THE RELATIONSHIP BETWEEN MARGINAL WILLINGNESS-TO-PAY IN THE HEDONIC AND DISCRETE CHOICE MODELS

MAISY WONG

ABSTRACT. Willingness-to-pay is important for welfare analysis. The two primary approaches to estimate marginal willingness-to-pay (MWTP) for differentiated goods are hedonics (Rosen, 1974) and discrete choice models (McFadden, 1974). For many years, researchers have alluded to the apparent duality between both models. The innovation in this paper is to show that the hedonic MWTP can be written as a function of choice probabilities in the discrete choice model. I find that the hedonic method estimates a *weighted average* of marginal utilities where higher weights are associated with consumer types whose choice probabilities indicate a high variance regarding their choice (*marginal* consumers). This variance decreases as choice probabilities approach 0 or 1. Therefore, the hedonic method gives more weight to the preferences of the marginal consumer relative to the discrete choice approach. We can use these probability weights to analyze how MWTP in the discrete choice model differs from MWTP in the hedonic model.

1. INTRODUCTION

Willingness-to-pay is important for welfare analysis. The two primary approaches to estimate willingness-to-pay (WTP) for differentiated goods are hedonics (Rosen, 1974) and discrete choice models (McFadden, 1974). These two approaches have been implemented to estimate parameters of interests in the fields of development, education, environmental, industrial organization, labor and urban economics, including the WTP for air quality, housing, automobiles and school quality, to name a few.[1] Despite a large body of literature that employs these two approaches, to my knowledge, there is no theoretical analysis of the relationship

I am indebted to Fernando Ferreira for his guidance and time. I thank Kenneth Chay, Han-Ming Fang, Alex Gelber, Michael Greenstone, Joe Gyourko, Mark Jenkins, Nicolai Kuminoff, Jeremy Tobacman and participants at the CSWEP Mentoring Workshop and the Wharton Applied Economics Workshop provided valuable feedback. Lee Hye Jin provided excellent research assistance. All errors are my own.

[1]See Bayer, Ferreira, & McMillan (2007); Cellini, Ferreira, & Rothstein (2008); Chay & Greenstone (2004); Kremer, Leino, Miguel, & Zwane (2009); Wong (2008), for a few examples in

between the two models.[2] Moreover, some papers that used both models to estimate WTP found different results. For example, Banzhaf (2002) finds that the WTP for the same change in air quality varies between \$8 (hedonics) to \$18-\$25 (discrete choice) using the same data.

Economists have alluded to the duality between both the hedonic and discrete choice models. Cropper, Deck, Kishor, & McConnell (1993) and Mason & Quigley (1990) use simulations to compare preference estimates from a traditional Rosen hedonic model to the traditional McFadden multinomial Logit. Bayer, Ferreira, & McMillan (2007) discuss parallels in the aggregate estimating equation in a discrete choice model with random coefficients to the estimating equation in a hedonic regression framework. These papers focus on aggregate consumer behavior in both models but I find that consumer heterogeneity will turn out to be an important distinguishing feature.

This paper shows that the hedonic MWTP can be written as a function of choice probabilities in the discrete choice model. I begin with a random coefficient, discrete choice Logit model where consumers have heterogeneous taste for products (the Logit error) and product characteristics.[3] They choose one among several discrete products to maximize utility. An equilibrium is a vector of equilibrium prices and an allocation of products such that no consumers have an incentive to deviate from their choices. This discrete choice Logit model has an aggregate probability function that summarizes the probability that consumers choose a product, given equilibrium prices.

Rosen's hedonic model is a dual way to describe equilibrium in a market with differentiated goods.[4] I investigate how probability functions in the random coefficients Logit model relates to the gradient of the hedonic price function, the first

---

different fields in the applied economics literature. There is a related literature on WTP estimation in general equilibrium models (Epple & Sieg (1997), Epple, Romer, & Sieg (2001) and Sieg, Smith, Banzhaf, & Walsh (2004)). These models are particularly suited for settings with non-marginal changes in product characteristics.

[2]Feenstra (1995) studies the theoretical properties of exact hedonic price indexes in a setting with discrete products but does not compare WTP in both models.

[3]There are multiple discrete choice frameworks, including models with a probit error and with no Logit error. I start with the Logit model because it is the most common framework in the literature on consumer WTP and the solutions have tractable functional forms.

[4]The seminal Rosen paper maps out a two-step methodology to estimate the average MWTP function of consumers. Recently, Ekeland, Heckman, & Nesheim (2004); Heckman, Matzkin, & Nesheim (2010) and Bajari & Benkard (2005) extend the Rosen model. There is a consensus that the second step of the Rosen model has not been implemented successfully (Deacon et al., 1998; Chay & Greenstone, 2004) while the recent papers are still relatively new. For these reasons, I

step to identifying the MWTP function in Rosen's hedonic model.[5]Rosen (1974) showed that a consumer choosing to buy a differentiated good will maximize his utility when his indifference curve is tangent to the hedonic price function. Using the first order conditions from a consumer optimization problem, he showed that the marginal rate of substitution (MRS) between a characteristic of the differentiated good and the numeraire good is equal to the gradient of the hedonic price function, evaluated at the optimal levels of product characteristics. Therefore, the gradient of the hedonic price function helps to identify MWTP for product characteristics from the tangencies.

