Universidad de
SanAndrés

Departamento de Economía
Ciclo de Seminarios

Nonparametric Estimation of Nonadditive Hedonic Models

Rosa Matzkin
Northwestern University - Universidad de San Andrés
(Con James Heckman y Lars Nesheim)

Martes 3 de septiembre de 2002
Sala de reuniones - CEDI

Departamento de Economía

Director
Mariano Tommasi
Ph.D. in Economics, University of Chicago

Coordinador del Ciclo de Seminarios
Martín Gonzalez Eiras
Ph.D. in Economics, MIT

En Internet:
http://www.udesa.edu.ar/deptodeeconomia/dtoeconomiappal.htm

NONPARAMETRIC ESTIMATION OF NONADDITIVE HEDONIC MODELS

James Heckman
Department of Economics
University of Chicago

Rosa Matzkin
Department of Economics
Northwestern University

Lars Nesheim
Institute for Fiscal Studies and
University College London

June 2002

Abstract

We present methods to estimate marginal utility and marginal product functions that are nonadditive in the unobservable random terms, using observations from a single hedonic equilibrium market. We show that nonadditive marginal utility and nonadditive marginal product functions are capable of generating equilibria that exhibit bunching, as well as other types of equilibria. We provide conditions under which these types of utility and production functions are nonparametrically identified, and we propose nonparametric estimators for them. The estimators are shown to be consistent and asymptotically normal.

1. INTRODUCTION

In hedonic models, the price of a product is a function of the vector of attributes characterizing the product. These models are used to study the price of a large variety of attributes, such as job safety, size of a house, school quality, distance of a house from an environmental hazard, and others.

In a seminal paper, Sherwin Rosen (1974) pioneered the study of hedonic models in perfectly competitive settings. An economy in these models is specified by a distribution of consumers and a distribution of firms. In equilibrium, consumers are matched with firms. In these models, each consumer is characterized by a utility function that depends on the attributes characterizing the product, as well as on some individual characteristics. Each firm is characterized by a production function that depends on the attributes characterizing the product, as well as on some characteristics of the firm. Given a price function for the attributes, each consumer demands the vector of attributes that maximizes his utility, and each firm supplies the vector of attributes that maximizes its profit. The equilibrium price function is such that the distribution of demand equals the distribution of supply, for all values of the attributes.

Rosen (1974) suggested a method to estimate hedonic models. First, estimate the price function. Second, use the equations for the first order conditions of the optimization of the consumers and firms to estimate the utility and production functions. When the utility and production functions are quadratic and the unobservable heterogeneity variables are normal, the model has a closed form solution, where the equilibrium marginal price function is linear in the attributes. (This particular specification was first studied by Tinbergen (1956).)

The influential papers by James Brown and Harvey Rosen (1982) and Brown (1983) strongly criticized the method of identification proposed by Rosen. (See also Epple (1987) and Kahn and Lang (1988).) Using the linear-quadratic model as an approximation, Brown and Rosen argued that hedonic models are not identified. They claimed that sorting implies that within a single market, there are no natural exclusion restrictions.

Recently, Ekeland, Heckman, and Nesheim (2001), building on previous work by Heckman (1991,1995,1999), analyzed Brown and Rosen's claim, and concluded that the nonidentification is specific to the linear case. Moreover, they showed that the linear case is nongeneric. Ekeland, Heckman, and Nesheim (2001) considered a model with additive marginal utility and additive marginal product function and showed that is identified from single market data. In the specification, the marginal utility was the sum of an unobservable random term, a nonparametric function of the attribute, and a nonparametric function of an observable individual characteristic. The marginal product function was specified in a similar way. The equilibrium price function as well as the conditional distributions of the attribute given the observable characteristics were assumed to be given. They presented two methods for recovering the functions. One was based on extensions of average derivative models (Powell, Stock, and Stoker (1989)) and transformation models (Horowitz (1996, 1998)). The other was based on nonparametric instrumental variables (Darolles, Florens, and Renault (2001), Blundell and Powell (2000), Newey and Powell (2000). The performance of those estimators and the ones presented in this paper are studied in Heckman, Matzkin, and Nesheim (2002).

Inspired by the positive identification result in Ekeland, Heckman, and Nesheim (2001), we investigate in this paper the possibility of relaxing the additive structure, which was used in that paper, for the marginal utility and the marginal product functions. The importance of such a study is not only to allow more flexibility in the specification of the utility and marginal product functions in the model, but, more importantly, to specify economies that can generate a wider variety of equilibrium price function.

In this paper, we consider hedonic equilibrium models where the marginal utility and marginal product functions are nonadditive in the unobserved heterogeneity variables. We show that these more general economies are capable of generating equilibria with bunching, in the sense that a positive mass of consumers and firms locate at a common location. (See Nesheim (2001) for analyses of various types of equilibrium price functions; also Wilson (1991).)

We provide conditions under which the nonadditive marginal utility and nonadditive marginal production function are identified from the equilibrium price function, the distribution of demanded attributes conditional on the observable characteristics of the consumers, and the distribution of supplied attributes conditional on the observable characteristics of the firms. The identification proceeds as follows. First, using the methods in Matzkin (2002), we show that from the conditional distributions we can identify the demand and supply functions, which are nonparametric, nonadditive functions of the observable and unobservable characteristics of, respectively, the consumers and firms. Second, we use the demand and supply functions, together with the equilibrium price function, and the restrictions imposed by the first order conditions to recover the marginal utility and marginal product functions. This last step requires making an assumption on the marginal utility and marginal product functions, which reduces by one the dimension of the domain of these functions.

We propose nonparametric estimators for the marginal utility and marginal product functions, and show that they are consistent and asymptotically normal.

Estimation of demand models generated by random utility functions have been studied in the past using parametric assumptions (Heckman (1974), McFadden (1974), Heckman and Willis (1977)), semiparametric assumptions (Manski (1975,1985), Cosslett (1983), Matzkin (1991), Horowitz (1992), Klein and Spady (1993), Ichimura and Thompson (1994), among others), and more recently, using nonparametric assumptions (Matzkin (1992,1993), Briesch, Chintagunta and Matzkin (1997), Brown and Matzkin (1998), Horowitz (2001), McFadden and Train (2000), among others). McElroy (1981,1987), Brown and Walker (1985,1995) and Lewbel (1996) considered inference of random utility and random production functions in perfectly competitive, non-hedonic situations.

Estimation of nonparametric models with nonadditive random terms has been previously studied in Olley and Pakes (1996), Altonji and Ichimura (1997), Altonji and Matzkin (1997, 2001), Briesch, Chintagunta and Matzkin (1997), Brown and Matzkin (1998), Heckman and Vytlacil (1999, 2001), Matzkin (1999, 2002), Vytlacil (2000), Blundell and Powell (2000), and, more recently, by Bajari and Benkard (2001), Chesher (2001), Hong and Shum (2001), and Imbens and Newey (2001). Bajari and Benkard (2001) consider hedonic price functions where some of the attributes are unobservable.

The outline of the paper is as follows. In the next section, we describe the hedonic model, for a univariate attribute. We provide two simple analytic examples of hedonic equilibria generated by nonadditive functions, one without bunching and the other with bunching. In Section 3, we study the identification of nonadditive marginal utility and nonadditive marginal product function. In Section 4, we present nonparametric estimators and their asymptotic properties.

2. THE HEDONIC EQUILIBRIUM MODEL

To describe the hedonic model, we will consider, for simplicity a labor market setting. Consumers (workers) match to single worker firms. Let z denote an attribute vector, characterizing jobs, assumed to be a disamenity for the consumers and an input for the firms. We will assume that z is unidimensional. Each consumer has a utility function $U^*(c, z, x, \varepsilon)$ where c is consumption, x is a vector of observable characteristics of the consumer and ε is an unobservable heterogeneity term. Each firm has a production function $\Gamma(z, y, \eta)$ where y is a vector of observable characteristics of the firm and η is an unobservable heterogeneity term. The function U^* will be assumed to be twice differentiable with respect to its first two arguments. The function Γ will be assumed to be twice differentiable with respect to z. The unobservable random terms, ε and η, will be assumed to be statistically independent of the vectors of observable characteristics, x and y.

Each consumer chooses (c, z) to maximize the utility function $U^*(c, z, x, \varepsilon)$ subject to the constraint

$$c = P(z) + R$$

where R denotes unearned income. Substituting the constraint into the utility function, we can describe the consumer's problem as the choice of z that maximizes the value of the function

$$U^*(P(z)+R,z,x,\varepsilon)$$

The first order condition for this maximization is

$$U_c^*\left(P(z)+R,x,z,\varepsilon\right)P_z\left(z\right)+U_z^*(P(z)+R,z,x,\varepsilon)=0$$

where U_c^* and U_z^* denote the partial derivatives of U^* with respect to, respectively, its first and second arguments. This can be expressed as

$$P_z\left(z\right)=v\left(P(z)+R,z,x,\varepsilon\right)$$

where

$$v\left(P(z)+R,z,x,\varepsilon)\right)\equiv\frac{-U_z^*(P(z)+R,z,x,\varepsilon)}{U_c^*\left(P(z)+R,x,z,\varepsilon\right)}$$

For simplicity, we will restrict our analysis to the case where

$$U_c^*\equiv1$$

so that

$$v\left(P(z)+R,z,x,\varepsilon\right)=-U_z^*(P(z)+R,z,x,\varepsilon)$$

We will further assume, also for simplicity, that $R=0$. Define

$$U_z(z,x,\varepsilon)\equiv-U_z^*(z,x,\varepsilon)$$

Then, the first and second order conditions for maximization of U^* over z become

FOC: $P_z(z)-U_z(z,x,\varepsilon)=0$

SOC: $P_{zz}(z)-U_{zz}(z,x,\varepsilon)<0$

By the Implicit Function Theorem and the SOC, there exists a function $z=s(x,\varepsilon)$ such that

$$P_z(s(x,\varepsilon))-U_z(s(x,\varepsilon),x,\varepsilon)=0$$

Moreover

$$\frac{\partial s(x,\varepsilon)}{\partial\varepsilon}=\frac{U_{z\varepsilon}(s(x,\varepsilon),x,\varepsilon)}{P_{zz}(s(x,\varepsilon))-U_{zz}(s(x,\varepsilon),x,\varepsilon)}$$

Hence,

$$\frac{\partial s(x,\varepsilon)}{\partial \varepsilon} > 0 \quad \text{if} \quad U_{z\varepsilon} < 0$$

Let $\widetilde{s}(z,x)$ denote the inverse of s with respect to ε. Then,

$$P_z(z) - U_z(z,x,\widetilde{s}(z,x)) = 0$$

$$\frac{\partial \widetilde{s}(z,x)}{\partial z} = \frac{P_{zz}(z) - U_{zz}(z,x,\widetilde{s}(z,x))}{U_{z\varepsilon}(z,x,\widetilde{s}(z,x))}$$

and

$$\frac{\partial \widetilde{s}(z,x)}{\partial z} > 0 \quad \text{if} \quad U_{z\varepsilon} < 0$$

From the other side of the market, each firm chooses z to maximize the profit function

$$\Gamma(z,y,\eta) - P(z)$$

The first and second order conditions of this optimization problem are

FOC: $\Gamma_z(z,y,\eta) - P_z(z) = 0$

SOC: $\Gamma_{zz}(z,y,\eta) - P_{zz}(z) < 0$

By the Implicit Function Theorem and SOC there exists a function $z = d(y,\eta)$ such that

$$\Gamma_z(d(y,\eta),y,\eta) - P_z(d(y,\eta)) = 0$$

Moreover

$$\frac{\partial d(y,\eta)}{\partial \eta} = \frac{\Gamma_{z\eta}(d(y,\eta),y,\eta)}{P_{zz}(d(y,\eta)) - \Gamma_{zz}(d(y,\eta),y,\eta)}$$

Hence,

$$\frac{\partial d(y,\eta)}{\partial \eta} > 0 \quad \text{if}\, \Gamma_{z\eta} > 0$$

