A Tractable Approach to Compare the Hedonic and Discrete Choice Frameworks

Maisy Wong*
University of Pennsylvania

Abstract

The two primary approaches to estimate marginal willingness-to-pay (MWTP) are hedonic (Rosen, 1974) and discrete choice (McFadden, 1974). While both approaches rely on revealed preference methods to estimate MWTP, the primitives underlying both models are different, making it difficult to compare them. This paper establishes the assumptions needed to develop a tractable framework to compare both approaches. I begin with a discrete choice model and show how to derive the gradient of the equilibrium price function implicitly. I then incorporate Rosen's insight that the price gradient is equal to the MWTP of the marginal individual whose indifference curve is tangent to the price function in equilibrium. However, with discrete choices, some individuals may be inframarginal and their indifference curves will not be tangent to the price function. The analytical mapping I derive formalizes this intuition and shows that the price gradient depends on weighted averages of marginal utilities where higher weights are assigned to individuals whose choice probabilities indicate more uncertain choices (*marginal* individuals). As this choice becomes more certain, the weights start to decrease. This result shows how choice probabilities and other moments of choice data can be used to distinguish marginal versus inframarginal individuals.[1]

JEL: C01, R21, J23

Keywords: Hedonic, discrete choice, sorting, demand

*Wharton Real Estate. 3620 Locust Walk, 1464 SHDH, Philadelphia, PA 19104-6302. Email: maisy@wharton.upenn.edu.

[1] I am indebted to Fernando Ferreira for his guidance and time. I am grateful for the feedback from the editor and two referees. I thank Kenneth Chay, Michael Greenstone, HanMing Fang, Alex Gelber, Mark Jenkins, Jeremy Tobacman and participants at the CSWEP Mentoring Workshop for comments. Hyejin Lee and Xuequan Peng provided excellent research assistance. I am grateful to the Research Sponsors Program of the Zell/Lurie Real Estate Center at Wharton for financial support. All errors are my own.

1 Introduction

The two primary approaches to estimate marginal willingness-to-pay (MWTP) are the hedonic (Rosen, 1974) and discrete choice (McFadden, 1974) models. While both approaches are used widely in many fields,[2] there is little formal analysis of the relationship between both models. Moreover, some papers that use both approaches to estimate MWTP find different results.[3] However, it is difficult to investigate why the estimates differ without a framework to relate MWTP from both approaches.

At first glance, there seems to be a relationship between the two frameworks since both rely on revealed preference and individuals' sorting behavior across products to infer MWTP. On the other hand, the hedonic method assumes a continuum of products while the discrete choice method does not. Furthermore, the hedonic model is associated with compensated demand while the discrete choice model is associated with uncompensated demand. These differences in the primitives underlying both models suggest that both approaches may not be comparable, in general.

This paper makes progress in developing a tractable framework to investigate the relationship between the two approaches. I begin with a discrete choice model and show that one can derive an implicit equilibrium price function which is akin to the hedonic price function in Rosen (1974). This is useful because I can then implement the insight from Rosen to use the gradient of the implicit price function to shed light on preferences. This approach was alluded to in the seminal article on the estimation of discrete choice models with heterogeneous individuals (Berry, 1994).[4]

I establish the conditions needed to circumvent the challenges above. First, for tractability, I assume quasi-linear utility so that the compensated and uncompensated demand functions are identical. With the different choice environments (discrete versus continuous), the equilibrium price function derived implicitly from the discrete choice model will generally not be comparable to the hedonic price function derived from Rosen's hedonic framework. Instead, my approach is to derive the equilibrium price function associated with the discrete choice model first, which then allows me to embed Rosen's insight within the discrete choice framework.

Section 2 describes a discrete choice model with a continuum of heterogeneous individuals choosing a discrete set of products to maximize utility. Individuals have quasi-linear utility with random coefficients and an idiosyncractic Logit error term. Products are differentiated by a vector

[2] See Bayer et al. (2007); Cellini et al. (2008); Chay and Greenstone (2004); Pakes (2003); Berry et al. (1995); Bitzan and Wilson (2007); Wong (2013), to name a few examples in the fields of labor economics, local public finance, environmental economics, industrial organization as well as urban and transportation economics.

[3] For example, Banzhaf (2002) finds that the MWTP for the same change in air quality varies between \$8 (hedonic) to \$18-\$25 (discrete choice) using the same data.

[4] See the first footnote in Berry (1994): "Indeed, my model implicitly produces a hedonic equilibrium pricing function that depends on product characteristics. However, the focus in this article on structural estimation with price-setting firms and unobserved demand characteristics differs from the typical focus in the hedonic literature."

of (exogenous) characteristics and by prices that are determined in equilibrium. Throughout the paper, I assume supply is fixed and focus on consumer demand. With a continuum of individuals choosing discrete products, some individuals could be inframarginal in that small changes in product characteristics will not change their utility-maximizing choice. This can arise due to mobility costs faced by individuals or fixed costs of production (Arnott, 1989; Bayer et al., 2009; Gavazza, 2011).

A market equilibrium in the discrete choice model is characterized by a vector of equilibrium prices and an allocation of individuals to products such that each individual has no incentive to deviate. Notably, in the discrete choice framework, choices made by individuals can be summarized by choice probabilities using a *share function* that indicates the share of individuals in a market who choose a product as a function of the characteristics and prices of all the products in the market.

The hedonic approach presents a dual way to characterize equilibria in markets with differentiated products using the *hedonic price function*. Rosen (1974) showed that a utility-maximizing individual choosing amongst a continuum of differentiated products satisfies the first order conditions of optimization when her indifference curve is tangent to the hedonic price function. This is the famous insight in Rosen (1974) that an individual's MWTP for a characteristic is equal to the gradient of the hedonic price function with respect to that characteristic. Overall, Section 2 highlights the hedonic price function and the share function as key equilibrium objects of interest in the hedonic and discrete choice models, respectively.

