IS MORE ALWAYS BETTER? THE EFFECT OF COMPLEXITY ON CHOICE OF BUSINESS PRODUCTS

SAGIT HAREL TAL

SUPERVISED BY: PROF. DAN HORSKY

SIMON SCHOOL OF BUSINESS

UNIVERSITY OF ROCHESTER

ROCHESTER, NY, 14627

JOB MARKET PAPER

JANUARY 1, 2010

IS MORE ALWAYS BETTER? THE EFFECT OF COMPLEXITY ON CHOICE OF BUSINESS PRODUCTS

Abstract

The technological advancements of the last decade brought with them digital products with a large number of features and options. This enhanced functionality, however, turned out to be a double-edged sword: using sophisticated options, keeping up to date with new features and sorting out the features that are needed from those that are just clutter requires customers to invest significant time and effort. Measuring the tradeoff between enhanced functionality and the increase in the complexity of business products is at the heart of this research. In this study a new methodology for quantifying this tradeoff in terms of market shares is proposed. A theoretical construct is developed, and a structural empirical model that lends itself to estimation techniques that utilize available product-level and aggregate consumer level data is presented. In a market validation example I empirically examine the multi- and single functional copiers market in order to demonstrate the effect of complexity on customer adoption as measured by device sales and market share. The findings reveal that the actual purchase behavior manifested in the specific market is concave in complexity and that an optimal level of complexity can be found. Additionally, I find that larger firms have a higher tolerance to complexity in multifunctional devices than smaller firms. Finally, the results of this research suggest that there are potential biases in price and features estimates when a customer preference model does not account for complexity.

Key Words: Random Coefficients Choice Models, BLP Estimation, Complex Products

"If technology exists in order to make our customers' lives easier and more productive, why is it that they often find it complex and frustrating?" A Fortune 100 manufacturer of large office equipment

1 Introduction

In the last decade, the idea of "One Device That Does It All" attracted almost all technology manufacturers. In a process that is often referred to as "feature creep", products with increasing numbers of features, functions, buttons and menus that come with book-length manuals became widely prevalent in the market. Microsoft Word 2003 for example, has 31 tool bars and more than 1500 commands and options. The feature creep process is further amplified by a related process known as convergence of products. A convergence product is a bundle of digital-platform technologies physically integrated onto a common product form that offers consumers alternatives to dedicated product forms. Most cell phones today, for instance, have a camera, GPS, music player, voice recorder, gaming device and other capabilities built in, each of these capabilities comes with its own set of features and functions. Recently, however, there is growing evidence that customers are challenged by the increasingly sophisticated products of the digital revolution. Using sophisticated options, keeping up to date with new features and sorting out the features that are needed from those that are just clutter requires customers to invest significant time and effort. And customers are reacting to this increasing complexity. In 2003 a research by Phillips Electronics found that at least half of returned products have nothing wrong with them, customers just couldn't figure out how to use them (Surowiecki, 2007). A preference for simplicity is revealed in the dominance of Google, whose uncluttered landing page stands in stark contrast to the busier pages of Yahoo or ask.com. Apple, frequently at the leading edge of design and usability, made simplicity its iPod's and iPhone's selling point. In fact, the iPhone Apps store allow users to choose and use the precise mix of applications they want on their devices.

In business and commercial environments, the complexity induced by feature richness issue is even more important as its effect on usage time and learning costs can be crucial. The overhead of time and efforts employees spend interacting with a huge set of features of production devices increases the underlying variable costs of firms. Indeed complexity may even cause threats to safety critical environments. For example, the accident at the Three Mile Island Unit 2 nuclear power plant near Middletown, Pa., on March

28, 1979 was caused by design deficiencies leading to personnel errors and component failures[1]. Thus, in business settings, the *total costs of ownership (TCO)*, some attributable to complexity, are accumulated into the firm's (often pre calculated) marginal costs. Clearly, more features and capabilities drive higher functionality in order to provide increasing benefits to the user. However, the complexity of the devices, an often unintended consequence of the abundance of features added to increase flexibility, may impose costs which are not easily observable at time of purchase and which reduce the anticipated benefit of the device. There may be, therefore, a tradeoff between the benefits to be derived from a device with advanced functionality and the overheads associated with the device's complexity. It is this tradeoff that this study investigates.

