Identification and Estimation of Hedonic Models[1]

Ivar Ekeland

Canada Research Chair in Mathematical Economics

University of British Columbia

James J. Heckman

University of Chicago and the American Bar Foundation

Lars Nesheim

Centre for Microdata Methods and Practice,

The Institute for Fiscal Studies and

University College London

July 21, 2003

Abstract

This paper considers the identification and estimation of hedonic models. We establish that in an additive version of the hedonic model, technology and preferences are generically identified up to affine transformations from data on demand and supply in a single hedonic market. For a very general parametric structure, preferences and technology are fully identified. This is true under a strong assumption of statistical independence of the error term. It is also true under the weaker assumption of mean independence of the error term. Much of the confusion in the empirical literature that claims that hedonic models estimated on data from a single market are fundamentally underidentified is based on linearizations that do not use all of the information in the model. The exact economic model that justifies widely used linear approximations has strange properties so the approximation is doubly poor. A semiparametric estimation method is proposed that is valid when a statistical independence assumption is valid. Alternatively, under the weaker condition of mean independence instrumental variables estimators can be applied to identify technology and preference parameters from a single market. They are justified by nonlinearities that are generic features of equilibrium in hedonic models.

JEL: C31

Ivar Ekeland
Department of Mathematics
University of British Columbia
Mathematics Annex 1210
1984 Mathematics Road
Vancouver, British Columbia
V6T 1Z2 CANADA
Phone: (604) 822-6532
Fax: (604) 822-6074
Email: ekeland@math.ubc.ca

James J. Heckman
Department of Economics
University of Chicago
1126 E. 59th Street
Chicago IL 60637
Phone: (773) 702-0634
Fax: (773) 702-8490
Email: jjh@uchicago.edu

Lars P. Nesheim
Institute for Fiscal Studies
7 Ridgmount Street
London WC1E 7AE
UNITED KINGDOM
Phone: 44 20 7291-4800
Fax: 44 20 7323-4780
Email: l.nesheim@ucl.ac.uk

1   Introduction

Sherwin Rosen pioneered the analysis of hedonic markets in a perfectly competitive set-
ting. He also proposed an econometric identification strategy for recovering preferences
and technology from hedonic markets. His hedonic model characterizes markets for het-
erogeneous goods (or characteristics or amenities) that implicitly price out the attributes
that characterize the goods (or characteristics or amenities).

Rosen's fundamental paper has shaped the way economists think about the pricing of
heterogeneous characteristics or attributes. Yet for two reasons, the full potential of his
method remains to be exploited. First, except for special cases, high dimensional hedonic
models with multiple characteristics require solutions of complicated partial differential
equations to fully characterize market equilibrium. This renders difficult theoretical analy-
ses which require computation of nonlinear implicit equations. Second, the method of
identification of preferences and technology proposed by Rosen has been severely criticized
in the literature. It is widely held that the preferences and technology generating hedonic
models are identified only through arbitrary functional form and exclusion assumptions,
especially when they are estimated on data from a single market.

This paper considers whether equilibrium in hedonic markets imposes any restrictions
on estimating equations and whether it is possible to identify technology and preferences
from data on a single hedonic market. We consider both parametric and nonparametric
versions of these questions.

We show that the hedonic model has empirical content even in a single market. For
very general parametric families, the hypothesis of equilibrium imposes very tight restric-
tions on the data. This is true under the assumption that the error term, representing

unobserved heterogeneity in productivity or preferences, is statistically independent of observed preference and productivity traits. It is also true under the weaker assumption of mean independence. Preferences and technology are generically identified from data on a single hedonic market. For the nonparametric case, we establish generic identification of technology and preference parameters up to affine transformations, the standard level of identification that can be obtained for market choice equations.

We establish that commonly used linearization strategies made to simplify estimation problems produce identification problems. The hedonic model is generically nonlinear. The nongeneric functional form assumptions made in the applied literature give rise to the identification problems that are widely thought to be fatal to Rosen's empirical methodology. We go on to show that the economic model which produces the widely-used linear estimating equations is implausible, so the approximation is doubly poor.

Based on our identification analysis we propose estimators that can recover the parameters of a hedonic model in a single market. The first is based on semi-parametric transformation model methods and requires estimation of conditional densities and their derivatives. This method is valid when the error term is independent of the observed exogenous variables. The second is based on instrumental variables techniques and simply requires estimation of conditional mean functions. This method is valid under the weaker assumption that the error term is mean independent of the observed exogenous variables.

Our identification analysis also applies to a broader class of empirical models of nonlinear pricing: models of the effects of taxes on behavior when taxes are set optimally (Mirrlees, 1971), and a model of monopoly pricing (Mussa and Rosen, 1978, Wilson, 1993). It also applies to the standard problem of taxes and labor supply (Heckman 1974; Hausman, 1980). For specificity, in this paper we focus on the hedonic model, briefly discussing other

applications in the conclusion.

This paper proceeds in the following way. In Section 2, we present the hedonic model and review an important special case, the linear-quadratic-normal model due to Tinbergen (1956) and used by Epple (1987). This model gives rise to closed form solutions and justifies widely used linear estimating equations for hedonic models as exact solutions. In Section 3, we discuss the peculiar properties of this model. While it widely recognized that this model has very special properties and is too restrictive for most applications, much of the hedonic literature focuses on linearized versions of more general hedonic models. These linear approximations however are rarely justified. Unless the true model is nearly linear or nearly additive, they cannot be justified. They implicitly assume that the true model is nearly linear or nearly additive. The influential criticism of Rosen's estimating strategy by James Brown and Harvey Rosen (1982) is based on a linear-quadratic approximation of the estimating equations from some true model. This approximation is exact when the true model is the Tinbergen model. However, when the Tinbergen model is slightly modified, the Brown-Rosen critique no longer applies. In a later section, we prove a theorem (Theorem 1) that establishes that for a general class of models, the Brown-Rosen critique only applies to a special, nongeneric, case. In Section 3 we go on to review standard discussions of instrumental variables methods applied to estimate preferences and technologies in hedonic markets: (a) that sorting implies that within a single market, there are no natural exclusion restrictions, so instrumental variables methods cannot be applied (Epple, 1987; Kahn and Lang, 1988) and (b) that identification can only be secured by multimarket data. We challenge both assertions in this paper and question conventional applications of the multimarket method.

In Section 4, we establish (a) the identifiability of the hedonic model within a single

market for a broad class of parametric models (polynomials of any finite order or any model belonging to a finite dimensional vector space); and (b) the identification of the hedonic model up to levels for a broad class of nonparametric models. We also establish that using all of the information from both sides of the hedonic market together adds nothing to what can be identified analyzing the supply side and demand side separately in conjunction with the hedonic pricing function. We show how extra information on levels of outcomes, rather than just pricing and demand equations, aids in identifying the missing level set information. In Section 5, we show that when the assumption of independence which is used in Section 4 is weakened to mean independence, the identification result still holds for hedonic models in a single cross section, as well as for a broad class of parametric models. Theorems 4 and 5 justify the application of instrumental variables in the general parametric case. We briefly discuss instrumental variables estimation and extensions of the existing literature to cover the nonparametric case. Section 6 presents some conclusions and suggestions for future research.

2 The Hedonic Model: General Results and An Important Special Case With A Closed Form Solution

We first present a general statement of the hedonic model. For simplicity, consider a labor market setting. The model is static. Consumers (workers) match to single worker firms. Let z be an attribute vector characterizing jobs. $P(z)$ is the earnings of workers supplying attribute vector z, which is a disamenity. Let R be unearned income. We define $U(c, z, x, \varepsilon, A)$ as the preferences of workers where x and ε represent observable (to

the econometrician) and unobservable characteristics of workers that vary across persons, A represents preference parameters common across persons and c is consumption where $c = P(z) + R$. Higher values of z lead to lower values of U. For ease of exposition, we assume $R = 0$. Given $P(z)$, a twice continuously differentiable price function, and assuming the utility function is twice differentiable[2], we obtain the following first and second order conditions for a maximum

$$U_c\left(c, z, x, \varepsilon, A\right) P_z\left(z\right) + U_z\left(c, z, x, \varepsilon, A\right) = 0 \tag{1}$$

$$U_{zz'} + U_c P_{zz'} + 2 P_z U_{cz'} + P_z U_{cc'}\left(P_z\right)' \text{ is negative definite.} \tag{2}$$

These characterize optimal job attribute choices for each worker. For each location z in attribute space, they characterize the set of workers who choose that location.

Firms demand attribute z and maximize profits which equal output $\Gamma(z, y, \eta, B)$ minus production costs $P\left(z\right)$ where y and η are observable and unobservable vectors of technology parameters that vary across firms and B is a common technology parameter shared by all firms. Observable and unobservable are defined with respect to what the econometrician observes. We assume that the production function is twice differentiable. Profits are

$$\Pi\left(z, y, \eta, B, P\left(z\right)\right) = \Gamma(z, y, \eta, B) - P(z)$$

and the first and second order conditions for a maximum are

_____

[2]For expositional convenience, we restrict our analysis to economies in which the equilibrium price function is smooth. Similar analyses can be done for economies in which the equilibrium price function is not smooth. For an example of an economy with smooth technologies and absolutely continuous distributions of consumer heterogeneity in which the equilibrium price function is piecewise twice continuously differentiable see Nesheim (2001). For other examples of sorting problems with non-smooth pricing functions see Wilson (1993).

$$\Gamma_z\left(z, y, \eta, B\right) - P_z\left(z\right) = 0 \tag{3}$$

$$\Gamma_{zz'} - P_{zz'} \text{ is negative definite.} \tag{4}$$

Throughout we follow the classical hedonic literature and we assume the regular case in which the second order conditions hold as strict inequalities, $\Gamma_{z\eta'}$ is positive definite, and $P_z U_{c\varepsilon'} + U_{z\varepsilon'}$ is positive definite. These conditions guarantee positive sorting on unobservables in the sense that in equilibrium $\frac{\partial \eta}{\partial z} > 0$ and $\frac{\partial \varepsilon}{dz} > 0$.

Workers differ in their preference vectors x and ε. Firms differ in their productivity vectors y and η. Let the densities of x and ε be f_x and f_ε and let x be independent of ε. x and ε have supports $\mathcal{X}$ and $\mathcal{E}$ respectively. The densities of y and η are f_y and f_η. y is independent of η and y and η have supports $\mathcal{Y}$ and $\mathcal{H}$ respectively. We assume that x, ε, y, and η are absolutely continuous random variables. In this paper, we focus on the case in which $\dim\left(\varepsilon\right) = \dim\left(\eta\right) = \dim\left(z\right)$ and in which there is no bunching in equilibrium. That is, in equilibrium every bundle of characteristics has population measure zero of demanders or suppliers. This is the classical case analyzed in Rosen (1974) and the subsequent literature. For an analysis of equilibria with bunching see Heckman, Matzkin, and Nesheim (2002).

Given the assumptions above, a local implicit function theorem applies and we can invert the first order conditions (FOC) (1) and (3) to obtain ε and η as functions of z and x and y, respectively. Inverting the FOC (1) for the worker we obtain

$$\varepsilon = \varepsilon\left(z, P_z, P(z), x, A\right).$$

Similarly, inverting the FOC (3) for the firm we obtain

$$\eta = \eta\left(z, P_z, y, B\right).$$

Using these relationships, we use f_x and f_ε to find the density of z supplied given $P(z)$, and we use f_y and f_η to find the density of z demanded given $P(z)$.

The **Supply Density** is:

$$\int_{\mathcal{X}} f_\varepsilon\left(\varepsilon\left(z, P_z, P(z), x, A\right)\right) \det\left[\frac{d\varepsilon\left(z, P_z, P(z), x, A\right)}{dz}\right] f_x\left(x\right) dx.$$

where the term in square brackets is the Jacobian matrix with respect to vector z (i.e., its effect on all arguments of ε that depend on z). This is the density of the amenity supplied as a function of the price function, preference parameters A, and the densities of x and ε.

The **Demand Density** is:

$$\int_{\mathcal{Y}} f_\eta\left(\eta\left(z, P_z, y, B\right)\right) \det\left[\frac{d\eta\left(z, P_z, y, B\right)}{dz}\right] f_y\left(y\right) dy.$$

Again, the term in square brackets is the Jacobian matrix with respect to vector z. This is the density of demand for a given price function, vector of technology parameters B, and pair of densities of y and η. From the second order conditions (2) and (4), respectively, the Jacobian terms are both positive.

Equilibrium in hedonic markets requires that supply and demand be equated at each point of the support of z. Hence, equilibrium prices $P\left(z\right)$ must satisfy the following second order partial differential equation

$$\int_{\mathcal{X}} f_\varepsilon\left(\varepsilon\left(z, P_z, P(z), x, A\right)\right) \det\left[\frac{d\varepsilon\left(z, P_z, P(z), x, A\right)}{dz}\right] f_x\left(x\right) dx = \tag{5}$$

$$\int_{\mathcal{Y}} f_\eta\left(\eta\left(z, P_z, B\right)\right) \det\left[\frac{d\eta\left(z, P_z, y, B\right)}{dz}\right] f_y\left(y\right) dy$$

The solution depends on U, the utility function of the workers, Γ, the technology of firms, and the pairs of density functions (f_x, f_ε) and (f_y, f_η) characterizing the population distributions of workers and firms respectively. Additionally, we impose that workers and

7

firms must receive wages and profits above reservation levels in order to participate in the market. This generates the boundary conditions that determine the solution of the partial differential equation. This entry condition also plays a role in the identification analysis. We examine the empirical content of the restrictions imposed by the equilibrium equation and these conditions in the rest of this paper.

When z is scalar and utility is quasi-linear so that $U(c, z, x, \varepsilon, A) = c - V(z, x, \varepsilon, A)$, $\frac{d\varepsilon}{dz} = \frac{P_{zz} - V_{zz}}{V_{z\varepsilon}}$ and $\frac{d\eta}{dz} = \frac{P_{zz} - \Gamma_{zz}}{\Gamma_{z\eta}}$. Since $V_{z\varepsilon} < 0$ and $\Gamma_{z\eta} > 0$, we can substitute these expressions into (5) to obtain an explicit expression for P_{zz}

$$P_{zz} = \frac{\int\limits_{\mathcal{X}} f_\varepsilon f_x \left(\frac{V_{zz}}{-V_{z\varepsilon}}\right) dx + \int\limits_{\mathcal{Y}} f_\eta f_y \left(\frac{\Gamma_{zz}}{\Gamma_{z\eta}}\right) dy}{\int\limits_{\mathcal{Y}} \frac{f_\eta f_y}{\Gamma_{z\eta}} dy + \int\limits_{\mathcal{X}} \frac{f_\varepsilon f_x}{-V_{z\varepsilon}} dx} \tag{6}$$

where the arguments of the functions have been suppressed for ease of exposition. In equilibrium, the curvature of the pricing function is a weighted average of the average curvature of the workers' utility and the average curvature of the firms' technology. The weights at any particular point in z space depend on the ratio of the densities of worker and firm heterogeneity.

Hedonic equilibrium is illustrated in Figure 1. The figure shows the optimal job sorting choices of three firm-worker pairs. The solid line depicts the equilibrium price function. The dotted lines depict firm output as a function of job type for three different firms.[3] Each firm chooses the job type z where the output function is tangent to the price function. The dashed lines depict worker disutility as a function of z for three different workers.[4]

---

[3] For each firm, this output function has been shifted vertically by substracting off each firm's equilibrium profits so that all three plots fit in the same figure.

[4] Each worker's disutility curve has been shifted vertically by subtracting off equilibrium utility so that all three plots fit in the same figure.

Each worker chooses the job type where disutility is tangent to the price function. Each worker matches with a firm so that the worker disutility function is tangent to the output function of his matched firm. In each case the curvature of firm output is less than the curvature of the price function which is less than the curvature of worker disutility. The curvature of the pricing function is a weighted average of the curvature of firm profits and the curvature of worker disutility. If this were not the case, then the firms and workers would not both be choosing optimal job types. In the special cases where all firms are alike or all consumers are alike the hedonic pricing function corresponds, respectively, to the firm profit or worker disutility functions. Otherwise, the curvature of the hedonic function differs from the curvature of technology or preference functions. This difference in curvature which is a fundamental characteristic of the equilibrium, provides the basis for the econometric identification results in this paper.

We next present a linear-quadratic model with normal heterogeneity due to Tinbergen (1956) that has a closed form expression. This is the model that justifies widely used empirical approximations as exact descriptions and provides an intuitive introduction to the hedonic model.

2.1 A Linear-Quadratic-Normal Example

Assume that preferences are quadratic in z and linear in c, unearned income $R = 0$, and that individual heterogeneity (x, ε) only affects utility through the single index $\theta = \mu_\theta(x) + \varepsilon$

where $\dim(\theta) = \dim(z)$ [5]. Workers maximize

$$U(c, z, \theta, A) = P(z) + \theta' z - \frac{1}{2} z' A z.$$

The conditions determining a worker's maximum are

$$P_z + \theta - Az = 0$$

where $P_{zz'} - A$ is negative definite. On the firm side, assume that the production function is quadratic in z and that firm heterogeneity only affects profits through the single index $\nu = \mu_\nu(y) + \eta$ where $\dim(\nu) = \dim(z)$. Profits are

$$\Pi(z, \nu, B, P(z)) = \nu' z - \frac{1}{2} z' B z - P(z)$$

and the conditions determining a firm's optimum are

$$\nu - Bz - P_z = 0$$

where $-(B + P_{zz'})$ is negative definite. The distributions of θ and ν in the population are normal. The distribution of θ is $\theta \sim N(\mu_\theta, \Sigma_\theta)$, and the distribution of ν is $\nu \sim N(\mu_\nu, \Sigma_\nu)$.

