Consumer Benefit From Air Quality Improvements

by

Dimitrios A Giannias

Working Paper No. 9

October 1988

ABSTRACT

This paper applies a simultaneous equations estimation technique to estimate a hedonic equilibrium model. The estimation results are used to compute consumer benefit from air quality improvements.

1. Introduction.

This paper applies a quality theory that is presented in Giannias (1987) to investigate the willingness to pay for improvements in the air quality of Houston. The theory specifies a hedonic equilibrium model that is offered for empirical work. This model introduces a housing quality index that maps housing characteristics into a scalar quality index. In other respects, this model is an improvement upon the previous work on hedonic equilibrium models, Tinbergen (1959) and Epple (1984), because i) it does not imply demand functions for differentiated goods that have a zero income elasticity and ii) it does not require a variance-covariance structure that needs to be diagonal.

All the previous applied work in this area uses a method that has been proposed by Harrison and Rubinfeld (1978) (or a variation of this method) to compute consumer benefit for changes in some of the characteristics of a differentiated good. This method empirically approximates the features of the hedonic price and willingness to pay functions using fitting criteria to derive them. This method provides more flexibility in letting the data determine the functional forms at the cost of not being able to test in a consistent way whether the assumed functional forms are consistent among themselves and with the underlying economic structure. In addition to that, this method cannot predict how non-marginal changes in exogenous parameters, e.g., the mean or the variance of the air quality distribution, will affect the equilibrium price distribution. As a result, this method cannot estimate the consumer benefit from such changes in exogenous parameters. The method that is followed by this paper makes prior assumptions about the characteristic of the economic agents interacting to form the hedonic equilibrium, uses that to derive the form of the hedonic function, and then estimates only that. Imposing these prior restrictions helps through the additional theoretical information that is essential for the estimation and the willingness to pay results.

Sections 2 and 3 present the economic and the econometric models respectively. The model is estimated in Section 4 and the structure of the economy is analyzed in Section 5. Concluding remarks are given in Section 6.

2. The Economic Model.

The differentiated product rental residential housing can be described by a vector of characteristics v, where $v = [v_1\ v_2\ v_3]$, v_1 is the size of a housing unit (number of rooms), v_2 is an air quality index, v_3 is the travel time to work (measured in minutes). The air quality variable is the inverse of the air pollution variable total suspended particulate matter (measured in microgram per cubic meter). It is assumed that v follows an exogenously given multi-normal distribution.

The quality of housing, h (a scalar), is a linear function of the vector of housing characteristics v, that is,

$$h = \epsilon\, v' , \tag{1}$$

where $\epsilon = [\epsilon_0\ \epsilon_1\ \epsilon_2]$ is a vector of parameters.

Consumer preferences are described by utility functions. A utility function, $U(h,x;a)$, depends on the quality of the house, h, on the numeraire good, x, and on the parameter a, where a is the number of persons in a family.

A consumer solves the following optimization problem:

$$\max \; U(h,x;a)$$

with respect to h, x

subject to $I = 12P(h) + 365x$ and

$$P(h) = \pi_0 + \pi_1 h$$

where I is the annual income of a consumer, P(h) is the equilibrium (monthly) gross rental price equation (it gives the gross monthly expenditure as a function of the housing quality h), and π_0 and π_1 are the parameters of the equilibrium price equation. The utility function is assumed to be a quadratic of the following form:

$$U(h,x;a) = \delta + (\zeta_0 + \zeta_1 a)h + 0.5\xi h^2 + xh \qquad (2)$$

where δ, ζ_0, ξ, ζ_1 are utility parameters.

The vector [a I] is assumed to follow an exogenously given multi-normal distribution.

