Measuring the Value of Product Characteristics in the Presence of Price Dispersion

Abstract

Modeling the price of multi-attribute products generally requires an assessment of each attributes' market value. In the presence of price dispersion, when similar products are sold at different prices, hedonic pricing models provide users with biased estimates of attribute value. This paper develops the hedonic pricing literature by proposing data envelopment analysis as a prior means of identifying a sub-sample of products which, after adjusting for attribute provision, display no price dispersion. These products then display a homogenous link between attributes and price, which can be modeled using hedonic pricing. This paper implements and evaluates this two-stage approach using 1000 observations from the UK mortgage market.

Key Words: Price dispersion, hedonic pricing, data envelopment analysis, bias, consumption efficiency, mortgages.

1. INTRODUCTION

This study proposes a means of measuring the value of product attributes when the link between the provision of differentiated product attributes and price is complicated by the presence of price dispersion.

Price dispersion is a violation of the economic law of one price. Products of comparable quality, in a perfectly functioning market, should sell for an identical price. In reality, the law of one price appears weak, with empirical evidence providing support for price dispersion across many industries. For recent examples see (Frank and Lamivard, 2009) the Swiss private medical insurance; (Martin-Olivier, Salas-Fuma and Saurina, 2008) Spanish bank loans; (Lee, Park, Oh and Kim, 2008) the Korean personal computer market; (Grover, Lim and Ayyagari, 2006) for a variety of internet based markets; and (Zhao, 2006) in the US grocery market.

While a significant amount of academic research, led by, (Stigler, 1961), (Reinganum, 1979), (Salop and Stiglitz, 1982), (Varian, 1980), (Borenstein and Rose, 1994) has focused on explaining price dispersion, less attention has been placed on the consequences of price dispersion within empirical investigations of differentiated products, attribute value and price.

A common means of assessing attribute value is to use hedonic pricing models, see (Court, 1938), (Lancaster, 1966) and (Rosen, 1974). Assuming that the law of one price holds, hedonic pricing models provide unbiased estimates of individual attribute price coefficients, (Rosen, 1974). Unfortunately, such an assumption is at odds with the plethora of evidence supporting the existence of price dispersion.

Business and economic academics, commercial pricing specialists and government agencies measuring inflation, (see Triplett, 1991), are all interested in measuring the value of differentiated attributes. Given the potential for price dispersion to generate bias within hedonic pricing models, then a hedonic approach capable of accommodating price dispersion would be extremely desirable.

Attempts to specify price dispersion as a stochastic element within a hedonic framework are evident within the literature. For example, (Caudill, 1993) and housing rent controls, (Munn and Palmquist, 1997) and lumber prices, (Kalita, Jagpal, and Lehmann, 2004) and consumer electronics; and (Lee et al., 2008) and personal computing, all apply the stochastic frontier developed by (Aigner, Lovell, and Schmidt, 1977). Unfortunately, the stochastic frontier specification also produces biased estimates of attribute price coefficients, see (Bardhan, Cooper, and Kumbhakar, 1998).

The purpose of this study is to accommodate price dispersion and avoid statistical bias in the hedonic pricing approach. To achieve this goal, this study applies hedonic pricing models to a data set pre-screened for price dispersion. This approach requires a prior analysis of the data using non-parametric efficiency measurement techniques. Such techniques are capable of identifying products where the quality adjusted price is the same and the law of one price holds. Focusing the hedonic price estimator on the observations where the law of one price holds results in unbiased estimates of the attribute price coefficients.

In order to illustrate this two-stage approach, the study investigates a sample of over 1,000 UK mortgage products, differentiated along 14 dimensions. The results show that a traditional one-stage hedonic model results in a poor fit of the data and a number of unexpected negative price effects. This study's preferred two-stage hedonic approach provides a better model of the data.

The important contributions of this study are: First, an effective empirical approach for dealing with price dispersion within hedonic pricing models, leading to unbiased results; which addresses a significant weakness in the existing approaches proposed by (Kalita et al., 2004) and (Lee et al., 2008). Second, the proposed two-stage approach is relatively simple to implement and accessible to researchers and practitioners. Non-parametric efficiency measurement is widely understood, appropriate software packages exists and problems tend not to be computationally demanding. Third, this study extends and supports the growing literature characterized by (Fernandez-Castro and Smith, 2002), (Lee et al., 2004, 2005 and 2008), (Ward, 2009) and (Chumpitaz, Kerstens, Paparoidamis and Staat, 2010), which investigate economic consumption problems from an efficiency perspective.

The discussion proceeds as follows. Section two develops a framework for addressing price dispersion when seeking to estimate hedonic pricing functions. Section three highlights the competing non-parametric efficiency approaches and provides a useful guide for future researchers. Section four provides an overview of the data and presents the results. Section five offers conclusions.

2. Hedonic Pricing, Price Dispersion And Consumption Efficiency

(Rosen, 1974) proposes the hedonic pricing formula (1), which links the price p_i of the ith product to the x_j attributes of the product.

$$p_i = f(x_{i1}, x_{i2} \ldots x_{ij}) \quad (1)$$

Unfortunately, price dispersion prevents the estimation of unbiased price coefficients from a hedonic model. Price dispersion as an empirically verified pervasive feature of markets, (Frank and Lamivard, 2009, Lee et al., 2008, Martin-Olivier, Salas-Fuma and

Saurina, 2008, Grover, Lim and Ayyagari, 2006 and Zhao, 2006), should therefore be recognized within empirical studies of price, value and differentiation.

