MANAGEMENT SCIENCE
Vol. 50, No. 5, May 2004, pp. 686–700
ISSN 0025-1909 | EISSN 1526-5501 | 04 | 5005 | 0686

DOI 10.1287/mnsc.1030.0200

Product Variety Under Brand Influence: An Empirical Investigation of Personal Computer Demand

Kai-Lung Hui
Department of Information Systems, National University of Singapore, Singapore, lung@comp.nus.edu.sg

Prior research suggests that brand may influence consumer preference for differentiated products. However, the extant literature does not measure how brand value affects product similarity and consumer choice. This paper examines demand response to the proliferation of personal computers (PCs). Using both the central processing unit (CPU) and brand as segmentation variables, I construct a two-level nested generalized extreme value (GEV) discrete choice model to estimate the brand values and product similarities of a set of PC vendors. With these estimates, I infer the relative efficacy of product variety for firms which possess different degrees of brand values. My results suggest that consumers treat PCs from the same firm as close substitutes, and the proximity of the PCs correlates positively with the firms' brand values. This finding suggests that there are decreasing demand returns to product variety for branded multiproduct firms. I discuss a few possible drivers of brand value and explore the significance of product line extension in building long-term brand reputation.

Key words: product variety; brand value; discrete choice; similarity; cannibalization
History: Accepted by Jagmohan S. Raju, marketing; received May 16, 2002. This paper was with the author 7 months for 3 revisions.

1. Introduction

Many durable goods manufacturers engage in extensive product proliferation to capture demand from multiple consumer segments. For example, Hyundai produces the *XG* premium sedan for luxury-car consumers, the *Accent* for small-car consumers, and the *Santa Fe* for consumers who prefer multipurpose vehicles. IBM sells three series of its *NetVista* desktop personal computers (PCs) to different home and organizational buyers, and it further allows buyers to customize certain components and peripherals within each series. Generally, multiproduct firms offer a variety of choices in a single product category to capture the surplus of consumers who may have heterogeneous quality valuations, tastes, or budget constraints. Past empirical studies have repeatedly found such high-variety strategies to be effective in expanding sales (see, e.g., Bayus and Putsis 1999, Chong et al. 1998, Kekre and Srinivasan 1990).

The revenue and cost implications of product variety have raised significant interest in economics, management science, and marketing research.[1] In theory, a firm prefers to differentiate its products and price schedules finely so that each consumer can find an offer that matches her preference—a case that corresponds to perfect price discrimination. However, notwithstanding the difficulty in identifying consumer preference, two factors limit the feasibility of such an extensive differentiation. First, on the supply side, increasing variety may raise production and operating costs because of the loss of scale economies and the imposition of supply-chain market mediation costs (Randall and Ulrich 2001). Second, on the demand side, because the firm sells a menu of similar products, consumer self-selection may result in significant cannibalization between the product variants of the firm. Anticipating that its products will cannibalize each other, the firm may prefer to aggregate its offers and sell a (much) smaller number of variants (Katz 1984, Moorthy 1984).

While previous research has generated substantial evidence and insights about the cost implications of product variety,[2] empirical work on demand response to variety and the extent of cannibalization within a product line is scant. Do consumers perceive the variants in a firm's product line as similar to each other? If so, does this perception vary across firms with different brand attributes or overall product qualities?

[1] I follow Lancaster's (1990) terminology and use the word variety to indicate the number of variants within a single product category. In subsequent discussions, I shall use the terms "high product variety" and "wide product line" interchangeably.

[2] For example, Fisher and Ittner (1999) and MacDuffie et al. (1996) find that the differentiation of automobile parts and options increases production cost and lowers overall productivity. Chakravarty and Balakrishnan (2001), Fisher et al. (1999), Lee and Tang (1997), and Swaminathan and Tayur (1998) suggest that modularization with common components and delayed differentiation may help firms exploit scale economies, and reduce design and production costs.

The answers to these questions are important to multiproduct firms in competitive industries as they provide insights into the benefits of product variety. Generally, a wide product line could help a firm preempt an entire market, protect entrenched niches, or develop brand reputation (Bayus and Putsis 1999, Putsis and Bayus 2001, Stavins 1995). Because product proliferation increases the likelihood of consumers finding their preferred products within a given product line, it may help expand a firm's sales and crowd out its competitors or future entrants (Schmalensee 1978). Further, product line extension allows a firm to send out credible quality signals that may have spillover effects on its existing products and hence strengthen its brand reputation (Balachander and Ghose 2003, Dacin and Smith 1994, Wernerfelt 1988). These effects may in turn contribute toward shaping the firm's brand value.[3]

Conversely, brand value could moderate the demand expansion effect of product variety. Previous research suggests that commonality, such as brand name or feature attributes, increases the perceived similarity between product variants in the same family (Kim and Chhajed 2000, 2001; Krishnan and Gupta 2001; Wanke et al. 1998). Ex ante, because consumers view products that share the same brand as rather similar, they tend to evaluate the products jointly when making purchase decisions. This could lead to undesirable local competition between the products, which limits the collective sales of the entire product line (i.e., the products may cannibalize each other more than they capture the demand from rival brands; Reddy et al. 1994). Because consumers frequently assimilate products that share a prominent brand (Park et al. 1991, Wanke et al. 1998), it is important to examine *whether* and *how* the brand value of a firm affects the perceived similarity of its products.

This study investigates the demand implications of brand and product variety interaction. Using a random-utility discrete choice model at the individual product level, I scrutinize the utility of consumers toward differentiated products that carry distinct brand and feature attributes in a single market. The discrete choice framework provides a direct linkage between the hypothetical model of consumer utility and empirical market share observations that is used as a proxy variable to measure the choice outcomes of consumers. By fitting the discrete choice model to product data and estimating its parameters, I examine on a firm-by-firm basis the similarity between a firm's products, and use it to infer the effect of product-line width on the studied firms' brand shares. The similarity estimates offer useful implications on cannibalization and the effectiveness of product variety in the presence of heterogeneous brand values.

I estimate the discrete choice model, which is derived from a generalized extreme value (GEV) demand specification, using a set of PC data from 1985 to 1993. I use a two-level nested structure for the GEV model to allow for rich similarity coefficients for the studied brands and central processing units (CPUs), the latter being a key quality attribute of most PC systems. To obtain consistent estimates of the model parameters, I transform the GEV model using the recent methodological advances by Berry (1994) to obtain a linear utility equation. This allows me to take into account endogenous prices in discrete choice product demand using linear instrumental variable estimations. My results suggest that consumers view the products of branded firms as highly similar to each other. This implies that cannibalization may be an imminent concern for firms that possess superior brand values. Compared to weakly branded rivals, these firms may derive less demand expansion from product proliferation.

Prior research has illustrated that new product extensions may benefit from having favorable parent brands and can at times enhance the market shares of multiproduct firms (Bayus and Putsis 1999, Kekre and Srinivasan 1990, Smith and Park 1992). By examining both brand value and product similarity in an integrated model, this paper adds to the existing literature a delicate assessment of the value of product variety across a large number of brands in a competitive industry where product proliferation is a common practice. Further, the assertion—that consumers appreciate brand when comparing differentiated products—has been a basic premise in the theoretical analyses of product differentiation and cannibalization (e.g., Desai 2001, Desai et al. 2001, Gilbert and Matutes 1993, Katz 1984). This study provides a pioneering estimate of whether such an assertion conforms to empirical demand pattern.

1.1. Product Differentiation, Similarity, and Brand Value

In an often-cited article, Moorthy (1984) studies a monopolist's problem of product-line design and pricing. His basic premise is that consumers have heterogeneous quality preferences that are not immediately observable by a seller, and they select products through a process of price-quality trade-offs. Because multiple products are offered in the same category, the (cheaper) lower-end products may cannibalize the sales of the (more expensive) higher-end ones, which limits the ability of the monopolist to extract surplus

[3] My conceptualization of brand value parallels that of Katz (1984), which includes both reputation and the commonality of product features. Reputation captures consumer tastes toward a firm, whereas feature commonality that affects the physical quality and performance of an entire product line.

from consumers in the high-value segments. Therefore, due to cannibalization, the monopolist may prefer to aggregate his offers and sell a small number of products that are well dispersed in the quality space.

Desai (2001) and Katz (1984) study a similar problem, but they incorporate brand in their models and extend the analysis into duopoly settings. They conclude that brand may moderate cannibalization and interfirm competition. If consumers have strong tastes in superior brands, then the price competition among lower-end products may make such products more attractive to the consumers. This intensifies cannibalization, because now the consumers may view the higher- and lower-end products as offering similar utilities.