Solving the utility maximization problem to obtain the demand for h and substituting it into the equilibrium condition, namely, Aggregate Demand for h = Aggregate Supply for h for all h, it can be proved[1] that the equilibrium price equation for the economy described above is[2]:

$$P(h) = \pi_0 + \pi_1 h \tag{3}$$

where $\pi_0 = (\frac{365}{12})[\varsigma_0 + \varsigma_1 \bar{a} + (\frac{\bar{I}}{365}) - A\bar{h}]$

$\pi_1 = (\frac{365}{24})(\xi + A),$

$\bar{a}$ is the mean size of a family, $\bar{I}$ is the mean consumer income, $\bar{h} = \epsilon \bar{v}'$ is the mean quality of residential housing, $\bar{v}$ is the mean of the vector of housing characteristics v.

$$A = \frac{\sqrt{t\Sigma_z t'}}{\sigma_s}$$

$$\sigma_s = \epsilon \Sigma_v \epsilon'$$

$t = [\varsigma_1 \; 1]$, and

Σ_v is the variance-covariance matrix of the exogenously given distribution of housing characteristics.

The above results imply that the equilibrium demand for h is given by the following function:

$$h = \bar{h} + \frac{\varsigma_1}{A}(a - \bar{a}) + \frac{1}{365A}(I - \bar{I}) \tag{4}$$

3. The Econometric Model.

The previous section implies that the complete model consists of the equations (1), (3), and (4).

For the residential housing market, I assume that the quality of housing is a latent variable. Without loss of generality, the quality of housing can be normalized by setting the parameter ϵ_0 equal to 1.

Substituting equation (1) in (3) and (4), I eliminate the quality of housing[3] and I obtain that the price equation and the first order condition for the consumer's optimization problem are respectively equivalent to:

$$P = (365/12)[\zeta_0 + \zeta_1 \bar{a} + (\bar{I}/365) - A\epsilon\bar{v}' + 0.5\ (\xi + A)\ \epsilon v'], \text{ and}$$

$$\epsilon(v' - \bar{v}') - \frac{\zeta_1}{A}\ (a - \bar{a}) - \frac{1}{365A}\ (I - \bar{I}) = 0.$$

I assume an additive error term on the above two equations. To be more specific, I assume that the equations that I will estimate are the following:

$$P = c + \beta_1 v_1 + \beta_2 v_2 + \beta_3 v_3 + u_1\ , \text{ and} \tag{5}$$

$$(v_1 - \bar{v}_1) + \epsilon_1(v_2 - \bar{v}_2) + \epsilon_2(v_3 - \bar{v}_3) + \epsilon_3(a - \bar{a}) + \epsilon_4(I - \bar{I}) + u_2 = 0 \tag{6}$$

where

$$c = (365/12)[\zeta_0 + \zeta_1 \bar{\alpha} + \bar{I}/365 - A\ \epsilon\ \bar{v}']\ , \tag{7}$$

$$\beta_{i+1} = (365/24)(\xi + A)\epsilon_1, \quad \text{for } i = 0,\ 1,\ 2, \tag{8}$$

$$\epsilon_3 = - \frac{\zeta_1}{A}\ , \tag{9}$$

$$\epsilon_4 = \frac{1}{365A}\ , \text{ and} \tag{10}$$

u_1, and u_2 are the econometric errors of the first and second equations respectively. They are assumed to satisfy the following: (A1) u_1 and u_2 are uncorrelated, (A2) a and I are uncorrelated to u_1 and u_2, (A3) v_2 and v_3 are uncorrelated to u_1, (A4) v_2 and v_3 are uncorrelated to u_2.

4. Estimation of the Reduced Form Equations.

I estimate the last two equations, (5) and (6), simultaneously via Maximum Likelihood. I also impose the restrictions that are implied by the structure of the model, namely,

$$\epsilon_1 = \frac{\beta_2}{\beta_1} \quad , \quad \text{and} \tag{11}$$

$$\epsilon_2 = \frac{\beta_3}{\beta_1} \quad . \tag{12}$$

I estimate the model using (1980) census tract data on rental prices, number of rooms, travel time to work, size of the family, and consumer income, and (1979) SAROAD based data on air quality. To obtain the annual arithmetic mean of total suspended particulate for each census tract, all the monitoring stations of the city were located according to census tract. The readings for these census tracts were used to represent pollution readings in adjacent census tracts since most cities contain a limited number of monitoring stations. If a census tract was adjacent to more than one census tract containing a monitoring station, then the average of readings was used.

