A Structural Approach to Hedonic Equilibrium Models

by

Dimitrios A. Giannias

Working Paper No. 6

August 1988

The author wishes to thank The Jerome Levy Economics Institute for its financial support.

ABSTRACT

This paper presents a quality theory for differentiated products. Analytical solutions for the equilibrium demand for quality equation and the equilibrium price equation are computed. The model is estimated and the willingness to pay for improvements in the air quality of Houston is computed. The empirical results show that the standard non-structural approach would seriously underestimate benefits for non-marginal and small changes in air quality.

I. Introduction and Summary.

In recent years, some economists have adopted a new approach to the theory of individual choices that helps to explain a number of phenomena that are difficult to understand within the confines of the traditional economic theory. This new approach argues that the characteristics of commodities provide (directly or indirectly) utility to individuals and/or services to production processes. Houthakker (1952) pioneered this approach to the problem of quality variation and to the theory of consumer behavior.

Becker (1965), Lancaster (1966), and Muth (1966) extended Houthakker's analysis to study consumer behavior but they did not work out the properties of market equilibrium. They assumed that commodities traded in the market do not possess final consumption attributes and that consumers are also producers. The consumers use the commodities purchased in the market as inputs into a self-production function for ultimate characteristics. Rosen (1976) studies both consumer and producer behavior and the properties of market equilibrium. Rosen assumes that consumers are not producers and that all the commodities with their ultimate characteristics are readily available and traded in the market.

With few exceptions, the hedonic approach has not been analyzed thoroughly and complete hedonic equilibrium models have not been estimated. Closed form solutions to hedonic equilibrium models have not been available for any class of economies that could serve as the foundation for empirical applications, a gap I hope to fill, in part here.

Tinbergen (1959) supplied the earliest contribution to the formulation and solution of hedonic equilibrium models. Epple (1984) generalizes Tinbergen's model to treat a commodity with an arbitrary number of attributes and introduces both endogenous demand and supply. However, these models have several restrictive features. Namely, the cross partial derivative of the utility function is zero, the marginal utility with respect to the numeraire good is constant (hence the income elasticity of demand for the product is zero), the variance-covariance matrices of the exogenously given distributions have to be diagonal or satisfy other restrictions, the number of consumer characteristics equals the number of product characteristics, and the price equation parameters are not unique.

I assume that a function maps physical characteristics into a scalar quality index and that economic agents care only about the quality of the commodity that they purchase. While this is a strong assumption, this quality index technology allows me to impose weaker a priori restrictions in other respects. The result is a class of models[1] with closed-form solutions that does not have the restrictive features enumerated above. In this paper, I present one of those models and an application. The theory characterizes market equilibrium and the application investigates how far one can go with closed form solutions and how well the resulting model fits the data. The application shows that it is feasible to estimate and test a closed-form model.

Section II reviews the non-structural approach. Section III introduces the theoretical model that I use to illustrate the kind of analysis that the structural approach can perform. This model assumes that the income and the supply distributions are exogenous and that consumers use the services of only one unit of the differentiated good. However, versions of the same basic model can relax these assumptions (see Giannias (1987)). An application is discussed in Section IV. Section V investigates extensions of the basic model. Concluding remarks are presented in Section VI.

II. The Non-Structural Approach.

Observed product prices and the specific amounts of characteristics can provide estimates for implicit or hedonic prices. The non-structural approach uses this information to derive the demand functions for the characteristics of a differentiated product. These demand functions can be used to compute the willingness to pay for marginal or non-marginal changes in the product characteristics, for example, Harrison and Rubinfeld (1978) and Ridker and Henning (1967). Harrison and Rubinfeld (1978) advanced the state-of-the-art by recognizing that the derivative of the price function is not a good approximation for assessing the benefit of changes in product characteristics.

As demonstrated in Epple (1987), most of the work that uses the hedonic approach is unsatisfying because the estimation methods do not yield consistent estimates. Bartik (1987) and Palmquist (1984) are two exceptions. No previous application contains a structural analysis. Depending on the structure of the economy and on the question that we are interested in, we do not always need to compute closed form solutions and make a structural analysis. For example, the standard approach can estimate the price equation and the parameters of the demand functions for product characteristics. However, structural analysis is needed to compute the effects of changes in exogenous parameters. Changes in exogenous parameters change the coefficients in the equilibrium hedonic price function and non-structural approaches cannot take account of such changes.

III. The Economic Model.

I consider a competitive economy in which individuals consume a differentiated good and the numeraire good, x. I assume that consumers use one unit of the differentiated good. For example, this differentiated good could be a house, an automobile, or a computer.

The differentiated good can be accurately described by a vector, v, of objectively measured characteristics. I assume that the consumers care only about the quality index, h, of the differentiated product. The quality, h, is a scalar and a function of the vector of physical characteristics, v. The model lets consumers have different utility functions and income. Each consumer can be described by a (1x2) vector z, where $z = [\varsigma \ I]$, I is the consumer income, and ς is a utility parameter. z is assumed to follow a multi-normal distribution with a mean $\bar{z}$ and a variance Σ_z. Let it be:

$$N(\bar{z},\Sigma_z) \tag{1}$$

Given ς, $U(h,x;\varsigma)$ is the utility that a consumer obtains from x and from the services of a differentiated good of h-quality. The utility function is assumed to be a quadratic of the following form.

$$U(h,x;\zeta) = \delta + \zeta h + \theta x + 0.5 \xi h^2 + \omega x h \tag{2}$$

where δ, ζ, ξ, ω, and θ are utility parameters (scalars).

A consumer with income I and a utility parameter ζ solves the following optimization problem:

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

$$\text{with respect to } h, x$$

$$\text{subject to } I = P(h) + x$$

where P(h) is the equilibrium price equation; it gives the price of the differentiated good as a function of the quality index, h. Eliminating x, the first order condition for the consumer's optimization problem is equivalent to:

$$P_h(h) = \frac{U_h(h, I-P(h);\zeta)}{U_x(h, I-P(x);\zeta)} \tag{3}$$

where $P_h(h)$ and $U_i(h,x;\zeta)$ are the first partial derivatives of P(h) and $U(h,x;\zeta)$ with respect to h and i respectively, i = h, x.