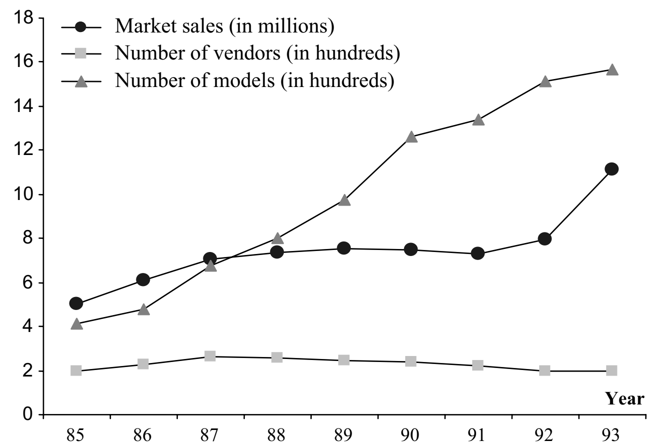
Theories in consumer psychology also suggest that brand could serve as a categorization attribute that could weaken the perceived differentiation of a firm's products (Boush and Loken 1991, Wanke et al. 1998). Consumers often extend their beliefs about and attitudes toward a brand to all of its products (Randall et al. 1998).[4] Hence, to them, the products may offer similar qualities, meaning that the products may substitute for each other in a consumer's choice (Desai et al. 2001, Robertson and Ulrich 1998). This might reduce the firm's ability to compete with rival firms outside its core market niches.

In an interesting study, Park et al. (1991) find that a prestigious brand is more extendable across products. This implies that even if some products carry dissimilar features, functions, or qualities, a common prestige brand concept may still govern their perceived qualities. This implication hints at the possibility that brand value may relate positively with product similarity. However, if consumers do categorize products by brand, then it could also be difficult for an inferior brand to extend its product line into premium market segments. On balance, whether product similarity increases or decreases with brand value is an important empirical question that remains to be answered. In the following sections, I shall address this question with large-scale sales data from the PC market.

1.2. The Personal Computer Industry

Following Bayus (1998, p. 766), I consider the PC as a general purpose, single-user machine that is microprocessor based. The PC industry has been growing rapidly since 1981, with continuously expanding sales, an increasing number of product models, and a large number of competing vendors (Bayus and Putsis 1999). In the early years, shortly after the introduction of the IBM PC in 1981, the supply of PCs was not able to keep up with the sudden upsurge in demand, which resulted in considerable order backlogs and brand-switching behavior (Langlois 1992). To prevent this initial turmoil from affecting my analysis of consumer choices, I have confined my scope to 1985–1993, a period when most PC vendors were able to fulfill purchase requests from consumers. There was also intense price and product competition in that period, and the assembly of PCs largely employed vertically disintegrated manufacturing processes, with most modular components provided by specialized suppliers (Angel and Engstrom 1995, Langlois 1992). Figure 1 shows a few descriptive trends of the PC market.[5]

Figure 1 PC Market Trends: 1985–1993



Several properties of the PC market suggest that it is suitable for studying the demand implications of brand and product variety interaction. First, the market is large enough to accommodate extensive product proliferation by vendors. Indeed, during the studied period, it was not uncommon for a vendor to concurrently sell more than 20 PC models that used various CPUs. Such a huge market capacity facilitates the analysis of outcomes due to product similarity and variety on a firm-by-firm basis.

Second, the large size and competitive nature of the PC market imply that consumer choice is not likely to depend heavily on the decision, entry, or exit of any particular vendor. Also, there is no evidence of PC vendors using product proliferation to deter entry (Bayus and Putsis 1999). Hence, the differential impact of product variety offers valuable insights to PC vendors who presumably are more interested in sales expansion than strategic preemption or entry deterrence.

[4] This is reflected in sayings such as: "No one was ever fired for buying IBM," or "It's a Sony!"

[5] The trends in Figure 1 are analogous to those reported by Bayus (1998) and Bayus and Putsis (1999). However, because I consider only the sales of desktop PCs in the U.S. domestic market (i.e., exclusive of laptop PCs and international shipments), the reported trends are slightly different from the cited studies.

Finally, the PC market is periodically stimulated by technological or software innovations. For instance, the introduction of the Intel 80386 CPU in 1986 and the 80486 CPU in 1989, together with improvements in hardware peripherals such as massive storage devices, networking interface, and graphic display units, has expanded the application scope of the PC. New software utilities, such as e-mail, presentation aid, and scheduling tools, have also increased the value of the PC as general home and office equipment. Advances such as these have led to noticeable variations in product-line width over time (Bayus et al. 2003), which is useful for inferring the firm-level performance of product variety from observed sales data.

Hartman and Teece (1990) find that brand reputation affects the market shares, pricing, and quality strategies of minicomputer vendors. Bresnahan et al. (1997) find evidence of brand segmentation in the PC market in the late 1980s. In their empirical model, they classify PCs as either branded or nonbranded, and find that PCs compete locally within their respective market segments. It is worthwhile to further explore the detailed interaction between brand, product similarity, and variety in the PC market.

The rest of this paper is organized as follows: Section 2 provides a sketch of the analysis strategy adopted in this study. It also presents the GEV choice model, explains its econometric properties, and discusses the roles of the included variables. Section 3 presents the research data and outlines the procedure of constructing exogenous instruments. Section 4 summarizes the estimation results and statistical tests. Section 5 discusses the research findings and implications. Finally, §6 concludes the paper.

2. The Model

2.1. Analysis Strategy

If brand value affects product similarity, then the correlation of consumer preferences for a firm's products should vary. Based on this premise, I devise an indirect inference procedure. First, I develop a discrete choice model that explains consumer preference for PC systems. I incorporate a separate similarity coefficient for each vendor to measure correlated consumer preferences (this corresponds to the elasticity of substitution between the firm's PCs). Next, I estimate brand values and similarity coefficients for a set of PC vendors, and then apply the estimates in a post hoc analysis to explicate their relationship.

2.2. Demand Structure

The PC consists of several major components, including the CPU, memory, storage devices, and monitor. Because most PC vendors produce multiple system configurations, the prices of PCs differ substantially both within and across vendors. For example, in 1993, an IBM Pentium PC could cost twice the price of an IBM PC that was driven by an 80486 chip. Because consumers have different processing needs and budgets, they may prefer a particular PC configuration irrespective of brand value. Therefore, other than brand, it is possible to segment the PC market along numerous feature dimensions.[6] It is important to control for such (physical) quality segmentation in my discrete choice model.

Among the feature attributes, the CPU is probably the most dominant characteristic that dictates a PC's computation capability. Most PC vendors devise product-line strategies and promotional campaigns using CPU class and speed. Empirical hedonic studies show that the CPU commands high implicit prices (Berndt et al. 1995, Nelson et al. 1994). These observations indicate the importance of the CPU to consumers buying and evaluating computers.

Further, Venkatesh and Mahajan (1997) show that the CPU could suppress brand in shaping the price premiums of PC systems. Bresnahan et al. (1997) also treat the CPU as one principle of differentiation in the PC market. In view of these suggestions, and because of the scarcity of detailed specification data, I allow for flexible similarity coefficients (substitution parameters) only for the studied brands and CPUs in my choice model. As long as the CPU is correlated with other quality attributes, the inclusion of the CPU variables and similarity coefficients should be sufficient in modeling feature segmentations.

Given that I have two segmentation variables—brand and the CPU—I need to find a demand structure that can capture consumer preference similarity along the two dimensions. I use a two-level nested GEV structure in my model, with the CPU at a higher-level nest than brand. The nesting level of the CPU is motivated by the evidence generated by Venkatesh and Mahajan (1997), which shows that the CPU is more important than brand in PC purchases. I have considered other possible demand specifications, such as the symmetric principle of differentiation GEV model of Bresnahan et al. (1997), the mixture logit model of Kamakura et al. (1996), and the ordered alternatives multinomial logit model of Small (1987). However, because I need to consider many brands and CPUs that consumers do not frequently repurchase and which cannot be sorted into a priori order, the cited models are neither appropriate nor feasible.[7]

[6] I thank two anonymous reviewers for pointing out this possibility.

[7] The symmetric principle of differentiation GEV model (Bresnahan et al. 1997) is particularly appealing because it does not impose an ex ante hierarchical nest structure. However, because I need to incorporate many similarity coefficients, it would be difficult to

2.3. Details

Consider a market with J firms, each of which produces n_j variants of PC models, $j = 1, \ldots, J$. I assume the existence of an outside group, $j = 0$, which corresponds to a no-purchase alternative. Consumers can always choose not to buy any of the PCs from the J firms if they find all the quality-price bundles unattractive. Note that the total number of firms, J, varies across years. I omit all year subscripts for brevity.

