the review of income and wealth

Review of Income and Wealth
Series 0, Number 0, 2022
DOI: 10.1111/roiw.12612



# AN INDEX APPROACH TO MEASURING PRODUCT DIFFERENTIATION: A HEDONIC ANALYSIS OF AIRFARES

BY CHARLES HOWELL*

*Department of Business, Universitat Autònoma de Barcelona*
*College of Business Administration, University of Nebraska Omaha*

AND

EMILI GRIFELL-TATJÉ

*Department of Business, Universitat Autònoma de Barcelona*

The main objective of this paper is to introduce an Allen-type index of differentiation based on cost functions. With this index, we create an economic measure of product differentiation that quantifies differences between products. Applied research has some generally accepted economic measures, for example, the Herfindahl–Hirschman Index for market concentration, or the Gini coefficient for inequality. Product differentiation, however, does not yet have an established measure. Our objective is to fill that gap and introduce a measure that can be used in market-related applied research such as market power, antitrust, price indexes, or market strategy. To operationalize the index, we introduce the concept of a core product and use cost functions to measure the degree of differentiation from the core product. To demonstrate the use of the index, we study the effect of product differentiation on price formation in the airline industry using an enhanced hedonic model. The model is empirically tested on 103,980 observations of quarterly US domestic airfare data between 2002 and 2016 and shows that product differentiation has a significant effect on both price and mark-up.

**JEL Codes:** D24, D40, D43, L93

**Keywords:** airline industry, core product, firm heterogeneity, hedonic function, pricing, product differentiation index

## 1. INTRODUCTION

The concept of product differentiation is generally understood, but not very well defined. Sharp and Dawes (2001) note that differentiation is a concept that has vague meanings and is often referenced without a formal definition. Because empirical work on the subject often uses non-complementary measures and operationalizations, differentiation is simplified to the point where it has little practical application or is very case specific. In this paper, we develop a model that uses an economic measurement of product differentiation that can be applied to a wide range of products and services and provides sufficient detail to answer

*Note*: We would like to give a special thanks to the two anonymous reviewers, who provided invaluable input to this study. We are also grateful to the Spain's Ministry of Economy, Industry and Competitiveness PID2020-114460GB-C31, for its generous financial support of this research.

*Correspondence to: Charles Howell, College of Business Administration, 6708 Pine Street, Omaha, NE 68182-0048, USA (cbhowell@unomaha.edu).

relevant economic and business questions. The index we develop can be applied to market-based questions on topics such as market power, antitrust (Milne, 1992), or price indexes (Ivancic and Fox, 2013). Our model synthesizes three elements, two from well-established theoretical literature streams, and the third our own methodological contribution. After developing the model, we demonstrate its usefulness in an empirical application of a modified hedonic approach to the study of airfares.

We use the term "characteristics" to refer to the properties of a good or service that create differentiation. Our use of this terminology is rooted in the seminal work of Lancaster (1966), who assumes that products are valued not only for their physical attributes but also for their non-physical attributes. In this work, utility is derived not from the product, but from the properties or characteristics of the product and other products and services consumed in conjunction. A different mix of properties or characteristics generates differentiation.

To operationalize the discussion of characteristics, we use a modified version of a product model proposed by Kotler (1967). In Kotler's model, the product is composed of different levels. We associate the concept of *core product* with the level that defines the most basic configuration of the product that can be purchased. In the terminology of Caves and Williamson (1985), this core product defines a product class of close substitutes that meet the consumer's needs. In our airline example, the core product would consist of a single departure time, a seat, a distance traveled, and a landing. The next level extends the product beyond core characteristics to include features that buyers expect when they purchase the product from a particular firm. These additional features are available to all buyers of the firm's product and can create differentiation between firms. For airlines, these features include characteristics such as on-time performance, or frequency of flights.

Our primary methodological contribution in this paper is the introduction of a product differentiation index based on cost functions that enables the application of theory to empirical studies. Applied research in areas such as market power, antitrust, or price indexes could employ this index. We demonstrate the use of the product differentiation index in the applied section with an empirical analysis of airfares, showing how the measure can be used to enhance the results of the hedonic approach. In this empirical contribution, we analyze key economic and business questions regarding product differentiation like: How much will it cost to provide the characteristic? How much do consumers value the characteristic? Can a profit be earned by changing the level of the characteristic?

The hedonic pricing model is frequently used to measure consumer valuation of product characteristics. Rosen (1974) defines hedonic prices as the set of implicit prices of characteristics that are revealed from the observed level of characteristics and prices of the product. However, empirical applications of the hedonic approach often provide contradictory results. The enhanced hedonic approach we present alleviates some of these issues and provides a more useful tool for practitioners. First, we include the index of product differentiation, and second, we use the concept of "core product" to disentangle customer valuation for characteristics correlated with the cost of the product. A key finding in the applied section is that having more competitors may reduce the basic price, but it increases the value of differentiation.

Our economic measurement method contributes to the line of research on product differentiation (Kotler, 1967; Caves and Williamson, 1985; Ivancic and Fox, 2013; Kostovetsky and Warner, 2020) and provides a path for future empirical applications. In the applied portion, we contribute to research into hedonic applications (Chwelos *et al.*, 2008; Cotteleer *et al.*, 2008; Heckman *et al.*, 2010; Bourassa *et al.*, 2016). Our inclusion of an economic measure of firm product differentiation is a unique contribution to this stream of literature. We also contribute to the literature using hedonic models to analyze airfares (Morrison and Winston, 1995; Good *et al.*, 2008) To the best of our knowledge, our use of product characteristics to explain mark-up and our definition of mark-up are unique contributions to the hedonic literature, as is our inclusion of estimated costs directly within the hedonic pricing model.

This paper is organized as follows. Section 2 provides background on hedonic functions. Section 3 covers methodology, developing the core product concept and the production differentiation index. Section 4 presents the data used in the study, Section 5 presents the results, and Section 6 concludes.

## 2. Background

In a seminal work, Rosen (1974) provides a theoretical foundation for hedonic analysis[1] of differentiated products by combining a family of utility-maximizing bid functions and profit-maximizing offer functions. A hedonic equation then maps the locus of consumer bid and producer offer functions for the characteristics that make up a product.[2] From the producer point of view, the equation parameters indicate the marginal reservation supply price, and from the consumer view, a marginal valuation. This theoretical foundation is expanded by Feenstra (1995) and Pakes (2003). They extend the model to imperfect competition and show that when mark-ups exist, the characteristic hedonic function is a sum of the marginal cost function and a function that summarizes the relationship between mark-ups and characteristics. When either of these "primitives" changes, we would expect the hedonic coefficient to change.

Defining the variables used in the following sections, we denote output quantity as $y \in \mathbb{R}_{++}$, the vector of inputs as $x \in \mathbb{R}^m_+$, the vector of input prices as $w \in \mathbb{R}^m_{++}$, the vector of output prices as $p \in \mathbb{R}^l_{++}$, and finally $z \in \mathbb{R}^n_+$ as a vector of $n$ characteristics. Following Rosen (1974), the hedonic model defines *a class of products* that can be described by the vector of characteristics $z$, where $z = \left(z_1, z_2, \dots, z_n\right)$. The components of $z$ are objective measures that are perceived by all consumers and producers identically but may be valued differently. Each product has an observed market price and associated characteristic vector $z$ such that the market for the product implicitly reveals a function $p(z) = p\left(z_1, z_2, \dots, z_n\right)$. The function $p(z)$ represents the minimum price for any bundle of characteristics. Rosen (1974) adopts the convention of measuring characteristics $z_i, i = 1, \dots n$, in a way that all can be considered "goods." To estimate the hedonic function, we define $p_j$ as the observed

[1]For surveys of the method to airfares see (Armknecht and Ginsburg, 1992; Good *et al.*, 2008) and for general surveys (Triplett, 2004; Hill, 2013).

[2]See Rosen (1974) for a full development of this model and additional details.

price of product $j$ with observed characteristic vector $z_j$. We relate $p_j$ with each of the characteristics $z_i, i = 1, \dots n$ and add an error term. A simple example of a linear form would be,

$$p_j = \beta_0 + \sum_{i=1}^{n} \beta_i z_i + \varepsilon. \tag{1}$$

In the model defined by Rosen (1974), the function defines the intersections of bid and offer curves for characteristics, forming an envelope linking together the equilibriums. Characteristics $z_i$ are considered both customer value driven, and producer cost or resource driven. Extending the model to include imperfect competition, Pakes (2003) shows that if we let $\left(z_j, p_j\right)$ denote the vector of characteristics and price of $j$ and $\left(z_{-j}, p_{-j}\right)$ the vector of characteristics and price for all other products, then demand for $j$ is $D_j(\cdot) = D\left(z_j, p_j, z_{-j}, p_{-j}, A\right)$ where $A$ is the distribution of consumer preferences for characteristics. Assuming a single product firm and marginal cost of $mc(\cdot)$, then $p_j = mc\left(z_j\right) + D_j(\cdot)/\left|\partial D_j/\partial p\right|$, where the second term is typically referred to as the mark-up. Simply put, the expected price of $j$ is the marginal cost of producing characteristics $z_j$ and a mark-up based on characteristics $z_j$ and other available products. We can define the hedonic function $h\left(z_j\right)$ for product $j$, as the expectation on price conditional on $z$ as

$$h\left(z_j\right) \equiv E\left[p_j \middle| z_j\right] = mc\left(z_j\right) + E\left(\frac{D_j(\cdot)}{\left|\partial D_j/\partial p\right|} \middle| z_j\right). \tag{2}$$

The mark-up can be described as a complex function that varies inversely with elasticity of demand and is dependent on characteristics of competing goods and consumer preferences. Interpretation of the coefficient can be difficult because it embodies both the marginal cost function and the mark-up function.

