Ann Oper Res (2010) 173: 145–161
DOI 10.1007/s10479-008-0354-x

Hedonic price function estimation in economics and marketing: revisiting Lancaster's issue of "noncombinable" goods

Ruben Chumpitaz · Kristiaan Kerstens · Nicholas Paparoidamis · Matthias Staat

Published online: 20 May 2008

Abstract Following Lancaster's (J. Political Econ. 74(1):132–157, 1966; Variety, equity and efficiency, 1979) interpretation of his characteristics approach to consumer theory, this contribution focuses on theoretical and empirical arguments questioning the smoothness of traditional hedonic price estimation techniques. Lancaster argued strongly against "combinability", i.e., that any efficient combination of characteristics is feasible and sensible. We therefore explicitly test the impact of convexity using a set of recent non-parametric estimators. The test is carried out on a sample of 114 digital cameras whose price evolution is followed over 6 months. The hypothesis of convexity is rejected using the Li (Econ. Rev. 15(3):261–274, 1996) test. The conclusions point out implications for economics and marketing.

Keywords Hedonic price · Indivisibility · Convexity

1 Introduction

In recent years, the analysis of price dispersion has witnessed an upsurge in the economics literature (see, e.g., the book of Blinder et al. 1998). The early literature was at least partly interested in the study of prices and their fluctuations (or lack thereof) because of the implications for macroeconomic theories of business cycles and unemployment (e.g., Carlton 1989). More recent studies have focused on a variety of issues related to firm strategies, industrial organization, and macro-economics (e.g., Warner and Barsky 1995). While a whole

We acknowledge the most constructive comments of three referees and the help of the editors. We retain responsibility for any remaining errors.

R. Chumpitaz · K. Kerstens (✉) · N. Paparoidamis
IESEG School of Management, CNRS-LEM (UMR 8179), 3 rue de la Digue, 59000 Lille, France
e-mail: k.kerstens@ieseg.fr

M. Staat
DSC GmbH, Carl-Banz-Str. 16a, 69198 Schriesheim, Germany
e-mail: matthias.staat@dsc-gmbh.com

series of alternative theories of price stickiness has been proposed (ranging from nominal contracts, implicit contracts, the dependency of quality on price, psychological pricing points, cost-based pricing rules of thumb, menu costs, etc.), it is fair to say that today there is no agreed theoretical explanation for the persistence of price dispersion observed in both goods and services, wholesale and retail trades.

In marketing there is also a tradition to study price dispersion using a variety of methodologies. For instance, computing price-quality-correlations of different products has been quite popular. However, as the exchange between Hjorth-Andersen (1992), Maynes (1992) and Ratchford and Gupta (1992) reveals, there is no reason to believe that these correlations provide any indication about the degree of market efficiency. While various measures have been interpreted as indicators of market efficiency, Ratchford and Gupta (1992) argue in favour of price characteristics frontiers to delineate the subset of efficient products (in line with, e.g., Kamakura et al. 1988), i.e., products worthwhile buying by fully informed consumers with according preferences. Even in the absence of information on consumer preferences, these measures at least provide an easily computable index of efficiency in markets with differentiated products. Firmly positioning this approach in line with the economics of information literature initiated by Stigler (1961), eventual inefficiencies are attributed to inappropriate search efforts. While economists have often declined explanations in terms of inefficiencies, some models of consumer choice focus on modelling the consumer's imperfect ability to choose due to limited information-processing capacities (see, e.g., de Palma et al. 1994).

This paper aims to refine the hedonic price function method by explicitly testing for the potentially enormous impact of the convexity assumption. We do this in the context of frontier specifications that focus on the lower boundary of the price distribution rather than traditional econometric methods estimating average relationships between price and characteristics. These frontier studies of hedonic price functions offer another way to study the same phenomenon of price dispersion by explicitly allowing for inefficiencies. These studies have mainly occurred in the marketing and operations research literature, not in the traditional economic literature. We see both a theoretical and an empirical reason to pursue this convexity matter.

The theoretical concern about non-convexity goes back to the seminal article of Lancaster (1966) who warned about the impact of indivisibilities ("combinable" and "noncombinable" goods in his parlance). On various occasions Lancaster noted that the issue of noncombinability of goods is central to his rethinking of traditional consumption theory and that it reflects the indivisibility of choosing among different goods. For instance, Lancaster (1979, p. 23) writes: "It is the author's view that, although the original characteristics analysis was based on the combinable and additive model and a considerable literature has grown up around this model, the true power of the characteristics approach is most evident in the analysis of goods, like consumer durables, in which noncombinability is the rule."

In the economic literature, Shephard (1978) is the first to provide an axiomatic analysis of Lancaster's (1966) household production theory and he stresses that the transformation of goods into the characteristics space need not result in a convex frontier, let alone a twice differentiable function (page 454). In marketing, Hjorth-Andersen (1983) is the first to use a simple vector dominance criterion to distinguish between efficient and inefficient product variants and to explicitly argue in favour of this approach as a test for Lancaster's Principle of Efficient Choice. Though no explicit efficiency measures are used, this is equivalent to adopting a non-convex frontier model. In a hedonic frontier setting, the need for assuming non-convexity has been argued for by Fernandez-Castro and Smith (2002) as well as by Lee et al. (2005).[1] However, none of these studies explicitly tests for the impact of convexity.

An empirical argument confirming the necessity to reconsider traditional specifications is the result documented in the industrial organisation literature that products tend to cluster together in characteristics space, both at a given point in time and over time. Classic examples are the case studies by Shaw (1982) on the fertiliser market and by Swann (1985) on microprocessors. Furthermore, Hjorth-Andersen (1988) and Grunewald et al. (1993) studied numerous markets and report the robust result that products tend to cluster around average and high quality, with producers providing rather consistent characteristics qualities (see also Lancaster 1990). This leads Triplett (2001, p. 148) to question the suitability of a smooth hedonic surface given that the commodity space is not dense.

Indeed, consider a producer serving a market where consumers prefer certain distinct combinations of characteristics over others. Several combinations may be attractive for different numbers of consumers. A producer may then either (i) offer a product with a combination of characteristics attractive for many consumers and face competition by other producers, or (ii) identify a niche combination of characteristics where he is able to serve the entire demand and could try to exert market power to some degree. Whereas a non-convex model considers the product efficient despite a possible price mark-up, the convex model is more likely to find a combination of other products that make the niche product appear inefficient. However, inefficiency only exists if there is a more efficient buy available for the consumer that puts him on a higher indifference curve. When convexity is assumed, this is inferred to be the case by construction of the model only.

