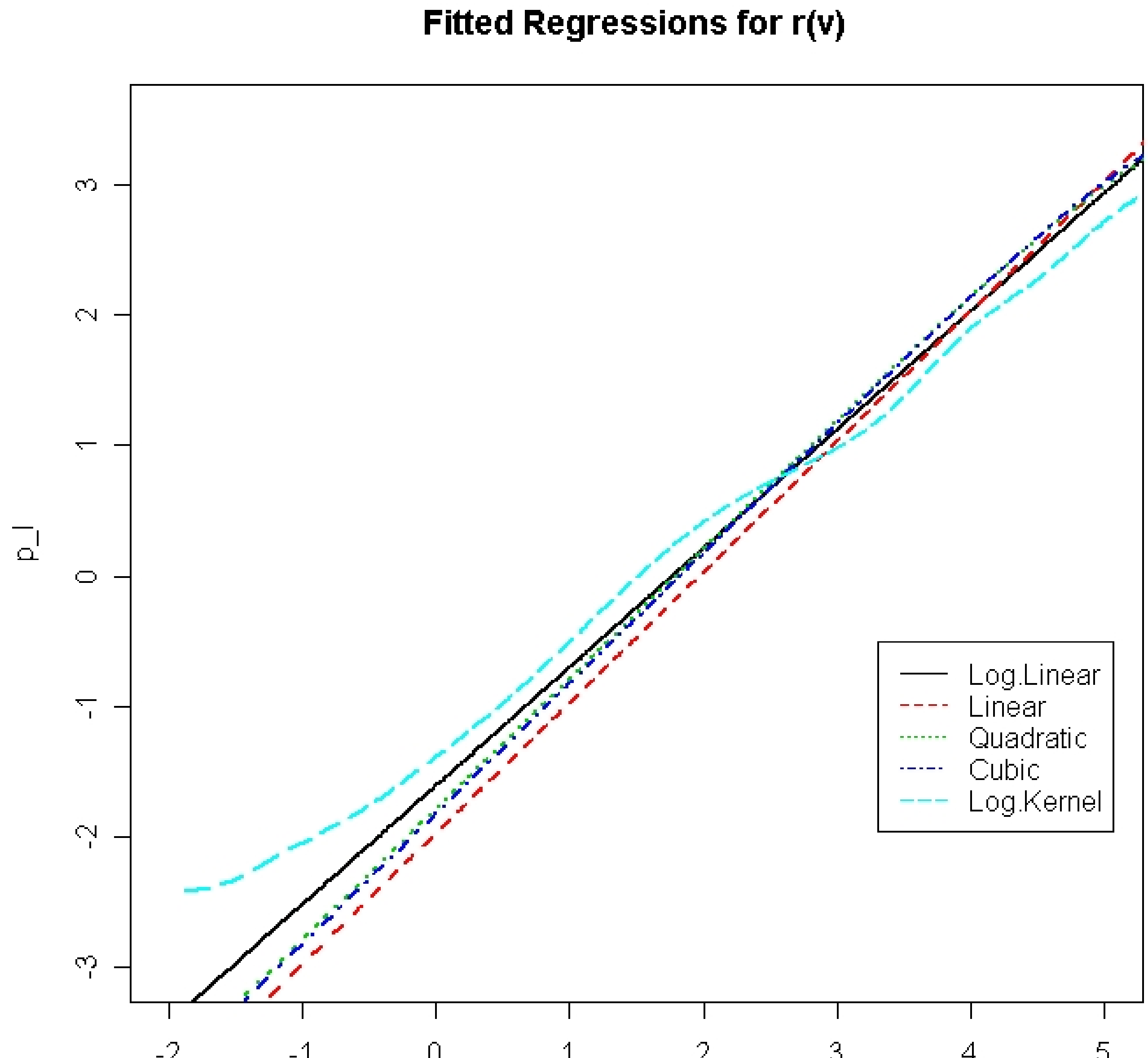
Empirical Results

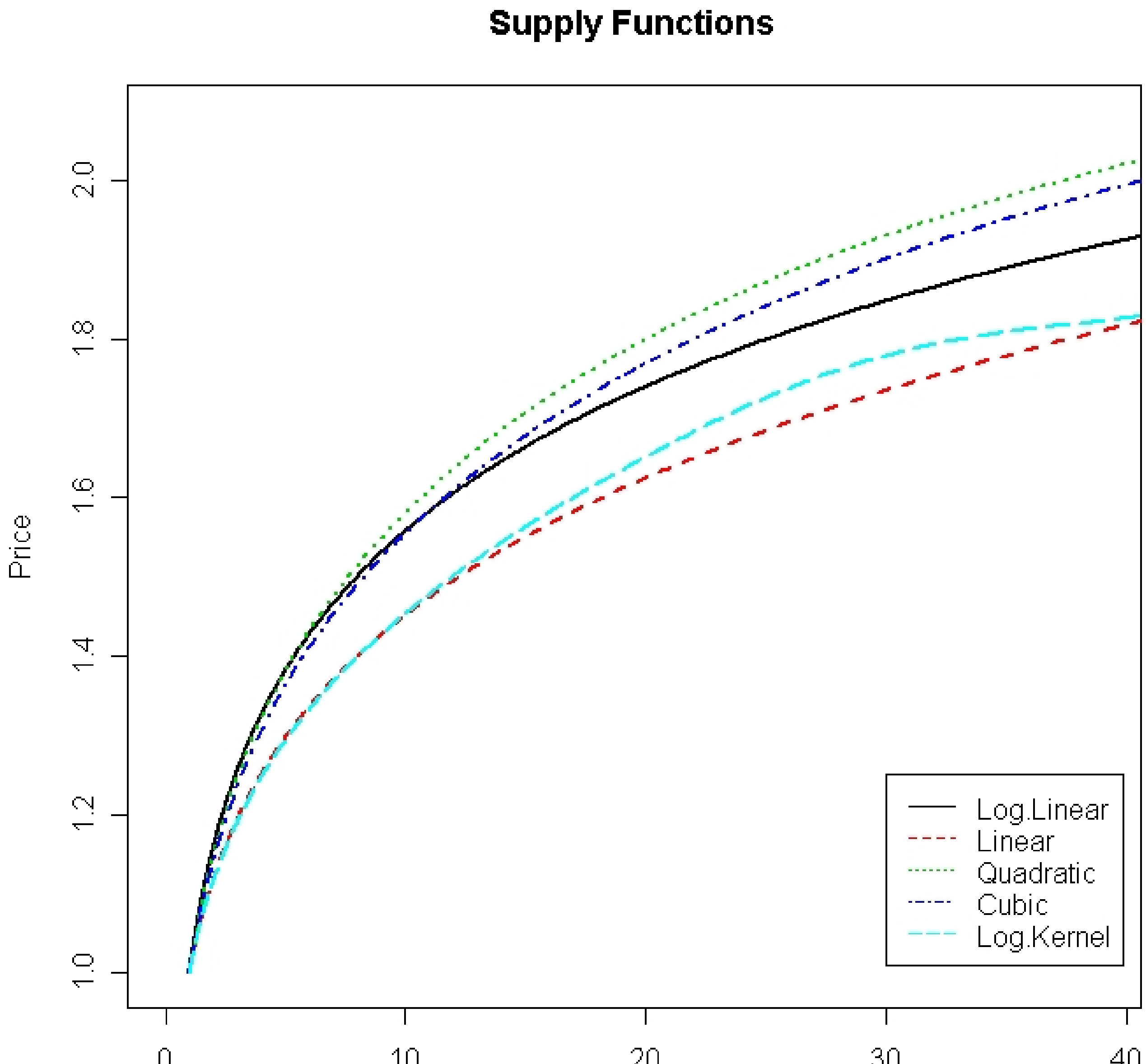
We find that the housing supply function per unit of land is price elastic with average price elasticities ranging between 4.3 and 6.6.

