When does variety increase with quality?

Basov, Suren and Danilkina, Svetlana and Prentice, David

La Trobe University, University of Melbourne

2 February 2009

When does variety increase with quality?

Suren Basov[1], Svetlana Danilkina[2], and David Prentice[3] [4]

February 2, 2009

[1]Department of Economics and Finance, LaTrobe University, Bundoora, Victoria 3086, Australia, s.basov@latrobe.edu.au.

[2]Department of Economics, The University of Melbourne, Melbourne, Victoria 3010, Australia, sdanilki@unimelb.edu.au

[3]Department of Economics and Finance, LaTrobe University, Bundoora, Victoria 3086, Australia, d.prentice@latrobe.edu.au

[4]Thanks to Damien Eldridge, Catherine de Fontenay, Tue Gørgens, Phillip Leslie, Phillip McCalman and Lawrence Uren for helpful comments. The first author acknowledges the support of ARC Discovery Grant DP0881381 "Mechanism design under bounded rationality: The optimal contracts in the complex world."

Abstract

Casual empiricism suggests higher quality is associated with greater variety. However, recent theoretical and empirical research has either not considered this link, or has been unable to establish unambiguous predictions about the relationship between quality and variety. In this paper we develop a simple model, which predicts that for low qualities variety should be positively correlated with quality and we establish conditions under which variety will either increase or decrease with quality at higher quality levels. The monopolist uses variety to increase the profitability of price discrimination across product lines of different qualities, by increasing the likelihood consumers choose high price products among products yielding the same utility. We show that the number of varieties offered by the monopolist is greater than the social optimum. The predictions of the model are supported by an analysis of the market for cars. A wide range of car manufacturers are found to offer a hump-shaped distribution of varieties.

Keywords and Phrases: Price discrimination, product variety, bounded rationality, cars.

JEL Classification Numbers: D4, D8, L11, L15, L62.

1 Introduction

A casual look at the shelves of a supermarket or at producer web sites reveals that many goods come in multiple "flavors" as part of a product line. Moreover, higher quality products often have a larger number of flavors. Branded products almost always have more varieties than generics. Further up the quality ladder, there are many more varieties of the premium Pickwick tea than of a basic Lipton brand. We demonstrate in this paper that this also holds for most of the product lines offered by car manufacturers. There is an extensive theoretical literature examining the interaction between quality and variety in imperfectly competitive markets where consumers' tastes are differentiated along one vertical dimension, quality, and one horizontal dimension, flavor.[1] In such models the correlation between quality and variety depends on the way they enter the consumer's utility function, their joint distribution, and the market structure. Therefore, it is impossible to come up with unambiguous predictions suitable for estimation or testing. Lancaster (1990) provides a review of such literature. One important conclusion that he draws from his analysis is the number of varieties produced by a monopolist is less than optimal.[2] We show that this conclusion can be overridden if there is vertical heterogeneity and consumers are, to a very small degree, boundedly rational, i.e. consumers are nearly rational.

[1] See, for example, Ansari, Economides, and Steckel (1998), Neven and Thisse (1990), and Shugan (1989).

[2] For reasons similar to those of a monopolist producing a lower than optimal quantity.

Recent more empirically oriented work by Draganska and Jain (2005) models varieties, as part of product-lines, as horizontal differentiation but does not explicitly consider vertical differentiation highlighted in models of price discrimination. As a consequence, this work does not analyse any link between quality and variety.

In this paper we develop a probabilistic choice model that predicts, first, that for low qualities variety should be positively correlated with quality, and, second, that for the higher quality levels the number of varieties should either increase more than exponentially with quality or increase and then decrease with quality. In the model, a second degree price-discriminating monopolist offers a menu of products of increasing quality. Consumers, modelled as caring about quality, but indifferent between flavors of the same quality, respond by randomizing across products that offer the same utility. The monopolist increases the profitability of price discrimination by offering a greater number of flavors at higher qualities. This increases both the likelihood consumers choose high-price products when randomizing and the expected profits. However, if at high quality levels the markets become sufficiently thin, the profit maximizing number of varieties will fall again. In this case, the distribution of varieties follows a hump shape.

These predictions are tested using unusually detailed data from the Australian car market. We find that the number of varieties increases over a substantial range of prices at the market level and at the make level. However, we do not find that the rate of increase is greater than exponential.

Instead we uncover a previously unidentified empirical regularity that the distribution of varieties offered by firms is hump-shaped. This is the case across a wide range of car makers.

One might object that most consumer product markets, including cars, are oligopolistic rather than monopolistic. However, the model can be generalized to this case along the lines of Champsaur and Rochet (1989). It can be shown that under some reasonable conditions there exists an equilibrium, where each producer specializes on a particular range of qualities. Qualitatively, the outcome is similar to the monopoly outcome. Proliferation of varieties in the oligopolistic case will be due to two effects: competition between the producers and price discrimination between consumers of different types who buy from the same producer. We find it interesting from the theoretical perspective that the second effect alone can lead to strong excessive flavor proliferation even if there is no preference for flavor and the degree of irrationality is small.

This paper makes several contributions to the literature. First, the model provides a new justification for observing different types of correlations between quality and variety. These correlations result from firms attempting to price discriminate rather than, as assumed earlier, solely from the distribution of preferences or costs — hypotheses that are notoriously difficult to test. Second, this paper also contributes to the literature on the modelling of new goods. Recent work, such as Berry and Pakes (2007), has pointed out the limitations of the standard random-utility model in handling new products. The model presented in this paper provides a very simple framework within which new products can be introduced. Third, the generality of the hump-shaped distribution of varieties offered within the car market suggests similar distributions of product varieties may occur in other markets. It also might help to distinguish between different sources of probabilistic choice (e.g. unobserved heterogeneity, bounded rationality) but more work is needed before attempting such differentiation in data.

The paper is organized in the following way. In Section 2 we provide some preliminary evidence to demonstrate a link between quality and variety for the car market. In Section 3 we introduce the simplest possible probabilistic choice model of consumer behavior with which to analyze the effect of increasing product variety on the profitability of price discrimination – the Luce model (Luce, 1959) of bounded rationality. The section closes with introducing the concept of a *nearly rational* consumer. In addition, we include Appendix 2 that reviews the similarities between this model, and other models of probabilistic choice, such as the random-utility model, and explains how a bounded rationality model provides a simpler framework to analyze the effects of introducing new varieties. In Section 4 we demonstrate that in a world populated by nearly rational consumers, the profitability of price discrimination will be lower compared with that predicted by the standard two-type screening model with fully rational consumers. In Section 5 we propose a way to overcome this problem by introducing multiple flavors for each quality level. In Section 6 we extend the model of Section 5 to multiple types and generate predictions about the relationship between quality and variety. Section 7 empirically analyzes these predictions using data from the Australian car market. Section 8 concludes.

2 Model proliferation in the car market

In this section, we provide some preliminary evidence on flavor proliferation from the car market. A comparison of automobiles classified into three groups, Small, Medium and Luxury, ascending in quality, demonstrates there are relatively more types of Luxury cars on offer, compared with Medium cars, despite sales of luxury cars being just one fifth those of medium and small cars. We then argue that this pattern is not obviously explained by differences in profitability in the different segments of the market.

Our data is composed of the price and characteristics for all cars sold as new in Australia in 1998, and registrations (sales) for cars, aggregated by model or make, for Australia in 1998.[3] Both datasets are compiled by a private data-collection firm, Glass's Guide.

Cars in the Glass's data set are first classified by make e.g. Ford or Toyota. Within each make there are models, e.g. (Ford) Falcon or (Toyota) Camry. Within each model there are variants. For the Ford Falcon the variant may have a name or just an initial, e.g. Futura, GL, Gli. The finest model classification code classifies some cars by series. Finally, within each of these groups, there may be multiple cars - each car with slightly

[3]Further details on the data are provided in Section 7.

different characteristics. Our dataset of prices and characteristics contains one observation for each distinct car offered for sale.

Though we will do a more refined analysis of this dataset in Section 7, for now we consider, in Table 1, averages for three groups of models of cars, as classified by Glass's Guide, representing three different levels of quality: Small, Medium and Luxury.

The first three columns of Table 1 summarize, for each group, the average number of features included with the car, average price, and total registrations (or sales). The first two columns suggest the ordering is consistent with ranking by quality. The third column demonstrates that the registrations for Luxury cars are just one fifth of those for Small and Medium cars.

The last four columns report the number of makes, models, variants and observations for each class. Note that the Luxury class features the largest number of makes, models, and variants. It also makes up 40% of the observations for new cars, despite making up less than 10% of registrations. The Small class also has the next largest number of makes, models and variants, but over five times the number of registrations for Luxury cars.

The second panel of Table 1 confirms what is suggested in the first panel. The first four columns report the number of registrations per Make, Model, Variant and Observation. This number is consistently lowest for the Luxury class. The last three columns report the number of Models, Variants and Observations per Make. The number of Models and Variants per Make is highest for Luxury cars, again consistent with our focus on a positive relationship between quality and variety.

A simple model of product variety is that multiple varieties are offered if it is profitable to do so. A simple way to think about profitability is that it is determined by the markup chargeable, the size of the market and the development costs. Hence, one reason we observe more Luxury cars than Medium cars is that it is more profitable to offer more types of cars. The sort of argument sometimes presented to us to justify this is that purchasers of Luxury cars have more distinctive preferences and can support more varieties. However, it seems unlikely, given the number of medium cars sold, that this explains why there are so many varieties of luxury cars compared with medium cars. With a market five times the size of that for luxury cars, there would seem to be scope for consumers in this market to also have distinct preferences and more than sufficient demand to support at least as many varieties rather than half as many.

