Volume Title: The Measurement of Labor Cost

Volume Author/Editor: Jack E. Triplett, ed.

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-81256-1

Volume URL: http://www.nber.org/books/trip83-1

Publication Date: 1983

Chapter Title: Structural Estimation in Implicit Markets

Chapter Author: James N. Brown

Chapter URL: http://www.nber.org/chapters/c7375

Chapter pages in book: (p. 123 - 152)

3 Structural Estimation in Implicit Markets

James N. Brown

3.1 Introduction

At least since the time of Adam Smith, economists have viewed the employment relation as a transaction in several dimensions, with employers and employees embodying multiple characteristics of interest to each other, and with the allocation of workers and wages across jobs the result of implicit markets for those characteristics.[1] Only recently, however, have economists begun to estimate the structural parameters of these implicit markets for characteristics. Although the labor economics literature contains a long line of empirical work relating differences in wages to differences in worker and job attributes, as yet there have been few attempts to go beyond these "hedonic" descriptions of labor market outcomes and estimate the underlying structural demand and supply functions for characteristics that generate these outcomes.[2]

To some extent, this scarcity of structural analyses may be attributable to lags in the development of the appropriate theory and methodology.[3] Such lags, however, cannot completely explain this scarcity, for several studies that are analogous in nature have now appeared in other fields, particularly in the field of urban economics.[4] It is more likely that the relative scarcity of structural hedonic studies of the labor market stems from the generally inconclusive results obtained by researchers who have estimated compensating wage differentials for various job or worker characteristics. Although these researchers have repeatedly found evidence consistent with the presence of compensating wage differentials for

James N. Brown is assistant professor of economics, Princeton University.

The author is grateful to Sherwin Rosen and Harvey S. Rosen for helpful discussions on the subject matter of this paper. Any remaining errors are entirely his own. He also wishes to thank the Hoover Institution for financial support during the completion of this paper.

jobs requiring additional schooling or postschool training, attempts to estimate compensating differentials related to other characteristics of the job-worker match have generated less clear-cut results.[5] In contrast, researchers in the field of urban economics, for example, have consistently found evidence of negative housing price differentials associated with air pollution, and estimates of these differentials have served as a basis for several "structural" analyses of the demand for clean air.[6]

The estimation of compensating differentials that have appropriate signs is clearly a convenient starting point, if not a necessary condition, for the estimation of a market structure that might have generated those differences. Given the weak and varied nature of the wage differentials estimated so far, it is therefore not surprising that so few structural hedonic analyses of labor market data have been carried out. Nevertheless, with future improvements in the accuracy and completeness with which total compensation and job and worker attributes are measured, one might reasonably hope that more "believable" differentials will yet be found. Consequently, it seems reasonable to expect more structural hedonic analyses of labor market data to appear in the future. This expectation seems especially well justified, moreover, when one considers the many policy issues that require information about market structure for their resolution.[7]

Given the very likely appearance of more structural hedonic analyses of labor market data in the future, the very limited appearance of such analyses in the past, the growing experience with analogous studies in other fields, and the importance of correct methodology in such applications, some assessment of the experience to date with structural estimation in hedonic price models appears worthwhile. This paper is intended to contribute to that assessment.

The general focus of this paper centers on the conditions under which one can estimate the structural equations that generate an observed hedonic price locus, as well as the methods one might use to do so. The more specific focus of this paper centers on the two-stage procedure for estimating structural equations in implicit markets that was first suggested by Rosen (1974).[8] The paper begins with a brief summary of this empirical procedure and notes that, although the procedure has now been applied by several researchers, there appears as yet to be only limited recognition of the restricted set of conditions under which this method actually will yield estimates of structural parameters.

In developing this point, the paper discusses three related subjects that seem to have received insufficient explicit attention in the past. The first of these subjects concerns the use of "constructed" marginal prices in the estimation of structural equations for markets in which no direct observations on marginal prices are available. Contrary to suggestions originally made by Rosen (1974), and also by Freeman (1974), it is argued here that the use of such constructed marginal prices may have fundamental effects on the identification of structural equations and on the statistical methods required for consistent estimation of structural parameters in implicit markets.

The second subject addressed in this paper concerns the behavioral endogeneity of marginal attribute prices at the level of the individual market participant and the special data requirements implied by this endogeneity. Contrary to recent assertions by some authors, it is argued here that structural parameters can be estimated using data from a single implicit market. However, it is also argued that, holding constant the number of observations, data from several markets will generally be preferable to data from a single implicit market.

The third subject addressed in this paper concerns the potential problems and consequences of specification error that are peculiar to structural estimation in implicit markets. The general conclusion of this section and, indeed, of the paper as a whole is that, while the two-stage procedure suggested by Rosen may provide consistent estimates of structural parameters in implicit markets, estimates based on this procedure should be viewed with particular caution.

3.2 The Two-Stage Procedure for Structural Estimation in Implicit Markets

Perhaps the best starting point for a discussion of structural estimation in hedonic price models is Rosen's (1974) article. Although not the earliest discussion of the structural determinants of observed hedonic price loci, this article probably has been the most influential, and it provides a useful context for the discussion to follow.[9]

In his 1974 article, Rosen considered the relation between the "hedonic" price equations that many researchers had estimated for various commodities (see, e.g., Griliches 1971) and the structural demand and supply functions for "characteristics" that in principle had generated those hedonic price loci. The fundamental question addressed by Rosen was the following: Given that one observes an empirical relation between the price of some product, P, and the vector of characteristics embodied by that product, Z, what structural interpretation can one attach to this relation? In particular, how is such a relation generated by and related to the underlying distributions of tastes and technologies among market participants, and can the parameters that characterize those tastes and technologies and their distributions be derived from knowledge of the $P(Z)$ locus itself?

In answering this question, Rosen emphasized two basic points: first, any observed $P(Z)$ locus, being a joint envelope of (compensated) marginal bid and offer functions for buyers and sellers, will not generally convey any direct structural information about the families of bid and offer functions from which it is generated; and second, as a general matter, in the absence of extreme simplifying assumptions regarding tastes, technologies, and the distributions of tastes and technologies, no simple analytic relation exists between the functional form and parameters of the $P(Z)$ relation and the functional forms, parameters, and distributions of consumers' tastes and producers' technologies—thus precluding any analytically based inference about structural equations and parameters simply from observations on a $P(Z)$ locus alone.[10]

For those interested in recovering the structural compensated demand and supply parameters underlying observed hedonic price loci, these two results offered little encouragement. However, as a by-product of his analysis, Rosen was also led to suggest a two-stage empirical procedure for estimating the structural parameters underlying observed hedonic price loci that did not require the derivation of an exact analytical relation between the structural parameters of interest and the observed market locus parameters.

Following Rosen's presentation of this procedure, assume that consumers' marginal willingness to pay for characteristic Z_i is some function $F_i(\cdot)$ of a vector of characteristics, Z, as well as a vector of exogenous shift variables, Y_1. Similarly, assume that the marginal supply price of Z_i is some function $G_i(Z_i, Y_2)$, where Y_2 denotes a vector of exogenous variables shifting supply. Letting $p_i(Z)$ denote the implicit marginal price for attribute Z_i, the tangency of compensated bid and offer functions at each level of characteristic Z_i implies the following model for the data (ignoring random terms):

$$p_i(Z) = F_i(Z, Y_1) \quad \text{(demand)}, \tag{1}$$

$$p_i(Z) = G_i(Z, Y_2) \quad \text{(supply)}, \tag{2}$$

for which Rosen (1974) suggested the following estimating procedure:

> First, estimate $P(Z)$ by the usual hedonic method, without regard to Y_1 and Y_2. That is, regress observed differentiated products' prices, P, on all their characteristics, Z, using the best-fitting functional form. This econometrically duplicates the information acquired by agents in the market, on the basis of which they make their decisions. Denote the resulting estimate of the function $P(Z)$ by $\hat{P}(Z)$. Next, compute a set of implicit marginal prices $\partial P(Z)/\partial Z_i = \hat{p}_i(Z)$ for each buyer and seller, evaluated at the amounts of characteristics (numerical values of Z) actually bought or sold, as the case may be. Finally, use estimated marginal prices $\hat{p}_i(Z)$ as endogenous variables in the second-stage simultaneous estimation of equations (1) and (2). Estimation of marginal prices plays the same role here as do direct observations on prices in standard theory and converts the second-stage estimation into a garden-variety identification problem. (p. 50)

This procedure has since been applied by a steadily growing number of researchers, but although several applications and discussions of the procedure have now appeared, there seems still to be only limited recognition of the conditions under which the method actually will yield estimates of structural parameters. The following discussion elaborates on these conditions and the problems that may arise when these conditions are not met.

3.3 The Role of Constructed Marginal Prices in Structural Estimation

In his original statement of the two-stage procedure, Rosen asserted that estimated marginal prices could play the same role in structural estimation that direct observations on marginal prices would play, if available. He went on to say that, as long as some sample variation in marginal attribute prices could be observed, the identifiability of equations (1) and (2) would be determined by the standard rank and order conditions applicable to any market for which direct observations on prices exist. Each of these statements, however, requires qualification. Without qualification, each statement could lead researchers applying the two-stage technique to misinterpret resulting estimates of structural parameters.