A popular rule of thumb by practitioners is to use the discrete choice framework when the number of products is small and the hedonic framework when the number of products is large. Since Rosen's model assumes a continuum of products, the more products in a market, the better the empirical setting approximates this assumption. There is also a practical reason-computational power-because discrete choice models with many products are computationally costly to estimate. Recently, with the rise in computational power, there are more settings where both the discrete choice model and the hedonic model have been employed (eg. housing). Therefore, it is important to know how the hedonic model performs in a setting with discrete products. To investigate this rule of thumb, I explore a hedonic model with fixed costs where consumers and producers still optimize over a continuous set of alternatives in product characteristics and indifference curves and isoprofit curves are differentiable. However, some products are no longer profitable to produce due to fixed costs, so the set of products in equilibrium is no longer a continuum.

With a discrete product space, consumer heterogeneity has important implications on whether the hedonic model can identify average MWTP in the population.(Nesheim, 2006) In the traditional Rosen framework, there is a continuum of consumers choosing amongst a continuum of products to maximize utility. In this setting, the indifference cuves of all consumers are tangent to the hedonic price function so that all consumers are just indifferent and all consumers are marginal consumers. In such a model, the average MWTP function of marginal consumers is also the average MWTP function of the population. With a discrete

focus on the one-step hedonic approach, which is the most common application of the hedonic method to date.

[5]The first step of Rosen's hedonic model estimates the hedonic price function, which is a necessary step to estimating the MWTP function in the second step.

product space, some consumers may be inframarginal in equilibrium (they do not have an incentive to deviate from their choice but their indifference curves are not tangent to the hedonic price function). Therefore, the gradient of the hedonic price function may not identify the average MWTP of the population.

The main finding in this paper is that the gradient of the hedonic price function is the ratio of a weighted average of individual marginal utilites, where the weights are a function of choice probabilities in the discrete choice Logit model. To give an example that relates this insight to the probability weights in the discrete choice model, consider a consumer type whose probability of choosing a product is one (zero). I find that the first step of the hedonic method cannot identify MWTP of these consumers (their weight is zero). This is because these are consumers who choose (not choose) a product with certainty (inframarginal consumers). More generally, I find that higher weights are associated with the marginal utilities of consumer types whose choice probabilities indicate more variation regarding their choices (marginal consumers). As this choice probability approaches 0 or 1, the weights start to decrease to zero.

Using the probability weights above, I show that average MWTP in the traditional McFadden Logit model is exactly equal to the gradient of the hedonic price function. However, I show that this is a special case. With heterogeneity in taste for product characteristics, both models are not necessarily duals of each other. This is because the ratio of a weighted average of marginal utilities (the hedonic price gradient) is not the same as the average ratio of marginal utilities (average MWTP in the random coefficients Logit model). Both are identical only when the probability weights and the marginal utilities are the same for all consumer types (this is the special case of the McFadden Logit model).

Many policy-relevant papers are interested in preferences of the marginal consumer, the mean consumer or both (Heckman & Vytlacil, 2005; Carneiro, Heckman, & Vytlacil, 2005). Chay & Greenstone (2004) estimate a hedonic model in the housing market with a correlated random coefficients framework to test whether the marginal consumer differs from the mean consumer in the presence of sorting. This paper shows that higher moments in choice data can be used to investigate how marginal consumers differs from the mean consumers.

The remainder of the paper is organized as follows: In the next section, I set up a random coefficients discrete choice Logit model and investigate how it relates to the Rosen hedonic model with fixed costs. Then, I derive the gradient of the hedonic price function as a function of choice probabilities in the Logit model. Finally, I conclude.

2. THEORY

2.1. **Discrete choice model with Logit.** *Consumer preferences*

There are $t = 1, ..., T$ markets and each market has J_t products. This is a multinomial discrete choice model where the probability that consumer i in market t chooses product j is π_{ijt}. These choice probabilities arise from a utility framework where consumer i's indirect utility from choosing product j in market t is,

$$u_{ijt} = V(x_{jt}, p_{jt}; \beta_i) + \varepsilon_{ijt} \tag{2.1}$$

where y_i is income, p_{jt} is product price, x_{jt} is a K-dimensional (row) vector of product characteristics, ε_{ijt} is a mean-zero stochastic term.[6] The price of the numeraire good, $y_i - p_{jt}$, is normalized to 1. For the main result in this paper, the important assumption is that ε_{ijt} is additive and separable from $V(\centerdot)$ and drawn from a Type I extreme value distribution. I will assume that each market is independent from other markets. To simplify the notation, I will drop the market subscript, t, from here. Each consumer of type i represents a population of measure 1 where each type has random taste parameters for products (drawn from $F(\varepsilon)$) and product characteristics (drawn from $F(\beta)$). The model is closed with an outside good, $j = 0$, and the utility from the outside good is normalized to 0.