I use subscript i to index the CPUs that are used in the PC models, with $i = 1, \ldots, I$, and k to index the model variants within each CPU category. That is, the products of i and k constitute all the n_j PC models produced by firm j. The quality of a PC is described by three attributes: (1) the brand value of the firm j, b_j; (2) the processing capability offered by the CPU i, c_i; and (3) the quality due to the *differentiation* attribute, ξ_{ijk}, that captures other unobservable quality variations. Further, I include an age variable, a_{ijk}, to capture consumer preference for PCs of different vintages. This could serve as a proxy variable for unobservable attributes that are age related (e.g., newer PCs may use better parts and components). Based on this setup, the utility of model ijk to consumer m is

$$u_{ijkm} = a_{ijk} + b_j + c_i + \xi_{ijk} - \alpha p_{ijk} + \varepsilon_{ijkm}. \tag{1}$$

By this specification, the unobservable attribute ξ_{ijk} captures the differential qualities of firm j's PCs that use CPU i, and it is mean independent of the "group" attributes, b_j and c_i. That is, $\mathrm{Cov}(\mathbf{b}, \boldsymbol{\xi}) = 0$ and $\mathrm{Cov}(\mathbf{c}, \boldsymbol{\xi}) = 0$.

In (1), p_{ijk} denotes the price charged by firm j for model k that uses CPU i, α represents the price elasticity of utility, and ε_{ijkm} is a unique taste parameter for consumer m, with $m = 1, \ldots, M$. M is a yearly constant indicating PC market potential, which includes consumers who choose the outside product because of insufficient utility. I assume consumer tastes ($\boldsymbol{\varepsilon}$) are correlated for PCs that use the same CPU, and that there is a second-order correlation for PCs produced by the same firm.[8] That is, $\mathrm{Cov}(\varepsilon_{ijkm}, \varepsilon_{abcm}) \geq 0$ when $i = a$, and it further increases when $j = b$. $\mathrm{Cov}(\varepsilon_{ijkm}, \varepsilon_{abcm}) = 0$ when $i \neq a$.

Consumers make independent choices of PC models. If each ε_{ijkm} is a mean-zero extreme-value random variable, then a collection of such variables follows a multivariate extreme-value distribution. By the generalized extreme-value model (GEV) (McFadden 1978), the conditional market share of selecting model k from firm j's PCs that use CPU i is

$$s_{k|j|i} = \frac{e^{\delta_{ijk}/[1-\sigma_j]}}{G_j}, \tag{2}$$

where

$$G_j = \sum_k e^{\delta_{ijk}/[1-\sigma_j]}, \tag{3}$$

and

$$\delta_{ijk} = a_{ijk} + b_j + c_i + \xi_{ijk} - \alpha p_{ijk} \tag{4}$$

is the *mean* utility provided by model k. Here, σ_j captures the choice similarity between firm j's PCs. If a particular σ_j approaches zero, then the within-group (i.e., brand) utility correlation goes to zero. If it approaches one, then firm j's PCs are highly similar to each other.

Similarly, the conditional share of selecting firm j from a group of PCs that use CPU i is

$$s_{j|i} = \frac{G_j^{[1-\sigma_j]/[1-\rho_i]}}{F_i}, \tag{5}$$

where

$$F_i = \sum_j G_j^{[1-\sigma_j]/[1-\rho_i]}. \tag{6}$$

Like σ_j in (2) and (3), ρ_i captures the choice similarity between all PCs that use CPU i.

Finally, the market share of CPU i is

$$s_i = \frac{F_i^{[1-\rho_i]}}{\sum_i F_i^{[1-\rho_i]}}. \tag{7}$$

With (2), (5), and (7), the (unconditional) market share of model ijk can be computed as

$$s_{ijk} = s_{k|j|i} \cdot s_{j|i} \cdot s_i = \frac{e^{\delta_{ijk}/[1-\sigma_j]}}{G_j^{1-[1-\sigma_j]/[1-\rho_i]} F_i^{\rho_i} \sum_i F_i^{[1-\rho_i]}}. \tag{8}$$

It is instructive to observe that if all the σ_js and ρ_is equal zero, then (8) becomes a standard logit formula with cross elasticities proportional to mean utilities; if they equal one, then the PCs are perfect substitutes. By estimating the similarity coefficients, I could measure the within-group (i.e., brand or CPU) utility correlations based on observed consumer choice data. These correspond to the similarities of the studied PCs.

estimate and identify the many weighting coefficients associated with the brand and CPU segments. To assess the generalizability of my findings, I shall reverse the nest order and verify the brand value and PC similarity relationship in one of the subsequent robustness tests (see Footnote 20).

[8] Because of the chosen hierarchical structure, firm-level preference correlations are always modeled within CPU nests. That is, I do not treat PCs produced by the same firm with different CPUs as closely related items. Later (in Footnote 20), I shall reverse the roles of brand and the CPU in the model. For ease of presentation, in subsequent discussions, firm-level preference correlation or product similarity refers to cross-substitution between a firm's PCs that use the same CPU.

The outside group contains only one member. By normalizing the mean utility of the outside product to zero, the market share of the outside group is

$$s_0 = \frac{1}{\sum_i F_i^{[1-\rho_i]}}. \qquad (9)$$

Following Berry (1994), I take logarithms and subtract (9) from (8) to arrive at

$$\log(s_{ijk}) - \log(s_0) = \frac{\delta_{ijk}}{1-\sigma_j} - \left[1 - \frac{1-\sigma_j}{1-\rho_i}\right]\log(G_j) - \rho_i \log(F_i). \qquad (10)$$

To estimate (10), I need to find $\log(G_j)$ and $\log(F_i)$. The (unconditional) share of firm j that uses CPU i is

$$s_{ij} = s_{j|i} \cdot s_i = \frac{G_j^{[1-\sigma_j]/[1-\rho_i]}}{F_i^{\rho_i} \sum_i F_i^{[1-\rho_i]}}. \qquad (11)$$

By taking logarithms and subtracting (9) from (11),

$$\log(G_j) = \frac{1-\rho_i}{1-\sigma_j}[\log(s_{ij}) - \log(s_0) + \rho_i \log(F_i)]. \qquad (12)$$

Substitute this into (10) and rearrange the terms:

$$\log(s_{ijk}) - \sigma_j \log(s_{ijk}/s_{ij}) - \rho_i \log(s_{ij}) - [1-\rho_i]\log(s_0) = \delta_{ijk} - \rho_i[1-\rho_i]\log(F_i). \qquad (13)$$

Finally, by taking logarithms and subtracting (9) from (7),

$$\log(F_i) = \frac{\log(s_i) - \log(s_0)}{1-\rho_i}. \qquad (14)$$

Substitute $\log(F_i)$ and δ_{ijk} in (13) by (14) and (4), respectively:

$$\begin{aligned}\log(s_{ijk}) - \log(s_0) &= \delta_{ijk} + \sigma_j \log(s_{ijk}/s_{ij}) + \rho_i \log(s_{ij}/s_i)\\ &= a_{ijk} + b_j + c_i - \alpha p_{ijk} + \sigma_j \log(s_{ijk}/s_{ij})\\ &\quad + \rho_i \log(s_{ij}/s_i) + \xi_{ijk}. \qquad (15)\end{aligned}$$

Let x_{ijk} denote the sales (in units) of PC ijk, and x_0 denote the size of the outside product. Then, it is straightforward to see that $s_{ijk} = x_{ijk}/M$ and $s_0 = x_0/M$. Also, $s_{ijk}/s_{ij} = s_{k|j|i}$ and $s_{ij}/s_i = s_{j|i}$. Substituting these into (15), the final equation becomes[9]

$$\begin{aligned}\log(x_{ijk}) &= a_{ijk} + b_j + c_i - \alpha p_{ijk} + \sigma_j \log(s_{k|j|i})\\ &\quad + \rho_i \log(s_{j|i}) + \log(x_0) + \xi_{ijk}. \qquad (16)\end{aligned}$$

(16) is a product-level equation that applies to all CPUs, firms, and PC models. That is, there are n_j rows of observations per firm per year. By stacking the J firms' PC data together, and with observations from multiple years, I construct a set of panel data to estimate (16). Both b_j and c_i can be extracted by firm- and CPU-specific dummy variables, respectively. The conditional shares, $\log(s_{k|j|i})$ and $\log(s_{j|i})$, can be readily computed from observed sales data. The size of the outside product, x_0, is not observable, but it is invariable within a particular year. Hence, I can encapsulate $\log(x_0)$ by year-specific dummy variables.