Related to price dispersion is the concept of consumption efficiency, (see Lee et al., 2004, Lee et al., 2005 and Lee et al., 2008), which measures the extent to which a consumer could achieve a higher level of utility for a given price with an alternative product. In contrast; and from the perspective of a firm seeking to understand the price that can be charged for a given combination of product characteristics, an efficient product maximizes the output (price) for a given level of input (product features).

In the efficiency literature radial approaches view inefficiency as linear or proportionate changes required in all inputs or outputs to achieve full efficiency. In contrast, non-radial approaches examine efficiency from a non-linear perspective by considering the possibility of non-proportionate improvements in some or all inputs and outputs simultaneously.

Following a similar approach to (Bardhan et al. 1998) $0 < \tau_i < 1$ represents the unknown radial inefficiency of the i^{th} product. An inefficient input can then be characterized as $\hat{x}_{ij} = \tau_i x_{ij}$; and similarly an inefficient output can be defined as $\hat{p}_i = \tau_i y_i$. Only when τ_i=1 is the i^{th} product efficient with $\hat{x}_{ij} = x_{ij}$ and $\hat{p}_i = p_i$. Introducing inefficiency into our hedonic pricing function, then (1) becomes:

$$\hat{p}_i = f(\hat{x}_{i1}\hat{x}_{i2}...\hat{x}_{ij}) \quad (2)$$

By allowing τ to vary by input j and be no longer common within the i^{th} observation, then $\hat{x}_{ij} = \tau_{ij} x_{ij}$ and (2) becomes a non-radial expression of possible inefficiencies.

For the purposes of exposition, assuming that *f* is linear, then an empirical estimable version of (2) is shown in (3), where the error term ε is $N(0,\sigma_{\varepsilon}^{2})$:

$$\hat{p}_i = \beta_0 + \beta_1\hat{x}_{i1} + \beta_2\hat{x}_{i2} ..... + \beta_j\hat{x}_{ij} + \varepsilon_i \ (3)$$

Due to researchers only observing $\hat{p}_i$ and the $\hat{x}_{ij}'s$, then ordinary least squares will produce biased estimates of the β price coefficients on the x_{ij}'s. This bias is because the β_j*'s* are not constant across efficient and inefficient products, varying by τ_i. Furthermore, since τ_i varies by product, then each observation also has its own intercept, $\beta_0 + \tau_{ij}$. Any application of (3) to a pooled data set of efficient and inefficient products will result in biased β estimates for the underlying technology that describes the relation between attributes and price for fully efficient products. This problem is alluded to by (Cubbin and Murfin, 1987) when undertaking market share analysis and (Bardhan et al., 1998) when examining production efficiency measurement.

The stochastic frontier approach developed by (Aigner et al., 1977) is an alternative specification for (3). Under this approach the error term $\varepsilon_i = v_i + u_i$, where v_i is $N(0,\sigma_{v}^{2})$; and $u_i \geq 0$ capturing the inefficiency.

$$\hat{p}_i = \beta_0 + \beta_1\hat{x}_{i1} + \beta_2\hat{x}_{i2} + ...v_i + u_i \ (4)$$

Used by (Kalita, Jagpal, and Lehmann, 2004) to examine the quality-price relationship in the durable and non-durable goods markets, (Caudill, 1993) to examine rent controls across differentiated properties; and, (Lee et al., 2008) to examine the personal computer market, the stochastic frontier approach is also known to produce biased estimates of the β coefficients. This bias is because the stochastic frontier assigns a statistical distribution to all the u_i's, not just those of the inefficient products, leading to

$E(u_i) > 0$. The estimated β_j coefficients from (4) are similar to the biased coefficients from (3) and the intercept value increases by $E(u_i) > 0$, see (Bardhan et al. 1998).

These statistical problems can be resolved if the empirical analysis focuses on fully efficient consumption choices, where the law of one price holds; and either τ_i=1 for radial inefficiency and τ_{ij}=1 for non-radial efficiency problems. In this study, a two-stage approach employs non-parametric efficiency measurements to categorize observations into efficient and inefficient groupings; and then uses only the efficient group of products when estimating the hedonic pricing function as in (5) below.

$$\hat{p}_i = \beta_0 + \beta_1 \hat{x}_{i1} + \beta_2 \hat{x}_{i2} .. + \beta_j \hat{x}_{ij} + \varepsilon_i \quad (5)$$

In such circumstances, the filtering out of inefficient observations in the first stage, means that $\hat{p}_i = p_i$ and $\hat{x}_{ij} = x_{ij}$. Therefore, an estimation of (5) will result in an unbiased estimate of the underlying efficient technology. Whilst not used in the hedonic pricing literature, this two-stage filtering approach is an accepted methodology in the production efficiency literature for finding a more statistically efficient estimator of the underlying technology, (see Arnold, Bardhan, Cooper, and Kumbhakar, 1996; Bardham, 1998; Simar, 1992; and Thiry and Tulkens, 1992).