Given that the air quality and travel time to work are census tract variables, assumptions A3 and A4 require that the consumer census tract locational choice is exogenous and uncorrelated to the econometric errors of the equations that I estimate[4]. Unlike other work, e.g., Harrison and Rubinfeld (1978), the model states that it is legitimate to use census tract data because 1) the price equation is linear in housing characteristics and 2) the equilibrium demand for housing quality is linear in consumer income and family size. To estimate the model, I use data on Houston, Texas. The results are given in Table 1.

To see if the model is of any value at all, I tested the hypothesis that all the parameters of the equation (5) equal zero. An F-test implies that this hypothesis is rejected at the 1% significance level. A similar F-test implies that I cannot accept the hypothesis that all the parameters of the second equation, equation (6), equal zero (at the 1% significance level).

The t-statistics (see Table 1) show that all the parameters are significant at the 10% significance level. Moreover, the size of a house (which is expected to be the main determinant of the rent), as well as the income and the size of the family (that are expected to be the main determinants of the demand for housing quality) are significant at the 1% significance level.

For the residential housing market, I expect the following: $\epsilon_1 > 0$, $\epsilon_2 < 0$, and $\zeta_1 > 0$. Therefore, the parameter estimates for ϵ_1, ϵ_2, and ζ_1 must satisfy these inequalities. The structural analysis that I make in the next section allows me to check whether my expectations are correct. This will be another test of the model.

5. Structural Analysis.

The empirical results of the previous section allow me to analyze the structure of the housing market of Houston, Texas. I can also specify how that structure depends on the mean of the air quality distribution. The latter enables me to address interesting questions that a non-structural approach cannot.

5.1. The Houston Housing Market.

The parameter estimates that are given in Table 1 and the theoretical model enable me to compute the rental price equation, the demand for housing quality, the quality index equation, the utility function, and the demand for the numeraire good.

Given the parameter estimates in Table 1, given that they satisfy (11) and (12), and given the relationships among the structural parameters and the reduced form equation parameters of the model, equations (7) - (10), I can solve for: A, ζ_1, ϵ_1, ϵ_2, ξ, and ζ_0. The solutions for these parameters follow: A = 25.10, ζ_1 = 12.66, ϵ_1 = 33.93, ϵ_2 = -0.068, ξ = -21.91, and ζ_0 = -1.13. (Note that in order to solve for ζ_0 from (7), I used the statistics that are given in Table 2).

Next, I use the parameter estimates that I have obtained so far to compute the rental price equation, the demand for housing quality, the housing quality index equation, and the utility function. They are respectively given by the following equations:

$$P = 93.51 + 48.59v_1 + 1648.67v_2 - 3.32v_3,$$

$$h = -0.167 + 0.505a + 0.000109I,$$

$$h = v_1 + 33.93v_2 - 0.068v_3, \text{ and} \tag{13}$$

$$U(h,x;a) = \delta + (-1.13 + 12.66a)h + xh - 10.95h_2 . \tag{14}$$

The price equation can also be written in the following way:

$$P = 93.51 + 48.59h.$$

I substitute the last equation and the demand for h into the budget constraint and solve for the demand for x. The demand for x is given next:

$$x = -2.81 - 0.81a + 0.0026I.$$

We can now see that: 1) the rent is positively related to number of rooms and air quality and negatively to travel time to work, e.g., an additional room increases the monthly rent by $48.60, 2) the rent is positively related to the quality of the house, 3) the demand for housing quality is positively related to the size of the family and income, 4) the housing quality is positively related to air quality and negatively to travel time to work, and 5) the marginal utility with respect to housing quality is positively related to the size of a family. These qualitative properties are consistent with my a priori expectations.