One of the more common applications of the hedonic method is in producing quality-adjusted price indexes. As products or services change over time, a hedonic function can separate price change resulting from inflation versus quality change. Triplett's (2004) handbook makes the point that accounting for quality change is one of the most difficult problems for price index compilers. The handbook also notes that the service characteristics we focus on in this paper are often overlooked in hedonic analysis and price indexes.

The hedonic method has also been applied to cost functions. Spady and Friedlaender (1978) suggest treating output as a function of physical output and its characteristics. Their method has been applied to industries such as freight transport (Spady and Friedlaender, 1978), airlines (Gillen *et al.*, 1990), and the military (Hanson, 2016). Our use of characteristics within the cost function contributes to this line of research.

## 3. Methodology

### 3.1. *The Core Product*

In this paper we look at the business literature in addition to the economic literature. The concept of the product plays a central role in business literature. In the introduction we referenced the product model theorized by Kotler (1967). This model, extended by Levitt (1980), has been used in studies such as credit card services (Goyal, 2006) and wine tourism (Duan *et al.*, 2018). However, applications of the model remain mostly at a theoretical and conceptual level due to the lack of a method to measure change between and within levels. Kotler (2011, p. 87) referred to himself as a *market economist*; our connection of the hedonic method to the product model is very much in that vein.

Consumer wants or needs for the benefit from a good or service are the starting point for this model. In the airline example, that would simply be transportation from one location to another. In the first level of the model, a product is defined that can satisfy the consumer's need and can be purchased. It has been variously termed as the *basic product*, *generic product*, or the term we have chosen, *core product*. At this first level, the firm has turned the consumer's need into a product that can be sold and defined by characteristics. This core product includes only the level of characteristics needed to fulfill the basic service or benefit. In the airline example, these characteristics would be a single departure time, a seat, a distance traveled, and a landing. We term the characteristics that define the core product as *core characteristics*.

The next level is where firms can introduce *differentiating characteristics* that can distinguish their product from those of competitors and create a *differentiated product*. These characteristics define what a consumer expects from a specific firm, and since consumers differ in expectations, these expectations can vary between firms. A consumer traveling for leisure purposes may have low expectations on the number of flights available, while the business traveler may expect more departure-time options.

In a previous section, we defined $z \in \mathbb{R}_+^n$ as a vector of $n$ characteristics that define a product. To operationalize the concept of the core product, we begin by defining $\bar{z} \in \mathbb{R}_+^n$ as the vector of core characteristics where $\bar{z} \subseteq z$, in other words $\bar{z}$ cannot have more elements than $z$. The approach to defining the core product is to consider the minimum values of characteristics available in the market for product $j$ at time $t$, that is $\bar{z} = \left\{\bar{z}_1, \ldots, \bar{z}_n\right\}$ and $\bar{z}_{ijt} = Min \left\{z_{ijt}, j = 1, .., k\right\}, i = 1, .., n$. We also define $d \in \mathbb{R}_+^n$ as the vector of differentiating characteristics or the difference between $z_{ijt}$ and $\bar{z}_{ijt}$. This definition implies $z_{ijt} = \bar{z}_{ijt} + d_{ijt}, i = 1, .., n; j = 1, .., k$. In other words, the characteristics for any product $j$ are equal to the value of the core characteristic, and the value of any differentiating characteristic(s). It is important to note that characteristics here are measured as levels and that the value of a core characteristic can be zero, meaning that the characteristic is not provided.[3]

Relating this definition of the core product to the Kotler (1967) product model, vector $\bar{z}$ defines the first level of the model, or a purchasable product. In concept, this core product represents the minimal product acceptable to the market. It can also be thought of as a benchmark product, with all other products being an extension

[3]As a numerical example, picture a market with two airlines, Delta and Southwest. Now consider $z$ to include three characteristics (*Flight Frequency*, *On-Time Performance*, *Seat Selection*) and that *Seat Selection* is measured as 1 if available and 0 if not. If characteristics $z$ are respectively (*4*, *75 percent*, *1*) and (*2*, *80 percent*, *0*), then $\bar{z}$ would include (*Flight Frequency*, *On-Time Performance*, *Seat Selection*) and values for this market would be (*2*, *75 percent*, *0*), $d_{Delta}$ would be (*2*, *0*, *1*) and $d_{Southwest}$ would be (*0*, *5 percent*, *0*).

of the benchmark. For a firm, the core product defines the basic product needed to enter the market at that time. The addition of time $t$ is part of operationalizing the product model, since what is considered the core product can change over time as consumer tastes and preferences change. In addition, we see that vector $z$ defines the second product level, the differentiated product, with vector $d$ defining the difference between the two levels.

Relating this operation definition to a hedonic function, we extend (Pakes, 2003) and let $\left(\bar{z}, d_j, p_j\right)$ denote the core characteristics, differentiating characteristics and price of $j$, and $\left(\bar{z}, d_{-j}, p_{-j}\right)$ for all other goods. Demand for $j$ now becomes $D_j(\cdot) = D\left(\bar{z}, d_j, p_j, d_{-j}, p_{-j}, A\right)$ where $A$ again is the distribution of consumer preference over characteristics. With this demand definition we extend equation (2) to include both the core product and differentiating characteristics as

$$(3) \qquad h\left(z_j\right) \equiv E\left[p_j \left| z_j\left(\bar{z}, d_j\right)\right.\right] = mc(\bar{z}) + mc\left(d_j\right) + E\left(\frac{D_j(\cdot)}{\left|\partial D_j / \partial p\right|} \left| z_j\left(\bar{z}, d_j\right)\right.\right).$$

In equation (3) we find the expectation on price of $j$ conditional on $z_j\left(\bar{z}, d_j\right)$ as the sum of the cost of producing core characteristics $\bar{z}$, the cost of producing differentiating characteristics $d_j$, and a mark-up conditional on $z_j\left(\bar{z}, d_j\right)$. This extension of Pakes (2003) separates cost into two components, the cost of producing the core product and the cost of producing differentiating characteristics beyond the core. We will extend equation (3) further in Section 3.4.

Generalizing this model beyond the industry studied in this paper, we can see how this core product can be generally applied. A good example, and one often analyzed by a hedonic function, is the personal computer industry. The *Consumer Price Index Manual* (ILO et al., 2004, p. 116) explains price differences in the UK market for personal computers. In the example, the results in table 7.4 of the publication present coefficients for various characteristics such as speed, RAM, brand, or other accessories. The model also allows for imputing prices in periods where a price is not available. Since $\bar{z}$ represents the minimum value of characteristics available, the core product may represent a product not yet available in the market. However, this is not an expected result. In the previous example from the *Consumer Price Index Manual*, the core product would be defined by the Compaq Prosignia with a Celeron processor, which has the minimum value of characteristics given by speed, RAM, and HD.

We focus on $z$ and $\bar{z}$ in operationalizing the product model and trying to understand the relative difference between the core product and the differentiated product. While defining the two vectors is relatively straightforward, creating a measurement that can define relative degrees of product differentiation requires another step. In the next section, we define the methodology to create this measurement.

### 3.2. *Product Differentiation Index*

To begin, we define the technology as production possibility set $T$, and define the associated product characteristics that can be produced for a given set of inputs as $T = \{(x, y, z) : x \text{ can produce } y \text{ with characteristics } z\}$, which is assumed to satisfy the usual properties of "no free lunch," closed, bounded for each $x$, monotonicity, and convexity. Corresponding to technology set $T$, we define $L$ as the input set of $x$ required to produce an output quantity $y$ with characteristics $z$ as $L(y,z) = \{x : (x,y,z) \in T\}$. In other words, output quantity $y$ with characteristics $z$ can be produced with the input vector $x$. Adding input prices $w$, we define the least expensive bundle of inputs to generate a given output quantity and level of characteristics. The minimum cost for a given output quantity $y$, characteristic level $z$, and input prices $w$ is defined as

$$c(w,y,z) = \min_{x} \left\{ w^T x : x \in L(y,z) \right\}. \tag{4}$$

From equation (4) we see that the minimum cost would be greater with an increase in output quantity $y$, input prices $w$, or characteristic level $z$. Function $c(w,y,z)$ is non-decreasing in $y$ and $z$ and non-decreasing, concave, and homogeneous of degree +1 in $w$. Recalling the core product definition, we would find that $z \geq \bar{z}$ and therefore $c(w,y,z) \geq c(w,y,\bar{z})$, where $c(w,y,\bar{z})$ defines the minimum total cost of providing output quantity $y$ with core characteristics $\bar{z}$ and input prices $w$.

To introduce and operationalize the product differentiation index, we need to refine our notation of the product. In the previous sections, the simplified notation "a product $j$" has been used. We now move to a notation that includes the firm $h$, a market $k$, and a period $t$. This notation allows us to compare the products available in a specified market during a particular time and facilitates the following definition of product differentiation index.