Interest in the statistical properties of the estimated efficiency measures and the appropriateness of various approaches taken to model efficiency in the second stage, (see Hoff, 2007; McDonald, 2009; Simar and Wilson, 2007; Badin, Dairo and Simar, 2012; and Johnson and Kuosmanen, 2012) are less relevant to this study. In this study the efficiency measures gained from stage-one are only used to select the stage-two sample. The determinants of efficiency are not modelled in stage-two of this study; and therefore, the efficiency measures and by consequence the statistical properties of the efficiency measures are not a consideration in stage-two of this study. However, the approach adopted here is still not without risk, with a possibility of the first-stage filtering process leading to a significant loss of observations. Under such circumstances, the one-stage approach proposed by (Johnson and Kuosmanen, 2012), which permits a simultaneous measurement of efficiency and estimation of the second-stage explanatory model, can be considered. However, this approach also has a significant drawback, which is that computational burden increases at a quadratic rate as sample size increases. Given that some of the consumption efficiency measures discussed below bring their own computational burden through the addition of numerous constraints, further computational burden from a simultaneous estimation procedure is likely to be undesirable.

The next section discusses the various ways in which non-parametric efficiency measures can categorize products as efficient, or inefficient.

3. Non-Parametric Efficiency Measures

Data envelopment analysis, DEA is the most well-known non-parametric efficiency measurement technique; and is closely related to the efficiency work of (Farrel, 1957). DEA consists of comparing a given decision making unit, (good, or service), to a frontier of piecewise linear combinations of all other units. An observation located on the frontier is deemed efficient, see (Coelli, 1996) and (Charnes, Cooper, and Rhodes, 1978).

Insert Figure 1 here

In figure 1 observations A, C, D and E define the (thick) frontier. The frontier observations are fully efficient, while B’s inefficiency is some function of its distance from the frontier.

Keeping inputs and outputs in the same proportions, the initial DEA approach proposed by (Charnes, Cooper, and Rhodes, 1978), CCR, provides a radial measure of efficiency by examining how much a unit has to increase all outputs, or lower all inputs, in order to be on the frontier. In figure 1 the radial expansion from B to b provides a radial measure of B's inefficiency. However, it is important to note that the comparator b does not exist in the sample of products under analysis. Instead, b is a hypothetical construction based on a linear combination of the features of products A and C. To ensure that B is only compared against observed products in the sample, then following (Tulkens, 1993), a free disposal hull, FDH, approach can be used, which creates the dotted frontier in figure 1.

The FDH approach is utilised by (Fernandez-Castro and Smith 2002) to examine consumption efficiency in the European car market. However, there are weaknesses in the approach. Significantly, the prevalence of dichotomous product features is a consideration in consumption efficiency problems. For example, in figure 1 if products A through to E are cars, price is the output, input 1 is engine size and input 2 is number of doors, then for car B, a radial expansion to an efficient frontier requires a proportionate increase in both engine size and number of doors. Size of engine can be reasonably considered a continuous variable, however, number of doors is either 2 or 4; and never 4.25, for example. Therefore, holding the number of doors constant and making a non-radial expansion from B to b' onto the dotted or thick frontier is a more appropriate means of measuring efficiency. The importance of non-linearities between product characteristics and price are also recongised in the hedonic pricing literature, see (Ekeland, Heckman and Nesheim, 2004) and (Chumpitaz et al. 2011).

By considering non-radial efficiency a further opportunity arises which is to allow simultaneous improvements in outputs and inputs, something which the input(output) only orientation of the radial efficiency approaches does not facilitate. Therefore, an approach which considers radial, non-radial, input and output orientations is advantageous. (For a theoretical discussion of these issues and empirical review see, e.g., De Borger, Ferrier and Kerstens, 1998).

Drawing on the work of (Cooper, Park and Pastor, 1999), the free disposal hull range adjusted measure, FDH-RAM is proposed by (Lee et al. 2005); and serves as a means of addressing the concerns outlined above. (Lee et al., 2004 and 2005) highlight many positives of the RAM approach. In particular, the RAM approach considers efficiency in outputs and inputs and thereby accommodates price and or product characteristic inefficiencies. The RAM approach also identifies both radial and non-radial inefficiencies, thereby reducing the misidentification of slack inefficiencies. Importantly, the RAM approach has attractive numerical qualities. As the name suggests, the efficiency measures from the RAM approach are also invariant to the units of measurement, efficiency measures lie strictly between 0 and 1. Furthermore, the method is more capable than other non-parametric efficiency measures of incorporating dummy variable characteristics, which are key features of goods and services, see (Lee et al. 2005). In addition, (Brockett, Cooper, Golden, Rousseau, and Wang, 2005) stress that unlike radial efficiency measures, such as CCR, the RAM approach generates monotonic efficiency measures, which are essential for any second-stage parametric estimation, such as hedonic pricing models. This final quality helps to address the concerns regarding second-stage statistical inference highlighted by (Simar and Wilson, 2008).