The supply for the differentiated product is exogenously given and the quality h follows an exogenously given normal distribution with a mean $\bar{h}$ and a variance σ^2. Let it be:

$$g(h) = N(\bar{h},\sigma^2) \tag{4}$$

The optimum decisions of consumers and sellers depend on the equilibrium price equation P(h). The price equation is determined so that buyers and sellers are perfectly matched. In equilibrium, no one of the economic agents can improve his position, all of their optimum decisions are feasible, and the price equation P(h) is determined by the distribution of consumer tastes and income, and by the supply for the differentiated product.

PROPOSITION. The price equation that equilibrates the market described above is

$$P(h) = \pi_0 + \pi_1 h \ , \tag{5}$$

where

$$\pi_1 = (\xi + A)/(2\ \omega) \tag{6}$$

$$\pi_0 = (-\ \theta\ \pi_1 + t\ \bar{z}' - A\ \bar{h})/\omega \tag{7}$$

$$t = [1 \quad \omega], \text{ and} \tag{8}$$

$$A = (\ t\ \Sigma_z\ t'\ /\ \sigma^2\ )^{0.5} \tag{9}$$

(Hereafter, a prime "'" will always denote the transpose of a vector or matrix).

Proof[2]

Substitute equations (2) and (5) into equation (3) and solve for h to obtain the equilibrium demand for h. The equilibrium demand for h, i.e., the demand function after substituting out P(h), is given by the following equation:

$$h = (-\ \omega\ \pi_0 - \theta\ \pi_1 + t\ z')/(2\ \omega\ \pi_1 - \xi) \tag{10}$$

where t is given in (8).

The equilibrium demand for h is linear in z. Therefore (1) and (10) imply that the aggregate equilibrium demand for h follows a normal distribution. Let it be: $f(h) = N(\bar{h}_d, \sigma_d^2)$, where $\bar{h}_d$ is the mean, σ_d^2 is the variance,

$$\bar{h}_d = (-\ \omega\ \pi_0 - \theta\ \pi_1 + t\ \bar{z}')/(2\ \omega\ \pi_1 - \xi), \text{ and} \tag{11}$$

$$\sigma_d^2 = t\ \Sigma_z\ t'/(2\ \omega\ \pi_1 - \xi)^2 \tag{12}$$

The assumption that the quality index h follows the exogenous distribution given in (4) implies that the equilibrium condition "Aggregate Demand = Aggregate Supply", that is, $f(h)\ dh = g(h)\ dh$, is equivalent to:

$$\bar{h}_d = \bar{h}, \text{ and} \tag{13}$$

$$\sigma_d^2 = \sigma^2 \tag{14}$$

where $\bar{h}_d$, σ_d^2, $\bar{h}$, and σ^2 are given in (11), (12), and (4).

The second order condition requires: $(2\ \omega\ \pi_1 - \xi) > 0$. From inspection of (6) it is seen that this second order condition is satisfied. It can be verified that the following equations, $\pi_1 = (\xi - A)/(2\ \omega)$ and $\pi_0 = (-\ \theta\ \pi_1 + t\ \bar{z}' - A\ \bar{h})/\omega$ also satisfy the equilibrium equations (13) and (14). However, this solution is ruled out because it does not satisfy the second order condition. QED

To illustrate that the price equation is an equilibrium relationship that incorporates features of tastes, supply, and the distributions of income and parameters of taste and supply, I present the following example.

Consider an economy in which consumers have identical preferences that can be described by the following utility function: $U = 100 + h + 2x + 0.5\ h^2 + xh$, where x and h are defined above. Consumers are assumed to use the services of one differentiated good. Let the consumer income follow an exogenously given normal distribution with a mean that is equal to 550 and a variance that is equal to 400. The quality of the differentiated good is assumed to follow an exogenously given normal distribution that is given in (4). Let this distribution have a mean equal to 2 and a variance that is equal to 1. The Proposition implies that the hedonic price equation is P(h) = 490 + 10.5 h. Suppose now that the mean of the product quality distribution decreases by one unit. The Proposition implies that the price equation becomes: P(h) = 510 + 10.5 h. If, in addition to the previous one unit change in the mean, the variance of the distribution of the product quality decreases by 0.75 of a unit, the price equation becomes: P(h) = 470 + 20.5 h.

IV. An Application.

The model that I presented in the previous section can be used for a study of the residential housing market. The empirical example that follows shows that it is feasible to estimate and test a closed-form model. The results of the empirical example will be used to investigate the willingness to pay for clean air.

IV.A. 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 the housing unit (number of rooms), v_2 is an air quality index, and v_3 is the travel time to work. v is assumed to follow an exogenously given multi-normal distribution.

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

$$h = \epsilon\ v' \tag{15}$$

where $\epsilon = [\epsilon_0\ \epsilon_1\ \epsilon_2]$ is a vector of parameters[3].

Consumer preferences are described by utility functions. The 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:

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

$$\text{with respect to } h,\ x$$

$$\text{subject to}\quad I = 12\ P(h) + 365\ x$$

where I is the annual income of a consumer, P(h) is the (monthly) rental price equation, 12 is the number of months in a year, and 365 is the number of days in a year. The utility function is:

$$U(h,x;a) = \delta + (\zeta_0 + \zeta_1\ a)\ h + 0.5\ \xi\ h^2 + x\ h \tag{16}$$

where δ, ζ_0, ζ_1, and ξ are utility parameters. The vector [a I] follows an exogenously given multi-normal distribution.