The limited estimates of markups on cars, from Berry et al. (1995), are also not consistent with the simple profitability story. Although the markup on the most expensive model they report, the BMW 735i, is much greater than the markups on other cars, the BMW 735i also has lower estimated variable profits than most of the Medium and Small cars. However, neither the profitability story, based on the results of Berry et al. (1995) or our proposed model explains the proliferation of models for small cars. We will see in the empirical analysis in Section 7 that this is a product of aggregation.

3 Consumer Behavior

In this section we develop a probabilistic choice model of consumer behavior to analyze the implications of introducing new varieties of a differentiated product. The seminal book of Anderson, de Palma and Thisse (1992) summarizes four types of models that lead to probabilistic choice. Drawing on their work, we argue that under reasonable conditions, one can move freely between these different interpretations of probabilistic choice models for a fixed set of alternatives (see Appendix 2). However, these models do differ in the simplicity of modelling the change in probabilities if new alternatives are introduced. The main property that drives our results is that adding new varieties leads to a *strict* decrease in the probabilities of choosing all previously available varieties. As we are agnostic about the source of probabilistic choice, for expositional simplicity we use the probabilistic choice model that provides the simplest framework which incorporates this feature. Note that this property is only a little stronger than the one that necessarily arises in random-utility models which require that the probabilities of existing varieties are non-increasing if new varieties are added.

The modelling is considerably simplified if choices between varieties depend solely on the utilities associated with each variety. As we argue in Appendix 2, while the random-utility and address models of product differentiation (with suitable restrictions) could be used to analyse our problem, this would significantly complicate the analysis as the choice probabilities in these models depend on the distribution of random shocks as well as utility. Instead, we work with the Luce model of bounded rationality which has the property we require and is particularly simple to work with. In the Luce model adding new varieties reduces probabilities of choosing existing varieties, and the probabilities depend solely on the utilities from the different alternatives. In the second subsection we introduce a concept of near-rationality that specifies the very small degree of bounded rationality required to generate our results.

3.1 The Luce model

In this subsection we present a simple probabilistic choice model, Luce's (1959) logit model, that incorporates the main property required for our results. As we argue in Appendix 2, this model can be derived from a random utility model, or a corresponding Markovian learning model, or using Machina's approach (see, Rockafellar (1970), Anderson, de Palma, and Thisse (1992), Fudenberg and Levine (1998)). We stress that the details of the probabilistic choice model used are not important for our results. The property required for our main theoretical results is that adding a new alternative strictly decreases the choice probabilities of all previously accessible alternatives. The Luce model incorporates this feature in a particularly clear way. Furthermore, it has the simplifying feature that the choice probabilities depend solely on the utilities of the different alternatives. The choice probabilities have the following form:

$$p_i = \frac{\exp(u_i/\lambda)}{\sum_{j=1}^{n} \exp(u_j/\lambda)}. \tag{1}$$

Note that according to the Luce model any two alternatives that have the same utility are selected with the same probability. Assuming identical probabilities makes the subsequent analysis relatively simple. Relaxing this assumption only changes the results quantitatively, not qualitatively. In this model parameter λ, which can take values from zero to infinity, can be thought of as representing the degree of irrationality of the economic agent. If $\lambda = 0$ then

$$p_i = \begin{cases} 1/k, & \text{if } u_i = \max\{u_1, ...u_n\} \\ 0, & \text{otherwise} \end{cases}, \tag{2}$$

where integer k is the cardinality of the set of the utility maximizers. Note that this is the first sense in which consumers randomize, even if fully rational, and that the probability of each choice falls as the number of optimal varieties increases. It is also important to note that this is the main feature of the Luce model that we use for our results. Other features of the model, which have been recently criticised, are not essential for our results.

If λ is greater than zero, then the probability of the consumer choosing a product other than their optimal product is positive. This is the second sense in which the consumers randomize across products. For $\lambda < \infty$ the probability of choosing a product increases with the utility provided. At the extreme case where $\lambda \to \infty$ the choice probabilities converge to $1/n$, i.e. the choice becomes totally random and independent of the utility level. Again, the probability of choosing each product falls as the number of varieties increases. Note that the logit formulation is not required for λ to play this role; for example, a probit formulation is also feasible.

3.2 A concept of a nearly rational consumer

In this subsection we introduce the concept of a nearly rational consumer — which summarizes the relatively minor degree of irrationality (or small value of λ) required to yield our results. Assume that the probability of different choices by the consumer is given by a continuously differentiable function:

$$p(\cdot) : R^n \times R_+ \to \Delta^n, \tag{3}$$

such that $u_i = u_j$ implies $p_i(u, \lambda) = p_j(u, \lambda)$ and $p(u, 0)$ is given by (2). Luce probabilities (1) satisfy these properties though the exact form of function $p(\cdot, \lambda)$ is not important for our purposes.

Let M be the set of the utility maximizers, i.e.

$$M = \{u_i : u_i = \max\{u_1, ...u_n\}\}. \tag{4}$$

Take any $u_j \in M$ and define

$$\Delta = \min_{u_k \notin M} (u_j - u_k). \tag{5}$$

Definition 1 *An economic agent whose choice probabilities are given by (3) is called nearly rational if* $\lambda << \Delta$.

In words, the definition says that an economic agent is nearly rational if her irrationality parameter λ is *much smaller* (sign $<<$ reads "much smaller") than the difference in utility between the optimal and the next to the optimal choice. The exact meaning of "much smaller" depends on the precision with which an econometrician wants to measure relative frequencies of different choices.

In the standard model of price discrimination with full rationality infinitely small price changes can be used to get each type of consumer to purchase the product designed for them. Introducing $\lambda > 0$ mutes the effect of price changes on the probability of choice, but we require only a small degree of bounded rationality to obtain our results.

4 The monopolistic screening model

Let us briefly review the basic screening model under full rationality[4] and discuss the consequences of offering the second best contract to the nearly rational consumers. Assume a risk neutral monopolist produces a unit of good with quality x at a cost $C(x)$, where $C(\cdot)$ is a strictly convex, twice differentiable function. Preferences of a consumer, of type θ, over a unit of good with quality x are given by a twice continuously differentiable utility function $u(\theta, x)$. Preferences of the consumer are quasilinear in money:

$$v(\theta, x, m) = u(\theta, x) + m.$$

[4]See Mas-Colell, Whinston and Green (1995), Section 13D, for details.

Each consumer wants to buy at most one unit of the monopolist's goods. Type θ is private information of the consumer. If the consumer does not purchase a good from the monopolist, she receives utility $u_0(\theta)$. For simplicity assume it does not depend on type and normalize it to be zero. Finally, assume

$$u_1 > 0,\ u_2 > 0,\ u_{12} > 0.$$

(Here u_i is the derivative of u w. r. t. the i^{th} argument, u_{12} is the cross partial derivative with respect to θ and x). The last of these conditions is known as *the Spence-Mirrlees condition or the single-crossing property.*

Let us assume that $\theta \in \{\theta_L, \theta_H\}$. Then the monopolist solves:

$$\max p_H(t_H - C(x_H)) + (1 - p_H)(t_L - C(x_L)),$$

where $p_H = \Pr(\theta = \theta_H)$, subject to the following constraints:

$$u(x_L, \theta_L) - t_L \geq 0 \tag{6}$$

$$u(x_L, \theta_L) - t_L \geq u(x_H, \theta_L) - t_H \tag{7}$$

$$u(x_H, \theta_H) - t_H \geq 0 \tag{8}$$

$$u(x_H, \theta_H) - t_H \geq u(x_L, \theta_H) - t_L. \tag{9}$$

Constraints (6) and (8) state that both types would like to participate in the contract and are known as the *individual rationality* constraints, and the constraints (7) and (9), known as the *incentive compatibility* constraints, ensure that no one would like to choose the contract meant for the other type. The basic result is Stole's *constraint reduction theorem* (Stole, 2000) that states that at the optimal allocation only two of these constraints bind: (6) and (9): that is the lowest type gets her reservation utility (in this case, zero) and the high type gets the *information rent*, that is just enough to prevent her from pretending to be the low type. This implies that

$$
\begin{aligned}
t_L &= u(x_L, \theta_L) && (10) \\
t_H &= u(x_H, \theta_H) - u(x_L, \theta_H) + u(x_L, \theta_L). && (11)
\end{aligned}
$$

Therefore, the monopolist's solves

$$\max p_H(u(x_H, \theta_H) - u(x_L, \theta_H) + u(x_L, \theta_L) - C(x_H)) + (1 - p_H)(u(x_L, \theta_L) - C(x_L)).$$

The first order conditions are

$$
\begin{aligned}
u_1(x_H, \theta_H) &= C'(x_H) && (12) \\
u_1(x_L, \theta_L) - C'(x_L) &= \frac{p_H}{1 - p_H}(u_1(x_L, \theta_H) - u_1(x_L, \theta_L)) > 0. && (13)
\end{aligned}
$$

Note that x_H is at the efficient level (no distortions at the top) and x_L is below the efficient level. At these prices the high value consumers are indifferent between the two products, and the low value consumers are indifferent between purchasing the low quality product and not purchasing at all.