However, it should be noted that despite all these attractions (Silva Portela, Castro Borges, and Thanassoulis, 2003) caution that the RAM approach can compare units with very distant peers, which in the product characteristics space means that products can be compared to very distant competitors/substitutes and could, therefore, distort the measure of efficiency.

The first-stage of the analysis in this study uses both the FDH and FDH-RAM to identify two sub-samples of efficient products. A hedonic pricing model is then estimated using the sub-samples of efficient products. The results from this two-stage process are compared to an estimated hedonic pricing model using the entire sample of efficient and inefficient products.

The next section estimates hedonic pricing functions for UK mortgage products using the two-stage approach. UK mortgage products are chosen because providers make strong use of differentiation, data is publically available on over 1,000 products and the key differentiated features are easy to measure, (Ward, 2009).

4. Data And Results

This study uses data on mortgage prices and characteristics from the UK Financial Services Authority's, FSA Consumer Tables, (http://www.moneymadeclear.org.uk/tables). An average borrower, characterized by an average house price and an average loan to value ratio in 2005 is drawn from the Council for Mortgage Lenders dataset. Mortgages available to the average borrower are taken from the FSA Consumer Tables. This approach results in 1075 usable observations.

The study uses the Annual Percentage Rate, APR, as the measure of price. The APR provides a measure of the overall cost of a mortgage for a typical borrower incorporating introductory and standard interest rates, as well as charges incurred in setting up the mortgage.

For each i^{th} mortgage, j product characteristics are collected from the Consumer Tables. The contents of Table 1 provide the definitions for each of these characteristics. As all product characteristics should be increasing in utility for a potential mortgage borrower a number of transformations are undertaken. For example, Setup Costs is transformed by calculating the difference between the maximum value in the sample and the observed value. Setup Costs is then increasing in utility and measures the discount on the maximum price for setting up a mortgage. Tie-in period and the valuation fee are subject to the same transformation.

Table 1 here.

Market share, provided by the Council of Mortgage Lenders, is used to measure brand. Because market share is only available at firm level, the study assumes that brand resides at the level of the firm, rather than the level of the product; and is therefore constant across each firm's *i* product offerings.

Table 2 lists the descriptive statistics for each variable. The large range and high standard deviation for each of the product outputs confirms a high degree of product differentiation across the sample. Table 3 reports the Pearson correlation coefficients between each of the product characteristics. The majority of the correlation values are small; and none are large enough to warrant multi-collinearity concerns when using regression analysis to estimate the hedonic pricing models.

Table 2 here

Table 3 here

The FDH approach identifies 641 efficient observations and the FDH-RAM finds 499 efficient observations. The FDH approach identifies more efficient units than the FDH-RAM approach. This is because in addition to the radial slacks considered by FDH, the FDH-RAM approach also considers non-radial slacks.

Figure 2 provides illustrations of the efficiency distributions for each of the efficiency measures. Both the FDH and the FDH-RAM identify large numbers of inefficient observations in the 80% and 90% range. At lower levels of overall efficiency, the FDH-RAM approach identifies more inefficiency observations than the FDH approach. The measures of inefficiency provide clear evidence for the existence of price dispersion in the UK mortgage market. The consequential risk of bias within a hedonic pricing analysis of UK mortgages is therefore likely.

Finally, the Spearman rank correlation between the two competing approaches of efficiency measurement is 0.84. Being a high correlation score, the competing efficiency measurement approaches appear to be selecting similar observations as efficient and inefficient, suggesting a good degree of comparability between the competing approaches see (Weill, 2004 and Bauer, Berger, Ferrier and Humphrey, 1998).

Figure 2 here

The hedonic pricing models using efficient observations only are estimated next. Following routine procedures within the hedonic literature see, (van Dalen and Bode, 2004), the hedonic pricing models are initially evaluated for the correct functional specification using a (Box and Cox, 1962) transformation; and are then corrected for heteroskedasticity using the (White, 1980) approach.

Table 4 reports the results from a traditional one-stage hedonic estimation as defined by (3), a stochastic frontier hedonic model, as in (4), where the efficiency term u_i is assumed to be half normal; and hedonic models using only efficient observations as in (5). (Estimates for the stochastic frontier hedonic model were also derived assuming a truncated normal and exponential distribution for u_i with no material difference in the estimated results).

Under a Box-Cox transformation the dependent variable Y is transformed $(Y^{\lambda} - 1)/\lambda$. If the estimated λ=1, then the model is best specified as linear. For all models the evidence for λ being nearer to one is strong and so a linear specification is adopted throughout.

Table 4 here

For the one-stage hedonic model and the stochastic frontier specification, a number of the estimated price parameters for the utility enhancing characteristics are negative. These results would indicate that utility enhancing features within a product reduce its value. These findings are illustrative of the potential bias generated by consumption inefficiency. Note that according to table 4 inefficient products are in the main the majority and are therefore likely to have a marked impact on the characteristics of the sample.

Turning to the results from the two-stage approach, all the models show a marked (adjusted) improvement in goodness of fit over the traditional one-stage OLS approach. This is an encouraging result for the proposed method and indicates that the stage-one filtering of inefficient observations has resulted in a more refined sample where the relationship between price and product characteristics is more consistent and easily modeled. The estimated coefficient for UNTIL has become positive and the measure of statistical significance for UNTIL, SETUP, INCENTIVE, RESTRICTION and BRAND have improved.