Let $\widetilde{d}(z,y)$ denote the inverse of d with respect to η. Then,

$$\Gamma_z(z,y,\widetilde{d}(z,y)) - P_z(z) = 0$$

$$\frac{\partial \widetilde{d}(z,y)}{\partial z} = \frac{P_{zz}(z) - \Gamma_{zz}(z,y,\widetilde{d}(z,y))}{\Gamma_{z\eta}(z,y,\widetilde{d}(z,y))}$$

and

$$\frac{\partial \widetilde{d}(z,y)}{\partial z} > 0 \ \text{ if } \Gamma_{z\eta} > 0$$

We will assume that $U_{z\varepsilon} < 0$ and $\Gamma_{z\eta} > 0$. In equilibrium, the density of the demanded z must equal the density of the supplied z for all values of z. To express this condition in terms of the primitive functions, consider the transformation

$$z = s(x,\varepsilon) \ \ \& \ \ x = x$$

The inverse of this transformation is

$$\varepsilon = \widetilde{s}(z,x) \ \ \& \ \ x = x$$

and the Jacobian determinant is

$$\begin{vmatrix} \frac{\partial \widetilde{s}(z,x)}{\partial z} & \frac{\partial \widetilde{s}(z,x)}{\partial x} \\ 0 & 1 \end{vmatrix} = \left| \frac{\partial \widetilde{s}(z,x)}{\partial z} \right| = \frac{\partial \widetilde{s}(z,x)}{\partial z}$$

Let f_x and f_ε denote the densities of the vector of observable and unobservable characteristics of the consumers. Let $\widetilde{X}$ denote the support of x. Then, the density of the supplied z is

$$\int_{\widetilde{X}} f_\varepsilon\left(\widetilde{s}(z,x)\right) \, f_x(x) \, \frac{\partial \widetilde{s}(z,x)}{\partial z} \, dx$$

To obtain the density of the demanded z, consider the transformation

$$z = d(y,\eta) \ \ \& \ \ y = y$$

The inverse of this transformation is

$$\eta = \widetilde{d}(z,y) \ \ \& \ \ y = y$$

and the Jacobian determinant is

$$\begin{vmatrix} \frac{\partial \widetilde{d}(z,y)}{\partial z} & \frac{\partial \widetilde{d}(z,y)}{\partial y} \\ 0 & 1 \end{vmatrix} = \left| \frac{\partial \widetilde{d}(z,y)}{\partial z} \right| = \frac{\partial \widetilde{d}(z,y)}{\partial z}$$

Let f_y and f_η denote the densities of the vector of observable and unobservable characteristics of the firms. Let $\widetilde{Y}$ denote the support of y. Then, the density of the demanded z is

$$\int_{\widetilde{Y}} f_\eta\left(\widetilde{d}(z,y)\right)\ f_y(y)\ \frac{\partial \widetilde{d}(z,y)}{\partial z}\ dy$$

The equilibrium condition is that the density of the demand equals the density of the supply, for all values of z ::

$$\int_{\widetilde{X}} f_\varepsilon\left(\widetilde{s}(z,x)\right)\ f_x(x)\ \frac{\partial \widetilde{s}(z,x)}{\partial z}\ dx = \int_{\widetilde{Y}} f_\eta\left(\widetilde{d}(z,y)\right)\ f_y(y)\ \frac{\partial \widetilde{d}(z,y)}{\partial z}\ dy$$

From the FOC of the consumer and firm, the functions $\widetilde{s}$ and $\widetilde{d}$ depend on the function P_z. Their derivatives depend then on P_z and P_{zz}. The equilibrium condition determines then a function P_z as a solution to a first order differential equation. This function will be the derivative of an equilibrium price function if the SOC of the consumer and firm are satisfied. To determine the conditions under which the SOC are satisfied, we substitute in the equilibrium equation the expression for the derivatives of the functions $\widetilde{s}$ and $\widetilde{d}$, to get

$$\begin{aligned}&\int_{\widetilde{X}} f_\varepsilon\left(\widetilde{s}(z,x)\right) f_x(x) \left[\frac{P_{zz}(z) - U_{zz}(z,x,\widetilde{s}(z,x))}{U_{z\varepsilon}(z,x,\widetilde{s}(z,x))}\right] dx \\ = &\int_{\widetilde{Y}} f_\eta\left(\widetilde{d}(z,y)\right) f_y(y) \left[\frac{P_{zz}(z) - \Gamma_{zz}(z,y,\widetilde{d}(z,y))}{\Gamma_{z\eta}(z,y,\widetilde{d}(z,y))}\right] dy\end{aligned}$$

or

$$\int_{\widetilde{X}}\int_{\widetilde{Y}} \left[\frac{f_\varepsilon P_{zz}\Gamma_{z\eta} - f_\varepsilon U_{zz}\Gamma_{z\eta} - f_\eta U_{z\varepsilon}P_{zz} + f_\eta U_{z\varepsilon}\Gamma_{zz}}{U_{z\varepsilon}\,\Gamma_{z\eta}}\right]\ f_x(x)\ f_y(y)\ dx\ dy = 0$$

or

$$\int_{\widetilde{X}}\int_{\widetilde{Y}} \left[\frac{P_{zz}\left(f_\varepsilon\Gamma_{z\eta} - f_\eta U_{z\varepsilon}\right) - f_\varepsilon U_{zz}\Gamma_{z\eta} + f_\eta U_{z\varepsilon}\Gamma_{zz}}{U_{z\varepsilon}\,\Gamma_{z\eta}}\right]\ f_x(x)\ f_y(y)\ dx\ dy = 0$$

So that

$$P_{zz} \int_{\widetilde{X}} \int_{\widetilde{Y}} \left[\frac{f_\varepsilon \Gamma_{z\eta} - f_\eta U_{z\varepsilon}}{U_{z\varepsilon}\,\Gamma_{z\eta}} \right] f_x(x)\; f_y(y)\; dx\; dy = \int_{\widetilde{X}} \int_{\widetilde{Y}} \left[\frac{f_\varepsilon U_{zz} \Gamma_{z\eta} - f_\eta U_{z\varepsilon} \Gamma_{zz}}{U_{z\varepsilon}\,\Gamma_{z\eta}} \right] f_x(x)\; f_y(y)\; dx\; dy$$

or

$$P_{zz} = \frac{\int_{\widetilde{X}} \int_{\widetilde{Y}} \left[\frac{f_\varepsilon U_{zz} \Gamma_{z\eta} - f_\eta U_{z\varepsilon} \Gamma_{zz}}{U_{z\varepsilon}\,\Gamma_{z\eta}} \right] f_x(x)\; f_y(y)\; dx\; dy}{\int_{\widetilde{X}} \int_{\widetilde{Y}} \left[\frac{f_\varepsilon \Gamma_{z\eta} - f_\eta U_{z\varepsilon}}{U_{z\varepsilon}\,\Gamma_{z\eta}} \right] f_x(x)\; f_y(y)\; dx\; dy}$$

The SOC of the consumer are satisfied if

$$P_{zz}(z) - U_{zz}(z, x, \widetilde{s}(z,x)) < 0$$

Substituting P_{zz} we get that the SOC of the consumer are satisfied when

$$\frac{\int_{\widetilde{X}} \int_{\widetilde{Y}} \left[\frac{f_\varepsilon U_{zz} \Gamma_{z\eta} - f_\eta U_{z\varepsilon} \Gamma_{zz}}{U_{z\varepsilon}\,\Gamma_{z\eta}} \right] f_x(x)\; f_y(y)\; dx\; dy}{\int_{\widetilde{X}} \int_{\widetilde{Y}} \left[\frac{f_\varepsilon \Gamma_{z\eta} - f_\eta U_{z\varepsilon}}{U_{z\varepsilon}\,\Gamma_{z\eta}} \right] f_x(x)\; f_y(y)\; dx\; dy} - U_{zz}(z, x, \widetilde{s}(z,x)) < 0$$

Similarly, for the firm, the SOC are satisfied when

$$\Gamma_{zz}(z, y, \widetilde{d}(z,y)) - \frac{\int_{\widetilde{X}} \int_{\widetilde{Y}} \left[\frac{f_\varepsilon U_{zz} \Gamma_{z\eta} - f_\eta U_{z\varepsilon} \Gamma_{zz}}{U_{z\varepsilon}\,\Gamma_{z\eta}} \right] f_x(x)\; f_y(y)\; dx\; dy}{\int_{\widetilde{X}} \int_{\widetilde{Y}} \left[\frac{f_\varepsilon \Gamma_{z\eta} - f_\eta U_{z\varepsilon}}{U_{z\varepsilon}\,\Gamma_{z\eta}} \right] f_x(x)\; f_y(y)\; dx\; dy} < 0$$

A necessary condition for the SOC of the consumer and firm to be satisfied for all z is that

$$\Gamma_{zz}(z, y, \widetilde{d}(z,y)) < U_{zz}(z, x, \widetilde{s}(z,x))$$

It is easy to verify that when there is only one heterogeneity variable, η, for the firm and only one heterogeneity variable, ε, for the consumer, the condition

$$\Gamma_{zz}(z, \widetilde{d}(z)) < U_{zz}(z, \widetilde{s}(z))$$

is necessary and sufficient for the SOC of both the firm and the consumer to be satisfied.

Consider, as a special case, the specification studied in Ekeland, Heckman, and Nesheim (2001) where, for some functions m_w, n_w, m_f, and n_f,

$$U_z(z, x, \varepsilon) = m_w(z) + n_w(x) - \varepsilon$$

$$\Gamma_z(z, y, \eta) = m_f(z) + n_f(y) + \eta$$

In this case,

$$\Gamma_{z\eta} = 1 \quad U_{z\varepsilon} = -1 \quad U_{zz} = m_w'(z) \quad and \;\; \Gamma_{zz} = m_f'(z)$$

$$\widetilde{s}(z,x) = m_w(z) - P_z(z) + n_w(x) \quad and \quad \widetilde{d}(z,y) = P_z(z) - m_f(z) - n_f(y)$$

Then,

$$P_{zz} = \frac{\int_{\widetilde{X}} \int_{\widetilde{Y}} \left[f_\varepsilon \left(\widetilde{s}\,(z,x) \right) m_w'(z) + f_\eta \left(\widetilde{d}\,(z,y) \right) m_f'(z) \right] \; f_x(x) \; f_y(y) \; dx \; dy}{\int_{\widetilde{X}} \int_{\widetilde{Y}} \left[f_\varepsilon \left(\widetilde{s}\,(z,x) \right) + f_\eta \left(\widetilde{d}\,(z,y) \right) \right] \; f_x(x) \; f_y(y) \; dx \; dy}$$

or

$$P_{zz} = \frac{m_w'(z) \int_{\widetilde{X}} f_\varepsilon \left(\widetilde{s}\,(z,x) \right) f_x(x) dx + m_f'(z) \int_{\widetilde{Y}} f_\eta \left(\widetilde{d}\,(z,y) \right) f_y(y) dy}{\int_{\widetilde{X}} f_\varepsilon \left(\widetilde{s}\,(z,x) \right) f_x(x) dx + \int_{\widetilde{Y}} f_\eta \left(\widetilde{d}\,(z,y) \right) f_y(y) dy}$$

Hence, the SOC of the consumer is satisfied when

$$\frac{m_w'(z) \int_{\widetilde{X}} f_\varepsilon \left(\widetilde{s}\,(z,x) \right) f_x(x) dx + m_f'(z) \int_{\widetilde{Y}} f_\eta \left(\widetilde{d}\,(z,y) \right) f_y(y) dy}{\int_{\widetilde{X}} f_\varepsilon \left(\widetilde{s}\,(z,x) \right) f_x(x) dx + \int_{\widetilde{Y}} f_\eta \left(\widetilde{d}\,(z,y) \right) f_y(y) dy} < m_w'(z)$$

Since the denominator is positive, this is equivalent to

$$m_f'(z) < m_w'(z)$$

Similarly, the SOC of the firm are satisfied when

$$m_f'(z) < \frac{m_w'(z) \int_{\widetilde{X}} f_\varepsilon \left(\widetilde{s}\,(z,x) \right) f_x(x) dx + m_f'(z) \int_{\widetilde{Y}} f_\eta \left(\widetilde{d}\,(z,y) \right) f_y(y) dy}{\int_{\widetilde{X}} f_\varepsilon \left(\widetilde{s}\,(z,x) \right) f_x(x) dx + \int_{\widetilde{Y}} f_\eta \left(\widetilde{d}\,(z,y) \right) f_y(y) dy}$$

which is also equivalent to

$$m_f'(z) < m_w'(z)$$

Hence, the SOC in the additive model are satisfied at any z if and only if

$$m'_f(z) < m'_w(z)$$

Clearly, the nonadditive model can generate a wider class of equilibria, since the condition for the SOC to be satisfied depends on the heterogeneity variables. In contrast, in the additive model, satisfaction of the SOC depends solely on a function of z. In Section 2.2, we present a nonadditive economy whose equilibrium exhibits bunching.