Section 3 presents the three main results of the paper. The first result shows how to derive an implicit equilibrium price function from the share function in the discrete choice model. This result relies on the Implicit Function Theorem and the assumption of a continuum of individuals. With discrete choices, the equilibrium price function may not be continuous and the gradient of the price function (which is needed to infer MWTP) may not be defined. However, the assumption of a continuum of individuals gives rise to a continuous share function, which describes the share of individuals that choose each (discrete) product in equilibrium. Since the share function is continuous, I can then apply the Implicit Function Theorem to derive the gradient of the equilibrium price function from the share function. Adopting Rosen's insight that the gradient of the implicit price function is equal to the MWTP of the marginal individual, Result 1 presents a way to relate MWTP (associated with the implicit price gradient) with the share function.

While Result 1 establishes an analytical relationship between the share function and an implicitly defined equilibrium price function, it also underscores several differences between the discrete and hedonic approaches. First, the equilibrium price function derived from the share function shows that the equilibrium price for each product depends on all the products in the market. With discrete products, some products could have "market power" in that inframarginal individuals' choices are inelastic and unresponsive to small changes in product characteristics. In particular, in the random coefficients Logit model, the idiosyncratic Type I extreme value error term (ε_{ij}) assumes each individual i has a non-zero preference shock for product j. Individuals with high values of ε_{ij} have strong preferences for product j and will be inframarginal. By contrast, in Rosen's hedonic model with a continuum of products, each individual is a marginal consumer and each product is unimportant relative to the market.

Since the implicitly derived equilibrium price function depends on all products in the market, the relationship in Result 1 is complex and difficult to interpret. Next, the second result shows how to simplify the relationship and presents an intuitive interpretation. Here, I rely on the assumption in the empirical literature that typically estimates a hedonic price function by regressing the price of a product on its own characteristics only. If we restrict that the equilibrium price function derived in Result 1 is such that the price of each product only depends on the characteristics of that product, then, Result 2 shows that the implicitly defined price gradient can be written as a ratio of weighted averages of individual marginal utilities. The weights are a function of choice probabilities in the discrete choice model with higher weights corresponding to individuals with more uncertain choices.

The simplified relationship in Result 2 hones in on a critical distinction between the hedonic and the discrete choice approaches. If preferences are inferred from the gradient of the equilibrium price function, only the preferences of marginal individuals will be identified. In equilibrium, prices adjust so that the marginal individual is just indifferent but inframarginal individuals' choices are inelastic. Put another way, the hedonic approach relies on tangencies between the indifference curves and the hedonic price function to identify MWTP using the hedonic gradient. However, only the *marginal* individuals' indifference curves are tangent to the hedonic price function. The indifference curves of *inframarginal* individuals are not necessarily tangent to the hedonic price function.

This analytical relationship also formalizes how choice probabilities and choice variances can be used to determine which individuals are likely to be marginal. To interpret the economic intuition behind the probability weights, consider an individual whose probability of choosing a product is one. I find that the MWTP derived from the price gradient assigns no weight to this individual. This is because she chooses a product with certainty (she is inframarginal and her indifference curve is not tangent to the hedonic price function). More generally, the MWTP derived from the implicit price function depends on a ratio of weighted averages of marginal utilities where higher weights are assigned to individuals whose choice probabilities indicate a higher degree of uncertainty regarding their choice (marginal individuals). As this choice becomes more certain (as the probability approaches 0 or 1), the weights start to decrease.

Moreover, the second result also provides a tractable way to identify conditions under which

MWTP from both approaches can be identical. It shows clearly that average MWTP derived from the implicit price function depends on a ratio of (weighted) averages of marginal utilities whilst average MWTP in the discrete choice model is an average of ratios of marginal utilities. In general, the ratio of averages will not equal the average of ratios except in special cases (for example, when the ratios are constant).

This intuition delivers the third result that the average MWTP for a characteristic is identical in the two approaches if the MWTP for that characteristic is constant across individuals. The traditional Logit model with no random coefficients satisfies this condition. This appears to be a special case when ratios of marginal utilities (MWTP's) are constant. With heterogeneous preferences for characteristics (for example, with random coefficients utility), only the slopes of the indifference curves of *marginal* individuals are equal to the gradient of the implicit price function. Therefore, the average MWTP derived from the implicit price function (which gives higher weights to marginal individuals) diverges from the average MWTP in the discrete choice model (which estimates an average MWTP, averaged across marginal and inframarginal individuals). In the special case when the MWTP for a characteristic is constant, the marginal individual and the average individual have the same MWTP, so, average MWTP estimated from the two approaches are the same.

Overall, the three results above highlight a series of assumptions needed to make the hedonic and the discrete choice approach comparable. As discussed above, both approaches have fundamental differences in model primitives and are generally difficult to compare. This paper establishes the conditions needed to develop a tractable framework to derive the gradient of an implicit price function associated with the share function in the discrete choice model. I then utilize Rosen's insight to connect the (implicit) price gradient with MWTP. Furthermore, I establish the stronger assumptions required in Result 2 (to simplify the relationship and to develop intuition) and in Result 3 (to show the conditions needed for the approaches to deliver identical average MWTP).

Finally, Section 4 discusses several extensions to the framework and highlights a few caveats. First, I discuss other functional forms, including other error distributions besides the Logit error. This paper also assumes quasi-linear utility which ignores income effects. Second, the framework above only considers marginal perturbations around the equilibrium point since the Implicit Function Theorem only holds locally around the equilibrium point. Additionally, throughout the paper, supply is fixed which ignores general equilibrium effects where supply adjusts as well. Third, I discuss how the findings in this paper inform empirical estimation and identification of preferences. This paper abstracts away from omitted variable concerns due to unobserved product quality to focus on selection concerns driven by the sorting of heterogeneous individuals.