The Proposition of Section III implies that the price equation is:

$$P = (365/12)\ [\zeta_0 + \zeta_1\ \bar{a} + (\bar{I}/365) - B\ \bar{h} + 0.5\ (\xi + B)\ h] \tag{17}$$

where a "¯" over a variable denotes the mean of the variable, $\bar{h} = \epsilon\ \bar{v}'$, $B = (\ s\ \Sigma_d\ s'\ /\ \sigma^2\ )^{0.5}$, $\sigma^2 = \epsilon\ \Sigma_v\ \epsilon'$, $s = [\zeta_1\ 1]$ and $d = [a\ I]$ are two (1x2) vectors, Σ_v is the variance-covariance matrix of the exogenously given distribution of the vector of housing characteristics, v, and Σ_d is the variance-covariance matrix of the (1x2) vector d.

The Proposition and equation (10) imply that the equilibrium demand for h is:

$$h = \bar{h} + \zeta_1 (a - \bar{a}) / B + (I - \bar{I}) / (365 B) \qquad (18)$$

IV.B. The Econometric Model.

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. Assuming an additive error term on the price equation and on the equilibrium demand for housing quality, I obtain:

$$P = c + \beta_1 v_1 + \beta_2 v_2 + \beta_3 v_3 + u_1 \text{ , and} \qquad (19)$$

$$h = \gamma - \epsilon_3 a - \epsilon_4 I + u_2 \qquad (20)$$

where

$$c = (365/12) (\zeta_0 - B \gamma) \qquad (21)$$

$$\beta_{i+1} = (365/24) (\xi + B) \epsilon_i \text{ , for } i = 0, 1, 2 \qquad (22)$$

$$\gamma = \bar{v}_1 + \epsilon_1 \bar{v}_2 + \epsilon_2 \bar{v}_3 + \epsilon_3 \bar{a} + \epsilon_4 \bar{I} \qquad (23)$$

$$\epsilon_3 = - \zeta_1/B \qquad (24)$$

$$\epsilon_4 = - 1/(365 B) \text{ , and} \qquad (25)$$

u_1 and u_2 are the econometric errors of the first and second equations respectively. They are assumed to satisfy: (A1) u_1 and u_2 are uncorrelated, (A2) a and I are uncorrelated to u_1 and u_2, and (A3) v_1, v_2, and v_3 are uncorrelated to u_1. These assumptions may be motivated, for example, by thinking of u_1 as a measurement error in price and u_2 as unmeasured buyer characteristics that are uncorrelated with measured buyer characteristics. The complete model consists of equations (15), (19), and (20).

IV.C. Estimation of The Reduced Form Equations.

To estimate the complete model, I introduce a four step estimation procedure. This estimation method yields consistent parameter estimates and uses the restrictions that are implied by the structure of the model, namely,

$$\epsilon_1 = \beta_2/\beta_1, \text{ and} \tag{26}$$

$$\epsilon_2 = \beta_3/\beta_1 \tag{27}$$

I estimate the model for Houston, Texas 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[4]. Unlike other work, e.g. Harrison and Rubinfeld (1978), the model implies that it is legitimate to use census tract data because 1) the price equation is linear in product characteristics, and 2) the equilibrium demand for product quality is linear in consumer income and family size. Hence, the model has convenient aggregation properties that allow mean values of census tract data to be used. The estimation method follows.

STEP 1: I estimate the price equation by ordinary least squares (which is appropriate under assumption A3). The parameter estimates are given in Table 1. They imply that the price equation is:

$$P = 172.2 + 45.77\ v_1 + 6701.21\ v_2 - 8.65\ v_3 \tag{28}$$

STEP 2: Given (26), (27), and the results of the previous step, I can obtain estimates for ϵ_1 and ϵ_2. This and the normalization $\epsilon_0 = 1$ enable me to obtain that the housing quality index equation is:

$$h = v_1 + 146.41\ v_2 - 0.189\ v_3 \tag{29}$$

STEP 3: I use the above specified housing quality equation to construct an estimated series for the housing quality for each census tract of my data set.

STEP 4: I use the housing quality indices that I obtained in step 3 to estimate the following equation by ordinary least squares:

$$h = \gamma - \epsilon_3\ a - \epsilon_4\ I + u_2 \tag{30}$$

Ordinary least squares is appropriate under assumptions A1 to A3. Deviations between the actual housing quality and its estimate (estimated from equation (29)) are measurement errors in the dependent variable in equation (30) and hence do not affect the consistency of ordinary least squares. The parameter estimates are given in Table 2. They imply that the equilibrium demand for housing quality is:

$$h = -1.52 + 0.026\ a + 0.000172\ I \tag{31}$$

IV.D. Test of the Model.

To see if the model makes a significant contribution to explaining the data, I tested the hypothesis that all the parameters of the equation (19) equal zero, that is, $\beta_1 = \beta_2 = \beta_3 = 0$. An F-test implies that this hypothesis is rejected at the 1% significance level. A similar F-test rejects the hypothesis that all the parameters of the second equation, equation (20), equal zero at the 1% significance level.

The t-statistics (see Tables 1 and 2) show that the size of a house and the travel time to work variable (which are expected to be the main determinants of the rent), as well as the income (which is expected to be the main determinant of the equilibrium demand for housing quality) are significant at the 1% significance level. Moreover, all coefficients have the anticipated signs in both equations.

For the residential housing market, I expect the parameters ϵ_1, ϵ_2, and ζ_1 to satisfy: $\epsilon_1 > 0$, $\epsilon_2 < 0$, and $\zeta_1 > 0$. That is, I expect 1) the housing quality to increase as air quality increases, 2) the housing quality to decrease as the travel time to work increases, and 3) the utility that is obtained from each additional unit of housing quality to increase as the size of the family increases. The parameter estimates obtained in this section show that the first two of the above inequalities are satisfied. In Section V.E., it is shown that the third inequality is also satisfied.