5.2. The Houston Housing Market as a Function of the Mean Air Quality.

In this subsection I want to compute how the structure of the Houston housing market depends on the mean air quality. To do that, I repeat the calculations of the previous subsection with the only difference that now I do not substitute 0.0141 imcm[5] for the mean air quality, $\bar{v}_2$, (see Table 2). The results follow.

The parameters A, ζ_0, ζ_1, ϵ_1, ϵ_2, and ξ do not change because they do not depend on the mean air quality. The housing quality index equation is given in (13), the utility function is given in (14), and the rental price equation is:

$$P = 459.36 - 25904.50\bar{v}_2 + 48.59v_1 + 1648.7v_2 - 3.32v_3 ,$$

or equivalently: $P = 459.36 - 25904.50\bar{v}_2 + 48.59h.$

The demand for housing quality equation is:

$$h = -0.647 + 33.93\bar{v}_2 + 0.505a + 0.000109I \qquad (15)$$

The demand for x follows:

$$x = -14.07 + 797.45\bar{v}_2 - 0.81a + 0.0026I \qquad (16)$$

I can now use the above results to illustrate the kind of questions that my analysis can address.

5.3. Illustration: The Willingness to Pay for an Improvement in Air Quality.

The purpose of this illustration is not to determine the precise dollar figure of the consumer benefit from an improvement in air quality. Rather, it is to illustrate how to perform a general equilibrium analysis that is accomodated by the model and to show that the previous (partial equilibrium) common practice for computing the consumer benefit from a non-marginal change in one of the characteristics of a differentiated good can yield a very different benefit figure. Much greater care would be necessary to estimate with confidence the precise dollar value of the consumer benefit from a change in the air quality distribution. Moreover, for a long run analysis the supply for houses should be made endogenous.

A consumer's willingness to pay for a y% improvement in air quality, W, is defined to be the solution to the equation:

$$V(a,I,t) = V(a,I + W,\ t + ty/100) \qquad (17)$$

where t is the mean air quality in Houston, and V(a,I,t) is the indirect utility function of an (a,I) - type consumer given that the mean air quality of the city of Houston equals t. That is, the consumer's benefit from a y% change in the mean air quality is the part of his income that he is willing to give up so that the utility after the y% change equals the utility before the y% change.

I will compute the benefit of the mean household in Houston[6] from a 1%, 5%, 10%, and 70% improvement in the mean air quality of the city. That is, I will compute W for y = 1, 5, 10, 70. The steps involved in the computation are explained next.

To obtain the indirect utility function, I substitute the demand for housing quality and the demand for numeraire good, equations (15) and (16) respectively, into the utility function, equation (14). Given the indirect utility function, I can specify the functional form of equation (17). Equation (17) is solved with respect to W, using procedures that are available in the TKSolver computer package, for W = 1, 5, 10, 70. The results are given in Table 3.

Table 3 shows that the benefit - percentage change in air quality ratio is slightly increasing in y. That is, if I multiply the percentage change y, by k, the consumer benefit increases by a factor greater than k. From a technical point of view, this result can be true within my framework because 1) the indirect utility function is quadratic in consumer income and mean air quality, and 2) the coefficients of these variables can be either positive or negative. In other words, 1) the equation that I solve for the willingness to pay, W, is non-linear in the mean air quality, and 2) the solution for W has a first derivative with respect to mean air quality that can be either increasing or decreasing in air quality. In the above application, it turned out that the relationship between mean consumer benefit and mean air quality improvements is increasing.

Ceteris paribus, changes in the mean air quality make a consumer happier because the quality of his house improves (this change in utility is decreasing in air quality improvements because the marginal utility with respect to housing quality is decreasing). However, increases in the mean air quality shift the price function for housing quality downward. This implies a redistribution of rents (from housing suppliers to consumers of housing) which lets consumers increase their utility even more. In the above application, the change in the distribution of rental prices and the assumed utility function imply that (after the mean air quality improvement) the mean consumer is able to buy a combination of goods that increases his utility at a rate that is greater than the percentage improvement in air quality.