To place a bound on the degree of irrationality required, according to equation (5), we consider the gaps in utility between the most preferred and next preferred products for each type of consumer. For type θ_H, the gap in utility between its preferred and next preferred option, equal to their information rent, is:

$$I_{21} = u(x_H, \theta_H) - t_H = u(x_L, \theta_H) - t_L = u(x_L, \theta_H) - u(x_L, \theta_L).$$

For type L, who is indifferent between purchasing x_L and not purchasing, the gap in utility to its next preferred option is:

$$\Delta_{IC} = t_H - u(x_H, \theta_L), \tag{14}$$

which also measures the slack in the incentive compatibility condition for the low type. Hence we can determine

$$\Delta = \min(I_{21}, \Delta_{IC}). \tag{15}$$

Let us call the contract (10)-(13) the *second best contract*. Assume that the monopolist offers the second best contract to the nearly rational consumers. By the definition of near rationality, the irrationality parameter λ is much less than both the high type information rent and the slack in the low type incentive compatibility constraint. Therefore, the fraction of high type consumers who decide not to participate or the fraction of the low type consumers who decide to choose the high quality product is negligibly small.

On the other hand, as a result of randomizing between equally preferred alternatives, similar to that described in equation (2), approximately half of the low type consumers decide to stay out of the market and approximately half of the high type consumers purchase the low quality product. This leads to a drop in the monopolist's profits which is higher order of magnitude than λ. If consumers were fully rational, then infinitely small price changes deal with this problem but, as we have argued, infinitely small price changes will not be sufficient if consumers are near-rational. Instead the monopolist must alter prices to violate the binding constraints by some finite amount.[5] This reduces the profits earned from both high and low types.

The alternative to a significant price cut is for the monopolist to create multiple flavors for each quality level. At each quality level, flavors differ only in ways that do not affect the utility from consuming the product. In other words, although flavors are products that are horizontally differentiated, consumers are indifferent between differentiated products of the same quality. Only vertical differences matter.[6] To see how the flavors are going to help, assume the monopolist sells m flavors of the high quality product and one flavor of the low quality product. Now high type consumers are faced with $(m+1)$ choices, each of which provides them with the same utility. The probability that the high type consumer purchases the high quality product is $m/(m+1)$. If the marginal cost of providing a new flavor is sufficiently low this way of ensuring participation may be preferable to leaving extra rents to the consumers. From the social point of view this is, however, a complete waste since consumers do not have preferences for flavor. In the next section we are going to investigate the idea of flavor proliferation in more detail.

[5] Basov(2008) shows that the required change in tariffs in an optimal contract is of the order of $\lambda/ln\lambda$, which corresponds to the probability of making a mistake, λ.

[6] This assumption has some empirical support. Indeed, Draganska and Jain (2005a) report that preferences for quality are much stronger than those for flavor.

5 The flavor proliferation model

In this section we analyze how flavor proliferation increases the profitability of price discrimination by overcoming the problem of consumers randomizing away from the most profitable product for their type. We demonstrate, for the two type case, that the number of flavors increases with the quality of the product.[7] Indeed, the effect is rather strong. If the cost of adding a new flavor converges to zero the ratio of the numbers of flavors for adjacent quality levels converges to infinity. Therefore, it is likely that this effect will override any supplementary pattern, which may emerge from the direct preferences for flavor.

If the monopolist offers n flavors of low quality and m flavors of high quality, the low quality good will be purchased by a fraction q_L of the consumers, where

$$q_L = (1 - p_H)\frac{n}{n+1} + p_H \frac{n}{n+m}, \tag{16}$$

while the high quality good will be purchased by a fraction q_H of the consumers, where

$$q_H = p_H \frac{m}{n+m}. \tag{17}$$

Before proceeding we note three assumptions. First, we assume that the marginal cost of adding a new flavor is $c > 0$ and does not depend on the quality. Second, we assume the utility function for each type is the prod-

[7] In principle, the monopolist can alter qualities, prices, and the number of flavors. However, under our parametric restrictions (to be specified below), the optimal prices and qualities are close enough to those optimal for rational consumers to neglect the differences.

uct of the quality of the good consumed and the type, θ. Finally, under conditions specified at equation (30), we can use the optimal qualities for rational consumers as an approximation for those optimal for nearly rational consumers.[8] Therefore, the monopolist solves:

$$\max_{m,n}((t_H - C(x_H))q_H + (t_L - C(x_L))q_L - c(n+m)). \tag{18}$$

Let us introduce the following notation:

$$\pi_L = t_L - C(x_L),\ \pi_H = t_H - C(x_H),\ p_L = 1 - p_H \tag{19}$$

i.e. π_i are the profits per consumer the monopolist can potentially earn on type i if all consumers of this type select the contract designed for them. Note that $\pi_H - \pi_L > 0$. Indeed,

$$\pi_H - \pi_L = t_H - t_L - C(x_H) + C(x_L). \tag{20}$$

Using expressions for the tariffs, based on our assumed utility function:

$$\pi_H - \pi_L = \theta_H(x_H - x_L) - C(x_H) + C(x_L) + \theta_L x_L. \tag{21}$$

Finally, since the optimal quality is efficient for the top type, $C'(x_H) = \theta_H$ and strict convexity of the cost implies:

$$\pi_H - \pi_L = C'(x_H)(x_H - x_L) - C(x_H) + C(x_L) + \theta_L x_L > 0. \tag{22}$$

[8] Basov (2008) demonstrates the difference in the optimal qualities for the two cases is at most of order $(\frac{\lambda}{ln\lambda})^2$ which, under these conditions, is very small. The differences in optimal tariffs in the two cases is O$(\frac{\lambda}{ln\lambda})$.

The monopolist's problem can be rewritten as:

$$\max_{m,n}(p_L\pi_L\frac{n}{n+1}+p_H\pi_H\frac{m}{n+m}+p_H\pi_L\frac{n}{n+m}-c(n+m)). \qquad (23)$$

Ignoring the constraint that m and n should be integers, one can write the first order conditions:

$$\begin{cases} \frac{p_L\pi_L}{(n+1)^2}-\frac{p_H(\pi_H-\pi_L)m}{(n+m)^2}=c \\ \frac{p_H(\pi_H-\pi_L)n}{(n+m)^2}=c \end{cases}. \qquad (24)$$

It is easy to observe that for small values of c

$$n = \frac{p_L^{2/3}\pi_L^{2/3}}{p_H^{1/3}(\pi_H-\pi_L)^{1/3}c^{1/3}}+O(c^{1/3}) \qquad (25)$$

$$m = \frac{p_L^{1/3}p_H^{1/3}\pi_L^{1/3}(\pi_H-\pi_L)^{1/3}}{c^{2/3}}+O(c^{1/3}). \qquad (26)$$

As $c\to 0$ both n and m go to infinity, but in a such way that

$$\frac{m}{n^2}\to\frac{p_H(\pi_H-\pi_L)}{p_L\pi_L}. \qquad (27)$$

As long as the profits from the high quality product are high enough, and the probability of the consumer being a high type is not too low, $\frac{m}{n}$ increases proportionally with n. Finally, flavor proliferation costs the monopolist:

$$F=c^{1/3}(\pi_H-\pi_L)^{1/3}\pi_L^{1/3}p_H^{1/3}p_L^{1/3}. \qquad (28)$$

Basov (2008) demonstrates that a monopolist who faces nearly rational consumers, as defined in Definition 1, and is restricted to offering a binary menu (i.e. no flavor proliferation is possible) will find it approximately optimal to leave qualities at the same level as for the rational consumer and to alter prices.[9] However, if flavor proliferation is allowed and

$$F << \lambda \tag{29}$$

holds, then it is even more profitable to proliferate flavors than to alter prices. Hence, under the following conditions, flavor proliferation, at the optimal qualities and prices derived earlier is approximately optimal:

$$F << \lambda << \Delta, \tag{30}$$

where Δ is defined by (15).

Several observations are due concerning these results. First, if c is sufficiently small both m and n are large, moreover, m/n is large. Therefore, the consumers will choose the options designed to them with probabilities close to one, as predicted by the screening model with fully rational consumers. Second, as one can see from equation (27), the number of flavors increase with quality at an *increasing rate*, i.e. the increase is faster than exponential. If one assumes that flavors are directly valued, this may offset some of this effect, but the wasteful proliferation effect is likely to determine the overall correlation as quality increases.

[9]Under the condition specified in equation (30), the difference between the optimal solution if consumers are rational and the optimal solution if consumers are near-rational is so small it can be ignored. Hence the optimal solution for rational consumers can be used without materially altering the results.

6 An extension of the model: multiple types.

In the previous two sections we discussed the monopolistic screening model with two types of consumers. We argued that if the consumers are not fully rational, it might be in the interest of the monopolist to introduce multiple flavors of a good of a given quality, even if the consumers care only about quality. In this section we extend the model to the case where there are more than two types of consumers. In the first subsection we demonstrate that the number of varieties increases faster than exponentially if per-consumer profits increase sufficiently fast with the quality level and markets for high quality varieties are sufficiently thick. In the second subsection, we allow for thinner markets at higher prices, and demonstrate that the relationship between quality and variety can take the form of a hump, with a maximum reached at a certain number.