Perhaps the most important thing to notice about equations (1) and (2) is that they are only part of a larger system of equations that also includes the equation used to define marginal prices. Consequently, when determining whether the parameters of equations (1) and (2) are identified, the rank and order conditions that must be considered are those that pertain to the entire three-equation system, and not just those that would pertain to equations (1) and (2) taken in isolation, as would be appropriate if equations (1) and (2) described a series of equilibria in separate, explicit markets for which direct observations on prices were available. The implication of this fact is that structural parameters which might otherwise be identified may not be identified when constructed marginal prices are used in place of direct observations on marginal prices. This fact seems to have gone unnoticed both by Rosen in his original statement of the two-stage procedure and by some researchers who subsequently have applied the technique. Neglect of this fact can lead to potentially serious misinterpretation of empirical estimates and in some cases appears to have done so.[11]

To illustrate the potential for such misinterpretation with an extreme case, suppose that the estimated first-stage equilibrium price locus for some implicit market is given by

(3) $$\hat{P}(Z) = \hat{g}_0 Z_i + \tfrac{1}{2}\hat{g}_1 Z_i^2 ,$$

so that the equilibrium marginal price function for Z in that market is estimated by

$$\hat{p}_i(Z) = \hat{g}_0 + \hat{g}_1 Z_i. \tag{4}$$

Suppose further that the structural demand and supply equations to be estimated are given by

$$p_i(Z) = a_0 + a_1 Z_i + a_2 Y_1 + u^d \qquad \text{(demand)}, \tag{5}$$

$$p_i(Z) = b_0 + b_1 Z_i + b_2 Y_2 + u^s \qquad \text{(supply)}, \tag{6}$$

where u^d and u^s denote random components of demand and supply, respectively.[12]

Looking only at equations (5) and (6) and interpreting them as if they described a series of equilibria in separate, explicit markets for which direct observations on prices were available, the parameters of these supply and demand functions would appear to be identified. Unfortunately, however, when one recognizes the presence of equation (4) as well in the structural model of this market, it becomes clear that the parameters of equations (5) and (6) are not identified. Because the variable $p_i(Z)$ must be replaced by $\hat{p}_i(Z)$ in the estimation of equations (5) and (6), and because $\hat{p}_i(Z)$ is an exact linear function of Z_i, observations on these marginal "prices" will not really provide any extra information beyond that already contained in observed sample values of Z_i. Indeed, it is easily verified that, as a result of this additional, mechanical dependence between marginal prices and observed values of Z_i, estimation of equations (5) and (6) using $\hat{g}_0 + \hat{g}_1 Z_i$ in place of $p_i(Z)$ will result in estimates of a_0 and b_0 that are both equal to $\hat{g}_0$, estimates of a_1 and b_1 that are both equal to $\hat{g}_1$, estimates of a_2 and b_2 that are both equal to zero, and values of R^2 equal to unity for either structural equation.[13]

More generally, in the presence of more than one characteristic, similar results emerge. Again taking an extreme example, if the estimated first-stage market locus were given by

$$\hat{P}(Z) = \hat{g}_1 Z_1 + \tfrac{1}{2}\hat{g}_{11} Z_1^2 + \hat{g}_2 Z_2 + \tfrac{1}{2}\hat{g}_{22} Z_2^2 + \hat{g}_{12} Z_1 Z_2, \tag{7}$$

so that the implicit marginal price for characteristic Z_i were given by

$$\hat{p}_i(Z) = \hat{g}_i + \hat{g}_{ii} Z_i + \hat{g}_{12} Z_j, \qquad (j \neq i), \tag{8}$$

then estimation of the following structural demand and supply equations:

$$p_i(Z) = a_{0i} + a_{1i} Z_i + a_{2i} Z_j + a_{3i} Y_1 + u_i^d, \qquad (i = 1, 2), \tag{9}$$

$$p_i(Z) = b_{0i} + b_{1i} Z_i + b_{2i} Z_j + b_{3i} Y_2 + u_i^s, \tag{10}$$

using $\hat{p}_i(Z)$ instead of direct observations on $p_i(Z)$ would lead to the following results:

(i) $\hat{a}_{01} = \hat{b}_{01} = \hat{g}_1$,

(ii) $\hat{a}_{02} = \hat{b}_{02} = \hat{g}_2$,

(iii) $\hat{a}_{11} = \hat{b}_{11} = \hat{g}_{11}$,

(iv) $\hat{a}_{22} = \hat{b}_{22} = \hat{g}_{22}$,

(v) $\hat{a}_{12} = \hat{a}_{21} = \hat{b}_{21} = \hat{b}_{12} = \hat{g}_{12}$,

(vi) $R^2 = 1$ for either structural equation.[14]

In this case, again, due to the presence of a third equation creating an exact link between marginal prices and observed values of Z, second-stage "structural" estimation would only reproduce first-stage estimated parameters.

It is worth emphasizing that results (i)–(v) would be obtained regardless of whether the researcher used ordinary least squares or some instrumental variables technique in attempting to estimate the structural supply and demand curves. Fundamentally, this problem arises from the exact, definitional dependence of the variable $\hat{p}_i$ on the set of regressors included in the structural equation to be estimated. As long as this exact dependence were present, the extreme results listed above would persist.[15]

The previous simple examples illustrate the potential for the use of constructed marginal prices to yield nonsense results in some cases. In extreme cases such as these, however, it is unlikely that the researcher would be unaware of the problem, given the extreme symptoms that are present. Nevertheless, although such extreme cases are unlikely to go unnoticed in practice, they are worth recognizing for two reasons.

First, these extreme examples emphasize the fact that structural estimation in implicit markets requires that marginal prices do more than simply vary—they must vary in a manner that is not collinear with the variables included on the right-hand side of the structural equations to be estimated. This point deserves emphasis, for it implies restrictions on the set of structural equations that can be estimated in conjunction with any given estimated marginal price function. Moreover, because there will generally be no guarantee that variables appearing in the estimated marginal price function for some implicit market should not also appear in the structural equations for that market, these extreme examples also illustrate the fact that it may often be impossible to estimate correctly specified structural equations using constructed marginal prices.[16]

Second, these extreme examples highlight the results toward which structural estimates may tend in less obvious cases, characterized by less than exact collinearity between constructed marginal prices and structural regressors. To explore these less obvious cases in more detail, suppose now that the marginal price function defining $\hat{p}_i(Z)$ includes some variable X not included in either of the structural equations to be estimated, so that the relevant three-equation system becomes

$$\hat{p}_i(Z) = a_0 + a_1 Z_i + a_2 Y_1 + u^d, \tag{11}$$

$$\hat{p}_i(Z) = b_0 + b_1 Z_i + b_2 Y_2 + u^s, \tag{12}$$

$$\hat{p}_i(Z) = \hat{g}_0 + \hat{g}_1 Z_i + \hat{g}_2 X. \tag{13}$$

In this case, the absence of exact collinearity between $\hat{p}_i(Z)$ and the set of structural regressors will allow the extreme results illustrated above to be avoided. Nevertheless, the additional relation between marginal prices and attribute values given by equation (13) must still be accounted for in any structural estimation of equations (11) and (12). Failure to do so could still result in the same sort of problems that arose in the more extreme case of exact collinearity between $\hat{p}_i(Z)$ and the set of structural regressors.

To illustrate this point most simply, suppose that the parameter b_1 in equation (12) is effectively infinite, so that equations (11) and (13) can be treated as a self-contained system of equations, and consider the results of estimating equation (11) by ordinary least squares. In this case, it is easily seen that ordinary least-squares estimation of equation (11) using values of $\hat{p}_i(Z)$ constructed from equation (13) will result in estimates of a_1 and a_2 with the following probability limits:

$$p\lim \hat{a} = a_1 + (\hat{g}_1 - a_1)\left[\frac{\sigma_d^2}{\hat{g}_2^2\sigma_X^2(1 - \rho_{XY_1}^2) + \sigma_d^2}\right],$$

$$p\lim \hat{a}_2 = \mathrm{a}_2 + (\hat{g}_2\beta_{XY_1} - a_2)\left[\frac{\sigma_d^2}{\hat{g}_2^2\sigma_X^2(1 - \rho_{XY_1}^2) + \sigma_d^2}\right],$$

where ρ_{XY_1} denotes the population correlation coefficient between X and Y_1, β_{XY_1} denotes the population regression coefficient for X as a function of Y_1, and σ_X^2 and σ_d^2 denote the population variances of X and u^d, respectively.[17] As these expressions show, even in this simple case for which ordinary least-squares estimation of equation (11) would normally be appropriate, the manner in which marginal price observations are constructed will cause ordinary least-squares estimates of a_1 and a_2 to be biased toward $\hat{g}_1$ and $\hat{g}_2\beta_{XY_1}$, respectively. This bias will be more extreme as the ratio $\sigma_X^2(1 - \rho_{XY_1}^2)/\sigma_d^2$ diminishes, with the extreme results initially discussed applying when that ratio equals zero (i.e., when marginal attribute prices embody no variation that is uncorrelated with the set of structural regressors in the equation estimated). Analogous results apply for ordinary least-squares estimation of b_1 and b_2.