Negative coefficients on the linked current accounts remain. These results may reflect an element of cross-subsidization. Linked mortgage and current account products are more valuable to banks when the customer frequently operates their current account in deficit. Higher fees and charges on the current account may then subsidize a reduced price on the mortgage product. However, negative coefficients do remain for valuation fees.

The results also show a degree of parameter stability in the second-stage regression results regardless of the non-parametric approach used in stage-one. This finding would suggest that researchers can feel reasonably assured that these competing non-parametric techniques generate consistent results, but that they should still bear in mind the arguments made in section 3 and assess the relevance of each approach within the context of the products under examination.

5. Conclusions

Hedonic pricing models are a simple and powerful tool for academic and commercial users seeking to understand the relationship between product characteristics and price.

(Lee et al., 2008) and (Kalita et al., 2004) try to develop hedonic pricing models by capturing price dispersion, or consumption efficiency, using stochastic frontier specifications. However, both traditional and stochastic hedonic pricing models can produce biased estimates of the fully efficient technology. The bias in this study is a possible cause of the estimated negative price coefficients for utility enhancing product characteristics. These results suggest that price dispersion can severely limit the insights offered by hedonic pricing models.

In seeking to address these concerns, this study proposes a two-stage approach which enables researchers to identify and focus upon fully efficient products, which after adjusting for quality, display no price dispersion. This approach is simple to implement, uses well developed and accepted non-parametric efficiency measures, requires no additional data; and in theory produces unbiased estimates of the price coefficients for each quality enhancing feature of a product.

Empirically, this study provides evidence of price dispersion and consumption efficiency within the UK mortgage market. Rank correlation scores from the competing DEA efficiency approaches show a good degree of comparability between the approaches. Finally, augmenting hedonic price models with efficiency information improves model estimation performance, as evidenced by improved goodness of fit and correctly estimated coefficient signs. These results, coupled with the simple to use approach, should be of significant value to users of hedonic models.

The approach developed in this study may also be subject to limitations and these could be examined by future research. In particular, in the measurement of inflation, hedonic pricing is commonly used to develop quality adjusted price indices for constantly changing technology products, such as computers. If the degree of price dispersion is common between periods, then the problem of price dispersion cancels out in the index. Such a situation could be reasonable for consecutive time periods, such as one or two years. But the notion that price dispersion remains constant for many consecutive periods is unlikely given changes in product technology, changes in competition and changes in consumers' understanding of the key product features. So, quality adjusted measures of inflation over longer time periods using hedonic pricing could be made more robust by a prior assessment of $\tau_{tij} = \tau_{sij}$, where τ_{tij} is the efficiency of the j^{th} characteristic of product *i*, in period *t*.

6. REFERENCES

Aigner DJ, Lovell CAK, Schmidt P. 1977. Formulation and estimation of stochastic frontier production functions. *Journal of Econometrics*, 6, 21-37.

Arnold VL, Bardhan IR, Cooper WW, Kumbhakar SC. 1996. New Uses of DEA and Statistical Regressions for Efficiency Evaluation and Estimation — With an Illustrative Application to Public Secondary Schools in Texas, *Annals of Operations Research*, 66, 255-277.

Bădin L, Daraio C, Simar L. 2012. How to Measure the Impact of Environmental Factors in a Nonparametric Production Model, *European Journal of Operational Research*, 223, 818-833.

Bardhan IR, Cooper WW, Kumbhakar SC. 1998. A simulation study of joint uses of data envelopment analysis and statistical regressions for production function estimation and efficiency evaluation. *Journal of Productivity Analysis*, 9, 249-278.

Borenstein S, Rose NL. 1994. Competition and price dispersion in the U.S. airline industry. *Journal of Political Economy* 102, 653–83.

Box GEP Cox DR. 1962. An analysis of transformations. *Journal of the Royal Statistical Society*, Series B, 1962, 211-243.

Brockett PL, Cooper WW, Golden L L, Rousseau J J, Wang Y. 2005. Financial Intermediary Versus Production Approach To Efficiency Of Insurance Companies. *Journal of Risk and Insurance,* 72, 393-412.

Bauer PW, Berger AN, Ferrier GD, Humphrey DB. 1998. Consistency Conditions for Regulatory Analysis of Financial Institutions: A Comparison of Frontier Efficiency Methods. *Journal of Economics and Business*, 50, 85-114.

Caudill SB. 1993. Estimating the Costs of Partial-Coverage Rent Controls: A Stochastic Frontier Approach. *Review of Economic and Statistics*, 75, 727-731.

Charnes A, Cooper WW, Rhodes E. 1978. Measuring the efficiency of decision making units. *European Journal of Operations Research*, 2, 429-444.

Chumpitaz R, Kerstens K, Paparoidamis N, Staat M. 2010. Hedonic price function estimation in economics and marketing: revisiting Lancaster's issue of "noncombinable" goods. *Annals of Operations Research*, 173, 145.

Coelli T J. 1996. A guide to DEAP Version 2.1: A data envelopment analysis (computer) program. *CEPA working paper 96/8*, Department of Econometrics, University of New England.