6.1 Variety Increasing with Quality

The assumptions on the fundamentals are the same as in Section 4, but now the consumer's type is given by $\theta \in \{\theta_1, ..., \theta_N\}$. Let $p_i = \Pr(\theta = \theta_i)$. Otherwise the analysis in this subsection proceeds in the same way as in Section 5. We assume $p_i > 0$ for all i and

$$\sum_{i=1}^{n} p_i = 1. \tag{31}$$

Denote by (x_i, t_i) the quality and tariff the monopolist would offer to type θ_i under the hypothesis of the full rationality. For $i = 1, ..., N$ define π_i by:

$$\pi_i = t_i - C(x_i) \tag{32}$$

and let $\pi_0 = 0$ be the profit the monopolist earns from the consumers who choose the outside option. Finally, define $\Delta\pi_i = \pi_i - \pi_{i-1}$. The first observation is that $\Delta\pi_i > 0$ for all i. Indeed,

$$\Delta\pi_i = t_i - t_{i-1} - C(x_i) + C(x_{i-1}). \tag{33}$$

The constraint reduction theorem (Stole, 2000) implies that the only binding constraints in the problem are the individual rationality for the lowest type and the incentive compatibility between types θ_i and θ_{i-1}. Hence we can exclude the tariffs to get:

$$\Delta\pi_i = \theta_i(x_i - x_{i-1}) - C(x_i) + C(x_{i-1}) + \theta_{i-1}x_{i-1}. \tag{34}$$

Finally, since the optimal quality is efficient for the top type and biased downward for the rest, $C'(x_i) \leq \theta_i$. Strict convexity of the cost implies:

$$\Delta\pi_i = C'(x_i)(x_i - x_{i-1}) - C(x_i) + C(x_{i-1}) + \theta_{i-1}x_{i-1} > 0. \tag{35}$$

Using the same approximation as in the two types case, the monopolist's problem can be written as:

$$\max_{\{n_i\}_{i=1}^N} (\sum_{i=1}^N (p_i\pi_i \frac{n_i}{n_i + n_{i-1}} + p_i\pi_{i-1}\frac{n_{i-1}}{n_i + n_{i-1}} - cn_i), \tag{36}$$

where we defined $n_0 = 1$. Ignoring the constraint that m and n should be integers, one can write the first order conditions:

$$\begin{cases} \frac{p_i \pi_i}{(n_i+n_{i-1})^2} - \frac{p_{i+1}\Delta\pi_i n_{i+1}}{(n_i+n_{i+1})^2} = c \\ \frac{p_i \Delta\pi_i n_i}{(n_i+n_{i+1})^2} = c \end{cases} . \tag{37}$$

As $c \to 0$ all n_i go to infinity, but in a such way that

$$\frac{n_{i+1}}{n_i^2} = \frac{p_{i+1}\Delta\pi_{i+1}}{p_i \Delta\pi_i}. \tag{38}$$

As n_{i+1}/n_i is proportional to n_i it is clear that as long as $p_{i+1}\Delta\pi_{i+1}/p_i\Delta\pi_i$ is greater than one that the number of varieties increases faster than exponentially. The value of n_1 is given by:

$$n_1 = \frac{(p_1\pi_1)^{2/(2N-1)}}{(p_2\Delta\pi_1 c)^{1/(2N-1)}} + O(c^{1/(2N-1)}) \tag{39}$$

and n_i for $i > 1$ can be calculated using (38). Finally, the flavor proliferation costs of the monopolist are:

$$F_N = c^{(N-1)/(2N-1)} \prod_{i-1}^{N} (\Delta\pi_i p_i)^{1/(2N-1)} \tag{40}$$

Once again, flavor proliferation is approximately optimal as long as:

$$F_N << \lambda << \Delta, \tag{41}$$

where Δ is defined by the maximal “slack” for the non-binding constraints (in the case of two types given by equation (15)).

6.2 Hump-shaped Relationship between Quality and Variety

In the previous section, we assumed that the expected profitability from each type increased with the type, i.e. $p_{i+1}\Delta\pi_{i+1}/p_i\Delta\pi_i > 1$. If we instead assume that

$$\frac{p_{i+1}\Delta\pi_{i+1}}{p_i\Delta\pi_i} < 1 \tag{42}$$

we demonstrate the relationship between the number of varieties and quality will follow a hump shape, increasing and then decreasing, rather than more than exponentially increasing over the whole range.

There are several reasons why we might wish to make an assumption like that for equation (42) which, for convenience, we will discuss in terms of our empirical application, cars. The most obvious cause could be as the price rises the number of consumers willing to pay this price falls sufficiently for expected profit to fall or that consumers with strong preferences for quality (or snobs) are sufficiently rare.

A possible reason for this may be concerns by consumers about quality. A company that makes relatively low quality cars can attempt to produce a high quality car, but consumers are unlikely to immediately accept the claimed quality and require a lower price, compared with the cars sold by established high-quality producers. This makes producing varieties outside of the range of quality currently accepted by consumers less profitable. This argument seems stronger, though, in explaining why higher than average quality cars, for a given maker, are less profitable, than lower than average quality cars.

An alternative explanation on the cost side is that if the firm produces a quality level less than or greater than the quality level at which the number of brands is greatest, production costs may be higher. Assume each carmaker can produce a particular quality of cars very cheaply but that there is a U-shaped average cost curve (as a function of quality). Hence, if the firm produces either higher or lower quality cars average costs are higher relative to the price that can be obtained for the car and profits are relatively small from producing them. Hence, the firm offers less brands at these quality levels. This is difficult to test without considerable data. However, the frequent rebadging of, particularly smaller, cars for sales in different markets suggests that production costs are unlikely to be the whole story. It may be too expensive for Mercedes to make a cheap small model in the same factories that make Mercedes, but they could license Daewoo to do so at the Daewoo factories.

To formally analyze the implications of equation (42) we introduce the following notation:

$$\begin{aligned} \xi_i &= \ln n_i & (43) \\ \beta_i &= \frac{1}{2} \ln \frac{p_i \Delta\pi_i}{p_{i+1}\Delta\pi_{i+1}} > 0 & (44) \\ \beta(x) &= \beta_i \text{ for } x \in (i-1, i]. & (45) \end{aligned}$$

Using this notation, equation (38) takes the form:

$$\xi_{i+1} - 2\xi_i = \beta_i \tag{46}$$

Then the solution to equation (38) can be written as:

$$\xi(x) = \xi_0 2^x - \int_0^x \beta(y) 2^{x-y} dy. \tag{47}$$

We also assume that $\xi_0 > \beta_0$ (this assumption means that the logarithm of the flavors at the lowest quality level is sufficiently big, which is consistent with our assumption of the small cost of flavor proliferation). To find maximum of $\xi(\cdot)$ w.r.t. x, note that

$$\xi'(x) = 2^x \ln 2(\xi_0 - \frac{\beta(x)2^{-x}}{\ln 2} - \int_0^y \beta(y)2^{-y} dy). \tag{48}$$

Therefore,

$$[\xi'(x) = 0] \Leftrightarrow [\xi_0 = \frac{\beta(n)2^{-n}}{\ln 2} + \sum_{i=0}^{n} 2^{-i}\beta_i]. \tag{49}$$

Since the right hand side of this expression increases in n there is at most one n for which this condition is satisfied. Hence, the relationship between quality and variety follows a hump shape rather than increasing over all qualities.

7 An empirical analysis of car varieties

In this section we analyze the models offered in the Australian car market to determine if there is support for the main prediction of the theoretical section: that the number of flavors increases with the quality level, at least for low qualities. We conclude that for a wide range of makes that the common pattern of model offerings is a hump shape, with a rapid increase in types followed by a similarly sharp decline past the mode.

The data we use for this analysis is the same data set, originally collected and compiled by the private data-collection firm Glass's Guide, used in Prentice and Yin (2005). The first component of the data set contains the prices and characteristics for all cars sold as new in Australia in 1998. The second component of the data set is registrations, usually by model or make, for Australia in 1998. [10]

For reporting purposes we divide all makes of cars into five groups, based on their average price in the data. These groupings are reported in Appendix 1.[11] The Low group includes makes such as Daewoo and Hyundai. The Medium group includes makes such as Holden and Toyota. The UpperMedium group includes makes such as Honda and Audi. The Prestige group includes makes such as BMW and Mercedes Benz and the High group includes makes such as Jaguar and Lamborghini. Table 2 and Graph 1 provide summary information about each group. Table 2 reports the descriptive statistics and Graph 1 includes two Box plots of prices by group.

All groups except, perhaps, Low, feature a positively skewed price distribution. The high group has a particularly highly skewed distribution though, as Table 2 demonstrates, this is over very few observations. Indeed the

[10] For higher priced cars, though, registrations may only be available for groups of models. For example, Mercedes Benz has six different models in the C-Class in 1998, whereas the registrations data is available only for C-Class, hence aggregating over the six models. Similarly, the Glass price data distinguishes between the Commodore and two higher quality versions of the Commodore, the Berlina and Calais, whereas in the registrations data, the Commodore, Berlina and Calais are aggregated as Commodores.

[11] The classification of makes may be different in other countries nearly all models offered in Australia are imports. Some firms, offering a complete range in their home market, may only export premium or smaller cars to Australia.

cheaper makes, Porsche and Jaguar, make up nearly half of the observations and 93% of the registrations for this class.

We now analyse the patterns of cars offered more closely. Based on our preliminary analysis in section 2 and following a large literature on hedonic pricing, we assume that price is correlated with quality. We then estimate kernel density functions of prices.

7.1 Market Level

In this section we compare the distribution of prices for all cars offered for sale in the market with an estimate of the distribution of cars sold in the market. The latter distribution is estimated by weighting each price by an estimate of the number of registrations for that car. Both distributions are estimated using a kernel density estimator. We compare the two distributions in three steps. First, we check if, and at what prices, the number of cars offered increases with price. Second, if the number of cars does increase with price, we check if the increase in the number of cars is at a greater rate than that of the weighted distribution. This would reject a simple model of there being a constant cost to adding a model. Third, we check if, and at what prices, there is greater weight in the distribution of cars compared with the distribution of sales. Graphically, we check if the distribution of cars is to the right of the distribution of sales.