An arbitrary price function induces a density of demand and a density of supply at every location z. The equilibrium price function can be found by equating these densities at every point z and solving the differential equation (5). However, in the linear-quadratic-normal case one can correctly guess that the solution to the problem is quadratic in z

$$P(z) = \pi_0 + \pi_1' z + \frac{1}{2} z' \pi_2 z$$

---

[5] The model in this example was first analyzed by Tinbergen (1956) and has been used by Epple (1987) and Tauchen and Witte (2001) among others.

and then find the coefficients (π_0, π_1, π_2) that satisfy the equilibrium equation. Assuming that the price function is quadratic, the first order condition for a worker is

$$\pi_1 + \pi_2 z + \theta - Az = 0. \tag{7}$$

For a firm, it is

$$\nu - Bz - \pi_1 - \pi_2 z = 0. \tag{8}$$

The second order conditions require that both $A - \pi_2$ and $B + \pi_2$ are positive definite. Thus we may solve for z from (7) to obtain

$$z = (A - \pi_2)^{-1}(\theta + \pi_1) \tag{9}$$

and from (8) to obtain

$$z = (B + \pi_2)^{-1}(\nu - \pi_1). \tag{10}$$

These equations define mappings from workers θ and firms ν to job types z. These mappings determine the density of supply and demand at every bundle of characteristics or attributes and the types of workers and firms at every location. Equilibrium is characterized by a vector π_1 and a matrix π_2 that equate demand and supply at all z. However, since both θ and ν are normally distributed, this only requires equating the mean and variance of supply and demand.

The mean supply $E^S(z)$ is obtained from (9):

$$(\text{Average Supply}) \quad E^S(z) = (A - \pi_2)^{-1} E(\theta + \pi_1) \; .$$

The mean demand is obtained from (10):

$$(\text{Average Demand}) \quad E^D(z) = (B + \pi_2)^{-1} E(\nu - \pi_1) \; .$$

Since $\mu_\theta = E(\theta)$ and $\mu_\nu = E(\nu)$, the condition $E^S(z) = E^D(z)$ implies that

$$\text{(Equality of means)} \quad (A - \pi_2)^{-1}(\mu_\theta + \pi_1) = (B + \pi_2)^{-1}(\mu_\nu - \pi_1) \ .$$

Rearranging terms, we obtain an explicit expression for π_1 in terms of $A, B, \mu_\theta, \mu_\nu$ and π_2 :

$$\pi_1 = [(A - \pi_2)^{-1} + (B + \pi_2)^{-1}]^{-1}[-(A - \pi_2)^{-1}\mu_\theta + (B + \pi_2)^{-1}\mu_\nu].$$

To determine π_2, compute the variances of supply and demand from (9) and (10) respectively to obtain:

$$\Sigma_z^S = (A - \pi_2)^{-1}\Sigma_\theta \left[(A - \pi_2)^{-1}\right]'$$

$$\Sigma_z^D = (B + \pi_2)^{-1}\Sigma_\nu \left[(B + \pi_2)^{-1}\right]'$$

where Σ_z^S is the variance of supply and Σ_z^D is the variance of demand. From equality of variances of the demand and supply distributions we obtain an implicit equation for π_2 :

$$\text{(Equality of variances)} \quad (A - \pi_2)^{-1}\Sigma_\theta \left[(A - \pi_2)^{-1}\right]' = (B + \pi_2)^{-1}\Sigma_\nu \left[(B + \pi_2)^{-1}\right]' \ .$$

We pin down initial conditions using the restrictions that $U \geq \bar{U}$, a reservation value, and profits are positive ($\Pi \geq 0$). After taking into account the equilibrium relationship between ν and z, equilibrium profits as a function of z are $\frac{1}{2}z'(B + \pi_2)z - \pi_0$. Hence nonnegativity of profits implies $-\pi_0 \geq 0$ since $(B + \pi_2)$ is positive definite by the second order conditions and we have to allow for the possibility of $z = 0$. Setting reservation utility equal to zero, a similar argument on the worker side implies $\pi_0 \geq 0$. Hence $\pi_0 = 0$.

Once we have solved for π_1 and π_2, (9) and (10) also define the equilibrium matching function linking the characteristics of suppliers (θ) to those of demanders (ν). Substituting out for z, this function is

$$(A - \pi_2)^{-1}(\theta + \pi_1) = (B + \pi_2)^{-1}(\nu - \pi_1)$$

so, the equilibrium relationship between θ and ν is

$$\theta = (A - \pi_2)(B + \pi_2)^{-1}(\nu - \pi_1) - \pi_1. \tag{11}$$

Because of sorting, equilibrium worker and firm characteristics are related. We discuss the implications of this relationship in Sections 3 and 4.

In the separable case where Σ_θ, Σ_ν, A, and B are diagonal, π_2 is diagonal. Effectively, this is a scalar case where each attribute is priced separately. In the scalar case, equality of variances implies that $(A - \pi_2)^2 \Sigma_\nu = (B + \pi_2)^2 \Sigma_\theta$. The second order conditions imply that $A - \pi_2 > 0$ and $B + \pi_2 > 0$. Defining $\sigma_\theta = (\Sigma_\theta)^{\frac{1}{2}}$ and $\sigma_\nu = (\Sigma_\nu)^{\frac{1}{2}}$, this implies that[6]

$$\pi_2 = \frac{A\sigma_\nu - B\sigma_\theta}{\sigma_\theta + \sigma_\nu}.$$

$$\pi_1 = \frac{-\mu_\theta \sigma_\nu + \mu_\nu \sigma_\theta}{\sigma_\theta + \sigma_\nu}.$$

π_2, the curvature of the price function, is a weighted average of the curvatures of workers' and firms' preference and technology functions. π_1 is a weighted average of the means of worker and firm distributions of heterogeneity. In both expressions, the weights depend on the relative variances of worker and firm heterogeneity. If workers are much more heterogeneous than firms $\sigma_\theta >> \sigma_\nu$, π_2 will approximately equal B, the curvature of firms' technology. If $\sigma_\theta = \sigma_\nu$ and $A = B$, $\pi_2 = 0$ is a solution and the equilibrium price function is linear in z. If $\sigma_\theta = \sigma_\nu$, but $A \neq B$, then $\pi_2 = \dfrac{A - B}{2}$. In the polar cases when $\sigma_\theta = 0$ or $\sigma_\nu = 0$ then there is effectively only one type of consumer or one type of firm respectively. If $\sigma_\theta = 0$ and $\sigma_\nu > 0$, then $\pi_2 = A$ and $\pi_1 = -\mu_\theta$. Then prices reveal the parameters of consumer preferences. If $\sigma_\nu = 0$ and $\sigma_\theta > 0$, $\pi_2 = B$ and $\pi_1 = \mu_\nu$. These two polar cases are discussed in Rosen (1974) and are the ones that dominate discussions in the

---

[6]The other root of the equation violates second order conditions.

empirical literature on hedonic models. Only in these two polar cases do prices directly reveal consumer preferences or firm productivities respectively. Similar results hold when z, θ, and ν are vectors.

3 Identifying and Estimating The Model

Sherwin Rosen stressed the importance of taking theory to data. He considered the problem of recovering technology and prefereo framed the empirical questions about hedonic models that have occupied the attention of economists for the past 30 years.

He analyzed the problem of estimating parameters of preference and technology functions using data on prices $P(z)$ and characteristics choices of agents in hedonic markets. Using the first order conditions (1) and (3) ((7) and (8) in the linear-quadratic-normal example) he proposed a two step method for estimating both preference and technology parameters. He did not consider direct estimation of production, profit or preference functions, a source of information we consider in section four. We simply note here that if there are no missing attributes, we can recover the production function directly from data on inputs and outputs using standard methods. Nevertheless, even if production (or profit) data are available, data on utility are not, so the problem considered by Rosen still remains for recovering the parameters of worker preferences.

>From our discussion of the linear-quadratic-normal case, the parameters π_1 and π_2 do not directly identify either preference or technology parameters except when $\Sigma_\theta = 0$ or $\Sigma_\nu = 0$ respectively. In general, the pricing function combines parameters of technology, preferences and distributions of heterogeneity, a result that is evident in equation (6).

The most direct approach to estimating the hedonic model would be to solve equation

14

(5) for $P(z)$ in terms of the parameters of preferences, technology and the distributions of tastes and productivity and to jointly estimate the demand functions and supply functions and distributions of preference and technology parameters exploiting all of the information in the equilibrium conditions including data on demand, supply and the pricing function. That approach is computationally complicated and does not transparently deliver identification of the deep structural parameters.

Rosen suggested an intuitively plausible and computationally simpler two step estimation procedure that has been widely used and widely criticized. In step 1 of his procedure, the analyst estimates $P(z)$ from market data. In step 2, the analyst uses first order conditions (1) and (3) in conjunction with the marginal prices obtained from step 1 to recover preferences and technology respectively.

In the context of the linear-quadratic-normal example, the analyst first estimates the pricing function $P(z)$ and forms the marginal prices from estimates of π_1 and π_2 and then estimates the curvature parameters of technology, and preferences using the estimated marginal prices in (7) and (8) respectively. Specifically, this method estimates A, B, $\mu_\theta(x)$, and $\mu_\nu(y)$ from the system of equations

$$\hat{\pi}_1 + \hat{\pi}_2 z = -\mu_\theta(x) + Az - \varepsilon \tag{12}$$

$$\hat{\pi}_1 + \hat{\pi}_2 z = \mu_\nu(y) - Bz + \eta \tag{13}$$

where as stated previously $\theta = \mu_\theta(x) + \varepsilon$, $\nu = \mu_\nu(y) + \eta$, and "^" denotes estimate.

In two influential papers, James Brown and Harvey Rosen (1982) and James Brown (1983) analyze the regression method based on (12) and (13). These papers contain most of the main ideas in the empirical literature on hedonics that emerged following Rosen's paper. They interpret (12) and (13) as linearized *approximations* to (1) and (3). The linear-

15

quadratic-normal model of Section 2 is the framework for which these approximations are exact.

In this approximation interpretation, the distributions of θ and ν (determined by the distributions of (x, ε) and (y, η) respectively) are kept in the background. Standard linear econometric methods are applied to identify the parameters of (12) and (13) and connections among the parameters of preferences, technology and the distributions of tastes and productivity are not made explicit. Issues of identification are confused with issues of estimation. Common to an entire genre of empirical economics, this literature focuses on finding "good instruments" and misses basic sources of identification in hedonic models.

Starting from (12) and (13), Brown (1983) and Brown and Rosen (1982) make three points which have frequently been reiterated in the subsequent empirical literature.

Point One: Identification Can Only Be Obtained Through Arbitrary Functional Form Assumptions

Since z is on both sides of (12) and (13), by a property of least squares, a regression using the constructed price $\hat{P}_z(z) = \hat{\pi}_1 + \hat{\pi}_2 z$ as the dependent variable in (12) or (13) only identifies π_2 even if $\mu_\theta(x)$ and $\mu_\nu(y)$ are functions of regressors. This argument is not necessarily fatal. In the special cases where there is no preference or no technology heterogeneity the method identifies preference or technology parameters respectively.

However, if the constructed price is a nonlinear function of z, this argument no longer holds. The nonlinear variation in $\widehat{P}_z(z)$ gives an added piece of information that can help to identify technology and preference parameters.[7] This identification strategy rules out collinearity between z and $\widehat{P}_z(z)$, but such nonlinearity is widely viewed as an artificial source of identification that is thought to be "arbitrary." In Theorem 1 in section 4, we

---

[7] See Fisher (1966) for an early discussion of the value of nonlinearities in identifying econometric models.

prove that this nonlinearity is a generic feature of equilibria in hedonic models. Generically, $P_z(z)$ is not a linear function of z. More generally, we prove that as a property of hedonic equilibria, generically, the curvature of z in $P(z)$ is different from the curvature of z in both technology and preference parameters, a result that might be anticipated from inspection of equation (6) characterizing the relationship between the curvature of technology, preference and price second partials.

Point Two: Absence of Instruments

Even if such "arbitrary" assumptions are made, so that we can use the nonlinearity in $\widehat{P}_z(z)$ to help identify the parameters and circumvent Point One, we still face standard endogeneity problems. As a property of sorting equilibria, z is correlated with ε and η in (12) and (13) respectively. Moreover, as discussed by Epple (1987), Bartik (1987), and Kahn and Lang (1988), exclusion restrictions from the other side of the market cannot be justified. In the notation of this section, the equilibrium matching condition (11) of Section 2 implies that

$$\mu_\theta(x) + \varepsilon = (A - \pi_2)(B + \pi_2)^{-1}(\mu_\nu(y) + \eta - \pi_1) - \pi_1.$$

In addition, conditional on z, the unobservable ε (or η) becomes stochastically dependent on the observables x and y, respectively even if they are independent in the underlying population.

Thus even if x is independent of ε, and y is independent of η, equilibrium y is a function of ε and equilibrium x is a function of η so x is endogenous in the supply equation and y is endogenous in the demand equation.

With data from a single market, one is forced to hunt for "clever" instruments which lack a solid economic foundation. Thus, even if "arbitrary" nonlinearities are invoked to

17

surmount point one, standard instruments appear to be lacking. In Sections 4 and 5 we show that the economics of the model guarantees valid instruments even though there are no exclusion restrictions.

Point Three: Use of Multimarket Data

Rosen (1974), Brown and Rosen (1982), Epple (1987), Kahn and Lang (1988), and Tauchen and Witte (2001) consider estimation of the first order conditions using multi-market data either across different markets (or economic regions), or across time in the same market. In this case, if we assume that preference parameters common across agents remain constant across markets while distributions of individual heterogeneity vary across markets, we can use cross market variation in prices and location choices to estimate the common preference parameters. This identification strategy relies on assumptions that can be tested if hedonic models can be identified in a single market. Using the techniques we develop in this paper, the structure of hedonic models can be estimated and identified using data from a single market for a class of additive parametric structures that includes as a special case the linear approximation version that has been the focus of nearly all empirical applications of hedonic models.

3.1   Using All Of The Economics of The Model

These criticisms are symptoms of a deeper problem with the current literature. The full economic content of the hedonic model is not being exploited. We argue that when it is exploited, the model is generically identified even within a single market without having to invoke assumptions about arbitrary functional forms. We develop this point formally in the next section. Here we develop the intuition for it using the linear-quadratic model.

18

Consider the restrictive assumptions of the linear-quadratic-normal model and its economic implications. Both consumer and firm heterogeneity have normal distributions (not approximately normal but exactly normal). Reservation profits equal reservation utilities equal zero identically so that all agents enter the market. Finally, a positive fraction of the population chooses equilibrium attribute levels where marginal prices are negative, marginal products are negative, and marginal disutilities are positive.[8] Minor alterations of the model to relax any of these restrictive features makes the marginal pricing function nonlinear in z and makes, Point One irrelevant for the linear-quadratic example of Section 2.

The linear-quadratic-normal model of Section 2 results in an equilibrium with a linear marginal price function. This equilibrium produces an econometric system that is not identified except for special cases (Brown-Rosen Point One). In this example, it would be incorrect to impose that the marginal price function is nonlinear. However, the model in Section 2 is very special. It belongs to a very small class of models that produce an equilibrium marginal price function that is linear in z. In the next section we prove as a special case of a more general theorem that most models "close" to the linear-quadratic models of Section 2 do not produce linear marginal price functions where we define "close" more precisely in the next section. In these models, it is not arbitrary to impose nonlinear marginal price functions.

To see how fragile Point One is, suppose that we change the scalar version of the model

---

[8]This last problem can be reduced by proper choice of parameters so that the fraction choosing locations with negative marginal prices is made small. The marginal price is $P_z = \pi_1 + \pi_2 z$. In equilibrium this is a normal random variable with mean $E(P_z) = \frac{A\mu_\nu - B\mu_\theta}{A+B}$ and standard deviation $STD(P_z) = \frac{|A\sigma_\nu - B\sigma_\theta|}{A+B}$.

to have non-normal θ and ν. Worker preferences are

$$U(z) = \theta z - \frac{A}{2}z^2 + P(z)$$

with first order condition

$$\theta - Az + P_z(z) = 0.$$

Profits are

$$\Pi(z) = \nu z - \frac{B}{2}z^2 - P(z)$$

with first order condition

$$\nu - Bz - P_z(z) = 0.$$

Unlike the Tinbergen model, θ and ν distributed as mixtures of two normals. We call this Model 1. Using the parameter values reported in appendix B.1, we solve the equilibrium differential equation (5) for this model numerically for three cases. Figures 2 and 3 show the equilibrium marginal price functions and curvatures of the price function for this model for each of the three cases. The bottom panels of each figure show the population density at every location z for each case. The weight (λ) on the first component of the mixtures of normal distributions is varied across cases. In the first case $\lambda = 1$, ν and θ are normally distributed. In this case, the marginal price is linear in z and the curvature of the price function is a constant. In the second case, $\lambda = 0.9$ and ν and θ now distributed as mixtures of normals with weights $\lambda = 0.9$ on the first component and $1 - \lambda = 0.1$ on the second component. With this minor perturbation, the marginal price function becomes nonlinear and the curvature is not constant. When $\lambda = 0.5$, the effect on the pricing equation is even more dramatic. The price function is nonlinear over a wide stretch of the z and the second derivative of the price function is far from constant. This is true for a model that is very

nearly the normal linear quadratic model. In general, for models that are not close to the non-generic linear-quadratic-normal Tinbergen model, Brown-Rosen Point One does not apply.