With this dataset, I can estimate (16) by treating the differentiation attribute, ξ_{ijk}, as a mean-zero random variable.[10] The parameters b_j, c_i, α, σ_j, ρ_i, and $\log(x_0)$ can then be uncovered by standard statistical procedures. Although I started from a logit demand specification, because of the linear transformations, I can apply an instrumental variable technique to account for endogenous prices and obtain consistent estimates of the price, CPU quality, brand value, and PC similarity coefficients. The latter two estimates serve as the apparatus for studying the effect of product variety under brand influence.

Before we proceed, it is worthwhile to discuss several econometric properties of the proposed model. First, although it is conceivable for brand value to vary with consumer preference, because (16) operates on mean utilities, the demand for each PC is determined solely by its characteristics.[11] Because b_j and c_i are mean independent of ξ_{ijk} (the only error variable in the equation), they can be consistently estimated by dummy variables.[12]

Second, p_{ijk} may correlate with ξ_{ijk}, as PCs with high-quality differentiation attributes may provide more value to consumers and hence carry higher prices. Because I treat ξ_{ijk} as least squares errors, exogenous price instruments are necessary to consistently estimate the model parameters. Similarly, the conditional market share of the PC models, $\log(s_{k|j|i})$, is also correlated with ξ_{ijk}, which means that more instrumental variables are required.

Finally, brand value may affect the (conditional) market shares of PC firms, which means that b_j and $\log(s_{j|i})$ may correlate with each other. This is a particular case of multicollinearity. It is well known that when right-hand-side variables are collinear, the least

[9] The market potential M is eliminated from (16) due to the logarithmic transformations.

[10] The mean-zero assumption amounts to rescaling the brand values, b_j, to a larger numerical domain. Because the interest here lies in the *relative* strengths of brand values, the transposition of b_j by a common distance is not important.

[11] However, it is individual tastes, $\boldsymbol{\varepsilon}$, that give rise to the GEV demand specification. For a thorough discussion on the role of $\boldsymbol{\varepsilon}$ in discrete choice models of product differentiation, see Berry (1994).

[12] To ensure that my results apply to endogenous (and hence varying) brand values, I shall allow the brand value parameters to change over time in one of the subsequent estimations.

Table 1 Summary Statistics—CPU

	First-generation		8088/86		80286		80386		80486		Pentium	
Year	AP+	%++	AP	%	AP	%	AP	%	AP	%	AP	%
1985	0.87	43.83	2.50	48.54	4.60	7.63	—	—	—	—	—	—
1986	0.66	34.54	2.32	47.75	3.54	17.30	6.80	0.41	—	—	—	—
1987	0.49	15.77	1.76	45.97	2.97	35.14	5.36	3.12	—	—	—	—
1988	0.54	8.72	1.42	39.11	2.65	42.47	4.98	9.70	—	—	—	—
1989	0.52	5.12	1.25	28.48	2.07	42.66	4.12	23.60	5.42	0.14	—	—
1990	0.52	2.02	1.33	18.48	1.76	31.56	3.32	46.66	6.08	1.28	—	—
1991	0.45	0.73	1.10	9.46	1.44	20.34	2.48	61.58	4.19	7.90	—	—
1992	0.64	0.18	0.97	2.77	1.24	4.24	1.85	55.52	2.53	37.29	—	—
1993	0.63	0.06	0.76	0.22	1.15	0.07	1.46	20.64	1.95	78.53	4.19	0.47

+ Sales-weighted average PC price (in thousand U.S. dollars).
++ Percentage of PCs that used CPUs in the class. A small number of observations are omitted because the CPUs are proprietary or unidentifiable (i.e., the technical specifications are not available). These constitute around 1% of the total population.

squares estimator is unbiased and consistent. Efficiency may be adversely affected, but as I shall illustrate below, my dataset contains many observations, which facilitates the large-sample convergence of least squares estimators.

3. Data

To estimate the research model, I acquired a set of comprehensive PC data from the International Data Corporation (IDC). The IDC regularly surveys computer vendors and their channel partners to compile data related to the PC industry. The data are then cross-checked with market analysts and selected end users for accuracy. Because the IDC has a long history of tracking the development of the computer industry, its data are considered reliable and are often used in empirical studies (e.g., Bayus 1998).

For each PC model, the IDC records its vendor, model name, date of first introduction, the CPU used, average retail price, and number of units sold in the U.S. domestic market on a yearly basis. As mentioned above, I confine the study window in this research to 1985–1993 because the supply of PCs was not stable before this period. Figure 1 shows that the number of vendors was quite steady, while the number of PC models rose during the studied period, which means that vendors expanded their PC line widths in general during that time.

Note that the IDC follows the naming conventions of PC vendors in defining models. Some vendors differentiate their models only by CPU class, while others create new models by varying finer feature attributes. To ensure consistency in measurement, I focus exclusively on vendors that defined new models by CPU speed and form factor (or peripheral features) in addition to CPU class.[13] Most vendors in the dataset followed such a differentiation principle within the studied window.

Further, the vendors used a wide variety of CPUs in their PCs. To avoid the unnecessary increase in model parameters, I compare the characteristics of the CPUs (data bus width, address space, and number of transistors), and group them into a smaller number of distinct classes. The Intel x86 family is arguably the most popular CPU in PC history. Hence, I map the other CPUs (produced by AMD, Cyrix, Motorola, etc.) into the x86 categories. Taken together, my dataset includes PCs that employed six classes of CPUs: first-generation, 8086/88, 80286, 80386, 80486, and Pentium. The first-generation class includes CPUs that had data bus widths of eight or fewer bits. These include once popular CPUs such as the 6502 and Z80, which were the dominant choices in the late 1970s and early 1980s.

Table 1 presents summary statistics related to the use of CPUs in the dataset. The price of PCs decreased gradually over the years, and it varied significantly across CPU classes. A PC that used an Intel 80486 CPU could cost 50% more than one that used an 80386 CPU. This indicates that the selection of the CPU as a high-level segmentation attribute is appropriate, as it clearly reflects discrete price ladders in the PC market.

Overall, during the studied period, there were more than 500 vendors in the PC market. Many of them had short life spans of only one or two years, and they sold very small numbers of units. Consequently, it is not possible to separately estimate a set of brand value and similarity coefficients for each of these vendors. Therefore, I restrict my attention to vendors that sold PCs consistently in every studied year. This helps

[13] For example, a vendor that used only one model for all 80486 PCs with various speeds would have been excluded from my analysis. The primary reason is that I do not know if the vendor really hadonly one model, or because it aggregated its numbers when reporting to the IDC. Only firms with reasonably similar differentiation principles (as exhibited in the IDC dataset) are included in the final analysis.

Table 2 Descriptive Statistics: 1985–1993

PC vendor	Average PC sales (000/year)	Average PC price (U.S.$000)	Average no. of models (no./year)	Average PC age+ (years)	Advertising expense (mil. US$/year)
Acma Computers	13.22	2.13	4.11	1.70	—
Advanced Logic Research	39.04	2.30	12.56	0.71	—
Apple Computer, Inc.	1077.67	2.08	11.89	1.45	32.68
Cambridge Graphics Systems	1.18	1.47	2.78	1.60	—
Compaq Computer Corporation	331.12	2.94	11.89	0.79	11.84
CompuAdd	83.17	1.84	12.11	1.05	—
Data General Corp.	8.02	3.64	5.11	2.63	—
Dell Computer Corporation	180.51	2.29	16.44	0.66	—
Digital Equipment Corporation	41.98	2.47	8.67	0.71	3.11
Epson America	104.65	1.65	10.00	1.37	6.42
Fountain Technologies, Inc.	66.35	1.41	8.89	1.34	—
Gateway 2000	153.40	2.10	5.89	1.28	—
Group Bull	195.47	2.64	17.67	0.91	—
IBM	1410.76	2.81	22.11	1.03	44.79
Leading Edge	101.71	1.41	8.56	0.81	—
Memorex-Telex	23.24	2.70	11.56	0.87	—
Micro Express	16.20	1.85	6.44	0.84	—
NCR	167.76	3.11	13.00	1.29	—
NEC Technologies, Inc.	97.48	2.49	15.11	1.13	4.41
Samsung	30.52	1.53	6.11	0.82	—
Tandy/Radio Shack	367.29	1.59	19.89	1.15	17.54
Vtech Computers, Inc.	121.37	1.20	13.22	1.32	—
Wyse	58.15	2.49	5.33	1.10	—
Average	203.92	2.18	10.84	1.15	17.26

+ Obtained by averaging the model ages across the years. This statistic indicates the "freshness" of the vendors' PC portfolios; higher average age means that the vendor generally sells older PCs.

screen out ad hoc vendors who might have failed to survive in the PC market, or those who had concentrated only on transient niche segments.[14] In the final dataset, I sample a total of 23 firms, including major PC vendors such as Apple, Compaq, Dell, and IBM. Table 2 shows several descriptive statistics of the selected vendors.