A common functional form for $V(x_{jt}, p_{jt}; \beta_i)$ in the discrete choice literature is the random coefficients utility function: $V(x_{jt}, p_{jt}; \beta_i) = x_{jt}\beta_i + \beta_{iP}(y_i - p_{jt})$[7] In this random coefficients discrete choice model, the average MWTP for characteristic k is

$$WTP_k^{DCM} = \int \frac{\beta_{ik}}{\beta_{iP}} dF(\beta) \tag{2.2}$$

[6] Some discrete choice models include a term, ξ_{jt}, which represents the unobserved quality of the product. This distinction is not important for the main result in this paper.

[7] Although this functional form is flexible (McFadden & Train, 2000), it places restrictions, especially on income effects. One can also model piecewise income effects or Cobb-Douglas utility (see Petrin (2002) and Berry, Levinsohn, & Pakes (1995)). The intuition behind the result will be the same.

Choice probabilities

Consumers purchase one unit of j that offers the highest utility.[8] The probability that type i chooses product j is derived from differences in utility, $Pr(u_{ij} - u_{ik} > 0, \forall j \neq k)$. Let A_j be the set of individuals who choose j

$$A_j(x., p., \delta) = \left\{\beta_i, \varepsilon_{i0}, \varepsilon_{i1}, ..., \varepsilon_{iJ} | u_{ij} - u_{ik} \geq 0, k = 0, ..., J\right\} \tag{2.3}$$

To obtain the aggregate probability that product j is chosen, π_j, we aggregate over all individuals in the market who chose product j,

$$\begin{aligned} \pi_j(\centerdot) &= \int_{A_j} dF(\beta) dF(\varepsilon) \\ &= \int \frac{exp(V_{ij})}{\sum_{j'=0}^{j'=J} exp\left(V_{ij'}\right)} dF(\beta) \\ &\equiv \int \pi_{ij}(\centerdot)\, dF(\beta) \end{aligned} \tag{2.4}$$

A consequence of the distributional assumption of the Logit error (ε) is that individual choice probabilities, π_{ij}, assumes the Logit form, $\pi_{ij} = \frac{exp(V_{ij})}{\sum_{j'} exp(V_{ij'})}$. Given a fixed supply, an equilibrium is represented by a vector of equilibrium prices and an allocation of products to consumers so that no one has an incentive to change their choices. The J-dimensional vector, π^*, summarizes the probability that consumers in the market choose product j, as a function of product characteristics and prices. Each element in this vector, π_j^*, is the probability function evaluated at the equilibrium quantities of product characteristics, $\pi(x_j^*, p_j^*)$.

2.2. **Hedonic model.** The hedonic model offers a dual way to describe a differentiated goods equilibrium. There is a continuum of products and a continuum of consumer types. To allow for a discrete product space in the Rosen model, I augment it to allow for fixed costs in the supply side. Consumers and producers optimize over a continuous set of alternatives so that indifference and isoprofit curves are continuous and differentiable but the set of products in equilibrium may not be a continuum because some products are not profitable due to fixed costs.

[8] While this is standard in the literature, this assumption can be relaxed easily.

Producer problem

Let the cost to produce a product with a vector of characteristics, x, be $C(x)+c^i$, where $C(.)$ is convex and the marginal cost of each characteristic, $\partial C/\partial x_k$, is positive and increasing. Without loss of generality, I assume that the first characteristic, x_1, has fixed cost of c^i for producer i. For example, the product could be local newspapers and the first characteristic could be the number of pages with color. The fixed cost would be the cost to purchase a color printer. There is a continuum of producer types where ω^i is type i's cost parameter. The producer's problem is to choose a product, x, to maximize profits.[9]

$$max_{\{x_k\}_k}\pi(x)=P(x)-(C(x_1,...,x_K;\omega^i)+c^i) \quad s.t. \quad \pi(x)\geqslant 0$$

A profit-maximizing producer i will produce x_k such that the marginal cost is equal to the marginal revenue, $\frac{\partial C}{\partial x_k}=\frac{\partial P}{\partial x_k}$. For characteristic 1, producers will also need to compare the variable profit holding other characteristics at the optimum level, $(P(x)-C(x_1,x^*_{2,}...,x^*_K;\omega^i))$, to the fixed cost. If $x^{'}_1$ satisfies the first order condition but $P((x^{'}_1,x^*_{-1})-C(x^{'}_1,x^*_{-1};\omega^i)<c^i$, then the producer produces x^*_1 such that total profits is at the boundary condition, $\pi=P(x^*)-C(x^*;\omega^i)-c^i=0$. Each producer takes the market price, $P(x)$, as the maximum price obtainable for model **x**.

Consumer problem

The consumer problem for differentiated goods is analogous. There is a continuum of consumer types. Type i will choose one unit of a product to maximize utility, U^i, subject to the budget constraint, $P(x)+numeraire\leqslant y^i$. Consumers take the market price, $P(x)$, as the minimum price needed for product **x**. Optimality is achieved when the ratio of the marginal utilities for x_k and the numeraire is equal to the ratio of the marginal costs for characteristic k and the numeraire (normalized to have a price of 1), $\frac{\partial P}{\partial x_k}=\frac{\partial U^i/\partial x_k}{\partial U^i/\partial P}$.[10]

Equilibrium

An equilibrium is characterized by consumers and producers who are maximizing utilities and profits so that no one has an incentive to change their choices.