These figures also reveal unattractive properties of the linear-quadratic model. Negative and positive quantities of z are demanded and supplied and marginal prices are negative for a large portion of the population. While the importance of these features can be reduced by appropriate choice of parameter values, they are intrinsic features of the linear-quadratic-normal model. One model that imposes a set of restrictions that eliminates these features is Model 2 displayed in Figures 4 and 5. In this model, we restrict marginal prices to be positive and restrict characteristics to be non-negative ($z > 0$) by writing down a transformation of the linear quadratic model. In this model, firms' profits are $\frac{exp(\nu)z^{1-B}}{1-B} - P(z)$ and workers' utility is $-\frac{\exp(-\theta)z^{1+A}}{1+A} + P(z)$ where ν and θ are distributed as mixtures of normals. On the firm side $\nu = \nu_0 + \nu_1'y + \eta$ where y and η are independent and distributed as mixtures of normals. The exact specification of parameter values used to construct this example is given in appendix B.2. In this model the firms' first order condition, after taking logarithms, is $\ln(P_z(z)) = -B\ln z + \nu_0 + \nu_1'y + \eta$. This is a simple logarithmic transformation of the linear-quadratic model. For model 2, we computed numerical solutions to the equilibrium differential equation (5). Now marginal prices are positive and only positive quantities of the characteristic are demanded and supplied. In this case, we require that $\ln P_z(z)$ not be a linear function of $\ln z$ in order to identify the parameters. Figures 4 and 5 show that there is a case where $\ln P_z(z)$ is a linear function of $\ln z$; namely the case where $\lambda = 1.0$ so that θ and ν are distributed as normal random variables. Figure 4 explicitly plots the derivative of $\ln P_z(z)$ in z, which is a straight line in the Tinbergen case. In the other two cases when $\lambda = 0.9$ or 0.5, however, $\ln P_z(z)$ is a non-linear function of $\ln z$.

Other examples can be generated in the context of the linear-quadratic-normal model by assuming that reservation utility does not equal zero so that not all agents enter the market or by imposing restrictions on the technology in other ways. By imposing economically plausible restrictions, Brown-Rosen Point One is shown to be less cogent. In Section 4 we show that these examples are generic.

Even though Point One is non-generic, Point Two remains. There are apparently no valid instruments for z on the right hand sides of (12) and (13). A strategy needs to be found to deal with the endogeneity of z. In the next two sections, we discuss two such strategies and present general results for a model with a single characteristic. We do not invoke arbitrary functional form restrictions or distributional assumptions and establish that the hedonic model is generically identified from data from a single market. Even though there are no conventional exclusion restrictions that justify the use of demand side variables as instrumental variables in the supply equations (or vice versa), instrumental variables that use non-linear transformations of the functions of the exogenous variables in each estimating equation are generically valid instruments.

4   Parametric and Nonparametric Analyses of a One Dimensional Model with Additively Separable First Order Conditions

This section analyzes a class of one dimensional hedonic models with additive separability in the first order conditions but with no specific functional form or distributional assumptions imposed. The one dimensional case allows us to abstract from a variety of problems

that we address in our other work: (a) questions of existence of solutions to partial differential equations and (b) questions about the proper treatment of missing attributes in a multidimensional model.[9] Both types of questions are important but they distract us from the basic questions of identification and testability of the hedonic model posed in the introduction to this paper.

We analyze a class of models in which z, ε, and η are one dimensional, preferences are linear in consumption, and both preferences and technologies belong to a restricted class so that monotonic transformations of consumer and firm first order conditions are additively separable in (z, x, ε) and (z, y, η) respectively.

Rewrite the consumer and firm first order conditions (1) and (3) as

$$P'(z) = -U_z(z, x, \varepsilon)$$

$$P'(z) = \Gamma_z(z, y, \eta)$$

where $P'(z) = P_z(z)$ where we assume that utility is linear in consumption. U_z and Γ_z are nonparametric functions. We assume the following type of separability for U_z and Γ_z

$$-U_z(z, x, \varepsilon) = M_1(\varphi_1(z) + \psi_1(g_1(x) - \varepsilon))$$

$$\Gamma_z(z, y, \eta) = M_2(\varphi_2(z) + \psi_2(\eta - g_2(y)))$$

where $M_1, M_2 \in C^2(R)$ and $\psi_1, \psi_2 \in C^2(R)$ are known and are strictly monotonically increasing in their arguments.[10] Linear-quadratic models are an element of this class of models in which M_1, M_2, ψ_1, and ψ_2 are identity functions and $\varphi_1(z) = Az$, $\varphi_2(z) = -Bz$,

---

[9]Existence conditions for ordinary differential equations are much easier to satisfy. See, e.g., Zachmanoglou and Thoe, 1986.

[10]$C^2(R)$ means that the functions are real valued, twice continuously differentiable.

$g_1(x) = \theta'x$, and $g_2(y) = \nu'y$. The model in our second example in section 3 is an element of this class in which M_1 and M_2 are the exponential functions and ψ_1 and ψ_2 are the identity functions.

With these restrictions, we can rewrite the consumer first order conditions as

$$\psi_1^{-1}\left(M_1^{-1}\left(P'(z)\right) - \varphi_1(z)\right) = g_1(x) - \varepsilon. \tag{14}$$

Similarly, we can rewrite the firms' maximization conditions as

$$\psi_2^{-1}\left(M_2^{-1}\left(P'(z)\right) - \varphi_2(z)\right) = \eta - g_2(y). \tag{15}$$

For ease of exposition we assume that M_1, M_2, ψ_1, and ψ_2 all equal the identity function, noting that the proofs presented in this section can easily be extended to treat the more general case for other known specifications of M_1, M_2, ψ_1, and ψ_2 [11]

We assume equilibrium prices so that equation (5) generates the hedonic price function. As in section 2, $f_x(x)$ is the density of x with support $\mathcal{X}$, f_ε is the density of ε with support $\mathcal{E}$, f_y is the density of y with support $\mathcal{Y}$, and f_η is the density of η with support $\mathcal{H}$. Assume $\mathcal{E} = \mathcal{H} = (-\infty, \infty)$. Also, x is independent of ε and y is independent of η. We assume that $(\varphi_1(z), \varphi_2(z)) \in C^2(R)$ and that both $g_1(x)$ and $g_2(y)$ are twice continuously differentiable in their arguments.

The first order conditions define mappings from (x, ε) to (x, z) and from (y, η) to (y, z) :

$$\varepsilon = \varphi_1(z) - P'(z) + g_1(x)$$

$$x = x$$

---

[11] These mappings cannot be identified. Thus (14) and (15) form an equivalence class of models which are indistinguishable.

and

$$\eta = P'(z) - \varphi_2(z) + g_2(y)$$

$$y = y.$$

The associated Jacobian terms are $dxd\varepsilon = [\varphi_1'(z) - P''(z)]dxdz$ and $dyd\eta = [P''(z) - \varphi_2'(z)]dydz$, respectively. >From the second order conditions, both terms in brackets are positive. Equilibrium condition (5) can be written as

$$\int_{\mathcal{X}} f_\varepsilon(\varphi_1(z) - P'(z) + g_1(x))(\varphi_1'(z) - P''(z))f_x(x)dx \qquad (16)$$

$$= \int_{\mathcal{Y}} f_\eta(P'(z) - \varphi_2(z) + g_2(y))(P''(z) - \varphi_2'(z))f_y(y)dy.$$

Initial conditions are provided by the requirements that $\Pi \geq 0$ and $U \geq \overline{U}$. We next state a genericity result for the hedonic model.

4.1   A Genericity Property for the Hedonic Model

A property $P(\theta)$, that characterizes parameter $\theta \in \Theta$, is called generic if the set $\Omega \subset \Theta$ of values of the parameter for which the property holds true contains a countable intersection of open dense subsets. In our context, θ indexes parameters that generate different specifications of a model. The property we are analyzing is identifiability of the model. If Θ is a complete metric space, such a set Ω will be dense in Θ, by Baire's theorem.[12] Moreover, the intersection of two such sets will also be dense in Θ. In other words, if a property is generic, and does not hold for a certain value $\bar{\theta}$ of the parameter, there will be in any neighborhood of $\bar{\theta}$ some other value θ of the parameter where the property holds true.

---

[12]See *e.g.* Royden (1968).

Genericity is a useful concept in our context because it provides us with a measure of the relative "density" of certain types of models, as the following discussion shows. Consider a model m. Typically, such a model depends on a set of parameters θ, the values of which are to be determined by the available data. Denoting by Θ the set of possible values for the parameters, define a map $\theta \rightarrow m(\theta)$ which associates with each $\theta \in \Theta$ the actual model to be used when the parameter values are θ. The parameter space Θ can be finite-dimensional, $\theta = (\theta_1, ..., \theta_n) \in R^n$. It can also be infinite-dimensional, for instance when θ is a function of a real variable, $\theta = \theta(x), \; x \in R$.

The most interesting properties of a model are those which are always true, that is, which hold for all versions of the model irrespective of the actual values of the parameters. Such a property is characterized formally by introducing the following set:

$$A(P) = \{\theta \mid P(\theta) \text{ is true}\}$$

and writing

$$A(P) = \Theta,$$

the entire set of all possible parameters. Such properties are rare. When $A(P) \neq \Theta$, it is natural to ask what kind of properties are true "in general" , where violations are "exceptional." For instance, we think that a "general" real number is irrational, and a "general" function is nonlinear. However, it is not so easy to give a precise mathematical content to this intuition.

One approach to making these intuitive ideas precise appeals to Bayesian notions. In this approach, one endows the parameter space Θ with a positive measure μ, and says that a property P holds "in general" if it holds μ-almost surely, that is, if:

$$\mu\left[\Theta - A(P)\right] = 0.$$

With this definition, taking Θ to be the real line, μ to be the Lebesgue measure, and $P(\theta)$ to be the property "θ is irrational", we will state (correctly) that real numbers are irrational in general.

The problem with this approach lies in choosing the measure μ. In the absence of a priori information, the Lebesgue measure may seem an appropriate starting point, because it is translation-invariant. Unfortunately, it is not a probability, so that it will not fit easily into a Bayesian framework. One may choose to overlook this problem and point out that the measure μ is not important, only the μ-negligible subsets are. One then runs into a second problem, namely the fact that there is no equivalent of Lebesgue measure in infinite-dimensional spaces such as function spaces. On such spaces, there is no translation-invariant measure that one could use to define negligible sets without a priori information.

If one does not ignore the problem that Lebesgue measure is not a probability, then one has to choose a probability. If Θ is finite-dimensional, one runs into the familiar problems of Bayesian theory. There is no reason why two different observers should share the same prior, or even why they should have priors which are absolutely continuous with respect to each other. If Θ is infinite-dimensional, there is the added difficulty that there are very few probabilities on such spaces; we have already noted that there is no equivalent of the Lebesgue measure. In fact, except for Dirac masses (point masses) and the like, the only known probabilities on spaces of continous functions are Gaussian, and derive from the Wiener measure (See Ito and MacKean, 1974 and Stroock, 1994 for comprehensive discussions). The Wiener measure is a very sophisticated mathematical tool, of great usefulness in other contexts, such as in the proper definition of Brownian motion, but there is no reason why it should serve as a universal standard to decide which functions are "general" and which are "exceptions".

There is a second approach to the problem of assessing the density of a property in a function space that we use in this paper, and that is due to Rene Thom (See Abraham and Robbin, 1967, and Aubin and Ekeland, 1984 for discussions of this approach). It consists in endowing the parameter space Θ, not with a positive measure, but with a complete metric.[13] This approach defines a property P to be *generic* if there is a sequence of subsets $U_n \subset \Theta$ such that

$$U_n \text{ is open and dense } \forall n$$

$$A(P) \supset \cap_n U_n$$

and defines a property as holding "in general" if it is generic.

This approach works nicely in our first example. The set of rational numbers is countable, so the set of irrational numbers is the intersection of a countable number of open dense sets. Let $\rho_n, n \in N$, be the rationals, then the set of irrationals will be $\cap_n U_n$, where U_n is the complement of $\{\rho_n\}$. This approach also works in our second example. Denote the space of continuous functions on $[0, 1]$ by $C^0([0, 1])$ and let $P(f)$ be true if f is affine. The subset of affine functions is

$$A_f(P) = \left\{ f \in C^0[0, 1] \,\middle|\, P(f) \text{ is true} \right\}.$$

$A_f(P)$ is closed and has empty interior so that its complement is an open dense subset. Thus, nonlinearity is a generic property in $C^0([0, 1])$.

If two properties P_1 and P_2 are generic, so is $P_1 \wedge P_2$. As a consequence, if P is generic, then its negation $\sim P$ cannot be generic. More generally, if a sequence of properties P_n are all generic, then so is $\wedge_n P_n$. In other words, generic properties behave in the same way as

---

[13]Meaning that all Cauchy sequences converge.

properties that are true almost surely, although there is no underlying measure to support them.

Thom's approach has the great advantage that in many cases of interest there is a natural metric on the parameter space Θ, even if there is no natural probability measure on it. So people will agree on what is generic or not, although they do not agree on which sets have measure zero. We now apply the tool of genericity to the analysis of hedonic models.

We first need a technical assumption to avoid some integrability problems. Assume there are some continuous probability densities h_x and h_y with $h_x > 0$ on $\mathcal{X}$ and $h_y > 0$ on $\mathcal{Y}$ such that f_x/h_x and f_y/h_y are bounded away from zero and infinity on $\mathcal{X}$ and $\mathcal{Y}$. Note that this will always be the case if $\mathcal{X}$ and $\mathcal{Y}$ are compact intervals. Set $f_x^0 = f_x/h_x$ and $f_y^0 = f_y/h_y$.

The "parameters" of our model are the functions: (φ_1, φ_2), (g_1, g_2), (f_ε, f_η), and $\left(f_x^0, f_y^0\right)$ .We have the following:

Theorem 1 *Generically with respect to any of the parameter pairs, the equilibrium equations have no solution of the form $P'(z) = a_1 + b_1\varphi_1(z)$, nor any solution of the form $P'(z) = a_2 + b_2\varphi_2(z)$, where a_1, a_2, b_1, and b_2 are constants.*

Proof. See Appendix A. ∎

The precise definitions of the parameter spaces and their respective topologies are given in Appendix A, together with the proof of the theorem. This theorem can easily be modified to prove that generically, the equilibrium equations have no solution $P'(z)$ that can be expressed as a polynomial in (φ_1, φ_2).

The idea motivating the proof is simple. Equilibrium equation (5) (or (28) below in the

notation of this section) is an infinite set of equations; one equation for each value of z. If $P'(z) = a_1 + b_1\varphi_1(z)$ (or alternatively $P'(z) = a_2 + b_2\varphi_2(z)$), then this equation must be satisfied at every value of z when $a_1 + b_1\varphi_1(z)$ is substituted into (5) for each $\varphi_1(z)$ or $\varphi_2(z)$. But making the substitution gives an infinite set of equations in two variables, a_1 and b_1 or a_2 and b_2 if the other substitution is made. While it is certainly possible that two variables can simultaneously satisfy an infinite set of equations (the Tinbergen linear-quadratic-normal economy is one example), in general, one would expect that the set of parameters (φ_1, φ_2), (g_1, g_2), (f_ε, f_η), and (f_x^0, f_y^0) that yields such cases is highly unlikely. In fact the set of parameters that yields such solutions is negligible relative to the set of potential parameters. The proof of the theorem formally establishes this claim.[14]

As a consequence of this theorem, Brown-Rosen Point One that regressions of $P'(z)$ on $\varphi_1(z)$ or $\varphi_2(z)$ only recover the marginal price is not generically correct. More generally, the curvature in z of the pricing function is generally distinct from the curvature in z in the demand or supply equations, a result that might have been anticipated from (6). The hedonic model is intrinsically (generically) nonlinear. The examples presented at the end of the section 3 are prototypical, not special. There is no arbitrariness in assuming that $P'(z)$ and $\varphi_1(z)$ do not lie in the same linear space.

---

[14]Mas-Colell (1985) discusses and illustrates the transversal theorem (p. 42-45) and genericity (chapter 8).

4.2 Can We Identify the Parameters of the Model from a Single Market ?

Even if Point One is not generic, Point Two remains. Within a single market, there are no natural exclusion restrictions. The sorting condition rules out use of shifter variables from the other side of the market from data on matched pairs even if they are available. The larger question considered in this paper is whether we can identify $(\varphi_1, \varphi_2, g_1, g_2, f_\varepsilon, f_\eta)$ from data on $P(z)$, z, x, and y from a single market. We focus on identifying $(\varphi_1, g_1, f_\varepsilon)$ from data on $P(z)$, z, and x since the analysis is symmetric for (φ_2, g_2, f_η) using data on $P(z)$, z, and y. We later consider what information, if any, is available from the joint density of $(z, x, y, P(z))$.

We present two methods for recovering these functions from data in a single market. One is based on extensions of average derivative models (Powell, Stock and Stoker, 1989) and closely related transformation models (see Horowitz, 1998). We develop these methods in this section. The other is based on nonlinear instrumental variables. (Kelejian, 1971; Amemiya, 1975). We discuss the second method in section five.

The trick in applying average derivative and transformation models to the hedonic problem is to exploit the separability of z, x and ε. Define

$$T_1(z) = \varphi_1(z) - P'(z).$$

This function combines price and preference data. This kind of function is called a transformation function and its nonparametric identification and estimation have received extensive theoretical attention. (See Horowitz (1998) for a survey and new results). These models extend average derivative models (Powell, Stock and Stoker, 1989) by considering nonlinear transformations of dependent variables.

Let F_ε be the cumulative distribution function corresponding to density f_ε and let $F_{z|x}(z,x)$ be the empirical cumulative distribution function of z conditional on x. $F_{z|x}(z,x)$ can be estimated directly from data on the joint distribution of (z,x). For the purposes of our analysis, we assume it is known. Since we have assumed x is independent of ε and taking account of the first order condition (14), we may write

$$F_{z|x}(z,x) = F_\varepsilon(T_1(z) + g_1(x)). \tag{17}$$

Assuming that $\lim_{z \to \infty} T_1(z) = \infty$, which follows from the assumption that the support of $\varepsilon = (-\infty, \infty)$, and further assuming that T_1 is twice continuously differentiable, and g_1 has continuous cross partials of order n_x, where $T_1'(z) > 0$ from the second order conditions.