Definition: Product Differentiation Index

Product differentiation of firm $h$, in market $k$, and period $t$ is measured by

(5)

$$\begin{aligned} Q\left(w_{hkt}, y_{hkt}, z_{hkt}, \bar{z}_{kt}\right) &= \frac{c\left(w_{hkt}, y_{hkt}, \bar{z}_{kt}\right)}{c\left(w_{hkt}, y_{hkt}, z_{hkt}\right)}; h = 1, .., l; k = 1, \dots, K; t = 1, \dots, T \\ &= \frac{w_{hkt}^T x'(y,\bar{z})_{hkt}}{w_{hkt}^T x''(y,z)_{hkt}}; h = 1, .., l; k = 1, \dots, K; t = 1, \dots, T \end{aligned}$$

This index of product differentiation takes the input price vector $w$ and the output quantity vector $y$ as given and compares the cost of producing $y$, with input prices $w$, under two characteristic vectors $z$ and $\bar{z}$. The second row shows an alternative way of expressing the product differentiation index in equation (5), where "$^{\mathrm{T}}$" defines the transpose of the input price vector and $x''(y,z)$ the vector of input quantities that minimizes cost associated with the vector of output quantities $y$ with characteristics $z$, where $c\left(w_{hkt}, y_{hkt}, z_{hkt}\right) = w_{hkt}^T x''(y,z)_{hkt}$. Similarly, $x'(y,\bar{z})$ expresses the vector of input quantities that minimize cost associated with the vector of output quantities $y$ with characteristics $\bar{z}$, where $c\left(w_{hkt}, y_{hkt}, \bar{z}_{kt}\right) = w_{hkt}^T x'(y,\bar{z})_{hkt}$. Note that in the context of expression (5), the output quantity $y$ can be defined by an output quantity vector.

In the case of one input, the product differentiation index in equation (5) is equal to $x'(y,\bar{z})/x''(y,z)$, the ratio of the single optimal input quantity needed to produce the output quantity $y$ with the observed characteristics $z$ to the corresponding optimal input quantity to produce the same amount of product with the core characteristics $\bar{z}$. It has the identity axiom because it takes the value equal to 1 when $\bar{z} = z$. Moreover, as the denominator of equation (5) cannot be lower than the numerator $c(w, y, z) \geq c(w, y, \bar{z})$ and the expression (5) ranges (0, 1]. Notice, it is bounded with the same rank of results as the Herfindahl–Hirschman Index (HHI) and the Gini coefficient. In a general case with multiple inputs, it measures the difference in cost between the observed level of characteristics of firm $h$, in market $k$, in period $t$, and the level in the core product. A product differentiation index equal to 1 indicates that the characteristics provided by the firm are equal to the core product, while a value less than 1 indicates differentiation. For example, a value of 0.90 would imply a core cost that is 10 percent less than the minimum cost with observed characteristics $z$ by firm $h$, in market $k$, and period $t$.

Allen (1949, p. 199), in a context defined by the consumer theory, states "to value the quantities purchased at certain prices. Fixing one set of prices (p), we define an expenditure line across the preference map joining the points (q) where a price plane with slopes (p) is tangential to the indifference surfaces. We can then measure the volume of consumption by taking indifference levels as ordered along this line and using distance between tangent planes. … A perfectly definite index of volume change is thus defined. But all this relates to the choice of prices (p). There is a measure of the volume index $\mathrm{Q}_{12}(\mathrm{p})$ for each price set (p) selected." We can transfer Allen's contribution from consumer theory to the production theory. In this new context, indifference levels are replaced by input isoquants and the points of tangency choose the optimal input quantities that minimize cost to produce different output quantities with an identical vector of input prices. Allen's volume index can be defined as $Q_{12}(w) = c\left(w, y^2\right) / c\left(w, y^1\right)$, which compares the cost of producing the output quantity $y^1$ with the output quantity $y^2$, using the same input prices $w$.[4] Therefore, the structure of Allen's quantity index is similar to that of the product differentiation index in equation (5). The former fixes input prices and measures the variation in cost associated with output quantity changes, while the latter fixes input prices and output quantities and measures the increase in cost associated with the provision of higher level of characteristics. Figure 1 visualizes the product differentiation index in equation (5).

It is worth noting that the product differentiation index in equation (5) does not define a common numerator for all firms in a market $k$ and period $t$. In fact, it is related to the choice of $w$ and $y$ and, for this reason, it can be firm specific. It may change from firm to firm, because it defines the minimum cost of producing the core characteristics $\bar{z}$ with the input prices and quantities specific to each firm. Therefore, what is common to all firms in a market $k$ and period $t$ is the provision of products with identical characteristics $\bar{z}$. This is one reason why the index in equation (5) can be interpreted as a measure of product differentiation. We can directly compare the results of the product differentiation index in equation (5) among firms of different sizes and between different markets and periods. Furthermore, a value of the product differentiation index equal to, e.g., 0.85 means that the product differentiation of this firm is higher than that of another firm of the highest value, e.g., 0.90. It is possible to rank firms based on the product differentiation index in equation (5)

[4]See also Diewert (2008 p. 16) and Grifell-Tatjé and Lovell (2015, Ch 7).

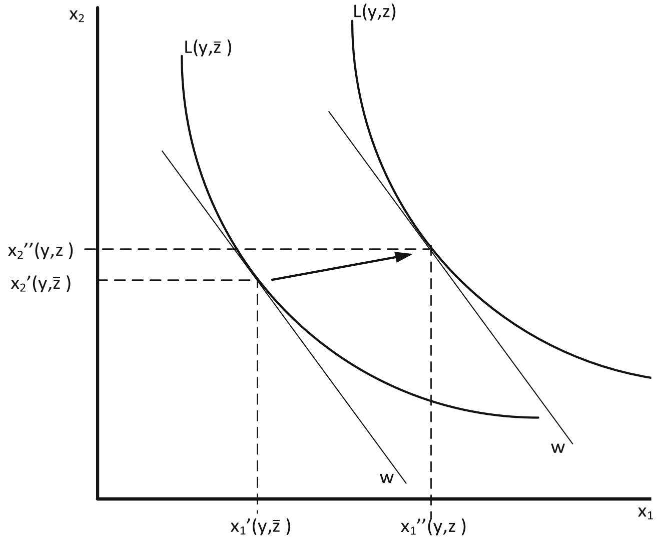


FIGURE 1. The Product Differentiation Index

from those with no product differentiation (equal or closest to 1) to those with the highest observed differentiation (the lowest value).

We would also like to highlight two additional things. The first is that the product differentiation index in equation (5) can be calculated using both parametric and non-parametric techniques, although the results of Section 3.5 are based on a parametric specification. The second is that the product differentiation index in equation (5) cannot be interpreted as a measure of product quality. Consider a market where all the firms provide a vector of identical product characteristics $z$, with the highest possible quality. In this market, the core characteristics $\overline{z}$ should be equal to those of $z$, and the product differentiation index in equation (5) takes the value 1, for all the firms, showing no product differentiation. This situation defines only one isoquant. We have $L(y, z) = L(y, \overline{z})$ in Figure 1.

## Properties of the Product Differentiation Index

The study of the axioms of tests of the product differentiation index in equation (5) requires the assumptions $w_{hkt} > 0, y_{hkt} > 0$, and $z_{hkt} > 0$ by firm $h$, in market $k$, and period $t$. The index is defined as the coefficient of two cost functions, and thus it should inherit the continuity and positivity axioms from a cost function, to which we can add the identity axiom mentioned earlier, i.e., when $\overline{z} = z$, we have $Q\left(w_{hkt}, y_{hkt}, z_{hkt}, z_{hkt}\right) = 1$. The product differentiation index in equation (5) has an additional list of interesting axioms or tests, which are:

i. Transitivity of core characteristics for fixed $y$ and $w$. It can be expressed as $Q\left(w_{hkt}, y_{hkt}, z_{hkt}, \overline{z}_{kt}\right) \times Q\left(w_{hkt}, y_{hkt}, \overline{z}_{kt}, \overline{\overline{z}}_{kt}\right) = Q\left(w_{hkt}, y_{hkt}, z_{hkt}, \overline{\overline{z}}_{kt}\right)$, if the observed characteristics are $z$ and the core characteristics $\overline{z}$, and the core characteristics change from $\overline{z}$ to $\overline{\overline{z}}$, then the value of $Q\left(w_{hkt}, y_{hkt}, z_{hkt}, \overline{\overline{z}}_{kt}\right)$ is the product of the product differentiation index from $z$ to $\overline{z}$ and the product differentiation index from $\overline{z}$ to $\overline{\overline{z}}$;

ii. Characteristics reversal test. Consider $\overline{\overline{z}} = z$ in (i), then $Q\left(w_{hkt}, y_{hkt}, z_{hkt}, \overline{z}_{kt}\right) \times Q\left(w_{hkt}, y_{hkt}, \overline{z}_{kt}, z_{hkt}\right) = 1$;

iii. *Circular test*. Eichhorn and Voeller (1976, p. 13) show that the circular property is derived from the properties of transitivity in (i) and reversal in (ii). It can be stated as $Q\left(w_{hkt}, y_{hkt}, z_{hkt}, \overline{z}_{kt}\right) \times Q\left(w_{hkt}, y_{hkt}, \overline{z}_{kt}, \overline{\overline{z}}_{kt}\right) \times, \dots, \times Q\left(w_{hkt}, y_{hkt}, \overline{\overline{\overline{z}}}_{kt}, z_{hkt}\right) = 1$;

iv. (Weak) monotonicity in core characteristics $\overline{z}$, where $Q\left(w_{hkt}, y_{hkt}, z_{hkt}, \overline{z}^{1}_{kt}\right) \leq Q\left(w_{hkt}, y_{hkt}, z_{hkt}, \overline{z}^{2}_{kt}\right)$, for $0 < \overline{z}^{1} \leq \overline{z}^{2}$;

v. (Weak) monotonicity in observed characteristics $z$, where $Q\left(w_{hkt}, y_{hkt}, z^{1}_{hkt}, \overline{z}_{kt}\right) \geq Q\left(w_{hkt}, y_{hkt}, z^{2}_{hkt}, \overline{z}_{kt}\right)$, for $0 < z^{1} \leq z^{2}$;

vi. Invariant to changes in the ordering of $w$, $y$, $z$, and $\overline{z}$. It implies $Q\left(w^{*}_{hkt}, y^{*}_{hkt}, z^{*}_{hkt}, \overline{z}^{*}_{kt}\right) = Q\left(w_{hkt}, y_{hkt}, z_{hkt}, \overline{z}_{kt}\right)$, where "*"denotes a permutation of the components of each vector;

vii. *Invariance to changes in the units of measurement* (commensurability axiom). That is, $Q\left(w_{hkt}, y_{hkt}, z_{hkt}, \overline{z}_{kt}\right)$ does not change if the units of measurement for each input quantity are changed. This axiom is due to the fact that the definition of the product differentiation index in equation (5) is based on cost, which is equal to the product of prices and quantities.