The strategy of adding a feature to products has been widely used almost all product categories. The examples are endless – cars, televisions, DVDs, microwaves and dishwasher to name a few, have each been supplied with a huge number of enhancing capabilities. Manufacturers' actions can be seen as reflection of customers' shift in preferences among other things. According to the Kano Model (Kano, 1965) yesterday's new and exciting features become today's necessities and tomorrow's basic requirements whose absence is unacceptable to customers. In order to satisfy migrating customer preferences, manufacturers have to constantly upgrade their products by adding more and more features. Fierce competition drives this process further as salespeople who sell these devices often push for extra features as ways to attract customers and stand up to competitors. Often, features offered in a competing device have to be quickly imitated and incorporated into a manufacturer's product. Differentiation becomes a race of features addition. The decision to add features became easier with the significant decline in production costs of the digital era. Adding new features to digital devices became technically easier, faster and cheaper, and this created an "internal audience" problem: engineers who design the products are usually tech savvy and features they think are necessary and important are not necessarily best for customers. As features are added to the devices, convergence strategies of electronic devices also play a substantial role in the chase after consumers' hearts and pockets. These products offer customers alternatives to the increasing number of dedicated high tech products that are available and are considered necessities in today's market. At the same time, convergence offers manufacturers another opportunity to cut costs by using same product platforms, and

[1] http://www.nrc.gov/reading-rm/doc-collections/fact-sheets/3mile-isle.html

create one device that can be used to target multiple segments. It is not surprising, then, that most digital products found in the market today are equipped with a large set of functions and capabilities.

In recent years however, there have been increasing number of reports in the popular media that inform on customers' satiation with and sometimes adverse reactions to feature rich products. For instance, the Economist's Strap In and Boot Up article (2006) describes the genuine desire of drivers to be done with technical support when they sit behind the wheels. The Wall Street Journal (Wingfield, 2006) published a special report advocating the need for simplicity in consumer electronic products. Few manufacturers are just now beginning to recognize and address the complexity issue in various ways like removing, grouping, hiding or displacing features. For example, The New York Times recently reported (Pogue, 2009) how the new operating systems of Apple and Microsoft are "cleaned-up, slimmed down versions of what came before" as the software giants attempt to correct their strategy of overloading their previous operating systems with, what is now called, "bloatware". However, the strategies applied by manufacturers lack quantitative or predictive tools that provide guidelines on how to identify the extent of the complexity effect on share, and how to produce products with an optimal blend of features.

In their efforts to understand customers' marketers and product designers have used classical theories of consumer choice. The traditional theory of choice behavior rests on the view that consumers attempt to maximize the utility they obtain from a good purchased in the marketplace. In the multi-attribute framework, the utility that a customer derives from a device can be expressed as a combination of the device's attributes and price as well as the customer's idiosyncratic tastes, needs and abilities. The attributes of a device contain **subjective attributes** (such as perceived style or brand name) and **objective attributes** which is the set of all the characteristics and features of the device. **Characteristics** are traits of the device that are set by the manufacturer and are not controlled by the user, such as memory capacity or size of a copying device. **Features** are functions or capabilities that the device offers that the user can manipulate and decide if and how to use while operating the device. For example, a copier may have a feature which enables duplexing (double sided capability), or a feature that enables the copying of an open book to two separate pages. The **complexity** examined in this research has its origin in the features that a device has and is therefore defined as a function of this features set. Graphically, the decision process of the customer is depicted in Figure 1.