While there are many empirical papers that use the hedonic and discrete choice methods to estimate MWTP, there is a relatively small literature that directly compares both approaches. Cropper et al. (1993) and Mason and Quigley (1990) use simulated data to compare MWTP estimates in hedonic and discrete choice models. Several papers allude to similarities and differences between both models (Ellickson, 1981; Ekeland et al., 2004; Bayer et al., 2007; Bajari and Benkard, 2005). The innovation in this paper is to provide a tractable framework that delivers an analytical relationship between the share function in the discrete choice model and the gradient of the implicit price function (derived from the share function). This exercise helps to crystallize the similarities and differences between both approaches.

The remainder of the paper is organized as follows: I briefly describe the discrete choice and hedonic models in Section 2. I derive the three results above in Section 3 and discuss their implications. I discuss some possible extensions in Section 4 and conclude in Section 5.

2 Discrete choice and hedonic models

I begin with a discrete choice model with random coefficients Logit as the underlying data generating process and describe an equilibrium in this model. Then, I characterize equilibrium in Rosen's hedonic model. This discussion highlights two equilibrium objects of interest: the share function (in the discrete choice framework) and the hedonic price function (in the hedonic framework). Throughout this paper, supply is fixed and I focus mainly on describing consumer preferences. For simplicity, I assume all characteristics other than price are exogenous.

2.1 Discrete choice model and the share function

There are $t = 1, \dots, T$ markets and each market has J_t differentiated products. Individual *i's* indirect utility from choosing product j in market t is,

$$u_{ijt} = V(x_{jt}, p_{jt}; \beta_i, y_i) + \varepsilon_{ijt} \tag{1}$$

where y_i is the income of individual i, p_{jt} is the price of product j in market t, x_{jt} is a K-dimensional (row) vector of exogenous characteristics of product j. The numeraire good, $y_i - p_{jt}$, has a normalized price of 1. Each individual i has heterogeneous taste parameters for product characteristics (β_i drawn from a cumulative distribution function, F_β) and a random taste parameter for product j (ε_{ijt} drawn from F_ε). The model is closed with an outside good, $j = 0$. The utility from the outside good is normalized to 0. Each market is independent from other markets. To simplify notation, I will drop the market subscript from here.

The empirical literature makes two common assumptions for equation (1). First, $V(x_{jt}, p_{jt}; \beta_i, y_i)$ is a random coefficients utility function. Second, ε is drawn from a Type I extreme value distribution. For example, a common functional form is

$$u_{ij} = \beta_{ip}(y_i - p_j) + x_j\beta_i + \varepsilon_{ij} \tag{2}$$

where β_{ip} is the marginal utility of income. This model assumes quasi-linear utility with no income effects. Under these assumptions, the MWTP of individual i for characteristic k is

$$MWTP^D_{ik} = \frac{\beta_{ik}}{\beta_{ip}} \tag{3}$$

and the average MWTP for characteristic k is $MWTP^D_k = \int \frac{\beta_{ik}}{\beta_{ip}} dF_\beta$. An individual chooses product j that offers the highest utility. Let A_j be the set of individuals who choose j:

$$A_j = \left\{ \left(\beta_i, \beta_{ip}, \varepsilon_{i0}, \varepsilon_{i1}, ..., \varepsilon_{iJ} | u_{ij} \geq u_{ik}, k = 0, ..., J \right) \right\} \tag{4}$$

The share of individuals in a market who choose product j (π_j) is obtained from aggregating across individuals in A_j,

$$\begin{aligned} \pi_j(\mathbf{x},\mathbf{p}) &= \int_{A_j} dF_\beta dF_\varepsilon \\ &= \int \frac{exp(V_{ij})}{\sum_{j'=0}^{j'=J} exp\left(V_{ij'}\right)} dF_\beta \equiv \int \pi_{ij} dF_\beta \end{aligned} \tag{5}$$

where the second row shows that the probability that i chooses j (π_{ij}) is $\frac{exp(V_{ij})}{\sum_{j'=0}^{j'=J} exp\left(V_{ij'}\right)}$ because the $\varepsilon's$ are drawn from a Type I extreme value distribution. The share function shows that the probability that individual i chooses product j, depends on how product j compares to all products in the market. If product j, has characteristics that are favored by individual i, then π_{ij} will be higher, since the utility will be higher from product j (V_{ij}) relative to other products $(V_{ij'})$.

An equilibrium is characterized by a vector of prices for each product and an allocation of individuals to products so that no one has an incentive to deviate. The *share function*, $\pi(.)$, can be used to concisely summarize the optimizing choices individuals make in the discrete choice model. Given a fixed supply, in equilibrium, each element (π^*_j) in the J-dimensional vector (π^*) summarizes the share of individuals in a market who choose product j, as a function of product characteristics and equilibrium prices, evaluated at $(\mathbf{x}^*, \mathbf{p}^*)$.

2.2 Hedonic model and the hedonic price function

Rosen's hedonic model offers a dual way to describe an equilibrium in a market. Each market has a continuum of heterogeneous individuals and a continuum of products, differentiated along a vector of characteristics, $\mathbf{x}$. Individual i takes the market price for products, $P(\mathbf{x})$, as given and chooses one unit of a product to maximize utility, u_i, subject to the budget constraint, $P(\mathbf{x}) + numeraire \leq y_i$. The numeraire is normalized to have a price of 1.