When the underlying technology of the cost function has constant returns to scale, c($w$, $y$, $z$) is non-decreasing and homogeneous of degree +1 in $w$, $y$, and $z$. In this situation, the product differentiation index in equation (5) has these additional axioms

viii. *Linear homogeneity in core characteristics* $\overline{z}$. If all core characteristics change by $\lambda > 0$, then the value of the product differentiation index in equation (5) is changed by $\lambda$, that is $Q\left(w_{hkt}, y_{hkt}, z_{hkt}, \lambda\overline{z}_{kt}\right) = \lambda Q\left(w_{hkt}, y_{hkt}, z_{hkt}, \overline{z}_{kt}\right)$;

ix. *Homogeneity of degree 0 in characteristics*. If all observed and core characteristics change by $\lambda > 0$, then the value of the product differentiation index in equation (5) does not change, that is $Q\left(w_{hkt}, y_{hkt}, \lambda z_{hkt}, \lambda\overline{z}_{kt}\right) = Q\left(w_{hkt}, y_{hkt}, z_{hkt}, \overline{z}_{kt}\right)$;

x. Mean value axiom.

$$min\left(\frac{x'_1(y,\overline{z})}{x''_1(y,z)} \dots, \frac{x'_M(y,\overline{z})}{x''_M(y,z)}\right) \leq Q\left(w_{hkt}, y_{hkt}, z_{hkt}, \overline{z}_{kt}\right) \leq max\left(\frac{x'_1(y,\overline{z})}{x''_1(y,z)} \dots, \frac{x'_M(y,\overline{z})}{x''_M(y,z)}\right)$$

Eichhorn and Voeller (1976, pp. 10, 27) demonstrate that every function that satisfies the axioms of identity, monotonicity in (iv) and (v), commensurability in (vii), linear homogeneity in (viii), and homogeneity of degree 0 in (ix) also satisfies the mean value axiom.

xi. *Proportionality in characteristics*. If $\overline{z} = z$ and all observed characteristics change by $\lambda > 0$, then the value of the product differentiation index in equation (5) is equal to $\lambda$, that is $Q\left(w_{hkt}, y_{hkt}, z_{hkt}, \lambda z_{hkt}\right) = \lambda$.

xii. *Homogeneity of degree minus one in observed characteristics* $z$. If all observed characteristics change by $\lambda > 0$, then the value of the product differentiation index in equation (5) is changed by $\lambda^{-1}$, that is $Q\left(w_{hkt}, y_{hkt}, \lambda z_{hkt}, \bar{z}_{kt}\right) = \lambda^{-1} Q\left(w_{hkt}, y_{hkt}, z_{hkt}, \bar{z}_{kt}\right)$.

See also the Appendix S1 for additional details.

### 3.3. *Contribution to the Hedonic Price Function*

The most common method of developing a price index is the "matched model" methodology, wherein the price for a product is collected at two periods of time, and the index computed from the two prices. This method can break down (Triplett, 2004) when the exact product is not available in both periods, or if there is a change in the product that makes the match inexact. Furthermore, there can be changes in the way a transaction occurs, beyond just the characteristics of the good. This is highlighted in the statement (Triplett, 2004, p. 15), "Even if the commodity is homogeneous (which is itself so infrequent empirically that it is of little practical importance), transactions are not homogeneous. It is transactions that matter in a price index."

A hedonic function can be used to alleviate this issue when the exact product is not available, or when there have been changes in the good, in effect creating a quality adjustment that helps separate true price difference from changes in the product or service. However, in an imperfect market, both Feenstra (1995) and Pakes (2003) have shown that hedonic coefficients can be biased where a price mark-up over costs exists. One method of controlling for this bias is by including a control for market power. Examples of this method can be found in Cotteleer *et al.* (2008), who use the number of potential buyers and sellers as a control, and in Harding *et al.* (2003), who use characteristics of the buyer as a proxy for bargaining power.

In markets where there are differentiated buyers, some individuals may value a particular bundle of characteristics more than others do. When preferences vary, it is not enough to consider only the level of market concentration. The level of available product differentiation must also be considered. By including our product differentiation index in the hedonic function, we contribute to the hedonic function by controlling for price variance arising from differing consumer preferences. In concept, the hedonic function would have three primary sections, characteristics of the product, a measure of product differentiation, and a measure of market power.

Beginning with the simple example of estimation presented in equation (1), we extend to include market power (*Market*) and the product differentiation index $Q\left(w_{hkt}, y_{hkt}, z_{hkt}, \bar{z}_{kt}\right)$ defined in Section 3.2. Applied to the hedonic function, we use the inverse of the product differentiation index for ease of interpretation and note it hereafter as $Q_{hkt}^{-1}$. This index can accommodate multiple outputs; however, in the context of a hedonic function, we use a single output. The definition of output $y$ in the index will vary depending on the product and market being examined. Adopting the standard double log functional form, a general version of the function is

$$(6) \qquad \ln p_{hkt} = \beta_0 + \sum_{i}^{n} \beta_{zi} \ln z_{ihkt} + \beta_Q \ln Q_{hkt}^{-1} + \sum_{l=1}^{n} \beta_{ml} \ln Market_{lkt} + \beta_t t + \varepsilon_{hkt}.$$

The price of firm *h's* product in market $k$ at time $t$ is estimated as a function of characteristics of the product $z$, market power as *Market*, and product differentiation as $Q_{hkt}^{-1}$.

### 3.4. *Hedonic Mark-Up Approach*

The hedonic function developed in equation (6) can be used in understanding price indexes. By controlling for product differentiation in a market, we can get a better estimate of how consumers value changes in characteristics. From the perspective of the firm, the result of equation (6) can be used to understand price levels and how characteristics and product differentiation affect them. There is however a factor lacking in equation (6) that is important: Providing a higher level of characteristics and differentiating from the market carry a cost. With the assumption of a profit motive, a firm would choose to differentiate only when the price increase received is at least equal to the cost of the differentiation. To examine this, we introduce a new version of the standard hedonic function, the hedonic mark-up function.

The first step in generating the hedonic mark-up function is defining a new dependent variable. Instead of price $p$, we now measure the ratio of the price to the cost of the core product. Given the log nature of the hedonic regression and the cost function, we find this as

$$(7) \qquad M_{hkt} = \ln\left(\frac{p_{hkt}}{c\left(w_{hkt}, y_{hkt}, \overline{z}_{kt}\right)}\right).$$

Conceptually, this is equivalent to moving $mc(\overline{z})$ to the left-hand side of equation (3).

Since expression (7) measures mark-up over the cost of the core product, the level of characteristics beyond the core is now the more relevant measure. To proxy for this difference, we use the product differentiation index that measures how much a firm's product differs from others in the market based on the cost of producing characteristics. Retaining the *Market* variables, the estimation of the hedonic mark-up function for airline $h$ on route $k$ at time $t$ is

$$(8) \qquad M_{hkt} = \beta_0 + \beta_Q \ln Q_{hkt}^{-1} + \sum_{l=1}^{l} \beta_{ml} \ln Market_{lkt} + \beta_t t + \varepsilon_{hkt}.$$

As in equation (7), we would expect increases in the *Market* variables to lead to reductions in mark-up. The index $Q_{hkt}^{-1}$ captures the effect of differentiation beyond the core product. We might expect coefficient $\hat{\beta}_Q$ to be positive, because when a firm differentiates it can create a mark-up over the cost of the core product. Coefficient estimate $\hat{\beta}_Q$ in equation (8) provides more information than just the direction and magnitude of change in mark-up. Holding all else constant, we can find the effect of a change in $Q_{hkt}^{-1}$ as $M_{hkt} = \hat{\beta}_Q \ln Q_{hkt}^{-1}$. Inserting the definition of $M_{hkt}$ from equation

(7) and the definition of $Q\left(w_{hkt}, y_{hkt}, z_{hkt}, \overline{z}_{kt}\right)$ from equation (5), we can rewrite $M_{hkt} = \widehat{\beta}_Q \ln Q_{hkt}^{-1}$ as

$$(9) \qquad \ln\left(\frac{p_{hkt}}{c\left(w_{hkt}, y_{hkt}, \overline{z}_{kt}\right)}\right) = \widehat{\beta}_Q \ \ln\left(\frac{c\left(w_{hkt}, y_{hkt}, z_{hkt}\right)}{c\left(w_{hkt}, y_{hkt}, \overline{z}_{kt}\right)}\right).$$

From equation (9), we see that higher levels of characteristics $z_{hkt}$ lead to a higher $c\left(w_{hkt}, y_{hkt}, z_{hkt}\right)$ and therefore a higher ratio $c\left(w_{hkt}, y_{hkt}, z_{hkt}\right)/c\left(w_{hkt}, y_{hkt}, \overline{z}_{kt}\right)$. The percentage change in that ratio, multiplied by $\widehat{\beta}_Q$, equals the expected percentage change increase in price over core cost $p_{hjt}/c\left(w_{hkt}, y_{hkt}, \overline{z}_{kt}\right)$. This percentage change does not directly tell us how a cost increase is reflected in a price increase. From equation (9), we can note that the denominators on the left- and right-hand sides are equal, meaning that the percentage change in $c\left(w_{hkt}, y_{hkt}, z_{hkt}\right)$ is directly reflected as a percentage change in $p_{hkt}$. When $\widehat{\beta}_Q > c\left(w_{hkt}, y_{hkt}, z_{hkt}\right)/p_{hkt}$ that change results in a price increase that exceeds the cost increase.