With respect to (16), I use average retail price as a measure of p_{ijk}, and I create three sets of dummy variables, one each for the studied vendors, CPUs, and years, to capture b_j, c_i, and $\log(x_0)$, respectively. The number of units sold for each model, together with the aggregated firm- and CPU-level sales, is used to calculate the two conditional shares, $s_{k|j|i}$ and $s_{j|i}$. The age of each PC (in years) is computed from the year in which it was first introduced. For instance, a model that was introduced in 1985 would have an age of five in the 1990 data, six in the 1991 data, and so on. This variable helps control for the vintage effect or gradual change in the use of PC parts/components that is not observable.

[14] This screening process may inevitably introduce sample selection bias. However, I believe selecting vendors by longevity is preferable to other size-related characteristics, such as sales, product lines, or prices. As Table 2 shows, the sampled vendors exhibited high variations in sales, which could serve as indicators of their scale of operation.

3.1. Instruments

As discussed in §2.3, the price of PCs (p_{ijk}) may correlate with their differentiation attribute (ξ_{ijk}). Also, the differentiation attribute directly affects the relative sales of a firm's PC models, which means that the attribute correlates with the market shares of the models (conditional on brand and the CPU, i.e., $s_{k|j|i}$). Hence, in (16), both p_{ijk} and $\log(s_{k|j|i})$ are correlated with random product-level errors. An instrumental variable (IV) regression is necessary to obtain consistent estimates of all the model parameters.

Generally, supply-side cost shifters are ideal candidates of instruments, but such data are not available in this study. Interestingly, the exogenous attributes of other products (offered by the same firm or competitors) are also appropriate instruments in discrete choice demand systems with product differentiation (Berry 1994). For each exogenous characteristic attribute z that is used by firm j in model k, an instrument vector includes (1) z itself; (2) the sum of z over all firm j's models, excluding model k; and (3) the sum of z over the models of other firms, excluding firm j (for a detailed discussion on this method of constructing exogenous instruments, see Berry et al. 1995; Bresnahan et al. 1997 apply this method to construct price instruments in their study of PC differentiation).

Table 3 Estimation Results+

Variable	OLS		2SLS	
Price (p_{ijk})	−0.183** (0.019)		−0.265** (0.058)	
Model age (a_{ijk})	−0.044* (0.020)		−0.033 (0.024)	
Vendor	**Brand value (b_j)**	**Similarity (σ_j)**	**Brand value (b_j)**	**Similarity (σ_j)**
Acma Computers	11.771** (0.475)	0.744** (0.100)	11.398** (0.476)	0.581** (0.172)
Advanced Logic Research	11.464** (0.475)	0.815** (0.046)	11.414** (0.480)	0.816** (0.065)
Apple Computer, Inc.	12.293** (0.460)	0.925** (0.069)	12.569** (0.454)	0.998** (0.089)
Cambridge Graphics Systems	11.665** (0.528)	0.531** (0.173)	11.265** (0.583)	0.554* (0.251)
Compaq Computer Corp.	12.090** (0.448)	0.797** (0.044)	11.974** (0.467)	0.726** (0.079)
CompuAdd	11.766** (0.449)	0.767** (0.049)	11.241** (0.460)	0.591** (0.080)
Data General Corp.	12.072** (0.496)	0.662** (0.133)	11.871** (0.526)	0.542** (0.189)
Dell Computer Corp.	11.937** (0.445)	0.850** (0.027)	11.383** (0.458)	0.704** (0.051)
Digital Equipment Corp.	12.239** (0.456)	0.794** (0.053)	12.168** (0.460)	0.776** (0.064)
Epson America	11.752** (0.448)	0.718** (0.050)	11.285** (0.469)	0.555** (0.091)
Fountain Technologies, Inc.	11.818** (0.457)	0.735** (0.047)	11.489** (0.468)	0.641** (0.062)
Gateway 2000	11.936** (0.470)	0.781** (0.061)	11.335** (0.523)	0.528** (0.149)
Group Bull	12.209** (0.431)	0.865** (0.024)	12.033** (0.460)	0.810** (0.058)
IBM	12.244** (0.446)	0.899** (0.032)	12.453** (0.467)	0.923** (0.060)
Leading Edge	12.290** (0.443)	0.978** (0.048)	11.975** (0.447)	0.898** (0.065)
Memorex-Telex	11.938** (0.452)	0.802** (0.038)	11.267** (0.472)	0.595** (0.073)
Micro Express	11.563** (0.513)	0.693** (0.129)	10.878** (0.559)	0.403* (0.163)
NCR	11.905** (0.467)	0.689** (0.070)	11.093** (0.736)	0.325 (0.257)
NEC Technologies, Inc.	11.962** (0.440)	0.805** (0.031)	11.691** (0.451)	0.724** (0.052)
Samsung	11.976** (0.441)	0.790** (0.038)	11.583** (0.461)	0.682** (0.081)
Tandy/Radio Shack	12.227** (0.404)	0.841** (0.035)	12.086** (0.384)	0.795** (0.061)
Vtech Computers, Inc.	11.702** (0.444)	0.819** (0.046)	11.396** (0.460)	0.743** (0.088)
Wyse	12.020** (0.456)	0.812** (0.060)	11.782** (0.472)	0.749** (0.103)
CPU	**CPU quality (c_i)**	**Similarity (ρ_i)**	**CPU quality (c_i)**	**Similarity (ρ_i)**
First-generation++	—	0.589** (0.085)	—	0.505** (0.091)
8088/86	2.335** (0.418)	0.910** (0.026)	2.397** (0.421)	0.849** (0.033)
80286	2.431** (0.422)	0.965** (0.042)	2.603** (0.447)	0.900** (0.048)
80386	2.686** (0.434)	0.927** (0.039)	3.053** (0.541)	0.887** (0.048)
80486	2.867** (0.436)	0.987** (0.034)	3.335** (0.588)	0.951** (0.051)
Pentium	−1.628** (0.518)	0.790** (0.153)	−0.986 (0.821)	0.650** (0.214)
N	2,244		2,244	
R^2	0.865		0.852	

** $p < 0.01$, * $p < 0.05$.

+ White's heteroscedasticity-consistent standard errors are in parentheses. Year-specific dummy variables, corresponding to the term $\log(x_0)$ in (16), are not reported for brevity. All significance levels are calculated using two-tailed tests.

++ The PCs with first-generation CPUs are treated as controls to avoid singularity of regressors.

In my dataset, the only available PC attributes are age and the CPU, the latter consisting of six dummy variables: one each for first-generation, 8086/88, 80286, 80386, 80486, and Pentium. Applying the above technique, each variable can be used to construct three instruments, which means that the instrument vector contains a total of 21 (=7 × 3) variables, excluding brand- and year-specific dummy variables. I then populate these variables with brand dummies to generate a sufficient number of instruments for the conditional shares, $\log(s_{k|j|i})$.[15] Later, I use these instruments in a two-stage least squares procedure to obtain consistent estimates of the demand parameters in (16). The estimation results are reported in the next section.

4. Results

4.1. Stage One: Demand (Utility) Equation

I first estimate (16) by both ordinary least squares (OLS) and two-stage least squares (2SLS) regressions. The results are reported in Table 3. The price coefficient is negative and significant, and its value drops substantially from OLS to 2SLS. This is consistent with previous studies that found more negative price

[15] Essentially, for each studied firm, the characteristics of its own PCs and other firms' PCs are used as a block of instruments forthe conditional share of its products (because σ_j is firm specific), while all the blocks collectively serve as instruments for price (α is common across all the observations).

elasticities using IV estimations (e.g., Berry et al. 1995, Bresnahan et al. 1997). Almost all the brand value, CPU quality, and similarity parameters are significantly different from zero, and they have wider spreads in the 2SLS regression. Because the IV estimator is consistent, I shall focus on the 2SLS results in subsequent discussions.

The age parameter is negative, but it is significant only in the OLS regression ($p < 0.17$ in 2SLS), providing weak evidence to suggest that consumers may rate new PC models more favorably. This could have been due to a vintage effect, but a more plausible explanation is the continuous improvement in computer quality that makes new PCs more appealing than their ancestors.

Among the PC vendors, a few early incumbents, such as Apple, Digital Equipment Corporation and IBM, commanded substantial brand values. The premium value of Apple is particularly noteworthy, because its PCs encompassed somewhat unique designs based on the 680x0 CPU architecture supplied by Motorola. Evidently, many consumers had strong tastes toward these PCs, which contributed to Apple's superior brand value.

Within the six classes of CPUs, except Pentium, the preference for the CPUs followed their order of processing qualities (i.e., highest for 80486 and lowest for first-generation computers). The negative preference for Pentium PCs could be due to their extraordinarily high prices during the initial launch in 1993 (see Table 1), which are not adequately captured by the price parameter. In any case, my dataset contains a very limited number of observations of Pentium PCs. Therefore, this result may not be general.