Thus, researchers cannot simply adopt the convexity assumption, but must test it against a non-convex price hedonic frontier, which makes minimal assumptions (essentially monotonicity) and reasons entirely in terms of vector dominance. The non-convex model leads to fewer inefficient products and lower amounts of inefficiency. This reflects the idea that product designs, despite being clustered, are in the large majority of cases efficient and explains any inefficiency detected in terms of dominance by other existing products rather than hypothetical combinations of products. This feature has proven useful in communicating efficiency results to managers and it is probably the simplest empirical model compatible with the original Lancaster (1966) theory (see Hjorth-Andersen 1983).

This short contribution is structured as follows. The next section develops the basic theoretical framework for the paper. After introducing the literature on hedonic price functions, we define the benefit function and the way it can be computed using nonparametric specifications of the price characteristics lower boundary. In analogy with the recent upsurge in discrete time productivity indices, we also introduce a new Luenberger hedonic price indicator to track efficiency changes over time. The empirical results on a sample of digital cameras sold in Germany are reported in Sect. 2. In particular, we formally test the null hypothesis that both distributions are identical using the Li (1996) test statistic, we employ violin plots summarising the differences between both convex and non-convex distributions, and we trace the rapid price evolution over time using a Luenberger hedonic price indicator. Finally, we revisit the Doyle and Green (1991) distinction between niche and all-round products and apply it to our sample. A concluding section summarises our findings.

[1] Fernandez-Castro and Smith (2002) partly follow Doyle and Green (1991) by imposing linearity together with non-convexity and calling this their favourite model. However, linearity is too strong an assumption for the hedonic price frontier, since it is incompatible with the nonlinear nature of the price characteristics relationship as set out in the theoretical literature on hedonic markets (see below).

2 Hedonic price functions and frontiers

2.1 Hedonic price functions: background

The "characteristics" approach to consumer theory developed by Lancaster (1966, 1979, 1990) writes utility not as a function of a vector of goods, but as a function of their characteristics. Characteristics are normally assumed to be objective, in contrast to the concept of attributes widely used in psychology and marketing. The approach builds upon activity analysis to model the combinations of characteristics in the household production process that can be achieved given assumptions on (i) whether combinations of goods are possible or not in a market, (ii) whether combinations can be made in a linear way or not, (iii) whether the number of characteristics is larger or smaller than the number of goods containing them, etc.

In economics, building upon the "characteristics" approach to consumer theory, Rosen (1974) developed a substantive theoretical framework to study market equilibria for differentiated commodities differing along multiple characteristics. Basically, one seeks to obtain an implicit price for the vector of observed characteristics to aggregate these into a measure of value for a given vector of characteristics (see Mendelsohn 1987 for an early review). A simple specification of this model, allowing for a closed-form solution, is the quadratic equilibrium hedonic price function, whereby implicitly a set of products is supplied exhibiting a range of characteristics such that products follow a normal distribution in characteristics space (see Heckman et al. 2005). However, the recent literature also shows that (i) the implicit price functions for characteristics are in general non-linear, and (ii) the market need not provide a continuum of products in equilibrium, but rather clusters of products exhibiting similar combinations of characteristics whilst products with certain other combinations of characteristics may well be sparsely represented (see, e.g., Ekeland et al. 2004). This recent analysis does not contribute to making the applied researchers' task any easier: apart from the traditional variable selection, specification and estimation problems, questions on the smoothness or not of the hedonic price function in characteristics space, among others, are being added to his/her agenda.

The utilisation of hedonic price functions to estimate, e.g., quality-adjusted price indices or the value of environmental externalities (e.g., air quality) is rather widespread in index theory and environmental economics (see the survey of Smith and Huang 1995). It is noteworthy that despite an early enthusiastic welcome to the use of hedonic price functions in marketing (e.g., Kristensen 1984), very few marketing applications seem to exist.[2]

2.2 Hedonic price frontiers and the benefit function

While the large majority of hedonic price functions have been estimated using traditional econometric methods focusing on the average relation between observed prices and characteristics, more recently a series of applications of frontier specifications to characterise the price quality correspondence and to explicitly measure the eventual presence of price

[2] A marketing application is Koschat and Putsis (2002) addressing the issue of media pricing and audience delivery on the basis of targeting audience segments with specific demographic characteristics. Silver (2000) analyses the UK video-recorder market using scanner data and innovates on several points of relevance to marketing (e.g., by including a proxy for the price-cost margin, by excluding models with low sales, etc.). While this selection may ignore other hedonic price studies relevant to marketing, the clear potential for price setting and communication strategies makes one wonder why so few hedonic price function applications have been around in marketing.

inefficiencies emerged.[3] While the main results of the traditional econometric approach are implicit prices, the frontier approach yields first and foremost efficiency measures indicating any deviations from the frontier. Furthermore, in the convex case this frontier approach can also generate implicit prices (see, e.g., Munn and Palmquist 1997). Obviously, these implicit prices are unavailable in case of a non-convex specification which is the main focus of our investigation.

Kamakura et al. (1988) seems to have been the seminal study of this kind in the marketing literature. Most of these studies employ nonparametric frontier models (see, e.g., Estellita Lins et al. 2005 or Mouchart and Vandresse 2007 for recent examples), though a few opt for stochastic parametric frontiers (e.g., Munn and Palmquist 1997).[4] The majority of studies explicitly refers to informational imperfections as a source of eventual inefficiencies, but few offer an external validation for this interpretation.[5]

In addition to all other advantages outlined above, we resort to nonparametric frontier models to test for the convexity assumption, which does not seem possible in a parametric approach.[6] Basically, the multidimensional characteristics are simply combined with the price dimension in a flexible piecewise linear frontier model, whereby optimal weights are objectively determined as a result of an optimization program seeking a minimum price for given minimal maintained assumptions on the possible combinations of observed characteristics. Exploiting the relation between efficiency measures and goodness-of-fit measures used for hypothesis testing (see Färe and Grosskopf 1995), the comparison of efficiency relative to both convex and non-convex nonparametric frontier models amounts to a test of convexity.