As should be obvious, the existence of a definitional relation linking $\hat{p}_i(Z)$ and Z_i contaminates ordinary least-squares efforts to estimate behavioral relations between $p_i(Z)$ and Z_i. Although this point seems obvious, it seems to have gone unnoticed in several discussions of the two-stage procedure and in some applications of that procedure as well. Freeman (1979), for example, has offered the following elaboration on the two-stage procedure as outlined above:

> There are three possibilities. First, if the supply of (commodities) with given bundles of characteristics is perfectly elastic at the observed prices, then the implicit price function of a characteristic can be taken as exogenous to individuals. A regression of observed levels of the characteristic against the observed implicit prices . . . incomes, and other socioeconomic indicators of individuals should identify the demand function. . . .
>
> Second, if the available quantity of each model is fixed, individuals can be viewed as bidding for fixed quantities of models with given bundles of characteristics. A regression of each individual's price against the quantity of the characteristic actually taken, incomes, and other variables should identify an inverse demand function. . . .
>
> Finally, if both the quantities demanded and quantities supplied of characteristics are functions of prices, a simultaneous equation approach can be used. (pp. 196–97)

Following these suggestions in their empirical study of the demand for clean air, Harrison and Rubinfeld (1978) assumed a completely inelastic supply curve for clean air at various residential sites and applied ordinary least squares in estimating the following inverse demand functions for reductions in air pollution as measured by nitrogen oxide content (table 3.1). Harrison and Rubinfeld defined log (W), the "marginal willingness to pay," as a constant plus the sum of log(NOX) and the logarithm of median housing values. However, if housing values are roughly proportional to income in Harrison and Rubinfeld's sample, as may be suggested by the simple correlation of .82 between median housing values and mean income in their data, the variable log(INC) in Harrison and Rubinfeld's demand equations may simply act as a proxy for the logarithm of median housing values in the definition of log(W). If so, then Harrison and Rubinfeld may simply have reproduced their definition of log(W). The suspicious pattern of Harrison and Rubinfeld's coefficients suggests this possibility.[18]

Given the obvious problems that result from ordinary least-squares estimation of structural equations in implicit markets, regardless of the true underlying market structure, consider now the use of instrumental variables in the estimation of structural supply and demand curves, assuming as before that estimated marginal prices contain some variation that is linearly independent of the regressors included in the structural equations to be estimated (as in equations [11]–[13]). It is easily determined that, due to the presence of X in the marginal price function given

Table 3.1 Partial Listing of Harrison and Rubinfeld's Estimated Inverse Demand Parameters

	Independent Variables[b]					
Dependent Variable[a]	Constant	log (NOX)	log (INC)	log (PDU)	Y_1log (NOX)	Y_2log (NOX)
log(W)	1.08	.87	1.00	—	—	—
log(W)	1.05	.78	1.01	−.24	—	—
log(W)	2.20	.97	.80	—	−.03	−.07

Source: Harrison and Rubinfeld (1978), p. 89. Observation units were census tracts. No standard errors were presented for these coefficient estimates, but all coefficients were statistically nonzero at a .01 level of significance.

[a]W = marginal willingness to pay, measured in dollars and calculated as a constant plus the sum of the logarithms of nitrogen oxide concentration and median value of owner-occupied homes for the corresponding census tract.

[b]NOX = nitrogen oxide concentration in pphm; INC = household income in hundreds of dollars; PDU = persons per dwelling unit; $Y_1 = 1$ when $95 \leq \text{INC} < 130$, 0 otherwise; $Y_2 =$ 1 when INC ≥ 130, 0 otherwise.

by equation (13), equations (11) and (12) are identified.[19] In this case, therefore, application of some instrumental variables procedure should generate consistent estimates of structural parameters.

To demonstrate this point, consider the two-stage least-squares estimators for the parameters a_1 and a_2 from the structural inverse demand function (11). These estimators can be viewed as deriving from a regression of constructed marginal prices on Y_1 and on fitted values of Z_i taken from an auxiliary regression of Z_i on Y_1 and Y_2, and are given by

(14) $$\hat{a}_1 = \frac{\operatorname{cov}(\hat{p}_i, \hat{Z}_i \mid Y_1)}{\operatorname{var}(\hat{Z}_i \mid Y_1)},$$

(15) $$\hat{a}_2 = \frac{\operatorname{cov}(\hat{p}_i, Y_1 \mid \hat{Z}_i)}{\operatorname{var}(Y_1 \mid \hat{Z}_i)},$$

where $\operatorname{cov}(\hat{p}_i, \hat{Z}_i \mid Y_1)$ denotes the sample partial covariance of $\hat{p}_i$ with fitted values of Z_i, holding Y_1 constant; $\operatorname{var}(\hat{Z}_i \mid Y_1)$ denotes the sample partial variance of fitted values of Z_i, holding Y_1 constant; and $\operatorname{cov}(\hat{p}_i, Y_1 \mid \hat{Z}_i)$ and $\operatorname{var}(Y_1 \mid \hat{Z}_i)$ are defined analogously. Using the definition of $\hat{p}_i$ from the estimated marginal price function (13), and expressing $\hat{Z}_i$ as $k_0 + k_1Y_1 + k_2Y_2$, where $k_1 = \operatorname{cov}(Z_i, Y_1 \mid Y_2)/\operatorname{var}(Y_1 \mid Y_2)$ and $k_2 = \operatorname{cov}(Z_i, Y_2 \mid Y_1)/\operatorname{var}(Y_2 \mid Y_1)$, these estimators can be rewritten as

(16) $$\hat{a}_1 = \hat{g}_1 + \hat{g}_2\left[\frac{\operatorname{cov}(X, Y_2 \mid Y_1)}{\operatorname{cov}(Z_i, Y_2 \mid Y_1)}\right] = \hat{g}_1 + \hat{g}_2\beta_{XZ_i} \mid Y_1,$$

(17) $$\hat{a}_2 = \hat{g}_2\left[\frac{\operatorname{cov}(X, Y_1 \mid \hat{Z}_i)}{\operatorname{var}(Y_1 \mid \hat{Z}_i)}\right] = \hat{g}_2\beta_{XY_1} \mid \hat{Z}_i,$$

where $\beta_{XZ_i|Y_1}$ denotes the estimated partial regression coefficient for X with respect to Z_i, holding Y_1 constant and using Y_2 as an instrument for Z_i; and where $\beta_{XY_1|Z_i}$ denotes the estimated partial regression coefficient for X with respect to Y_1, holding $\hat{Z}_i$ constant.

Given the presence of $\hat{g}_1$ and $\hat{g}_2$ in these expressions, one might expect that instrumental variables estimates of a_1 and a_2 would be biased by the use of constructed marginal prices, as was the case in the extreme examples initially discussed. This expectation would not be correct, however.

In interpreting the above estimators for a_1 and a_2, it is helpful to notice that for the system of equations given by

$$p_i = a_0 + a_1 Z_i + a_2 Y_1, \tag{18}$$

$$p_i = g_0 + g_1 Z_i + g_2 X, \tag{19}$$

variations in Z_i, Y_1, and X must be related according to the following equation:

$$(a_1 - g_1)\Delta Z_i + a_2 \Delta Y_1 - g_2 \Delta X = 0. \tag{20}$$

Thus, given any two values of the vector (Z_i, X, Y_1) that satisfied equations (18) and (19) and for which Y_1 remained constant, a_1 could be derived from the relation

$$a_1 = g_1 + g_2 \left.\frac{\Delta X}{\Delta Z_i}\right|_{\Delta Y_1 = 0}. \tag{21}$$

Similarly, given any two values of the vector (Z_i, X, Y_1) that satisfied equations (18) and (19) and for which Z_i remained constant, a_2 could be derived from the relation

$$a_2 = g_2 \left.\frac{\Delta X}{\Delta Y_1}\right|_{\Delta Z_i = 0}. \tag{22}$$

Holding Y_1 constant, equations (18) and (19) imply that marginal prices will vary (as measured by $g_2 \Delta X$) as Z_i varies only to the extent that a_1 differs from g_1. Thus, a_1 can be measured as differing from g_1 by the extent that marginal prices vary as Z_i varies, holding Y_1 constant. Similarly, holding Z_i constant, equations (18) and (19) imply that marginal prices will vary (as measured by $g_2 \Delta X$) as Y_1 varies only to the extent that a_2 differs from zero. Thus, a_2 can be measured as differing from zero by the extent that marginal prices vary as Y_1 varies, holding Z_i constant.

This reasoning clearly applies regardless of whether one interprets equations (18) and (19) as deterministic or as stochastic. In the latter case, this reasoning provides the conceptual basis for the estimators given by equations (16) and (17). Although these estimators will be influenced by the definitional relation linking marginal prices and attribute levels, this influence has a legitimate theoretical interpretation. As long as $\hat{g}_1$ and

$\hat{g}_2$ are consistent estimates of the true equilibrium relation between marginal prices and values of Z_i and X, consistent estimation of a_1 and a_2 requires only that $\beta_{XZ_i|Y_1}$ and $\beta_{XY_1|Z_i}$ be estimated consistently. Given the structure of equations (11)–(13), moreover, it is clear that the use of Y_2 as an instrument for Z_i in equation (11) would implicitly provide the consistent estimates of $\beta_{XZ_i|Y_1}$ and $\beta_{XY_1|Z_i}$ required. Thus, conditional on the presence in the equilibrium marginal price function of some variable X that is not perfectly collinear with the set of structural regressors, and conditional on consistent estimates of the equilibrium marginal price function, the application of instrumental variables procedures can generate consistent estimates of structural parameters in implicit markets.[20]

To summarize the results of this section, consistent estimation of structural parameters in implicit markets is possible, and constructed marginal prices can play the same role in structural estimation that direct observations on marginal prices would play if they were available, but only if three conditions are met (in addition to the usual requirement that structural equations be correctly specified): First, constructed marginal prices must embody some variation that is orthogonal to the set of structural regressors in the equation estimated. Second, constructed marginal prices must be consistent estimates of true marginal prices. Third, constructed marginal attribute prices and observed attribute levels must be treated econometrically as jointly endogenous variables, regardless of the true underlying market structure.[21] The following sections elaborate on the first two of these conditions.