A different way of expressing the equilibrium condition is by using distribution functions instead of density functions. Let Z_w denote the supplied z and Z_f denote the demanded z. The equilibrium condition is that for all values z,

$$\Pr(Z_w \leq z) = \Pr(Z_f \leq z)$$

Assume that $U_{z\varepsilon} < 0$ and $\Gamma_{z\eta} > 0$. Then, for the consumer (worker),

$$\begin{aligned}
\Pr(Z_w \leq z) &= \Pr\left(s(X,\varepsilon) \leq z\right) \\
&= \int_{\widetilde{X}} \Pr\left(s(X,\varepsilon) \leq z | X = x\right) f_x(x)\, dx \\
&= \int_{\widetilde{X}} \Pr\left(\varepsilon \leq \widetilde{s}\,(z,x)\, |X = x\right)\, f_x(x)\, dx \\
&= \int_{\widetilde{X}} \Pr\left(\varepsilon \leq \widetilde{s}\,(z,x)\right)\, f_x(x)\, dx \\
&= \int_{\widetilde{X}} F_\varepsilon\left(\widetilde{s}\,(z,x)\right)\, f_x(x)\, dx
\end{aligned}$$

while for the firm

$$\begin{aligned}
\Pr(Z_f \leq z) &= \Pr\left(d(Y,\eta) \leq z\right) \\
&= \int_{\widetilde{X}} \Pr\left(d(Y,\eta) \leq z | Y = y\right) f_y(y)\, dy \\
&= \int_{\widetilde{X}} \Pr\left(\eta \leq \widetilde{d}\,(z,y)\, |Y = y\right)\, f_y(y)\, dy \\
&= \int_{\widetilde{X}} \Pr\left(\eta \leq \widetilde{d}\,(z,y)\right)\, f_y(y)\, dy \\
&= \int_{\widetilde{X}} F_\eta\left(\widetilde{d}\,(z,y)\right)\, f_y(y)\, dy
\end{aligned}$$

Hence, the equilibrium condition becomes

$$\int_{\widetilde{X}} F_\varepsilon\left(\widetilde{s}(z,x)\right)\, f_x(x)\, dx = \int_{\widetilde{X}} F_\eta\left(\widetilde{d}(z,y)\right)\, f_y(y)\, dy$$

which is a functional equation in P_z. If it were the case that $\Gamma_{z\eta} < 0$, the equilibrium condition would be

$$\int_{\widetilde{X}} F_\varepsilon\left(\widetilde{s}(z,x)\right)\, f_x(x)\, dx = \int_{\widetilde{X}} \left[1 - F_\eta\left(\widetilde{d}(z,y)\right)\right]\, f_y(y)\, dy$$

while if it were the case that $U_{z\varepsilon} > 0$, the equilibrium condition would be

$$\int_{\widetilde{X}} \left[1 - F_\varepsilon\left(\widetilde{s}(z,x)\right)\, f_x(x)\, dx\right] = \int_{\widetilde{X}} F_\eta\left(\widetilde{d}(z,y)\right)\, f_y(y)\, dy$$

2.1. AN ANALYTIC EXAMPLE

To provide a very simple analytic example of a nonadditive economy, suppose that all the heterogeneity across firms is represented by a scalar variable η and all the heterogeneity across consumers is represented by a scalar variable ε. Suppose that the consumer problem is

$\text{Max}_z\ P(z) - \frac{z^\beta}{\varepsilon}$

and the firm problem is

$\text{Max}_z\ z^\alpha \eta - P(z)$.

Suppose that ε is distributed $U(\varepsilon_l, \varepsilon_u)$, η is distributed $U(\eta_l, \eta_u)$, $\varepsilon_l = \eta_l$, and $\varepsilon_u = \eta_u$. Then, the first and second order conditions for the consumer's problem are

FOC: $P_z - \frac{\beta\, z^{\beta-1}}{\varepsilon} = 0$

SOC: $P_{zz} - \frac{\beta\,(\beta-1)\, z^{\beta-2}}{\varepsilon} < 0$

The first and second order conditions for the firm's problem are

FOC: $\alpha\, z^{\alpha-1}\eta - P_z = 0$

SOC: $\alpha\,(\alpha-1)\, z^{\alpha-2}\eta - P_{zz} < 0$.

The inverse supply and demand functions are

$$\varepsilon = \widetilde{s}(z) = \frac{\beta\, z^{\beta-1}}{P_z} \qquad \textit{and} \qquad \eta = \widetilde{d}(z) = \frac{P_z}{\alpha\; z^{\alpha-1}}$$

The equilibrium condition is

$$F_\varepsilon\left(\frac{\beta\, z^{\beta-1}}{P_z(z)}\right) = F_\eta\left(\frac{P_z(z)}{\alpha\; z^{\alpha-1}}\right)$$

which, using the assumption about the distributions of ε and η becomes

$$\frac{\beta\, z^{\beta-1}}{P_z(z)} = \frac{P_z(z)\, z^{1-\alpha}}{\alpha}$$

for all z such that $\varepsilon_l \leq \frac{\beta\, z^{\beta-1}}{P'(z)} \leq \varepsilon_u$. Hence, the equilibrium price function is

$$P_z(z) = \left(\alpha\, \beta\, z^{\alpha+\beta-2}\right)^{1/2}$$

for all z such that $\varepsilon_l^{\frac{2}{\beta-\alpha}} \left(\frac{\alpha}{\beta}\right)^{\frac{1}{\beta-\alpha}} \leq z \leq \varepsilon_u^{\frac{2}{\beta-\alpha}} \left(\frac{\alpha}{\beta}\right)^{\frac{1}{\beta-\alpha}}$. Substituting this equation into the first order conditions of the consumer and firm, it is easy to verify that the supply function of the consumer is

$$z = \left(\frac{\alpha}{\beta}\right)^{\frac{1}{\beta-\alpha}} \varepsilon^{\frac{2}{\beta-\alpha}}$$

for $\varepsilon_l \leq \varepsilon \leq \varepsilon_u$, and the demand function of the firm is

$$z = \left(\frac{\alpha}{\beta}\right)^{\frac{1}{\beta-\alpha}} \eta^{\frac{2}{\beta-\alpha}}$$

for $\eta_l \leq \eta \leq \eta_u$. Hence, in equilibrium, for each t between $\varepsilon_l = \eta_l$ and $\varepsilon_u = \eta_u$, each consumer with $\varepsilon = t$ gets matched with a firm with $\eta = t$. Using these equations and P_{zz} into the SOC of the consumer and firm, it is easy to verify that the SOC's are satisfied if and only if

$$\alpha < \beta$$

2.2. AN EXAMPLE OF AN EQUILIBRIUM WITH BUNCHING

Hedonic equilibrium models where the heterogeneity enters into the marginal utility and marginal product functions in nonadditive ways are capable of generating different types of equilibrium. In the simple example presented in the previous section, the heterogeneity variables of the consumer and the firm were continuously distributed and the resulting equilibrium z was also continuously distributed. We next present an example where the resulting equilibrium z is a mixed, continuous-discrete, random variable, even though the heterogeneity variables of the consumer and the firm are continuously distributed.

Suppose that each firm has a production function

$$\Gamma(z,\eta) = z^{\alpha}\,\eta$$

where $\alpha = .5$ and η is distributed $U(0,1)$. Each firm's problem is then

$Max_z\ z^{\alpha}\,\eta - P(z)$

FOC: $\alpha\, z^{\alpha-1}\eta - P_z(z) = 0$

SOC: $\alpha(\alpha-1)z^{\alpha-2}\eta - P_{zz}(z) < 0$

The FOC implies

$$\eta(z) = \frac{P_z(z)\ z^{1-\alpha}}{\alpha}$$

Suppose that each consumer has a disutility of z given by

$$V(z,\varepsilon) = z^{\varepsilon}$$

where ε is a random variable distributed $U(.25,.75)$. Each consumer's problem is then

$Max_z\ P(z) - V(z,\varepsilon)$

FOC: $P_z(z) - \varepsilon\ z^{\varepsilon-1} = 0$

SOC: $P_{zz}(z) - \varepsilon\,(\varepsilon-1)\,z^{\varepsilon-2} < 0.$

Applying the results from the previous section, we get that the SOC of the consumer and firm are satisfied if and only if

$$\Gamma_{zz}(z,\eta(z)) < V_{zz}(z,\varepsilon(z))$$

or, equivalently, when

$$\Gamma_{zz}(z(\eta),\eta) < V_{zz}(z(\varepsilon),\varepsilon) \quad \& \quad z(\eta) = z(\varepsilon) = z$$

Hence, in the given specification, the SOC are satisfied if and only if

$$\alpha(\alpha-1)z^{\alpha-2}\eta < \varepsilon\,(\varepsilon-1)\,z^{\varepsilon-2}$$

Using the FOC of the firm and the consumer, this last condition becomes

$$(\alpha-1)z^{-1}P_z(z) < (\varepsilon-1)\,z^{-1}P_z(z)$$

which is satisfied when

$$\alpha < \varepsilon$$

Since $\Pr(\varepsilon > \alpha) = .5$, a positive proportion of the market will locate at corner solutions. More specifically, in an equilibrium, a typical consumer with $\varepsilon \in (\alpha, \frac{3}{2}\alpha]$ will supply

$$z = \left(1 - \frac{\alpha}{2\varepsilon}\right)^{\frac{1}{\varepsilon-\alpha}}$$

and will get matched with a firm η such that

$$\eta = \frac{\varepsilon}{\alpha}\left(1 - \frac{\alpha}{2\varepsilon}\right)$$

Firms with $\eta < .5$ and $\varepsilon < .5$ will locate at $z = 0$.

3. IDENTIFICATION

In this section, we analyze the identification of the random marginal utility and marginal production functions in hedonic equilibrium models. We assume that the equilibrium price function and the distributions of (z,x) and (z,y) are given, where z denotes the observed location, x denotes the vector of observable characteristics of a typical consumer, and y denotes the vector of observable characteristics of a typical firm. We will consider here the identification of the marginal product function, $\Gamma_z(z,y,\eta)$ and of the distribution of η. The identification of the utilty $U^*(z,x,\varepsilon)$ and of the distribution of ε can be established in an analogous way, and is therefore omitted. We will consider the cases that require the minimal number of coordinates of y.