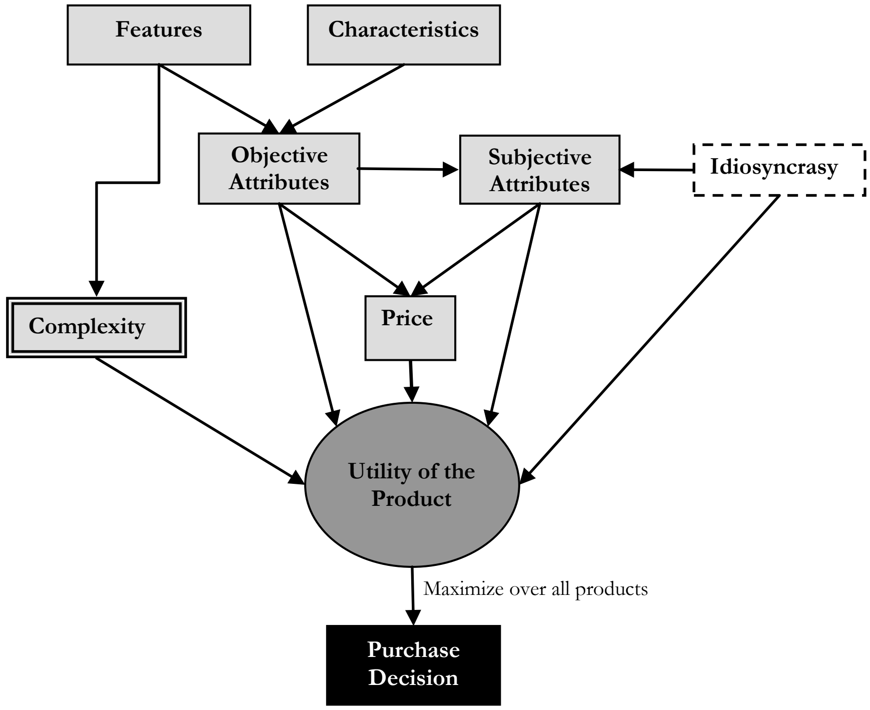

Figure 1 – A View of The Customer's Decision Process

Formally, the customer's decision process can be specified as the following maximization problem:

$$\underset{j \in J}{Max}\, U = U\left(x_j, y_j, z_j, G(x_j), P(x_j, y_j, z_j), e\right) \tag{1.1}$$

where *J* is the set of available devices, U is the utility that a customer derives from purchasing device j, x_j is the set of features the device offers the user, y_j is the characteristics set of device j, z_j is the set of subjective attributes, $G(x_j)$ is the device's complexity level, induced by it's capabilities or the blend of features and functions the device offers, and P is the device's price which, through the manufacturing cost mechanism, is a function of features, characteristics and the subjective attributes. Lastly, e represents the customer's idiosyncratic tastes, needs, and abilities that affect the perceived utility. Graphically, the decision process is depicted in Figure 1.

In light of the formulation of the decision process, the widely observed persistence of the feature/function addition strategy is a puzzling phenomenon. Manufacturers usually pay careful attention to what customers say they want. They use sophisticated marketing strategies and integrated product design tools to target the right consumers for the right product. Decisions about product design are often based on multi-attribute utility models. Why is it then that customers so often find these products too complex, overloaded and unmanageable? Was the strategy of adding features good for a while and then turned sour as the number of features became too large for customers to handle? When does the marginal utility from adding another feature offset the marginal price and usage costs incurred by this addition? Is there an optimal number of features that should be incorporated into a device? The purpose of this study is to provide theoretical and empirical frameworks for quantifying the tradeoff between complexity and enhanced functionality.