3.4 The Role of Cross-Market Data in Structural Estimation

The preceding section emphasized the requirement for structural estimation that constructed marginal prices embody some variation orthogonal to the set of structural regressors. Little was said, however, about the possible sources of such variation. This section addresses that subject, focusing in particular on the assertion made by some researchers (see, e.g., G. Brown and Mendelsohn 1980) that this variation must reflect differences across separate implicit markets in the marginal price functions facing market participants. It is argued here that structural identification in implicit markets does not necessarily require the presence of cross-market variation in marginal prices, although such cross-market variation will generally be preferable to an equivalent amount of within-market price variation, given the limited ability to test for specification error with data taken from a single implicit market.

To provide a context for the assertion that cross-market price variation is necessary for structural identification in implicit markets, consider the data requirements for the estimation of a demand function in a standard market model. Because only one price can be observed within a single market, it is clear that data from more than one market will be necessary to estimate any response of quantity demanded to changes in prices. Given such multimarket data, the ideal experiment for identifying the effect of price on quantity demanded might then involve a comparison of quantities demanded across several markets having identical demand curves (identical levels of income, for example) but different supply curves and, consequently, different prices. In the absence of such an ideal data set, essentially the same sort of comparison could be made statistically by comparing the covariation of quantities and prices that is orthogonal to income, for example, with the variation in prices that is orthogonal to income.

Now, consider instead a single implicit market. Price variation can be observed within such a market, so it may appear that the same statistical method can be applied within a single implicit market as is applied in the case of several separate explicit markets. However, the price variation observed within a single implicit market, unlike the price variation observed across separate explicit markets, cannot possibly be exogenous to shifts in the demand curves being estimated, since marginal prices within a single implicit market can vary across consumers only if demand curves vary across consumers. Thus, although one might observe variation in marginal prices and quantities demanded within a single implicit market, such variation does not clearly correspond to the basic conceptual experiment underlying the estimation of demand curves in standard markets. It is therefore not clear that making use of this variation just as one would for a set of ordinary markets will yield coefficients with structural content.

This behavioral endogeneity of marginal prices at the level of the individual market participant has led some researchers to assert that data from a single implicit market cannot be sufficient to estimate structural demand and supply parameters. G. Brown and Mendelsohn (1980), for example, state that

> data from a single market, producing necessarily one set of prices, are inadequate for estimating the demand functions for characteristics. Each consumer faces the same relative prices of characteristics in one market so no demand function can be estimated. . . . To estimate demand, variation in the price at each level is necessary. . . . The way to obtain suitable price variations is clear, if tedious. Each location is regarded as a separate market. Price variations across markets form the essential ingredients for estimating demand functions for characteristics, along with associated quantities of characteristics and other socioeconomic demand determinants. (pp. 3–4)

The analysis of the previous section, however, suggests that this assertion may be incorrect, since there appeared in that analysis no obvious requirement that X embody such cross-market variation.

To investigate this issue, consider a single implicit market for which the underlying structural inverse supply and demand functions are given by equations (11) and (12). Suppose further that for this market the equilibrium sorting of buyers and sellers leads to an equilibrium marginal price function that can be written as

$$p_i(Z) = g_0 + g_1 Z_i + u, \tag{23}$$

where u is a zero-mean disturbance term uncorrelated with all variables except Y_1 and Y_2.[22]

In this case, it is easily seen that structural estimation using the two-stage procedure would not be possible with data from only this market. As discussed earlier, the estimated equilibrium marginal price function must include some variable orthogonal to Z_i, Y_1, and Y_2 in order for structural estimation to be feasible, but given the present assumptions regarding u, no such function could be estimated. Thus, structural estimation would not be possible with data taken from this one market alone.

In contrast, suppose now that data are available from several such markets. In this case structural estimation may be possible if g_0 and g_1 vary across markets.[23] In effect, the availability of cross-market data allows market-specific dummy variables to play the role of X in an augmented equilibrium marginal price function, and these market-specific dummy variables may have nonzero coefficients in that function, even though no variable other than Z_i has a nonzero coefficient within any single market. In cases such as this, multimarket data will be necessary and may be sufficient for structural estimation.

Although necessary in some cases, however, cross-market data will not be necessary in all cases. To illustrate, consider the estimation of equation (11) using data from a single implicit market in which X denotes the square of Z_i. From a conceptual or sample design viewpoint, identification by this nonlinearity can be viewed as consistent with a hypothetical comparison of observed *differences* in quantities demanded and observed *differences* in marginal prices across *pairs* of consumers with identical differences in quantities demanded at given marginal prices (i.e., identical differences in Y_1). In order for this conceptual experiment to be valid, marginal price differences must vary across pairs of consumers, and consumers must respond identically to differences in marginal prices, even though they implicitly choose different levels of marginal prices. But these requirements amount to nothing more than the inclusion of Z_i^2 (or some higher order term) in the equilibrium marginal price function and exclusion of Z_i^2 (or that higher order term) from the structural inverse demand function. Thus, as long as one can assume an equilibrium marginal price function that is quadratic in Z_i, one can in principle estimate an inverse demand function that is linear in Z_i using data from a single implicit market. More generally, as long as one can assume an equilibrium marginal price function that is of order m in Z_i, one can in principle estimate an inverse demand function that is of order $m-1$ in Z_i using data from a single implicit market.[24]

Nevertheless, although one can in principle estimate an inverse demand function of order $m-1$ in Z_i by first estimating an equilibrium marginal price function of order m in Z_i, there is no guarantee that the data taken from any single market actually will support such estimation, either in the sense of generating a sufficiently nonzero coefficient on Z_i^m in the estimated marginal price function, or in the sense of justifying the restriction that Z_i^m be excluded from the inverse demand function. Furthermore, the appropriateness of this latter restriction can never be tested using data from a single market alone, since the inclusion of Z_i^m on both sides of the inverse demand function would then lead to the extreme results discussed earlier.

It is in this regard that cross-market data will generally be preferable to single-market data. By allowing a broader set of structural equations to be estimated than would an equivalent amount of within-market data, cross-market data provide a greater opportunity to test the restrictions on which structural estimation is based. However, although the opportunity for such testing is extended by the availability of cross-market data, and although cross-market data may allow the estimation of structural equations that could not be estimated with single-market data, cross-market data will not always be sufficient for structural estimation, nor will cross-market data allow statistical testing of this sufficiency.

To demonstrate that cross-market data may not be sufficient for structural estimation in implicit markets, one need only note in the context of equations (11), (12), and (23) that if a_0, b_0, a_1, and b_1 also vary across markets as g_0 and g_1 were assumed to vary, structural estimation again would be impossible, even with cross-market data.[25] Moreover, as in the case previously discussed, the researcher could never test the appropriateness of imposing constancy on these coefficients, since allowing them to vary in estimation would once again result in the extreme problems discussed initially. Thus, structural estimation, whether on the basis of single-market or cross-market data, ultimately must rest on a priori restrictions that may not be met by the data and that cannot all be tested. Given this fact, it is worthwhile to consider the potential problems and consequences of specification error that may affect structural estimation in implicit markets. The following section addresses this issue.

3.5 Specification Error in Implicit Markets

In contrast to the case of ordinary markets for which direct observations on prices are available, structural estimation in implicit markets requires not only that structural equations be correctly specified, but also that the first-stage equation used to construct marginal price "observations" itself be correctly specified. Because the estimated first-stage $P(Z)$ function fundamentally determines the "data" on which second-stage structural estimation is based, any error made in the estimation of that function will generally be translated into errors in the estimation of structural equations. This point is surely not surprising, but it is especially important to emphasize in the context of implicit markets, where theory provides little basis for the specification of either the first-stage market locus or the second-stage structural equations, and where the "constructed" nature of the dependent variable creates an inherent risk that second-stage structural estimation may only reproduce parameters of the estimated marginal price function.

To illustrate some of the problems of specification that are peculiar to structural estimation in implicit markets, consider the consequences that arise when some variable is incorrectly excluded from the estimated marginal price function for an implicit market. Suppose, for example, that the true equilibrium marginal price function for this market is given by

$$p_i(Z) = g_0 + g_1 Z_i + g_2 X + g_3 W, \tag{24}$$

but that the researcher instead constructs marginal prices using the relation

$$\hat{p}_i(Z) = \hat{g}_0 + \hat{g}_1 Z_i + \hat{g}_2 X, \tag{25}$$

with the $\hat{g}_i$ derived from a first-stage regression of P on Z_i, Z_i^2, and Z_iX. Suppose further that the true structural equations for this market are those given by equations (11) and (12), and that the researcher estimates correctly specified versions of these equations. Finally, suppose that the omitted variable W is orthogonal to all variables in the structural supply and demand functions, so that its omission from the marginal price function does not cause any direct bias in estimated structural parameters.