Theorem 3.1. *Let $y = (y_1, y_2) \in R^2$. Suppose that for some unknown twice differentiable function $m : R^2 \to R$, which is strictly increasing in its second argument, and some known differentiable function $q : R^2 \to R$*

$$(2) \quad \Gamma_z(z, y_1, y_2, \eta) = m\left(q(z, y_1), y_2 - \eta\right)$$

Normalize the function m, fixing its value at one point, so that for some values $\overline{z}$ of z, $\overline{y}_1$ of y_1, and $\alpha \in R$,

$$(3) \quad m\left(q(\overline{z}, \overline{y}_1), \alpha\right) = P_z\left(\overline{z}\right)$$

Let $\left[t_2^l, t_2^u\right]$ denote the support of $y_2 - \eta$, and for any $t_2 \in \left[t_2^l, t_2^u\right]$, let $[t_1^l(t_2), t_1^u(t_2)]$ denote the support of $q\left(d(y_1, y_2, \eta), y_1\right)$ conditional on $y_2 - \eta = t_2$. Then, for any (z, y_1, y_2, η) such that $y_2 - \eta \in \left[t_2^l, t_2^u\right]$ and $q(z, y_1) \in [t_1^l(t_2), t_1^u(t_2)]$

$$\Gamma_z(z, y_1, y_2, \eta) \text{ is identified}$$

Proof of Theorem 3.1: Since Γ_z is weakly separable in $y_2 - \eta$, the function $z = s(y_1, y_2, \eta)$, which satisfies the FOC is also weakly separable in $y_2 - \eta$. Hence, for some unknown function v

$$(4) \quad d(y_1, y_2, \eta) = v(y_1, y_2 - \eta)$$

Let y_2 and η be such that $y_2 - \eta = \alpha$. Then, by (2) and (3)

$$(5) \quad \Gamma_z(\overline{z}, \overline{y}_1, y_2, \eta) = P_z(\overline{z})$$

Hence, $\overline{z}$ satisfies the FOC when $y_1 = \overline{y}_1$ and $y_2 - \eta = \alpha$. It then follows that

$$(6) \quad v(\overline{y}_1, \alpha) = d(\overline{y}_1, y_2, \eta) = \overline{z}$$

By the definition of d,

$$(7) \; \frac{\partial s}{\partial y_2} = \frac{\Gamma_{zy_2}}{P_{zz} - \Gamma_{zz}}$$

By the SOC of the firm, $P_{zz} - \Gamma_{zz} > 0$. By (3) and the assumption that the function m is strictly increasing in its second coordinate, $\Gamma_{zy_2} > 0$. Hence, by (7) and (4), $v_2 > 0$, where v_2 denotes the derivative of v with respect to its second argument. Summarizing, the unknown function v that relates y_1, y_2, and η to the value of z that satisfies the FOC is such that

(8) $z = v(y_1, y_2 - \eta)$, v is strictly increasing in its second argument & $v(\overline{y}_1, \alpha) = \overline{z}$

By (8) and Matzkin (2002), the function v and the distribution of η are identified from the conditional cdf of z given (y_1, y_2). More specifically, let $\Gamma_{Z|Y=(y_1,y_2)}(z)$ denote the conditional cdf of z given (y_1, y_2). Let $F_\eta(\eta)$ denote the cdf of η. Then, for any e

$$(9) \quad F_\eta(e) = 1 - F_{Z|Y=(\overline{y}_1, \alpha+e)}(\overline{z})$$

and for any $\widetilde{y}_1, \widetilde{y}_2, \widetilde{e}$

$$(10) \quad v(\widetilde{y}_1, \widetilde{y}_2 - \widetilde{e}) = F^{-1}_{Z|Y=(\widetilde{y}_1, \widetilde{y}_2)}\left(1 - F_\eta(\widetilde{e})\right)$$

That is, the value of the cdf of η at e equals 1 minus the value of the conditional cdf of z given (y_1, y_2), when $z = \overline{z}$, $y_1 = \overline{y}_1$, and $y_2 = \alpha + e$. For any values $\widetilde{y}_1, \widetilde{y}_2, \widetilde{e}$, the value of the function v at $(\widetilde{y}_1, \widetilde{y}_2 - \widetilde{e})$ equals the value of the inverse of the conditional cdf of z given that $y_1 = \widetilde{y}_1$ and $y_2 = \widetilde{y}_2 - \widetilde{e}$, where this inverse function is evaluated at $1 - F_\eta(e)$.

Next, to show that the function m is identified, let (t_1, t_2) be any vector such that $t_2 \in \left[t_2^l, t_2^u\right]$ and $t_1 \in [t_1^l(t_2), t_1^u(t_2)]$. Let y_1^* denote a solution to

$$(11) \quad q(v(y_1^*, t_2), y_1^*) = t_1$$

Since q is a known function and v can be recovered from the conditional cdf of z given (y_1, y_2), the only unknown in (11) is y_1^*. Since $t_2 \in \left[t_2^l, t_2^u\right]$ and $t_1 \in [t_1^l(t_2), t_1^u(t_2)]$, y_1^* exists. Since $v(y_1^*, t_2)$ satisfies the FOC,

$$\begin{aligned}(13) \qquad m(t_1, t_2) &= m\left(q\left(v(y_1^*, t_2), y_1^*\right), t_2\right) \\ &= P_z\left(v(y_1^*, t_2)\right) \\ &= P_z\left(d(y_1^*, y_2^*, \eta^*)\right)\end{aligned}$$

for any y_2^* and η such that $y_2^* - \eta = \alpha$. In (13), the first equality follows because $q(v(y_1^*, t_2), y_1^*) = t_1$; the second equality follows because when z is substituted by the value that satisfies the first order conditions, the value of the marginal product function m equals the value of the marginal price function at the particular value of z that satisfies the first order conditions. The third equality follows by the restriction on the function d. Since the function P_z is known and the function d can be recovered without knowledge of m, (13) implies that the function m is identified.. This completes the proof of Theorem 3.1.

The statement and the proof of Theorem 1 can be easily modified to show that the function Γ_z is also identified when for some unknown function $m : R^2 \to R$ and some known function $q : R^2 \to R$, such that $F_{zy_2} > 0$

$$(14) \qquad F_z(z, y_1, y_2, \eta) = m(q(z, y_2 - \eta), y_1)$$

To see this, normalize the function m by requiring that for some values $\overline{z}$ of z, $\overline{y}_1$ of y_1, and $\alpha \in R$,

$$(15) \quad m\left(q(\overline{z}, \alpha), \overline{y}_1\right) = P_z\left(\overline{z}\right)$$

then, by (14), the demand function $z = d(y_1, y_2, \eta) = v(y_1, y_2 - \eta)$ for some unknown function v, and by (15),

$$(16) \quad v(\overline{y}_1, \alpha) = \overline{z}$$

By the assumption that $F_{zy_2} > 0$ and the SOC, v is strictly increasing in α. Using (16) together with the monotonicity of v with respect to α, we can use the results in Matzkin (2002) to identify and estimate the function v and the distribution of η. Next, to identify and estimate the function m at a point (t_1, t_2), find the value k^* that satisfies

$$(17) \ q\left(v(t_2, k^*), k^*\right) = t_1$$

Then,

$$(18) \quad m(t_1, t_2) = P_z\left(v(t_2, k^*)\right)$$

The following Theorem is based on a similar identification strategy, but it uses a different set of normalizations for the function F_z. It requires only one observable heterogeneity variable y.

Theorem 3.2. *: Let $y \in R$. Suppose that for some unknown function $m : R^2 \to R$ and some known function : $R^2 \to R$,*

$$(21) \qquad \Gamma_z(z, y, \eta) = m\left(q(z, y), \eta\right)$$

Assume that $\Gamma_{z\eta} > 0$. Use the function P_z to fix the value of the unknown function Γ_z at one value $\overline{y}$ of y, and on the 45 degree line on the (z, η) space, by requiring that for all t,

$$(22) \qquad \Gamma_z(t, \overline{y}, t) = P_z(t)$$

Let η be given. Let $q \in (q_l(\eta), q_u(\eta))$, the support of $q(d(y, \eta), y)$. Then,

$$\Gamma_z(z, y, \eta) \ \text{ is identified}$$

Proof of Theorem 3.2: By (22) it follows that the value of z that satisfies the FOC when $y = \overline{y}$ and $\eta = t$ is $z = t$. Hence, the demand function, $d(y, \eta)$, satisfies

$$(23) \qquad d(\overline{y}, \eta) = \eta$$

By the SOC and the assumption that $\Gamma_{z\eta} > 0$

$$(24) \qquad \frac{\partial d}{\partial \eta} = \frac{\Gamma_{z\eta}}{P_{zz} - \Gamma_{zz}} > 0$$

Then, we can use the methods in Matzkin (2002) to determine that

$$(25) \qquad F_\eta(e) = F_{Z|Y=\overline{y}}(e)$$

and

$$(26) \qquad d(\widetilde{y}, e) = F^{-1}_{Z|Y=\widetilde{y}}(F_\eta(e))$$

That is, the value of the cdf of η at any point e is given by the value of the conditional cdf of z given that $y = \overline{y}$, when this cdf is evaluated at $z = e$. The value of the function d at any point $(\widetilde{y}, e)$ is given by the value of the inverse function of the conditional cdf of z given that $Y = \widetilde{y}$, when the inverse function is evaluated at $F_\eta(e)$.

Next, to see that the function m is identified, let y^* denote the solution to

$$(27) \qquad q\,(d(y^*, t_2), y^*) = t_1$$

Hence,

$$(29)\quad m(t_1, t_2) = m\left(q\left(d(y^*, t_2), y^*\right), t_2\right)$$

Since d is the demand function, and the value of m equals that of the marginal product function, Γ_z, it follows by the FOC that

$$(30)\qquad m\left(q\left(d(y^*, t_2), y^*\right), t_2\right) = P_z\left(d(y^*, t_2)\right)$$

Hence,

$$(31)\qquad m(t_1, t_2) = P_z\left(s\left(y^*(t_1, t_2), t_2\right)\right)$$

This shows that the function m is identified and it completes the proof of Theorem 3.2.

Instead of specifying in proposition 2 that

$$(21)\qquad \Gamma_z(z, y, \eta) = m\left(q(z, y), \eta)\right)$$

and (22) hold, we could have specified, instead, that

$$(33)\qquad \Gamma_z(z, y, \eta) = m\left(q(z, \eta), y\right)$$

and (22) hold. The identification analysis would have been very similar. More specifically, if (33) and (22) hold, then, also in this case, the demand function satisfies

$$d(\overline{y}, \eta) = \eta$$

and is strictly increasing in η. So, the function d and the distribution of η are both identified from the conditional cdf of z given y. Next, to identify $m(t_1, t_2)$, let y^* denote a solution to

$$q\left(d(y^*, t_2), t_2\right) = t_1$$

Making use of the FOC,

$$m(t_1, t_2) = P_z\left(d(y^*, t_2)\right)$$

Using similar arguments, it is also possible to show that the function Γ_z is identified when a normalization is achieved by completely specifying the distribution of η, and where the function Γ is specified to equal either $m\left(q(z,y),\eta\right)$ or $m\left(q(z,\eta),y\right)$ for known q and unknown m.

4. ESTIMATION

The proofs of Theorems 3.1 and 3.2 provide ways of nonparametrically estimating the distribution of η, the demand function d, and the marginal product function Γ. Suppose that the assumptions in Theorem 1 are satisfied, so that the demand function has the form $v(y_1, y_{2-}\eta)$. To obtain an estimator for Γ, first estimate the distribution of η and the demand function v using the conditional cdf of z given (y_1, y_2), as described in Matzkin (2002). Then, use the estimated function $\widehat{v}$ and the known function q to calculate the value y_1^* that satisfies

$$q(\widehat{v}(y_1^*, t_2), y_1^*) = t_1$$

The estimator $\widehat{m}(t_1, t_2)$ of $m(t_1, t_2)$, is then given by the equation

$$\widehat{m}(t_1, t_2) = P_z\left(\widehat{v}(y_1^*(t_1, t_2), t_2)\right) \tag{13}$$

A similar procedure can be described using the steps in the proof of Theorem 3.2.