An equilibrium in the hedonic model is characterized by individuals who are maximizing utilities given their budget constraints. Graphically, individual i's taste for x_k can be illustrated using bid functions (indifference curves in $P - x_k$ space) with steeper bid functions representing stronger taste (higher MWTP) for x_k. Each individual chooses a product that corresponds to the bid function that maximizes her utility. Under the first order conditions, optimality is achieved when the MWTP for x_k (the ratio of the marginal utility for x_k and the marginal utility for the numeraire) is equal to the ratio of the marginal cost for x_k and the marginal cost for the numeraire (normalized to 1). Under these assumptions, individual $i's$ MWTP for characteristic k is

$$MWTP_{ik}^{H} = \frac{\frac{\partial u_i}{\partial x_k}}{\frac{\partial u_i}{\partial P}} = \frac{\partial P}{\partial x_k} \tag{6}$$

In equilibrium, prices adjust so that each product is sold to the highest bidder and the marginal individual is just indifferent between a marginal gain in utility from choosing an additional unit of x_k and incurring a marginal cost for it (relative to the numeraire). Locally around the equilibrium, individuals are considering trade-offs along the indifference curves, with utility held constant. Therefore, the hedonic model is associated with compensated demand and compensated elasticities. Notably, with the assumption of quasi-linear utility for the discrete choice model, there is no difference between compensated and uncompensated demand.

Equilibrium interactions in the market trace out a price-characteristic $(P - x_k)$ locus that implicitly defines a market clearing, *hedonic price function*, $P(\mathbf{x})$. The hedonic price function is also the upper envelop of bid functions. Importantly, equation (6) delivers the famous insight from Rosen (1974) that the gradient (with respect to x_k) of the hedonic price function at each point is equal to individual $i's$ MWTP for x_k (the bid function for individual i is tangent to the hedonic price function at that point).

A major distinction between the hedonic and discrete choice approaches is the number of products. With a continuum of products, each individual is marginal. If the first order condition is not satisfied for individual i (if $MWTP_{ik} > \frac{\partial P}{\partial x_k}$ so that the relative marginal benefit from characteristic k is greater than the relative marginal cost, for example), she can always find a close substitute with marginally more characteristic k since there is a continuum of products. However, with discrete choices, some individuals could be inframarginal ($MWTP_{ik} > \frac{\partial P}{\partial x_k}$ in equilibrium). This can arise due to adjustment costs faced by consumers (mobility costs associated with changes in location choices, for example) or fixed costs of production faced by firms.

The insight that relates the gradient of the hedonic price function to MWTP is the foundation underlying many empirical applications of preference estimation. Most of the empirical estimations focus on estimating the hedonic price function, known as the first stage of the hedonic estimation method. It is well-known that the second stage of the hedonic estimation method is generally not identified (see Epple (1987) and Bartik (1987), and also Bishop and Timmins (2015) and Banzhaf (2015) for more recent discussions). For simplicity, this paper abstracts away from omitted variable concerns by assuming all characteristics are exogenous and assuming no unobserved product quality. I discuss identification concerns in Section 4.

3 Results

This section builds on the discrete choice and hedonic approaches described in Section 2 to provide a tractable framework to compare both approaches. The results in this section are derived holding the functional forms and distributional assumptions described in Section 2.1 fixed. I discuss some caveats to this model and possible extensions in Section 4. The analysis delivers the three results below. The first result is an analytical mapping between the share function and the gradient of the equilibrium price function associated with the discrete choice model.

Result 1: *Let the share function $\pi_1, ..., \pi_J : \mathbf{R}^{\mathbf{J(K+1)}} \to \mathbf{R}^{\mathbf{1}}$ be a C^1 function around the equilibrium point, $(\mathbf{x}^*, \mathbf{p}^*)$, which satisfies the system of J equations described in equation (7) below. Then, the gradient of the equilibrium price function can be expressed implicitly using the share function.*

$$
\begin{aligned}
\pi_1(p_1, ..., p_J, x_{11}, ..., x_{1K}, ..., x_{jk}, ..., x_{J1}, ..., x_{JK}) &= \pi_1^* \\
&\vdots \\
&\vdots \\
\pi_J(p_1, ..., p_J, x_{11}, ..., x_{1K}, ..., x_{jk}, ..., x_{J1}, ..., x_{JK}) &= \pi_J^*
\end{aligned} \tag{7}
$$

Proof. This result is an application of the Implicit Function Theorem (Theorem 15.7 in Simon and Blume (1994)). Consider the system of equations above as possibly defining $p_1, ..., p_J$ as implicit functions of $x_{11}, ..., x_{JK}$. The left hand side of each equation j is the share function for product j and the right hand side is the share of individuals in the market choosing that product. Suppose that (**p*, x***) is a solution of (7). 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 (**p*, x***) is nonzero (ie. the matrix is invertible), then there exist C^1 functions in $\mathbf{R^{J(K+1)}}$

$$\begin{aligned} P_1(x_{11},...,x_{JK}) &= p_1 \\ \\ P_J(x_{11},...,x_{JK}) &= p_J \end{aligned} \tag{8}$$

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

$$\begin{aligned} \pi_1(P_1(\mathbf{x}),...,P_J(\mathbf{x}),x_{11},...,x_{1K},...x_{jk},...,x_{J1},...,x_{JK}) &= \pi_1^* \\ &\vdots \\ \pi_J(P_1(\mathbf{x}),...,P_J(\mathbf{x}),x_{11},...,x_{1K},...x_{jk},...,x_{J1},...,x_{JK}) &= \pi_J^* \end{aligned} \tag{9}$$

for all $\mathbf{x} = (x_{11},...,x_{JK})$ in B and the gradient of this implicit function is

$$\begin{bmatrix} \frac{\partial P_1}{\partial x_{jk}} \\ \vdots \\ \frac{\partial P_J}{\partial x_{jk}} \end{bmatrix} = - \begin{bmatrix} \frac{\partial \pi_1}{\partial p_1} & \cdots & \frac{\partial \pi_1}{\partial p_J} \\ \vdots & \ddots & \vdots \\ \frac{\partial \pi_J}{\partial p_1} & \cdots & \frac{\partial \pi_J}{\partial p_J} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial \pi_1}{\partial x_{jk}} \\ \vdots \\ \frac{\partial \pi_J}{\partial x_{jk}} \end{bmatrix} \tag{10}$$