9 I assume that each producer considers producing 1 product only. See Rosen (1974) for a discussion on relaxing this assumption.

10 As discussed in the original Rosen paper, second order conditions are satisfied provided utility functions follow the standard assumptions and the price function is not too concave.

Prices adjust so that the marginal consumer and the marginal producer are just indifferent in equilibrium and each point on the hedonic price function is a consumer/producer indifference condition ($\frac{\partial U^i/\partial x_k}{\partial U^i/\partial P} = \frac{\partial P}{\partial x_k} = \frac{\partial C}{\partial x_k}$). Equilibrium interactions of consumers and producers in a market trace out this price-characteristic locus that defines a market clearing, implicit (hedonic) price function, $P(x)$.

In principle, the hedonic approach can be used to recover consumer's MWTP function (also known as the bid function). Rosen proposed a 2-step method. The idea behind the first step is that the hedonic locus traces out the tangencies between the marginal consumers' bid functions and the marginal producers' offer functions. In the first step, the hedonic price function is estimated by projecting product prices onto the space of product characteristics. The estimated function is used to predict the product specific marginal (hedonic) prices for characteristic k, $\partial P_j/\partial x_{jk}$. The first order conditions show that this estimated gradient identifies MWTP at the equilibrium points. To estimate MWTP away from the equilibrium points, the second step uses consumer and producer data as demand and cost shifters to estimate the MWTP function and the compensated supply function simultaneously. However, there is a consensus that empirical applications have not identified a situation in which the identification assumptions for the second step holds. [11] Therefore, the literature has instead focused on the first step, estimating the hedonic price function. This one-step hedonic method is necessary to identify MWTP, and with additional assumptions, is also sufficient to estimate MWTP for the representative consumer (see Muelbauer (1974) and Rosen (1974) and a discussion of applications to price indexes in Pakes (2003)).

In the Rosen model with no fixed costs and a continuum of producer and consumer types, all consumers and producers are located on the tangencies in equilibrium. In other words, each consumer and producer is indifferent and hence is a marginal consumer/producer. If there exists a consumer i and a producer i' such that $\frac{\partial U^i/\partial x_1}{\partial U^i/\partial P} > \frac{\partial C^{i'}}{\partial x_1}$, then both have an incentive to deviate. The consumer gains by consuming more x_1 as long as $\frac{\partial U^i/\partial x_k}{\partial U^i/\partial P} - \frac{\partial P}{\partial x_k} > 0$ and it is profitable for the producer to produce more x_1 as long as $\frac{\partial P}{\partial x_k} - \frac{\partial C}{\partial x_k} > 0$. Therefore, in equilibrium, all consumers and producers are tangent to the hedonic price function so

[11] See Deacon, Brookshire, Fisher, Kneese, Kolstad, Scrogin, Smith, Ward, & Wilen (1998) and Chay & Greenstone (2004). Recently, some researchers have revisited this question (see Bajari & Benkard (2005); Bishop & Timmins (2008); Heckman, Matzkin, & Nesheim (2010); Kuminoff, Parmeter, & Jaren Pope)

that the gradient of the hedonic price function at each point, $\frac{\partial P}{\partial x_k}$, identifies each consumer's MWTP for that characteristic and each producer's marginal cost of producing that characteristic around the equilibrium points. With a continuum of producer types but only one type of consumer, the family of bid functions degenerates to a single surface and $P(x)$ must be everywhere identical with a unique family of MWTP functions for the representative consumer (who is also the marginal consumer and the average consumer).

With fixed costs, however, some consumers may not be tangent to the hedonic price function. For example, the utility-maximizing choice for some consumers could be where $\frac{\partial U^i/\partial x_1}{\partial U^i/\partial P} > \frac{\partial P}{\partial x_1}$, so that they would like to consumer more of characteristic 1. Even if there exists producers i' such that $\frac{\partial C^{i'}}{\partial x_1} < \frac{\partial U^i/\partial x_1}{\partial U^i/\partial P}$, it may not be profitable for producers to produce more x_1 due to fixed costs. Since the bid function that maximizes their utility is not tangent to the hedonic price function, the gradient of the hedonic price function cannot identify MWTP for these consumers. They are inframarginal in that they are not indifferent but have no incentive to change their choices.

Therefore, the hedonic model with fixed costs has implications for the interpretation of the hedonic gradient as average MWTP in the population.[12] If an equilibrium is associated with a set of inframarginal consumers whose utility-maximizing indifference curves are not tangent to the hedonic price function, the one step hedonic method can only identify the average MWTP for marginal consumers. This could be problematic because MWTP of the marginal consumer could be different than MWTP of the average consumer but the econometrician cannot observe who is marginal.