Cooper WW, Park KS, Pastor J T. 1999. RAM: A Range Adjusted Measure of Inefficiency for Use with Additive Models, and Relations to Other Models and Measures in DEA *Journal of Productivity Analysis* 11, 5–42.

Court A. 1938. Hedonic price indexes with automotive examples. *The dynamics of automobile demand*, New York, General Motors Corporation.

Cubbin JS, Murfin AJ. 1987. Regression analysis versus linear programming in the analysis of price-quality relationships: an application to the determination of market shares. *Oxford Bulletin of Economics and Statistics*, 49, 385-399.

De Borger B, Ferrier GD, Kerstens K. 1998. The choice of a technical efficiency measure on the free disposal hull reference technology: A comparison using US banking data, *European Journal of Operational Research*, 105, 427-446

Ekeland I, Heckman JJ, Nesheim L. 2004. Identification and Estimation of Hedonic Models, *Journal of Political Economy*, 112, 60–109.

Farrell MJ, 1957. The measurement of productive efficiency. *Journal of the Royal Statistical Society,* A 120, 3, 253- 290.

Fernandez-Castro AS, Smith PC. 2002. Lancaster's Characteristics Approach Revisited: Product Selection using Non-parametric Methods. *Managerial and Decision Economics, 23, 83-91.*

Frank RG, Lamiraud K. 2009. Choice, price competition and complexity in markets for health insurance. *Journal of Economic Behaviour and Organization*, 71, 550-562.

Grover R, Lim J, Ayyagari R. 2006. The dark side of information and market efficiency in e-markets. *Decision Sciences*, 37, 297-324.

Hoff A, 2007. Second stage DEA: Comparison of Approaches for Modelling the DEA Score, *European Journal of Operational Research*, 181, 425-435.

Johnson AL, Kuosmanen T. 2012. One-Stage and Two-Stage DEA Estimation of the Effects of Contextual Variables, *European Journal of Operational Research*, 220, 559-570.

Kalita JK, Jagpal S, Lehmann DR. 2004. Do high prices signal high quality? A theoretical model and empirical results. *Journal of Production and Brand Management*, 13, 279-288.

Lancaster KJ, 1966. A new approach to consumer theory. *Journal of Political Economy*, 74, 132-157.

Lee J-D, Repkine A, Hwang S-W, Kim T-Y. 2004. Estimating consumers' willingness to pay for individual quality attributes with DEA. *Journal of the Operational Research Society, 55, 1064-1070.*

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, 65-83.

Lee J-D, Park CO, Oh D-H, Kim T-Y. 2008. Measuring consumption efficiency with utility theory and stochastic frontier analysis. *Applied Economics*, 40, 2961-2968.

McDonald J, 2009 Using Least Squares and Tobit in Second Stage DEA Efficiency Analyses, *European Journal of Operational Research*, 197, 792-798.

Martin-Oliver A, Salas-Fumas V, Saurina J. 2008. Search Cost and Price Dispersion in Vertically Related Markets: The Case of Bank Loans and Deposits. *Review of Industrial Organization*, 33, 297-323.

Munn IA, Palmquist RB. 1997. Estimating hedonic price equations for timber stumpage market using stochastic frontier estimation procedures. *Canadian Journal of Forestry Research*, 27, 1276-1280.

Reinganum J. 1979. A simple model of equilibrium price dispersion. *Journal of Political Economy* 87, 851–58.

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

Salop S, Stiglitz J. 1982 The theory of sales: A simple model of equilibrium price dispersion with identical agents. *American Economic Review* 72, 1121–30.

Silva Portela MCA, Castro Borges P, Thanassoulis E. 2003. Finding closest targets in non-oriented DEA models: The case of convex and non-convex technologies *Journal of Productivity Analysis,;* 19, 251 – 269.

Simar L. 1992. Estimating Efficiencies from Frontier Models with Panel Data: A Comparison of Parametric, Non-Parametric and Semi-Parametric Methods with Bootstrapping. *Journal of Productivity Analysis*, 3, 171-203.

Simar L, Wilson PW. 2007. Estimation and inference in two-stage semi-parametric models of production processes. *Journal of Econometrics*, 136, 31-64.

Simar L, Wilson PW. 2008. Statistical Inference in Nonparametric Frontier Models: Recent Developments and Perspectives, in: Fried H, Lovell CAK, Schmidt S. (eds) *The Measurement of Productive Efficiency and Productivity Change*, New York, Oxford University Press, 421-521.

Stigler G. 1961. The economics of information. *Journal of Political Economy* 69, 213–25.

Thiry B, Tulkens H. 1992. Allowing for Inefficiency in Parametric Estimation of Production Functions for Urban Transit Firms. *Journal of Productivity Analysis*, 3, 45-65.

Triplett JE 1991. Hedonic Methods in Statistical Agency Environments: An Intellectual Biopsy, in:Berndt ER and Triplett JE. (eds.) *Fifty Years of Economic Measurement: The Jubilee of the Conference on Research in Income and Wealth*, Chicago, University of Chicago Press, 207-238.