Since ε is Type I extreme value and independent from F_β, we know from (5) that $\pi_j = \int \frac{exp(V_{ij})}{\sum_{j'=0}^{j'=J} exp(V_{ij'})} dF_\beta$. Moreover, with random coefficients utility from (2), $\frac{\partial V_{ij}}{\partial x_{jk}} = \frac{\partial V_{ij}}{\partial x_{j'k}} = \beta_{ik}$ and $\frac{\partial V_{ij}}{\partial P_j} = \frac{\partial V_{ij}}{\partial P_{j'}} = \beta_{ip}$.

Therefore, the partial derivatives on the right hand side of equation (10) are:

$$\frac{\partial \pi_j}{\partial x_{jk}} = \begin{cases} \int \beta_{ik} \pi_{ij}(1-\pi_{ij}) dF_\beta & \\ \int \beta_{ik} \pi_{ij} \pi_{ij'} dF_\beta & j \neq j' \end{cases} \quad \text{and} \quad \frac{\partial \pi_j}{\partial P_j} = \begin{cases} \int \beta_{ip} \pi_{ij}(1-\pi_{ij}) dF_\beta & \\ \int \beta_{ip} \pi_{ij} \pi_{ij'} dF_\beta & j \neq j' \end{cases}$$

■

This result delivers an analytical relationship between the gradient of the share function and the gradient of the implicit price function. The steps from equations (7) to (9) use equilibrium conditions and the share function, $\pi(\mathbf{x}, \mathbf{p})$, to implicitly define price as a function of $\mathbf{x}$, $P(\mathbf{x})$. In general, with discrete choices, the equilibrium price function may not be continuous and the gradient may not be defined. However, the assumption of a continuum of agents with a distribution of Type I extreme value preference shocks and a discrete number of products gives rise to continuous share functions. The Implicit Function Theorem can then be used to derive the gradient of the implicit price function from the share function (equation (10).

Locally around the equilibrium point, ($\mathbf{p^*}, \mathbf{x^*}$), a small change in x_{jk} will induce individuals' choices to change, which in turn, leads to changes in market shares and equilibrium prices. One implication of using the Implicit Function Theorem for Result 1 is that the price function can only be defined locally around the equilibrium. This is valid for marginal changes in characteristics around the equilibrium, but not applicable for general equilibrium analysis with non-marginal changes in product characteristics.

Nevertheless, the analytical relationship in (10) is useful because it represents a mapping between the gradient of the equilibrium price function and the share function in the discrete choice model. Using Rosen's insight that connects the price gradient to the gradient of bid functions, this mapping can be used relate MWTP inferred from the equilibrium price function to MWTP in the discrete choice model.

While the analytical relationship in (10) is useful, it is still difficult to interpret because it is complicated by the inverse of the JxJ matrix in (10). This inverse shows that the equilibrium price for each product depends on all products in the market. This is reflected in the share function as π_{ij} depends on the ratio of individual i's utility from product j relative to the sum of her utility from all products in the market. With discrete choices, some products can have "market power" because the inframarginal individuals' choices are inelastic. By contrast, in Rosen's hedonic framework with a continuum of products, each product is small relative to the market and no one is inframarginal because each individual can always find a product so that her indifference curve is exactly tangent to the hedonic price function.

Next, the second result shows that the analytical relationship in Result 1 can be simplified if the equilibrium price function, $P(\mathbf{x})$, is only a function of own-product characteristics, so that $\frac{\partial P_j}{\partial x_{j'k}} = 0$ for $j \neq j'$.[5] It is a common assumption made in the empirical literature. For example, the hedonic price function is typically estimated by regressing the price of a product on the characteristics of that product only (but rarely on the characteristics of other products).

Result 2: *If the implicit price function defined in Result 1 is a function of own-product characteristics only, then, the gradient of this price function can be expressed as a ratio of weighted*

[5]For example, if k represents the square footage of a house, this assumption states that locally around the equilibrium, the price of house j depends on its square footage but the square footage of other houses will not affect its price.

averages of marginal utilities where the weights depend on choice probabilities in the discrete choice model.

Proof. The own-product assumption reduces the dimensionality of the price function from $\mathbf{R}^{\mathbf{J(K+1)}}$ to $\mathbf{R}^{\mathbf{(K+1)}}$. This is the major simplifying step that reduces the JxJ matrix in (10). To derive the second result, differentiate each row j of (9) with respect to x_{jk},

$$\begin{aligned} \frac{\partial \pi_1}{\partial P_1}\frac{\partial P_1}{\partial x_{1k}} &= -\frac{\partial \pi_1}{\partial x_{1k}} \\ &\vdots \\ \frac{\partial \pi_J}{\partial P_J}\frac{\partial P_J}{\partial x_{Jk}} &= -\frac{\partial \pi_J}{\partial x_{Jk}} \end{aligned} \tag{11}$$

where the additional terms on the left hand side of (11) are 0 now because $\frac{\partial P_j}{\partial x_{j'k}} = 0$. Therefore, we can re-write (11) as

$$\begin{bmatrix} \frac{\partial P_1}{\partial x_{1k}} \\ \vdots \\ \frac{\partial P_J}{\partial x_{jk}} \end{bmatrix} = -\begin{bmatrix} \frac{\frac{\partial \pi_1}{\partial x_{1k}}}{\frac{\partial \pi_1}{\partial P_1}} \\ \vdots \\ \frac{\frac{\partial \pi_J}{\partial x_{Jk}}}{\frac{\partial \pi_J}{\partial P_J}} \end{bmatrix} \tag{12}$$

where $\frac{\partial \pi_j}{\partial x_{jk}} = \int \beta_{ik}\pi_{ij}(1-\pi_{ij})dF_\beta$ and $\frac{\partial \pi_j}{\partial P_j} = \int \beta_{ip}\pi_{ij}(1-\pi_{ij})dF_\beta$.