In this case, one might expect the omission of W from the marginal price function (or, more precisely, the omission of the product of Z_i and W from the first-stage estimated $P(Z)$ locus) not to induce bias in structural estimates, since this "measurement error" would be confined to the dependent variable alone and would not be directly correlated with the variables included in the structural equations estimated. Nevertheless, because the omission of Z_iW from the estimated first-stage $P(Z)$ locus will generally lead to inconsistent estimates of g_1 and g_2, and because errors in the estimation of g_1 and g_2 will lead to "measurement errors" in the estimation of p_i that are correlated with Z_i, structural parameter estimates will be made inconsistent by this omission, even though the structural equations themselves are correctly specified. In general, only if the product of Z_i and W were orthogonal to the variables included in the first-stage estimated $P(Z)$ locus, and W orthogonal to the instrumental variables used in estimating the structural demand and supply equations, would structural parameter estimates not be made inconsistent by such omission.[26]

Alternatively, suppose again that the variable W is incorrectly omitted from the estimated marginal price function, but now suppose also that W is incorrectly included in the structural inverse demand function. In this case again, it is obvious that, because estimated coefficients in the marginal price function will generally be made inconsistent by the exclusion of W from that function (or, more precisely, by the exclusion of Z_iW from the first-stage estimated $P(Z)$ locus), estimates of a_1 and a_2 also will be made inconsistent by this exclusion. Furthermore, it is a straightforward matter to see that the resulting estimated structural coefficient for W in this case will be biased toward the coefficient for W in the true marginal price function.[27] Thus, even though W does not appropriately belong in the structural inverse demand function, it may appear statistically significant in that function, and the researcher may be given no warning that the inclusion of W in the structural demand function is inappropriate, as would generally be provided by a low t-statistic if the marginal price function were correctly specified.

The potentially serious consequences of incorrectly excluding some variable from the estimated marginal price function for an implicit market may appear to warrant the inclusion of possibly extraneous variables in that function. The incorrect inclusion of such variables, however, may also have potentially serious consequences. To illustrate this fact, suppose now that W no longer belongs in the true equilibrium marginal price function for the implicit market discussed above, but that W is incorrectly included in the estimated version of the marginal price function for that market.

In this case, as before, even if estimated structural equations are correctly specified, specification error in the marginal price function can lead to inconsistent estimates of structural parameters by way of measurement error in the dependent variable that is correlated with the arguments of the structural equations estimated. Unlike the case where W is incorrectly excluded from the estimated marginal price function, however, incorrect inclusion of W in the estimated marginal price function will cause inconsistent estimates of correctly specified structural equations only if W is correlated with the arguments of those structural equations. Assuming that W truly is an extraneous variable, such inconsistency would therefore appear to be unlikely. Nevertheless, given the ad hoc manner in which the estimated $P(Z)$ locus is usually specified, the possibility of such bias should not be overlooked.[28]

Moreover, the potential consequences of incorrectly including W in the estimated marginal price function may become more serious when the estimated structural equations themselves are misspecified. Given that W has been incorrectly included in the estimated marginal price function, if W is also incorrectedly included in the estimated structural demand function, it can be seen that the estimated structural coefficient for W will be biased toward the coefficient for W in the estimated marginal price function, with exact quality holding when W is orthogonal to X, given $\hat{Z}_i$ and Y_1.[29] Thus, as before, the incorrect inclusion of an irrelevant variable in an estimated structural equation can result in statistically significant estimated structural coefficients for that variable, and the researcher may be given no warning that such inclusion is inappropriate, as would generally be provided by a low t-statistic if the marginal price function were correctly specified.

As a final example, suppose again that the irrelevant variable W is incorrectly included in the estimated marginal price function, and suppose now that the variable X is incorrectly included in the structural demand equation. In this case, the presence of W in the marginal price function and absence of W from the structural demand function will allow estimates of a_1 and a_2 to be calculated, but given that W is an irrelevant variable, it can easily be shown that the estimated structural coefficients for Z_i, Y_1, and X will be biased toward the coefficients for those variables in the marginal price function, with exact equality holding when W is orthogonal to the variables included in the structural demand and supply functions.[30] Once again, misspecification may result in estimated structural parameters that merely reflect estimated parameters of the marginal price function, and once again there may be no clear statistical evidence of such misspecification.

This last example is relevant not only to cases in which structural equations have been misspecified, but also to cases in which correctly specified structural equations are not identified but nonetheless estimated on the basis of an extraneous variable included in the estimated equilibrium marginal price function. As this example indicates, the presence of such bogus identification will generally result in estimated structural parameters that mimic previously estimated parameters of the marginal price function. In cases where estimated structural parameters and estimated parameters of the marginal price function appear to coincide, therefore, one might be tempted to infer that such bogus identification is present. Unfortunately, this inference would not be without risk, for it is always possible that the two sets of parameters could be similar for legitimate reasons. Nevertheless, given the ex-post, curve-fitting nature of the process by which first-stage specification generally occurs, an extra burden of proof might reasonably be expected to fall on the researcher, especially when structural and marginal price function parameters appear to coincide. In such cases particularly, one should be wary that irrelevant variables or inappropriate variables have been included in both the first and second stages of the estimation procedure.

Considering the potential for structural parameter estimates to mimic first-stage locus parameter estimates when both the marginal price function and structural demand or supply functions are misspecified, it is worth noting that, in many instances, the "structural" parameter estimates implied by such inadvertent reproduction of the equation used to construct marginal prices may be qualitatively similar to those implied by demand theory. For example, if one first estimates the market locus given by equation (7) and then uses equation (8) to construct marginal prices, inadvertent reproduction or near reproduction of equation (8) would lead to estimated demand curves that tended to display symmetry of cross-price effects and that also tended to display negative own-price effects for characteristics in which the estimated version of equation (7) was concave. This tendency suggests that one should interpret with caution studies that present negative estimated own-price effects and symmetry of estimated cross-price effects as evidence that structural demand curves really have been estimated.[31]

In this regard, consider the estimates reported by Witte, Sumka, and Erekson (1979, hereafter Witte et al.) in their application of the two-stage procedure to the housing market.[32] In their study, Witte et al. first estimated, for each of four cities, a quadratic market locus relating housing values to various characteristics, including dwelling quality, dwelling size, and lot size (see table 3.2).

Using these estimates to construct marginal characteristic prices, Witte et al. then estimated a set of linear (inverse) demand and supply functions, imposing constancy of structural coefficients across markets (see table 3.3).

Upon inspection, the following characteristics of Witte et al.'s estimates become apparent. First, there is a general similarity in magnitude between estimated own-price effects on demand and on supply. In only two of the nine cases shown in table 3.3 are the two estimated effects not similar in magnitude. Second, the estimates in table 3.3 display the symmetrical pattern implied by the equations used to construct estimated marginal prices. On the demand side this pattern might be explained by Slutsky symmetry, but on the supply side it seems unlikely that anything other than the method by which marginal prices were constructed can account for this pattern. Third, there is a general similarity in magnitude between the coefficients on squared values of characteristics from the first-stage equation and Witte et al.'s estimated own-price effects on supply and demand from the second-stage estimation. In particular, there is a tendency for δ_{11} to exceed δ_{22}, which exceeds δ_{33} (in absolute value), and there appears to be a corresponding decline in the absolute value of

Table 3.2 **Partial Listing of Witte, Sumka, and Erekson's Estimated Market Locus Parameters**

	Estimated Parameters					
City	δ_{11}	δ_{22}	δ_{33}	δ_{12}	δ_{13}	δ_{23}
Greenville	−7.40	−3.23	[a]	[a]	[a]	0.65
	(3.05)	(1.21)	—	—	—	(0.27)
Kinston	8.53	−0.78	−0.001	−2.00	0.75	−0.17
	(6.75)	(1.25)	(0.02)	(4.95)	(0.83)	(0.31)
Lexington	9.29	−0.40	−0.011	6.13	0.19	−0.05
	(2.77)	(1.41)	(0.01)	(2.29)	(0.24)	(0.10)
Statesville	[a]	−2.47	[a]	14.18	[a]	[a]
	—	(1.02)	—	(3.94)	—	—

Source: J. Brown and H. Rosen (1981), p. 10. Numbers in parentheses are standard errors. These estimates are based on an estimating equation of the form

$$R = a + \sum_{i=1}^{5} \delta_i Z_i + \sum_{i=1}^{5} \sum_{j=1}^{5} \delta_{ij} Z_i Z_j + \sum_{i=1}^{5} \gamma_i D_i + U,$$

where: R denotes annual contract rent; Z_1 denotes dwelling quality; Z_2 denotes dwelling size; Z_3 denotes lot size; Z_4 denotes neighborhood quality; Z_5 denotes accessibility; D_1 denotes a dummy variable = 1 if heat charges included in rent; D_2 denotes a dummy variable = 1 if furnishings included in rent.