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

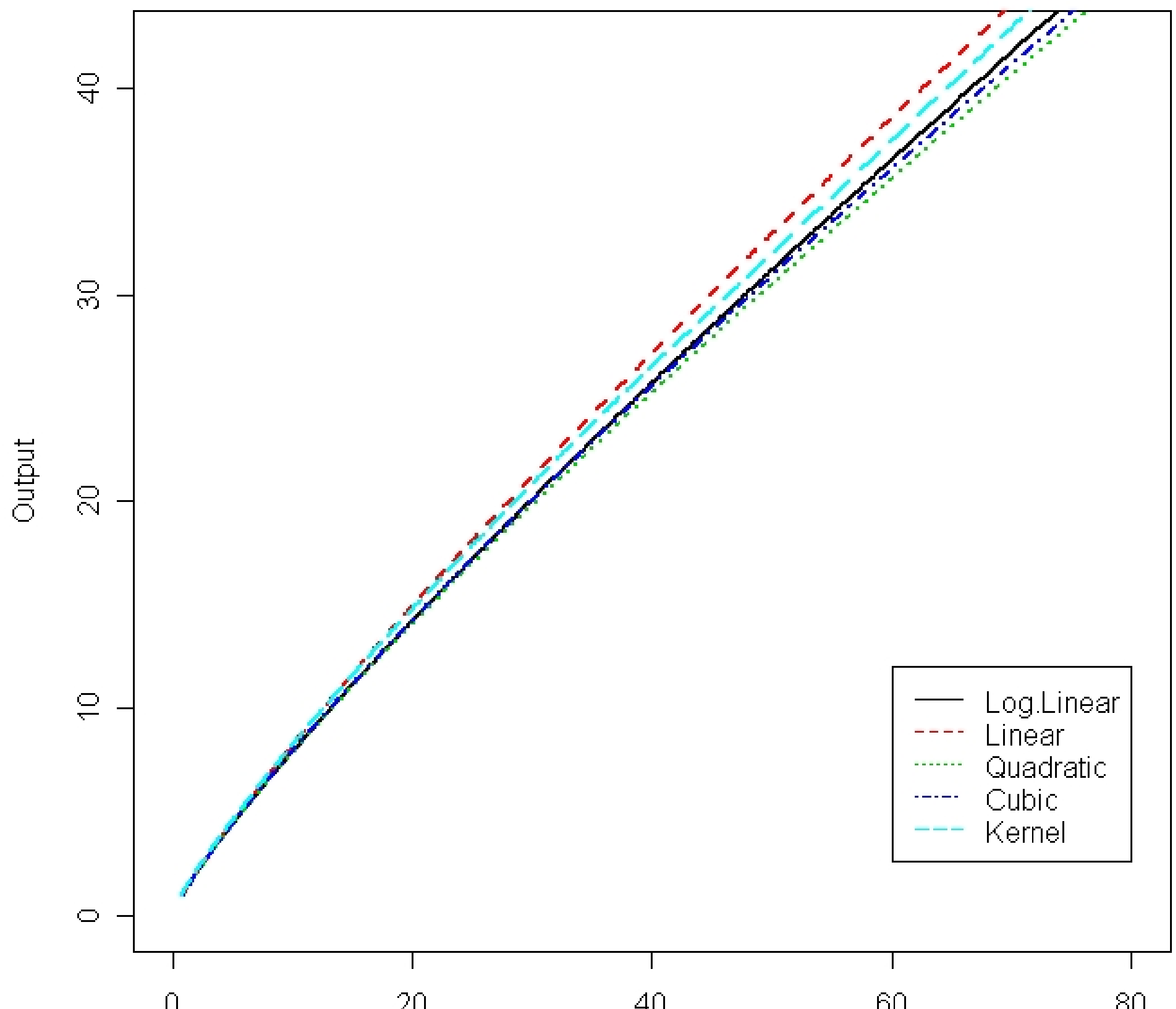
We find that the elasticity of substitution between land and non-land factors ranges between 1 in the linear case and 0.84 in the quadratic case.

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)
p_l
3
2
1
0
-1
-2
-3
-2
-1
0
1
2
3
4
5
Log.Linear
Linear
Quadratic
Cubic
Log.Kernel


Supply Functions
Price
2.0
1.8
1.6
1.4
1.2
1.0
0
10
20
30
40
Log.Linear
Linear
Quadratic
Cubic
Log.Kernel

Production Functions



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.

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).

Application #2: Car Repair Services

We obtained a unique data set based on surveys conducted for Underhood Service Magazine. Survey was conducted in 2004 using a random sample of 4,000 subscribers, yielding 241 responses.

- Nearly all respondents operate a single repair shop.
- Large majority are family owned and in business for 20 years.
- Each has an average of 4.4 (2.84) repair bays.
- 32% are in cities with population below 15,000.
- Average of 3.7 full-time employees.
- Mean hourly wage rate for technicians is $58.54 (14.01).

Price Dispersion in Car Repair Services

Data set shows how output prices for similar goods can vary substantially both within and across states.

Table 4: Price Dispersion in Car Repair Services

state	min	max	state	min	max
Diagnostic Check only					
California	20	300	Ohio	35	98
Brake Repair					
California	110	600	Pennsylvania	75	400
Spark Plug Replacement					
Florida	60	400	Indiana	50	300
Oil and Lube Job					
Florida	25	100	Wisconsin	20	36

Production Function Estimates

Estimate revenue production function per (experienced) repair technician. Analogous variables:

- $L \equiv$ Number of technicians [mean = 2.39 (1.17)]
- $p_\ell \equiv$ Annual salary for experienced technician [$38,016 (15,610)]
- $v \equiv$ Annual revenues per technician [$144,364 (62,305)]

After eliminating incomplete observations, we are left with 97 observations. Supply function estimates imply mean price elasticity of 4.58 and standard deviation of 1.95. Mean substitution elasticity of 1 (due to Cobb-Douglas).


Density of Price Measures
Density
0.0
0.1
0.2
0.3
0.4
0.5
1
2
3
4

Conclusions

- Differences in factor prices are essential to obtain differences in factor inputs and thus to identifying production functions.

- We have demonstrated how to estimate production functions when output prices are unobserved.

- We have illustrated the usefulness of the approach using two applications.

- We think that this approach can be helpful in measuring productivity and technological progress across industries and countries.