Moreover

$$\frac{\partial F_{z|x}(z,x)}{\partial x_i} = f_\varepsilon(T_1(z) + g_1(x)) \cdot \frac{\partial g_1}{\partial x_i}. \tag{18}$$

>From (18), assuming $\dfrac{\partial g_1(x)}{\partial x_j} \neq 0$ for all j,

$$\frac{\frac{\partial F_{z|x}(z,x)}{\partial x_i}}{\frac{\partial F_{z|x}(z,x)}{\partial x_j}} = \frac{\frac{\partial g_1(x)}{\partial x_i}}{\frac{\partial g_1(x)}{\partial x_j}} \quad \text{for all } i,j.$$

This ratio determines the level sets of $g_1(x)$. More generally, taking the ratio of $(\textbf{??})$ to (18) for an arbitrary argument i, we obtain

$$\left(\frac{\frac{\partial F_{z|x}(z,x)}{\partial z}}{\frac{\partial F_{z|x}(z,x)}{\partial x_i}} \right) = \frac{T_1'(z)}{\frac{\partial g_1(x)}{\partial x_i}}. \tag{19}$$

>From (18), sign $(\frac{\partial F_{z|x}(z,x)}{\partial x_i})$ = sign $\left(\dfrac{\partial g_1}{\partial x_i} \right)$. Assume, without loss of generality, that

32

$$\frac{\partial g_1}{\partial x_i} > 0.^{15}$$

Then the left hand side of (19) is positive and (recall that $T_1'(z) > 0$ from second order conditions)

$$\frac{\partial}{\partial z} \log \left[\frac{\frac{\partial F_{z|x}(z,x)}{\partial z}}{\frac{\partial F_{z|x}(z,x)}{\partial x_i}} \right] = \frac{T_1''(z)}{T_1'(z)}. \tag{20}$$

Define $h_i(z,x) = \log \left[\frac{\frac{\partial F_{z|x}(z,x)}{\partial z}}{\frac{\partial F_{z|x}(z,x)}{\partial x_i}} \right]$. Since $h_i(z,x)$ satisfies equation (20), then $h_i(z,x)$ must be of the form

$$h_i(z,x) = h_{0i} + h_1(z) + h_{2i}(x)$$

where h_{0i} is a constant and without loss of generality we make the normalizations that $h_1(0) = 0$ and $h_{2i}(0) = 0$. h_{0i}, $h_1(z)$, and $h_{2i}(x)$ are known empirically. Further, equation (20) can be written as

$$\frac{dh_1(z)}{dz} = \frac{T_1''(z)}{T_1'(z)}.$$

This equation has the solution

$$T_1'(z) = K_1 \exp(h_1(z)) \tag{21}$$

where K_1 is a constant of integration which must be positive since the second order conditions imply that $T_1'(z) > 0$. This implies that

$$T_1(z) = C_1 + K_1 \int_0^z \exp(h_1(z'))dz'$$

where C_1 is a new constant of integration. Thus, from data on the joint distribution of (z,x) alone, we can identify the function $T_1(z)$ up to two constants.

---

[15] If $\frac{\partial g_1}{\partial x_i}$ is not strictly positive, then the arguments below can be carried out separately on each region where it does not change sign.

33

This solution enables us to solve for $g_1(x)$. Substituting (21) into (19),

$$\frac{\partial g_1}{\partial x_i} \exp(h_{0i} + h_1(z) + h_{2i}(x)) = K_1 \exp(h_1(z)), \quad i = 1, ..., n_x$$

$$\frac{\partial g_1(x)}{\partial x_i} = K_1 \exp(-h_{0i} - h_{2i}(x)), \quad i = 1, ..., n_x \qquad (22)$$

Thus, the partial derivative of g_1 with respect to each x_i $i = 1, ..., n_x$ is identified. Further, this defines $g_1(x)$ as the solution of a set of partial differential equations. The solution of this set of equations is

$$g_1(x) = R_1 + K_1 \widetilde{g}_1(x)$$

where R_1 is a constant of integration and

$$\widetilde{g}_1(x_1, ..., x_{n_x}) = \qquad (23)$$

$$\sum_{i=1}^{n_x} \int_{-\infty}^{x_i} \exp(-h_{0i} - h_{2i}(x_1, ..., x_i', ..., x_{n_x})) dx_i'$$

$$+ \sum_{k=2}^{n_x} (-1)^{k-1} \times \left(\sum_{i_1=1}^{(1+n_x-k)} \sum_{i_2=i_1+1}^{(2+n_x-k)} \cdots \sum_{i_k=i_{k-1}+1}^{n_x} \left(\int_{-\infty}^{x_{i_1}} \cdots \int_{-\infty}^{x_{i_k}} \frac{\partial \exp\left(-h_{0i_1} - h_{2i_1}\left(x_1, ..., x_{i_1}', ..., x_{i_k}', ..., x_n\right)\right)}{\partial x_{i_2} \cdots \partial x_{i_k}} \times \right. \right.$$
$$\left. \left. dx_{i_1}' \cdots dx_{i_k}' \right) \right).$$

For the case where $n_x = \dim(x) = 1$, this implies

$$g_1(x) = R_1 + K_1 \int_{-\infty}^{x} \exp(-h_0 - h_2(x')) dx'.$$

For the case where $n_x = 2$,

$$g_1(x_1, x_2) = R_1 + K_1 \sum_{i=1}^{2} \int_{-\infty}^{x_i} \exp(-h_{0i} - h_{2i}(x_i', x_{-i})) dx_i' -$$

$$K_1 \int_{-\infty}^{x_2} \int_{-\infty}^{x_1} \frac{\partial \exp(-h_{01} - h_{21}(x_1', x_2'))}{\partial x_2} dx_1' dx_2'.$$

34

Thus we can identify

$$\widetilde{g}_1(x) = \frac{g_1(x) - R_1}{K_1}$$

and

$$\tilde{T}_1(z) = \frac{T_1(z) - C_1}{K_1} = \int_0^z \exp(h_1(z'))dz' \qquad (24)$$

where $\widetilde{g}_1(x)$ is given in (23) and the right hand sides of (23) and (24) are constructed entirely from the data.

In this notation

$$\varepsilon = T_1(z) + g_1(x) = (C_1 + R_1) + K_1(\tilde{T}_1(z) + \widetilde{g}_1(x)).$$

If we substitute this into (17) we obtain

$$F_{z|x}(z, x) = F_\varepsilon(C_1 + R_1 + K_1(\tilde{T}_1(z) + \widetilde{g}_1(x))). \qquad (25)$$

This identifies the distribution function F_ε up to the unknown parameters C_1, R_1, and K_1. Using a normalization such as setting the median or the mean of ε to 0, we can identify the constant $C_1 + R_1$ up to scale K_1. Finally, since $P'(z)$ can be estimated nonparametrically from data on $(z, P(z))$, for the purposes of our identification argument, we can treat it as known. Hence, we can identify $\varphi_1(z) = C_1 + K_1\tilde{T}_1(z) + P'(z)$ up to the unknown constants C_1 and K_1.

Separability plays an important role in this proof. Because of separability $h_i(z, x)$ is separable in z and x meaning that one more derivative (with respect to z or x) depends only on z or x, and we can integrate back to recover $\varphi_i(z)$ and $g_i(x)$ each up to two constants (a location and scale parameter). Recall that $h_i(z, x)$ is itself formed from the ratio of derivatives.

In summary, the joint distribution of $(z, x, P(z))$ in conjunction with the additive structure contains enough information to identify the structure of preferences and the distribution of preferences in the population up to the three constants C_1, R_1, and K_1. An algorithm for implementing this procedure when $n_x = 1$ is as follows. First estimate $\widehat{P}(z)$, $\widehat{P}'(z)$, and $\widehat{F}_{z|x}(z, x)$ non-parametrically. Then calculate

$$
\begin{aligned}
\widehat{h}_i(z, x) &= \widehat{h}_{0i} + \widehat{h}_1(z) + \widehat{h}_{2i}(x) \\
&= \log\left(\frac{\frac{\partial \widehat{F}_{z|x}(z,x)}{\partial z}}{\frac{\partial \widehat{F}_{z|x}(z,x)}{\partial x_i}}\right).
\end{aligned}
$$

Use this to estimate $\widehat{\widetilde{T}}_1(z) = \int\limits_0^z \exp\left(\widehat{h}_1(z')\right) dz'$ and $\widehat{\widetilde{g}}_1(x)$ using equation (23). Then fix a value of K_1, say $K_1 = 1$. Also, fix R_1 since C_1 and K_1 are not separately identified. Then using (25), for each fixed C_1 we can trace out an estimator of F_ε by varying $\left(\widetilde{T}_1(z) + \widetilde{g}_1(x)\right)$. Doing this we can fix C_1 so that the median (or mean) of ε is zero. Finally, estimators of φ_1 and g_1 are $\widehat{\varphi}_1(z) = C_1 + K_1 \widehat{\widetilde{T}}_1(z) + \widehat{P}'(z)$ and $\widehat{g}_1(x) = R_1 + K_1 \widehat{\widetilde{g}}_1(x)$.

Analogous arguments on the demand side can be used to identify the structure of technology and the distribution of productivities using data on $(z, y, P(z))$. In particular, if we define

$$
T_2(z) = P'(z) - \varphi_2(z)
$$

we can identify φ_2, g_2, and f_η up to a set of constants so that $\varphi_2(z) = P'(z) - C_2 - K_2 \widetilde{T}_2(z)$, $g_2(y) = R_2 + K_2 \widetilde{g}_2(y)$, F_η satisfies an equation analogous to (25), and $\widetilde{T}_2(z)$ and $\widetilde{g}_2(y)$ are defined like $\widetilde{T}_1(z)$ and $\widetilde{g}_1(x)$ respectively. In these expressions, $K_2 > 0$ from the second order condition.

36

4.3   Using Output and Profit Information

The lack of identification of the scale of the utility function is a classical result. We do not observe utility so we can only identify level sets connected with utility. If we observe utility or output or profits on the production side, we can determine the missing parameters by using direct analysis of the utility or production or profit functions. Suppose profit is observed directly and that profit associated with the first order conditions (15) is

$$\Pi\left(z,y\right) = \Phi_2\left(z\right) + z\left(\eta - g_2\left(y\right)\right) - P\left(z\right) \tag{26}$$

where $\Phi_2\left(z\right) = \int_0^z \varphi_2\left(z'\right)dz'$. Direct estimation of (26) taking into account that characteristic level z is chosen by a firm with observable and unobservable characteristics y and η, entails identification of a correlated random coefficient model in a semiparametric setting.[16] Using (15) as a replacement function in the sense of Heckman and Robb (1985) or as a control function in the sense of Blundell and Powell (2001), we may solve for η and substitute in (26) to obtain

$$\Pi\left(z,y,\eta\right) = \Phi_2(z) + zP'\left(z\right) - z\varphi_2\left(z\right)$$

so

$$\psi(z) = \Pi(z,y,\eta) - zP'(z) = \int_0^z \varphi_2(z')dz' - z\varphi_2(z).$$

But, since profits and prices are observed, $\psi(z)$ is observed. We may estimate the derivative of the right hand side

$$\frac{\partial\psi(z)}{\partial z} = -z\varphi_2'(z).$$

---

[16] See Heckman and Vytlacil (1998) for a discussion of correlated random coefficient models.

Integrating up we obtain

$$C_0 + \int_0^z \left[-\frac{1}{z'} \frac{\partial \psi(z')}{\partial z} \right] dz' = \varphi_2(z)$$

so we determine $\varphi_2(z)$ up to an additive constant. Combining this with the information that $\varphi_2(z) = P'(z) - C_2 - K_2 \widetilde{T}_2(z)$ provides determination of K_2.

4.4   A General Parametric Approach

With additional (weak) parametric structure, we can determine the scaling constants without using the output data. Using only demand, supply and pricing data, we can stay within the Rosen program which does not contemplate using output data. We now assume that there is a finite-dimensional vector space V which contains both φ_1 and φ_2 and which is known *ex ante*. In other words, both φ_1 and φ_2 can be described by a finite set of parameters $(a_1, ..., a_K)$ and $(b_1, ..., b_K)$ that enter linearly. That is, $\varphi_1(z) = \sum_{k=0}^{M} a_k \overline{\varphi}_k(z)$ and $\varphi_2(z) = \sum_{k=0}^{M} b_k \overline{\varphi}_k(z)$ where $\overline{\varphi}_0(z) = 1$ and the $\overline{\varphi}_k(z)$, $k = 1, ..., M$, are known C^1 functions and M is known. For example, V could be the set of polynomials of degree less than or equal to M where M is a known integer.

To see how this restriction can be used to determine the scaling constants, consider the supply side of the market. Using the arguments above $P'(z)$ and $\widetilde{T}_1(z)$ are functions of the data alone. In large samples we can treat them as known. Then, using the definition

of $T_1(z)$ and the restriction $\varphi_1(z) = \sum_{k=0}^{M} a_k \overline{\varphi}_k(z)$ we can write

$$
\begin{aligned}
\widetilde{T}_1(z) &= -\frac{C_1}{K_1} - \frac{1}{K_1} P'(z) + \frac{1}{K_1} \sum_{k=0}^{M} a_k \overline{\varphi}_k(z) \\
&= \frac{a_0 - C_1}{K_1} - \frac{1}{K_1} P'(z) + \frac{1}{K_1} \sum_{k=1}^{M} a_k \overline{\varphi}_k(z).
\end{aligned}
\tag{27}
$$

This equation has $M+3$ unknown parameters, $(C_1, K_1, a_k, k = 0, ..., M)$. If we considier $M + 3$ values of z, we have $M+3$ equations. Can this system be inverted to solve for the parameters? C_1 and a_0 are not independently identified since $\overline{\varphi}_0(z) = 1$. If $P'(z)$ is an element of V, then the system of equations has rank $M+1$ since $P'(z)$ is then linearly dependent on the functions $\overline{\varphi}_k$. This is precisely the problem that worried Brown and Rosen (1982). However, if $P'(z)$ is not an element of V then $P'(z)$ is linearly independent of the functions $\overline{\varphi}_k$. In this case, the system of equations has rank $M+2$ and K_1 is identified. Theorem 2 proves that generically $P'(z)$ is not an element of V and thus the a_k, $k = 1, ..., M$ and K_1 are identified. Thus, the Brown and Rosen point is generically irrelevant within this flexible class of parametric models.

Theorem 2 *Generically with respect to any of the parameter pairs in Theorem 1, no solution P' of the equilibrium equation belongs to V, and φ_1, φ_2 are identified up to additive constants*

Proof. As shown above, we have:

$$
\widetilde{T}_1(z) = -\frac{P'(z)}{K_1} - \frac{C_1}{K_1} + \frac{\varphi_1(z)}{K_1}
$$

$$
\widetilde{T}_2(z) = \frac{P'(z)}{K_2} - \frac{C_2}{K_2} - \frac{\varphi_2(z)}{K_2}.
$$

39

Arguing as in Theorem 1, we can show that generically $P' \notin V$. As a consequence, there must be some continuous function f such that $\int_{\mathcal{Z}} f(z) h(z) dz = 0$ for all $h \in V$, but $\int_{\mathcal{Z}} f(z) P'(z) dz \neq 0$. Applying such a function to both sides of the preceding equations, we obtain:

$$\int_{\mathcal{Z}} \widetilde{T}_1(z) f(z) dz = \frac{-1}{K_1} \int_{\mathcal{Z}} P'(z) f(z) dz$$

$$\int_{\mathcal{Z}} \widetilde{T}_2(z) f(z) dz = \frac{1}{K_2} \int_{\mathcal{Z}} P'(z) f(z) dz$$

which determines K_1 and K_2. Plugging back into the equations, we find that φ_1 and φ_2 are determined up to the additive constants C_1 and C_2. ∎

This theorem suggests an estimation strategy for recovering K_1 (or K_2) when it is known that φ_1 and/or φ_2 belong to a finite-dimensional space. First estimate $P'(z)$ and $\widetilde{T}_1'(z)$ by nonparametric methods. Then estimate the sample analogue of equation (27). Thus, for a very general class of finite dimensional models including polynomial models, we obtain identification of the (φ_1, φ_2) functions from single market data.

Thus far we have considered identification only using data from one side of the market or the other. Sometimes, information on the equilibrium pairs is available. We next consider the information available in the joint densities.

4.5   Is There Identifying Information In The Joint Densities?

So far we have only considered identification using data from one side of the market. We have used information on the densities of (x, z) and (y, z) and have shown how to identify everything except K_1 and K_2. We obtain $\frac{C_i + R_I}{K_i}$ by a normalization of the errors to mean or median zero, but cannot separately identify $\frac{C_i}{K_i}$ and $\frac{R_i}{K_i}$. In the parametric case covered by

Theorem 2, we identify K_1 and K_2. There is one potentially powerful piece of information that we have not yet used; the joint distribution of (x, y, z). We now consider whether additional identifying information can be extracted from this joint density. This joint distribution may have identifying power because the distribution of z conditional on x is not the same as the distribution of z conditional on x and y. Where there is sorting on both sides of the market, this full joint density contains information that might be exploited.[17] We show that there is no more identifying information available beyond what is in the marginal densities.

Recall the first-order conditions from the previous section. On the worker side we have $\varepsilon = T_1(z) + g_1(x)$ and on the firm side we have $\eta = T_2(z) + g_2(y)$. The joint density of $(x, y, \varepsilon, \eta)$ is $f_x(x) f_y(y) f_\varepsilon(\varepsilon) f_\eta(\eta)$, since by assumption x, y, ε, and η are jointly independent. Note that this independence does not hold conditional on characteristic z. The hedonic equilibrium maps the joint distribution of $(x, y, \varepsilon, \eta)$ to the joint distribution of $(x, y, \varepsilon, \eta, z)$. This mapping does not change the distribution of $(x, y, \varepsilon, \eta)$. This distribution is exogenous.[18]

To derive the restrictions that equilibrium places on the observable data; i.e. the joint distribution of (x, y, z), note that the dimension of random vector $(x, y, \varepsilon, \eta)$ is $n_x + n_y + 2$ where n_x is the dimension of x, n_y is the dimension of y, and ε and η are each of dimension 1. The equilibrium maps this random vector into the observable random vector (x, y, z).

Suppose that we ignore equilibrium and assume that worker and firm choices are made independently of each other. Let z_1 be the choice of workers and z_2 the choice of firms.

---

[17]Epple (1987) discusses the potential importance of using the full joint density but his discussion is not complete.

[18]A more complete dynamic analysis would model how this marginal distribution changes over time.