If $u(x)$ is a utility function defined over a choice set X of combinations of prices p and vectors of characteristics z (i.e., $x = (p, z) \in X$, whereby $p \in R_{++}$ and $z \in R^n_+$), and g is a vector or reference bundle used for utility comparisons ($g \in R_{++} \times -R^n_+$, with $g \neq 0$), then the *benefit function* with reference g is defined for $x \in X$ and reference utility value u by:

$$B(g; x) = \begin{cases} \sup_\beta\{\beta : u(x - \beta g) \geq u, x - \beta g \in X\} & \text{if} \\ \quad x - \beta g \in X \text{ and } u(x - \beta g) \geq u \text{ for some } \beta \\ -\infty \quad \text{otherwise.} \end{cases} \tag{1}$$

This benefit function (see Luenberger 1992) or distance function is naturally dual to the expenditure function in consumer theory and shows the (normally) semi-positive benefit that is possible by moving from a given bundle x into the direction of g while maintaining

[3] In marketing, Maynes (1975) seminal work introduces the “perfect information frontier” (PIF) measuring excess prices paid for given quality attributes due to informational imperfections. The PIF is defined in terms of piecewise linear segments connecting goods in a two dimensional space of price and an aggregate quality index (based on subjective weightings) and the lowest price for a given quality is determined by visual inspection. This construct has been criticised right from the outset by, e.g., Triplett (1975) who contests the subjective and cardinal nature of the quality index.

[4] See Lovell (1993) for an overview of these various methods.

[5] In a related literature on wage inefficiencies, Polachek and Robst (1998) report an external validation of this informational asymmetry interpretation. They define incomplete information in a residual way as the difference between a worker’s wage and the maximum potential wage estimated via a stochastic frontier and compare these estimates to a direct measure of worker information about the functioning of labour markets (e.g., questions about matching job titles with job descriptions, educational requirements for certain jobs, and selecting the highest paying job among two): they find a significant positive link.

[6] These nonparametric frontier models are also known under the moniker Data Envelopment Analysis models (DEA) in the operations research literature (see Cooper et al. 2000 or Ray 2004 for an overview).

the reference utility level u.[7] Thus, the benefit function geometrically indicates the farthest one can move away from a given commodity vector x into the direction of the reference commodity vector g while still achieving the reference utility u. Notice that in the hedonic price frontier setting the benefit function is new. It seeks for utility improvements in the direction of lower prices and higher values of the characteristics vector. The benefit function can be conceived as an auxiliary function derived from the utility function having a cardinal meaning. It is traditionally defined at the level of the individual consumer similarly to the utility function. Details on its properties are found in Luenberger (1992).

It is important to define efficiency in the price characteristics space in this general manner to allow maintaining and developing links with traditional economic welfare analysis. In particular, it is well-known that the individual benefit functions can be aggregated to define social welfare measures (see, e.g., Luenberger 1992). Notice that Fernandez-Castro and Smith (2002) define a partial efficiency measure only seeking improvements in characteristics space at the cost of ignoring the price dimension, while Lee et al. (2005) employ a multi-dimensional, range adjusted, slack-based efficiency measure, that—to the best of our knowledge—has never been explicitly linked to a consumer value function.

In practice, most frontier studies do not analyse individual consumer choices, but rather notional choices based upon list prices. This is also the case of our study where transactions are unavailable and only list prices are known. In such cases, one can think of the efficiency analysis in terms of the representative consumer comparing available heterogeneous products.

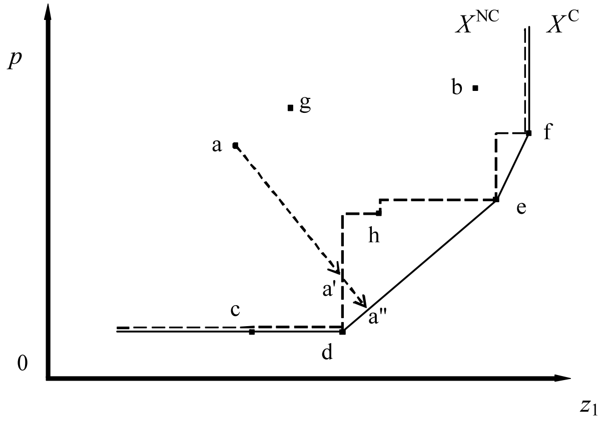
To present the intuition of this approach, we draw a two-dimensional Fig. 1 representing the price on the vertical axis and one characteristic on the horizontal axis. In Fig. 1, for a given sample of observed characteristics and prices denoted by points a to h, and assuming that the characteristic is desirable (more is better), a piecewise linear boundary envelops these data on the southeast side. For this sample, points c, d, e and f define the linear segments of the boundary and offer the best combinations of quality and price from the viewpoint of the consumer. Points in the interior of this boundary are inefficient. When convexity is questioned, it is always possible to look at the discrete observations solely and to reason in terms of vector dominance. This leads to the staircase shaped price-characteristics frontier in the same Fig. 1 which defines a smaller set of possible combinations. Now points c, d, e, h and f define the boundary of a staircase shaped price-characteristics frontier. Fewer observations are found to be inefficient. For instance, point h is now part of the frontier and no longer inefficient.

When measuring efficiency using the benefit function for observation a, for instance, we project this unit onto the hedonic price frontier into the direction of its own characteristic vector and the negative of its price. This leads to projection points a′ on the non-convex frontier and a″ on the convex frontier. Obviously, the distance to the convex frontier is longer and the product thus appears more inefficient.

Assume one observes the price and a vector of desirable characteristics for $k = 1, \ldots, K$ products. To measure price reductions and characteristic improvements using the benefit function (β) relative to this sample of observations, one simply needs to solve for each

[7] Chambers et al. (1996) explore the link between the benefit function and the directional distance function in production theory. In the empirical application, the direction vector g is the evaluated observation itself.

Fig. 1 Hedonic price frontier using nonparametric frontier estimators



observed product j the following mathematical programming problem:

$$
\begin{aligned}
&\max_{\beta,\lambda^k} \quad \beta \\
&\text{s.t.} \quad z_n^j + z_n^j\beta \le \sum_{k=1}^{K} \lambda^k z_n^k, \quad n = 1, \dots, N, \\
&\qquad p^j - p^j\beta \ge \sum_{k=1}^{K} \lambda^k p^k, \\
&\qquad \sum_{k=1}^{K} \lambda^k = 1, \\
&\qquad \beta \ge 0, \quad \lambda \in \Gamma,
\end{aligned}
\tag{2}
$$

where the set $\Gamma \in \{\Gamma^{C}, \Gamma^{NC}\}$ has two subsets, with: (i) $\Gamma^{C} = \{\lambda \in R_+^K\}$ and (ii) $\Gamma^{NC} = \{\lambda \in (0, 1)\}$ representing the hypotheses of convexity and non-convexity, respectively. Notice that in this formulation the direction vector is the price and characteristics vector of the product being evaluated (i.e., $g = (p^j, z^j)$), which leads to a proportional interpretation.