To investigate the internal consistency of the theory (given additive error terms), I test the joint normality of prices and product characteristics, and of product quality, family size and income. To be more specific, I test the null hypothesis that u_i is normally distributed, i = 1, 2. An omnibus test[5] using $X^2(\sqrt{b_{1i}}) + X^2(b_{2i})$ provides evidence in favor of the null hypothesis, where $X^2(\sqrt{b_{1i}})$ and $X^2(b_{2i})$ are standardized normal equivalents to the sample skewness, $\sqrt{b_{1i}}$, and kurtosis, b_{2i}, i = 1, 2, $\sqrt{b_{11}} = -0.174$, $\sqrt{b_{12}} = 0.046$, $b_{21} = 2.288$, and $b_{22} = 2.791$. The usual Kolomogorov D statistic implies the same results[6]. These normality tests imply that the price equation, given v, and the demand for housing quality, given [a I], are linear in v and [a I] respectively.

Figures 1 and 2 provide a graphical assessment of normality. The normal probability plot for both residuals results in a reasonably straight line indicating that both residuals are normal, see Srivastava and Carter (1983).

In Section IV.A., it is assumed that the quality index equation is linear in v, that the utility parameter ζ is linear in family size, and that the utility function is quadratic. These assumptions imply (19) and (20). To investigate whether the model is misspecified by the omission of some variables, a Ramsey test, see Ramsey (1969), is applied on (19) and (20). This test provides evidence that there are not any variables omitted from either (19) or (20).

To investigate further the above issue, the following price equation is estimated: $P = \beta_0 + \Sigma_i \beta_i v_i + \Sigma_{i,j} \beta_{ij} v_i v_j$. An F-test provides evidence in favor of the null hypothesis that $\beta_{ij} = 0$ for all i and j, where i = 1, 2, 3 and j = 1, 2, 3. Moreover, the four step estimation method of Section IV.C. is repeated with equation (30) being replaced by: $h = \gamma_0 + \gamma_1 a + \gamma_2 I + \gamma_3 a I + \gamma_4 a^2 + \gamma_5 I^2$. An F-test provides evidence in favor of the null hypothesis that $\gamma_3 = \gamma_4 = \gamma_5 = 0$.

In addition to the above, the following regression model is considered: $P^{(\lambda)} = \beta_1 v_1^{(\lambda)} + \beta_2 v_2^{(\lambda)} + \beta_3 v_3^{(\lambda)}$, where $z^{(\lambda)} = (z^\lambda - 1)/\lambda$ for $\lambda \neq 0$ and $z^{(\lambda)} = \log(z)$ for $\lambda = 0$, z = P, v_1, v_2, v_3, h, a, I.

Considering the more practically interesting cases of $\lambda = 1$ and $\lambda = 0$ and applying the Box-Cox procedure, see Box and Cox (1962), it is obtained that $\lambda = 1$ yields a smaller residual variance. The four step estimation procedure of Section IV.C. is then repeated with (30) being replaced by: $h^{(\lambda)} = \gamma_1 a^{(\lambda)} + \gamma_2 I^{(\lambda)}$. Application of the Box-Cox procedure on the last equation for $\lambda = 1$ and $\lambda = 0$ implies that $\lambda = 1$ yields a smaller residual variance. These estimation results indicate that a linear specification is preferred to a log-log for both equations.

In the Appendix, models that assume Cobb-Douglas utility functions, a Tinbergen (1959)-Epple (1984) type of quadratic utility function, and log-log and log-linear in product characteristics quality index (and price) equations have been tested and found to be inconsistent with the data.

The above tests provide evidence in favor of the internal consistency of the theory of Section III and of the additional assumptions of Section IV.A with the data. These tests and the qualitative properties of the estimated model suggest that our formulation is not inappropriate for analyzing the structure of the housing market of Houston.

IV.E. Structural Analysis.

The parameter estimates that I obtained in the previous section allow me to analyze the structure of the housing market of Houston and specify how the structure depends on the mean of the air quality distribution. The latter enables me to address interesting questions that a non-structural approach cannot.

IV.E.(i). The Houston Housing Market.

Given the parameter estimates obtained in Section IV.C. and equations (21)-(27), I can compute the parameters of the utility function and the equilibrium demand for the numeraire good[7]. They are respectively given by the following equations:

$$U(h,x;a) = \delta + (-18.55 + 0.41\ a)\ h + x\ h - 6.46\ h^2\ , \text{ and} \qquad (32)$$

$$x = -3.39 - 0.039\ a + 0.0025\ I.$$

We can now see that 1) the rent is positively related to the quality of a house[8], 2) the equilibrium demand for housing quality is positively related to the size of the family and income (see equation (31)), 3) the housing quality is positively related to air quality and negatively to travel time to work (see equation (29)), and 4) the marginal utility with respect to housing quality is positively related to the size of a family (see equation (32)). These qualitative properties are as one would intuitively expect.

IV.E.(ii) The Houston Housing Market and the Mean Air Quality.

In this subsection, I illustrate how the preceding results can be used in a structural analysis of the value of a change in air quality. To do that, I first repeat the calculations of the previous subsection treating mean air quality, $\bar{v}_2$, as a variable rather than fixing it at 0.0141 imcm[9] its sample mean of. The results follow.

The parameters B, 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 and the utility function are given in (29) and (32) respectively. The equilibrium rental price equation, the equilibrium demand for housing quality, and the equilibrium demand for the numeraire good are functions of the mean air quality. They are respectively equal to:

$$P = 1172.47 - 70941.13\ \bar{v}_2 + 45.77\ h\ ,$$

$$h = -3.58 + 146.41\ \bar{v}_2 + 0.026\ a + 0.000172\ I, \text{ and} \tag{33}$$

$$x = -33.18 + 2112.7\ \bar{v}_2 - 0.039\ a + 0.0025\ I \tag{34}$$

The above results are used to illustrate the kind of questions that a structural analysis can address. The purpose of the illustration is not to determine the precise dollar figure of the willingness to pay for 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 willingness to pay for a non-marginal change in one of the characteristics of a differentiated good can yield a very different benefit figure.