3. COMPARING HEDONIC AND DISCRETE CHOICE MODELS

Theorem 1. *Let consumer utility from good j be $U_{ij} = V(X_j;\beta_i) + \varepsilon_{ij}$ where X_j is a vector of characteristics of product j; β_i indexes consumer i's taste for product characteristics, drawn from the distribution $F(\beta)$ and ε_{ij} is an idiosyncratic taste shock drawn from a Type I extreme value distribution, assumed independent from*

[12]Instead of fixed costs, other sources of friction can produce the same result where some consumers are inframarginal. For example, with transaction costs, some consumers' indifference curve may not be tangent to the hedonic price function but still may not have an incentive to change their choices because of transaction costs. Regulations that restrict choice sets could also cause non-tangencies in equilibrium(Rosen, 1974).

$F(\beta)$. Let π_{ij} denote the probability that consumer i chooses product j. If an equilibrium exists that can be represented by a hedonic price function, then, the gradient of the hedonic price function for that equilibrium is a function of the choice probabilities, π_{ij}.[13]

Proof. [see Theorem 15.7 in Simon & Blume (1994)] Let $\pi_1, ..., \pi_J : \mathbf{R}^{\mathbf{J(K+1)}} \to \mathbf{R}^{\mathbf{1}}$ be C^1 functions. Consider a system of equations in an equilibrium, with constants $\pi_1^*, ..., \pi_J^*$,

$$
\begin{aligned}
\pi_1(p_1, ..., p_J, x_{11}, ..., x_{1K}, ..., x_{jk}, ..., x_{J1}, ..., x_{JK}) &= \pi_1^* \\
&\vdots \\
&\vdots \\
\pi_J(p_1, ..., p_J, x_{11}, ..., x_{1K}, ..., x_{jk}, ..., x_{J1}, ..., x_{JK}) &= \pi_J^*
\end{aligned}
\tag{3.1}
$$

as possibly defining $p_1, ..., p_J$ as implicit functions of $x_{11}, ..., x_{JK}$. Suppose that ($\mathbf{p^*}$,$\mathbf{x^*}$) is a solution of Equation 3.1. If the determinant of the JxJ matrix

$$
\begin{bmatrix}
\frac{\partial \pi_1}{\partial p_1} & \cdots & \frac{\partial \pi_1}{\partial p_J} \\
\vdots & \ddots & \vdots \\
\frac{\partial \pi_J}{\partial p_1} & \cdots & \frac{\partial \pi_J}{\partial p_J}
\end{bmatrix}
$$

evaluated at ($\mathbf{p^*}$,$\mathbf{x^*}$) is nonzero so that the matrix is invertible, then there exist C^1 functions

$$
\begin{aligned}
P_1(x_{11}, ..., x_{JK}) &= p_1 \\
P_J(x_{11}, ..., x_{JK}) &= p_J
\end{aligned}
\tag{3.2}
$$

defined on a ball B about $\mathbf{x}^*$ such that

[13]When the utility function has a Logit error, consumers have idiosyncratic taste shocks for products, ε_{ij}. An equilibrium could exist with two products that share the same characteristics but different prices. Consumers would still demand the product with the higher price if they have a high taste shock for that product.Therefore, a hedonic price function may not exist for every equilibria in the discrete choice Logit model because the characterisic-price mapping may not be 1-1.

(3.3)
$$\begin{aligned}\pi_1(P_1(\mathbf{x}),...,P_J(\mathbf{x}),x_{11},...,x_{1K},...x_{jk},...,x_{J1},...,x_{JK}) &= \pi_1^*\\ &\vdots\\ \pi_J(P_1(\mathbf{x}),...,P_J(\mathbf{x}),x_{11},...,x_{1K},...x_{jk},...,x_{J1},...,x_{JK}) &= \pi_J^*\end{aligned}$$

for all $\mathbf{x} = (x_{11},...,x_{JK})$ in B.

And the gradient of the implicit price function with respect to x_{jk} is

(3.4)
$$\begin{bmatrix} \partial P_1/\partial x_{jk} \\ \vdots \\ \partial P_J/\partial x_{jk} \end{bmatrix} = - \begin{bmatrix} \frac{\partial \pi_1}{\partial p_1} & \cdots & \frac{\partial \pi_1}{\partial p_J} \\ \vdots & \ddots & \vdots \\ \frac{\partial \pi_J}{\partial p_1} & \cdots & \frac{\partial \pi_J}{\partial p_J} \end{bmatrix}^{-1} \begin{bmatrix} \partial \pi_1/\partial x_{jk} \\ \vdots \\ \partial \pi_J/\partial x_{jk} \end{bmatrix}$$

Since ε is Type I exteme value, additive and independent from $F(\beta)$, $\pi_j = \int \frac{exp(V_{ij})}{\sum_{j'} exp(V_{ij'})} dF(\beta)$, $\frac{\partial \pi_j}{\partial x_{jk}} = \int \frac{\partial V_{ij}}{\partial x_{jk}} \pi_{ij}(1-\pi_{ij})\, dF(\beta)$ and $\frac{\partial \pi_j}{\partial x_{j'k}} = \int \frac{\partial V_{ij}}{\partial x_{j'k}} \pi_{ij}\pi_{ij'}\, dF(\beta)\, \forall j \neq j'$. Similarly for $\frac{\partial \pi_j}{\partial P_j}$. □