### 3.5. *Hedonic Mark-Up Approach by Characteristic*

Calculating product differentiation index $Q_{hkt}^{-1}$ requires estimation of the two underlying cost functions. The approach outlined in Section 3.4 can be used with any estimation method[5] and is independent of the functional form of the underlying cost functions. In this paper, we specify a stochastic parametric cost function and a translog functional form[6] because they have characteristics that allow us to extract cost elasticity estimates on product characteristics, and to extend the approach developed in Section 3.4. Estimating the translog, we add a time variable ($t$) to output quantity, input prices, and characteristics to capture cost changes due to technical progress or regress over time.[7] Product characteristics $z$ are treated as cost shifters and are not included in interaction effects. With $C_{ht}$ as the total cost for firm $h$ at time $t$, the translog cost function to be estimated is

(10)

$$\ln C_{ht} = \sum_{k=1}^{m} \alpha_k \ln w_{kht} + \tfrac{1}{2} \sum_{k=1}^{m} \sum_{j=1}^{m} \alpha_{kj} \ln w_{kht} \ \ln w_{jht} + \alpha_y \ln y_{ht} + \tfrac{1}{2}\, \alpha_{yy} \left(\ln y_{ht}\right)^2 + \sum_{k=1}^{m} \alpha_{ky} \ln w_{kht} \ln y_{ht} + \sum_{i=1}^{n} \alpha_i \ln z_{iht} + \alpha_t t + \varepsilon_{ht}; h = 1, .., l, t = 1, .., T.$$

[5]As examples, stochastic parametric frontiers (Gillen *et al.*, 1990; Ahn and Sickles, 2000; Assaf, 2009; Johnston and Ozment, 2013), deterministic non-parametric frontiers (Barros *et al.*, 2013; Arjomandi and Seufert, 2014), and stochastic non-parametric frontier (Tsionas, 2003).

[6]The translog form, proposed by Christensen *et al.* (1973), has been widely used in airline cost functions since Caves *et al.* (1984) study on the economies of density and scale.

[7]We considered including a variable to indicate the airline; however, in most cases, product characteristics and the airline are very closely correlated.

To ensure the estimated cost function is homogeneous of degree one in input prices, and to ensure symmetric cross effects, the following typical restrictions are imposed.

$$\sum_{k=1}^{m} \alpha_k = 1, \sum_{k=1}^{m} \alpha_{kj} = \sum_{k=1}^{m} \alpha_{jk} = \sum_{k=1}^{m} \alpha_{ky} = 0$$

We also restrict linear homogeneity in input price by normalizing cost and price by one of the input prices. Following Shephard's lemma, which implies that efficient input shares be equal to the partial derivative of the cost function with respect to input prices, we obtain the share of input $k$ in total cost ($S_k$) as

$$S_k = \frac{\partial \ln C}{\partial \ln w_k} = \alpha_k + \sum_{k=1}^{m} \alpha_{kj} \ln w_{jht} + \alpha_{ky} \ln y_{ht}.$$

We use the cost elasticities found in equation (10) to predict cost at the observed level of the characteristic $z_{iht}$ and at the level of the core product in the market at the firm output level.

The effect captured in equation (8) for index $Q_{hkt}^{-1}$ is the effect of all product characteristics combined. To move to an estimation for individual characteristics, we can take advantage of the log nature of the translog cost function. The value of $Q_{hkt}^{-1}$ can be calculated from the cost function, expression (10), after parameters $\widehat{\alpha} = \left\{\widehat{\alpha}_k, \widehat{\alpha}_{kj}, \widehat{\alpha}_y, \widehat{\alpha}_{yy}, \widehat{\alpha}_{ky}, \widehat{\alpha}_i, \widehat{\alpha}_t\right\}$ are estimated, and noting the definition of $Q_{hkt}^{-1}$ we have

$$\text{(11)} \quad \left[Q\left(w_{hkt}, y_{hkt}, z_{hkt}, \overline{z}_{kt}\right)\right]^{-1} = \exp\left[\ln c\left(w_{hkt}, y_{hkt}, z_{hkt}\right) - \ln c\left(w_{hkt}, y_{hkt}, \overline{z}_{kt}\right)\right]$$
$$= \exp\left[\sum_{i=1}^{n} \widehat{\alpha}_i \ln z_{ihkt} - \sum_{i=1}^{n} \widehat{\alpha}_i \ln \overline{z}_{ikt}\right]$$
$$= \exp\left[\sum_{i=1}^{n} \widehat{\alpha}_i \ln\left(\frac{z_{ihkt}}{\overline{z}_{ikt}}\right)\right] \quad h = 1, .., l, t = 1, .., T.$$

Replacing $\ln\left[Q\left(w_{hkt}, y_{hkt}, z_{hkt}, \overline{z}_{kt}\right)\right]^{-1}$ in estimation (8) with the definition from equation (11) creates an estimation that measures the effect of each characteristic separately:

$$\text{(12)} \qquad M_{hkt} = \beta_0 + \sum_{i=1}^{n} \beta_i \left[\widehat{\alpha}_i \ln\left(\frac{z_{ihkt}}{\overline{z}_{ikt}}\right)\right] + \sum_{l=1}^{3} \beta_{ml} \ln Market_{lkt} + \beta_t t + \varepsilon_{hkt}.$$

We can interpret $\widehat{\beta}_i$ on individual characteristics the same way as we interpret $\widehat{\beta}_Q$. Rewriting equation (9) in terms of individual characteristics, we have

$$\text{(13)} \qquad \ln\left(\frac{p_{hkt}}{c\left(w_{hkt}, y_{hkt}, \overline{z}_{kt}\right)}\right) = \sum_{i=1}^{n} \beta_i \left[\widehat{\alpha}_i \ln\left(\frac{z_{ihkt}}{\overline{z}_{ikt}}\right)\right].$$

Coefficient $\widehat{\beta}_i$ from equation (13) relates price over core cost $p_{hkt}/c\left(w_{hkt}, y_{hkt}, \overline{z}_{kt}\right)$ to the portion of the difference between total minimum cost and minimum core cost due to characteristic $z_i$. If provided characteristic $z_i$ is higher than core characteristic $\overline{z}_i$, ratio $z_{ihkt}/\overline{z}_{ikt}$ is greater and we expect a larger mark-up over core cost $p_{hkt}/c\left(w_{hkt}, y_{hkt}, \overline{z}_{kt}\right)$. As in equation (9), the denominators on the left-hand and right-hand sides are effectively equal, meaning that higher levels of $z_i$ are directly reflected as higher levels of $p_{hkt}$. When $\widehat{\beta}_i > c\left(w_{hkt}, y_{hkt}, z_{hkt}\right)/p_{hkt}$, we expect that an increase in $z_i$ would result in a price increase that exceeds the cost increase.

## 4. The Data

All data for this study are from the Bureau of Transportation Statistics (BTS), an independent statistical agency within the U.S. Department of Transportation (DOT). We specify the unit of analysis as a route, which is defined as a nonstop, one-way, or round-trip flight between two airport pairs[8] for a single carrier. Routes are considered bidirectional, meaning a flight from Atlanta (ATL) to Denver (DEN) or the return trip, DEN to ATL, constitutes a single route. Nonstop flights are those that have no intermediate stop between the airport pairs. Descriptive characteristics of the data, including the mean, standard deviation, and the 25th and 75th percentiles, for the average route and quarter are outlined in Table 1.

Fare data are in the first row of Table 1 and come from the Airline Origin and Destination Survey (DB1B). Fares are in 2002 dollars and prices are for a one-way ticket, or half the price of a round-trip ticket. Defining a market as the combination of all routes, we see that many markets are highly concentrated, with an average of 2.17 competitors per market. However, the larger number of airport competitors, 8.93, should serve to provide some level of competition and hold down prices.

We include four product characteristics for vector $z$: (1) distance; (2) performance; (3) convenience; (4) service level. Distance is simply the number of miles between the two end points on a route. We define performance as the percentage of flights that arrive within 15 min of the scheduled arrival time.[9] Convenience is measured as the number of daily departures, or flight frequency. For a similar volume level, having more frequent flights adds to the cost by reducing load factors and raising landing fees.[10] Service level, measured as the percentage of first- or business-class tickets, increases the cost by increasing the use of resources (food, materials, labor) and through increasing seating space used in the cabin.