[a]Witte et al., excluded these variables because they did not add significantly to the explanatory power of the regression (see Witte et al., 1979, p. 1151, note 13).

the coefficients in table 3.3 as one reads along the main diagonal from northwest to southeast. Similarly, Witte et al.'s estimated values of δ_{13} tend to be small in absolute value, as do the estimated coefficients for Z_3 in the demand and supply equations for Z_1, and for Z_1 in the demand and supply equations for Z_3. Finally, although not reproduced here, in only eight out of twenty-four cases were Witte et al.'s estimated coefficients on demand and supply shift variables statistically nonzero at less than a .10 level of significance.[33] Thus, although one cannot reject the hypothesis that these estimates accurately reflect structural parameters, the patterns they display suggest that these estimates may reflect the construction of marginal prices more than they reflect any true market structure.

The discussion in this section emphasizes the misinterpretation of structural estimates that may result from specification error in implicit markets. Like structural estimation in ordinary markets, structural estimation in implicit markets ultimately rests on a priori restrictions that may not be met by the data and that cannot all be tested. Nevertheless, certain types of misspecification in implicit markets will result in structural estimates that, through their similarity to estimated marginal price function parameters, offer at least circumstantial evidence that such misspecification is present. Given this fact, and given also the limited theoretical basis for identifying restrictions imposed in hedonic structural estimation, it seems especially important that structural studies of im-

Table 3.3 **Partial Listing of Witte, Sumka, and Erekson's Estimated Structural Parameters**

	Independent Variables[a]		
Dependent Variables	Z_1	Z_2	Z_3
Demand price for Z_1	−8.65	5.00	0.41
	(4.78)	(4.63)	(0.36)
Supply price for Z_1	11.08	7.83	−0.74
	(2.87)	(2.76)	(0.49)
Demand price for Z_2	8.12	−6.97	0.41
	(2.49)	(2.41)	(0.19)
Supply price for Z_2	6.41	−0.71	0.28
	(1.16)	(1.12)	(0.20)
Demand price for Z_3	−0.28	0.38	−0.03
	(0.19)	(0.19)	(0.01)
Supply price for Z_3	0.12	−0.02	0.01
	(0.09)	(0.09)	(0.02)

Source: J. Brown and H. Rosen (1981), p. 9. Numbers in parentheses are standard errors.
[a]Z_1 = dwelling quality; Z_2 = dwelling size; Z_3 = lot size.

plicit markets provide sufficient information to assess the likelihood of such misspecification. Of the structural hedonic studies that have been carried out, however, few have provided such information. Considering the questions that have been raised in this section, one would hope that future structural studies of implicit markets will not be similar in this regard.

3.6 Summary and Conclusion

Structural estimation in implicit markets differs from structural estimation in explicit markets in one fundamental respect: the absence of directly observed prices for the good implicitly traded and the consequent presence in implicit markets of a third equation linking prices and quantities, in addition to the usual demand and supply functions. Due to the required use of constructed marginal attribute prices in implicit markets, a complete description of the process by which "observed" data are generated in such markets must include this third equation. Failure to consider this third equation can lead the researcher to use inappropriate data or inappropriate statistical methods in the estimation of structural parameters.

The use of constructed marginal attribute prices in implicit markets imposes additional restrictions on the research methods required for structural estimation in implicit markets. Constructed marginal prices may play the same role in structural estimation that direct observations on marginal prices would play if they were available, but they will not necessarily play that role, and their ability to play that role is less general than many discussions and applications of the two-stage procedure might lead one to expect.

First, constructed marginal prices must embody some variation that is orthogonal to the set of regressors included in the structural equations estimated. This requirement applies to ordinary markets as well as implicit markets, but in ordinary markets with directly observed prices, the required variation can be purely random. In contrast, in implicit markets this variation must be generated by some observable variable not included in the set of structural regressors. Relative to the case of ordinary markets for which direct observations on prices are available, therefore, the requirement that constructed marginal prices not be perfectly collinear with the set of structural regressors limits the set of structural equations that can be estimated in conjunction with any given equilibrium marginal price function, and may require that estimated structural equations omit some variable that would not have to be omitted from those equations if marginal prices were directly observable. Consequently, because there will generally be no guarantee that all variables included in the equilibrium marginal price function for some implicit market should not also appear in the underlying structural demand and supply functions for that market, there will generally be no guarantee that structural estimation using constructed marginal prices will not suffer from potentially serious omitted variables bias that would not be present if marginal attribute prices were directly observable. Moreover, relative to the case of ordinary markets, the researcher may have little opportunity to test statistically for the structural significance of omitted variables, since the inclusion of these variables in the structural equations to be estimated could result in exact duplication of the estimated marginal price function or near duplication of that function, depending on the variables in question and the true underlying structure of the implicit market studied.

Second, constructed marginal prices must be treated as jointly endogenous with observed attribute levels in implicit markets, regardless of the true parameters of the structural equations estimated (except, of course, when one side of the market is characterized by complete homogeneity). In contrast to the case of ordinary markets, therefore, the use of constructed marginal prices prevents the researcher from exploiting, for example, the assumption of vertical or horizontal structural demand or supply curves in order to identify parameters of interest. Consequently, structural parameters that might be identified in the context of ordinary markets with directly observable prices might not be identified in the context of implicit markets with constructed marginal prices.

Third, marginal attribute prices must be constructed without error if potentially serious misinterpretation of estimates is to be avoided. Unlike the case of ordinary markets, measurement error in the dependent variable cannot generally be assumed to be uncorrelated with structural regressors. Consequently, such measurement error can generally be expected to lead to inconsistent parameter estimates. Incorrect exclusion or inclusion of variables from the estimated marginal price function may lead to economically reasonable and statistically significant structural coefficients for structurally irrelevant variables when structural equations have been misspecified. The use of constructed marginal prices therefore creates the inherent risk that structural estimation will be biased by the definitional relation linking marginal prices and observed attribute levels in a manner not statistically discernible to the researcher. Given this fact, structural estimates in implicit markets should be viewed with particular caution.

Notes

1. The standard reference in this area, of course, is Adam Smith's statement that "the whole of the advantages and disadvantages of the different employments of labor and stock must, in the same neighborhood, be either perfectly equal or continually tending toward equality" (Smith 1937, p. 99).

2. There are several studies of labor market data that interpret observed "hedonic" relationships as structural on the basis of an assumed homogeneity of preferences or technologies. There are far fewer studies (see, e.g., R. Smith 1974; Woodbury 1983; Atrostic 1982; and Sider 1981) that estimate structural equations in a manner that allows for heterogeneity on both sides of the market. It is this latter type of analysis to which the statement in the text refers.

3. A selective chronology of theoretical and methodological work relevant to the development of structural analyses in implicit markets would include the following: A. Smith (1937); Court (1941); Roy (1950); Houthakker (1952); Tiebout (1956); Tinbergen (1956); Griliches (1961); Alonso (1964); Becker (1965); Lancaster (1966); Muth (1966); Lewis (1969); Griliches (1971); S. Rosen (1974); Freeman (1974); Sattinger (1975); Lucas (1975); Epple (1980); J. Brown and H. Rosen (1981). Although several theoretical and empirical papers on the subject of implicit markets were written prior to 1974, it was not until S. Rosen's (1974) exposition that an empirical procedure for estimating structural demand and supply functions in implicit markets was clearly spelled out.

4. In the urban economics literature, the technique has been applied by McDougall (1976) in estimating the demand for local school and police services; by Harrison and Rubinfeld (1978) and Nelson (1978) in estimating demand and supply functions for clean air; and by Witte, Sumka, and Erekson (1979), Linneman (1980, 1981), and Blomquist and Worley (1981) in estimating demand and supply functions for various housing and neighborhood attributes.

5. On this subject, see R. Smith (1979), C. Brown (1980), and the papers cited therein.

6. See, for example, Harrison and Rubinfeld (1978) and Nelson (1978).

7. At the macrolevel, any evaluation of the potential effects of policies applied to entire markets would generally require knowledge of structural parameters. Knowledge only of market equilibrium, compensating wage loci would not be sufficient, since any policy applied to entire markets would generally alter those loci in a manner that could be predicted only with knowledge of those markets' underlying structural demand and supply parameters, as well as the distributions of tastes and technologies within those markets.

At the microlevel, any assessment of the potential success of efforts to alter the specific bundles of job characteristics jointly chosen by workers and firms, whether by the monetary inducements of taxes and subsidies or by imposed restrictions on quantities, would also require knowledge of the structural bid and offer functions for the relevant characteristics. In general, only with such knowledge could one predict the likely substitution among various job characteristics induced by those policies.

8. This procedure was discussed and applied also by A. M. Freeman in papers dating approximately from the time of Rosen's original contribution (see, e.g., Freeman 1974, 1979).

9. For an earlier paper on the subject of equilibrium in implicit markets, see Lewis (1969). For an analysis similar to and contemporary with Rosen's, see Freeman (1974).

10. The obvious exception to this statement, as noted by Rosen and Freeman, occurs when one side of the market is characterized by complete homogeneity, so that the observed $P(Z)$ locus is equivalent to the compensated marginal bid or offer function for that side of the market.

11. With the exception of the recent papers by Epple (1980) and J. Brown and H. Rosen (1981), I have found no explicit discussion of this fact in the implicit markets literature. Moreover, at least two empirical applications of the two-stage procedure (Harrison and Rubinfeld 1978; Witte, Sumka, and Erekson 1979) appear to suffer from misinterpretation due to neglect of this fact. Several other studies may suffer from such misinterpretation, but the authors of those studies present insufficient information for the reader to determine whether this is so.