Corollary. *If the hedonic price function, $P(\centerdot)$, is only a function of own-product characteristics, so that $\frac{\partial P_j}{\partial x_{j'k}} = \frac{\partial P(\mathbf{x})}{\partial x_{j'k}} \mid_{\mathbf{x}=\mathbf{x}_j^*} = 0 \forall j' \neq j$, then the gradient of the hedonic price function is the ratio of a weighted average of marginal utilities, where the weights are a function of choice probabilities in the discrete choice model.*

Proof. Differentiating each row j of equation (3.3) with respect to x_{jk},

(3.5)
$$\begin{aligned}\frac{\partial \pi_1}{\partial P_1}\frac{\partial P_1}{\partial x_{1k}} &= -\left(\frac{\partial \pi_1}{\partial x_{1k}}\right)\\ &\vdots\\ \frac{\partial \pi_J}{\partial P_J}\frac{\partial P_J}{\partial x_{Jk}} &= -\left(\frac{\partial \pi_J}{\partial x_{Jk}}\right)\end{aligned}$$

Therefore,

(3.6)

$$\begin{bmatrix} \partial P_1/\partial x_{1k} \\ \vdots \\ \partial P_J/\partial x_{Jk} \end{bmatrix} = -\begin{bmatrix} \frac{\partial \pi_1}{\partial p_1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \frac{\partial \pi}{\partial p_J} \end{bmatrix}^{-1} \begin{bmatrix} \partial \pi_1/\partial x_{1k} \\ \vdots \\ \partial \pi_J/\partial x_{Jk} \end{bmatrix} = -\begin{bmatrix} \frac{\int \frac{\partial V_{i1}}{\partial x_{1k}} w_{i1} dF(\beta)}{\int \frac{\partial V_{i1}}{\partial P_1} w_{i1} dF(\beta)} \\ \vdots \\ \frac{\int \frac{\partial V_{iJ}}{\partial x_{Jk}} w_{iJ} dF(\beta)}{\int \frac{\partial V_i}{\partial P_J} w_{iJ} dF(\beta)} \end{bmatrix}$$

where $w_{ij} = \pi_{ij} * (1-\pi_{ij}) \forall j$. □

Just as equilibrium conditions in the hedonic model defines the hedonic price function and its gradient implicitly, so does a discrete choice model. Theorem 1 uses the Implicit Function Theorem to relate the equilibrium price functions in both models using choice probabilities.

The weights are intuitive. They are the variance in a multinomial probability distribution, with probability type i choosing product j, π_{ij}. When π_{ij} is 0 or 1, none or all consumers of type i are not choosing/choosing product j such that the variance is 0. When $\pi_{ij} = 0.5$, half of type i consumers choose product j and the variance is the highest because marginal improvements in x_{jk} can tip the consumers to change their choices.

Equilibrium conditions in the hedonic model provides economic content to the probability weights. When the set of products is discrete, not all consumers are at the tangency. The hedonic gradient is only identified for consumers at the tangency (marginal consumers). Since these consumers are just indifferent, their choices have the highest variance. Consumer types with probabilities close to 0 or 1 have a small variance and are not likely to be at the tangency (these are inframarginal consumers).

The functional form of these weights is related to the distributional assumption of ε. The weights depend on how π_{ij} changes for marginal changes to x_{jk}. Due to this assumption, the individual choice probability, π_{ij}, is a sigmoid-shaped function of individual utility, $\pi_{ij} = \frac{exp(V_{ij})}{\sum_{j'} exp(V_{ij'})}$.[14] When V_{ij} is low relative to the utility for other choices, π_{ij} is close to 0 and a small improvement in x_{jk} may increase V_{ij} but is unlikely to make it higher than the utility for other choices. Therefore, this will have little effect on the probability of it being chosen. The

[14] Note also that this distributional assumption ensures that choice probabilities are strictly positive always and the partial derivatives will be well-defined.

sigmoid shape implies that a marginal improvement in x_{jk} leads to a small change in π_{ij} if the choice probability is close to 0 or 1, but the change is steepest when π_{ij} is 0.5.

Equation 3.6 suggests one could use numerical derivatives to explore the relationship between hedonics and discrete choice models with other distributional assumptions (Probit models or no Logit errors) even though these models do not have choice probabilities with analytical functional forms. The model with no Logit error (Berry & Pakes, 2007) is particularly interesting because standard utility functions in hedonics do not have the Logit error term.