Then the joint density without imposing equilibrium is

$$f_x\left(x\right)f_y\left(y\right)f_\varepsilon\left(T_1\left(z_1\right)+g_1\left(x\right)\right)f_\eta\left(T_2\left(z_2\right)+g_2\left(y\right)\right)T_1'\left(z_1\right)T_2'\left(z_2\right)$$

where $T_1'(z_1)$ and $T_2'(z_2)$ are the Jacobians of transformation for z_1 and z_2 respectively.

Imposing the equilibrium condition $(z_1 = z_2 = z)$ determines the density of (x,y) conditional on z :

$$
\begin{aligned}
& f_{xy|z}\left(x,y,z\right) \\
&= \frac{f_x\left(x\right)f_y\left(y\right)f_\varepsilon\left(T_1\left(z\right)+g_1\left(x\right)\right)f_\eta\left(T_2\left(z\right)+g_2\left(y\right)\right)T_1'\left(z\right)T_2'\left(z\right)}{\left(\int_{\mathcal{X}}f_x\left(x'\right)f_\varepsilon\left(T_1\left(z\right)+g_1\left(x'\right)\right)T_1'\left(z\right)dx'\right)\left(\int_{\mathcal{Y}}f_y\left(y'\right)f_\eta\left(T_2\left(z\right)+g_2\left(y'\right)\right)T_2'\left(z\right)dy'\right)} \\
&= \frac{f_x\left(x\right)f_y\left(y\right)f_\varepsilon\left(T_1\left(z\right)+g_1\left(x\right)\right)f_\eta\left(T_2\left(z\right)+g_2\left(y\right)\right)T_1'\left(z\right)T_2'\left(z\right)}{\left(f_z\left(z\right)\right)^2}
\end{aligned}
$$

where $f_z\left(z\right)$ is the marginal density of z and by the equilibrium condition (5) we have

$$
\begin{aligned}
f_z(z) &= \int_{\mathcal{X}}f_x\left(x'\right)f_\varepsilon\left(T_1\left(z\right)+g_1\left(x'\right)\right)T_1'\left(z\right)dx' \qquad (28) \\
&= \int_{\mathcal{Y}}f_y\left(y'\right)f_\eta\left(T_2\left(z\right)+g_2\left(y'\right)\right)T_2'\left(z\right)dy'.
\end{aligned}
$$

Multiplying through by the marginal density, the joint density of (x,y,z) is

$$f_{xyz}\left(x,y,z\right) = \frac{f_x\left(x\right)f_y\left(y\right)f_\varepsilon\left(T_1\left(z\right)+g_1\left(x\right)\right)f_\eta\left(T_2\left(z\right)+g_2\left(y\right)\right)T_1'\left(z\right)T_2'\left(z\right)}{f_z\left(z\right)}. \qquad (29)$$

In the appendix we prove that there is no more identifying information in the joint densities than in the marginal densities.

Theorem 3 *Joint density (29) provides no more identifying information than the marginal densities $f_{xz}\left(x,z\right)$ and $f_{yz}\left(y,z\right).$*

Proof. See Appendix A. ∎

Note however that these conditions can be used to improve the efficiency of estimation. Define $\widetilde{\varepsilon} = \frac{\varepsilon}{K_1}$ and $\widetilde{\eta} = \frac{\eta}{K_2}$ and define $f_{\widetilde{\varepsilon}}$ and $f_{\widetilde{\eta}}$ to be the respective densities of the scaled error terms as in the proof of Theorem 3. From the arguments in the previous section $f_{\widetilde{\varepsilon}}$ and $f_{\widetilde{\eta}}$ are identified. If we substitute all identified parameters into (28) we can write the equilibrium condition in terms of them. We obtain

$$\int_{\mathcal{X}} f_x\left(x'\right) f_{\widetilde{\varepsilon}}\left(\tau_1 + \widetilde{T}_1\left(z\right) + \widetilde{g}_1\left(x'\right)\right) \widetilde{T}_1'\left(z\right) dx' = \tag{30}$$

$$\int_{\mathcal{Y}} f_y\left(y'\right) f_{\widetilde{\eta}}\left(\tau_2 + \widetilde{T}_2\left(z\right) + \widetilde{g}_2\left(y'\right)\right) \widetilde{T}_2'\left(z\right) dy'$$

where as in the appendix we define $\tau_1 = \frac{C_1 + R_1}{K_1}$ and $\tau_2 = \frac{C_2 + R_2}{K_2}$. In empirical applications, each of the parameters denoted by "$\sim$" can be estimated and this equation establishes a functional relationship between the supply and demand side parameter estimates. This relationship can be used to improve the efficiency of estimation. Since the slope of the equilibrium price $P'\left(z\right)$ generally must also be estimated from data on $\left(z, P\left(z\right)\right)$, a similar argument can be used to improve efficiency in estimating it since its estimate must "fit" the data and satisfy the equilibrium equation (28). We develop these efficiency gains in another paper.

5 Instrumental Variables

Theorems 1 and 2 and the arguments in Section 4 show how separability and the independence between x and ε can be used to identify the structure of the hedonic model. The nonlinearity inherent in the model and weak parametric structure were then used

to recover the parameter K_1. Much of the argument does not rely on the independence assumption. In this section, we relax the independence assumption and show that when $E(\varepsilon|x) = 0$ and we impose general parametric structure on φ_1 and g_1, then the parameters of the model can be identified by the method of instrumental variables. Nonlinearity in the equilibrium hedonic model is crucial for this identification result. Theorem 5 below justifies the application of instrumental variables for general parametric versions of model (14). Instrumental variables are generically valid. The exclusion restrictions we exploit to show this identification do not use variables from the other side of the market which are endogenous because of the sorting condition. Rather, the exclusion restrictions we exploit arise naturally from the parametric structure of the model and the nonlinearity emerging from the equilibrium pricing equation.

We analyze the supply side first order condition

$$P'(z) = \varphi_1(z) + g_1(x) - \varepsilon \tag{31}$$

under the conditions stated in section 4 but with $(x, \varepsilon) \sim f_{x\varepsilon}(x, \varepsilon)$ where $f_{x\varepsilon}$ is a strictly positive density. We assume $E(\varepsilon|x) = 0$ and $E(g_1^2(x)) < \infty$. We also assume that $(y, \eta) \sim f_{y\eta}(y, \eta)$.

The literature reviewed in Section 3 establishes that in a single market setting there are no exclusion restrictions that justify the use of demand side variables (y) as instrumental variables in estimating this equation. Variables from the demand side of the market are stochastically dependent on ε given z. However, instruments for $\varphi_1(z)$ are still available. If $E(\varphi_1(z)|x)$ is not collinear with $g_1(x)$, then it is possible to construct $E(\varphi_1(z)|x)$ and use it as an instrument for $\varphi_1(z)$ in (31). Kahn and Lang (1988) make this point by way of an example for a particular functional form. In this section we establish that generically

$E\left(\varphi_1\left(z\right)|x\right)$ is a valid instrument for any arbitrary parametric functional form that satisfies the conditions required to prove Theorem 5. This result highlights the main themes of our paper: that the hedonic model is intrinsically nonlinear, that nonlinearity is an important source of identifying information and that intuitions developed in linear econometrics when applied to a nonlinear model are misleading. We can use our result to justify the choice of parametric nonlinear IV as in Amemiya (1975).

Theorem 4 *Generically with respect to any of the parameter pairs* $\left(\varphi_1, \varphi_2\right)$, $\left(g_1, g_2\right)$, *and* $\left(f_{x\varepsilon}, f_{y\eta}\right)$, *the equilibrium equations have no solution of the form* $P'\left(z\right) = a_1 + b_1\varphi_1\left(z\right)$, *nor any solution of the form* $P'\left(z\right) = a_2 + b_2\varphi_2\left(z\right)$ *where* a_1, a_2, b_1, *and* b_2 *are constants.*

Proof. See Appendix A. ∎

This theorem extends Theorem 1 to the case where ε is not independent of x. Using this result we can show the following.

Theorem 5 *Generically with respect to any parameter pairs in Theorem 4,* $E\left(\varphi_1\left(z\right)|x\right)$ *cannot be collinear with* $g_1\left(x\right)$.

Proof. See Appendix A. ∎

As a consequence of this theorem, we can use $E\left(\varphi_1\left(z\right)|x\right)$ as an instrument for $\varphi_1(z)$ using parametric nonlinear IV (Amemiya, 1975). As an example, consider the case where it is known that $\varphi_1\left(z\right) = \varphi_0 + \varphi_1 z + \varphi_2 z^2$ and $g_1\left(x\right) = g_1\left(x, \theta\right)$ where θ is a vector of parameters. In this case, first estimate $\widehat{P}'\left(z\right)$. Then estimate the regressions $E\left(z|x\right)$ and $E\left(z^2|x\right)$. Finally estimate

$$\widehat{P}'\left(z\right) = \varphi_0 + \varphi_1 z + \varphi_2 z^2 + g_1\left(x, \theta\right) - \varepsilon \tag{32}$$

using $E\left(z\left|x\right.\right)$ and $E\left(z^2\left|x\right.\right)$ as instruments for z and z^2. While these instruments are functions of x, Theorem 5 guarantees that they are linearly independent of $g_1\left(x,\theta\right)$.

More generally, using nonlinear IV we can estimate the parameters generating $\varphi_1(z)$ and the parameters generating $g_1\left(x,\theta\right)$. The separability implies that we can vary $E\left(\varphi_1\left(z\right)\left|x\right.\right)$ separately from $g_1(x)$ and hence identify both.

We conjecture that this condition also justifies the application of nonparametric IV (Darolles et. al, 2001, Florens, Heckman, Meghir and Vytlacil, 2000, or Newey and Powell, 2000) when these parametric restrictions are not imposed. However, as currently formulated those papers require an exclusion restriction which is not intrinsic to the model and it is necessary to extend their arguments to impose Theorem 5 as an identifying condition in the estimation. This is a task we leave for the future.

6   Summary, Conclusions and Proposed Extensions

This paper considers identification and estimation of technology and preference parameters using data on choices made in a single hedonic market. The general hedonic problem is formulated, a normal-linear-quadratic version of the model is developed and its advantages and peculiarities are exposed.

Standard criticisms directed against Sherwin Rosen's two stage estimation procedure for hedonic models are shown to be misleading. Generically, a separable nonparametric version of the linear-quadratic-normal model is identified up to levels. When it is not identified, the pricing function alone identifies technology or preference parameters. With mild functional form assumptions, the model is completely identified in both the case where ε is independent of x and under the weaker assumption that $E\left(\varepsilon\left|x\right.\right)=0$. Two estimation

procedures are presented: (a) nonparametric transformation methods, and (b) IV in a general nonlinear (but parametric) setting.

The analysis developed here applies to closely related problems of estimating preferences and technology when taxes are set optimally (Mirrlees, 1971 and 1986), when monopolists price discriminate (Mussa and Rosen, 1978; Wilson, 1993) and for the standard problem of taxes and labor supply (Heckman, 1974; Hausman 1980) when tax schedules are nonlinear and continuous.

Our presentation of the hedonic model is for the vector case. Yet our basic proofs are only for the scalar case. Extensions for the scalar nonseparable and the vector cases are underway in joint work with Rosa Matzkin. That work also considers the case of identification for a nonseparable hedonic model with vector attributes when some of the attributes are missing (Heckman, Matzkin and Nesheim, 2002).

References

[1] Abraham, Ralph and Joel Robbin, *Transversal Mappings and Flows*, New York: Benjamin, 1967.

[2] Amemiya, Takeshi. "The Nonlinear Limited Information Maximum Likelihood Estimator and the Modified Nonlinear Two Stage Least Squares Estimator," *Journal of Econometrics* 31 (1975): 375-386.

[3] Aubin, Jean-Pierre and Ivar Ekeland. *Applied Nonlinear Analysis*, New York: John Wiley and Sons, 1984.

[4] Bartik, Timothy. "The Estimation of Demand Parameters in Hedonic Price Models," *Journal of Political Economy* 95 (1987): 81-88.

[5] Blundell, Richard and James Powell. "Endogeneity of Single Index Models," unpublished manuscript, University College London, 2001.

[6] Brown, James and Harvey Rosen. "On the Estimation of Structural Hedonic Price Models," *Econometrica* 50 (1982): 765-769.

[7] Brown, James. "Structural Estimation In Implicit Markets," in *The Measurement of Labor Cost*, edited by J. Triplett. Chicago: University of Press for NBER, 1983.

[8] Darolles, Serge, Jean-Pierre Florens and Eric Renault. "Nonparametric Instrumental Variable Regression," unpublished manuscript, University of Toulouse, (2001), June 14.

[9] Epple, Dennis. "Hedonic Prices and Implicit Markets: Estimating Demand and Supply Functions for Differentiated Products." *Journal of Political Economy* 95 (1987): 59-80.

[10] Epple, Dennis and Holger Sieg. "Estimating Equilibrium Models of Local Jurisdictions," *Journal of Political Economy* 107 (1999): 645-681.

[11] Fisher, Franklin. *The Identification Problem in Economics*, New York: McGraw Hill, 1966.

[12] Florens, Jean-Pierre, James Heckman, Costas Meghir, Edward Vytlacil. "Instrumental Variables, Local Instrumental Variables and Control Functions," unpublished manuscript, University of Toulouse, 2000.

[13] Hausman, Jerry. "The Effects of Wages, Taxes, and Fixed Costs on Women's Labor Force Participation." *Journal of Public Economics* 14 (1980): 161-194.

[14] Heckman, James. "The Effect of Day Care Programs on Women's Work Effort," *Journal of Political Economy* (1974). Reprinted in T. W. Schultz (ed), *Economics of the Family: Marriage, Children, and Human Capital,* (University of Chicago Press, 1974), 491-517.

[15] Heckman, James. "Notes on Hedonic Models," unpublished lecture notes on Hedonic Models. University of Chicago, Department of Economics, revised from 1991 and 1995 versions, 1999.

[16] Heckman, James, Rosa Matzkin and Lars Nesheim. "Identification and Estimation of Hedonic Models: the Vector Nonseparable Case with Missing Attributes", under preparation, University of Chicago, November 2002.

[17] Heckman, James and Edward Vytlacil. "Instrumental Variables Methods for the Correlated Random Coefficient Model: Estimating the Average Rate of Return to Schooling When The Return is Correlated with Schooling," *Journal of Human Resources* 33 (1998): 974-987.

[18] Heckman, James and Robb, Richard. "Alternative Methods for Evaluating the Impact of Interventions." In *Longitudinal Analysis of Labor Market Data*, edited by James Heckman and Burton Singer. Cambridge University Press for Econometric Society Monograph Series, New York, NY, 1985.

[19] Horowitz, Joel. *Semiparametric Methods in Econometrics*, New York and Berlin: Springer-Verlag, 1998.

[20] Ito, K. and H.P. MacKean, *Diffusion Processes and Their Sample Paths*, Berlin: Springer-Verlag, 1974.

[21] Kahn, Shulamit and Kevin Lang. "Efficient Estimation of Structural Hedonic Systems," *International Economic Review* 29 (1988): 157-166.

[22] Kelejian, Harry. "Two Stage Least Squares and Econometric Models Linear in the Parameters but Nonlinear in the Endogenous Variables," *Journal of the American Statistical Association*, 1971.

[23] Mas-Colell, Andreu. *The Theory of General Economic Equilibrium: A Differentiable Approach*, Cambridge, England: Cambridge University Press, 1985.

[24] Mirrlees, James. "The Theory of Optimal Taxation," in K. J. Arrow and M.D. Intriligator, eds., *Handbook of Mathematical Economics*, Vol. III, New York: Elsevier, Chapter 24; 1197-1249, 1991.

[25] Newey, Whitney and James Powell, "Instrumental Variables For Nonparametric Models," unpublished manuscript, MIT, (2000).

[26] Nesheim, Lars. "Equilibrium Sorting of Heterogeneous Consumers Across Locations: Theory and Empirical Implications," Ph.D. dissertation, University of Chicago, (2001).

[27] Powell, James, James Stock and Thomas Stoker. "Semiparametric Estimation of Index Coefficients," *Econometrica* 57 (1989): 1403-1430.

[28] Rosen, Sherwin. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *Journal of Political Economy* 82 (1974): 34-55.

[29] Royden, H. L., *Real Analysis*, New York: Macmillan (1968).

[30] Stroock, Daniel, *Probability Theory, An Analytic View*, New York: Cambridge University Press, 1994.

[31] Tauchen, Helen and Anne Witte. "Estimating Hedonic Models: Implications of the Theory," unpublished manuscript, University of North Carolina Department of Economics, (April 2001).

[32] Tinbergen, Jan. "On The Theory of Income Distribution," *Weltwirtschaftliches Archiv* 77 (1956): 155-73.

[33] Wilson, Robert, *Nonlinear Pricing* Oxford,: Oxford University Press (1993)

[34] Zachmanoglou, E. C. and Dale Thoe, *Introduction to Partial Differential Equations.* New York: Dover, 1986.

Appendix A   Proofs

Recall that we have denoted by $\mathcal{X}$ and $\mathcal{Y}$ the supports of f_x and f_y, so that we may assume that $x \in \mathcal{X}$ and $y \in \mathcal{Y}$. Denote by $\mathcal{Z}$ the domain of z, so that $z \in \mathcal{Z}$; both φ_1 and φ_2 map $\mathcal{Z}$ into R. For the sake of simplicity, it will be assumed that $\mathcal{Z}$ is an interval, possibly unbounded.

For a rectangle $\mathcal{A} \subseteq R^n$, denote by $C^1(\mathcal{A})$ the space of continuously differentiable[19] functions on $\mathcal{A}$ endowed with the following topology: $\varphi_n \to \varphi$ iff φ_n converges to φ and the derivatives φ'_n converge to φ', uniformly on all compact subsets of $\mathcal{A}$. It is known that this topology turns C^1 into a complete metric space.[20] $C^1(\mathcal{Z})$ is the natural space for φ_1 and φ_2. $C^1(\mathcal{X})$ and $C^1(\mathcal{Y})$ are the natural spaces for g_1 and g_2 respectively.