■

Equation (12) indicates that the gradient of the implicit price function (defined in Result 1) can be written as a ratio of weighted averages of marginal utilities ($\frac{\partial P_j}{\partial x_{jk}} = \frac{\int w_{ij}\beta_{ik}dF_\beta}{\int w_{ij}\beta_{ip}dF_\beta}$), where the weights, w_{ij}, are a function of choice probabilities in the discrete choice model ($w_{ij} = \pi_{ij}(1-\pi_{ij})$). These weights represent the variance of individual $i's$ choices. Equation (12) gives 0 weight to individuals whose choice probabilities are 1 or 0. This is because these are individuals who will choose (not choose) a product with certainty (the variance of their choice is 0). Conversely, equation (12) gives the maximum weight to individuals whose choice probability is 0.5.[6] These

[6]The max at 0.5 is a consequence of the Type I extreme value assumption. This distributional assumption implies that the choice probabilities are drawn from a logistic distribution. This is because choices are driven by differences in random utilities and the difference between two random variables of Type I extreme value distribution is a random variable drawn from the logistic distribution. Logistic distributions have a cumulative distribution function that is sigmoid shape with the steepest slope at 0.5.

are individuals who have the highest choice variance and are on the margin of choosing or not choosing a product.

The key insight is that the hedonic method relies on the tangency between the bid functions and the price function (see first order conditions in (6)) but only the marginal individual's bid function is tangent to the price function. Therefore, the hedonic method gives a higher weight to marginal individuals whose choice probabilities indicate a higher degree of uncertainty regarding their choices. As this choice becomes more certain (π_{ij} approaching 0 or 1), the weights decrease.

Result 2 formalizes why the implications are different with discrete versus continuous products and shows how we can use other moments from choice data to determine which individuals are marginal versus inframarginal. Since the hedonic approach assumes a continuum of products, in principle, each individual is marginal because she can always find a product where her indifference curve is tangent to the price function. Therefore, the theory does not provide guidance on how to determine which types of individuals are more likely to be marginal. The analytical relationship between the share function and the price gradient provides a theoretical justification for using choice data and other moments (choice variance) to determine which individuals are more likely to be marginal. This complements the hedonic approach which typically only utilizes data on prices but not choice data.

In the case of random coefficients Logit, each individual is matched to a choice so everyone is inframarginal and experiences a choice-specific surplus, captured in the ε_{ij} term. In theory, the inframarginal individual's MWTP for characteristic k ($\frac{\beta_{ik}}{\beta_{ip}}$) should be greater than the marginal cost or the implicit price of characteristic k ($\frac{\partial P}{\partial x_k}$). Therefore, the gradient from the implicit price function with respect to characteristic k can serve as a lower bound for the inframarginal individual's MWTP for that characteristic ($\frac{\partial P}{\partial x_k} < \frac{\beta_{ik}}{\beta_{ip}}$). In practice, this intuition will tend to hold in settings where there are many choices spanning the support for characteristic k, so that each choice is relatively less important. However, the inequality may not hold if the choice-specific taste (ε_{ij}) is important. For example, an individual may still choose a house j even if her MWTP with respect to the square footage of the house is *less* than the implicit price for square footage ($\frac{\beta_{ik}}{\beta_{ip}} < \frac{\partial P}{\partial x_k}$), if she has a strong taste for that house (ε_{ij}). Small changes in square footage will not change her choice if the choice-specific term is important. Such an inframarginal individual is less likely to change her choices, and will have lower weights as a result.

Next, Result 3 shows how to use the simplified relationship in Result 2 to identify the conditions needed to compare MWTP from both approaches. To fix ideas, I focus on comparing the average MWTP associated with the discrete choice model in Section 2.1 and the average MWTP inferred from the gradient of the equilibrium price function. The average MWTP is typically the primary empirical object associated with the first stage of any hedonic analysis. In principle, the intuition should carry through as well for other moments of the MWTP distribution.

Result 3: *The average MWTP for characteristic k from the discrete choice model is equal to the average MWTP for characteristic k from the implicit price gradient if the ratios of marginal utilities* $\frac{\beta_{ik}}{\beta_{ip}}$ *are constant across all individuals. The traditional Logit model with no random coefficients satisfies this condition.*

Proof. This result compares the average MWTP for characteristic k estimated using the price gradient to the average MWTP in the discrete choice model, averaged across $MWTP_{ik}^D$ and $MWTP_{ik}^H$, as defined in (3) and (6), respectively

$$MWTP_k^D = \int \frac{\beta_{ik}}{\beta_{ip}} dF_\beta \tag{13}$$

$$MWTP_k^H = \frac{1}{J}\sum_j \frac{\partial P_j}{\partial x_{jk}} = \frac{1}{J}\sum_j \frac{\int w_{ij}\beta_{ik} dF_\beta}{\int w_{ij}\beta_{ip} dF_\beta} \tag{14}$$

since the average MWTP estimated from the price gradient is $\frac{\partial P_j}{\partial x_{jk}} = \frac{\int w_{ij}\beta_{ik} dF_\beta}{\int w_{ij}\beta_{ip} dF_\beta}$ from equation (12).