Equation (3.6) also nests the traditional McFadden model ($U_{ij} = V(X_j) + \varepsilon_{ij}$) as a special case. In this case, $\pi_j = \frac{exp(V_j)}{\sum_{j'} exp(V_{j'})}$ so that $\frac{\partial P_j}{\partial x_{jk}} = -\frac{\partial \pi_j / \partial x_{jk}}{\partial \pi_j / \partial P_j} = -\frac{\left(\partial V_j / \partial x_{jk}\right)\pi_j(1-\pi_j)}{\left(\partial V_j / \partial P_j\right)\pi_j(1-\pi_j)} = -\frac{\left(\partial V_j / \partial x_{jk}\right)}{\left(\partial V_j / \partial P_j\right)}$. Therefore, a regression estimate of the average gradient of the hedonic price function would equal the average MWTP in the discrete choice model. This is because in the traditional McFadden Logit model, there is no heterogeneity in the taste for product characteristics and only heterogeneity in the taste for products (ε_{ij}). Therefore, the taste for characteristic k is represented by only one family of bid functions for characteristic k, with gradient $-\frac{\left(\partial V_j / \partial x_{jk}\right)}{\left(\partial V_j / \partial P_j\right)}$. In this setting with a distribution of products but no heterogeneity in the taste for product characteristics, the hedonic price function is the bid function and MWTP in the discrete choice model is the same as the MWTP in the hedonic model.

Beyond the McFadden Logit model, I do not find evidence supporting duality between MWTP in both models. This is easiest to see in a random coefficients utility model where the marginal utilities are $\partial V_{ij} / \partial x_{jk} = \beta_{ik}$. In this case, the population average MWTP for characteristic k is

$$MWTP_k^D = \int MWTP_{ik} dF(\beta) = \int \frac{\beta_{ik}}{\beta_{iP}} dF(\beta)$$

and each element in the vector (3.6) is

$$\partial P_j / \partial x_{jk} = \frac{\int \beta_{ik} \pi_{ij}(1-\pi_{ij}) dF(\beta_i)}{\int \beta_{iP} \pi_{ij}(1-\pi_{ij}) dF(\beta_i)}$$

Unless the average of the ratio of marginal utilities ($MWTP_k^D$) is equal to the average of the ratio of weighted averages of marginal utilities (equation (3.6)), it is unlikely that the MWTP in both models will be the same. Even in the case where the weights are the same for each individual, so that the ratio of the weighted average is just the ratio of the average ($\frac{\int \beta_{ik} dF(\beta_i)}{\int \beta_{iP} dF(\beta_i)}$), it is still not the average of the ratios ($MWTP_k^D$) unless the marginal utilities are the same for all individuals. This is the special case of the McFadden Logit model where $U_{ij} = V(X_j) + \varepsilon_{ij}$ and the marginal utility for characteristic k is $\partial V_j / \partial x_{jk}$.

In principle, given the distributional assumptions, the discrete choice model identifies a full distribution of MWTP, $F(WTP_i^D)$ but it is computationally costly to estimate such a model. One goal in this paper is to see whether there exists a mapping between hedonic MWTP and discrete choice MWTP ($MWTP_k^H = f(MWTP_k^D)$ and $MWTP_k^D = f^{-1}\left(MWTP_k^H\right)$). This mapping would be helpful because the hedonic model requires less computational time to estimate. The results above suggest that such a mapping is unlikely to be 1-1, except in the special case where the choice variance and marginal utilities are equal for all consumers.

How do the MWTP's differ?

The marginal consumers in an equilibrium are the consumers who satisfy the set of indifference conditions in the equilibrium ($\frac{\partial U^i / \partial x_k}{\partial U^i / \partial P} = \frac{\partial P}{\partial x_k}$). In the one-step Rosen method, only product characteristics data is used to estimate the hedonic price function. Econometricians cannot observe who the marginal consumer is but Equation (3.6) shows how the econometrician can use an estimate of the variance around the choices to estimate which type is likely to be the marginal consumer. With data on observed heterogeneity (eg. income, age, race and education), the sample can be stratified into different consumer types. A consistent estimate of the choice probability, $\hat{s_{ij}}$, is the percent of consumers of type i choosing product j and the variance is $\hat{s_{ij}} * (1 - \hat{s_{ij}})$. Equation (3.6) shows that the first step of the Rosen method is associated with higher weights for the preferences of consumer types whose sample choice variance is the highest. These are the consumers who are most likely to be the marginal consumers. If the weights are unequal across consumer types, then not all consumers are likely to be marginal and hence the hedonice gradient is unlikely to identify the average MWTP of the population. Furthermore, the MWTP estimates for both models will likely differ.

4. CONCLUSION

Willingness-to-pay is important for welfare analysis. The two primary approaches to estimate willingness-to-pay (WTP) are hedonics (Rosen, 1974) and discrete choice models (McFadden, 1974). For many years, researchers have alluded to the apparent duality between both theories. The innovation in this paper is to show that the hedonic gradient can be written as a function of choice probabilities in the discrete choice model. The main finding is that the gradient of the hedonic price function is a ratio of a weighted average of individual marginal utilites, where the weights are a function of choice probabilities in the discrete choice model. Consistent with previous literature, higher weights are associated with marginal consumers who are apt to have higher choice variances. Beyond the McFadden Logit model, I find that both models are not likely to be duals of each other.

In on-going research, I extend the relationship to other discrete choice models, including Probit and a model with no Logit error.

References

Bajari, P., & Benkard, C. L. (2005). Demand Estimation with Heterogeneous Consumers and Unobserved Product Characteristics: A Hedonic Approach. *Journal of Political Economy*, *113*(6), 1239–1276.