Providing some additional interpretation on this mathematical programming problem, the first N constraints compare the given characteristics of product j along each of its dimensions (on the left hand side (LHS)) with all possible combinations of characteristics of the K products in the sample along each of the same dimensions (on the right hand side (RHS)). The activity vector (λ) allows searching for optimal combinations of characteristics. The inequality sign indicates that we are looking for better combinations of characteristics.[8] The second constraint compares the price of product j on the LHS to all possible combinations of prices of the K products in the sample on the RHS. The inequality now indicates that we are looking for the lowest combinations of observed prices. The sum of the activity vectors equalling unity in the third constraint is essential to allow for a flexible envelopment of the price-characteristics data. This is necessary given the non-linear nature of the

[8]When some characteristics are undesirable (less is better), then the first constraint needs to be rewritten in that the weak inequality must be reversed for the dimensions corresponding to these undesirable characteristics.

price-characteristics relationship developed in recent hedonic theory (see above). The latter convex frontier fits the non-convex frontier closest (i.e., it is its convex closure).[9] This implies that if the convexity assumption is rejected in the tests reported on below, it would be rejected for all other convex frontiers (e.g., the ones employing traditional parametric methods).

When for a given product j represented by a price p^j and a characteristics vector z^j (hence, the pair (p^j, z^j)) the efficiency measure equals zero ($\beta = 0$), the product is considered efficient. When the efficiency measure for a product j is greater or equal to zero ($\beta \geq 0$), the product is considered inefficient. In such a case, the efficiency measure indicates the % reduction in price and the % expansion of characteristics that is possible. Furthermore, the optimal activity vector (λ) indicates the combinations of products (convex case) or the single product (non-convex case) that do better in that lower or equal prices are available with better characteristics.

2.3 Hedonic price frontiers over time: a Luenberger hedonic price indicator

To provide some basic idea about the evolution of efficiency over time and in the spirit of Chambers (2001, 2002), we compute a new Luenberger hedonic price indicator with constant quality dimensions based on the benefit function specified above. In general, a Luenberger hedonic price indicator measures the change of the "consumption technology" for heterogeneous goods, i.e., goods whose product characteristics and price may vary over time. Denoting the time-related versions of the static benefit function as

$$B_{T(b)}(g^a; x^a) = \sup_{\beta}\{\beta : u(x^a - \beta g^a) \geq u, x^a - \beta g^a \in X^{T(b)}\}$$

where $(a, b) \in \{t, t+1\} \times \{t, t+1\}$, this Luenberger hedonic price indicator $L(p^t, z^t, p^{t+1}, z^{t+1}; g^t, g^{t+1})$ is in general defined as follows:

$$\begin{aligned} &L(p^t, z^t, p^{t+1}, z^{t+1}; g^t, g^{t+1}) \\ &\quad = \frac{1}{2}[(B_{T(t)}(p^t, z^t; g^t) - B_{T(t)}(p^{t+1}, z^{t+1}; g^{t+1})) \\ &\qquad + (B_{T(t+1)}(p^t, z^t; g^t) - B_{T(t+1)}(p^{t+1}, z^{t+1}; g^{t+1}))]. \end{aligned} \tag{3}$$

In a traditional consumer context, Chambers (2001, p. 114) talks about a Luenberger commodity indicator. It is perhaps easier to conceive it as the difference-based version of the popular Malmquist productivity index, but now phrased in terms of prices and characteristics using a benefit function defined over a consumption technology. This general formulation allows for changes in both the characteristics vector and the price.

This Luenberger hedonic price indicator is additively decomposed as follows:

$$\begin{aligned} &L(p^t, z^t, p^{t+1}, z^{t+1}; g^t, g^{t+1}) \\ &\quad = [B_{T(t)}(p^t, z^t; g^t) - B_{T(t+1)}(p^{t+1}, z^{t+1}; g^{t+1})] \\ &\qquad + \frac{1}{2}[(B_{T(t+1)}(p^{t+1}, z^{t+1}; g^{t+1}) - B_{T(t)}(p^{t+1}, z^{t+1}; g^{t+1})) \\ &\qquad + (B_{T(t+1)}(p^t, z^t; g^t) - B_{T(t)}(p^t, z^t; g^t))], \end{aligned} \tag{4}$$

[9] The convex hedonic price frontier model is formally similar to the convex variable returns to scale production frontier introduced by Banker et al. (1984) in the DEA literature.

where the first difference (inside the first square brackets) measures efficiency change of the benefit function between periods t and $t+1$, while the arithmetic mean of the two last differences (inside the second square brackets) captures the change of the consumption technology.

Positive values indicate improvements in the consumption technology for heterogeneous goods (in this case: same characteristics for lower price); negative values signify declines (in this case: same characteristics for higher price). In our empirical application, the characteristics remain constant and only the price changes over time. In this way, it is in principle possible to trace price changes over the life cycle of a given product. When it is possible to link successive models over time and to handle both disappearing old products and emerging new products, also simultaneous price and quality changes could be captured with this new Luenberger hedonic price indicator (we refer the reader to discussions on hedonic price indices for more details: e.g., Triplett 2001).

3 Empirical analysis

Our sample consists of 114 models of digital cameras from 17 different brands observed over 6 months from March to August 2005 on websites in Germany. The quality of these cameras ranges from simple 99 € snapshot devices to prosumer types of cameras selling for 750 €. Apart from price information, five output characteristics are recorded: (i) quality of the picture, (ii) features, (iii) ergonomics, (iv) documentation, and (v) service and support.[10]

The characteristics data are expressed in %, i.e., the best performance in each respective dimension is set at 100% and the other observations are scaled as fractions of this best performing observation. Of these five characteristics, three are (nearly) continuous. (i) Quality ranges between 63 and 100, with 32 distinct values; (ii) features between 32 and 100, with 43 distinct values; (iii) while ergonomics varies between 29 and 100 and has 46 distinct values. This is different for documentation and service. While in theory, these features could have this many distinct values as well, it is not the case in practice. Documentation ranges between 29 and 100 and service between 17 and 100, both variables take only 6 distinct values. Obviously, many companies follow only a few number of standardized service concepts, resulting in only a small number of distinct values. The market seems to have settled at these focal points. Therefore, we treat all characteristics as cardinal data.[11]

Characteristics data come from the website of the computer magazine Chip (www.chip.de), while the price data have been added from the Evendi website (www.evendi.de), a price comparison service. Prices are monthly lows of "online" prices. These data allow an analysis of the efficiency of web-based markets.