To describe the estimators suppose that the equilibrium price function is known, and that the available data is $\{Z^i, Y^i\}$ for each of N_1 firms, and $\{Z^i, X^i\}$ for each of N_2 consumers. For simplicity, we will concentrate on the estimation of the marginal product function for the case where the assumptions in Theorem 1 are satisfied. Let $f(z, y_1, y_2)$ and $F(z, y_1, y_2)$ denote, respectively, the joint pdf and cdf of (Z, Y_1, Y_2). Let $\hat{f}(z, y_1, y_2)$ and $\hat{F}(z, y_1, y_2)$ denote the corresponding kernel estimators. Let $\hat{f}_{Z|Y=(y_1,y_2)}(z)$ and $\hat{F}_{Z|Y=(y_1,y_2)}(z)$ denote the kernel estimators of, respectively, the conditional pdf and conditional cdf of Z given $Y = (y_1, y_2)$. Then,

$$\hat{f}(z, y_1, y_2) = \frac{1}{N\sigma_N^3} \sum_{i=1}^{N} K\left(\frac{z-Z^i}{\sigma}, \frac{y-Y^i}{\sigma}\right),$$

$$\hat{F}(z, y_1, y_2) = \int_{-\infty}^{z} \int_{-\infty}^{y} \hat{f}_N(s, y)\, ds\, dy,$$

$$\hat{f}_{Z|Y=y}(z) = \frac{\hat{f}_N(z,y)}{\int_{-\infty}^{\infty} \hat{f}_N(s,y)\, ds}, \text{ and}$$

$$\hat{F}_{Z|Y=y}(z) = \frac{\int_{-\infty}^{z} \hat{f}_N(s,y)\, ds}{\int_{-\infty}^{\infty} \hat{f}_N(s,y)\, ds}$$

where $y = (y_1, y_2)$, and where $K : R \times R^L \to R$ is a kernel function and σ_N is the bandwidth. Analogous equations hold when Y is substituted with X and $y = (y_1, y_2)$ is substituted with $x = (x_1, x_2)$. The above estimator for $F(z, y)$ was proposed in Nadaraya (1964). When $K(s, y) = k_1(s)k_2(y)$ for some kernel functions $k_1 : R \to R$ and $k_2 : R^2 \to R$,

$$\hat{F}_{Z|Y=y}(z) = \frac{\int_{-\infty}^{z} \hat{f}_N(s,y)\, ds}{\int_{-\infty}^{\infty} \hat{f}_N(s,y)\, ds} = \frac{\sum_{i=1}^{N} \widetilde{k}_1(\frac{z-z^i}{\sigma})\, k_2(\frac{y-Y^i}{\sigma})}{\sum_{i=1}^{N} k_2(\frac{y-Y^i}{\sigma})}$$

where $\widetilde{k}_1(u) = \int_{-\infty}^{u} k_1(s)\, ds$. Note that the estimator for the conditional cdf of Z given Y is different from the Nadaraya-Watson estimator for $F_{Z|Y=y}(z)$. The latter is the kernel estimator for the conditional expectation of $W \equiv 1[Z \le z]$ given $Y = y$. For any t and y, $\hat{F}^{-1}_{Z|Y=y}(t)$ will denote the set of values of Y for which $\hat{F}_{Z|Y=y}(z) = t$. When the kernel function k_1 is everywhere positive, this set of values will contain a unique point.

Suppose that the marginal product function is such that for some unknown function m

$$\Gamma_z(z, y_1, y_2, \eta) = m\left(q(z, y_1), y_2 - \eta\right)$$

where $q : R^2 \to R$ is some *known* function. Normalize the value of the function m at one point by requiring that at some values $\overline{z}$ of z, $\overline{y}_1$ of y_1, and $\alpha \in R$,

$$m\left(q(\overline{z}, \overline{y}_1), \alpha\right) = P_z\left(\overline{z}\right)$$

Let $d(y_1, y_2, \eta)$ be the function that satisfies, for each (y_1, y_2, η), the FOC of the firm. Then, as argued in the proof of Theorem 1,

$$d(y_1, y_2, \eta) = v(y_1, y_2 - \eta)$$

for some unknown function v, which is strictly increasing in its second coordinate and is such that

$$v(\overline{y}_1, \alpha) = d(\overline{y}_1, y_2, \eta) = \overline{z}$$

Using Matzkin (2002) it follows that for any e

$$\widehat{F}_{\eta}(e) = 1 - \widehat{F}_{Z|Y=(\overline{y}_1,\alpha+e)}(\overline{z})$$

and for any $\widetilde{y}_1, \widetilde{y}_2, \widetilde{e}$

$$\widehat{v}(\widetilde{y}_1, \widetilde{y}_2 - \widetilde{e}) = \widehat{F}^{-1}_{Z|Y=(\widetilde{y}_1,\widetilde{y}_2}\left(1 - \widehat{F}_{\eta}(\widetilde{e})\right)$$

As described above, to obtain an estimator for $m(t_1, t_2)$, we first calculate $\widehat{y}_1^*$ such that

$$(11) \quad q(\widehat{v}(\widehat{y}_1^*, t_2), \widehat{y}_1^*) = t_1$$

and then let

$$\widehat{m}(t_1, t_2) = P_z\left(\widehat{v}(\widehat{y}_1^*, t_2)\right)$$

We will establish the asymptotic properties of this estimator for the case where the function $q(z, y) = z \cdot y$. We will make the following assumptions:

Assumption A.1: *The sequence* $\left\{Z^i, Y^i\right\}$ *is i.i.d.*

Assumption A.2: $f(Z, Y)$ *has compact support* $\Theta \subset R^3$ *and is twice continuously differentiable.*

Assumption A.3: *The kernel function* $K(\cdot,\cdot)$ *is Lipschitz, vanishes outside a compact set, integrates to 1, and is of order 2.*

Assumption A.4: *As* $N \to \infty$, $\ln(N)/N\sigma_N^3 \to 0$ *and* $\sigma_N^2\sqrt{N\sigma_N^4} \to 0$

Assumption A.5: $0 < f(y_1, y_2), f(\overline{y}_1, \alpha + e) < \infty$; *there exist* $\delta, \xi > 0$ *such that* $\forall s \in N(v(y_1^*, y_2 - e), \xi)$, $f(s, y) \geq \delta$; *there exist* $\delta', \xi' > 0$ *such that* $\forall (s, y_1) \in N\left((v(y_1^*, y_2 - e), y_1^*), \xi'\right)$ $f(s, y) \geq \delta$, $\partial v(y_1^*, t_2) \neq 0$, $F_{Z|Y=(y_1^*,y_2)}(v(y_1^*, y_2 - e)) \neq 0$.

Assumption A.6: $t_2 = y_2 - e$ for some y_2 in the interior of the support of Y_2 and some e in the interior of the support of η; t_1 belongs to the interior of the support of $q\left(d(y_1, y_2, \eta), y_1\right)$ conditional on $Y_2 = y_2$ and $\eta = e$..

Let $\int K(z)^2 = \int \left(\int K(s,y)\, ds\right)^2 dy$, where $s \in R$. When assumptions A.1-A.5 are satisfied, Theorems 1 and 2 in Matzkin (2002) imply that for any e and (y_1', y_2'),

$$\sup_{e \in R} \left| \widehat{F}_\eta(e) - F_\eta(e) \right| \to 0 \quad \& \quad \widehat{v}(y_1', y_2' - e) \to v(y_1', y_2' - e) \quad \text{in probability}$$

and that

$$\sqrt{N}\sigma \left(\widehat{F}_\eta(e) - F_\eta(e) \right) \to N\left(0, V_F\right) \quad \&$$

$$\sqrt{N}\sigma_N \left(\widehat{v}(y_1', y_2' - e) - v(y_1', y_2' - e) \right) \to N\left(0, V_v\right)$$

where

$$V_F = \left\{ \int K(z)^2 \right\} \left[F_\eta(e)\, (1 - F_\eta(e) \right] \left[\frac{1}{f(\overline{y}_1, \alpha + e)} \right]$$

and

$$V_n = \left\{ \int K(z)^2 \right\} \left[\frac{F_\eta(e)\, (1 - F_\eta(e))}{f_{Z|Y=y}(v(y_1', y_2' - e))^2} \right] \left[\frac{1}{f(\overline{y}_1, e + \alpha)} + \frac{1}{f(y_1', y_2')} \right]$$

The next theorem uses assumptions A.1-A.6 to establish the asymptotic properties of $\widehat{m}(t_1, t_2)$. Let $y = (y_1^*, y_2)$ and $v^* = v(y_1^*, y_2 - e)$ for y_2 and e such that $y_2 - e = t_2$. Let $\widetilde{y} = (\overline{y}_1, \alpha + e)$. Define the constant C by

$$C = \left(\frac{1}{f_{Z|Y=y}\left(v^*\right)} \right) \left(1 + \left[\frac{y_1^* \, \frac{\partial F_{Z|Y=y}(v^*)}{\partial y_1}}{f_{Z|Y=y}\left(v^*\right) \, \frac{\partial q(v(y_1^*, t_2), y_1^*)}{\partial y_1}} \right] \right)$$

Theorem 4.1. *Suppose that Assumptions A.1-A.6 are satisfied. Then, $\widehat{m}(t_1, t_2)$ converges in probability to $m(t_1, t_2)$ and*

$\sqrt{N\sigma^4} \left(\widehat{m}(t_1, t_2) - m(t_1, t_2) \right) \to N(0, V_m)$ *in distribution, where*

$V_m = \left\{ \int K(z)^2 \right\} [C]^2 \left(P_{zz}\left(v^*\right) \right)^2 \left(\frac{1}{f(\widetilde{y})} + \frac{1}{f(y)} \right) \left(F_{Z|Y=\widetilde{y}}(\overline{z})(1 - F_{Z|Y=\widetilde{y}}(\overline{z})) \right)$

The proof of this Theorem is presented in the Appendix.

5. SUMMARY

We have considered hedonic equilibrium models where the marginal utility of each consumer and the marginal product of each firm are both nonadditive functions of the attribute and a random vector of individual characteristics, which are different for the consumers and firms. We have demonstrated that this type of specification is capable of generating equilibria of different types, with and without bunching. We have shown that when the vector of individual characteristics contains an observable characteristic, it is possible to identify the nonadditive random marginal utility and nonadditive random marginal product. We have provided nonparametric estimators for these functions and have shown that they are consistent and asymptotically normal.

6. APPENDIX

Proof of Theorem 3: We use a Delta Method, like the ones developed in Ait-Sahalia (1994) and Newey (1994). (See Matzkin (1999) and Altoniji and Matzkin (1997) for other applications of this method.) Let $F(z,y)$ denote the distribution function (cdf) of the vector of observable variables (Z,Y), $f(z,y)$ denote its probability density function (pdf), $f(y)$ denote the marginal pdf of Y, and $F_{Z|Y=y}$ denote the conditional cdf of Z given $Y=y$. Let $L=3$. For any function $G: R^{1+L} \to R$, define $g(z,y) = \partial^L G(z,y)/\partial z \partial y$, $g(y) = \int g(z,y)\, dz$, $g(y) = \int g(z,y)\, dz$, $G_{Z|Y=y'}(z') = \left(\int_{-\infty}^{z'} g(z,y')dz \right) / g(y')$, and $\widetilde{G_Z}(z,y) = \int^z g(s,y)ds = \int 1[s \leq z]\; g(s,y)\; ds$ where $1[\cdot] = 1$ if $[\cdot]$ is true, and it equals zero otherwise. Let $\underline{\mathrm{C}}$ denote a compact set in R^L that strictly includes Θ. Let D denote the set of all functions $G: R^L \to R$ such that $g(z,y)$ vanishes outside $\underline{\mathrm{C}}$. Let D denote the set of all functions $\widetilde{G_Z}$ that are derived from some G in D. Since there is a 1-1 relationship between functions in D and functions in D, we can define a functional on D or on D without altering its definition. Let $\|G\|$ denote the sup norm of $g(z,y)$. Then, if $H \in$D, there exists $\rho_1 > 0$ such that if $\|H\| \leq \rho_1$ then, for some $0 < a, b < \infty$, all y and all $s \in N(v(y_1, y_2 - e), \xi)$,

$$(1)\; |h(y)| \leq a\,\|H\|, \left| \int_{-\infty}^{y} h(s,y)ds \right| \leq a\,\|H\|,$$

$$|f(y) + h(y)| \geq b\,|f(y)|, \text{and} \;\; f(s,y) + h(s,y) \geq b\,|f(s,y)|.$$