Banzhaf, H. S. (2002). Quality Adjustment for Spatially-Delineated Public Goods: Theory and Application to Cost-of-Living Indices in Los Angeles. Resources for the Future Discussion Paper 02-10.

Bayer, P., Ferreira, F., & McMillan, R. (2007). A Unified Framework for Measuring Preferences for Schools and Neighborhoods. *Journal of Political Economy*, *115*(4), 588–638.

Berry, S. T., Levinsohn, J., & Pakes, A. (1995). Automobile Prices in Market Equilibrium. *Econometrica*, *63*(4), 841–890.

Berry, S. T., & Pakes, A. (2007). The Pure Characteristics Demand Model. *International Economic Review*, *48*(4), 1193–1225.

Bishop, K., & Timmins, C. (2008). Simple, Consistent Estimation of the Marginal Willingness to Pay Function: Recovering Rosen's Second Stage without Instrumental Variables. Duke University Department of Economics Working Paper.

Carneiro, P., Heckman, J. J., & Vytlacil, E. (2005). Estimating Marginal and Average Returns to Education. Forthcoming, *American Economic Review*.

Cellini, S., Ferreira, F., & Rothstein, J. (2008). The Value of School Facilities: Evidence from a Dynamic Regression Discontinuity Design. NBER Working Paper No. 14516.

Chay, K., & Greenstone, M. (2004). Does Air Quality Matter? Evidence from the Housing Market. *Journal of Political Economy*, *113*(2), 376–424.

Cropper, M. L., Deck, L., Kishor, N., & McConnell, K. E. (1993). Valuing Product Attributes Using Single Market Data: A Comparison of Hedonic and Discrete Choice Approaches. *The Review of Economics and Statistics*, *75*(2), 225–232.

Deacon, R., Brookshire, D., Fisher, A., Kneese, A., Kolstad, C., Scrogin, D., Smith, V. K., Ward, M., & Wilen, J. (1998). Research Trends and Opportunities in Environmental and Natural Resource Economics. *Environmental and Resource Economics*, *11*(3-4), 383–397.

Ekeland, I., Heckman, J., & Nesheim, L. (2004). Identification and Estimation of Hedonic Models. *Journal of Political Economy*, (S2), S60–S109.

Epple, D., Romer, T., & Sieg, H. (2001). Interjurisdictional Sorting and Majority Rule: An Empirical Analysis. *Econometrica*, *69*(6), 1437–1466.

Epple, D., & Sieg, H. (1997). Estimating Equilibrium Models of Local Jurisdictions. *Journal of Political Economy*, *107*(4), 645–681.

Feenstra, R. C. (1995). Exact Hedonic Price Indexes. *Review of Economic Statistics*, *77*(4), 634–653.

Heckman, J. J., Matzkin, R., & Nesheim, L. (2010). Nonparametric Identification and Estimation of Nonadditive Hedonic Models. *Econometrica*, *78*(5), 1569–1591.

Heckman, J. J., & Vytlacil, E. (2005). Structural Equations, Treatment Effects, and Econometric Policy Evaluation. *Econometrica*, *73*(3), 669–738.

Kremer, M., Leino, J., Miguel, E., & Zwane, A. P. (2009). Spring Cleaning: Rural Water Impacts, Valuation and Property Rights Instructions. NBER Working Paper No. w15280.

Kuminoff, N., Parmeter, C., & Jaren Pope, t. . W. n. . F., year = 2010 (????).

Mason, C., & Quigley, J. (1990). Comparing the Performance of Discrete Choice and Hedonic Models. In M. M. Fischer, P. Nijkamp, & Y. Papageorgiou (Eds.) *Spatial Choices and Processes*. Amsterdam: North Holland Publishing Company.

McFadden, D. (1974). The Measurement of Urban Travel Demand. *Journal of Political Economy*, *3*(4), 303–328.

McFadden, D., & Train, K. (2000). Mixed MNL Models for Discrete Response. *Journal of Applied Econometrics*, *15*, 447–470.

Muelbauer, J. (1974). Household Production Theory, Quality, and the Hedonic Technique. *American Economic Review*, *64*(6), 977–994.

Nesheim, L. (2006). Hedonic Price Functions. Cemmap Working Paper CWP18/06.

Pakes, A. (2003). A Reconsideration of Hedonic Price Indexes with an Application to PC's. *AER*, *93*(5), 1578–1596.

Petrin, A. (2002). Quantifying the Benefits of New Products: The Case of the Minivan. *Journal of Political Economy*, *110*(4), 705–729.

Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy*, *82*(1), 34–55.

Sieg, H., Smith, V. K., Banzhaf, H. S., & Walsh, R. (2004). Estimating the General Equilibrium Benefits of Large Changes in Spatially Delineated Public Goods. *International Economic Review*, *45*(4), 1047–1077.

Simon, C. P., & Blume, L. (1994). *Mathematics for Economists*. W.W. Norton.
Wong, M. (2008). *Estimating the Impact of the Ethnic Housing Quotas in Singapore*. Ph.D. thesis, Massachusetts Institute of Technology.

MAISY@WHARTON.UPENN.EDU, WHARTON REAL ESTATE DEPARTMENT