Notice that differences in efficiency across products may be due to a variety of reasons, like brand and retailer attributes, the phase of the product life cycle, market structure, etc. Efficiency results are furthermore conditional upon the correct specification of the price characteristics hedonic relationship. It is well-known that when this relationship is misspecified due to unknown characteristics, then the interpretation of these efficiency estimates is problematic (see Varian 1988).

[10] The features are the same in other German magazines that we have checked. The presentation of the results may differ and some magazines prefer to test fewer items and describe them at length. It is our impression that identical to very similar characteristics are used internationally.

[11] Frontier benchmarking methods have been developed to handle variables at various scales of measurement (see, e.g., Cook et al. 1993 and more recent contributions).

Table 1 Descriptive statistics of benefit function per model and per month

		Mean	Trimmed mean*	Stand. dev.	Max	# Efficient observations
Convex model	Month 1	0.189	0.186	0.166	0.576	37
	Month 2	0.182	0.179	0.159	0.546	33
	Month 3	0.175	0.172	0.157	0.529	35
	Month 4	0.186	0.183	0.162	0.530	33
	Month 5	0.206	0.203	0.169	0.594	31
	Month 6	0.222	0.219	0.185	0.642	31
Non-convex model	Month 1	0.078	0.074	0.118	0.410	66
	Month 2	0.080	0.075	0.122	0.453	65
	Month 3	0.079	0.074	0.125	0.455	68
	Month 4	0.083	0.077	0.121	0.464	67
	Month 5	0.109	0.104	0.141	0.503	59
	Month 6	0.119	0.114	0.153	0.528	61

*5% observations excluded

Since we only observe 6 months of prices, an analysis of price variation per product is of little use.[12] Instead, in Table 1 we report the descriptive statistics of the levels of the benefit function for both the convex and non-convex models over the six months in our sample. Average efficiency is higher and the dispersion of the distribution is much smaller for the non-convex model. The average levels of the benefit function per month have the tendency to increase over time for both the convex and non-convex frontier models, indicating that inefficiency levels increase over time. The amount of efficient observations nearly doubles in the non-convex model. Note that the non-convex model seeks for complete vector dominance and is therefore much more lenient in this respect.

We depict a kernel estimate for the densities of the benefit function estimated using both convex and non-convex estimators (see (2)) in Fig. 2 for the first month (March). Clearly, the density of the non-convex model is markedly more skewed. The densities for each of the other months have typically the same basic shapes.

An inspection of the violin plots provided in Fig. 3 demonstrates how the distribution of inefficiencies differs between both models. The violin plot essentially provides visual information by a local density estimator that is rotated by 90° and mirrored symmetrically. Figure 3 consists of six pairs of violin plots (one for each month), whereby the first violin plot of each pair pertains to the convex model and the second one to the non-convex model.

To formally test for the difference between both densities, we employ a test-statistic developed by Li (1996) and refined by Fan and Ullah (1999) that is valid for both dependent and independent variables and that is entirely non-parametric in nature. Recently, the development of statistical representations of nonparametric frontier estimators has led to the development of batteries of rigorous statistical tests of various hypotheses with respect to returns to scale, frontiers shifts and the like (this development is nicely surveyed in, e.g., Banker and Natarajan 2004). However, none of the available tests is particularly designed to

[12] Bayliss and Perloff (2002) analyse price dispersion for a single digital camera over 14 weeks at 41 firms.

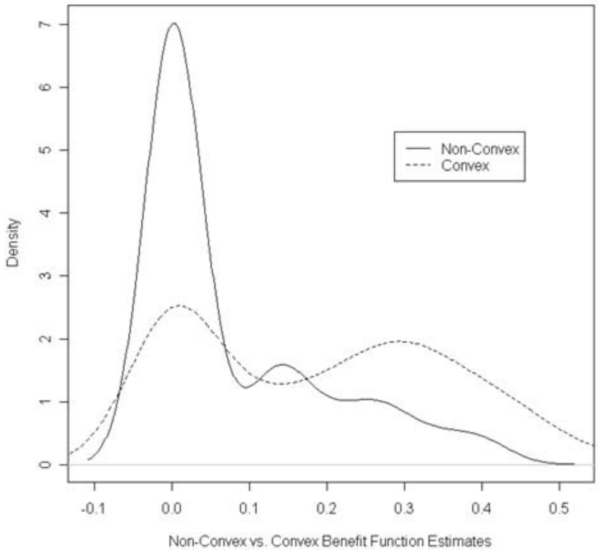


Fig. 2 Kernel densities of benefit function on convex and non-convex models (March 2005)

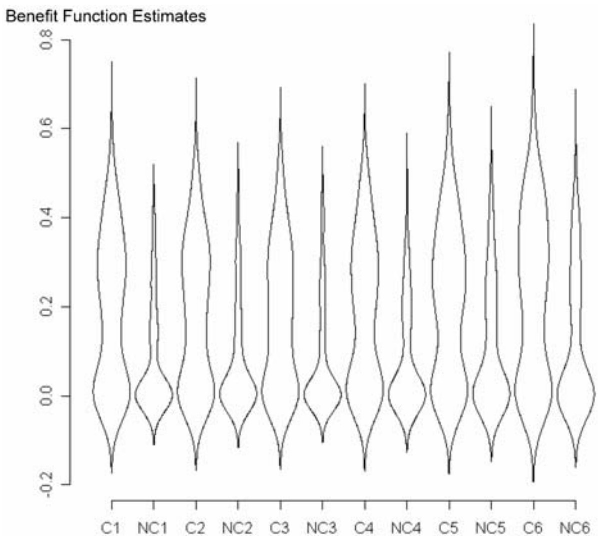


Fig. 3 Violin plots per month: convex vs. non-convex models

tackle the issue of convexity. Therefore, we have opted in the end to employ a conservative approach by using the fully non-parametric Li (1996) test.

The null hypothesis for this Li-test is simply that both distributions are identical; the alternative hypothesis is that they are different. This test statistic asymptotically follows a standard normal-distribution (for small samples, a bootstrap approximation can be employed).[13] To conserve space, we simply list the test statistic per month: 7.95, 7.95, 9.01, 9.49, 7.29 and 7.70. Thus, it turns out that the null hypothesis can be rejected at the 1% significance level (critical value is 2.33) for each month. It is therefore safe to conclude that

[13] Further details on its application in a similar setting can be found in Kumar and Russell (2002).

convex and non-convex estimates yield markedly different results and that non-convexity fits the price characteristics consumption technology for differentiated goods better.