For an interval $\mathcal{A} \subseteq R$, denote by $C_1^2(\mathcal{A})$ the space of twice differentiable functions f on $\mathcal{A}$, satisfying $\int_\mathcal{A} f = 1$ and $f > 0$ everywhere on $\mathcal{A}$, with f, f', and f'' continuous and uniformly bounded. It is endowed with the topology of uniform convergence of f, f', f'' which turns it into a complete metric space. $C_1^2(\mathcal{E})$ and $C_1^2(\mathcal{H})$ are the natural spaces for f_ε and f_η respectively.

Denote by $C_1^0(\mathcal{X})$ the space of continuous functions f such that $\int_\mathcal{X} f h_x = 1$ and $f > 0$ everywhere on $\mathcal{X}$, endowed with the uniform norm, which turns it into a complete metric space. This is the natural space for f_x^0 where we recall that $f_x^0 = f_x / h_x$.

Finally, denote by $C_1^0(\mathcal{Y})$ the space of continuous functions f such that $\int_\mathcal{Y} f h_y = 1$ and $f > 0$ everywhere on $\mathcal{Y}$, endowed with the uniform norm, which turns it into a complete metric space. This is the natural space for f_y^0 where we recall that $f_y^0 = f_y / h_y$.

---

[19]If $z_0 \in \mathcal{Z}$ is the left (or right) extremity of $\mathcal{Z}$, a derivative at z_0 will be understood to mean a right (or left) derivative.

[20]And, even a Banach space if $\mathcal{Z}$ is compact.

We now restate Theorem 1 more precisely:

Theorem 1 Restated *Generically with respect to any of the parameters pairs $(\varphi_1, \varphi_2) \in$
$C^1(\mathcal{Z}) \times C^1(\mathcal{Z}), (f_\varepsilon, f_\eta) \in C_1^2(\mathcal{E}) \times C_1^2(\mathcal{H}), (g_1, g_2) \in C^1(\mathcal{X}) \times C^1(\mathcal{Y}), (f_x^0, f_y^0) \in C_1^0(\mathcal{X}) \times$
$C_1^0(\mathcal{Y})$ the equilibrium equations have no solution of the form $P'(z) = a_1 + b_1\varphi_1(z)$, nor
any solution of the form $P'(z) = a_2 + b_2\varphi_2(z)$, where a_1, a_2, b_1, and b_2 are constants.*

Proof of Theorem 1:

Set $(\varphi_1, \varphi_2, f_\varepsilon, f_\eta, g_1, g_2, f_x^0, f_y^0) = \theta$ and $C^1(\mathcal{Z}) \times C^1(\mathcal{Z}) \times C_1^2(\mathcal{E}) \times C_1^2(\mathcal{H}) \times C^1(\mathcal{X}) \times$
$C^1(\mathcal{Y}) \times C_1^0(\mathcal{X}) \times C_1^0(\mathcal{Y}) = \Theta$.

Define a map $\Phi : \Theta \times R^2 \to C^0(R)$ derived from equilibrium condition (16) by substituting $P'(z) = a_1 + b_1\varphi_1(z)$

$$
\begin{aligned}
\Phi(\theta, a_1, b_1)(z) &= (1 - b_1)\varphi_1'(z)\int_{\mathcal{X}} f_\varepsilon(-a_1 + (1 - b_1)\varphi_1(z) + g_1(x))f_x^0(x)h_x(x)\,dx \\
&\quad - (b_1\varphi_1'(z) - \varphi_2'(z))\int_{\mathcal{Y}} f_\eta(a_1 + b_1\varphi_1(z) - \varphi_2(z) + g_2(y))f_y^0(y)h_y(y)\,dy.
\end{aligned}
$$

Saying that the equilibrium equation has a solution of the form $P'(z) = a_1 + b_1\varphi_1(z)$ means
that there is a pair $(a_1, b_1) \in R^2$ such that $\Phi(\theta, a_1, b_1) = 0$ for every z. We want to show
that generically in θ this cannot happen.

To do so, fix three points z_1, z_2, and z_3 in $\mathcal{Z}$, pairwise distinct, and define a map
$\Psi : \Theta \times R^2 \to R^3$ by:

$$\Psi(\theta, a_1, b_1) = (\Phi(\theta, a_1, b_1)(z_i))_{1 \leq i \leq 3}.$$

We will show that the map Ψ is C^1 and that its derivative $D\Psi$, which is a linear map from
$\Theta \times R^2$ into R^3, is onto. We will then apply Thom's transversality theorem which states
that if $\Psi = 0$ implies that $D\Psi$ is onto, then generically in θ, $\Psi(\theta, a_1, b_1)$ is transversal to

54

the origin. This means that generically in θ either

$$\gneqq (\theta, a_1, b_1) \neq 0$$

or

$$[\Psi (\theta, a_1, b_1) = 0] \implies \left[\left(\frac{\partial \Psi}{\partial a_1} (\theta, a_1, b_1), \frac{\partial \Psi}{\partial b_1} (\theta, a_1, b_1) \right) \text{ is onto} \right]$$

But the linear map $\left(\frac{\partial \Psi}{\partial a_1}, \frac{\partial \Psi}{\partial b_1} \right)$ can never be onto, because it sends a two dimensional space into a three dimensional one. It follows that $\Psi (\theta, a_1, b_1) \neq 0$ generically.

Lemma: The map Ψ is C^1.

Proof: The Gateaux derivative $D\Psi$ of Ψ at (θ, a_1, b_1) is easily expressed. Set $\delta\theta = \left(\delta\varphi_1, \delta\varphi_2, \delta f_\varepsilon, \delta f_\eta, \delta g_1, \delta g_2, \delta f_x^0, \delta f_y^0 \right)$, where the components of $\delta\theta$ belong to the appropriate vector spaces, $\delta f_\varepsilon, \delta f_\eta, \delta f_x^0, \delta f_y^0$ being subject to the additional requirement of integrating to zero. Similarly, set $(\delta a_1, \delta b_1) \in R^2$, and compute the first variation of Ψ:

$$D\Psi (\delta\theta, \delta a_1, \delta b_1) = \tag{A-1}$$

$$\left[\begin{array}{c} D\Psi (\delta\varphi_1) + D\Psi (\delta\varphi_2) + D\Psi (\delta f_\varepsilon) + D\Psi (\delta f_\eta) + \\ D\Psi (\delta g_1) + D\Psi (\delta g_2) + D\Psi (\delta f_x^0) + D\Psi (\delta f_y^0) + D\Psi (\delta a_1) + D\Psi (\delta b_1) \end{array} \right]_{i=1,2,3}$$

where we use the notation

$$D\Psi (\delta\varphi_1) =$$

$$\left[\begin{array}{l} \delta\varphi_1'(z_i) \left(\begin{array}{l} (1 - b_1) \int\limits_{\mathcal{X}} f_\varepsilon(-a_1 + (1 - b_1)\varphi_1(z_i) + g_1(x))f_x^0(x)h_x(x)\, dx \\ -b_1 \int\limits_{\mathcal{Y}} f_\eta(a_1 + b_1\varphi_1(z_i) - \varphi_2(z_i) + g_2(y))f_y^0(y)h_y(y)\, dy \end{array} \right) + \\ \delta\varphi_1(z_i) \left(\begin{array}{l} (1 - b_1)^2 \varphi_1'(z_i) \int\limits_{\mathcal{X}} f_\varepsilon'(-a_1 + (1 - b_1)\varphi_1(z_i) + g_1(x))f_x^0(x)h_x(x)\, dx \\ -b_1^2 \varphi_1'(z_i) \int\limits_{\mathcal{Y}} f_\eta'(a_1 + b_1\varphi_1(z_i) - \varphi_2(z_i) + g_2(y)) f_y^0(y) h_y(y)\, dy \\ +b_1\varphi_2'(z_i) \int\limits_{\mathcal{Y}} f_\eta'(a_1 + b_1\varphi_1(z_i) - \varphi_2(z_i) + g_2(y)) f_y^0(y) h_y(y)\, dy \end{array} \right) \end{array} \right]_{i=1,2,3}$$

$$D\Psi\left(\delta\varphi_2\right) =$$

$$\left[\begin{array}{c} \delta\varphi_2'(z_i) \int\limits_{\mathcal{Y}} f_\eta(a_1 + b_1\varphi_1(z_i) - \varphi_2(z_i) + g_2(y))f_y^0(y)h_y(y)\,dy \\[2mm] +\delta\varphi_2(z_i)\left(b_1\varphi_1'(z_i) - \varphi_2'(z_i)\right)\int\limits_{\mathcal{Y}} f_\eta'(a_1 + b_1\varphi_1(z_i) - \varphi_2(z_i) + g_2(y))f_y^0(y)\,h_y(y)\,dy \end{array}\right]_{i=1,2,3}$$

$$D\Psi\left(\delta f_\varepsilon\right) =$$

$$\left[(1 - b_1)\,\varphi_1'(z_i)\int\limits_{\mathcal{X}} \delta f_\varepsilon(-a_1 + (1 - b_1)\,\varphi_1(z_i) + g_1(x))f_x^0(x)\,h_x(x)\,dx\right]_{i=1,2,3}$$

$$D\Psi\left(\delta f_\eta\right) =$$

$$\left[-\left(b_1\varphi_1'(z_i) - \varphi_2'(z_i)\right)\int\limits_{\mathcal{Y}} \delta f_\eta(a_1 + b_1\varphi_1(z_i) - \varphi_2(z_i) + g_2(y))\,f_y^0(y)\,h_y(y)\,dy\right]_{i=1,2,3}$$

$$D\Psi\left(\delta g_1\right) =$$

$$\left[(1 - b_1)\,\varphi_1'(z_i)\int\limits_{\mathcal{X}} f_\varepsilon'(-a_1 + (1 - b_1)\,\varphi_1(z_i) + g_1(x))\delta g_1(x)\,f_x^0(x)\,h_x(x)\,dx\right]_{i=1,2,3}$$

$$D\Psi\left(\delta g_2\right) =$$

$$\left[-\left(b_1\varphi_1'(z_i) - \varphi_2'(z_i)\right)\int\limits_{\mathcal{Y}} f_\eta'(a_1 + b_1\varphi_1(z_i) - \varphi_2(z_i) + g_2(y))\,\delta g_2(y)\,f_y^0(y)\,h_y(y)\,dy\right]_{i=1,2,3}$$

$$D\Psi\left(\delta f_x^0\right) =$$

$$\left[(1 - b_1)\,\varphi_1'(z_i)\int\limits_{\mathcal{X}} f_\varepsilon(-a_1 + (1 - b_1)\,\varphi_1(z_i) + g_1(x))\delta f_x^0(x)\,h_x(x)\,dx\right]_{i=1,2,3}$$

$$D\Psi\left(\delta f_y^0\right) =$$

$$\left[-\left(b_1\varphi_1'(z_i) - \varphi_2'(z_i)\right)\int\limits_{\mathcal{Y}} f_\eta(a_1 + b_1\varphi_1(z_i) - \varphi_2(z_i) + g_2(y))\,\delta f_y^0(y)\,h_y(y)\,dy\right]_{i=1,2,3}$$

$$D\Psi\left(\delta a_1\right) =$$

$$\left[\delta a_1 \left(\begin{array}{c} (b_1 - 1)\,\varphi_1'\left(z_i\right) \int\limits_{\mathcal{X}} f_\varepsilon'(-a_1 + (1-b_1)\varphi_1(z_i) + g_1(x))f_x^0(x)h_x\left(x\right)dx \\[2ex] -\left(b_1\varphi_1'\left(z_i\right) - \varphi_2'\left(z_i\right)\right) \int\limits_{\mathcal{Y}} f_\eta'(a_1 + b_1\varphi_1(z_i) - \varphi_2(z_i) + g_2(y))f_y^0(y)h_y\left(y\right)dy \end{array} \right) \right]_{i=1,2,3}$$

$$D\Psi\left(\delta b_1\right) =$$

$$\left[\delta b_1 \left(\begin{array}{c} -\varphi_1'\left(z_i\right)\int\limits_{\mathcal{X}} f_\varepsilon(-a_1 + (1-b_1)\varphi_1(z_i) + g_1(x))f_x^0(x)h_x\left(x\right)dx \\[2ex] -\varphi_1'\left(z_i\right)\int\limits_{\mathcal{Y}} f_\eta(a_1 + b_1\varphi_1(z_i) - \varphi_2(z_i) + g_2(y))f_y^0(y)h_y\left(y\right)dy \\[2ex] -\varphi_1\left(z_i\right)\left(1-b_1\right)\varphi_1'\left(z_i\right)\int\limits_{\mathcal{X}} f_\varepsilon'(-a_1 + (1-b_1)\varphi_1(z_i) + g_1(x))f_x^0(x)h_x\left(x\right)dx \\[2ex] -\varphi_1\left(z_i\right)b_1\varphi_1'\left(z_i\right)\int\limits_{\mathcal{Y}} f_\eta'(a_1 + b_1\varphi_1(z_i) - \varphi_2(z_i) + g_2(y))f_y^0(y)h_y\left(y\right)dy \\[2ex] +\varphi_1\left(z_i\right)\varphi_2'\left(z_i\right)\int\limits_{\mathcal{Y}} f_\eta'(a_1 + b_1\varphi_1(z_i) - \varphi_2(z_i) + g_2(y))f_y^0(y)h_y\left(y\right)dy \end{array} \right) \right]_{i=1,2,3}$$

Since the functions f_ε and f_η are uniformly bounded, as are their first derivatives, all the integrals in these formulas are well-defined. Since the functions f_ε and f_η are uniformly continuous, as are their first derivatives, these integrals depend continuously on (a_1, b_1) and on θ. So the function Ψ is C^1. ∎

This ends the proof of the lemma. To prove the theorem, we have to vary each pair of parameters singly. This amounts to considering, instead of Ψ, the partial maps obtained by keeping all parameter values fixed except two, and showing that the corresponding derivative is onto. This gives four different cases.

Genericity with respect to (φ_1, φ_2) We consider the partial map $\Psi\left(\varphi_1, \varphi_2, a_1, b_1\right)$ and the derivative of the partial map $D\Psi(\delta\varphi_1, \delta\varphi_2, \delta a_1, \delta b_1)$, where it is understood that all the other parameters $f_\varepsilon, f_\eta, g_1, g_2, f_x^0, f_y^0$ are set at fixed values. Hence the derivative of

the partial map is given by $(A - 1)$ with all variations other than $(\delta\varphi_1, \delta\varphi_2, \delta a_1, \delta b_1)$ set to zero. The derivative of the partial map is

$$
D\Psi\left(\delta\varphi_1, \delta\varphi_2, \delta a_1, \delta b_1\right) =
$$

$$
\left[
\begin{array}{c}
\delta\varphi_1'(z_i)
\begin{pmatrix}
(1-b_1)\int\limits_{\mathcal{X}} f_\varepsilon(-a_1 + (1-b_1)\varphi_1(z_i) + g_1(x))f_x^0(x)h_x\,(x)\,dx \\
-b_1\int\limits_{\mathcal{Y}} f_\eta(a_1 + b_1\varphi_1(z_i) - \varphi_2(z_i) + g_2(y))f_y^0(y)h_y\,(y)\,dy
\end{pmatrix}
+ \\[2em]
\delta\varphi_1(z_i)
\begin{pmatrix}
(1-b_1)^2\varphi_1'(z_i)\int\limits_{\mathcal{X}} f_\varepsilon'(-a_1 + (1-b_1)\varphi_1(z_i) + g_1(x))f_x^0(x)h_x\,(x)\,dx \\
-b_1^2\varphi_1'(z_i)\int\limits_{\mathcal{Y}} f_\eta'\,(a_1 + b_1\varphi_1(z_i) - \varphi_2(z_i) + g_2(y))\,f_y^0\,(y)\,h_y\,(y)\,dy \\
+b_1\varphi_2'(z_i)\int\limits_{\mathcal{Y}} f_\eta'\,(a_1 + b_1\varphi_1(z_i) - \varphi_2(z_i) + g_2(y))\,f_y^0\,(y)\,h_y\,(y)\,dy
\end{pmatrix}
+
\end{array}
\right]_{i=1,2,3}
$$

$$
\left[
\begin{pmatrix}
\delta\varphi_2'(z_i)\int\limits_{\mathcal{Y}} f_\eta(a_1 + b_1\varphi_1(z_i) - \varphi_2(z_i) + g_2(y))f_y^0(y)h_y\,(y)\,dy \\[1em]
+\delta\varphi_2(z_i)\,(b_1\varphi_1'(z_i) - \varphi_2'(z_i))\int\limits_{\mathcal{Y}} f_\eta'(a_1 + b_1\varphi_1(z_i) - \varphi_2(z_i) + g_2(y))f_y^0\,(y)\,h_y\,(y)\,dy \\[1em]
+D\Psi\,(\delta a_1) + D\Psi\,(\delta b_1)
\end{pmatrix}
\right]_{i=1,2,3}
$$

Since the points z_i are pairwise distinct, we can choose the $(\delta\varphi_1, \delta\varphi_2)$ so that

$$
(\delta\varphi_1(z_i), \delta\varphi_2(z_i)) = (0, 0)
$$

for all i. Choosing in addition $(\delta a_1, \delta b_1) = (0, 0)$ cancels all the terms on the right-hand side except the first and third ones. Since the remaining integrals are non-zero, the coefficients of $\delta\varphi_1'(z_i)$ and $\delta\varphi_2'(z_i)$ cannot vanish together. So the image by $D\Psi$ of vectors such that $(\delta\varphi_1(z_i), \delta\varphi_2(z_i)) = (0, 0)$ and $(\delta a_1, \delta b_1) = (0, 0)$ must be all of R^3.

Saying that $D\Psi$ is onto means that the partial map Ψ is transversal to every point in R^3, in particular to the origin. By Thom's transversality theorem, generically in (φ_1, φ_2),

the partial map

$$(a_1, b_1) \rightarrow \Psi(\varphi_1, \varphi_2, a_1, b_1)$$

is transversal to the origin. This means that generically in (φ_1, φ_2) either $\Psi(\varphi_1, \varphi_2, a_1, b_1) \neq 0$ or $\Psi(\varphi_1, \varphi_2, a_1, b_1) = 0$ and the partial derivative $D_{a_1,b_1}\Psi$ is onto. The latter case is impossible since $D_{a_1,b_1}\Psi$ sends a two-dimensional space into a three-dimensional one. So generically $\Psi(\varphi_1, \varphi_2, a_1, b_1) \neq 0$.