Let $\widetilde{z} = \overline{z}$, $\widetilde{y} = (\overline{y}_1, e+\alpha)$ and $v^* = v(y_1^*, t_2)$.Let y_2 and e be such that $y - e = \alpha$. Define the functionals

$$\kappa(G) = y_1 \text{ such that } \ v(y_1, t) \cdot y_1 = t_1$$

$$\Phi(G) = v(y_1^*, t) = G^{-1}_{Z|Y=(y_1^*,t)} \left(G^{-1}_{Z|Y=\widetilde{y}}(\widetilde{z}) \right)$$

$$\Lambda(G) = v(\kappa(G), t) = G^{-1}_{Z|Y=(\kappa(G),t)} \left(G_{Z|Y=\widetilde{y}}(\widetilde{z}) \right)$$

$$\delta(G) = \Lambda(G) \cdot \kappa(G)$$

$$\beta(G) = t_1$$

$$\eta(G) = G_{Z|Y=(\kappa(G),t)}(\Lambda(G))$$

$$\nu(G) = G_{Z|Y=\widetilde{y}}(\widetilde{z})$$

Then, $v(y_1^*, t_2) = F^{-1}_{Z|Y=(y_1^*,y_2)} \left(F_{Z|Y=\widetilde{y}}(\widetilde{e}) \right)$ and $\eta(G) = \nu(G)$ for all G.

$$\eta(F+H) - \eta(F) = \frac{\int^{\Lambda(F+H)} f(s,\kappa(F+H),y_2) + h(s,\kappa(F+H),y_2)}{f(\kappa(F+H),y_2) + h(\kappa(F+H),y_2)} - \frac{\int^{\Lambda(F)} f(s,\kappa(F),y_2)}{f(\kappa(F),y_2)}$$

$$= \frac{f(\kappa(F),y_2) \int^{\Lambda(F+H)} f(s,\kappa(F+H),y_2)ds + f(\kappa(F),y_2) \int^{\Lambda(F+H)} h(s,\kappa(F+H),y_2)ds}{f(\kappa(F+H),y_2)(f(\kappa(F+H),y_2) + h(\kappa(F+H),y_2))}$$

$$- \frac{f(\kappa(F+H),y_2) \int^{\Lambda(F)} f(s,\kappa(F),y_2) + h(\kappa(F+H),y_2) \int^{\Lambda(F)} f(s,\kappa(F),y_2)}{f(\kappa(F+H),y_2)(f(\kappa(F+H),y_2) + h(\kappa(F+H),y_2))}$$

By the Mean Value Theorem

$$f\left(\kappa(F+H), y_2\right) = f\left(\kappa(F), y_2\right) + \frac{\partial f(\widetilde{y}_1,y_2)}{\partial y_1} \left(\kappa(F+H) - \kappa(F)\right)$$

$$h\left(\kappa(F+H), y_2\right) = h\left(\kappa(F), y_2\right) + \frac{\partial h\left(\widetilde{y}_1^a,y_2\right)}{\partial y_1} \left(\kappa(F+H) - \kappa(F)\right)$$

$$f\left(s, \kappa(F+H), y_2\right) = f\left(s, \kappa(F), y_2\right) + \frac{\partial f\left(s,\widetilde{y}_1'(s),y_2\right)}{\partial y_1} \left(\kappa(F+H) - \kappa(F)\right)$$

and $h\left(s, \kappa(F+H), y_2\right) = h\left(s, \kappa(F), y_2\right) + \frac{\partial h\left(s,\widetilde{y}_1''(s),y_2\right)}{\partial y_1} \left(\kappa(F+H) - \kappa(F)\right)$

for some $\widetilde{y}_1, \widetilde{y}_1^a, \widetilde{y}_1'(s)$, and $\widetilde{y}_1''(s)$ between $\kappa(F+H)$ and $\kappa(F)$. Hence,

$$\eta(F+H) - \eta(F) = \frac{f(\kappa(F),y_2) \int^{\Lambda(F+H)} f(s,\kappa(F),y_2)ds + f(\kappa(F),y_2)(\kappa(F+H) - \kappa(F)) \int^{\Lambda(F+H)} \frac{\partial f\left(s,\widetilde{y}_1'(s),y_2\right)}{\partial y_1} ds}{f(\kappa(F),y_2)(f(\kappa(F+H),y_2) + h(\kappa(F+H),y_2))}$$

$$+ \frac{f(\kappa(F),y_2) \int^{\Lambda(F+H)} h(s,\kappa(F),y_2)ds + f(\kappa(F),y_2)(\kappa(F+H) - \kappa(F)) \int^{\Lambda(F+H)} \frac{\partial h\left(s,\widetilde{y}_1''(s),y_2\right)}{\partial y_1} ds}{f(\kappa(F),y_2)(f(\kappa(F+H),y_2) + h(\kappa(F+H),y_2))}$$

$$-\frac{f(\kappa(F),y_2)\int^{\Lambda(F)} f(s,\kappa(F),y_2)ds+\frac{\partial f(\widetilde{y}_1,y_2)}{\partial y_1}(\kappa(F+H)-\kappa(F))\int^{\Lambda(F)} f(s,\kappa(F),y_2)ds}{f(\kappa(F),y_2)(f(\kappa(F+H),y_2)+h(\kappa(F+H),y_2))}$$

$$-\frac{h(\kappa(F),y_2)\int^{\Lambda(F)} f(s,\kappa(F),y_2)ds+\frac{\partial h(\widetilde{y}_1^a,y_2)}{\partial y_1}(\kappa(F+H)-\kappa(F))\int^{\Lambda(F)} f(s,\kappa(F),y_2)ds}{f(\kappa(F),y_2)(f(\kappa(F+H),y_2)+h(\kappa(F+H),y_2))}$$

Also by the Mean Value Theorem,

$$\int^{\Lambda(F+H)} f\left(s,\kappa(F),y_2\right)ds=\int^{\Lambda(F)} f\left(s,\kappa(F),y_2\right)ds+f\left(v,\kappa(F),y_2\right)\left(\Lambda(F+H)-\Lambda(F)\right)$$

$$\int^{\Lambda(F+H)} h\left(s,\kappa(F),y_2\right)ds=\int^{\Lambda(F)} h\left(s,\kappa(F),y_2\right)ds+h\left(v',\kappa(F),y_2\right)\left(\Lambda(F+H)-\Lambda(F)\right)$$

$$\int^{\Lambda(F+H)} \frac{\partial f\left(s,\widetilde{y}_1'(s),y_2\right)}{\partial y_1}ds=\int^{\Lambda(F)} \frac{\partial f\left(s,\widetilde{y}_1'(s),y_2\right)}{\partial y_1}ds+\frac{\partial f\left(v'',\widetilde{y}_1'(v''),y_2\right)}{\partial y_1}\left(\Lambda(F+H)-\Lambda(F)\right)$$

and

$$\int^{\Lambda(F+H)} \frac{\partial h\left(s,\widetilde{y}_1''(s),y_2\right)}{\partial y_1}ds=\int^{\Lambda(F)} \frac{\partial h\left(s,\widetilde{y}_1''(s),y_2\right)}{\partial y_1}ds+\frac{\partial h\left(v^a,\widetilde{y}_1'(v^a),y_2\right)}{\partial y_1}\left(\Lambda(F+H)-\Lambda(F)\right)$$

for some v,v',v'',v^a between $\Lambda(F+H)$ and $\Lambda(F)$. Hence,

$$\eta(F+H)-\eta(F)$$

$$=\frac{f(\kappa(F),y_2)\int^{\Lambda(F)} f(s,\kappa(F),y_2)ds+f(\kappa(F),y_2)(\Lambda(F+H)-\Lambda(F))f(v,\kappa(F),y_2)ds}{f(\kappa(F),y_2)(f(\kappa(F+H),y_2)+h(\kappa(F+H),y_2))}$$

$$+\frac{f(\kappa(F),y_2)(\kappa(F+H)-\kappa(F))\int^{\Lambda(F)} \frac{\partial f(s,\widetilde{y}_1'(s),y_2)}{\partial y_1}ds+f(\kappa(F),y_2)(\kappa(F+H)-\kappa(F))(\Lambda(F+H)-\Lambda(F))\frac{\partial f(v'',\widetilde{y}_1'(v''),y_2)}{\partial y_1}}{f(\kappa(F),y_2)(f(\kappa(F+H),y_2)+h(\kappa(F+H),y_2))}$$

$$+\frac{f(\kappa(F),y_2)\int^{\Lambda(F)} h(s,\kappa(F),y_2)ds+f(\kappa(F),y_2)(\Lambda(F+H)-\Lambda(F))h(v',\kappa(F),y_2)}{f(\kappa(F),y_2)(f(\kappa(F+H),y_2)+h(\kappa(F+H),y_2))}$$

$$+\frac{f(\kappa(F),y_2)(\kappa(F+H)-\kappa(F))\int^{\Lambda(F)} \frac{\partial h(s,\widetilde{y}_1''(s),y_2)}{\partial y_1}ds+f(\kappa(F),y_2)(\kappa(F+H)-\kappa(F))(\Lambda(F+H)-\Lambda(F))\frac{\partial h(v^a,\widetilde{y}_1'(v^a),y_2)}{\partial y_1}}{f(\kappa(F),y_2)(f(\kappa(F+H),y_2)+h(\kappa(F+H),y_2))}$$

$$-\frac{f(\kappa(F),y_2)\int^{\Lambda(F)} f(s,\kappa(F),y_2)ds+\frac{\partial f(\widetilde{y}_1,y_2)}{\partial y_1}(\kappa(F+H)-\kappa(F))\int^{\Lambda(F)} f(s,\kappa(F),y_2)ds}{f(\kappa(F),y_2)(f(\kappa(F+H),y_2)+h(\kappa(F+H),y_2))}$$

$$-\frac{h(\kappa(F),y_2)\int^{\Lambda(F)} f(s,\kappa(F),y_2)ds+\frac{\partial h(\widetilde{y}_1^a,y_2)}{\partial y_1}(\kappa(F+H)-\kappa(F))\int^{\Lambda(F)} f(s,\kappa(F),y_2)ds}{f(\kappa(F),y_2)(f(\kappa(F+H),y_2)+h(\kappa(F+H),y_2))}$$

$$=\frac{f(\kappa(F),y_2)(\Lambda(F+H)-\Lambda(F))f(v,\kappa(F),y_2)ds}{f(\kappa(F),y_2)(f(\kappa(F+H),y_2)+h(\kappa(F+H),y_2))}$$

$$+\frac{f(\kappa(F),y_2)(\kappa(F+H)-\kappa(F))\int^{\Lambda(F)} \frac{\partial f(s,\widetilde{y}_1'(s),y_2)}{\partial y_1}ds+f(\kappa(F),y_2)(\kappa(F+H)-\kappa(F))(\Lambda(F+H)-\Lambda(F))\frac{\partial f(v'',\widetilde{y}_1'(v''),y_2)}{\partial y_1}}{f(\kappa(F),y_2)(f(\kappa(F+H),y_2)+h(\kappa(F+H),y_2))}$$

$$+\frac{f(\kappa(F),y_2)\int^{\Lambda(F)}h(s,\kappa(F),y_2)ds+f(\kappa(F),y_2)(\Lambda(F+H)-\Lambda(F))h(v',\kappa(F),y_2)}{f(\kappa(F),y_2)(f(\kappa(F+H),y_2)+h(\kappa(F+H),y_2))}$$

$$+\frac{f(\kappa(F),y_2)(\kappa(F+H)-\kappa(F))\int^{\Lambda(F)}\frac{\partial h(s,\widetilde{y}_1''(s),y_2)}{\partial y_1}ds+f(\kappa(F),y_2)(\kappa(F+H)-\kappa(F))(\Lambda(F+H)-\Lambda(F))\frac{\partial h(v^a,\widetilde{y}_1'(v^a),y_2)}{\partial y_1}}{f(\kappa(F),y_2)(f(\kappa(F+H),y_2)+h(\kappa(F+H),y_2))}$$

$$-\frac{\frac{\partial f(\widetilde{y}_1,y_2)}{\partial y_1}(\kappa(F+H)-\kappa(F))\int^{\Lambda(F)}f(s,\kappa(F),y_2)ds}{f(\kappa(F),y_2)(f(\kappa(F+H),y_2)+h(\kappa(F+H),y_2))}$$

$$-\frac{h(\kappa(F),y_2)\int^{\Lambda(F)}f(s,\kappa(F),y_2)ds+\frac{\partial h(\widetilde{y}_1^a,y_2)}{\partial y_1}(\kappa(F+H)-\kappa(F))\int^{\Lambda(F)}f(s,\kappa(F),y_2)ds}{f(\kappa(F),y_2)(f(\kappa(F+H),y_2)+h(\kappa(F+H),y_2))}$$