The marketing literature that addresses the effect of complexity induced by features richness on customer's behavior is surprisingly scarce. The behavioral studies were the first to investigate the existence of the adverse reaction to the increase in the number of features and to offer some explanations to this phenomenon. This body of work focuses only on partial sets of features and on the complexity induced by increasing the number of these features. The original work of Thompson et. al (2005) shows that too many features can make the consumer overwhelmed, thus suggesting there exists an optimum for the preferred number of features in a device. The explanation to this feature fatigue lies, according to this study, in the gap between what people say or think they want *before* they own the product, and what people really want *after* they use it. This research finds that customers tend to choose overly complex products because a product with more features is considered of better quality. Only after using the product do they find that it does not maximize their satisfaction. Other behavioral works such as Zhao et. al (2005) and Meyer et. al. (2008) support this stream of thought finding an upwardly-biased valuation for the new sets of features in the context of video game platforms. They show that customers fail to fully use the features after purchase and this use pattern is due to the difference in valuation of new capabilities at the time of purchase and after use. Specifically, Meyer explains that while usage decisions are driven by short term aspects like reducing the learning costs, purchase decisions don't take these usage costs into account and are usually driven by an optimistic belief about the usefulness of the new capabilities. A different validation for customer's reaction to complexity is offered by Zhang et. al (2007) who report that when means are connected to multiple goals they are less likely to be chosen and are pursued only when one of these goals is activated. In the context of feature rich devices, Zhang's findings suggest that a device that 'does it all' may be less attractive than its simpler counterpart. Additional support to customers' preference for simplicity is found in the recent study by Hans. et. al (2009) who find that technological performance is inversely related to the preference for converged products over the dedicated options. All of these studies support the argument that too many features and functions have a negative effect on customer satisfaction.

The main shortcoming of this new stream of behavioral research is that all of these studies are held in lab and experimental settings with measurements through surveys and do not capture the actual behavior of consumers in the marketplace. It is important to examine how actual purchase behavior is affected by complexity through sales analyses because there may be differences between findings implied by stated preference and real purchase behavior. In fact, Horsky and Nelson (1990) and Horsky et. al (2004) suggest that there are differences between customers' attribute weights generated by value elicitations through stated preference and the attribute weights measured by revealed preference choice analyses. Horsky et. al report that tangible attributes are weighted more heavily than intangible attributes in revealed choice as compared to in value elicitation. Since complexity has its origin in tangible attributes, it may be an important factor in the customer's decision process at the time of purchase.

One important aspect that is further neglected in the behavioral literature is price. Feature rich complex products tend to be more costly to produce and are therefore priced higher than products with fewer features. Another drawback of the existing behavioral literature is that while all researches find that more is not necessarily better, they do not provide an answer to a crucial question: When should manufacturers stop adding features to their products? Clearly, the strategy of adding features is not always wrong; it might be so just after a certain optimal point of a functionality and complexity blend. Yet, none of the behavioral studies provide any exploration for the presence of an "optimum" let alone a general method for deriving it. Heterogeneity in customers' sensitivity to complexity is one more aspect that received little consideration in the behavioral literature. It may be that products of different complexity levels will be suited for different customers, perhaps with different skill sets, and this may be an appropriate segmentation strategy. For example, an academic administrative assistant can operate a more complex copying device better than a student who seldom uses the device. Finally, the behavioral studies that examine different aspects of complexity effect on product choice focus on consumer products, mostly electronic. Very little attention has been given to the effect of complexity in industrial or commercial product domains, which is surprising since these domains make up the largest component of the US economic market. In sum, while the behavioral researchers were the first to approach the topic of complexity induced by enhanced functionality, their studies present conservative tests of their hypotheses and further research on actual purchase behavior is warranted.

In the quantitative literature, the issue of complexity has not yet been explicitly addressed. Some subtle aspects of the tradeoff between ease of use and effectiveness can be found in the traditional multi-attribute literature. These early choice models were based on a parsimonious set of subjective attributes, price and economic constraints. Often, perceptual attributes such as 'ease of use' or 'comfort' as well as 'performance' or 'effectiveness' were found to have significant effects on choice (Horsky and Nelson, 1992; Hauser and Shugan, 1980), and different techniques suggested ways to link these identified primary customer needs with engineering characteristics of new products (see for example, Hauser and Clausing's 1988 House of Quality). These findings may point to the coexistence of two factors – capabilities and complexity, that these days are reported to have opposing effects on utility in certain scenarios. However, the classic methods lack the ability to quantify this tradeoff into a metric that translates into sales or share. Moreover, these methods do not examine how the number of features offered by a product is factored into this tradeoff. In particular, since the "ease of use" subjective attribute is related to the whole product rather than to each of its functions, the impact that *each* feature has on this construct is not examined. In fact, treating "ease of use" as another independent component of a utility function misses any interaction effects which may exist in the presence of proliferation of features that may lead to increasing complexity.