Before discussing the results of this comparison, we first describe how we construct the weights. As we do not have registrations data matching each price observation we construct weights as follows. First, we match all price data to the model categories in the registration data. Then we divide the registrations for the model by the number of observations. In effect, we assume that each observation for a model sells an equal number. This likely overstates the number of registrations for higher price variants within each models.[12] This will bias this examination against finding the differences in the distributions suggested above. Finally, note that we estimate the kernel density at 50 points for the distribution of prices, and then use the same fifty points to estimate the distribution for sales.

Graph 2a confirms the positive skewness of the distribution of prices, and demonstrates a similar skewness in the distribution of registrations. Because this skewness may cause problems for the kernel density estimator, we re-estimate using observations with prices below \$200,000. First, at low prices, both the distribution of prices and distribution of registrations increases as price goes up. However, it is not obvious that the distribution of prices rises at a faster rate than that for sales. Indeed, it is more likely the distribution of sales rises at a faster rate. However, at prices above \$50,000 it is clear that there is more mass in the distribution of prices than the distribution of

[12]The extent of this is indicated by examining the ratio of the range between the maximum and minimum prices for each model. The median range is 6862 and the median range relative to mean is 0.16 which do not seem large. However about twenty-five percent of observations have a range over mean of 0.34 or greater. Looking at the top 10 percent of observations, though they are drawn from various price levels, there are probably more from the UpperMedium and Prestige cars. If we look at the range over mean for the observations with larger weights, most ranges are fairly small, though the Holden Commodore has a substantial price range from 18 to 49 thousand dollars (and the registrations data appears to include observations for the more pricey Berlina and Calais models).

sales. This confirms the impressions suggested by our analysis of Table 1. At prices above $50,000 there are more types sold relative to registrations. While there is also greater mass in the distribution for the very cheapest cars, it is relatively small and suggests some of the puzzling features of Table 1 resulted from aggregation. Hence, at the market level, there is some support for our hypothesis, particularly among the more expensive cars.

7.2 Make Level

The next step is to examine if there is support for the hypotheses at the Make level. We use the data on the price by observation for each make available in Australia. We then estimate a kernel density function for each make and examine the distribution of prices. Note we do not use the registrations data to weight these estimates. Because we need sufficient observations to construct the density functions, we consider only makes that have more than 30 observations. In reference to the classes used to summarize the Makes in Table 2 and Graph 1 the excluded makes (indicated with a * in Appendix 1) tend to have much lower registrations in their class. This criterion leaves us with 17 makes, with all Makes in the High Class being excluded.

Each kernel density function follows the same general pattern as for the market as a whole. Nearly all distributions are positively skewed. The number of observations rises very rapidly from the lowest level followed by a slower decline. Graphs 3A and 3B demonstrate this takes place for all makes - from the relatively cheap Daewoo to the highest price Mercedes Benz. This suggests that the range of offerings is not solely a function of price, but a common strategy across car-makers.

One aspect of the importance of this phenomenon is how many sales this pattern applies to. So, using the weights described earlier we estimate the number of registrations for cars to the left of the modal point of each distribution. The results, by the class described earlier, are presented in Table 3. For just over 50% of the Makes the share covered is greater than 50%. For the Makes in the Low group, the share of registrations covered is typically quite high ranging from 0.59 to 0.63. The range is greater for Medium but note the maximum and minimum cases are quite distinct from the mass of observations which fall between 0.36 and 0.65. Similarly for UpperMedium, the two middle observations, between the maximum and minimum, are 0.37 and 0.43. For Prestige there is a wide range between 0.28 and 0.56.

The last test we perform is to determine whether the second prediction of the model holds i.e. that the number of cars will increase greater than exponentially. To carry out this test we extract the densities at 50 points and take the natural logarithm of them. Each point is assigned a rank. We then estimate the derivatives of the log density with respect to rank. If the second hypothesis holds, we should see these derivatives increasing. As it turns out a formal hypothesis test is unnecessary. We report two representative graphs (Graphs 4A and 4B) for Mercedes Benz and Daewoo. The derivatives fall over almost the whole range.

7.3 Robustness

In this subsection, we examine the sensitivity of our results to choice of the bandwidth parameter in the kernel density estimation - as this is cited to be the main concern for this type of estimation (Silverman, 1986). To examine if the hump shape is consequence of over-smoothing, we approximately halve the bandwidth used and perform the same steps for each of the makes.

When we examine the estimated kernel densities about half of them have not changed very much. Of the remaining nine, the most common changes is that additional peaks emerge to the right of the largest peak or small steps emerge on either side. In general the modal price moves to the left. In only two cases are there changes that run against the main conclusions we draw from the standard analysis. The summary statistics are presented in Table Four. While the mean and medians have not changed very much, the results are probably a little more dispersed.

So, to summarize, the main results for the makes do not appear to change substantially with the bandwidth.

8 Conclusions

In this paper we have analyzed theoretically and empirically the links between quality and variety. The model has two components. First, consumers are modelled as caring about quality but indifferent between varieties of the same quality. We use a model of near-rational consumers which is required for flavor proliferation to be more profitable than price changes. For expositional simplicity, we use the Luce model of bounded rationality as this enables introducing new goods more easily than some other models, e.g. random-utility models. Second, we assume a monopolist who engages in second-degree price discrimination. When a price-discriminating monopolist offers products at prices that yield identical utility, consumers randomize across the different products. The profitability of price discrimination is increased by the monopolist offering more varieties at higher qualities as the likelihood that consumers choose high price products increases. If, at high quality levels, the markets become sufficiently thin, though, the profit maximizing number of varieties will fall again, yielding a hump-shaped relationship between variety and quality. We then examine these predictions using data from the Australian car industry and find the number of varieties does increase over substantial ranges of prices at both the market and make level. However, the overall relationship, across a wide range of car makes, is hump-shaped — a previously unidentified empirical regularity.

Another important contribution of the paper is that the number of varieties offered by the market can be higher than the social optimum even in the case of a single monopoly, which is opposite to the conclusions drawn from the previous literature (see, Lancaster (1990)).

References

Anderson, S.P., dePalma, A. and Thisse, J.-F. Discrete choice theory of product differentiation, MIT Press, Cambridge, MA, (1992).

Anderson S. P., Goeree, J. K. and Holt, C. A. Noisy directional learning and the logit equilibrium, Scandinavian Journal of Economics, 106(3), 581-602, (2004).

Ansari, A., Economides, N. and Steckel, J. The max-min-min principal of product differentiation, Journal of Regional Science, 38, 207-230, (1998).

Basov, S. Incentives for boundedly rational agents, Topics in Theoretical Economics, 3, Article 2, 1-14, (2003).

Basov, S. The mechanics of choice, In: Progress in Economics Research, v.10, A. Tavidze (ed.), 55-64, Nova Science Publishers, NY, (2006).

Basov, S. Monopolistic screening with boundedly rational consumers, Australasian Economic Theory Workshop, Bond University, (2008).

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

Berry, S. and Pakes, A. The Pure Characteristics Demand Model, International Economic Review, 48(4), 1193 - 1225, (2007).

Bush, R. and Mosteller, F. Stochastic models for learning. Wiley, New York, (1955).

Champsaur, P. and Rochet, J.-C. Multiproduct duopolists, Econometrica, 57(3), 533-557, (1989).

Draganska, M. and Jain, D. C. Product-Line length as a competitive tool, Journal of Economics and Management Strategy, 14, 1-28, (2005).

Draganska, M. and Jain, D. C. Consumer Preferences and Product-Line Pricing Strategies: An Empirical Analysis, Marketing Science, 25, 164-174, (2005a).

Falmange, J.-C. A Representation Theorem for Finite Random Scale Systems, Journal of Mathematical Psychology, 18, 52-72, (1978).

Foster, D. and Young, P. Stochastic evolutionary game theory. Theoretical Population Biology, 38, 219-232, (1990).

Friedman, D. The evolutionary game model of financial markets, Quantitative Finance, 1, 177-185, (2000).

Friedman, D. and Yellin, J. Evolving landscapes for population games, University of California Santa Cruz, mimeo, (1997).

Fudenberg, D. and Harris, C. Evolutionary dynamics with aggregate shocks. Journal of Economic Theory, 57, 420-41, (1992).

Fudenberg, D. and Levine, D. K., The theory of learning in games, Cambridge, MA, MIT Press, (1998).

Glass's Guide. Black and White Data Book, Glass's Guide, Melbourne, (1998).

Hofbauer, J. and Sandholm, W. H. On the global convergence of stochastic fictitious play, Econometrica, 70, 2265-2294, (2002).

Kandori, M., Mailath, G. and Rob, R. Learning, mutation and long run equilibria in games. Econometrica, 61, 29-56, (1993).

Lancaster, K. The economics of product variety: A survey, Marketing Science, 9, 189-206, (1990).

Luce R. D. Individual choice behavior, Wiley, New York, (1959).

Machina, M.J. Stochastic choice functions generated from deterministic preferences over lotteries, The Economic Journal, 95, 575-594, (1985).

Mas-Colell, A., Whinston, M. D. and Green, J. R. Microeconomic theory, Oxford University Press, (1995).

Neven, D. and Thisse, J. On quality and variety in competition, in Gabszewicz, J., Richard, J. and Wolsley, J., (eds): Economic Decision-Making: Games, Econometrics, and Optimization, Elsevier, Amsterdam, (1990).

Prentice, D. and Yin, X. Constructing a quality-adjusted price index for a heterogeneous oligopoly, The Manchester School, 72, 423-442, (2005).

Rockafellar, T. R. Convex analysis, Princeton University Press, NJ, (1970).

Shugan, S.M. Product assortment in triopoly, Management Science, 35, 304-320, (1989).