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

$$V(a,I,t) = V(a,I+W,t+y/100) \tag{35}$$

where t is the mean air quality in Houston, and V(a,I,t) is the equilibrium 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, taking account of equilibrium price adjustments, equals the utility before the y% change.

I compute the benefit to the mean household in Houston of a 1%, 2.5%, 5%, 7.5%, 10%, 12.5%, and 15% improvement in the mean air quality of the city. That is, I compute W for y = 1, 2.5, 5, 7.5, 10, 12.5, 15. The steps involved in the computation are explained next.

To obtain the equilibrium indirect utility function, I substitute the equilibrium demands for housing quality and numeraire good, equations (33) and (34) respectively, into the utility function, equation (32). Into this equilibrium indirect utility function I substitute the mean income, the mean number of persons in a household, and the mean air quality of Houston (see Table 3). With these substitutions equation (35) can be written as: $A W^2 + B(y) W + G(y) = 0$, where W is the willingness to pay (in thousands of dollars), A = 0.2388873, and the parameter values of B(y) and G(y) depend on y (the percentage improvement in the mean air quality) and they are given in Table 4.

Solving the last equation[10] with respect to W for y = 1, 2.5, 5, 7.5, 10, 12.5, and 15, I obtain the benefit figures that are given in Table 5. Next, I contrast these results to the ones obtained using the non-structural approach and I compare the results. To compute benefits using the non-structural approach, 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[11] for a y%

change in the mean air quality of Houston. To illustrate this method, I use the price equation given in Table 1.

Given a rental price equation that is linear in air quality (see Section IV.D.), the non-structural approach defines the willingness to pay in the following way[12]: W = 12 (AQC) (DV), where DV is the change in the mean air quality of Houston, and AQC = 6701.2 is the coefficient of the air quality variable in the rental price equation (see Table 1). Calculating the benefit of the mean household using the latter definition for the willingness to pay, I obtain the estimates given in Table 6.

From Tables 5 and 6, we can now see that the two methods give very different benefit figures even for small changes in the mean air quality (e.g., a 1% change). The benefit figures of Table 6 are 90.5% below the benefit figure based on the structural model (given on Table 5). This difference arises only because of differences in method of calculation, since the same price equation parameters were used for both calculations. Hence, the non-structural approach does not give a good approximation to the currently calculated measure of willingness to pay.

For a complete investigation of the problem I study, the effects of air quality improvements on the suppliers should be examined. Air quality improvements shift the price equation for housing quality downwards and the housing quality distribution changes. Total Net Benefit equals the sum of Total Consumer Benefit and Total Supplier Benefit. To get an idea about the magnitude of that figure, I multiplied the sum of the mean consumer benefit and the annual change in rent revenues of the mean house by the total number of households in Houston (602,696) for several air quality improvements. These results, as well as the Net Social Benefit per Household, are given in Table 7. The results imply that an Automobile Emission Control Policy that improves air quality by 10% is justified if it does not cost more than $106.87 per household. These results also show that the biggest effect of a uniform increase in air quality is a distributional effect that is implied by a drop in rental prices (the net social benefit per household is approximately 9% of the net tenant benefit in Table 5).

V. Extensions of the Basic Model.

To apply the theory of Section III, assumptions about the utility parameter ς and the quality index equation must be introduced. In Section IV, for example, it is assumed that ς is linear in family size and that the quality index equation is linear in the vector of product characteristics, v.

In general, the utility parameter ς can be a polynomial function of a, $\varsigma(a)$, of degree n, and the quality index equation can be a polynomial function of product characteristics, v, and a, h(v,a), of degree m, where a is a vector of characteristics that specifies the type of a consumer. Independently of the degrees of the polynomial functions $\varsigma(a)$ and h(v,a), n and m respectively, the four step estimation method of Section IV can be applied and all the structural parameters of interest can be identified.

The latter generalization (i) allows some flexibility in letting the data determine the appropriate functional forms, (ii) makes the equilibrium price equation non-linear in product characteristics, and (iii) allows the first derivative of the utility function with respect to a product characteristic be a function of the consumer type and/or non-linear in product characteristics. The advantage of this approach is that it allows researchers to test the internal consistency of the assumed structure and functional forms, for example, the normality of h and of $[\varsigma\ I]$ and the assumed functional forms must be consistent with the data.

The methodology used to prove the Proposition of Section III can be used to specify the restrictions among the parameters of the equilibrium price equation and the rest of the parameters of the model in cases that some of the assumptions of Section III are relaxed; for example, if the utility function is a cubic or of a higher degree, if in addition to ς the rest of the utility parameters are functions of the type of the consumer, or if the equilibrium price equation is non-linear in product quality. However, the four step estimation method will not be applicable in those cases.

VI. Conclusions.

To estimate the willingness to pay for a non-marginal change in air quality or another attribute, the non-structural approach takes the marginal willingness to pay schedule as given. That (implicitly) assumes that the air quality distribution does not change; a change in the air quality distribution shifts the price equation and the marginal willingness to pay curve. Consequently, this method cannot be used to estimate the benefit from a policy that implies a non-marginal change in the air quality distribution or other exogenous parameters. The empirical results show that this method could even miscalculate benefits of small changes in the air quality distribution.

The structural approach can provide an estimate for the consumer utility function and can compute the changes in the equilibrium demand for housing quality and numeraire good that are implied by changes in exogenous parameters. Consequently, it can compute the willingness to pay for such changes. The model that I present in this paper is offered for this kind of structural analysis. It is also an important result that we do not need data for more than one city or time series data in order to estimate the proposed structural model. This is not necessarily the case with a non-structural approach. For example, Witte's and al (1979) experiment cannot be replicated with data from only one city, see Brown and Rosen (1981).