In view of these results it would seem hard to justify the current practice of using convex models. Product designs are efficient in the majority of cases and can probably be sold with a certain mark up relative to the convex frontier, despite the widely observed clustering. To argue otherwise, one must assume that producers prefer offering—relative to the convex frontier—overpriced products in the hope that consumers are ill-informed enough to actually buy these products and that the company would enjoy higher profits from such a strategy than from alternative product strategies. Clearly, such a conglomerate of rather speculative hypotheses makes the convexity assumption loose much of its appeal.

Figure 4 (part a) plots the evolution of the Luenberger hedonic price indicator (see (3)) computed relative to the convex consumption technology over time. While the Luenberger hedonic price indicator is positive (except between the 4th and 5th month), the frontier change is always positive and the technical efficiency change moves from positive to negative over time (see (4)). This means that some products are pushing the price characteristics function downward, but that over time the price dispersion with respect to this shifting function increases and some products tend to lag behind. The negative value for the Luenberger indicator between June and July is probably due to small price increases in anticipation of a rising demand for cameras shortly before the summer holiday season. The evolution of this indicator for the non-convex case is highly similar and therefore discarded.

In addition to this general evolution, we redo this same analysis per cluster of products. Indeed, we performed a hierarchical cluster analysis on the price and characteristics of the cameras (Ward's method) resulting in three clusters representing top-, middle- and low-priced products and brands (see Milligan and Cooper 1985). In Fig. 4, these top-, middle- and low-priced segments are represented in parts b, c and d respectively. The negative value of the Luenberger indicator in June/July only shows up in the low-priced cluster. For the other clusters the Luenberger indicator remains slightly positive, indicating that the price decrease slows down in anticipation of increased demand. A striking result is the negative technical efficiency change for the lower segments, indicating that price dispersion actually increases over time. Thus, it seems that different pricing tactics appear to be relevant for the low segment (part d) compared to the others.

Turning to the analysis of the role models (i.e., products functioning as part of the benchmark for some inefficient product), we are inspired by Doyle and Green (1991) who count the appearance of each product in the optimal activity vector of all other products and distinguish between niche and all-round products depending on whether an efficient product does not or does appear in any other optimal activity vector. Table 2 offers some elementary descriptive statistics on the presence of such all-round products in the model technology of inefficient products per month. The upper part of Table 2 refers to the convex model, the lower part to the non-convex model.

Several observations are noteworthy. First, on average the convex all-round products provide role models for far more other products (see first column), while the standard deviation is, relatively speaking, larger for the non-convex model (see second column). Second, the product corresponding to the maximum number reported in Table 2 always coincides between both models, except in one month. However, it turns out that the convex model attributes importance to some all-round products that do not appear at all under the non-convex model.[14] This difference is entirely due to convexity, whereby two products with extreme but complementary product designs can be combined into one "average" product.

[14] Detailed results are suppressed for reasons of space.

Fig. 4 Luenberger hedonic price indicator: evolution in general and by segment

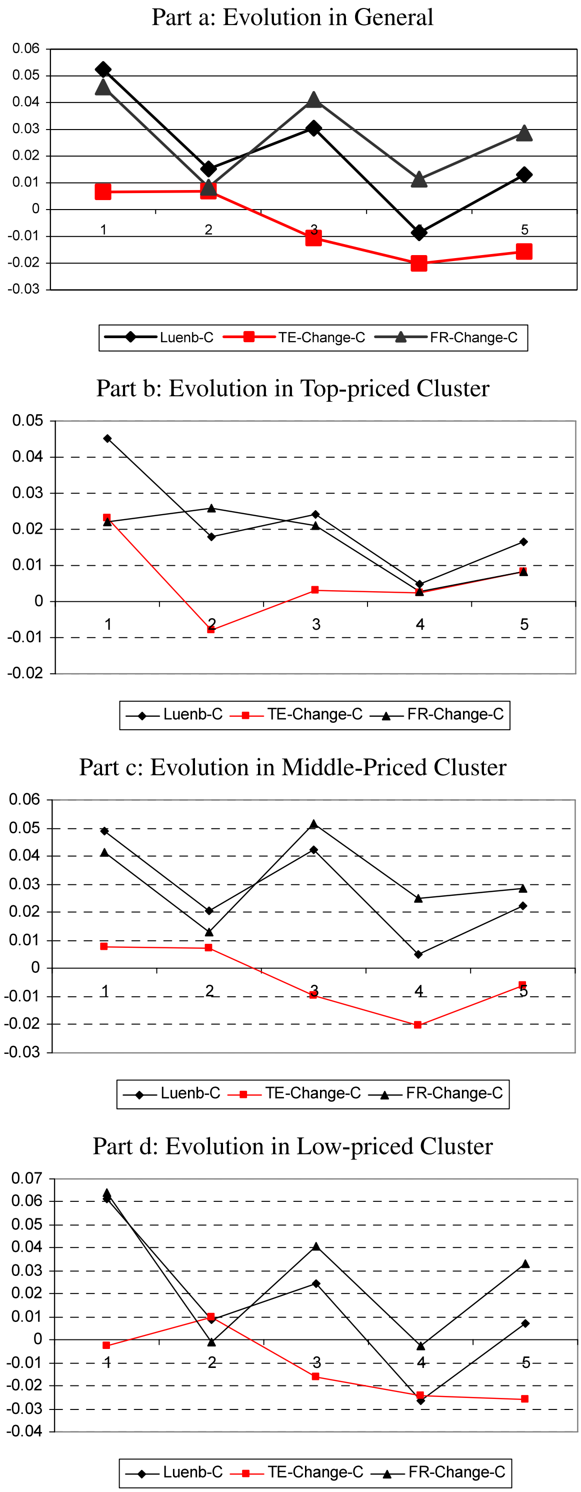


Table 2 Descriptive statistics for the all-round products per model and per month

		Mean	Stand. dev.	Min	Max
Convex model	Month 1	9.555	11.463	1	36
	Month 2	11.167	12.903	1	47
	Month 3	9.593	11.466	1	41
	Month 4	10.32	11.287	1	38
	Month 5	11	12.604	1	42
	Month 6	11.167	12.211	1	42
Non-convex model	Month 1	3.571	4.415	1	18
	Month 2	3.923	6.238	1	24
	Month 3	3.2	5.557	1	23
	Month 4	3	3.950	1	16
	Month 5	4.75	5.971	1	21
	Month 6	4.154	5.728	1	21