Silverman, B. W. Density Estimation for Statistics and Data Analysis, Chapman & Hall, London, (1986).

Stole, L. Lectures on contracts and organizations, (2000), http://gsblas.uchicago.edu/papers/lectures.pdf.

Young, P. The evolution of conventions, Econometrica, 61, 57-84, (1993).

Appendix 1

List of Makes included in each Group. Makes with an asterisk have less than thirty observations and are not included in the analysis of models by make. None of the High group have more than thirty observations.

Low Daewoo, Daihatsu, Hyundai, Kia*, Proton*, Seat*, Suzuki*.

Medium Chrysler*, Citroen*, Ford, Holden, Mazda, Mitsubishi, Nissan, Peugeot*, Subaru, Toyota, Volkswagen

UpperMedium Alfa Romeo*, Audi, Honda, MG*, Saab, Volvo

Prestige BMW, Lexus*, Lotus*, Mercedes Benz

High* Aston Martin, Bentley, Ferrari, Jaguar, Lamborghini, Maserati, Porsche, Rolls Royce

Appendix 2: Models of Probabilistic Choice

In this Appendix we compare four models of probabilistic choice. This comparison demonstrates two points. First, under some reasonable assumptions, one can move freely between different interpretations of random choice models. Second, that the Markovian and Machina-type models provide an easier framework than the random-utility and address models, to analyse the effect of introducing new varieties. Hence in the body of the paper we use the Luce model of probabilistic choice as a particularly simple foundation for our model of consumer behavior. Though one could use a random-utility or address model, modelling would just be more complex.

Markovian learning models

Though these models originated in mathematical psychology in the work of Bush and Mosteller (1955), they have been widely used in economics (e.g. Foster and Young (1990), Fudenberg and Harris (1992), Kandori, Mailath, Rob (1993), Young (1993), Friedman and Yellin (1997), Anderson, Goeree, and Holt (2004), Friedman (2000), and Basov (2003)). Though most economic applications assume a continuous choice space, for simplicity of presentation and for consistency with our application, we assume a finite choice space.

Suppose an individual faces a choice among n different options. A boundedly rational individual is assumed to start with a random choice and adjust her choice over time in a way that appears beneficial given her current experience. From time to time the individual may also experiment. This kind of behavior usually leads to a Markov process over the choice space, which can be described as:

$$p_{t+1} = f(p_t, u), \tag{50}$$

where $p_\tau \in \Delta^n$ is the vector of choice probabilities at time τ, $u \in R^n$ is the vector of utilities associated with different choices, $f : \Delta^n \to \Delta^n$ is a continuous function and Δ^n is the $n-$ dimensional unit simplex. The steady states of equation (50) can be interpreted as long-run choice probabilities.[13] A simple form of this relationship occurs if the transition probabilities between the states of the system are constant:

$$f(p_t, u) = A(u)p_t, \tag{51}$$

where A is an $n \times n$ matrix.[14]

The steady state probabilities for equation (50) can be written as:

$$p^* = p^*(u) \tag{52}$$

Let us also impose the following symmetry condition: for any permutation $\delta : \{1, 2, ..., n\} \to \{1, 2, ..., n\}$ one obtains

$$q_{t+1} = f(q_t, v), \tag{53}$$

where

$$q_\tau^i = p_\tau^{\delta(i)}, v_\tau^i = u_\tau^{\delta(i)}, \tag{54}$$

[13]Note that a steady of system (50) exists according to the Brouwer fixed point theorem.
[14]Condition (51) is usually violated in social learning models. For examples of such models, see Basov (2006).

i.e. dynamics (50) remains invariant under the relabelling of the choices. If the symmetry condition holds system (50) should generically possess at least one symmetric steady state,[15] i.e. a steady state where the choices that have equal utilities will be made with equal probabilities.[16] Symmetry also implies there is a simple transformation of the relationship in equation (50) when a new alternative is added such that the probabilities of previously available options fall. Also, note that the probability with which each outcome is chosen depends solely on the vector of utilities.

Finally, note that, as argued by Anderson et al (2004), the Luce model can be derived directly from a Markovian learning model.

Random utility and Address Models

In random utility models it is assumed that the utility of each option is affected by a random idiosyncratic shock, which is unobservable to an econometrician. Individuals are rational and choose the option with the highest total utility, which is the sum of the observable and unobservable components. However, from the point of view of an econometrician the choice is probabilistic.

In address models probabilistic choice on market level arises from unobserved heterogeneity of the consumers in the horizontal direction. Anderson, de Palma, Thisse (1992) establish that under broad assumptions these models

[15] This follows from the index theorem (see, for example, Section 17D of Mas-Colell, Whinston, Green, 1995), which implies that generically the number of the fixed points is odd.

[16] If condition (51) is satisfied then the steady state is generically unique, and therefore symmetric.

are equivalent to random utility models.

Hence, in the standard random utility (or address) model, as described by Anderson, de Palma and Thisse (1992), an individual chooses one of the n alternatives, with which the payoffs $u_1 + \varepsilon_1, ..., u_n + \varepsilon_n$ are associated. Vector u is publicly observable and common among the individuals. For example, if the alternatives are jobs vector y, u can refer to wages. Vector ε, on the contrary, is the private information of the individual. We assume that it is distributed over R^n with some strictly positive density, which does not depend on the base payoff vector u. An econometrician will observe the following choice probabilities:

$$p_i(y) = \Pr(u_i = \max u_j) = \Pr(\varepsilon_j \le u_i + \varepsilon_i - u_j,\ \forall j = \overline{1, n}). \tag{55}$$

First, note that any random utility model can be re-interpreted as resulting from Markovian learning. Indeed, if $q \in \Delta^n$ is the vector of probabilities generated by the random utility model, consider a Markovian model with

$$f(p, u) = (qq^T)p, \tag{56}$$

where q^T is a row vector transposed to the column vector q. It is easy to see that vector $p = q$ is a steady state of system (50) (recall that since q is a probability vector, $q^T q = 1$). The reverse, however, is not always true. Falmange (1978) proved that a system of choice probabilities has a random utility representation if and only if its Block-Marschak polynomials are non-negative. Intuitively, non-negative Block-Marschak polynomials imply that adding an alternative never increases the probabilities of the remaining choices and the marginal effect of adding an alternative (as well as the marginal effect it has on the marginal effect, etc.) decreases as choice set shrinks (see, Anderson, de Palma, Thisse, 1992).

Second, note that the choice probabilities depend on both the utilities and the distribution of the random shocks (as there are infinitely many joint distributions of $(n+1)$ random variable with the same $n-$dimensional marginals).

Finally, the Luce model can be derived from a random-utility model if the unobserved components of utility are independently, identically distributed according to the extreme value distribution, with parameter λ (see Anderson, de Palma and Thisse, 1992). However, generating the specific result that probabilities of existing varieties decline as new varieties are introduced requires extra assumptions on the joint probability distribution of random components of utility for the new and old varieties, as the general random-utility requires only that probabilities are non-increasing.

Machina's approach

Machina (1985) proposes that individuals have direct preferences over probability distributions summarized by a function $V(p)$. Machina's approach can be interpreted as a model of bounded rationality by associating with each probability distribution a numerical function, the cost of computation.

Definition *A continuously differentiable convex function* $c : \Delta \rightarrow R_{+}$ *is* *called a cost of computation if*

$$\lim_{y \to x} \|\nabla c(y)\| = \infty \tag{57}$$

for any x on the boundary of Δ.

This definition implies that definitely excluding even one alternative (as represented by a point on the boundary of the unit-simplex) as a possible solution is prohibitively costly. On the other hand, selecting some distributions may entail very low cost. For example, I may look at my watch and select the alternative whose number in the list equals the number of minutes past after the last whole hour. This rule will produce some distribution of choices that depends on my behavioral habits. Though the quality of choice will be very poor, since the rule has no relation to the actual payoffs, the cost of computation in this case is minimal. A boundedly rational individual selects the vector of choice probabilities $p^*(u)$ to solve:

$$p^*(u) = \arg\max(\sum_{i=1}^{n} p_i u_i - c(p)), \tag{58}$$

i.e. Machina's utility function has a form:

$$V(p) = \sum_{i=1}^{n} p_i u_i - c(p). \tag{59}$$

For a given set of choices, any choice probabilities derived from a random utility model can be obtained from (58) for an appropriate choice of $c(\cdot)$. More precisely, the following theorem holds:

Theorem *(Hofbauer and Sandholm, 2002) Let $p^*(u)$ be the vector of choice probabilities obtained from (55), where the components of vector ε are i.i.d.*

over R^n with some strictly positive density, which does not depend on the payoff vector π. Then there exists a convex function $c : \Delta \to R$, continuous on Δ and continuously differentiable on its interior such that

$$p^*(\pi) = \arg\max(\sum_{i=1}^{n} p_i u_i - c(p)) \tag{60}$$

$$\lim_{p \to q} \|\nabla c(p)\| = \infty \tag{61}$$

for any q on the boundary of Δ.

Note that, as with the Markovian model, probabilities are solely a function of the utilities. Furthermore, if we specify the cost of computation as

$$c(p) = \eta(\ln n + \sum_{i=1}^{n} p_i \ln p_i). \tag{62}$$

we can directly derive the Luce model (Fudenberg and Levine, 1998).