12. This example, along with the accompanying discussion, is taken from J. Brown and H. Rosen (1981). Harvey S. Rosen deserves equal credit for the points made here.

It should be noted that the equilibrium price locus and marginal price function for an hedonic market will not generally be independent of the structural demand and supply functions underlying that market. Indeed, the distributions of shift variables and random elements in the structural functions will, by way of those functions and the condition of market equilibrium, fully determine the equilibrium price locus and marginal price function. Thus, one cannot arbitrarily choose any set of structural functions that might correspond to any given equilibrium price locus and marginal price function (and vice versa). Strictly speaking, therefore, there is no guarantee that structural functions such as (5) and (6) would appropriately correspond to equilibrium functions such as (3) and (4). Nevertheless, for present purposes, this point need not be developed. The present discussion seeks only to explore the consequences of estimating, for whatever reason, equations (5) and (6) using marginal prices constructed according to equation (4). No claim is made here that such estimation would be generally appropriate.

13. These results are easily demonstrated by considering first the ordinary least-squares estimator for the column vector $(a_0 a_1 a_2)'$ from the regression $p_i = Xa + u$, where X denotes the row vector $(1, Z_i Y_1)$. This estimator is given by the familiar expression $\hat{a} = (X'X)^{-1} X'p_i$. Given that marginal price "observations" are constructed according to equation (4), this expression for a can be rewritten as $\hat{a} = (X'X)^{-1}X'X\hat{g}$, where $\hat{g}$ denotes the column vector $(\hat{g}_0, \hat{g}_1, 0)'$. Carrying out the multiplication, the result is that $\hat{a} = \hat{g}$. Furthermore, because such a regression would simply reproduce an identity, the value of R^2 corresponding to such a regression would necessarily be unity. Similar results apply for the estimation of equation (6).

To generalize this result, consider next the estimation of equation (5) using two-stage least squares. In this case, the estimator for the parameter vector a would be given by $\hat{a} = (\hat{X}'\hat{X})^{-1}\hat{X}'p_i$, which can be rewritten as $\hat{a} = (\hat{X}'\hat{X})^{-1}\hat{X}'\hat{X}\hat{g}$, or $\hat{a} = (\hat{X}'X)^{-1}\hat{X}(\hat{X} + e)\hat{g}$, where e denotes the vector of residuals from the first-stage auxiliary regression of X on a set of instrumental variables, and where $\hat{g}$ is defined as before. Noting that e must be orthogonal to the elements of X, the result that $\hat{a} = \hat{g}$ is once again derived. Unlike the case of ordinary least-squares estimation, however, the value of R^2 corresponding to this estimated equation will not equal unity, since $\hat{X}$ will not match X perfectly. Similar results apply for the estimation of equation (6).

14. The proof here is identical to that in note 13, with the obvious redefinition of X, $\hat{a}$, and $\hat{g}$.

15. See note 13 for a discussion of this point.

16. In extreme cases such as those just discussed, it may appear that the researcher can always avoid the extreme results mentioned simply by including some additional variable in the estimated marginal price function. This solution may not always be possible, however, since there is no guarantee that the data will allow the inclusion of that variable to make any effective difference in constructed marginal prices.

Because the use of constructed marginal prices may require estimated structural equations to exclude some variable that would not have to be excluded if marginal prices were directly observable, the use of constructed marginal prices may prevent estimation of structural equations that could be estimated if direct observations on marginal prices were available. To elaborate, consider an implicit market for which the structural demand and supply functions are given by equations (11) and (12) and for which Y_1 and Y_2 are matched in this market such that the following equilibrium marginal price function results: $p_i = g_0 + g_1 Z_i + u$. If p_i were directly observable, and if Y_1 and Y_2 were not collinear, the matrix of reduced form coefficients for the system given by equations (11) and (12) would be nonsingular, and those equations would be identified. But given that p_i (or equivalently, u) is not observable, that matrix will be singular when $\hat{p}_i$ is used in place of p_i, unless some other variable is included in the estimated marginal price function. It is entirely possible, however, that u might be uncorrelated with all other variables. Thus, the lack of observability of p_i may prevent the identification and estimation of equations that would otherwise be identified.

17. These expressions follow from application of the standard expression for ordinary least-squares bias in the presence of simultaneity. See, for example, Dhrymes (1974), p. 168.

18. It is unlikely that Harrison and Rubinfeld are alone in reporting biased estimates of structural parameters in implicit markets. Unfortunately, only one other structural hedonic study (Witte, Sumka, and Erekson 1974) presents sufficient information for the reader to assess the possibility of bias due to the use of constructed marginal prices. This other study is discussed in section 3.5.

19. This statement follows from the fact that equations (11) and (12) each exclude two exogenous variables and include two endogenous variables, thus satisfying the order condition for identification, while the pattern of the exclusion restrictions embodied in equations (11)–(13) allows the rank condition for equations (11) and (12) to be met as well.

20. As will be seen, differences in functional form between the equilibrium marginal price function and the structural equation to be estimated also can allow estimation of structural parameters.

21. This statement assumes that neither side of the market is characterized by complete homogeneity.

22. The comments made in the second paragraph of note 12 apply here also.

23. As will be discussed, variation in g_0 or g_1 across markets will allow identification of structural parameters in this case only if those parameters do not also vary across markets.

In the absence of such cross-market variation in structural parameters, cross-market variation in the parameters of equilibrium marginal price functions may result from differences across markets in the joint distributions of X_1, Y_2, u^d, and u^s.

24. Noting that any function can be approximated arbitrarily closely by a polynomial of a suitably chosen order, it is clear from this discussion that differences in functional form between the equilibrium marginal price function and the structural equation to be estimated also can allow estimation of structural parameters.

25. In this case, the interacted set of market-specific dummy variables implicitly included in $\hat{p}_i$ by way of cross-market variation in g_0 and g_1 would also appear in the set of structural regressors, leading to the extreme results initially discussed. It should be noted, however, that if a_0 and b_0 (or a_1 and b_1) were constant across markets, variation in g_1 (or g_0) would allow identification of structural parameters.

26. If the estimated marginal price function were correctly specified, the two-stage least-squares estimators for a_1 and a_2 in this case would be given by

$$\hat{a}_1 = \hat{g}_1 + \hat{g}_2\left[\frac{\text{cov}(X, Y_2 \mid Y_1)}{\text{cov}(Z_i, Y_2 \mid Y_1)}\right] + \hat{g}_3\left[\frac{\text{cov}(W, Y_2 \mid Y_1)}{\text{cov}(Z_i, Y_2 \mid Y_1)}\right],$$

$$\hat{a}_2 = \hat{g}_2\left[\frac{\text{cov}(X, Y_1 \mid \hat{Z}_i)}{\text{var}(Y_1 \mid \hat{Z}_i)}\right] + \hat{g}_3\left[\frac{\text{cov}(W, Y_1 \mid \hat{Z}_i)}{\text{var}(Y_1 \mid \hat{Z}_i)}\right],$$

and $\hat{a}_1$ and $\hat{a}_2$ would provide consistent estimates of a_1 and a_2. With the product of Z_i and W omitted from the first-stage estimated $P(Z)$ locus, however, the resulting estimators for a_1 and a_2 would be given by

$$\hat{\hat{a}}_1 = \hat{\hat{g}}_1 + \hat{\hat{g}}_2\left[\frac{\text{cov}(X, Y_2 \mid Y_1)}{\text{cov}(Z_i, Y_2 \mid Y_1)}\right],$$

$$\hat{\hat{a}}_2 = \hat{\hat{g}}_2\left[\frac{\text{cov}(X, Y_1 \mid \hat{Z}_i)}{\text{var}(Y_1 \mid \hat{Z}_i)}\right],$$

where $\hat{\hat{g}}_1$ and $\hat{\hat{g}}_2$ are derived from a first-stage regression that omits the product of W and Z_i from the estimated $P(Z)$ locus.

Upon comparison of these expressions with those given above, it is clear that, in general, $\hat{\hat{a}}_1$ and $\hat{\hat{a}}_2$ will be consistent for a_1 and a_2 only if W is orthogonal to Y_1, given Y_2, and to Y_2, given Y_1; and if $\hat{\hat{g}}_1$ and $\hat{\hat{g}}_2$ are consistent for g_1 and g_2. In general, this latter condition will require that Z_iW be orthogonal to the variables included in the $P(Z)$ locus.

27. This result is most easily seen by considering the ordinary least-squares estimator for a_3 in the "true" demand equation

$$p_i(Z) = a_0 + a_1 Z_i + a_2 Y_1 + a_3 W + u^d.$$

Given that the true marginal price function is equal to $g_0 + g_1 Z_i + g_2 X + g_3 W$, but that the researcher has incorrectly specified $\hat{p}_i(Z)$ as $\hat{g}_0 + \hat{g}_1 Z_i + \hat{g}_2 X$, the resulting regression of $\hat{p}_i$ on $\hat{Z}_i$, Y_1, and W can be written as

$$\begin{aligned}\hat{p}_i(Z) = (a_0 + \hat{g}_0 - g_0) + (a_1 + \hat{g}_1 - g_1)\hat{Z}_i + a_2 Y_1 + (a_3 - g_3)W \\ + [(\hat{g}_2 - g_2)X + u^d].\end{aligned}$$

Assuming that a_3 is truly zero, the ordinary least-squares estimate of a_3 will tend toward $-g_3 + (\hat{g}_2 - g_2)\beta_{XW \mid Y_1, Y_2}$, rather than zero.