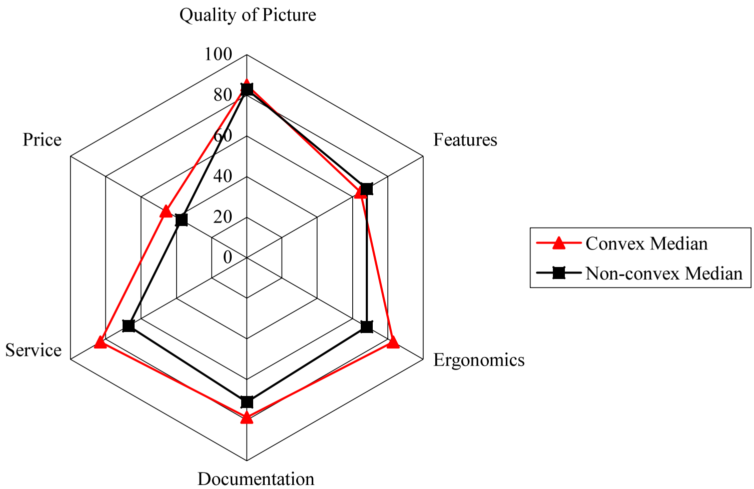


Fig. 5 Radar plot of median characteristics of all-round products per model

Figure 5 is a radar plot showing the medians of each of the five characteristics and price for the all-round products identified under convex and non-convex models. This plot lists each of these six dimensions on a ray originating from the centre to form the hexagon displayed. The amounts were standardized such that the farther away from the centre the "better" the value for the dimension. It may appear surprising that the medians derived under the convex model dominate the non-convex medians, except for the characteristic "features". Note, however, that to be an all-round product under the non-convex model, the product has to offer consistently high parameter values for all dimensions. This need not be the case under the convex model: products that have extreme values for one characteristic and only average values for others may still enter the reference technology of inefficient products. This drives up the median of the characteristics of the all-round products identified under the convex model. In general, however, the shape of the polygon reveals that the medians of all-round products under convex and non-convex models are qualitatively very similar.

Since the non-convex results depend on dominance reasoning solely, they are probably easier to communicate with marketing professionals, not the least because they allow for the existence of efficient niches which play a major role in marketing strategy formulation. In conclusion, since we reject convexity as a hypothesis it may be worthwhile reconsidering the use of convexity for distinguishing between niche and all-round products following the seminal Doyle and Green (1991) analysis.

4 Conclusions

In this contribution, starting from the problem of indivisibilities in the Lancaster (1966) household production model and the empirical regularity of product clustering, we have developed a case for not imposing convexity when estimating hedonic price functions. Making use of the benefit function for the first time in this context, a recent variation on the more traditional distance function in consumer theory, we estimate comparable convex and non-convex hedonic price frontiers using non-parametric extremum estimators.

The empirical analysis uses a short panel of digital cameras whose price evolution is recorded over a half year period, while their other characteristics remain constant. Analysing the densities and formally testing the null hypothesis that both distributions are identical using the Li (1996) test leads us to reject this null hypothesis for each month. Violin plots summarising the differences between both convex and non-convex distributions confirm these formal statistical tests. The rapid price evolution over time is also traced using a Luenberger hedonic price indicator. While the frontier is pushed downwards (towards lower prices), the price dispersion itself seems to increase over time. Finally, revisiting the Doyle and Green (1991) distinction between niche and all-round products shows that this classification may indicate an important share of all-round products whose status entirely depends on the convexity hypothesis. Using dominance analysis to define all-round products makes the results not only easier to communicate, but it also contributes to making them more robust. By allowing for strategic product positioning into niche segments of markets and by reducing the number of seemingly inefficient products, these results point out a small number of problematic product designs dominated by others. This allows focusing on cases in need of a change in design/strategy and makes it possible to identify new niches in the price/characteristics space that marketing professionals could target for future product launches.

Given the formal test results, it is clear this study reinforces the few earlier existing studies employing non-convex rather than convex hedonic price frontiers. Apart from its statistical significance, also the economic relevance of the non-convexity assumption has been underlined in this consumer context. We hope this work contributes to improving the analysis of price dispersion for heterogeneous goods and services in both the areas of economics and marketing. In principle, the methodology developed in this contribution could be applied in a wide variety of hedonic markets. Especially, it seems that the potential of these methods have not yet been fully exploited in the marketing area.

References

Banker, R. D., & Natarajan, R. (2004). Statistical tests based on DEA efficiency scores. In W. W. Cooper, L. M. Seiford, & J. Zhu (Eds.), *Handbook on data envelopment analysis* (pp. 299–321). Boston: Kluwer.

Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, *30*(9), 1078–1092.

Bayliss, K., & Perloff, J. M. (2002). Price dispersion on the Internet: good firms and bad firms. *Review of Industrial Organization*, *21*(3), 305–324.

Blinder, A., Canetti, E. R. D., Lebow, D. E., & Rudd, J. B. (1998). *Asking about prices: a new approach to understanding price stickiness*. New York: Russel Sage Foundation.

Carlton, D. W. (1989). The theory and facts about how market clears: is industrial organization valuable for understanding macroeconomics? In R. Schmalensee & R. D. Willig (Eds.), *Handbook of industrial organization* (Vol. 1, pp. 909–946). Amsterdam: North-Holland.

Chambers, R. G. (2001). Consumers' surplus as an exact and superlative cardinal welfare indicator. *International Economic Review*, *42*(1), 105–119.

Chambers, R. G. (2002). Exact nonradial input, output, and productivity measurement. *Economic Theory*, *20*(4), 751–765.

Chambers, R. G., Chung, Y., & Färe, R. (1996). Benefit and distance functions. *Journal of Economic Theory*, *70*(2), 407–419.

Cook, W. D., Kress, M., & Seiford, L. (1993). On the use of ordinal data in data envelopment analysis. *Journal of the Operational Research Society*, *44*(2), 133–140.

Cooper, W. W., Seiford, L. M., & Tone, K. (2000). *Data envelopment analysis: a comprehensive text with models, applications, references and DEA-solver software*. Boston: Kluwer.

de Palma, A., Myers, G. M., & Papageorgiou, Y. Y. (1994). Rational choice under an imperfect ability to choose. *American Economic Review*, *84*(3), 419–440.

Doyle, J. R., & Green, R. H. (1991). Comparing products using data envelopment analysis. *Omega*, *19*(6), 631–638.

Ekeland, I., Heckman, J. J., & Nesheim, L. (2004). Identification and estimation of hedonic models. *Journal of Political Economy*, *112*(1), 60–109.

Estellita Lins, M. P., de Lyra Novaes, L. F., & Loureiro Legey, L. F. (2005). Real estate appraisal: a double perspective data envelopment analysis approach. *Annals of Operations Research*, *138*(1), 79–96.