Table 1
SUMMARY OF DATA

Category	Features	Price	Reg'ns	Makes	Models	Variants	Observations
All	11	59.6	584.3	37	175	484	1081
Small	7	22.6	251.4	19	39	133	397
Medium	11	34.6	268.6	9	16	54	207
Luxury	15	114.3	47.9	26	97	251	390

	Reg'ns per				Per Make		
Category	Make	Model	Variant	Obs.	Models	Variants	Obs.
All	15.79	3.33	1.21	0.54	4.73	13.08	29.22
Small	13.23	6.44	1.89	0.63	2.05	7.00	20.89
Medium	29.84	16.78	4.97	1.30	1.78	6.00	23.00
Luxury	1.84	0.49	0.19	0.12	3.73	9.65	15

Note: Price and Features data supplied by Glass's Guide. Registrations data from Glass's Guide (1998). Price and Registrations are in thousands. All includes Small, Medium, Luxury, Sports and People Movers. Small combines the Glass categories of Small and Light. Medium combines the Glass categories of Medium and Upper Medium. Luxury combines the Glass categories of Prestige and Luxury

Table 2
SUMMARY BY MAKE

Category	Observations	Makes	Reg'ns	Price
Low	187	7	94.4	20.3
Medium	544	11	443.2	34.2
UpperMedium	155	6	27.1	63.1
Prestige	135	4	18.1	107.6
High	60	9	1.5	294.7

Note: Price and Options data supplied by Glass's Guide. Registrations data from Glass's Guide (1998). Price and Registrations are in thousands. See Appendix 1 for assignment of makes to groups.

Table 3

SUMMARY OF KERNEL DENSITIES BY MAKE

Category	Number of Makes	Mean Share	Median Share	Minimum Share	Maximum Share
Low	3	0.61	0.62	0.59	0.63
Medium	8	0.49	0.51	0.14	0.76
UpperMedium	4	0.45	0.4	0.27	0.73
Prestige	2	0.42	0.42	0.28	0.56

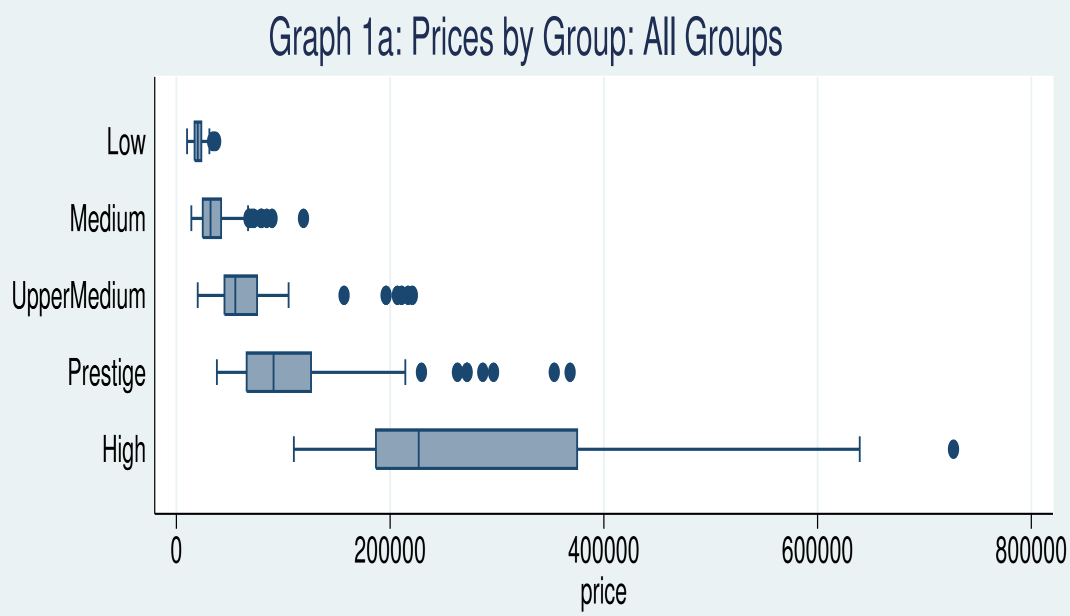
Note: Price and Options data supplied by Glass's Guide. Registrations data from Glass's Guide (1998). See Appendix 1 for assignment of makes to groups.

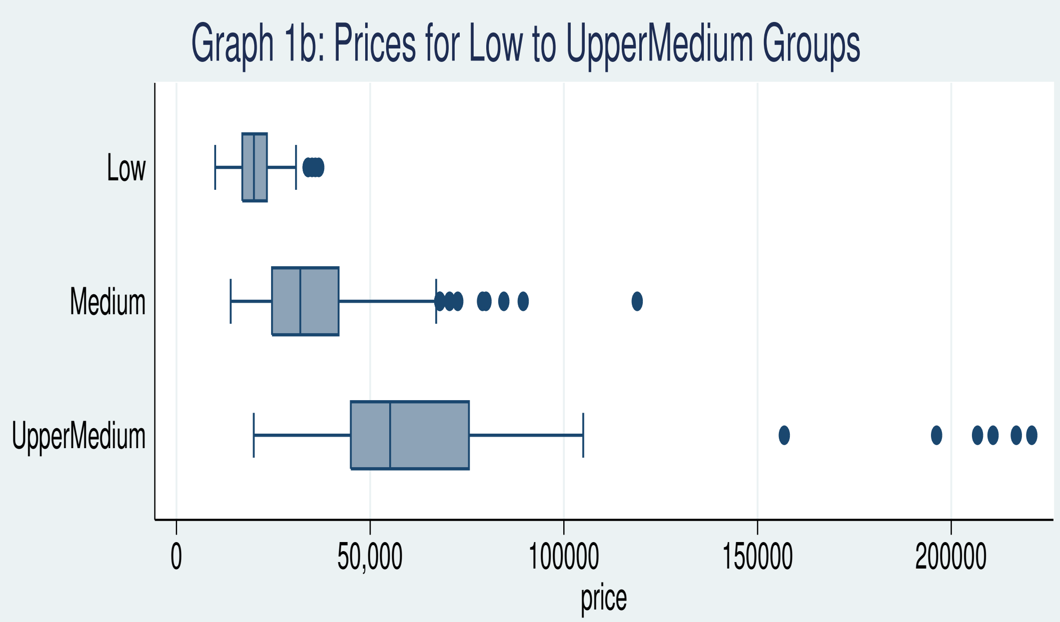
Table 4

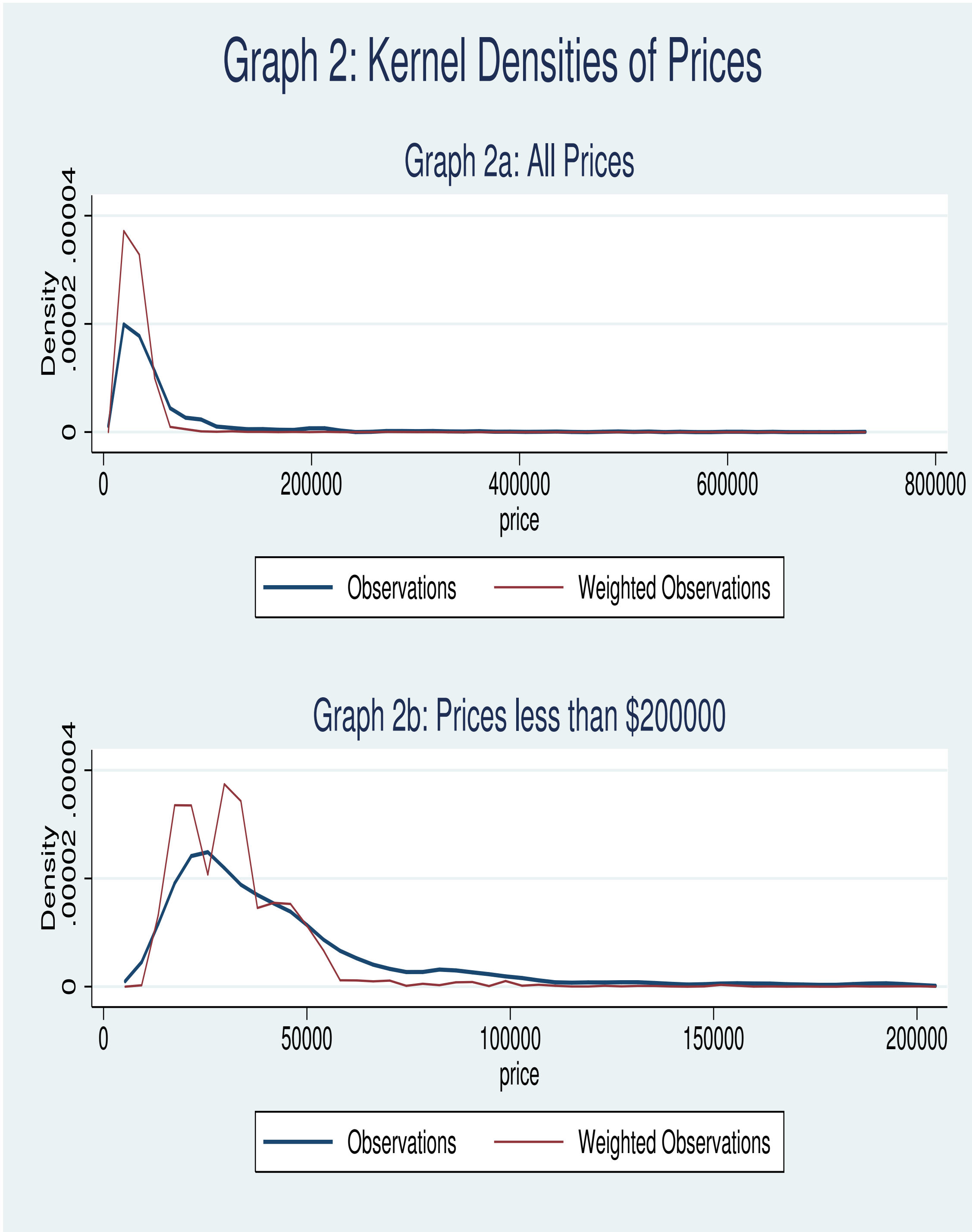
SUMMARY OF KERNEL DENSITIES BY MAKE - SMALLER BANDWITH

Category	Number of Makes	Mean Share	Median Share	Minimum Share	Maximum Share
Low	3	0.58	0.70	0.32	0.73
Medium	8	0.48	0.48	0.11	0.81
UpperMedium	4	0.43	0.3	0.16	0.7
Prestige	2	0.44	0.44	0.31	0.56