Strikingly, the two estimates of average MWTP will generally be different because the discrete choice method estimates the average of ratios ($\int \frac{\beta_{ik}}{\beta_{ip}} dF_\beta$) and the price gradient approach depends on the ratio of (weighted) averages. For $MWTP_k^D = MWTP_k^H$, we need the ratios ($\frac{\beta_{ik}}{\beta_{ip}}$) to be constant across all $i's$ so that the average of ratios equals the ratio of averages. If $\beta_{ik} = c\beta_{ip}$ for all i and for some constant c, then, $MWTP_k^D = \int \frac{\beta_{ik}}{\beta_{ip}} dF_\beta = c$ and $\frac{\partial P_j}{\partial x_{jk}} = \frac{\int w_{ij}\beta_{ik} dF_\beta}{\int w_{ij}\beta_{ip} dF_\beta} = \frac{\int c w_{ij}\beta_{ip} dF_\beta}{\int w_{ij}\beta_{ip} dF_\beta} = c$. So, $MWTP_k^H = c$ also.

The traditional Logit model with no random coefficients satisfies this condition. Without random coefficients, equation (2) reduces to $u_{ij} = \bar{\beta}_p(y_i - p_j) + x_j\bar{\beta} + \varepsilon_{ij} = V_j + \varepsilon_{ij}$. So, $MWTP_k^D = \int \frac{\beta_{ik}}{\beta_{ip}} dF_\beta = \frac{\bar{\beta}_k}{\bar{\beta}_p}$. Also, the share function simplifies from $\pi_j = \int \frac{exp(V_{ij})}{\Sigma_{j'=0}^{j'=J} exp(V_{ij'})} dF_\beta$ to $\frac{exp(V_j)}{\Sigma_{j'=0}^{j'=J} exp(V_{j'})}$.

Therefore, applying (12) to the simplified share function, $\frac{\partial \pi_j}{\partial x_{jk}} = \bar{\beta}_k \pi_j(1-\pi_j)$ and $\frac{\partial \pi_j}{\partial P_j} = \bar{\beta}_p \pi_j(1-\pi_j)$. And, $\frac{\partial P_j}{\partial x_{jk}} = -\frac{\frac{\partial \pi_j}{\partial x_{jk}}}{\frac{\partial \pi_j}{\partial P_j}} = \frac{\bar{\beta}_k \pi_j(1-\pi_j)}{\bar{\beta}_p \pi_j(1-\pi_j)} = \frac{\bar{\beta}_k}{\bar{\beta}_p}$ for all j. Therefore, $MWTP_k^D = MWTP_k^H = \frac{\bar{\beta}_k}{\bar{\beta}_p}$. ■

Intuitively, without random coefficients, there is no heterogeneity in the taste for product characteristics and only heterogeneity in the taste for products (ε_{ij}). So, $MWTP_{ik}$ is constant across individuals ($\frac{\beta_{ik}}{\beta_{ip}} = \frac{\bar{\beta}_k}{\bar{\beta}_p}$). Individuals have bid functions with identical slopes with respect to k. This is akin to having a representative consumer, a special condition discussed in Rosen (1974) where all the bid functions effectively collapse to one bid function for the representative consumer, and this bid function is also the hedonic price function.

Implementing Rosen's insight here, the implicit price function has a constant gradient with respect to k equal to the constant, $\frac{\bar{\beta}_k}{\bar{\beta}_p}$. In other words, the representative consumer is also the average consumer *and* the marginal consumer so there is no wedge between $MWTP_k^D$ (which estimates the MWTP for the average consumer) and $MWTP_k^H$ (which gives higher weights to marginal consumers).

The empirical object of interest is the average of the ratios $(\frac{\beta_{ik}}{\beta_{ip}})$ rather than the ratio of the average. Result 3 clarifies when the two will diverge. This will depend on (i) the covariance between the distribution for β_{ik} and the distribution for β_{ip} and (ii) the choice weights and how different are the average versus the marginal individual. There is no wedge if the marginal consumer is the same as the average consumer. Otherwise, the wedge will be greater if the inframarginal individual is different from the marginal individual. For example, when a characteristic is in limited supply, the marginal individual will have a higher MWTP than the average individual (Bayer et al., 2007).

4 Discussion

So far, the results above have been derived from the discrete choice model specified in Section 2.1. This section discusses some additional caveats and possible extensions of the framework above.

Functional Forms: There are several functional form assumptions to highlight. First, besides the Logit model and the Type I extreme value distribution for the idiosyntractic taste shock, the intuition above should carry through for other distributions as well, as long as the share functions in the discrete choice model will be continuous. In principle, the exercise above can be applied to a Probit model but the gradients and partial derivatives will be more complicated. Second, I have focused on the average MWTP, following the empirical literature, but the findings above can be generalized to other moments of the MWTP distribution as well.[7]

Third, the quasi-linear utility assumption is important so that there are no income effects and compensated and uncompensated demand will be the same. However, the intuition that the price gradient approach will generally underweight inframarginal individuals should still hold with income effects. For example, if a characteristic is normal in that demand for it increases with income, then, all else equal, MWTP based on uncompensated demand would generally be stronger than MWTP based on compensated demands due to the income effect. From an applied perspective, the choice of compensated versus uncompensated demand will depend on the setting. Public finance applications that estimate the deadweight loss of taxation tend to focus on compensated demand (Hausman, 1981). Research on consumption inequality would rely on uncompensated de-

[7] See Chay and Greenstone (2004) for more discussion on the importance of the average MWTP and how to generalize beyond the average MWTP.

mand functions to model the implications of income effects (see, for example, Handbury (2013) and Diamond (2016)).