Fan, Y., & Ullah, A. (1999). On goodness-of-fit tests for weakly dependent processes using kernel method. *Journal of Nonparametric Statistics*, *11*(1–3), 337–360.

Färe, R., & Grosskopf, S. (1995). Nonparametric tests of regularity, Farrell efficiency and goodness-of-fit. *Journal of Econometrics*, *69*(2), 415–425.

Fernandez-Castro, A. S., & Smith, P. C. (2002). Lancaster's characteristics approach revisited: product selection using non-parametric methods. *Managerial and Decision Economics*, *23*(2), 83–91.

Grunewald, O., Faulds, D. J., & McNulty, M. S. (1993). Evidence on agglomeration in quality space-revisited. *Journal of Industrial Economics*, *41*(2), 205–214.

Heckman, J., Matzkin, R., & Nesheim, L. (2005). Estimation and simulation of hedonic models. In T. Kehoe, T. N. Srinivasan, & J. Whalley (Eds.), *Frontiers in applied general equilibrium: in honor of Herbert scarf* (pp. 277–340). Cambridge University Press: Cambridge.

Hjorth-Andersen, C. (1983). Lancaster's principle of efficient choice: an empirical note. *International Journal of Industrial Organization*, *1*(2), 287–295.

Hjorth-Andersen, C. (1988). Evidence on agglomeration in quality space. *Journal of Industrial Economics*, *37*(2), 209–223.

Hjorth-Andersen, C. (1992). Alternative interpretations of price-quality relations. *Journal of Consumer Policy*, *15*(1), 71–82.

Kamakura, W. A., Ratchford, T. B., & Agrawal, J. (1988). Measuring market efficiency and welfare loss. *Journal of Consumer Research*, *15*(3), 289–302.

Koschat, M. A., & Putsis, W. P. (2002). Audience characteristics and bundling: a hedonic analysis of magazine advertising rates. *Journal of Marketing Research*, *39*(2), 262–273.

Kristensen, K. (1984). Hedonic theory, marketing research, and the analysis of complex goods. *International Journal of Research in Marketing*, *1*(1), 17–36.

Kumar, S., & Russell, R. R. (2002). Technological change, technological catch-up, and capital deepening: relative contributions to growth and convergence. *American Economic Review*, *92*(3), 527–548.

Lancaster, K. (1966). A new approach to consumer theory. *Journal of Political Economy*, *74*(1), 132–157.

Lancaster, K. J. (1979). *Variety, equity and efficiency*. Oxford: Basil.

Lancaster, K. J. (1990). The economics of product variety: a survey. *Marketing Science*, *9*(3), 189–206.

Lee, J.-D., Hwang, S.-W., & Kim, T.-Y. (2005). The measurement of consumption efficiency considering the discrete choice of consumers. *Journal of Productivity Analysis*, *23*(1), 65–83.

Li, Q. (1996). Nonparametric testing of closeness between two unknown distribution functions. *Econometric Reviews*, *15*(3), 261–274.

Lovell, C. A. K. (1993). Production frontiers and productive efficiency. In H. Fried, C. A. K. Lovell, & S. Schmidt (Eds.), *The measurement of productive efficiency: techniques and applications* (pp. 3–67). Oxford: Oxford University Press.

Luenberger, D. G. (1992). Benefit functions and duality. *Journal of Mathematical Economics*, *21*(5), 461–481.

Maynes, E. S. (1975). The concept and measurement of product quality. In N. E. Terleckyj (Ed.), *Household production and consumption* (pp. 529–560). New York: National Bureau of Economic Research.

Maynes, E. S. (1992). Salute and critique: remarks on Ratchford and Gupta's analysis of price-quality relations. *Journal of Consumer Policy*, *15*(1), 83–96.

Mendelsohn, R. (1987). A review of identification of hedonic supply and demand functions. *Growth and Change*, *18*(1), 82–92.

Milligan, G. W., & Cooper, M. C. (1985). An examination of procedures for determining the number of clusters in a data set. *Psychometrika*, *50*(2), 159–179.

Mouchart, M., & Vandresse, M. (2007). Bargaining powers and market segmentation in freight transport. *Journal of Applied Econometrics*, *22*(7), 1295–1313.

Munn, I. A., & Palmquist, R. B. (1997). Estimating hedonic price equations for a timber stumpage market using stochastic frontier estimation procedures. *Canadian Journal of Forest Research*, *27*(8), 1276–1280.

Polachek, S., & Robst, J. (1998). Employee labor market information: comparing direct world of work measures of workers' knowledge to stochastic frontier estimates. *Labour Economics*, *5*(2), 231–242.

Ratchford, B. T., & Gupta, P. (1992). On estimating market efficiency. *Journal of Consumer Policy*, *15*(3), 275–293.

Ray, S. C. (2004). *Data envelopment analysis: theory and techniques for economics and operations research*. Cambridge: Cambridge University Press.

Rosen, S. (1974). Hedonic prices and implicit markets: production differentiation in pure competition. *Journal of Political Economy*, *82*(1), 34–55.

Shaw, R. W. (1982). Product proliferation in characteristics space: the UK fertiliser industry. *Journal of Industrial Economics*, *31*(1–2), 69–91.

Swann, G. (1985). Product competition in microprocessors. *Journal of Industrial Economics*, *34*(1), 33–53.

Shephard, R. W. (1978). On household production theory. In W. Eichhorn et al. (Eds.), *Theory and applications of economic indices* (pp. 449–485). Würzburg: Physica Verlag.

Silver, M. (2000). Hedonic regressions: an application to VCRs using scanner data. *Omega*, *28*(4), 399–408.

Smith, V. K., & Huang, J. C. (1995). Can markets value air quality? A meta-analysis of hedonic property value models. *Journal of Political Economy*, *103*(1), 209–227.

Stigler, G. J. (1961). The economics of information. *Journal of Political Economy*, *69*(3), 213–225.

Triplett, J. E. (1975). Comments on "the concept and measurement of product quality". In N. E. Terleckyj (Ed.), *Household production and consumption* (pp. 567–574). New York: National Bureau of Economic Research.

Triplett, J. E. (2001). Hedonic indexes and statistical agencies, revisited. *Journal of Economic and Social Measurement*, *27*(3–4), 131–153.

Varian, H. (1988). Revealed preference with a subset of goods. *Journal of Economic Theory*, *46*(1), 179–185.

Warner, E. J., & Barsky, R. B. (1995). The timing and magnitude of retail store markdowns: evidence from weekends and holidays. *Quarterly Journal of Economics*, *110*(2), 321–352.