In Table 1 we present statistics for observed product characteristics and for the level of the core product. Recall that the core product is the lowest level of characteristics available in a market. On the average route, the percentage of first-class

[8]Several cities have airports that are close enough together to be considered substitutes. In those cases, we have combined the airports. The following close-by airports are combined in the results: DFW (Dallas–Fort Worth) and DAL (Love Field); LGA (La Guardia), EWR (Newark), and JFK (J. F. Kennedy); AZA (Phoenix-Mesa Gateway) and PHX (Phoenix Sky Harbor); TPA (Tampa) and PIE (St. Petersburg Clearwater); DCA (Reagan) and IAD (Washington Dulles); ORD (O'Hare) and (MDW) Midway.

[9]Gayle and Yimga (2018) found that travelers are willing to pay \$1.56 per minute to avoid late arrival. A study by Zou and Hansen (2012) found that improving performance can increases cost.

[10]Flight frequency is a characteristic valued by consumers; both Borenstein (1989) and Douglas and Miller (1974) note it as a product differentiator that increases brand value.

TABLE 1
Quarterly Descriptive Statistics by Carrier and Route

| Statistic | Mean | St. Dev. | Pctl (25) | Pctl (75) |
|---|---|---|---|---|
| *Fares and marker* | | | | |
| Mean fare ($) | 174.72 | 74.99 | 121.71 | 214.9 |
| RPM (# 000s) | 63, 105.22 | 85, 147.32 | 15, 725.43 | 73, 575.28 |
| Distance (miles) | 1, 043.71 | 719.17 | 490.00 | 1, 440.00 |
| Route passengers (# 000s) | 65, 063.00 | 69, 916.00 | 21, 170.00 | 82, 810.00 |
| Market passengers (# 000s) | 136, 346.00 | 167, 471.00 | 30, 475.00 | 176, 333.00 |
| Airport competitors (#) | 8.93 | 2.54 | 7.50 | 10.50 |
| Route competitors (#) | 2.17 | 0.14 | 1.00 | 3.00 |
| *Product characteristic* | | | | |
| Business/first class (%) | 0.04 | 0.06 | 0.00 | 0.07 |
| Flight frequency (#) | 6.49 | 6.42 | 2.02 | 8.30 |
| On-time arrival (%) | 0.81 | 0.09 | 0.76 | 0.87 |
| *Core characteristic* | | | | |
| Business/first class (%) | 0.00 | 0.00 | 0.00 | 0.00 |
| Flight frequency (#) | 4.78 | 4.67 | 1.97 | 6.03 |
| On-time arrival (%) | 0.78 | 0.10 | 0.73 | 0.85 |
| *Input prices* | | | | |
| Labor price ($) | 23, 339.34 | 5, 272.13 | 19, 801.80 | 26, 885.70 |
| Fuel price ($) | 2.04 | 0.84 | 1.42 | 2.87 |
| Capital price ($) | 12, 767.02 | 3, 750.19 | 10, 200.65 | 13, 838.20 |
| Other materials price ($) | 147.90 | 15.08 | 137.40 | 160.67 |

travelers is 4 percent, while the average core level is 0.00 percent; this reflects the fact that several carriers (e.g., Southwest, JetBlue, and Spirit) do not market first- or business-class tickets. Regarding flight frequency, we see that the average carrier provided almost two more flights per day than carriers providing only the core level of service. Finally, for on-time arrival performance, the average for all carriers is 81 percent, while the core level is 78 percent. All three characteristics exhibit a high degree of variation.

To measure output ($y$) we use revenue passenger miles (RPM), instead of available seat miles (ASM) or revenue ton miles (RTM), because RPM matches up well with our secondary use of cost in a hedonic regression and the characteristic of distance. Table 1 also presents descriptive statistics for input prices $(w)$. We use a standard KLEM model for inputs and include capital, labor, energy, and materials.[11] Fuel price is calculated as total expense over total gallons and labor price as total expense over the number of full-time equivalents per quarter. The price of capital has been measured in a number of ways and is dependent on how capital quantity is measured. We are following a process similar to that of Färe *et al.* (2007), who develop capital price based on capital expense and the total number of seats available. Capital expense comes from actual leasing rates and capital depreciation, while the number of seats available is calculated from the number of planes in service and the seat configuration. As a proxy for all other materials, we use the producer price index (PPI) collected by the BLS. This index varies by quarter, but not by airline.

[11]These data are provided by the BTS in various schedules and allow us to disaggregate system-wide inputs and specify only the inputs used on domestic flights.

TABLE 2
Mean Effective Coefficient Across Observations

| Variable | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|
| Capital | 0.07 | 0.03 | 0.02 | 0.18 |
| Labor | 0.30 | 0.06 | 0.12 | 0.43 |
| Fuel | 0.23 | 0.07 | 0.03 | 0.40 |
| Other materials | 0.40 | na | na | na |
| Output | 1.00 | 0.17 | 0.65 | 1.50 |
| First/business-class share | 0.01 | 0.00 | 0.01 | 0.01 |
| Percentage on-time arrival | 0.06 | 0.00 | 0.06 | 0.06 |
| Flight frequency | 0.38 | 0.00 | 0.38 | 0.38 |

## 5. Results

### 5.1. *Cost Function*

Since our focus is on the product differentiation index and its contribution to a hedonic function, discussion of the cost function estimation is limited to the main results. More detail on estimation technique and results are available on request. For estimation, all values are in log form, and dollar figures have been converted to 2002 dollars using the relevant producer price index (PPI). Average estimation coefficients are presented in Table 2. We find the input price shares for capital, labor, fuel, and other materials to be 7 percent, 30 percent, 23 percent, and 40 percent, respectively, results that are in line with those of other cost studies cited in this paper. We use the average route for each carrier for each quarter between 2002 and 2016 to obtain coefficient estimates. The carriers included are those that carry at least 1 percent of the domestic US passenger volume and include 16 different airlines over the period.[12] We find cost per passenger mile, 14.7 cents on average, to be in line with industry reported values. Applying the elasticity coefficients at the observed route-level data, we find a mean cost of 11.5 cents per mile on the core product and 15.0 cents per mile based on observed characteristics. We also find the average value of the index of product differentiation $Q_{hkt}^{-1}$ to be 0.86, indicating that on average, the minimum cost of the core product is 14 percent less than the minimum cost of observed characteristics.

### 5.2. *Estimation Results—Hedonic Price Function*

In Table 3 we present the results of two estimations of equation (6), the first as a cross-section estimation and the second as a panel estimation that includes fixed effects for market $k$. The fixed-effects estimation accounts for market characteristics that are not included in product characteristics or the *Market* variables, market characteristics such as the percentage of business versus leisure travelers, available substitutes, and size of market catchment. In the fixed-effects estimation, we exclude

[12]Carriers are American Airlines, Alaska Airlines, JetBlue, Continental, Delta, Frontier, AirTran Airways, Hawaiian Airlines, America West, Spirit, Northwest, ATA Airlines, United Airlines, US Airways, Virgin America, and Southwest.

TABLE 3
Dependent Variable: Mean Route Fare

| Estimation Formula | Cross Section $p_{hkt}$ (6) | Fixed Effects $p_{hkt}$ (6) |
|---|---|---|
| Distance ln ($z$) | 0.331*** | |
| | (0.001) | |
| First/business-class share ln ($z$) | 0.033*** | 0.000 |
| | (0.000) | (0.001) |
| Percentage on-time arrival ln ($z$) | −0.108*** | 0.012* |
| | (0.009) | (0.006) |
| Flight frequency ln ($z$) | 0.192*** | 0.067*** |
| | (0.002) | (0.002) |
| $[Q(w_{hkt}, y_{hkt}, z_{hkt}, \bar{z}_{kt})]^{-1}$ | −0.090*** | −0.025*** |
| | (0.003) | (0.002) |
| Market passengers (*Marker*) | −0.188*** | −0.138*** |
| | (0.002) | (0.002) |
| Airport competitors (*Marker*) | 0.087*** | −0.023*** |
| | (0.004) | (0.006) |
| Route competitors (*Marker*) | 0.003 | −0.100 |
| | (0.002) | (0.002) |
| Observations | 103,980 | 103,980 |
| Adjusted $R^2$ | 0.57 | 0.03 |

Notes: All regressions include qtr and merger dummies and all variables logged. Std error in parentheses.*Significant at the 10 percent level.***Significant at the 1 percent level.

the characteristic of *Distance* because by definition, it is perfectly correlated with the market.

Beginning with our contribution to the hedonic price estimation, we find that the coefficient on $Q^{-1}_{hkt}$ is negative and highly significant in both the cross-section and panel estimates. The economic intuition behind this result is that when there are carriers in the market providing a low level of characteristics, the overall market price is moderated. This effect has been documented by others in the more specific example of Southwest. Vowles (2001) documents this "Southwest effect" on four specific routes, showing an increase in traffic and decrease in airfares. Our work in this paper allows this to be applied as a more general concept. Finding that the coefficient on $Q^{-1}_{hkt}$ is statistically significant, we also show that it can be useful as a control where a hedonic is being used to generate a quality adjusted prices index.