Tulkens H. 1993. On FDH efficiency analysis: some methodological issues and applications to retail banking, courts and urban transit. *Journal of Productivity Analysis* 4: 183–210.

Van Dalen J, Bode B. 2004. Quality-corrected price indices: the case of the Dutch new passenger car market, 1990-1999. *Applied Economics*, 36, 1169-1197.

Varian H. 1980. A model of sales. *American Economic Review* 70, 651–59.

Ward DR. 2009. Product differentiation and consumption efficiency in mortgage markets. *Journal of Business Research,* 62, 805-809.

Weil W L. 2004. Measuring Cost Efficiency in European Banking: A Comparison of Frontier Techniques. *Journal of Productivity Analysis*, 21, 133 – 152.

White H. 1980 A heteroskedasticity consistent covariance matrix estimator and a direct test of heteroskedasticity. *Econometrica,* 48, 817 – 838.

Zhao Y. 2006. Price dispersion in the grocery market. *Journal of Business*, 79, 1175-1192.

Table 1: Mortgage Product Characteristics

Name	**Product Characteristics**	**Description**
UNTIL	Until/for period	The period of time in months over which the initial introductory discount rate is applied.
TIEIN	Tie-in period	The period of time in months during which early redemption of the mortgage results in penalties being applied to the borrower.
VALUATION	Valuation fee	The fee to be paid by the borrower for the mortgaged property to be valued for the interest of the lender.
SETUP	Set-up fees	Administration costs associated with checking, evaluating and advancing mortgages and funds.
INCENTIVE	Incentives	Fixed level financial incentives paid by the lender to third parties on behalf of the borrowers, typically legal fees, or valuation fees.
RESTRICTION	Restricted availability	The product is only available to customers who are new to the lender, or are current to the lender, (depending on lender)
FLEXIBLE	Flexible features	The provision of an option by the lender which allows borrows to increase payments, redeem the mortgage early, or take payment holidays at no further cost.
LINKED_C	Linked current account	The provision of balance off setting by the lender, where the mortgage account is offset by credit balances on linked current and savings accounts, thus saving interest charges
CONDITIONAL	Conditional insurance	The borrower does not have to take out buildings, contents or life insurance with the lender to be eligible for the loan
CAT_STAN	CAT standard	The mortgage is certified as low cost by the regulator
SELFCERT	Self-certification	The lender requires no proof of earnings from borrower.
DISCT	Discount	The size of the initial discount off the lender's standard variable rate.
CASHBACK	Cashback	Cash incentives remitted to borrower at the beginning of the mortgage. These payments are proportionate to the size of the loan and are typically worth 3-5% of the mortgage value.
BRAND	Brand	The overall market share of the mortgage provider.

Table 2: Descriptive Statistics

	Units	Minimum	Maximum	Mean	Std. Deviation
APR	Percentage	5.1	8.5	6.4	0.5
IVR	Percentage	1.5	7.7	5.4	0.8
UNTIL	Months	0.0	120.0	20.4	20.9
TIEIN	Months	180.0	300.0	278.3	21.8
VALUATION	£'s	0.0	435.0	248.7	118.5
SETUP	£'s	0.0	996.0	610.3	152.6
INCENTIVE	£'s	0.0	750.0	175.9	214.7
RESTRICTION	Yes, No	0.0	1.0	0.2	0.4
FLEXIBLE	Yes, No	0.0	1.0	0.4	0.5
LINKED_C	Yes, No	0.0	1.0	0.1	0.4
CONDITIONAL	Yes, No	0.0	1.0	0.03	0.2
CAT_STAN	Yes, No	0.0	1.0	0.03	0.2
SELFCERT	Yes, No	0.0	1.0	0.2	0.4
DISCT	Percentage	0.0	4.8	1.0	0.7
CASHBACK	£'s	0.0	10300.0	212.3	1193.4
BRAND	Percentage	0.3	26.0	2.6	4.3
N		1075			

Table 3 Correlation Coefficients

PRICE	UNTIL	TIEIN	VALUATION	SETUP	INCENTIVE	RESTRICTION	FLEXIBLE	LINKED_C	CONDITIONAL	CAT_STAN	SELFCERT	DISCT	CASHBACK	BRAND
PRICE	0.01	-0.08	-0.14	0.00	0.00	0.07	-0.17	-0.24	0.11	0.00	0.10	0.20	0.18	0.11
UNTIL		-0.42	0.14	0.04	0.12	-0.17	0.16	-0.26	-0.06	-0.03	-0.10	0.32	-0.07	0.06
TIEIN			-0.09	0.05	-0.14	-0.17	0.40	0.25	-0.01	0.17	0.02	-0.34	-0.42	0.10
VALUATION				0.11	0.14	0.09	-0.12	0.09	-0.02	-0.03	-0.21	0.04	-0.02	-0.08
SETUP					-0.02	0.07	-0.03	0.11	0.01	-0.06	-0.08	-0.02	-0.05	-0.09
INCENTIVE						0.06	0.12	0.05	0.07	0.02	0.07	0.02	0.16	0.10
RESTRICTION							-0.21	-0.10	-0.06	-0.09	-0.21	0.05	0.04	-0.10
FLEXIBLE								0.44	0.11	0.19	0.08	-0.09	-0.11	0.11
LINKED_C									0.07	-0.05	0.07	-0.05	-0.07	-0.06
CONDITIONAL										-0.03	-0.01	-0.01	-0.02	-0.05
CAT_STAN											-0.06	-0.04	0.00	0.23
SELFCERT												0.02	0.06	0.04
DISCT													-0.17	0.06
CASHBACK														0.00
BRAND														