We have thus proved that, generically in (φ_1, φ_2), we must have $\Phi(\varphi_1, \varphi_2, a_1, b_1)(z_i) \neq 0$ for one i at least. This implies that $\Phi(\varphi_1, \varphi_2, a_1, b_1)(z)$ cannot be identically zero, and hence that the equilibrium equation does not have a solution of the form $P'(z) = a_1 + b_1\varphi_1(z)$. A parallel argument shows that, generically in (φ_1, φ_2), the equilibrium equation does not have a solution of the form $P'(z) = a_2 + b_2\varphi_2(z)$. Since the intersection of two generic properties is generic, the theorem follows for the pair (φ_1, φ_2).∎

Genericity with respect to (f_ε, f_η) We consider the partial map $\Psi(f_\varepsilon, f_\eta, a_1, b_1)$, where it is understood that all the other parameters are pegged to fixed values. The derivative of the partial map is given by

$$D\Psi(\delta f_\varepsilon, \delta f_\eta, \delta a_1, \delta b_1) =$$
$$\left[\begin{array}{c} (1 - b_1)\varphi_1'(z_i) \int_{\mathcal{X}} \delta f_\varepsilon(-a_1 + (1 - b_1)\varphi_1(z_i) + g_1(x))f_x^0(x) h_x(x) dx \\ - (b_1\varphi_1'(z_i) - \varphi_2'(z_i)) \int_{\mathcal{Y}} \delta f_\eta(a_1 + b_1\varphi_1(z_i) - \varphi_2(z_i) + g_2(y)) f_y^0(y) h_y(y) dy \\ + D\Psi(\delta a_1) + D\Psi(\delta b_1) \end{array} \right]_{i=1,2,3}$$

Introduce the distribution functions μ_1 and μ_2 of the random variables g_1 and g_2. They are probability measures on the real line. Setting $(\delta a_1, \delta b_1) = (0, 0)$, the above formula can be

rewritten as:

$$D\Psi\left(\delta f_\varepsilon, \delta f_\eta\right) =$$

$$\left[\begin{array}{c} (1 - b_1)\, \varphi_1'\left(z_i\right) \int \delta f_\varepsilon(-a_1 + (1 - b_1)\, \varphi_1\left(z_i\right) + t)d\mu_1 - \\ (b_1\varphi_1'\left(z_i\right) - \varphi_2'\left(z_i\right)) \int \delta f_\eta\left(a_1 + b_1\varphi_1\left(z_i\right) - \varphi_2\left(z_i\right) + t\right) d\mu_2 \end{array}\right]_{i=1,2,3}$$

Setting $c_i = -a_1 + (1 - b_1)\, \varphi_1\left(z_i\right)$ and $d_i = a_1 + b_1\varphi_1\left(z_i\right) - \varphi_2\left(z_i\right)$, and denoting by μ_1^i and μ_2^i the translates of μ_1 and μ_2 by $-c_i$ and $-d_i$ we rewrite the partial derivative again as:

$$D\Psi\left(\delta f_\varepsilon, \delta f_\eta\right) = \left[(1 - b_1)\, \varphi_1'\left(z_i\right) \int \delta f_\varepsilon(t)d\mu_1^i - (b_1\varphi_1'\left(z_i\right) - \varphi_2'\left(z_i\right)) \int \delta f_\eta\left(t\right) d\mu_2^i\right]_{i=1,2,3}.$$

We pick the probability measures μ_1^i and $\mu_2^i, i = 1, 2, 3$, so they are pairwise different, and the values of the z, so that $\varphi_1'\left(z_i\right)$ and $\varphi_2'\left(z_i\right)$ do not vanish. Then the coefficients of the integrals cannot vanish simultaneously, and the right-hand side clearly spans R^3. We conclude as in the preceding case by applying Thom's transversality theorem.∎

Genericity with respect to $\left(f_x^0, f_y^0\right)$ We consider the partial map $\Psi\left(f_x^0, f_y^0, a_1, b_1\right)$, where it is understood that all the other parameters are pegged to fixed values. The partial derivative is given by

$$D\Psi(\delta f_x^0, \delta f_y^0, \delta a_1, \delta b_1) =$$

$$\left[\begin{array}{c} (1 - b_1)\, \varphi_1'\left(z_i\right) \int_{\mathcal{X}} f_\varepsilon(-a_1 + (1 - b_1)\, \varphi_1\left(z_i\right) + g_1\left(x\right))\delta f_x^0\left(x\right) h_x\left(x\right) dx \\ - (b_1\varphi_1'\left(z_i\right) - \varphi_2'\left(z_i\right)) \int_{\mathcal{Y}} f_\eta\left(a_1 + b_1\varphi_1\left(z_i\right) - \varphi_2\left(z_i\right) + g_2\left(y\right)\right) \delta f_y^0\left(y\right) h_y\left(y\right) dy \\ + D\Psi\left(\delta a_1\right) + D\Psi\left(\delta b_1\right) \end{array}\right]_{i=1,2,3}$$

60

We claim that the partial map obtained by setting $(\delta a_1, \delta b_1) = 0$ is onto. At these values

$$D\Psi(\delta f_x^0, \delta f_y^0) =$$

$$\left[\begin{array}{c} (1 - b_1)\, \varphi_1'\,(z_i) \int\limits_{\mathcal{X}} f_\varepsilon(-a_1 + (1 - b_1)\, \varphi_1\,(z_i) + g_1\,(x)) \delta f_x^0\,(x)\, h_x\,(x)\, dx \\[2mm] - (b_1 \varphi_1'\,(z_i) - \varphi_2'\,(z_i)) \int\limits_{\mathcal{Y}} f_\eta\,(a_1 + b_1 \varphi_1\,(z_i) - \varphi_2\,(z_i) + g_2\,(y))\, \delta f_y^0\,(y)\, h_y\,(y)\, dy \end{array} \right]_{i=1,2,3}.$$

We choose the z_i so that the $\varphi_1'\,(z_i)$ and the $\varphi_2'\,(z_i)$ do not vanish, and so that the functions $f_\varepsilon(-a_1 + (1 - b_1)\, \varphi_1\,(z_i) + g_1\,(x))$ and $f_\eta\,(a_1 + b_1 \varphi_1\,(z_i) - \varphi_2\,(z_i) + g_2\,(y))$ are pairwise different on a set of positive measure. The claim then follows, and genericity obtains as in the preceding cases.■

Genericity with respect to (g_1, g_2) We consider

$$D\Psi(\delta g_1, \delta g_2, \delta a_1, \delta b_1) =$$

$$\left[\begin{array}{c} (1 - b_1)\, \varphi_1'\,(z_i) \int\limits_{\mathcal{X}} f_\varepsilon'(-a_1 + (1 - b_1)\, \varphi_1\,(z_i) + g_1\,(x)) \delta g_1\,(x)\, f_x^0\,(x)\, h_x\,(x)\, dx \\[2mm] - (b_1 \varphi_1'\,(z_i) - \varphi_2'\,(z_i)) \int\limits_{\mathcal{Y}} f_\eta'\,(a_1 + b_1 \varphi_1\,(z_i) - \varphi_2\,(z_i) + g_2\,(y))\, \delta g_2\,(y)\, f_y^0\,(y)\, h_y\,(y)\, dy \\[2mm] + D\Psi\,(\delta a_1) + D\Psi\,(\delta b_1) \end{array} \right]_{i=1,2,3}$$

and we argue as in the preceding case.

Proof of Theorem 3:

The strategy of the proof is to determine whether taking the objects determined from the marginal densities as demonstrated in the previous subsection and plugging them into (29) provides any more information about the parameters that are not identified.

First we write all the unknown objects in (29) in terms of known functions. We have

$$T_1(z) = K_1 \tilde{T}_1(z) + C_1; \ T_2(z) = K_2 \tilde{T}_2(z) + C_2$$

$$f_{\tilde{\varepsilon}}(\tilde{\varepsilon}) = f_\varepsilon(K_1 \tilde{\varepsilon}) K_1; \ f_{\tilde{\eta}}(\tilde{\eta}) = f_\eta(K_2 \tilde{\eta}) K_2$$

$$g_1(x) = K_1 \tilde{g}_1(x) + R_1; \ g_2(x) = K_2 \tilde{g}_2(x) + R_2$$

where

$$\tilde{\varepsilon}(z,x) = \frac{\varepsilon}{K_1} = \tilde{T}_1(z) + \tilde{g}_1(x) + \frac{C_1 + R_1}{K_1}$$

$$\tilde{\eta}(z,x) = \frac{\eta}{K_2} = \tilde{T}_2(z) + \tilde{g}_2(x) + \frac{C_2 + R_2}{K_2}.$$

and we define $\tau_1 = \frac{C_1 + R_1}{K_1}$ and $\tau_2 = \frac{C_2 + R_2}{K_2}$. All objects denoted with "$\sim$" are known from the analysis of the marginal densities of (z,x) and (z,y). Additionally, τ_1 and τ_2 are known from imposing a mean or median 0 restriction on $\tilde{\varepsilon}$ and $\tilde{\eta}$ respectively.

Substituting these expressions into (29) we obtain

$$f(x,y,z) = \frac{f_x f_y f_{\tilde{\varepsilon}}\left(\tilde{T}_1(z) + \tilde{g}_1(x) + \tau_1\right) \tilde{T}_1'(z) f_{\tilde{\eta}}\left(\tilde{T}_2(z) + \tilde{g}_2(y) + \tau_2\right) \tilde{T}_2'(z)}{\int_{\mathcal{X}} f_x(x') f_{\tilde{\varepsilon}}\left(\tilde{T}_1(z) + \tilde{g}_1(x') + \tau_1\right) \tilde{T}_1'(z) \, dx'}.$$

It is apparent from the expression for the joint density that it conveys no new identifying information on K_1 and K_2 beyond what is obtained from the marginal distribution, since all terms involving these two parameters cancel and K_1 and K_2 do not explicitly appear in the final expression.

Proof of Theorem 4:

As in Theorem 1, we assume there are some continuous probability densities h_x and h_y with $h_x > 0$ on $\mathcal{X}$ and $h_y > 0$ on $\mathcal{Y}$ such that $f_{x\varepsilon}/h_x$ and $f_{y\eta}/h_y$ are bounded away from zero and infinity on $\mathcal{X}$ and $\mathcal{Y}$. Note that this will always be the case if $\mathcal{X}$ and $\mathcal{Y}$ are compact intervals. Set $f_{x\varepsilon}^0 = f_{x\varepsilon}/h_x$ and $f_{y\eta}^0 = f_{y\eta}/h_y$.

The notation and parameter spaces are as in Theorem 1 except that now the natural spaces for $f_{x\varepsilon}^0$ and $f_{y\eta}^0$ are $C_1^2 (\mathcal{X} \times \mathcal{E})$ and $C_1^2 (\mathcal{Y} \times \mathcal{H})$. Consider the partial map

$$\Psi \left(f_{x\varepsilon}^0, f_{y\eta}^0, a_1, b_1 \right) =$$
$$\left[\begin{array}{c} (1 - b_1) \, \varphi_1' (z) \int\limits_{\mathcal{X}} f_{x\varepsilon}^0 (x, -a_1 + (1 - b_1) \, \varphi_1 (z) + g_1 (x)) h_x (x) \, dx - \\ (b_1 \varphi_1' (z) - \varphi_2' (z)) \int\limits_{\mathcal{Y}} f_{\eta y}^0 (y, a_1 + b_1 \varphi_1 (z_i) - \varphi_2 (z_i) + g_2 (y)) \, h_y (y) \, dy \end{array} \right]_{i=1,2,3}$$

and its derivative

$$D\Psi \left(\delta f_{x\varepsilon}, \delta f_{y\eta}, \delta a_1, \delta b_1 \right) =$$
$$\left[\begin{array}{c} (1 - b_1) \, \varphi_1' (z_i) \int\limits_{\mathcal{X}} \delta f_{x\varepsilon}^0 (x, -a_1 + (1 - b_1) \, \varphi_1 (z_i) + g_1 (x)) h_x (x) \, dx - \\ (b_1 \varphi_1' (z_i) - \varphi_2' (z_i)) \int\limits_{\mathcal{Y}} \delta f_{y\eta}^0 (y, a_1 + b_1 \varphi_1 (z_i) - \varphi_2 (z_i) + g_2 (y)) \, h_y (y) \, dy \\ +D\Psi \left(\delta a_1 \right) + D\Psi \left(\delta b_1 \right) \end{array} \right]_{i=1,2,3}$$

We need to show that the right-hand side spans R^3. Consider perturbations of the type $\delta f_{x\varepsilon} (x, \varepsilon) = \delta f_{x\varepsilon} (x', \varepsilon)$ for all x, x' and $\delta f_{y\eta} (y, \eta) = \delta f_{y\eta} (y', \eta)$ for all y, y'. Then argue as in the proof of genericity with respect to (f_ε, f_η) as in Theorem 1. Then we can conclude the hypothesis of Theorem 4 as in the preceding theorems by applying Thom's transversality theorem.∎

Proof of Theorem 5:

By assumption,

$$P' (z) - \varphi_1 (z) = g_1 (x) - \varepsilon.$$

Because of the second-order condition $P'' (z) - \varphi_1' (z) < 0$ so the left side can be inverted uniquely (globally) to obtain

$$z = \Lambda \left(g_1 (x) - \varepsilon \right)$$

63

where by the implicit function theorem $\Lambda'(q) = [P''(\Lambda(q)) - \varphi_1'(\Lambda(q))]^{-1}$. Define the mapping

$$h = \varphi_1(\Lambda(g_1(x) - \varepsilon))$$

$$x = x,$$

where

$$E_z(\varphi_1(z)|x) = E_z(h|x) = \int_{\mathcal{E}} \varphi_1(\Lambda(g_1(x) - \varepsilon)) f_{\varepsilon|x}(\varepsilon|x) d\varepsilon.$$

This conditional expectation is a functional of g_1. We must prove that generically $E_z(\varphi_1(z)|x)$ is not a linear function of $g_1(x)$. We will show that, generically with respect to $f_{x\varepsilon}^0$, if the conditional expectation is a linear function of $g_1(x)$ then $P'(z) = a_1 + b_1\varphi_1(z)$ is a solution of the equilibrium equation. Generically, this cannot happen, and the result follows.

Assume that $E_z(\varphi_1(z)|x)$ is a linear function of $g_1(x)$. Pick a direction δg_1, and define a function $q(t)$ on the real line by

$$q(t) = \int_{\mathcal{E}} [\varphi_1(\Lambda(g_1(x) + t\,\delta g_1(x) - \varepsilon)) - \varphi_1(\Lambda(g_1(x) - \varepsilon))] f_{\varepsilon|x}(\varepsilon|x) d\varepsilon.$$

Since $E_z(\varphi_1(z)|x)$ is a linear function of $g_1(x)$, then $q(t)$ is linear, so that $q''(0) = 0$. Performing the computations, we get the equation:

$$[\delta g_1(x)]^2 \int_{\mathcal{E}} \lambda(g_1(x) - \varepsilon) f_\varepsilon(\varepsilon|x) d\varepsilon = 0$$

where the composite function $\lambda = \left[(\varphi_1''(P'' - \varphi_1') - \varphi_1'(P''' - \varphi_1''))(P'' - \varphi_1')^{-3}\right] \circ \Lambda$. This reduces to:

$$\int_{\mathcal{E}} \lambda(g_1(x) - \varepsilon) f_{\varepsilon|x}(\varepsilon|x) d\varepsilon = 0 \text{ a.e } x$$

which is the same as

$$\int_{\mathcal{E}} \lambda(\varepsilon_1) f_{\varepsilon|x}(g_1(x) - \varepsilon_1|x) d\varepsilon_1 = 0 \text{ a.e } x.$$

The function λ has the property that, a.e. x, it integrates to zero against the density $f_{\varepsilon|x}(g_1(x) - \varepsilon_1|x)$. This can only be the case either λ equals zero or $f_{\varepsilon|x}(g_1(x) - \varepsilon_1|x)$ does not depend on x. This latter condition can only obtain if

$$f(x,\varepsilon) = f_x(x) f_\varepsilon(g_1(x) - \varepsilon) \text{ a.e } (x,\varepsilon). \tag{A-2}$$

That is, $f_{\varepsilon|x}(g_1(x) - \varepsilon_1|x)$ does not depend on x if and only if the joint distribution of (x,ε) has the property that $\varepsilon_1 = g_1(x) - \varepsilon$ is independent of x. The set of functions that $f_{x\varepsilon}^0$ that satisfies $(A-2)$ is clearly not generic in $C_1^2(\mathcal{X} \times \mathcal{E})$.

Hence, it follows that generically with respect to $f_{x\varepsilon}^0$, we have $\lambda = 0$ a.e. As a result,

$$\left[(\varphi_1'' - \varphi_1'(P''' - \varphi_1''))(P'' - \varphi_1')^{-2}\right] \circ \Lambda$$

vanishes, meaning that $(\varphi_1'' - \varphi_1'(P''' - \varphi_1''))(P'' - \varphi_1')^{-2}$ vanishes on the range of Λ, which is precisely the domain $\mathcal{Z}$ of z. This proves that

$$(\varphi_1''(P'' - \varphi_1') - \varphi_1'(P''' - \varphi_1'')) = 0 \text{ on } \mathcal{Z}$$

so that

$$\frac{\varphi_1''}{\varphi_1'} = \frac{P'''}{P''}$$

and $P'' = b_1 \varphi_1'$. Thus $P' = a_1 + b_1 \varphi_1$ solves the equilibrium equation and the proof is concluded.∎

65

Appendix B   Model Specifications

B.1   Model 1

Figures 2 and 3 display the slope and curvature of the equilibrium price function in the unrestricted linear-quadratic hedonic economy in three cases. In all three cases we use the parameters $A = 2.0$ and $B = 2.0$. In addition, each case assumes that both ν and θ are distributed as mixtures of two normals. ν is distributed as a mixture of normals with weight λ on the first component (mean 0.0 and variance 0.5) and weight $1 - \lambda$ on the second component (mean 1.0 and variance 1.0). θ is distributed as a mixture of normals with weight λ on the first component (mean 1.0 and variance 1.0) and weight $1 - \lambda$ on the second component (mean -1.0 and variance 0.1).