The modern quantitative choice literature does not explicitly account for complexity or complexity related attributes. Starting in the early 1990's, observed choice data was widely available to marketers in their studies of purchase behavior. Individual and aggregate, cross sectional and panel, scanner and market data sets became industry standard in calibrating choice models and nailing down choice attributes. The availability of the revealed preference data (objective product attributes, prices, sales, and demographics) and the increasing computational power led to innovative enhancements that build on the early work of

McFadden's (1974) discrete choice mode. Structural and NEIO[2] models along with fast estimation techniques were further developed and deployed throughout the field (Berry, 1994; BLP 1995; Dube et. al, 2009). These models, however, focus only on objective features, characteristics and price as main effects; subjective attributes were no longer used, and even the "ease of use" effect disappeared. Researchers, moving into the objective attribute space, assign the subjective aspects to different components of the unobserved error terms (Dube et. al, 2002) or argue that there exists a high correlation between the perceptual dimensions and physical characteristics (Agarwal and Ratchford, 1980). The technological advancements of the last decade also drove the developments of new conjoint methods that were capable of dealing with products that have a massive number of features and characteristics (see Rao et. al, 2009 for a review of these methods). These methods try to explore parsimonious techniques that avoid the need to employ a large number of product profiles in conjoint settings. However, in all of these conjoint methods the underlying assumption is that each additional feature has a positive effect on utility. Two way interaction effects between some features are infrequently considered in different conjoint designs, but there are very few studies with designs of three or higher order interactions. Consequently, neither the complexity induced by feature richness, nor optimal number of features are explicitly taken into consideration in the process of product design through conjoint studies.

Recently, new studies report that focusing only on objective attributes can be insufficient since customers' preferences are in fact driven by **both** groups of factors (Ashok et. al, 2002; Luo et. al, 2008; Horsky et. al, 2009; Temme et. al, 2008). It is not surprising then that the long recognized need to incorporate subjective factors into product choice and design models has been recently revisited in the marketing literature. The trend of using both subjective and objective attributes is further strengthened by new techniques that utilize existing datasets augmented by individual level stated preference data (Horsky et. al 2006). The new techniques offer ways to integrate subjective constructs ('ease of use' among them), features, characteristics and price into quantitative choice and conjoint models. Nonetheless, none of these innovative techniques investigate the effect of complexity, in particular the complexity induced by feature richness, on choice.

[2] New Empirical Industrial Organization

Complexity therefore is an important yet elusive decision factor that has been overlooked by most of the marketing literature. Only a handful of behavioral studies examine limited aspects of the influence of complexity on customer stated preferences and satisfaction. The quantitative literature touches this topic only marginally and a complexity construct does not explicitly appear in these models. Indeed none of the existing studies exhibit a complete approach that facilitates a thorough investigation of the complexity effect and its actual manifestation in the marketplace. Incorporating complexity into quantitative models is likely to improve our understanding of the underlying customer purchase behavior.

In this paper, I develop a comprehensive model that examines the effects of features, complexity and price in a heterogeneous demand system. I provide a theoretical framework, formulate a structural choice model, offer an appropriate empirical methodology and demonstrate a market validation example. The empirical methodology I employ utilizes aggregate data that is augmented by individual level data to better represent customers' preferences. This research is the first to quantitatively examine how the complexity effect is manifested in preference for industrial devices in business environments. In this study, I show that the revealed purchase behavior of business products demonstrates a concave complexity effect: as the number of features increases, the expected utility (and share) increases and then decreases after some optimal point. Additionally, I suggest that the failure to account for the complexity effect in analytical methodologies may produce biased preference estimates. For instance, price induces a negative effect on utility. Since feature rich products tend to be more complex and more expensive, the price effect may be compromised if complexity is not accounted for. Lastly, the methodology I propose is can be used for both industrial and consumer markets and provides a tool set for finding the optimal profit maximizing level of complexity. It can therefore be very useful to manufacturers in product design stages.