TABLE 1

THE PRICE EQUATION

VARIABLE	COEFFICIENT	STANDARD ERROR	T-STATISTIC
v_1	45.76947	10.76549	4.251498
v_2	6701.209	3365.305	1.991263
v_3	-8.650030	1.74207	-4.965374
INTERCEPT	172.2020	66.23065	2.600035

N = 57

MULTIPLE CORRELATION COEFFICIENT = 0.40

NOTE : N is the number of observations.

TABLE 2

THE DEMAND FOR HOUSING QUALITY EQUATION

VARIABLE	COEFFICIENT	STANDARD ERROR	T-STATISTIC
I	0.000172258	0.00002530637	6.806901
a	0.02581847	0.1270296	0.2032476
INTERCEPT	-1.522414	0.5712655	-2.664985

N = 57

MULTIPLE CORRELATION COEFFICIENT = 0.51

TABLE 3

HOUSTON STATISTICS.

Mean air quality:	0.0141
Mean number of persons in a family:	2.5
Mean income:	15954

TABLE 4

B( ) AND G( ) PARAMETER VALUES

y	B(y)	G(y)
1	3.6626292	0.438379
2.5	3.7480956	1.101240
5	3.8905211	2.239821
7.5	4.0329465	3.420860
10	4.1753720	4.644353
12.5	4.3177916	5.910249
15	4.4602505	7.218973

TABLE 5

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

AIR QUALITY IMPROVEMENT	ANNUAL BENEFIT
1 %	$ 120.64
2.5%	$ 299.53
5 %	$ 597.64
7.5 %	$ 895.76
10 %	$ 1193.87
12.5%	$ 1491.97
15%	$ 1791.51

TABLE 6

THE BENEFIT OF THE MEAN HOUSEHOLD
ESTIMATES IMPLIED BY THE NON - STRUCTURAL METHOD

AIR QUALITY IMPROVEMENT	ANNUAL BENEFIT
1 %	$ 11.34
2.5%	$ 28.35
5 %	$ 56.69
7.5%	$ 85.04
10 %	$ 113.38
12.5 %	$ 141.73
15%	$ 170.08

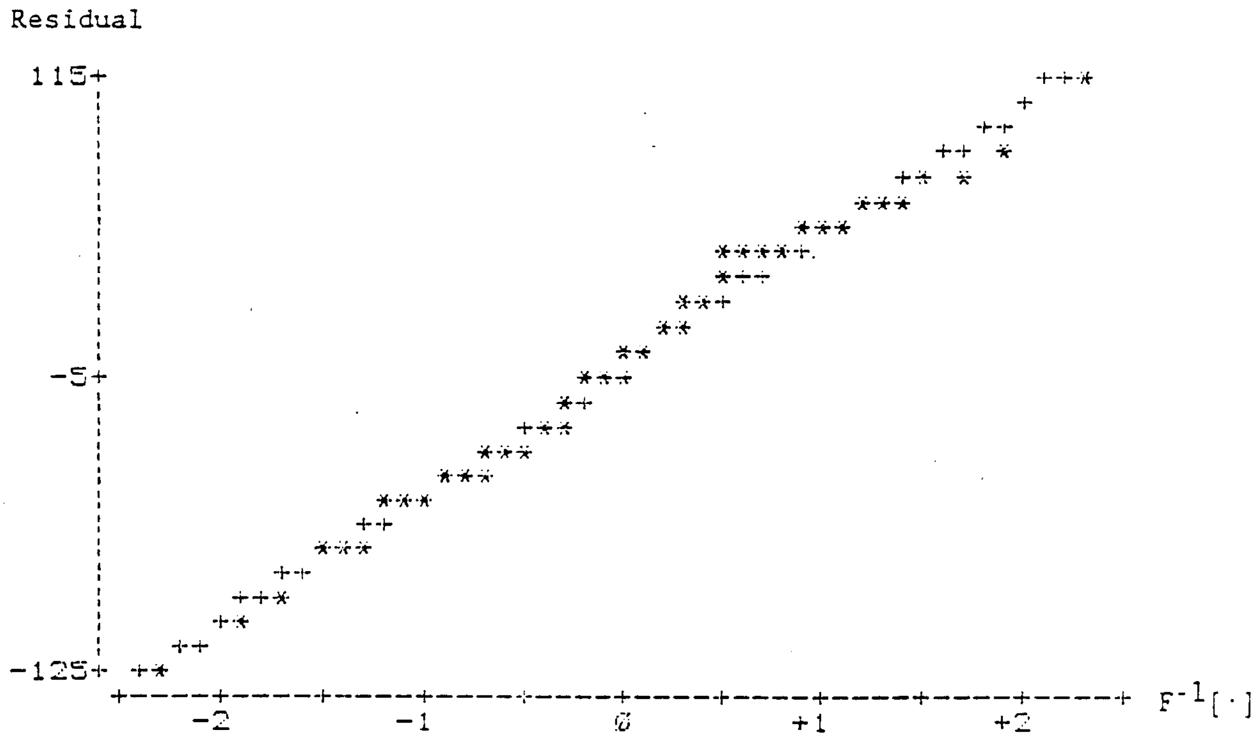
TABLE 7

TOTAL NET BENEFIT AND NET SOCIAL BENEFIT PER HOUSEHOLD IN HOUSTON, TEXAS

AIR QUALITY IMPROVEMENT	TOTAL NET BENEFIT	NET SOCIAL BENEFIT PER HOUSEHOLD
1%	$ 7,205,291	$ 11.96
2.5%	$ 16,800,633	$ 27.88
5%	$ 32,666,183	$ 54.20
7.5%	$ 48,177,650	$ 79.94
10%	$ 64,411,387	$ 106.87
12.5%	$ 80,348,598	$ 133.32
15%	$ 97,111,201	$ 161.13