Table 4 Estimated Hedonic Pricing Coefficients From Alternative Estimation Approaches

		BASIC	Stochastic	FDHRAM	FDH
	β_0	4.7064	4.4850	4.1291	4.0800
		(0.0000)	(0.0000)	(0.000)	(0.0200)
UNTIL	β_1	-0.0004	-0.0002	0.0027	0.0028
		(0.5412)	(0.8227)	(0.0015)	(0.000)
TIEIN	β_2	0.0057	0.0057	0.0058	0.0058
		(0.0000)	(0.0000)	(0.000)	(0.0000)
VALUATION	β_3	-0.0005	-0.0005	-0.0002	-0.0002
		(0.0001)	(0.0001)	(0.0003)	(0.1060)
SETUP	β_4	0.0001	0.0001	0.0003	0.0003
		(0.1044)	(0.1282)	(0.0002)	(0.0073)
INCENTIVE	β_5	0.0001	0.0001	0.0001	0.0002
		(0.8660)	(0.8822)	(0.0001)	(0.0324)
RESTRICTION	β_6	0.0995	0.1037	0.2833	0.2340
		(0.0013)	(0.0134)	(0.000)	(0.0000)
FLEXIBLE	β_7	-0.1555	-0.1540	0.0503	-0.0056
		(0.0000)	(0.0000)	(0.0538)	(0.9038)
LINKED_C	β_8	-0.2723	-0.2655	-0.3893	-0.0619
		(0.0000)	(0.0000)	(0.000)	(0.0000)
CONDITIONAL	β_9	0.3943	0.3932	0.2720	0.1853
		(0.0000)	(0.0000)	(0.000)	(0.0543)
CAT_STAN	β_{10}	-0.0563	-0.0582	-0.1824	-0.0726
		(0.5433)	(0.3863)	(0.1431)	(0.5231)
SELFCERT	β_{11}	0.1230	0.1261	0.0758	0.1012
		(0.0008)	(0.0022)	(0.0659)	(0.0596)
DISCT	β_{12}	0.2037	0.2076	0.3590	0.3795
		(0.0000)	(0.0000)	(0.000)	(0.000)
CASHBACK	β_{13}	0.0001	0.0001	0.0002	0.0002
		(0.0000)	(0.0000)	(0.000)	(0.000)

BRAND	β_{14}	0.7669	0.7876	1.6800	1.5529
		(0.0005)	(0.0952)	(0.000)	(0.000)
	λ	1.34 (0.0000)		1.09 (0.0481)	1.04 (0.0362)
	R^2	0.2249		0.4208	0.4172
	Adj R^2	0.2147		0.41117	0.4094
	F(j-1, n)	21.97 (0.0000)		28.55 (0.0000)	41.82 (0.0000)
	Breusch-Pagan	199.92 (0.0000)		192.42 (0.0000)	176.18 (0.0000)
	Log-liklihood		-607.221		
	s(v)		0.3965		
	s(u)		0.2575		
	S		0.4728		
	N	1075	1075	641	499

p values are in parentheses

Figure 1Data Envelopment Analysis

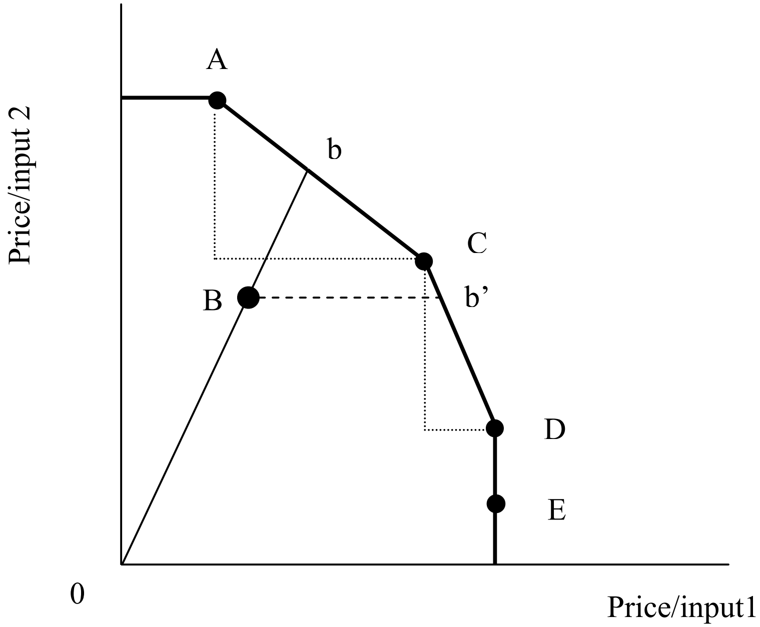

Figure 2 Distribution of Consumption Efficiency by type of DEA Approach