28. Given that W has been incorrectly included in the estimated marginal price function, the two-stage least-squares estimators for the parameters a_1 and a_2 will be given by

$$\tilde{a}_1 = \tilde{g}_1 + \tilde{g}_2\left[\frac{\text{cov}(X, Y_2 \mid Y_1)}{\text{cov}(Z_i, Y_2 \mid Y_1)}\right] + \tilde{g}_3\left[\frac{\text{cov}(W, Y_2 \mid Y_1)}{\text{cov}(Z_i, Y_2 \mid Y_1)}\right],$$

$$\tilde{a}_2 = \tilde{g}_2\left[\frac{\text{cov}(X, Y_1 \mid \hat{Z}_i)}{\text{var}(Y_1 \mid \hat{Z}_i)}\right] + \tilde{g}_3\left[\frac{\text{cov}(W, Y_1 \mid \hat{Z}_i)}{\text{var}(Y_1 \mid \hat{Z}_i)}\right],$$

where $\tilde{g}_1$ and $\tilde{g}_2$ are derived from a first-stage regression that incorrectly includes the product of W and Z_i in the estimated $P(Z)$ locus. Assuming that $Z_i W$ is truly an extraneous variable in the first-stage estimated $P(Z)$ locus, $\tilde{g}_3$ will have a probability limit of zero, and $\tilde{g}_1$ and $\tilde{g}_2$ will remain consistent estimators for the true marginal price function. Thus $\tilde{a}_1$ and $\tilde{a}_2$ will remain consistent for a_1 and a_2. Given the ad hoc nature in which the $P(Z)$ locus is usually specified, however, with W chosen on the basis of a nonzero estimated value of $\tilde{g}_3$, it is not unlikely that the final terms in these two expressions will be nonzero in any given sample.

29. In this case, the estimated coefficient for W in the structural inverse demand function will be given by

$$\tilde{g}_3 + \tilde{g}_2\left[\frac{\text{cov}(X, W \mid \hat{Z}_i, Y_1)}{\text{var}(W \mid \hat{Z}_i, Y_1)}\right].$$

Thus, although the true structural coefficient for W may be zero, the estimated structural coefficient for W will not generally be zero and will approach the coefficient for W in the marginal price function as the sample "effect" of W on X, holding $\hat{Z}_i$ and Y_1 constant, diminishes.

It is worth noting that when W is incorrectly included in a structural equation as well as in the estimated marginal price function, the incorrect inclusion of W in the marginal price function will no longer affect the coefficient estimates for the other structural regressors.

30. In this case, the estimated structural coefficients for Z_i, Y_1, and X will be given by

$$\tilde{\tilde{a}}_1 = \tilde{g}_1 + \tilde{g}_3\left[\frac{\text{cov}(W, Y_2 \mid Y_1, X)}{\text{cov}(Z_i, Y_2 \mid Y_1, X)}\right],$$

$$\tilde{\tilde{a}}_2 = \tilde{g}_3\left[\frac{\text{cov}(Y_1, W \mid \hat{Z}_i, X)}{\text{var}(Y_1 \mid \hat{Z}_i, X)}\right],$$

$$\tilde{\tilde{a}}_X = \tilde{g}_2 + \tilde{g}_3\left[\frac{\text{cov}(X, W \mid Y_1, \hat{Z}_i)}{\text{var}(X \mid Y_1, \hat{Z}_i)}\right].$$

In the event that the extraneous variable W is uncorrelated with Y_1, Y_2, and X, when Y_1, $\hat{Z}_i$, and X are held constant, these estimators will reduce to

$$\tilde{\tilde{a}}_1 = \tilde{g}_1,$$

$$\tilde{\tilde{a}}_2 = 0,$$

$$\tilde{\tilde{a}}_X = \tilde{g}_2.$$

31. See, for example, the papers by Harrison and Rubinfeld; Linneman; McDougall; Nelson; and Witte, Sumka, and Erekson cited in note 4.

32. The following discussion is taken from J. Brown and H. Rosen (1981). Harvey S. Rosen deserves equal credit for the points that follow.

33. One would expect such coefficients to be near zero if Witte et al. had nearly reproduced their marginal price function.

References

Alonso, W. 1964. *Location and land use*. Cambridge, Mass.: Harvard University Press.

Atrostic, B. K. 1982. The demand for leisure and nonpecuniary job characteristics. *American Economic Review* 72, no. 3: 428–40.

Becker, G. S. 1965. A theory of the allocation of time. *Economic Journal* 75: 493–517.

Blomquist, G., and L. Worley. 1981. Hedonic prices, demands for urban housing amenities, and benefit estimates. *Journal of Urban Economics* 9: 212–21.

Brown, C. 1980. Equalizing differences in the labor market. *Quarterly Journal of Economics* 94: 113–34.

Brown, G., and R. Mendelsohn. 1980. The hedonic-travel cost method. University of Washington, Department of Economics. Paper presented at the American Economic Association meeting, Denver, Colorado, 1980.

Brown, J., and H. S. Rosen. 1981. On the estimation of structural hedonic price models. Princeton University, Econometric Research Program Research Memorandum no. 284. (An abbreviated version of the paper appears in *Econometrica* 50: 765–68.)

Court, L. M. 1941. Entrepreneurial and consumer demand theories for commodity spectra. *Econometrica* 9: 135–62, 241–97.

Dhrymes, P. J. 1974. *Econometrics*. New York: Springer-Verlag.

Epple, D. 1980. Hedonic prices and implicit markets: Estimating demand and supply functions for differentiated products. Graduate School of Industrial Administration, Carnegie-Mellon University. Typescript.

Freeman, A. M. III. 1974. On estimating air pollution control benefits from land value studies. *Journal of Environmental Economics and Management* 1: 74–83.

———. 1979. The hedonic approach to measuring demand for neighborhood characteristics. In *The economics of the neighborhood*, ed. David Segal, 191–217. New York: Academic Press.

Griliches, Z. 1961. Hedonic price indexes for automobiles: An econometric analysis of quality change. In *The price statistics of the federal government.* General Series no. 73, 137–96. New York: National Bureau of Economic Research.

Griliches, Z., ed. 1971. *Price indexes and quality change*. Cambridge, Mass.: Harvard University Press.

Harrison, D., and D. Rubinfeld. 1978. Hedonic housing prices and the demand for clean air. *Journal of Environmental Economics and Management* 5: 81–102.

Houthakker, H. S. 1952. Compensated changes in quantities and qualities consumed. *Review of Economic Studies* 19: 155–64.

Lancaster, K. J. 1966. A new approach to consumer theory. *Journal of Political Economy* 74: 132–56.

Lewis, H. G. 1969. Employer interests in employee hours of work. *Cuadernos de Economia* (Catholic University, Chile) 18: 38–54.

Linneman, P. 1980. An empirical methodology for analyzing the properties of public goods. *Economic Inquiry* 18: 600–616.

———. 1981. The demand for residence-site characteristics. *Journal of Urban Economics* 9: 129–48.

Lucas, R. E. B. 1975. Hedonic price functions. *Economic Inquiry* 13: 157–78.
McDougall, G. S. 1976. Hedonic prices and the demand for local public goods. *Public Finance* 31: 265–77.
Muth, R. F. 1966. Household production and consumer demand functions. *Econometrica* 34: 699–708.
Nelson, J. P. 1978. Residential choice, hedonic prices, and the demand for urban air quality. *Journal of Urban Economics* 5: 357–69.
Rosen, S. 1974. Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy* 82: 34–55.
Roy, A. D. 1950. The distribution of earnings and of individual output. *Economic Journal* 60: 489–505.
Sattinger, M. 1975. Comparative advantage and the distribution of earnings and abilities. *Econometrica* 43: 455–68.
Sider, H. 1981. Work-related accidents and the production process. BLS Working Paper 117 (April). Washington, D.C.: U.S. Bureau of Labor Statistics, Office of Research and Evaluation.
Smith, A. 1937. *An inquiry into the nature and causes of the wealth of nations*. New York: Modern Library.
Smith, R. S. 1974. The feasibility of an "injury tax" approach to occupational safety. *Law and Contemporary Problems* 38: 730–44.
———. 1979. Compensating wage differentials and public policy: A review. *Industrial and Labor Relations Review* 32: 339–52.
Tiebout, C. M. 1956. A pure theory of local expenditure. *Journal of Political Economy* 64: 416–24.
Tinbergen, J. 1956. On the theory of income distribution. *Weltwirtschaftliches Archiv* 77, no. 2: 155–75.
Witte, A. D., H. J. Sumka, and H. Erekson. 1979. An estimate of a structural hedonic price model of the housing market: An application of Rosen's theory of implicit markets. *Econometrica* 47: 1151–73.
Woodbury, S. A. 1983. Substitution between wage and nonwage benefits. *American Economic Review* 73, no. 1: 166–82.

This Page Intentionally Left Blank