In the cross-section estimate of Table 3, we see that *Distance* is positive and highly significant, indicating that the fare price increases by roughly 0.33 percent for each 1 percent increase in miles traveled. These results are similar to those in the works previously cited (Borenstein, 1989; Morrison and Winston, 1995). Continuing with Table 3 equation (6), we notice differences between the cross-section panel estimates. *First-Class/Business-Class Share* and *Flight Frequency* are positively correlated with price in both models, although the effect weakens in economic and statistical significance in the panel estimation. The reversal of sign on *On-Time Arrival* is likely due to differences between markets, with leisure markets placing a lower value on arrival time.

Examining our variables measuring market power, we see that the number of *market passengers* is the most consistent measure, with an increase in the market size leading to a reduction in price under both estimation forms. The significant positive sign on *airport competitors* was not expected but might be explained by hub and spoke networks. Hub airports are likely to have more airlines present, and one of the findings by Borenstein (1989) was that at hubs the dominant airline can charge higher prices than it does throughout its system. The reversal of sign on this variable when the market is controlled for strengthens that argument. The number of *route competitors* was not significant to price formation after controlling for all other factors. We also examined the effects of a merger or acquisition on price and find that in the four quarters following a merger of two of the airlines on a route, the mean fare price increases. We do not report on these results, but all models include this variable as a control.

### 5.3. *Estimation Results—Hedonic Mark-Up Function*

In Table 4, we present the results of estimations for equations (8) and (12), again presenting both cross-section and fixed-effects results. We do not include the characteristic of *Distance* in Table 4, because it is a core characteristic that does not differ by carrier in a market.

In mark-up estimations (8) and (12), we see positive and significant results on the product differentiation index $Q_{hkt}^{-1}$ with coefficient values of 0.969 in the cross-section estimate and 0.897 under the fixed-effects estimate. These results indicate that carriers can increase mark-up over the core cost by differentiating from the core. Recalling that a $Q_{hkt}^{-1}$ value of 1 indicates homogeneity with other carriers

TABLE 4
Dependent Variable: Markup over the Core Cost

| | Cross Section | | Fixed Effects | |
|---|---|---|---|---|
| Estimation Formula | $M_{hkt}$ (8) | $M_{hkt}$ (12) | $M_{hkt}$ (8) | $M_{hkt}$ (12) |
| $[Q(w_{hkt}, y_{hkt}, z_{hkt}, \bar{z}_{kt})]^{-1}$ | 0.969*** | | 0.897*** | |
| | (0.006) | | (0.004) | |
| First/business-class share $\hat{\alpha}\ \ln(z/\bar{z})$ | | 2.918*** | | 1.938*** |
| | | (0.044) | | (0.134) |
| Percentage on-time arrival $\hat{\alpha}\ \ln(z/\bar{z})$ | | −5.098*** | | 2.129*** |
| | | (0.383) | | (0.179) |
| Flight frequency $\hat{\alpha}\ \ln(z/\bar{z})$ | | 0.944*** | | 0.890*** |
| | | (0.006) | | (0.004) |
| Market passengers (*Market*) | −0.239*** | −0.247*** | −0.662*** | −0.663*** |
| | (0.003) | (0.003) | (0.003) | (0.003) |
| Airport competitors (*Market*) | −0.690*** | −0.714*** | 0.034*** | 0.035*** |
| | (0.009) | (0.009) | (0.010) | (0.010) |
| Route competitors (*Marker*) | 0.215*** | 0.237*** | 0.132*** | 0.128*** |
| | (0.005) | (0.005) | (0.004) | (0.004) |
| Observations | 103,980 | 103,980 | 103,980 | 103,980 |
| Adjusted $R^2$ | 0.30 | 0.31 | 0.44 | 0.44 |

*Notes*: All regressions include qtr and merger dummies and all variables logged. Std error in parentheses.***Significant at the 1 percent level.

on the route, an increase in $Q_{hkt}^{-1}$ signals an increase in product differentiation.[13] Similar positive and significant results are reflected in the individual characteristics of *First-Class/Business-Class Share* and *Flight Frequency* under both estimations. However, in the characteristic of *Percentage On-Time Arrival*, we see the same reversal of sign that occurred with the price hedonic. To examine this reversal further, we repeated estimation (12) for the 15 annual periods in our sample separately[14] and found that up until 2012, the coefficient on *Percentage On-Time Arrival* was positive, but negative afterward. An explanation for this change could be the growth of ultra-low-cost carriers (ULCC), and the conversion of Frontier Airlines from a low-cost carrier (LCC) to a ULCC. These carriers placed a lower priority on arriving on time, reducing the cost of the core product in many markets.

Looking at market power, we see that the coefficient on *market passengers* is negative and highly significant, indicating that an increase in the number of passengers leads to a reduction in the ability to mark-up price over the core cost. In contrast to the price hedonic, the coefficient on *route competitors* is positive and significant in both estimations. This result indicates that having more competitors on the route leads to a greater mark-up in price over the core cost. This result may be partly due to the construction of the measure and estimation, because having more competitors allows for greater differentiation. The coefficient on the *airport competitors* has the expected negative sign in the cross-section estimation, but this turns positive when we control for the market.

The mark-up hedonic in Table 4 explains the direction and magnitude of the relationship between the ratio of price over core cost $p_{hkt}/c\left(w_{hkt}, y_{hkt}, \overline{z}_{kt}\right)$ and the ratio of the cost of characteristics provided over core cost $c\left(w_{hkt}, y_{hkt}, z_{hkt}\right)/c\left(w_{hkt}, y_{hkt}, \overline{z}_{kt}\right)$. It does not directly measure the relationship between any changes in price and cost that come from an increase in differentiation. To understand this relationship, we recall that the ability to increase price over the cost increase is based on the condition $\widehat{\beta}_Q > c\left(w_{hkt}, y_{hkt}, z_{hkt}\right)/p_{hkt}$ for $Q_{hkt}^{-1}$ in equation (8) and $\widehat{\beta}_i > c\left(w_{hkt}, y_{hkt}, z_{hkt}\right)/p_{hkt}$ in equation (12) for individual characteristics. Noting that the mean weighted value of $c\left(w_{hkt}, y_{hkt}, z_{hkt}\right)/p_{hkt}$ for our sample is 0.77, and comparing the sample value to the $\widehat{\beta}_Q$ estimates in Table 4, we see that the expected price increase is greater than the cost increase. This result implies that airlines can earn profit by differentiating from the core product.

Based on the unreported annual estimates, the individual characteristic of *First-Class/Business-Class Share* most consistently earns a price increase greater than the cost increase. This result is intuitive because *First-Class/Business-Class Share* is the only characteristic included that allows for price discrimination. The second-most consistently profitable characteristic is *Flight Frequency*; here we find that price increases exceed the cost increases in approximately 80 percent of the observations. The least profitable characteristic is *Percentage On-Time Arrival*, which increases price over the cost increase in only 60 percent of the observations.

[13]An example of this can be constructed at our sample means. Mean fare price is \$174.72, mean cost $c\left(w_{hkt}, y_{hkt}, z_{hkt}\right)$ is \$110.63, and mean core cost $c\left(w_{hkt}, y_{hkt}, \overline{z}_{kt}\right)$ is \$92.50, and mean markup over base cost $M_{hkt}$ is equal to 1.89. At these values, a 10% increase in $c\left(w_{hkt}, y_{hkt}, z_{hkt}\right)$ is an increase of \$11.06 in cost. The 0.969 estimate of $\widehat{\beta}_Q$ indicates that the 10% increase in $Q_{hkt}^{-1}$ results in a 9.79% increase in mark-up $M_{hkt}$ to 2.07, which implies a fare price of \$191.76, which is a \$17.04 increase in price.

[14]Results are not published, but are available on request.

Analyzing profit effects by market characteristics, we observe an interesting effect related to the number of market passengers. We find that in larger markets, carriers are better able to convert characteristic-driven cost increases into price increases, in spite of the negative effect of market size on price. In almost all cases, as market size grows, profit earned from an increase in characteristics grows. We believe this increased profit is because carriers can capitalize on customers' willingness to pay when consumers can compare carriers. As stated before, having more competitors may reduce the basic price, but it increases the value of differentiation.

## 6. Conclusion

We have introduced an Allen-type index number based on cost functions that measures firm product differentiation and have shown how it can be used to extend a hedonic regression. This product differentiation index can be used in many types of empirical studies, such as those on market power, antitrust, or price indexes. With a similar rank of results as the HHI or Gini coefficient, our product differentiation can fill a similar role, and can fill a gap in the existing literature. In Section 3.2 we develop the properties of the index and how it can be used and interpreted, and in the Appendix S1 we provide additional details. We also mention that the index can be calculated using parametric, or non-parametric methods and through econometric or linear programming techniques.

The product differentiation index we develop in this paper compares observed product characteristics to a core product. We extend Kotler's (1967) model to operationalize this concept. Future extensions of this Allen-type index, a ratio of cost functions, could develop this method further. For example, this methodology could be used to look at how a core product has changed over time, or how a firm has changed its product characteristics to enter a market. It could also be used to understand the level of product heterogeneity in a market.

This paper also opens up a new line of research based on the definition of an innovative concept of mark-up over the cost of the core product. We generate a hedonic estimation that controls for both the cost of the core product and product differentiation, allowing us to understand the effect of differentiation on mark-up and price. Applying this concept to the US domestic airline market, we analyze the cost and price effects of product differentiation. We find that differentiating by providing a higher level of service, as measured by first−/business-class share and more frequent flights, an airline can consistently improve its mark-up over the core product. Furthermore, it can earn a premium over the cost increase due to the higher level of product characteristics. As a tool for firms and pricing professionals, this type of analysis can enable a firm to make a more informed and strategic business decision on additions, subtractions, or changes to the characteristics of its product(s).