The remainder of the paper is organized as follows: In Section 2 I develop a theoretical model of complexity effect on choice in business environments. In Section 3 I translate the theoretical model into a structural empirical model and describe the estimation methodology I use. Section 4 presents the particular business market that is investigated in this study and details the data and variables used to calibrate the empirical model. I discuss the results in Section 5 and conclude with future research ideas and managerial implications in Section 6.

2 A Theoretical Model of Complexity Effect on Choice of Business Products

2.1 A Model of Business Choice

Businesses are constantly looking for ways to improve their production processes, reduce their costs and differentiate their outputs from their competitors'. One way to do so is to purchase production inputs that aid in achieving these goals. Consequently, the organizational buying models reflect demand which is *derived* demand: purchases are made in order to meet customers' demand with minimum costs. The benefits from purchasing the optimal industrial device include an increase in the efficiency of the production process and improvement of the quality of end product - both of which are separately linked to an increase in the expected profit of the firm (Horsky and Mohanty, 2009). The efficiency boost can be achieved by purchasing a device that has the potential to make the operations of the firm faster, easier, less labor intensive and therefore less expensive. The improvement of the end product quality can either be a direct product of the device (for example, brochures printed on a color laser printer are better than ones printed on a color inkjet device), the way the device works (for example, automation of booklet creation by a multifunctional copier decreases the probability of errors), or the way the device's outputs are incorporated into the firm's operations (speeding up the business process by printing receipts faster). Devices with superior characteristics and more features are intuitively associated with a higher potential to increase efficiency and improve output quality thus driving the expected derived profits upward.

The benefits, however, do not come without costs. When a business purchases an industrial device, the total cost of ownership (henceforth TCO) needs to be taken into account. The TCO includes fixed costs (price of the device, software and hardware costs, floor space, etc..) and variable costs (supplies costs, training costs, help desk support, users' work time, waste disposal, costs of faulty output, etc..). While some of these costs are universally accepted as real costs, others are considered 'soft' costs where standard accounting may not be easily applied to them. User's interaction time with the device is a natural example of a 'soft cost' that is not easily quantified. A comparison between the TCO of more sophisticated devices and the TCO of simpler ones reveals that the former are usually associated with higher costs. Not only are they more expensive, but the numerous features they offer usually impose a heavier load on their processing unit which means that these devices are often slower. More importantly, there is a significant yet elusive increase in usage costs in feature rich devices. There are more functions to maintain, the probability for feature malfunctions increases and supplies of more sophisticated devices tend to be more expensive. Additionally, the cost of employees' time interacting with or working on the device increases: the training time associated with learning how to operate the device and the time it takes to perform different tasks is higher as the number of functions increases.

The higher costs associated with feature richness may mitigate the potential gains from more sophisticated devices. Thus, when a firm decides which device to buy, given that a decision to buy a device has already been made, it needs to consider all the benefits and costs of each device, and purchases the device that maximizes its profits. Simply put, the firm's decision process can be specified as:

$$\underset{j \in J}{Max} \quad \Pi_j = R_j - C_j \tag{2.1}$$

where *J* is the set of all devices in the choice set of the firm, R_j is the revenue, or benefit, from device *j*, C_j is the TCO of device *j*, and Π_j is the profit from device *j*. In the optimization specified above, the revenue and costs represent an average of the perceived revenues and costs by the individuals who are involved in the institutional decision process, as well as the perceptions of the actual users who might not participate in the purchase decision. In other words, I assume that the firm's profit function represents the aggregate preferences of the individuals who are driving the purchase decision.