Next, I compute the benefit from the same air quality improvements using the alternative approach[7] and I compare the results.

To compute benefits using the non-structural approach, I first estimate a marginal willingness to pay schedule, and then I integrate the marginal willingness to pay from $\bar{v}_2$ to $\bar{v}_2+\bar{v}_2 y/100$ to obtain a measure of the willingness to pay[8] for a y% change in the mean air quality of Houston. To illustrate this method, I use a price equation that is linear in air quality. The parameter estimates[9] are given in Table 4.

Given a rental price equation that is linear in air quality, the non-structural approach would define the willingness to pay in the following way[10]:

$$W = 12(AQC)(DV)$$

where DV is the change in the mean air quality of Houston, AQC is the coefficient of the air quality variable in the rental price equation, and AQC = 6701.2 (see Table 4).

Calculating the benefit of the mean household using the latter definition for the willingness to pay, I obtained the estimates[11] given in Table 5.

6. Conclusions.

From Tables 3 and 5, it can be seen that the two methods imply very different benefit figures. The reason is that if there is a change in one of the exogenous parameters of the model, the latter method (by its nature) is not appropriate for computing the willingness to pay for a change in one of the characteristics of a differentiated good.

To compute benefits, the non-structural approach uses a different method and different estimates of the price equation (the ones given in Table 4). To separate those two issues and to show the difference that arises because of differences in methods of calculation, I compute benefits using AQC = 1648.7 (the estimate of the coefficient of the air quality variable in the price equation that is given in Table 1). I obtained the benefit estimates that are given in Table 6.

The benefit figures of Table 5 are approximately 71% below the benefit figure based on the structural model (given in Table 3); this difference arises because of differences in method of calculation as well as differences in coefficients. The benefit figures of Table 6 are approximately 93% below the benefit figures of Table 3; this difference arises because of differences in method of calculation (using the same estimated coefficients). The results show that the non-structural approach can give very different benefit figures even for small changes in the mean air quality (e.g. a 1% change).

TABLE 1

ESTIMATION RESULTS - HOUSTON, TEXAS - THE RESTRICTED MODEL

	VARIABLE	COEFFICIENT	STANDARD ERROR	T-STATISTIC
EQUATION 1:	v_1	48.59252	14.452640	3.362189
	v_2	1648.671	949.02110	1.737233
	v_3	-3.318238	0.9485320	-3.498288
	INTERCEPT	93.51385	49.934550	1.872728
	SIGMA	59.95461	5.9348540	10.10212
EQUATION 2:	$v_2-\bar{v}_2$	33.92849	19.899270	1.705012
	$v_3-\bar{v}_3$	-0.06828701	0.01574382	-4.337385
	$a - \bar{a}$	-0.5044743	0.09248937	-5.454403
	$I - \bar{I}$	-0.0001091461	0.0000153241	-7.122514
	SIGMA	0.3736063	0.03698296	10.10212
	RHO	-0.2432288	12.19337	-0.01994762

N = 57

FUNCTION = 76.48

NOTE: N is the number of observations.
SIGMA is the standard deviation of the model error.
RHO is the estimate of the inter-equation error correlation.
FUNCTION is the negative of the loglikelihood function.

TABLE 2

HOUSTON STATISTICS

Mean number of rooms:	4.1281
Mean air quality:	0.014123
Mean travel time to work:	25.956
Mean number of persons in a family:	2.4998
Mean income:	15954

TABLE 3

THE BENEFIT OF THE MEAN HOUSEHOLD
ESTIMATES IMPLIED BY THE STRUCTURAL ANALYSIS

AIR QUALITY IMPROVEMENT	ANNUAL BENEFIT
1%	$ 36.56
5%	$ 209.73
10%	$ 426.19
70%	$ 3023.65

TABLE 4

PRICE EQUATION LINEAR IN AIR QUALITY

VARIABLE	COEFFICIENT	STANTARD DEVIATION	T-STATISTIC
v_1	45.76947	11.22621	4.077019
v_2	6701.209	2578.897	2.598479
v_3	-8.650030	1.574976	-5.492168
INTERCEPT	172.2020	58.04411	2.966743
SIGMA	52.08527	4.878232	10.67708