One limitation of applying this method to the airline industry is the availability of cost data at the route level. For the US domestic market cost data are only at the system level, while fare data were gathered at the route level. We took many steps to ensure estimates developed with system data were applicable to the route level, but the possibility of discrepancies does exist.

Our empirical results provide specific strategic implications for airlines and other firms providing differentiated products. The first implication is that differentiation should not be a one-size-fits-all decision. The rivals and substitute products in a market affect the resulting price and mark-up. When possible, the firm should adjust the service offering by market based on rivals and available substitutes. For example, in markets with fewer competitors, firms should emphasize characteristics that allow for price discrimination between their own customers and should deemphasize those characteristics that differentiate them from their competitors. The opposite is true in competitive markets.

## References

Ahn, S. C. and R. C. Sickles, “Estimation of Long-Run Inefficiency Levels: A Dynamic Frontier Approach,” *Econometric Reviews*, 19(4), 461–92, 2000.

Allen, R. G. D., “The Economic Theory of Index Numbers,” *Economica*, 16(63), 197, 1949.

Arjomandi, A. and J. H. Seufert, “An Evaluation of the World’s Major Airlines’ Technical And Environmental Performance,” *Economic Modelling*, 41, 133–44, 2014.

Armknecht, P. A. and D. H. Ginsburg, “Improvements in Measuring Price Changes in Consumer Services: Past, Present, And Future,” in Z. Griliches (ed), *Output Measurement in the Service Sectors*, University of Chicago Press, Chicago, IL, 109–57, 1992.

Assaf, A., “Are U.S. Airlines Really in Crisis?” *Tourism Management*, 30(6), 916–21, 2009.

Barros, C. P., Q. B. Liang and N. Peypoch, “The Technical Efficiency of US Airlines,” *Transportation Research Part A: Policy and Practice*, 50, 139–48, 2013.

Bourassa, S. C., E. Cantoni and M. Hoesli, “Robust Hedonic Price Indexes,” *International Journal of Housing Markets and Analysis*, 9(1), 47–65, 2016.

Borenstein, S., “Hubs and High Fares: Dominance and Market Power in the U.S. Airline Industry,” *The RAND Journal of Economics*, 20(3), 344–65, 1989.

Caves, D. W., L. R. Christensen and M. W. Tretheway, “Economies of Density versus Economies of Scale: Why Trunk and Local Service Airline Costs Differ,” *The RAND Journal of Economics*, 15(4), 471, 1984.

Caves, R. E. and P. J. Williamson, “What is Product Differentiation, Really?,” *The Journal of Industrial Economics*, 34(2), 113, 1985.

Christensen, L. R., D. W. Jorgenson and L. J. Lau, “Transcendental Logarithmic Production Frontiers,” *The Review of Economics and Statistics*, 55(1), 28–45, 1973.

Chwelos, P. D., E. R. Berndt and I. M. Cockburn, “Faster, smaller, cheaper: an hedonic price analysis of PDAs,” *Applied Economics*, 40(22), 2839–2856, 2008. https://doi.org/10.1080/00036840600993924

Cotteleer, G., C. Gardebroek and J. Luijt, “Market Power in a GIS-Based Hedonic Price Model of Local Farmland Markets,” *Land Economics*, 84(4), 573–92, 2008.

Diewert, W. E., “Index Numbers,” in P. Macmillan (ed), *The New Palgrave Dictionary of Economics*, Palgrave Macmillan, London, 1–32, 2008. https://doi.org/10.1057/978-1-349-95121-5_940-2.

Douglas, G. W. and J. C. Miller, “American Economic Association Quality Competition, Industry Equilibrium, and Efficiency in the Price-Constrained,” *The American Economic Review*, 64(4), 657–69, 1974.

Duan, B. (. Y.)., C. Arcodia, E. Ma and A. Hsiao, “Understanding Wine Tourism in China Using an Integrated Product-Level and Experience Economy Framework,” *Asia Pacific Journal of Tourism Research*, 23(10), 949–60, 2018.

Eichhorn, W. and J. Voeller, “Theory of the Price Index. Fisher’s Test Approach and Generalizations,” *Lecture Notes in Economics and Mathematical Systems*, Springer-Verlag, Berlin, 1976.

Färe, R., S. Grosskopf and R. C. Sickles, “Productivity? of US Airlines After Deregulation,” *Journal of Transport Economics and Policy*, 41, 93–112, 2007.

Feenstra, R. C., “Exact Hedonic Price Indexes,” *The Review of Economics and Statistics*, 77(4), 634–53, 1995.

Gayle, P. G. and J. O. Yimga, “How much do consumers really value air travel on-time performance, and to what extent are airlines motivated to improve their on-time performance?,” *Economics of Transportation*, 14(June), 31–41, 2018.

Gillen, D. W., H. Oum and M. W. Tretheway, “Airline Cost Structure and Policy Implications: A Multi-Product Approach for Canadian Airlines,” *Journal of Transport Economics and Policy*, 24(1), 9–34, 1990.

Good, D. H., R. C. Sickles and J. C. Weiher, "A Hedonic Price Index for Airline Travel," *Review of Income and Wealth*, 54(3), 438–65, 2008.
Goyal, A., "Consumer Perception Towards the Purchase of Credit Cards," *Journal of Services Research*, 6, 179–90, 2006.
Grifell-Tatjé, E. and C. A. K. Lovell, *Productivity Accounting. The Economics of Business Performance*, Cambridge University Press, New York, 2015.
Hanson, T., "Efficiency and Productivity in the Operational Units of the Armed Forces: A Norwegian Example," *International Journal of Production Economics*, 179, 12–23, 2016.
Harding, J. P., S. S. Rosenthal and C. F. Sirmans, "Estimating Bargaining Power in the Market for Existing Homes," *Review of Economics and Statistics*, 85(1), 178–88, 2003.
Heckman, J. J., R. L. Matzkin and L. Nesheim, "Nonparametric Identification and Estimation of Nonadditive Hedonic Models," *Econometrica*, 78(5), 1569–91, 2010.
Hill, R. J., "Hedonic Price Indexes for Residential Housing: A Survey, Evaluation And Taxonomy," *Journal of Economic Surveys*, 27(5), 879–914, 2013.
ILO, IMF, OECD, Eurostat, UNECE, and the World Bank, *Consumer Price Index Manual: Theory and Practice*, 2004.
Ivancic, L. and K. J. Fox, "Understanding Price Variation Across Stores and Supermarket Chains: Some Implications for CPI Aggregation Methods," *Review of Income and Wealth*, 59(4), 629–47, 2013.
Johnston, A. and J. Ozment, "Economies of Scale in the US Airline Industry," *Transportation Research Part E: Logistics and Transportation Review*, 51, 95–108, 2013.
Kostovetsky, L. and J. B. Warner, "Measuring Innovation and Product Differentiation: Evidence from Mutual Funds," *Journal of Finance*, 75(2), 779–823, 2020.
Kotler, P., Marketing Management: Analysis, Planning, and Control, I e XI ed., Prentice-Hall, Upper Saddle River, NJ, 1967.
Kotler, P., "Philip Kotler's Contributions to Marketing Theory And Practice," *Review of Marketing Research*, 8, 87–120, 2011.
Lancaster, K., "A New Approach to Consumer Theory," *Journal of Political Economy*, 74(2), 132–57, 1966.
Levitt, T., *Marketing Success Through Differentiation—of Anything*. Harvard Business Review, January, 1980.
Milne, G. R., "A Marketing Approach for Measuring Product Market Differentiation and Concentration in Antitrust Cases," *Journal of Public Policy & Marketing*, 11(2), 90–100, 1992.
Morrison, S. and C. Winston, "The Evolution of the Airline Industry–Steven Morrison, Clifford Winston–Google Books," *The Evolution of the Airline Industry*, The Brookings Institution, Washington, DC, 1995.
Pakes, A., "A Reconsideration of Hedonic Price Indexes with an Application to PC's," *American Economic Review*, 93(5), 1578–96, 2003.
Rosen, S., "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *Source Journal of Political Economy*, 82(1), 34–55, 1974.
Sharp, B. and J. Dawes, "What is Differentiation and How Does it Work?," *Journal of Marketing Management*, 17(7–8), 739–59, 2001.
Spady, R. H. and A. F. Friedlaender, "Hedonic Cost Functions for the Regulated Trucking Industry," *The Bell Journal of Economics*, 9(1), 159, 1978.
Triplett, J. E., *Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes* (OECD Science, Technology and Industry Working Papers, No. 2004/09), 2004.
Tsionas, E. G., "Combining DEA and Stochastic Frontier Models: An Empirical Bayes Approach," *European Journal of Operational Research*, 147(3), 499–510, 2003.
Vowles, T. M., "The "Southwest Effect" in Multi-Airport Regions," *Journal of Air Transport Management*, 7, 251–8, 2001.
Zou, B. and M. Hansen, "Impact of Operational Performance on Air Carrier Cost Structure: Evidence from US Airlines," *Transportation Research Part E: Logistics and Transportation Review*, 48(5), 1032–48, 2012.

## Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's web site:

**Appendix S1:** Supporting Information