FIGURE 1

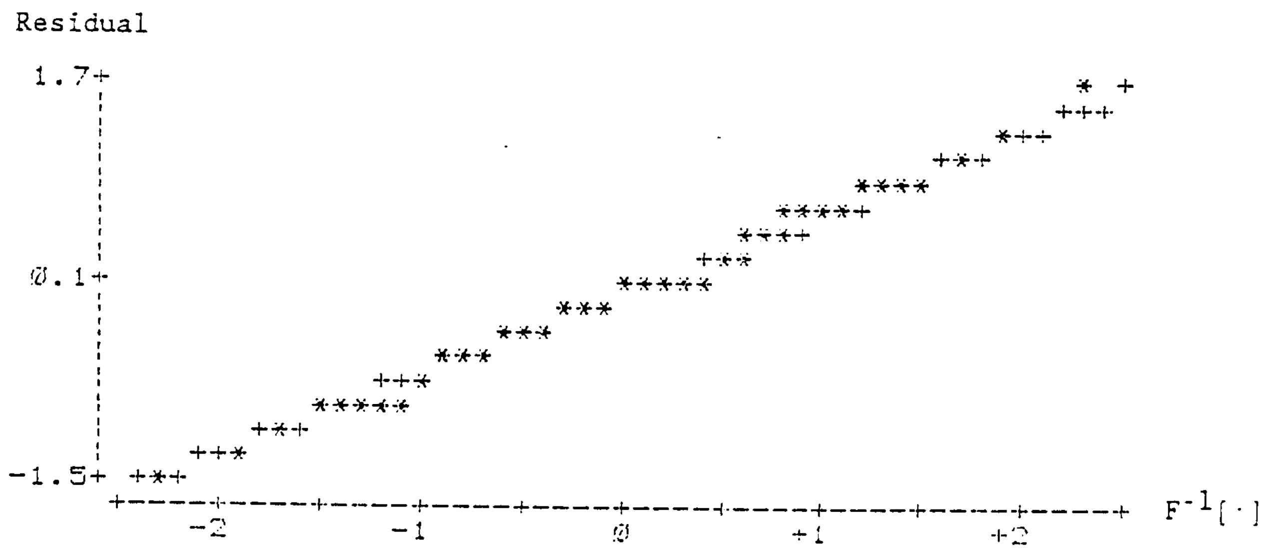
NORMAL PROBABILITY PLOT OF THE RESIDUALS OF THE PRICE EQUATION



NOTE: Asterisks (*) mark the data values and the horizontal coordinate is $F^{-1}[(r_i - 3/8)/(n + 1/4)]$, where r_i is the rank of the data value, F^{-1} is the inverse of the standard normal distribution function, and n is the number of non-missing data values. The plus signs (+) provide a reference straight line that is drawn using the sample mean and standard deviation.

FIGURE 2

NORMAL PROBABILITY PLOT OF THE RESIDUALS OF THE DEMAND FOR HOUSING QUALITY EQUATION



APPENDIX

In the Appendix it is investigated whether the data is consistent with other structural models that either generate observationally equivalent behavior or assume alternative specifications for the utility function, quality index equation, and price equation. Some of the following models introduce the product characteristics directly into the utility function.

First it is assumed that the housing price equation is linear in v, that is, $P = \pi_0 + \Sigma_i \pi_i v_i$, and that heterogeneous consumers have preferences that can be described by the following utility function: $U(v,x;a) = \ln(D) + \gamma_4 \ln(x) + \Sigma_i \gamma_i \ln(v_i)$, where x, a, and v_i, i = 1, 2, 3, are defined in Section IV, $\gamma_i = \gamma_{0i} + \gamma_{1i}$ a, i = 1, 2, 3, $\gamma_4 = 1 - \Sigma_i \gamma_i$, and D and γ_{ij} are parameters for all i and j.

The above formulation implies that the demand for v_i, i = 1, 2, 3, is the following: $v_i = c_{0i} + c_{1i} a + c_{2i} I + c_{3i} a I$ (36)

where the parameters satisfy: $c_{1i} + c_{ji} \pi_0 = 0$, j = 2, 3 (37)

The null hypothesis that the parameters of (36) satisfy (37) is rejected at the 1% significance level.

Alternatively, it is assumed that consumers have identical preferences that can be described by the following utility function:

$$U(h,x) = \ln(G) + b \ln(x) + (1 - b) \ln(h) \quad (38)$$

and that the price equation is: $P(h) = \pi h$ (39)

where b, G, and π are parameters, x is the numeraire good, and the housing quality is a function of v, h = h(v).

First it is assumed that the housing quality index equation is $\ln(h) = \epsilon_1 \ln(v_1) + \epsilon_2 \ln(v_2) + \epsilon_3 \ln(v_3)$, where v_i is defined in Section IV and ϵ_i is a parameter, i = 1, 2, 3. This specification implies that the price equation and the demand for housing quality are equivalent to:

$$\ln(P) = m_0 + m_1 \ln(I) + m_2 \ln(v_1) + m_3 \ln(v_2) + m_4 \ln(v_3) \qquad (40)$$

$$h = b\ I/\pi \qquad (41)$$

where

$$m_{i+1} = \epsilon_i \quad , i = 1, 2, 3 \qquad (42)$$

$$m_0 = \ln[F(U^*(I))] \qquad (43)$$

$$F(U^*(I)) = [G/U^*(I)]^{1/b}\ b\ (1-b)^{(1-b)/b} \quad , \text{ and} \qquad (44)$$

$U^*(I)$ is the maximum utility of a consumer with income I.