N = 57

FUNCTION = 43.70

TABLE 5

THE BENEFIT OF THE MEAN HOUSEHOLD
ESTIMATES IMPLIED BY THE PREVIOUS METHOD

AIR QUALITY IMPROVEMENT	ANNUAL BENEFIT
1%	$ 11.34
5%	$ 56.69
10%	$ 113.38
70%	$ 793.69

TABLE 6

THE BENEFIT OF THE MEAN HOUSEHOLD
ESTIMATES IMPLIED BY THE PREVIOUS METHOD
AND THE PARAMETER ESTIMATES OF TABLE 1

AIR QUALITY IMPROVEMENT	ANNUAL BENEFIT
1%	$ 2.79
5%	$ 13.95
10%	$ 27.90
70%	$ 195.27

REFERENCES

- D. Epple, Closed Form Solutions to a Class of Hedonic Equilibrium Models, Carnegie-Mellon University, mimeo, (1984).

- D.A. Giannias, Differentiated Goods: Solutions to Equilibrium Models, Carnegie-Mellon University, Ph.D. Dissertation, 1987.

- D. Harrison and D. Rubinfeld, Hedonic Housing Prices and the Demand for Amenities, J. of Environ. Econ. and Man. 5, 81-1-2 (1978).

- J. Tinbergen, On the Theory of Income Distribution, in Selected Papers of Jan Tinbergen. (Klassen, Koyck, and Witteveen, eds.), Amsterdam: North-Holland, 243-263, (1959).

Endnotes

1. The proof is similar to the proof of Proposition 1 of Giannias (1987). The general strategy of the proof was introduced by Tinbergen (1959) and extended by Epple (1984).

2. There are two solutions that satisfy the equilibrium condition. The one of them is rejected because it does not satisfy the second order condition for utility maximization.

3. I do this because the quality of housing is unobservable by econometricians.

4. According to the theory, a utility maximizing consumer chooses the quality of a differentiated good and there may exist many goods that can provide that quality. The theory was not meant to specify how the consumer chooses among those equal quality differentiated goods. In terms of the housing market application, the theory does not specify how the consumer makes a locational choice. The consumer is indifferent to all the (housing-location) combinations that can provide the quality of housing that maximizes his happiness. Since a consumer cares only about the quality of the differentiated good housing, I assume that he moves randomly to any census tract and picks a house of the quality that he is looking for. However, if in a census tract demand does not equal supply, the theory suggests that consumers do not bid prices up (that would make the price of a specific house to be different in different census tracts) but they move into another census tract; since there are no moving costs, a consumer will move into another area where he can find a house of the quality that he is looking for. (A3) and (A4) assume that this random locational choice is uncorrelated to the econometric error terms of the equations that I estimate.

5. imcm = 1/(micrograms per cubic meter).

6. The mean household of Houston is described in Table 2.

7. Harrison and Rubinfeld (1978) is an example of the non-structural approach that I am referring to in the next paragraph. Here, I am not referring to the four step estimation procedure which is the main contribution of that paper. The four step estimation procedure is the way that they apply that method (namely, the way that they estimate the marginal willingness to pay), given a quadratic "price" equation that they consider. If the price equation is linear in attributes, they use another method to compute benefit.

8. That would assume a uniform improvement in air quality. That is, an improvement in each census tract that equals the mean air quality improvement.

9. The parameter estimates have been obtained using a single equation estimation technique.

10. For example, Harrison and Rubinfeld (1978), page 92, footnote 28. For a price equation that is linear in air quality, they define willingness to pay, as well as average benefit, in exactly the same way. I recall that if the price equation is linear in attributes, they do not use their four step procedural model to compute benefit (see also footnote 7).

11. To obtain these estimates, I assumed a uniform improvement in air quality. That is, the improvement in each census tract equals the improvement in the mean air quality of the whole city.