All firm and CPU similarity coefficients lie within the theoretical range of zero to one and, except for NCR, they are significantly different from zero. That is, consumers regarded PCs that used the same class of CPUs or that were produced by the same vendor as somewhat similar. This finding is consistent with the theoretical models of self-selection. Because consumers select PCs by trading quality with price, PCs that carry the same brand or CPU (i.e., have similar perceived or actual qualities) may locate closer to each other in the product space. This could lead to localized competition (cannibalization), the degree of which is implied by the estimated similarity coefficients.[16]

[16] In a standard vertical differentiation model with self-selection, consumers choose products by maximizing the utility function $u = vq - p$, where q denotes product quality, p denotes price, and v parameterizes the taste for quality. With this stylized setting, it is immediately obvious that the extent of competition depends on how closely the products are located in the quality space (Katz 1984, Moorthy 1984).

4.2. Stage Two: Firm-Level Similarity Analysis

Once the parameters in (16) are identified, I can use the values of b_j and σ_j to infer the relationship between brand and consumer perception of PC similarity. I continue to use the least squares procedure to make statistical inference. However, because both b_j and σ_j are point estimates, it is necessary to account for their errors.

Specifically, I compute the joint variation of b_j and σ_j from the estimated covariance matrix in Stage 1), and use it as weight to perform generalized least squares (GLS). The idea is that point estimates (b_j and σ_j) with smaller standard errors are more reliable and should be allocated larger weights so that they could exert a greater influence on the analysis. A GLS regression of the similarity coefficients (σ_j) on the brand values (b_j) then gives

$$\text{similarity} = -2.5572 + \underset{(0.0409)}{0.2795} \times \text{brand value}. \tag{17}$$

The regression is significant with $F = 46.61$ ($p < 0.01$) and $R^2 = 0.88$. The parameter of brand value is positive and significant ($p < 0.01$), which means that it is more likely for consumers to treat a firm's PCs as similar (closer substitutes) when it has a *high* brand value.

Interestingly, a similar GLS analysis reveals that brand value and the average price range of PCs are positively correlated ($p < 0.02$). That is, in my sample, a firm that had a higher brand value actually sold PCs with wider price dispersions, which presumably indicates that its PCs were more varied in quality. Therefore, it appears that even if a branded firm offers a wide range of products, consumers may still treat its products as close substitutes—a finding that is surprising but remarkable to multiproduct firms.

It is instructive to note that (17) requires parametric assumptions, and it is fitted using information contained in all the b_js and σ_js. Table 4 reports the comparison of brand values across the studied vendors. Evidently, not all vendors had statistically different brand values, which implies that using the estimated values of b_js in (17) may be too optimistic.[17] Instead,

[17] Another limitation of the least squares approach is that the dependent and independent variables (b_j and σ_j) contain correlated measurement errors (due to the common residuals from the first-stage regression). Recall from (16) that b_j is estimated with dummy variables, whereas σ_j is the coefficient of $\log(s_{k|j|i})$, and the latter is negative by definition. Hence, the measurement errors of b_j and σ_j are negatively correlated, which tends to bias the least squares coefficient, obtained by regressing σ_j on b_j, downwards. That is, the true coefficient should be more positive than the one reported in (17). I refer interested readers to Hyslop and Imbens (2001) for a thorough discussion on the implications of mismeasured variables. In this study, because my interest lies in the signs but not magnitudes of the GLS coefficients, the endogeneity of b_j and σ_j should not pose a significant threat to the conclusions.

Table 4 Brand Value and Similarity Comparisons[+]

Vendor		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Acma Computers	[1]			○						○				○	○	○						○		
Advanced Logic Research	[2]			○						×					○	○						×		
Apple Computer, Inc.	[3]	+	+		○	○	○		○		○	○	○	○		○	○	○	○	○	○		○	○
Cambridge Graphics Systems	[4]			−						○					○									
Compaq Computer Corp.	[5]			−			○		○		○						○	○						
CompuAdd	[6]			−		−				○				○	○	○						○		
Data General Corp.	[7]																	○						
Dell Computer Corp.	[8]			−		−				○				○	○	○						○		
Digital Equipment Corp.	[9]	+	+		+		+		+		○	○	○				○	○		○	○		○	
Epson America	[10]			−		−				−				○	○	○						○		
Fountain Technologies, Inc.	[11]			−						−					○	○						○		
Gateway 2000	[12]			−						−					○	○						○		
Group Bull	[13]	+		−			+		+		+						○	○					○	
IBM	[14]	+	+		+		+		+		+	+	+				○	○	○	○	○		○	○
Leading Edge	[15]	+	+	−			+		+		+	+	+				○	○					○	
Memorex-Telex	[16]			−		−				−				−	−	−						○		
Micro Express	[17]			−		−		−		−				−	−	−				○	○	○		○
NCR	[18]			−											−									
NEC Technologies, Inc.	[19]			−						−					−			+						
Samsung	[20]			−						−					−			+						
Tandy/Radio Shack	[21]	+	+				+		+		+	+	+				+	+					○	
Vtech Computers, Inc.	[22]			−						−				−	−	−						−		
Wyse	[23]			−											−			+						

[+] The lower triangle tabulates the comparison of brand values. The plus (minus) entries indicate that the row vendor had a significantly higher (lower) brand value than the column vendor. The upper triangle shows the corresponding comparison of similarity coefficients. Only the similarity coefficients of vendors with significantly different brand values are compared. The entry "○" indicates that the comparison of the similarity coefficients follows the same sign as that of the brand values. The entry "×" indicates otherwise.

a more conservative approach would be to perform a nonparametric sign test, which concerns only the direction of similarity with respect to brand value.

Assuming that similarity is independent of brand value, if the vendors are compared individually, around half of the high-value vendors would have bigger similarity coefficients. In Table 4, 82 pairs of vendors had significantly different brand values (i.e., there were 82 effective "brand treatments"). Among these, the high-value vendors in 80 cases also had bigger similarity coefficients than their counterparts. The probability for this to occur is less than 0.01 (based on a binomial distribution with "success" probability of 0.50; z-score $= 8.50$). Clearly, the (null) hypothesis that similarity is independent of or decreases with brand value should be rejected.

Finally, because my model also groups PCs by their CPUs, the same analysis can be applied to examine PC similarity with respect to differences in CPU quality. The nonparametric sign test gives a two-tailed probability of 0.016.[18] That is, PCs that used higher-quality CPUs were perceived as more similar by consumers. This result is consistent with my overall finding that consumers view high-quality (perceived or actual) products as closer variants.

So far, my analysis has revealed that a firm's PCs are treated as similar variants by consumers, and the degree of similarity varies with its brand value. Because similarity corresponds directly to cannibalization, my findings would have profound implications on the sales and market shares of multiproduct firms. Recall from (16) that $\log(x_{ijk})$ increases with $\sigma_j \log(s_{k|j|i})$, the latter term being always negative. That is, when σ_j is large, the sales of each individual product would drop—a result due to the cannibalization of local products. Hence, the value of σ_j gives a good indication of the sales expansion effect of product-line extension. Because σ_j increases with b_j, a firm that has a high brand value may gain proportionally *less* from product proliferation.

[18] I exclude Pentium PCs from this test because their CPU-quality parameter is not significantly different from zero, and the parameter appears less reliable than those of the other CPUs. Also, I do not perform the GLS analysis because there are too few CPU categories.

4.3. Extension

To further assess the above findings, I estimate an alternative specification that incorporates the firms' advertising expenditures. Prior research suggests that advertising may signal product quality and enhance a brand's premium (Kirmani and Wright 1989, Randall et al. 1998). More importantly, if advertising affects the popularity of a firm's products, ignoring it in Equation (16) might cause systematic bias in the brand value estimates. To address this, I collect a set of firm-level advertising expense data on computers from the Leading National Advertisers *Ad $ Summary* report. I compile the advertising expenditures of seven PC vendors (Apple, Compaq, Digital Equipment Corporation, Epson, IBM, NEC, and Tandy/Radio Shack; see the last column of Table 2 for a summary) from this report, and repeat the analysis by adding them into the first-stage regression.[19]

The coefficient of advertising expenditure is positive, but not statistically significant ($p < 0.72$). A possible explanation is that the firms had used advertising more for promoting old PCs than new PCs (Bayus et al. 2003). Indeed, the correlation between advertising expense and PC age is positive and significant ($p < 0.05$), indicating that firms with older PCs (which are likely to be less favorable to consumers) tend to spend more on advertising. Interestingly, although advertising may not significantly raise utility, it does seem to affect brand value—the (weighted) correlation is positive and marginally significant ($p < 0.10$). These results are consistent with previous studies (e.g., Balachander and Ghose 2003, Randall et al. 1998).