We next obtain an expression for $\kappa(F+H)-\kappa(F)$. By the definition of κ, for all G,

$$\left[G^{-1}_{Z|Y=(\kappa(G),y_2)}(G_{Z|Y=\widetilde{y}}(\widetilde{z}))\right]\kappa(G)=t_1$$

Denote $(F+H)^{-1}_{Z|Y=(\kappa(F+H),y_2)}((F+H)_{Z|Y=\widetilde{y}}(\widetilde{z}))$ by $F'_{K'}$, $(F)^{-1}_{Z|Y=(\kappa(F+H),y_2)}(F)_{Z|Y=\widetilde{y}}(\widetilde{z}))$ by $F_{K'}$, $(F)^{-1}_{Z|Y=(\kappa(F),y_2)}(F)_{Z|Y=\widetilde{y}}(\widetilde{z}))$ by F_K, $\kappa(F+H)$ by K', and $\kappa(F)$ by K. Then, since $\beta(F+H)-\beta(F)=0$,

$$F'_{K'}\cdot K'-F_K\cdot K=0.$$

Hence

$$(*)\ \ 2\cdot\left(F'_{K'}-F_{K'}\right)\cdot(K'-K)+\left(F'_{K'}-F_{K'}\right)\cdot K$$

$$+F_K\cdot(K'-K)+(F_{K'}-F_K)\cdot K=0$$

By the proof of Theorem 2 in Matzkin (2002), it follows that, for $y=(y_1^*,y_2)$

$$F'_K-F_K=\Phi(F+H)-\Phi(F)=D\Phi+R\Phi$$

where

$$D\Phi=\frac{f(y)}{f(\widetilde{y})^2f(\Phi(F),w)}A\widetilde{y}-\frac{f(y)}{f(y)^2f(\Phi(F),y)}Ay,$$

$$A\widetilde{y}=f(\widetilde{y})\int^{\widetilde{z}}h(s,\widetilde{y})ds-h(\widetilde{y})\int^{\widetilde{z}}f(s,\widetilde{y})ds$$

$$Ay=f(y)\int^{\Phi(F)}h(s,y)ds-h(y)\int^{\Phi(F)}f(s,y)ds$$

and where for some large enough constants a_4 and a_6

$$|R\Phi(F,H)| \leq a_4 \|H\|^2 \text{ and } |D\Phi(F,H| \leq a_6 \|H\|.$$

Using similar arguments, it is easy to show that for some reminder R that is bounded by a constant times $\|H\|^2$

$$F'_{K'} - F_{K'}$$

$$= (F+H)^{-1}_{Z|Y=(\kappa(F+H),t)} \left((F+H)_{Z|Y=\widetilde{y}}(\widetilde{z})\right) - (F)^{-1}_{Z|Y=(\kappa(F+H),t)} \left((F)_{Z|Y=\widetilde{y}}(\widetilde{z})\right)$$

$$= \frac{f(y)}{f(\widetilde{y})^2 \ f(v^*,y)} \left[f(\widetilde{y}) \int^{\widetilde{z}} h(s,\widetilde{y})ds - h(\widetilde{y}) \int^{\widetilde{z}} f(s,\widetilde{y})ds\right]$$

$$- \frac{f(y)}{f(y)^2 \ f(v^*,y)} \left[f(y) \int^{v^*} h\,(s,y)\,ds - h\,(y) \int^{v^*} f(s,y)ds\right] + R$$

where the first two terms are bounded by the multiplication of a constant multiplied by $\|H\|$. This together with the assumption that $F_K \neq 0$ and the expression in (*) imply that

$$\|K' - K\| \leq a \|H\|$$

for some positive constant a. By the Mean Value Theorem, there exists c between $\kappa(F+H)$ and $\kappa(F)$ such that

$$F_{K'} - F_K$$

$$= (F)^{-1}_{Z|Y=(\kappa(F+H),t)} \left((F)_{Z|Y=\widetilde{y}}(\widetilde{z})\right) - (F)^{-1}_{Z|Y=(\kappa(F),t)} \left((F)_{Z|Y=\widetilde{y}}(\widetilde{z})\right)$$

$$= v\,(K(F+H), t_2) - v\,(K(F), t_2)$$

$$= \frac{\partial v(c,t_2)}{\partial y_1} \left(\kappa(F+H) - \kappa(F)\right)$$

Hence, for some reminder term R that is bounded by a scalar times $\|H\|^2$,

$$\frac{f(y)}{f(\widetilde{y})^2 \ f(v^*,y)} \left[f(\widetilde{y}) \int^{\widetilde{z}} h(s,\widetilde{y})ds - h(\widetilde{y}) \int^{\widetilde{z}} f(s,\widetilde{y})ds\right] \cdot y_1^*$$

$$- \frac{f(y)}{f(y)^2 \ f(v^*,y)} \left[f(y) \int^{v^*} h\,(s,y)\,ds - h\,(y) \int^{v^*} f(s,y)ds\right] \cdot y_1^*$$

$$+ v^* \cdot \left(\kappa(F+H) - \kappa(F)\right) + y_1^* \frac{\partial v(c,t_2)}{\partial y_1} \left(\kappa(F+H) - \kappa(F)\right) + R = 0$$

It follows that, for some R such that for some $a > 0$, $|R| \leq a \|H\|^2$

$$\kappa(F+H) - \kappa(F)$$

$$= \frac{(Ay - A\widetilde{y}) \frac{y_1^*}{f_{Z|Y=y}(v^*)} \cdot}{v^* + y_1^* \frac{\partial v(c,t_2)}{\partial y_1}} + R$$

Substituting this into the expression for $\eta(F+H) - \eta(F)$, rewriting the resulting expression, and putting all terms of order less than $\|H\|^2$ into the term R, we get that

$$\eta(F+H) - \eta(F)$$

$$= f_{Z|Y=y}(v^*) (\Lambda(F+H) - \Lambda(F)) + R$$

$$+Ay$$

$$+ \left[\frac{f(y) \int^{v^*} \frac{\partial f(s,y)}{\partial y_1} ds - \frac{\partial f(\widetilde{y}_1, y_2)}{\partial y_1} \int^{v^*} f(s,y) ds}{f(y)^2} \right] \left[\frac{(Ay - A\widetilde{y}) \frac{y_1^*}{f_{Z|Y=y}(v^*)} \cdot}{v^* + y_1^* \frac{\partial v(c,t_2)}{\partial y_1}} \right]$$

By the definition of ν,

$$\begin{aligned}
\nu(F+H) - \nu(F) &= (F+H)_{Z|Y=\widetilde{y}}(\widetilde{z}) - F_{Z|Y=\widetilde{y}}(\widetilde{z}) \\
&= \frac{\int^{\widetilde{z}} f(s,\widetilde{y})\, ds + \int^{\widetilde{z}} h(s,\widetilde{y})\, ds}{f(\widetilde{y}) + h(\widetilde{y})} - \frac{\int^{\widetilde{z}} f(s,\widetilde{y})\, ds}{f(\widetilde{y})} \\
&= \frac{f(\widetilde{y}) \int^{\widetilde{z}} h(s,\widetilde{y})\, ds - h(\widetilde{y}) \int^{\widetilde{z}} f(s,\widetilde{y})\, ds}{f(\widetilde{y})^2} + R_4 \\
&= A\widetilde{y} + R_4
\end{aligned}$$

where R_4 is of order $\|H\|^2$. Hence, since

$$\eta(F+H) - \eta(F) = \nu(F+H) - \nu(F),$$

$$f_{Z|Y=y}(v^*) (\Lambda(F+H) - \Lambda(F)) + R$$

$$+Ay$$

$$+\left[\frac{f(y)\int^{v^*}\frac{\partial f(s,y)}{\partial y_1}ds-\frac{\partial f(\widetilde{y}_1,y_2)}{\partial y_1}\int^{v^*}f(s,y)ds}{f(y)^2}\right]\left[\frac{(Ay-A\widetilde{y})\frac{y_1^*}{f_{Z|Y=y}(v^*)}\cdot}{v^*+y_1^*\frac{\partial v(c,t_2)}{\partial y_1}}\right]$$
$$=A\widetilde{y}+R_5$$

Hence,

$$\Lambda(F+H)-\Lambda(F)$$
$$=(A\widetilde{y}-Ay)\left(1+\left[\frac{f(y)\int^{v^*}\frac{\partial f(s,y)}{\partial y_1}ds-\frac{\partial f(\widetilde{y}_1,y_2)}{\partial y_1}\int^{v^*}f(s,y)ds}{f(y)^2}\right]\left[\frac{\frac{y_1^*}{\left(f_{Z|Y=y}(v^*)\right)^2}\cdot}{v^*+y_1^*\frac{\partial v(c,t_2)}{\partial y_1}}\right]\right)+R_5$$

for some R_5 of order $\|H\|^2$. Note that

$$\frac{\partial F_{Z|Y=y}(v^*)}{\partial y_1}=\left[\frac{f(y)\int^{v^*}\frac{\partial f(s,y)}{\partial y_1}ds-\frac{\partial f(\widetilde{y}_1,y_2)}{\partial y_1}\int^{v^*}f(s,y)ds}{f(y)^2}\right]$$

and that

$$\frac{\partial q(v(y_1^*,t_2),y_1^*)}{\partial y_1}=v^*+y_1^*\frac{\partial v(c,t_2)}{\partial y_1}$$

Hence,

$$\Lambda(F+H)-\Lambda(F)$$
$$=\frac{(A\widetilde{y}-Ay)}{f_{Z|Y=y}(v^*)}\left(1+\left[\frac{\frac{\partial F_{Z|Y=y}(v^*)}{\partial y_1}\frac{y_1^*}{f_{Z|Y=y}(v^*)}}{v^*+y_1^*\frac{\partial v(c,t_2)}{\partial y_1}}\right]\right)+R_5$$

It follows that

$$\Lambda(F+H)-\Lambda(F)=D\Delta(F;H)+R\Delta(F;H)$$

where for some scalar b,

$|D\Delta(F;H)|\leq b\|H\|$, $|R\Delta(F;H)|\leq b\|H\|^2$ and

$$D\Delta(F;H) = (A\widetilde{y} - Ay)\left(\frac{1}{f_{Z|Y=y}(v^*)}\right)\left(1 + \left[\frac{\frac{\partial F_{Z|Y=y}(v^*)}{\partial y_1}\frac{y_1^*}{f_{Z|Y=y}(v^*)}}{v^* + y_1^* \frac{\partial v(c,t_2)}{\partial y_1}}\right]\right)$$