The model can be estimated[13] using the following four step method. STEP 1: estimate the price equation, equation (40), by ordinary least squares. STEP 2: use (42) and the results of the previous step to obtain estimates for the parameters of the quality index equation. STEP 3: use the specified housing quality equation to construct an estimated series for the housing quality for each census tract of the data. STEP 4: use the housing quality indices that are obtained in step 3 to estimate (41). In step 1 the following assumptions about m_0 were considered:

$$m_0 = m_{01} + m_{02}\ I \quad , \text{ and} \qquad (45)$$

$$m_0 = c + \Sigma_i\ m_{0i}\ D_i \qquad (46)$$

where c and m_{0i} are parameters for all i, and $D_1 = 1$ if income is in [0,5500) and 0 else, $D_2 = 1$ if income is in [5500,10500) and 0 else, $D_3 = 1$

if income is in [10500,15500) and 0 else, ...
The null hypothesis that the demand for h is proportional to income is tested and rejected at the 1% significance level under both alternative specifications (45) and (46).

The latter model, given in (38) and (39), is also considered under the assumption that the housing quality equation is: $\ln(h) = \epsilon_1 v_1 + \epsilon_2 v_2 + \epsilon_3 v_3$. This specification implies that the demand for housing quality and the price equation are given respectively in (41) and the following equation: $\ln(P) = m_0 + m_1 \ln(I) + m_2 v_1 + m_3 v_2 + m_4 v_3$, where the relationships among the parameters of the model are given in (42), (43), and (44). The four step method that is described above is used to estimate the model. The null hypothesis that the demand for h is proportional to income is tested and rejected at the 1% significance level under both alternative specifications (45) and (46).

Finally, experimentation with the Tinbergen (1959)-Epple (1984) formulation shows that this is not an appropriate formulation because the null hypothesis that the income elasticity of the demand for v_i, $i = 1, 2, 3$, is zero is rejected at the 1% significance level.

ENDNOTES

1. This class of models is studied in Giannias (1987).

2. The general strategy of the proof of Proposition was introduced by Tinbergen (1959) and extended by Epple (1984).

3. In general, a linear quality index equation is less restrictive than might at first appear since the elements of v can be arbitrary functions of measured product characteristics. (15) does not imply that consumers have to agree on a ranking of housing units because they are not assumed to have identical preferences.

4. The air quality variable, v_2, is assumed to be equal to the inverse of the air pollution variable (Particulate matter).

5. See D' Agostino and Pearson (1973). For both tests the following composite test statistic is used: $(N/6)\ (\sqrt{b_{1i}})^2 + (N/24)\ (b_{2i} - 3)^2$, where N is the number of observations, i = 1, 2. The statistic is distributed as a χ^2 with v = 2 degrees of freedom.

6. This statistic is available using procedures in the SAS computer package.

7. The equilibrium demand for the numeraire good is obtained from the budget constraint after substituting out the equilibrium price equation and the

equilibrium demand for housing quality.

8. To see this, note that (26), (27), (28), and (29) imply the following equation: $P = 172.2 + 45.77\ h$.

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

10. For each equation, there are two solutions. For each equation, the one of the two solutions is rejected because it indicates a willingness to pay that is greater than the mean consumer income.

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

12. For example, this approach is used by Harrison and Rubinfeld (1978), page 92, footnote 28. That is, for a price equation that is linear in air quality, Harrison and Rubinfeld do not use their four step procedural model to compute benefit.

13. Note that all the parameters of the model can be identified.

REFERENCES

- Bartik, T.J., "The Estimation of Demand Parameters in Hedonic Price Models", J.P.E. 95 (1985): 81-95.

- Becker, G.S., "A Theory of the Allocation of Time", Econ. J. (Sept. 1965): 493-517.

- Box, G.E.P. and Cox, D.R., "An Analysis of Transformations", Journal of the Royal Statistical Society, Series B (1962): 211-243.

- Brown, J.N. and Rosen, H.S., "On the Estimation of Structural Hedonic Price Models", Princeton University, mimeo, 1981.

- D' Agostino, R.B. and Pearson, E.S., "Tests for Departures from Normality. Empirical Results for the Distribution of b_2 and $\sqrt{b_1}$", Biometrika 60 (1973): 613-22.

- Epple, D., "Hedonic-Prices and Implicit Markets: Estimating Demand and Supply Functions for Differentiated Products", J.P.E. 95 (1987): 59-80.

- .... , "Closed Form Solutions to a Class of Hedonic Equilibrium Models", Carnegie-Mellon University, mimeo, 1984.

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

- Harrison, D. and Rubinfeld, D., "Hedonic Housing Prices and the Demand for Amenities", Journal of Environmental Econonomics and Manangement 5 (1978): 81-102.

- Houthakker, H.S., "Compensated Changes in Quantities and Qualities Consumed", Review of Economic Studies 19, no. 3 (1952): 155-64.

- Lancaster, K.J., "A New Approach to Consumer Theory", J.P.E. 74 (1966): 132-56.

- Muth, R.F., "Household Production and Consumer Demand Functions", Econometrica 34 (1966): 699-708.

- Palmquist, R.B., "Estimating Demand for the Characteristics of Housing", Review of Economics and Statistics 66 (1984): 694-404.

- Ramsey, J.B., "Tests for Specification Errors in Classical Linear Least Squares Regression Analysis", Journal of the Royal Statistical Society, Series B, vol. 31 (1969): 350-371.

- Ridker, R.G. and Henning, J.A., "The Determinants of Residential Property Values with Special Preferences to Air Pollution", Review of Economics and Statistics (1967).

- Rosen, S., "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition", J.P.E. 82 (1974): 34-55.

- Srivastava, M.S. and Carter, E.M., An Introduction to Applied Multivariate Statistics, Elsevier Science Publishing Co. Inc. (1983): 64-75.

- Tinbergen, J., "On the Theory of Income Distribution", in Selected Papers of Jan Tinbergen (Klassen, Koyck, and Witteveen, eds.), North-Holland (1959): 243-63.

- Witte, A.D., Sumka, H.J., and Erekson, O.H., "An Estimate of a Structural Hedonic Price Model of the Housing Market. An Application of Rosen's Theory of Implicit Markets", Econometrica 47 (1979): 1151-73.