The second-stage GLS and nonparametric analyses produce similar findings; that is, high brand value firms generally have bigger similarity coefficients.[20]

5. Discussion and Implications

Based on empirical data in the PC industry, this study generates two main findings. First, consumers treat PCs of the same brand or those using the same CPU as similar variants. This implies that for a PC vendor, in terms of demand expansion, there may be decreasing returns to product proliferation. In particular, Table 3 shows that consumers view the PCs of several vendors as very close substitutes within their respective product lines. I do not have detailed specification data of all the studied PCs, but I have been able to collect the descriptions of a few PC models from the Ziff Communications *Computer Select* database.

In 1992, IBM produced three models of 80286 PCs—*PS*/1, *PS*/2 25, and *PS*/2 30. A close examination of their specifications reveals that they were almost identical, except for minor features such as hard-disk size and overall weight. Similarly, CompuAdd sold two 80286 PCs, the *CompuAdd* 212 and *CompuAdd* 216, which shared the same platform and peripheral attributes. Only clock speed and the number of expansion slots varied between these two models. Given this anecdotal information, it is not surprising for consumers to show correlated preferences for a firm's PC models. Indeed, the PCs were very similar in both configurations and functionalities.

My second and more important finding is that from the demand perspective, product similarity is positively related to brand value; the higher a firm's brand value, the more likely consumers will see its products as offering comparable utilities. This finding suggests that the short-term, incremental demand gain arising from product proliferation could be limited for high-value brands. Because of local competition, introducing a new variant may not necessarily extend market coverage. Rather, it could lead to a redistribution of consumers, and cannibalize the sales of the existing products. In accordance with Desai et al. (2001) and Katz (1984), the threat of cannibalization is indeed imminent for branded multiproduct firms.

In contrast, for a new firm whose brand value is yet to be established, product-line extension could be a practical strategy for raising short-term sales. A slight degree of cannibalization is unavoidable. However, for such a firm, its products are treated as more distant variants (substitutes). This means that a new line extension may steal few customers from its sister products, and may stand a better chance of attracting marginal consumers who previously would not buy from the firm.

Given the importance of brand value in shaping product similarity, it is worthwhile to explore how such value is formed among multiproduct firms. Recall that brand value consists of two components: reputation and common feature quality. Many extrinsic factors, such as store image, after-sales service, and distribution channels, can affect a firm's reputation. If data on these variables are available, we could either enter them directly into the choice model, or

[19] The product characteristics of other firms continue to serve as instruments for advertising expenditure. I restrict the scope to seven PC vendors because the *Ad $ Summary* report does not record the advertising expenses of the remaining vendors in all the years. Three more vendors (Dell, Leading Edge, and Samsung) can be included if the study window is adjusted to 1988–1993, but the conclusions do not change with this sample. The results are also similar if I add in other non-computer-related advertising expenses.

[20] I have also tested a few alternative specifications. These include: (1) splitting the sample into two sets, 1985–1989 and 1990–1993, and allowing the brand value parameters to change over time; (2) leaving out the small number of Pentium observations to focus on CPUs that were, at that time, more stable and mature; (3) reversing the GEV nest structure by grouping the CPU within brand; and (4) allowing the price elasticity parameter, α, to vary with the studied brands and CPUs. In all these cases, the second-stage GLS and nonparametric tests give similar results, with p values almost always less than 0.01. I omit the detailed statistics for brevity.

regress the estimated brand values on these firm-level attributes to see if they play a significant role in building reputation. I have included one such attribute—advertising expenditure—in this study, and found some evidence that it might help enhance a firm's reputation. It is certainly challenging, but valuable for future research, to use similar attributes to calibrate empirical models of product differentiation with a brand component.

Brand value may also relate to common feature quality and vertical product-line extent. Using hedonic analysis, Randall et al. (1998) find that a brand's price premium is related to its product quality; consumers tend to attach favorable brand values to firms that produce high-quality products. I could not directly replicate the hedonic analysis of Randall et al. because of the lack of PC specification data. However, in a vertically differentiated market, product qualities should increase with prices to sustain positive market shares for all products (Moorthy 1984). By treating price as an indicator of product quality, I could assess the relationship between brand value, common feature quality, and vertical product-line extent using the brand value estimates obtained from the GEV choice model and the IDC PC price data.

In theory, because brand value includes reputation and feature commonality, it should correlate positively with the firm-level average PC price. Further, if vertical product-line extent does contribute to brand value, then the price of a firm's most expensive PC model should positively influence the firm's brand value, and this positive relationship should persist even if the common feature component is removed from the price data (Randall et al. 1998).

To test these hypotheses, I subtract the average from maximum PC prices to remove any common feature qualities for each studied vendor, and use the transformed variable as an indicator for vertical product-line extent. I then regress the estimated brand values (Table 3) separately on the average PC price and this transformed variable for each CPU category. I exclude the first-generation CPUs and Pentium from this analysis because very few vendors sold PCs with these two types of processors. As before, the standard errors of brand values obtained from the GEV model are used as least squares weights.

The regression results largely support the above analysis. Brand value and the average PC price are positively correlated ($p < 0.01$ for 80286 and 80386; $p < 0.06$ and 0.02 for 8086/88 and 80486, respectively). Also, in line with the study of Randall et al. (1998), the price of the most expensive PC model positively affects brand value ($p < 0.01$ for 80286; $p < 0.02$ for all other categories), and some of these positive influences persist even after the common feature quality is removed from the price data ($p < 0.05$ and 0.08 for 8086/88 and 80486; $p < 0.41$ and 0.11 for 80286 and 80386, respectively). These preliminary analyses provide face validity that my brand value estimates capture common feature quality. They also support the hypothesis that premium products help enhance brand value.

Finally, as discussed earlier, a firm may purposely extend its product line for strategic reasons, such as market preemption or to signal product quality (Bayus and Putsis 1999, Schmalensee 1978, Wernerfelt 1988). These may help improve the firm's reputation; therefore, it is theoretically possible that multiproduct firms may proliferate their products so as to build up brand value. To test this conjecture, I perform a split-sample analysis as in Footnote 20 (1985–1989 versus 1990–1993), and examine the differences in brand value and the average number of unique PC models between the two periods of observations.

Because every GEV choice model has a scale parameter that is not identifiable (McFadden 1978), the brand value estimates of the two periods are not directly comparable. Instead, I regress the second-period brand values on the first-period ones (with the corresponding standard errors as weights), and use the residuals as an indicator for brand value changes. The residuals are then regressed on the *change* in the average PC line width, which gives

$$\begin{aligned}\Delta(\text{brand value}) \\ &= -1.262 + \underset{(0.044)}{0.104} \cdot \Delta(\text{average no. of models.}) \qquad (18)\end{aligned}$$

The coefficient of the independent variable is positive and significant ($p < 0.03$), which indicates that product proliferation helps firms enhance their brand values. Therefore, although a branded multiproduct firm may receive decreasing returns in the short run from product variety, a proliferation strategy may yield long-term benefit in strengthening brand value. It is important for firms to recognize the multiple, qualitatively different demand implications of product-line extension.

6. Conclusions

This study shows that product similarity varies with brand value; the higher a firm's brand value, the more similar are its products from the demand perspective. It is imperative to note that my study has focused solely on the demand side of the market. A complete analysis of the benefit to product variety should take into account supply-side factors, such as manufacturing overhead or variable production costs (Lancaster 1990). Previous research has found that a wide product line may decrease the return on investment of manufacturing plants; product variation may increase production costs and reduce overall productivity (Fisher and Ittner 1999, MacDuffie et al. 1996, Swamindass et al. 1999). These seem to imply an unfavorable cost implication of product variety for tangible goods manufacturers.

Interestingly, setting aside manufacturing costs, Bayus et al. (2003) find that product-line extension (new product introduction) increases PC firms' profitability through reductions in selling, general, and administrative (SG&A) expenses; and other marketing and advertising costs. Their findings add another valuable perspective to the continuing debate on the value of product variety. To recapitulate, although product proliferation may increase production costs, it could potentially help a firm establish or maintain brand reputation, reduce marketing costs, and most importantly, expand demand and market share. This last benefit could be considerably impeded by cannibalization, and my study augments this fruitful area of research by highlighting the (often ignored) role of brand value in shaping product similarity. Clearly, future research needs to examine more closely the trade-off between these counteracting effects of product proliferation in an integrated analysis.

Acknowledgments

The author thanks the department editor, the associate editor, two anonymous reviewers, and Teck Ho, Tom Lee, and Ivan Png for their helpful comments and suggestions. This research was supported by the National University of Singapore research grant R253-000-017-112.