Let $h(s,\widetilde{y}) = \widehat{f}(s,\widetilde{y}) - f(s,\widetilde{y})$ and $h(\widetilde{y}) = \widehat{f}(\widetilde{y}) - f(\widetilde{y})$, then

$$\left[\frac{f(\widetilde{y})\int^{\widetilde{z}} h(s,\widetilde{y})ds - h(\widetilde{y})\int^{\widetilde{z}} f(s,\widetilde{y})ds}{f(\widetilde{y})^2}\right] = \int \frac{\left[1(s\leq\widetilde{z}) - F_{Z|\widetilde{Y}=\widetilde{y}}(\widetilde{z})\right]}{f(\widetilde{y})}\left(\widehat{f}(s,\widetilde{y}) - f(s,\widetilde{y})\right)ds$$

and

$$\left[\frac{f(y)\int^{v^*} h(s,y)\,ds - h(y)\int^{v^*} f(s,y)ds}{f(y)^2}\right] = \int \frac{\left[1(s\leq v^*) - F_{Z|Y=y}(v^*)\right]}{f(y)}\left(\widehat{f}(s,y) - f(s,y)\right)ds$$

so that

$$D\Delta(F;\widehat{F} - F)$$

$$= C\int \frac{\left[1(s\leq\widetilde{z}) - F_{Z|\widetilde{Y}=\widetilde{y}}(\widetilde{z})\right]}{f(\widetilde{y})}\left(\widehat{f}(s,\widetilde{y}) - f(s,\widetilde{y})\right)ds + C\int \frac{\left[1(s\leq v^*) - F_{Z|Y=y}(v^*)\right]}{f(y)}\left(\widehat{f}(s,y) - f(s,y)\right)ds$$

where

$$C = \left(\frac{1}{f_{Z|Y=y}(v^*)}\right)\left(1 + \left[\frac{\frac{\partial F_{Z|Y=y}(v^*)}{\partial y_1}\frac{y_1^*}{f_{Z|Y=y}(v^*)}}{v^* + y_1^* \frac{\partial v(c,t_2)}{\partial y_1}}\right]\right)$$

Following the same arguments as in Matzkin (2002), it is easy to show that this implies that

$\widehat{v}(y_1^*, t_2) - v(y_1^*, t_2)$ converges in probability to 0 and that

$\sqrt{N\sigma^4}\,(\widehat{v}(y_1^*, t_2) - v(y_1^*, t_2)) \to N(0, V_v)$ in distribution where

$$V_v = \left\{\int K(z)^2\right\}[C]^2\left(\frac{1}{f(\widetilde{y})} + \frac{1}{f(y)}\right)\left(F_{Z|Y=\widetilde{y}}(\widetilde{z})(1 - F_{Z|Y=\widetilde{y}}(\widetilde{z}))\right)$$

Since

$$\widehat{m}(t_1, t_2) = P_z\left(\widehat{v}(y_1^*, t_2)\right)$$

it follows by the delta method that

$$\sqrt{N\sigma^4}\left(\widehat{m}(t_1,t_2)-m(t_1,t_2)\right)\rightarrow N(0,V_m)$$

in distribution where

$$V_m=\left\{\int K(z)^2\right\}[C]^2\left(P_{zz}\left(v(y_1^*,t_2)\right)\right)^2\left(\frac{1}{f(\widetilde{y})}+\frac{1}{f(y)}\right)\left(F_{Z|Y=\widetilde{y}}(\widetilde{z})(1-F_{Z|Y=\widetilde{y}}(\widetilde{z}))\right)$$

7. REFERENCES

AÏT-SAHALIA, Y. (1994) "The Delta and Bootstrap Methods for Nonlinear Functionals of Nonparametric Kernel Estimators Based on Dependent Multivariate Data," mimeo, Princeton University..

BROWN, J. and H. ROSEN (1982) "On the Estimation of Structural Hedonic Price Models," Econometrica, 50, 765-769.

ALTONJI, J.G. and H. ICHIMURA (1997) "Estimating Derivatives in Nonseparable Models with Limited Dependent Variables", mimeo.

ALTONJI, J.G. and R.L. MATZKIN (2001) "Panel Data Estimators for Nonseparable Models with Endogenous Regressors," mimeo, Northwestern University.

BAJARI, P. and l. BENKARD (2001) "Demand Estimation with Heterogeneous Consumers and Unobserved Product Characteristics: A Hedonic Approach," mimeo, Stanford University.

BARTIK, T. (1987) "The Estimation of Demand Parameters in Hedonic Price Models," Journal of Political Economy, 95, 81-88.

BLUNDELL, R. and J.L. POWELL (2000) "Endogeneity in Nonparametric and Semiparametric Regression Models," invited lecture, 2000 World Congress of the Econometric Society.

BLUNDELL, R. and J.L. POWELL (2001) "Endogeneity of Single Index Models," mimeo, University College London.

BRIESCH, R., P. CHINTAGUNTA, and R.L. MATZKIN (1997) "Nonparametric Discrete Choice Models with Unobserved Heterogeneity," mimeo, Northwestern University.

BROWN, B.W. and M.B. WALKER (1989) "The Random Utility Hypothesis and Inference in Demand Systems", Econometrica, 57, 815-829.

BROWN, B.W. and M.B. WALKER (1995) "Stochastic Specification in Random Production Models of Cost-Minimizing Firms", Journal of Econometrics, 66, (1-2).

BROWN, D.J. and R.L. MATZKIN, (1998) "Estimation of Nonparametric Functions in Simultaneous Equations Models, with an Application to Consumer Demand," mimeo, Northwestern University.

BROWN, J. (1983) "Structural Estimation in Implicit Markets," in The Measurement of Labor Cost, editied by J. Triplett. Chicago: University of Press for NBER.

CHESHER, A. (2001) "Quantile Driven Identification of Structural Derivatives," mimeo, Institute for Fiscal Studies and University College London.

COSSLETT, S. R. (1983) "Distribution-free Maximum Likelihood Estimator of the Binary Choice Model, " Econometrica, 51(3), 765-782.

DAROLLES, S., J-P FLORENS, AND E. RENAULT (2001) "Nonparametric Instrumental Variable Regression," mimeo, University of Toulouse (2001)

EKELAND, I., J. HECKMAN and L. NESHEIM (2001) "Identification and Estimation of Hedonic Models," mimeo, University of Chicago

EPPLE, D. (1985) "Hedonic Prices and Implicit Markets: Estimating Demand and Supply Functions for Differentiated Products," Journal of Political Economy, 95, 59-80.

HECKMAN, J.J. (1974) "Effects of Day-Care Programs on Women's Work Effort," Journal of Political Economy.

HECKMAN, J. (1991,1995,1999) "Notes on Hedonic Models," mimeo, University of Chicago.

HECKMAN, J., R. MATZKIN, R. AND L. NESHEIM (2002) "Estimation and Simulation of Hedonic Models," presented at the Conference on Frontiers in Applied General Equilibrium, Yale University, April 5-6, 2002.

HECKMAN, J.J. and E.J. VYTLACIL (1999) "Local Instrumental Variables and Latent Variable Models for Identifying and Bounding Treatment Effects," Proceedings of the National Academy of Science, Vol 96.

HECKMAN, J.J. and E.J. VYTLACIL (2001) "Local Instrumental Variables," in Nonlinear Statistical Inference: Essays in Honor of Takeshi Amemiya, by C. Hsiao, K. Morimune, and J. Powell 9editors), Cambridge University Press.

HECKMAN, J.J. and R. WILLIS (1977) "A Beta-Logistic Model for the Analysis of Sequential Labor Force Participation by Married Women," Journal of Political Economy.

HONG H. and M. SHUM (2001) "A Semiparametric Estimator for Dynamic Optimization Models, with an Application to a Milk Quota Market," mimeo, Princeton University.

HOROWITZ, J.L. (1992) "A Smooth Maximum Score Estimator for the Binary Choice Model," Econometrica, 60, 505-531.

HOROWITZ, J.L. (1996) "Semiparametric Estimation of a Regression Model with an Unknown Transformation of the Dependent Variable," Econometrica, 64, 103-137.

HOROWITZ, J.L. (1998) Semiparametric Methods in Econometrics, New York and Berlin: Springer-Verlag.

HOROWITZ, J.L. (2001) "Nonparametric Estimation of a Generalized Additive Model with an Unknown Link Function," Econometrica, 69 (2), 499-513.

ICHIMURA, H. and T.S. THOMPSON (1994) "Maximum Likelihood Estimation of a Binary Choice Model with Random Coefficients of Unknown Distribution," mimeo, University of Minnesota.

IMBENS, G.W. AND W.K. NEWEY (2001) "Identification and Estimation of Triangular Simultaneous Equations Models Without Additivity," mimeo, UCLA.

KAHN, S. and K. LANG (1988) "Efficient Estimation of Structural Hedonic Systems," International Economic Reveiew, 29, 157-166.

KLEIN, R.W. and R.H. SPADY (1993) "An Efficient Semiparametric Estimator for Discrete Choice Models," Econometrica, 61, 387-422.

LANCASTER, T. (1979) "Econometric Methods for the Analysis of Unemployment" Econometrica, 47, 939-956.

LEWBEL, A. (1996) "Demand Systems With and Without Errors: Reconciling Econometric, Random Utility, and GARP Models," mimeo, Brandeis University, September 1996.

MANSKI, C. (1975) "Maximum Score Estimation of the Stochastic Utility Model of Choice," Journal of Econometrics, 3, 205-228.

MANSKI, C. (1985) "Semiparametric Analysis of Discrete Response: Asymptotic Properties of the Maximum Score Estimator," Journal of Econometrics, 27, 313-333.

MATZKIN, R.L. (1991) "Semiparametric Estimation of Monotone and Concave Utility Functions for Polychotomous Choice Models," Econometrica, 59, 1315-1327.

MATZKIN, R. (1992) "Nonparametric and Distribution-free Estimation of the Binary Choice and the Threshold Crossing Models," Econometrica, 60, 239-270.

MATZKIN, R. (1993) "Nonparametric Identification and Estimation of Polychotomous Choice Models," Journal of Econometrics, 58,137-168.

MATZKIN, R. (2002) "Nonparametric Estimation of Nonadditive Random Functions," mimeo, Northwestern University.

McELROY, M.B. (1981) "Duality and the Error Structure in Demand Systems," Discussion Paper #81-82, Economics Research Center/NORC.

McELROY, M.B. (1987) "Additive General Error Models for Production, Cost, and Derived Demand or Share Systems", Journal of Political Economy, 95, 737-757.

McFADDEN, D. (1974) "Conditional Logit Analysis of Qualitative Choice Behavior," in P. Zarembka, ed. Frontiers in Econometrics, pp105-142, New York: Academic Press.

McFADDEN, D. and K. TRAIN (2000) "Mixed MNL Models for Discrete Response" mimeo, University of California at Berkeley.

NADARAYA, E.A. (1964) "Some New Estimates for Distribution Functions," Theory Prob. Applic., 15, 497-500.

NEWEY, W.K. (1994) "Kernel Estimation of Partial Means and a General Variance Estimator," Econometric Theory, 10, 233-253.

NEWEY, W.K. and J. POWELL (2000) "Instrumental Variables for Nonparametric Models," mimeo, MIT.

NESHEIM, L. (2001) "Equilibrium Sorting of Heterogeneous Consumers Across Locations: Theory and Empirical Implications," Ph.D. Thesis, University of Chicago.

OLLEY, G.S. and A. PAKES (1996) "The Dynamics of Productivity in the Telecommunications Equipment Industry," Econometrica, Vol. 64, 6, 1263-1297.

POWELL, J., J. STOCK, and T. STOKER (1989) "Semiparametric Estimation of Index Coefficients," Econometrica, 57, 1403-1430.

ROEHRIG, C.S. (1988) "Conditions for Identification in Nonparametric and Parametric Models", Econometrica, 56, 433-447.

ROSEN, S. (1974) "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," Journal of Political Economy, 82, 34-55.

TINBERGEN, J. (1956) "On the Theory of Income Distributions," Weltwirtschaftliches Archiv, 77, 155-73.

VYTLACIL, E. (2001) "Semiparametric Identification of the Average Treatment Effect in Nonseparable Model," mimeo, Stanford University.

WILSON, (1991)