Each of the three cases is distinguished by the weights on the two components. The benchmark case assumes that λ, the weight on component one of the mixture is 1.0. Hence, case 1 is the benchmark linear-quadratic-normal model. The other two cases use different values of λ to show how the slope and curvature of the price function vary with λ. The other two cases use $\lambda = 0.9$ and $\lambda = 0.5$.

B.2   Model 2

Figures 4 and 5 display the slope and curvature of the equilibrium price function in a model that restricts $z > 0$ and $P_z(z) > 0$. In this model, firms' profits are $\Pi(z) = \frac{\exp(\nu) z^{1-B}}{1-B} - P(z)$ where $\nu = \nu_0 + \nu_1' y + \eta$. y and η are independent and are distributed as mixtures of normals. Workers have utility $V(z) = -\frac{\exp(-\theta) z^{1+A}}{1+A} + P(z)$ where $\theta = \theta_0 + \theta_1' x + \varepsilon$ and x and ε are independent and distributed as mixtures of normals. Figures (2) and (3) display the slope

and curvature of the equilibrium price function in this model for three cases. All three cases use the parameters $A = 2.0$ and $B = 2.0$. In addition, each case assumes that both ν and θ are distributed as mixtures of two normals. ν is distributed as a mixture of normals with weight λ on the first component (mean 0.0 and variance 0.5) and weight $1 - \lambda$ on the second component (mean 1.0 and variance 1.0). θ is distributed as a mixture of normals with weight λ on the first component (mean 1.0 and variance 1.0) and weight $1 - \lambda$ on the second component (mean -1.0 and variance 0.1).

Each of the three cases is distinguished by the weights on the two components. The benchmark case assumes that λ, the weight on component one of the mixture is 1.0. In this case $\ln P_z(z)$ is a linear function of $\ln z$ as shown in figure 4. The other two cases use different values of λ to show how the slope and curvature of the price function vary with λ. The other two cases use $\lambda = 0.9$ and $\lambda = 0.5$.

Figure 1: Optimal job choice for three worker-firm pairs

Dollars

z

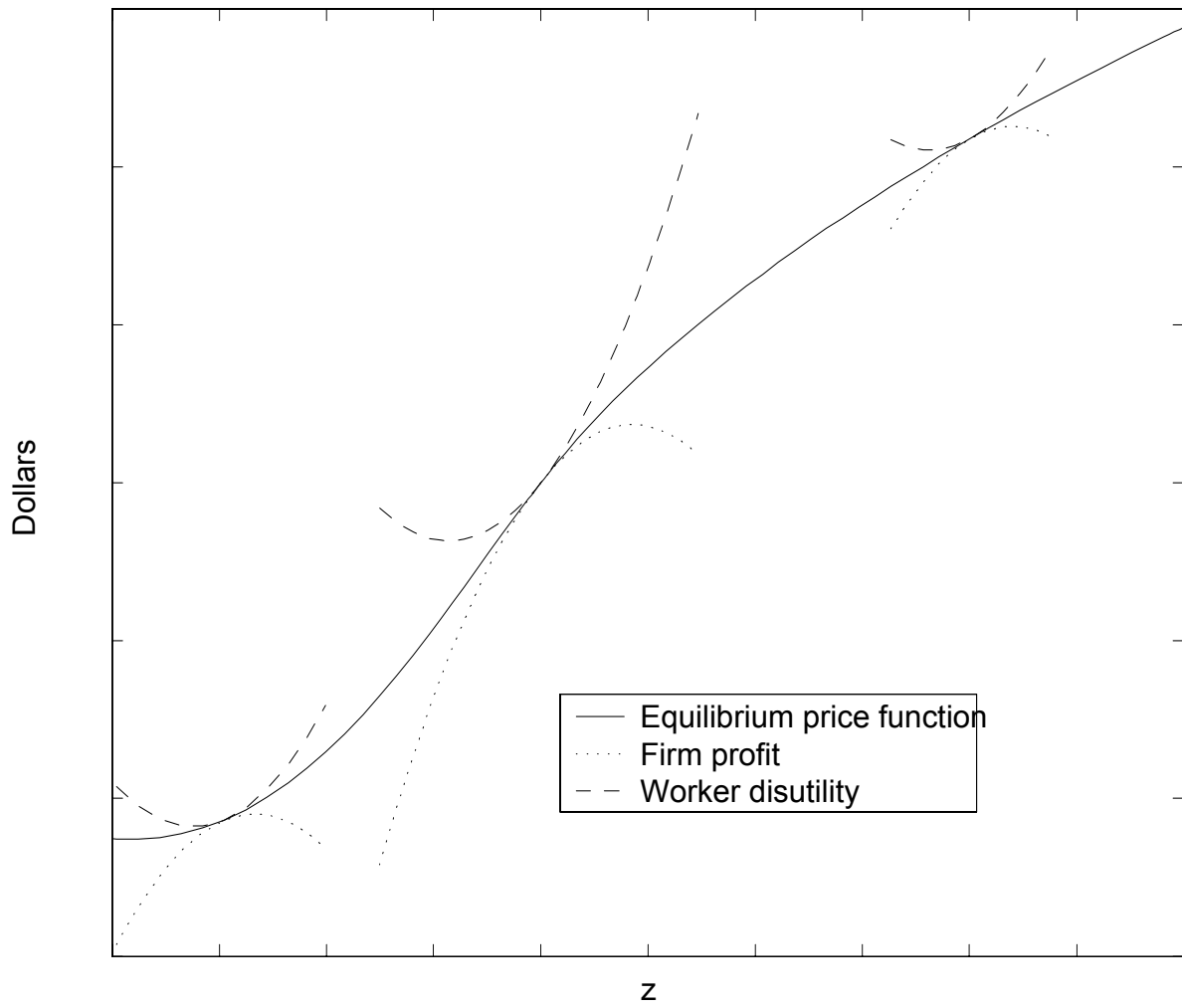Equilibrium price function
Firm profit
Worker disutility

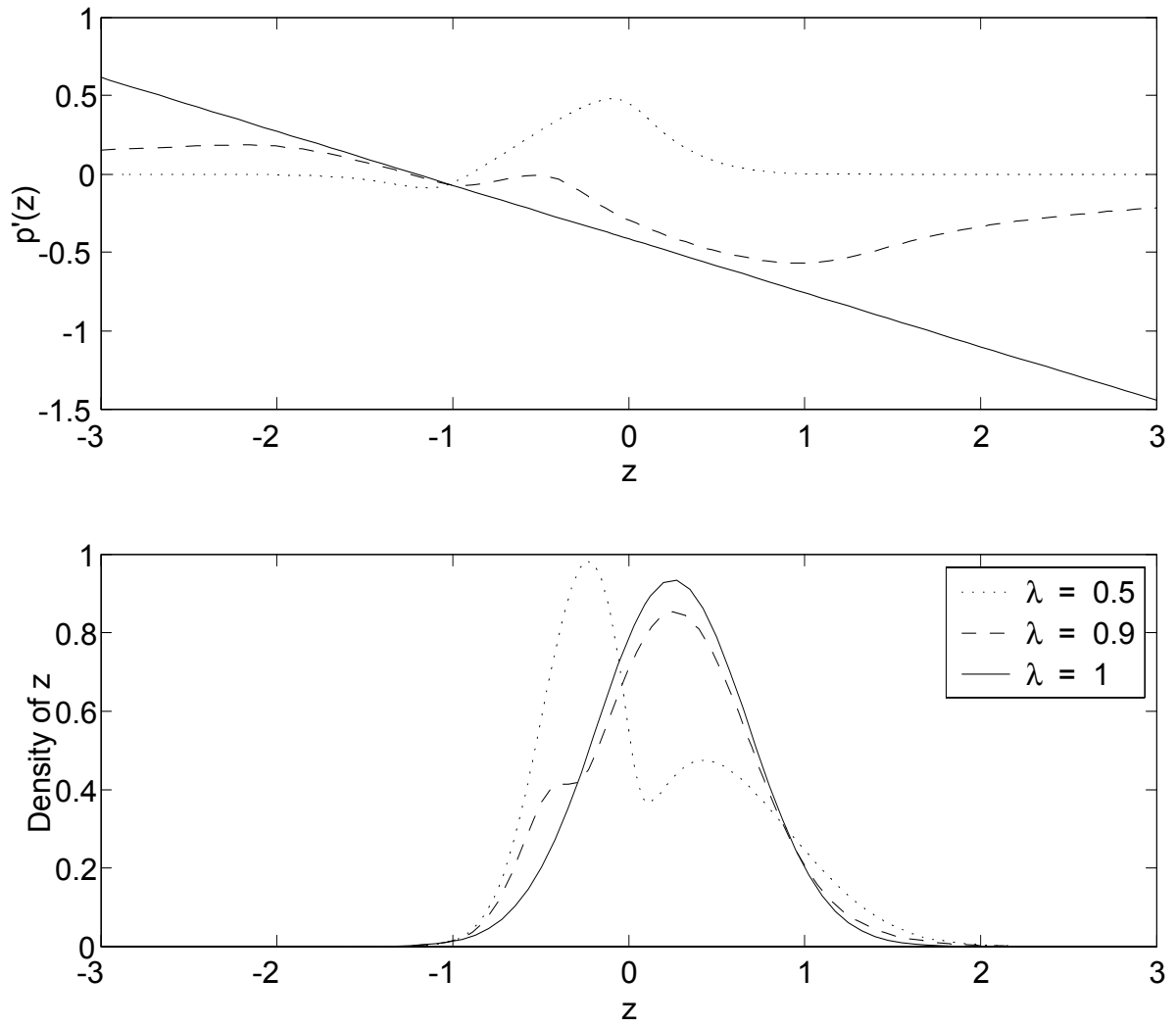Figure 2: Slope of Price Function: Model 1

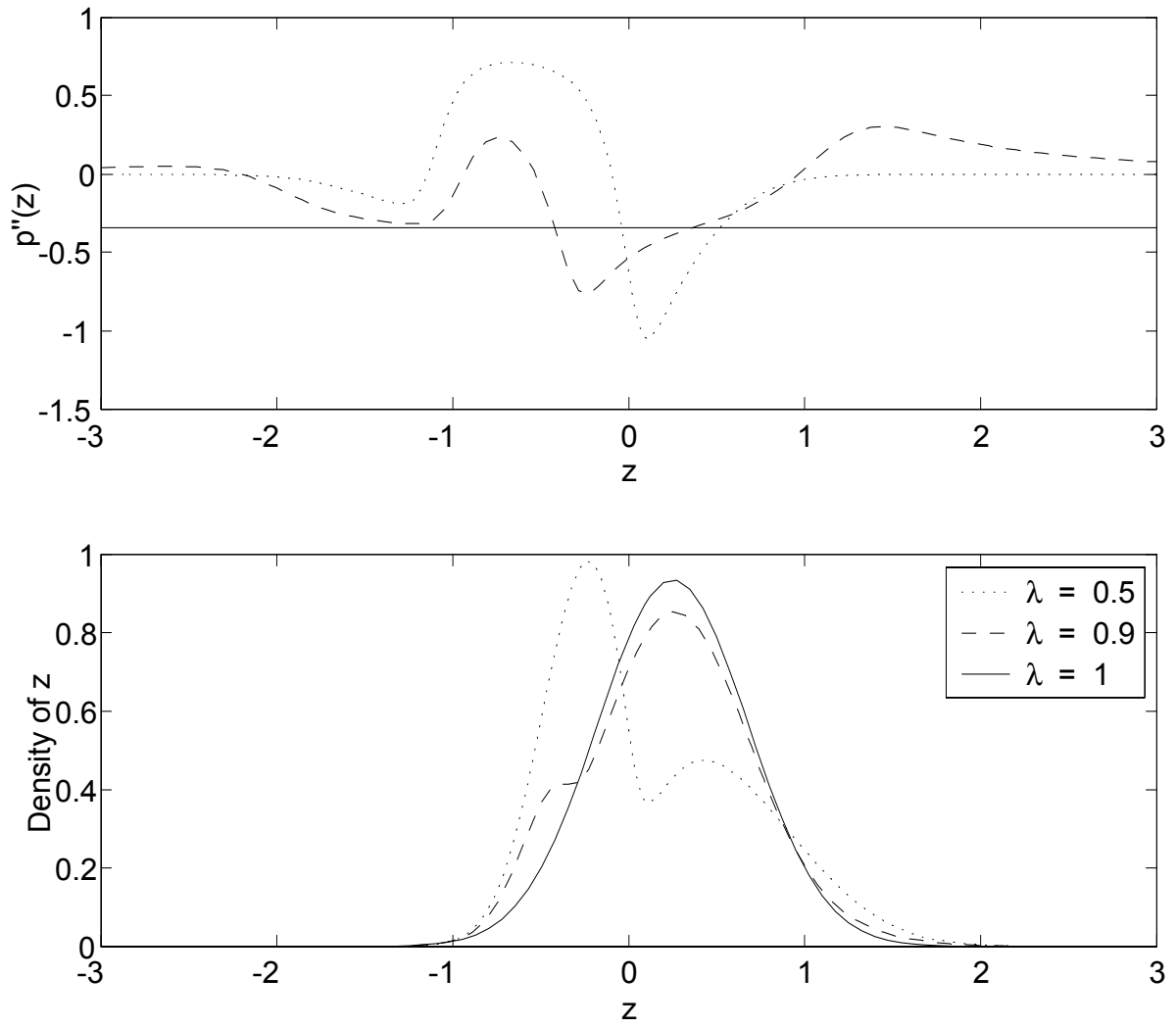Figure 3: Curvature of Price Function: Model 1

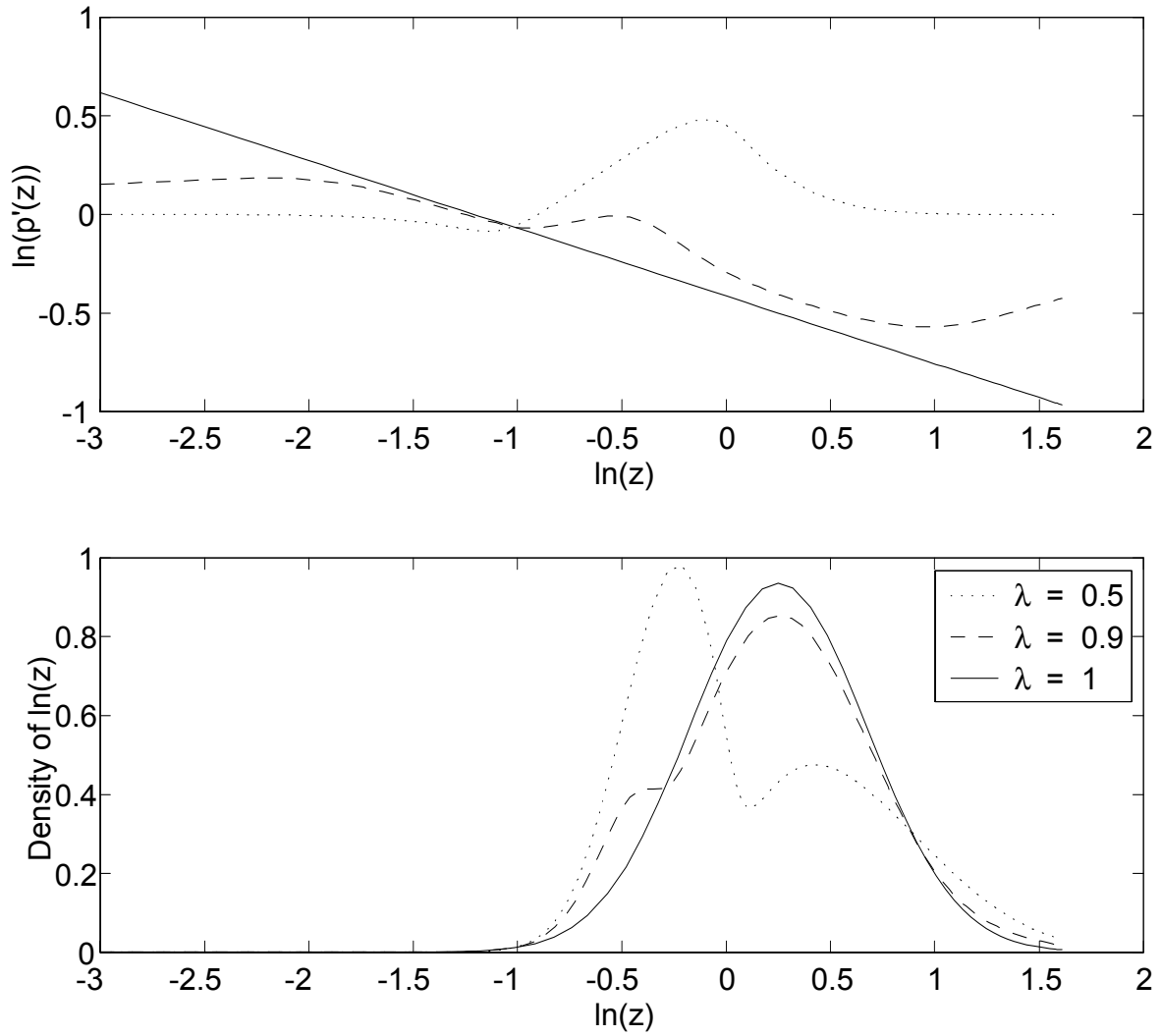Figure 4: Logarithm of Slope of Price Function: Model 2

Figure 5: Elasticity of Slope of Price Function with Respect to z