References

Ad $ Summary Report. 1985–1993. Leading National Advertisers.

Angel, David P., James Engstrom. 1995. Manufacturing systems and technological change: The U.S. personal computer industry. *Econom. Geography* **71**(1) 79–102.

Balachander, Subramanian, Sanjoy Ghose. 2003. Reciprocal spillover effects: A strategic benefit of brand extensions. *J. Marketing* **67**(1) 4–13.

Bayus, Barry L. 1998. An analysis of product lifetimes in a technologically dynamic industry. *Management Sci.* **44**(6) 763–775.

Bayus, Barry L., William P. Putsis, Jr. 1999. Product proliferation: An empirical analysis of product line determinants and market outcomes. *Marketing Sci.* **18**(2) 137–153.

Bayus, Barry L., Gary Erickson, Robert Jacobson. 2003. The financial rewards of new product introductions in the personal computer industry. *Management Sci.* **49**(2) 197–210.

Berndt, Ernst R., Zvi Griliches, Neal J. Rappaport. 1995. Econometric estimates of price indexes for personal computers in the 1990's. *J. Econometrics* **68**(1) 243–268.

Berry, Steven T. 1994. Estimating discrete-choice models of product differentiation. *RAND J. Econom.* **25**(2) 242–262.

Berry, Steven T., James Levinsohn, Ariel Pakes. 1995. Automobile prices in market equilibrium. *Econometrica* **63**(4) 841–890.

Boush, David M., Barbara Loken. 1991. A process-tracing study of brand extension evaluation. *J. Marketing Res.* **28**(1) 16–28.

Bresnahan, Timothy F., Scott Stern, Manuel Trajtenberg. 1997. Market segmentation and the sources of rents from innovation: Personal computers in the late 1980s. *RAND J. Econom.* **28**(0) S17–S44.

Chakravarty, Amiya K., Nagraj Balakrishnan. 2001. Achieving product variety through optimal choice of module variations. *IIE Trans.* **33**(7) 587–598.

Chong, J., Teck Ho, Chris S. Tang. 1998. Product structure, brand width, and brand share. Teck Ho, Chris S. Tang, eds. *Research Advances in Product Variety*. Kluwer Academic Publishers, Boston, MA.

Dacin, Peter A., Daniel C. Smith. 1994. The effect of brand portfolio characteristics on consumer evaluations of brand extensions. *J. Marketing Res.* **31** 229–242.

Desai, Preyas S. 2001. Quality segmentation in spatial markets: When does cannibalization affect product line design. *Marketing Sci.* **20**(3) 265–283.

Desai, Preyas S., Sunder Kekre, Suresh Radhakrishnan, Kannan Srinivasan. 2001. Product differentiation and commonality in design: Balancing revenue and cost drivers. *Management Sci.* **47**(1) 37–51.

Fisher, Marshall L., Christopher D. Ittner. 1999. The impact of product variety on automobile assembly operations: Empirical evidence and simulation analysis. *Management Sci.* **45**(6) 771–786.

Fisher, Marshall L., Kamalini Ramdas, Karl Ulrich. 1999. Component sharing in the management of product variety: A study of automotive braking systems. *Management Sci.* **45**(3) 297–315.

Gilbert, Richard J., Carmen Matutes. 1993. Product line rivalry with brand differentiation. *J. Indust. Econom.* **41**(3) 223–240.

Hartman, Raymond S., David J. Teece. 1990. Product emulation strategies in the presence of reputation effects and network externalities: Some evidence from the minicomputer industry. *Econom. Innovation New Tech.* **1**(1/2) 157–182.

Hyslop, Dean R., Guido W. Imbens. 2001. Bias from classical and other forms of measurement error. *J. Bus. Econom. Statist.* **19**(4) 475–481.

Kamakura, Wagner A., Byung-Do Kim, Jonathan Lee. 1996. Modeling preference and structural heterogeneity in consumer choice. *Marketing Sci.* **15**(2) 152–172.

Katz, Michael L. 1984. Firm-specific differentiation and competition among multiproduct firms. *J. Bus.* **57**(1 part 2) S149–S166.

Kekre, Sunder, Kannan Srinivasan. 1990. Broader product line: A necessity to achieve success?" *Management Sci.* **36**(10) 1216–1231.

Kim, Kilsun, Dilip Chhajed. 2000. Commonality in product design: Cost saving, valuation change and cannibalization. *Eur. J. Oper. Res.* **125** 602–621.

Kim, Kilsun, Dilip Chhajed. 2001. An experimental investigation of valuation change due to commonality in vertical product line extension. *J. Product Innovation Management* **18** 219–230.

Kirmani, Amna, Peter Wright. 1989. Money talks: Perceived advertising expense and expected product quality. *J. Consumer Res.* **16**(3) 344–353.

Krishnan, V., Saurabh Gupta. 2001. Appropriateness and impact of platform-based product development. *Management Sci.* **47**(1) 52–68.

Lancaster, Kelvin. 1990. The economics of product variety: A survey. *Marketing Sci.* **9**(3) 189–206.

Langlois, Richard N. 1992. External economies and economic progress: The case of microcomputer industry. *Bus. Hist. Rev.* **66**(1) 1–50.

Lee, Hau L., Chris S. Tang. 1997. Modelling the costs and benefits of delayed product differentiation. *Management Sci.* **43**(1) 40–53.

MacDuffie, John Paul, Kannan Sethuraman, Marshall L. Fisher. 1996. Product variety and manufacturing performance: Evidence from the international automotive assembly plant study. *Management Sci.* **42**(3) 350–369.

McFadden, Daniel. 1978. Modelling the choice of residential location. A. Karlqvist, L. Lundqvist, F. Snickars, J. Weibull, eds. *Spatial Interaction Theory and Planning Models*. North-Holland, Amsterdam, The Netherlands, 75–96.

Moorthy, K. Sridhar. 1984. Market segmentation, self-selection, and product line design. *Marketing Sci.* **3**(4) 288–307.

Nelson, Randy A., Tim L. Tanguay, Christopher D. Patterson. 1994. A quality-adjusted price index for personal computers. *J. Bus. Econom. Statist.* **12**(1) 23–31.

Park, C. Whan, Sandra Milberg, Robert Lawson. 1991. Evaluation of brand extensions: The role of product feature similarity and brand concept consistency. *J. Consumer Res.* **18**(2) 185–193.

Putsis, William P., Barry L. Bayus. 2001. An empirical analysis of firms' product line decisions. *J. Marketing Res.* **38**(1) 110–118.

Randall, Taylor, Karl Ulrich. 2001. Product variety, supply chain structure, and firm performance: Analysis of the U.S. bicycle industry. *Management Sci.* **47**(12) 1588–1604.

Randall, Taylor, Karl Ulrich, David Reibstein. 1998. Brand equity and vertical product line extent. *Marketing Sci.* **17**(4) 356–379.

Reddy, Srinivas K., Susan L. Holak, Subodh Bhat. 1994. To extend or not to extend: Success determinants of line extensions. *J. Marketing Res.* **31**(2) 243–262.

Robertson, David, Karl Ulrich. 1998. Planning for product platform. *Sloan Management Rev.* **39**(4) 19–31.

Schmalensee, R. 1978. Entry deterrence in the ready-to-eat breakfast cereal industry. *Bell J. Econom.* **9**(2) 305–327.

Small, Kenneth A. 1987. A discrete choice model for ordered alternatives. *Econometrica* **55**(2) 409–424.

Smith, Daniel C., C. Whan Park. 1992. The effects of brand extensions on market share and advertising efficiency. *J. Marketing Res.* **29**(3) 296–313.

Stavins, Joanna. 1995. Model entry and exit in a differentiated-product industry: The personal computer market. *Rev. Econom. Statist.* **77**(4) 571–584.

Swamidass, Paul M., Satish S. Nair, Sanjay I. Mistry. 1999. The use of neural factory to investigate the effect of product line width on manufacturing performance. *Management Sci.* **45**(11) 1524–1538.

Swaminathan, Jayashankar M., Sridhar R. Tayur. 1998. Managing broader product lines through delayed differentiation using vanilla boxes. *Management Sci.* **44**(12 part 2) S161–S172.

Venkatesh, R., Vijay Mahajan. 1997. Products with branded components: An approach for premium pricing and partner selection. *Marketing Sci.* **16**(2) 146–165.

Wanke, Michaela, Herbert Bless, Norbert Schwarz. 1998. Context effects in product line extensions: Context is not destiny. *J. Consumer Psych.* **7**(4) 299–322.

Wernerfelt, Birger. 1988. Umbrella branding as a signal of new product quality: An example of signalling by posting a bond. *RAND J. Econom.* **19**(3) 458–466.