Non-marginal changes: Since the Implicit Function Theorem only applies locally around the equilibrium points, the revealed preference arguments above center around a given equilibrium, and use the equilibrium objects associated with that equilibrium (share functions and price functions) to infer preferences. This method cannot be readily extended to analyses with non-marginal changes (that might cause the equilibrium to change) or general equilibrium analyses with endogenous supply responses as well. The distinction between welfare analyses for marginal and non-marginal changes in product characteristics has been highlighted in the literature (see Chay and Greenstone (2004) and Bartik (1988) on similar limitations arising from Rosen's hedonic framework).

Identification: This paper assumes no unobserved product quality and that all product characteristics are exogenous. The goal of this exercise is to abstract away from omitted variable bias due to unobserved product quality (which has been studied extensively in the literature), to highlight selection concerns that will arise in a context with unobserved heterogeneity in *individuals*. Here, the issue is that the empiricist cannot determine which individuals are marginal and which individuals are not, without observing their preferences. The results above show how to use choice data to determine which individuals are more likely to be marginal. Instead of relying on functional forms, if there was repeated choice data, it may also be possible to use this data to estimate the variance of choices.

5 Conclusion

Marginal willingness-to-pay (MWTP) is important for welfare analysis. The two primary approaches to estimate MWTP are hedonic (Rosen, 1974) and discrete choice (McFadden, 1974). The innovation in this paper is to provide a tractable framework that delivers a novel analytical mapping between MWTP in the discrete choice model and MWTP inferred from the price gradient. The first result shows how to use the share function in the discrete choice model to define the gradient of the equilibrium price function implicitly. Second, if we further assume that the equilibrium price function is only a function of own-product characteristics (a common assumption in the empirical literature), then, the gradient of the price function depends on a ratio of weighted averages of marginal utilities with higher weights for individuals with more uncertainty in choices (marginal individuals). Third, the average MWTP's will only be identical if MWTP is constant across individuals. The traditional Logit model without random coefficients satisfies this condition. My analysis relies on distributional and functional form assumptions commonly used in the empirical literature and holds these assumptions fixed throughout the paper. However, the main insights can be readily generalized to other settings.

The framework presented here resonates with previous work that have alluded to the duality between both the hedonic and discrete choice approaches. To my knowledge, this paper provides the first tractable framework that directly compares both approaches and establishes the (strong) assumptions required. Moreover, the framework illustrates why MWTP can differ and how one can use choice data to examine which types of individuals are more likely to be marginal, thereby complementing the hedonic approach that typically uses price data only. Empiricists have traditionally chosen one approach over another based on the availability of data, computational costs, and rules-of-thumb, such as the number of products in the market. The exercise in this paper highlights other differences between the two approaches to be considered. Important directions for future work include analyses using other functional forms, incorporating supply adjustments, welfare analyses with non-marginal changes in product attributes, and incorporating unobserved product quality.

References

Arnott, R. (1989). Housing Vacancies, Thin Markets, and Idiosyncratic Tastes. *The Journal of Real Estate Finance and Economics*, 2(1):5–30.

Bajari, P. and 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.

Banzhaf, S. (2015). Panel Data Hedonics: Rosen's First Stage and Difference-in-Differences as "Sufficient Statistics". Working paper.

Bartik, T. (1988). Measuring the Benefits of Amenity Improvements in Hedonic Price Models. *Land Economics*, 64(2):172–183.

Bartik, T. J. (1987). The Estimation of Demand Parameters in Hedonic Price Models. *Journal of Political Economy*, 95(1):81–88.

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

Bayer, P., Keohane, N., and Timmins, C. (2009). Migration and Hedonic Valuation: The Case of Air Quality. *Journal of Environmental Economics and Management*, 58(1):1–14.

Berry, S. T. (1994). Estimating Discrete-Choice Models of Product Differentiation. *The Rand Journal of Economics*, 25(2):242–262.

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

Bishop, K. and Timmins, C. (2015). Estimating the Marginal Willingness to Pay Function without Instrumental Variables. Working paper.

Bitzan, J. and Wilson, W. W. (2007). A Hedonic Cost Function Approach to Estimating Railroad Costs. In Dennis, S. and Talley, W. K., editors, *Research in Transport Economics: Railroad Economics*, volume 20, pages 119–152.

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

Chay, K. and 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., and 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.

Diamond, R. (2016). The determinants and welfare implications of us workers' diverging location choices by skill: 1980-2000. *American Economic Review*, 106(3):479–524.

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

Ellickson, B. (1981). An Alternative Test of the Hedonic Theory of Housing Markets. *Journal of Urban Economics*, 9(1):56–79.

Epple, D. (1987). Hedonic Prices and Implicit Markets: Estimating Demand and Supply Functions for Differentiated Products. *Journal of Political Economy*, 95(1):59–80.

Gavazza, A. (2011). The Role of Trading Frictions in Real Asset Markets. *American Economic Review*, 101(4):1106–43.

Handbury, J. (2013). Are Poor Cities Cheap for Everyone? Non-Homotheticity and the Cost of Living Across US Cities. Working Paper.

Hausman, J. A. (1981). Exact consumer's surplus and deadweight loss. *The American Economic Review*, 71(4):662–676.

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

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

Pakes, A. (2003). A Reconsideration of Hedonic Price Indexes with an Application to PC's. *American Economic Review*, 93:1578–96.

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

Simon, C. P. and Blume, L. (1994). *Mathematics for Economists*. W.W. Norton, New York, NY.

Wong, M. (2013). Estimating Ethnic Preferences Using Ethnic Housing Quotas in Singapre. *Review of Economic Studies*, 80(3):1178–1214.