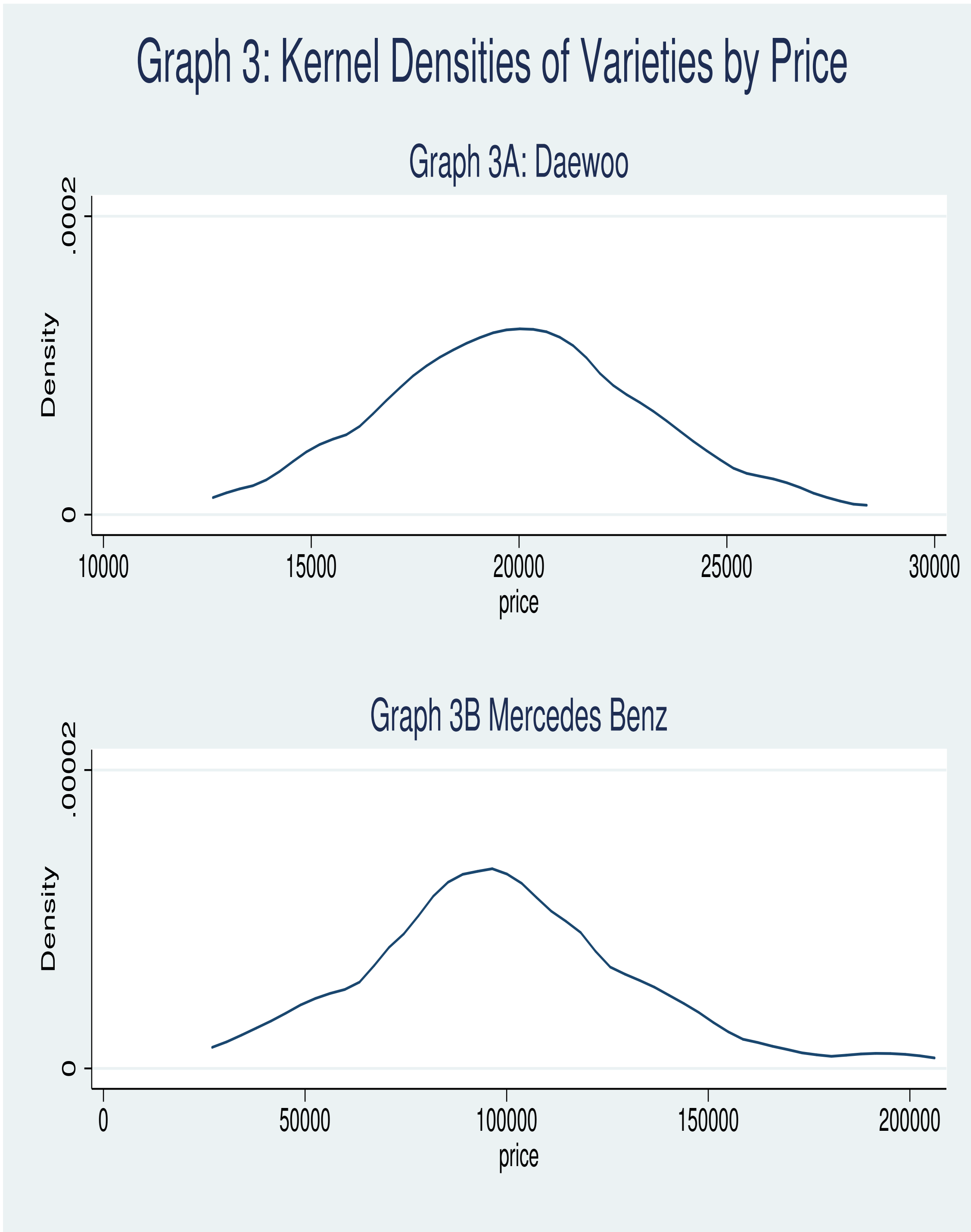
Note: Price and Options data supplied by Glass's Guide. Registrations data from Glass's Guide (1998). See Appendix 1 for assignment of makes to groups.

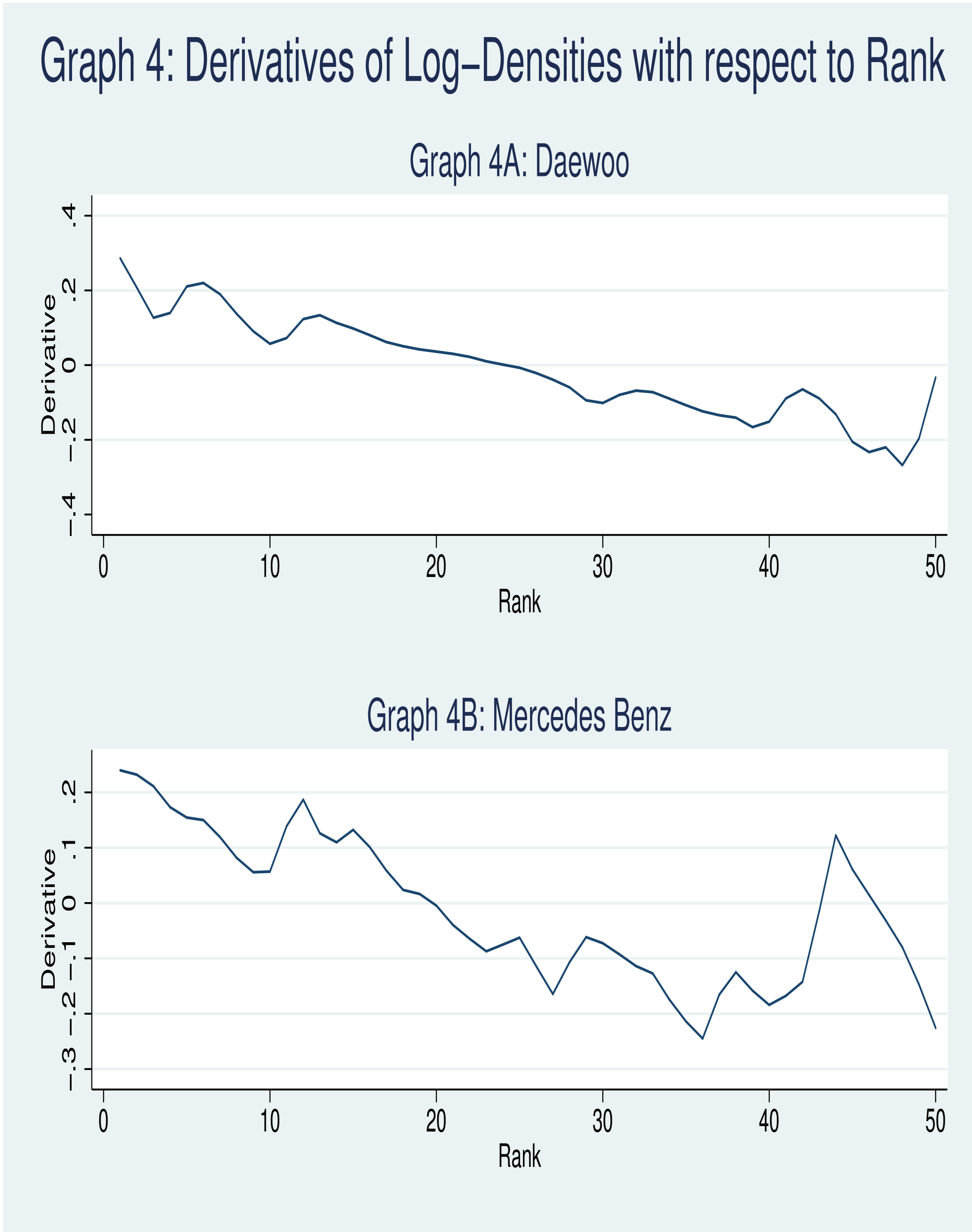

Graph 1a: Prices by Group: All Groups
Low
Medium
UpperMedium
Prestige
High
0
200000
400000
600000
800000
price


Graph 1b: Prices for Low to UpperMedium Groups
Low
Medium
UpperMedium
0
50,000
100000
150000
200000
price



Graph 2: Kernel Densities of Prices
Graph 2a: All Prices
Density
0
.00002
.00004
0
200000
400000
600000
800000
price
Observations
Weighted Observations
Graph 2b: Prices less than $200000
Density
0
.00002
.00004
0
50000
100000
150000
200000
price
Observations
Weighted Observations


Graph 3: Kernel Densities of Varieties by Price
Graph 3A: Daewoo
Density
.0002
0
10000
15000
20000
25000
30000
price
Graph 3B Mercedes Benz
Density
.00002
0
0
50000
100000
150000
200000
price


Graph 4: Derivatives of Log−Densities with respect to Rank
Graph 4A: Daewoo
Derivative
−.4
−.2
0
.2
.4
0
10
20
30
40
50
Rank
Graph 4B: Mercedes Benz
Derivative
−.3
−.2
−.1
0
.1
.2
0
10
20
30
40
50
Rank