To accommodate decision choices in markets with large choice sets, I adopt Rosen's indirect utility type model (1974) which extends Lancaster's (1971) multi-attribute approach to indivisible and infrequently purchased items. The revenue and costs of each device are therefore defined as functions of the device's features, characteristics, price and other intangible factors. Specifically, the benefits, or revenue from buying device *j* are a function of what it does (i.e. its capabilities or features), how it does it (its characteristics set) and other intangible factors brand name, style, etc…Formally, the revenue of a device is defined as:

$$R_j = R\left(x_j, y_j, z_j\right) \tag{2.2}$$

where x_j is a vector of N_1 zero-one values, each representing the presence or absence of a specific feature in device *j*. For example, the first element in x_j can represents the presence or absence of duplexing capability in a copying device. Similarly, y_j is a vector of N_2 values that represent the magnitude or intensity of each characteristic of device *j*. For instance, the first element in y_j can represent the size of a device; the second element can represent the speed, etc… Lastly, z_j is a set of N_3 intangible attributes which not captured by the tangible features or characteristics.

Similarly, the TCO incurred by the firm from purchasing device *j* can be expressed as:

$$C_j = C\left(P(x_j, y_j, z_j), T(x_j), x_j, y_j, w_j\right) \tag{2.3}$$

where $P(x_j, y_j, z_j)$ is the device's price, which, due to manufacturing costs is a function of the devices features, characteristics and intangibles. $T(x_j)$ represents the usage time costs, or the cost of the time employees spend interacting with the device in order to complete different tasks required for the production process of the firm. It is modeled as a function of all the features the device offers and includes the costs of all time-consuming tasks related to the device such as training and learning, help desk time, service, maintenance and repair, and, most importantly, the actual interaction and activation time. The TCO is also a function of the device's features, x_j, (for example, if a multifunctional printer has a photo mode feature, then supplies costs includes the price of special photo paper required for this capability), the device characteristics, y_j, (for example, the device's size incurs a rent for office space), and other intangible factors that effect cost such as waste disposal and costs of faulty output (w_j). In the static setting, the benefits and costs are totaled over the life of the device. Combining Equations (2.1) - (2.3), the profit from device *j* is:

$$\underset{j \in J}{Max} \quad \Pi_j = R\left(x_j, y_j, z_j\right) - C\left(P(x_j, y_j, z_j), T(x_j), x_j, y_j, w_j\right) \tag{2.4}$$

2.2 Modeling the Effect of Complexity on Choice

While usage time concept is explicitly defined in Equation (2.3), it is also implicitly present in the benefit part of the profit function (R_j) through the way the device enhances efficiency of production, simplifies operations and shortens tasks' time. Time, therefore, introduces two opposing forces that influence the firm's choice. On one hand, each feature has the potential to increase the efficiency of the production process and shorten interaction time. On the other hand, task time may increase as the number of functions increases or as the operation of each feature becomes more complicated. The idea of the device's complexity factors into the model through the time component that is present in the cost function. It may mitigate or even outweigh the benefits of features which are accounted for in the benefit part of the utility. To explicitly model the complexity of a device, the time component is decomposed into two parts:

$$T\left(x_j\right) = T_1\left(x_j\right) + \Delta T\left(G(x_j)\right) \tag{2.5}$$

Under the assumption that each feature of the device has a basic activation or operating time net of the other features in the device, the first part of the time component, $T_1(x_j)$, represents these basic activation times. Additionally, the time it takes to operate a feature may be increased by the presence of other features. In other words, complexity due to proliferation of features may impose additional usage time. For instance, to create multiple copies, the user has to sort out through different menus or disable other features such as stapling and duplexing. To model this additional time, I assume that each device has a complexity level which is a function of its features, $G(x_j)$. Thus, $\Delta T\left(G(x_j)\right)$ represents the additional time incurred by this complexity level.

The prospective business customer's maximization problem is to choose the device with the right mix of characteristics, features and complexity that maximizes its profits:

$$\begin{aligned} &\underset{j \in J}{Max} \quad \Pi_j = R\left(x_j, y_j, z_j\right) - C\left(P(x_j, y_j, z_j), T(x_j), x_j, y_j, w_j\right) \\ &where \\ &\qquad T\left(x_j\right) = T_1\left(x_j\right) + \Delta T\left(G(x_j)\right) \end{aligned} \tag{2.6}$$

3 The Empirical Model of the Effect of Complexity on Choice

The theoretical model needs to be transformed into an empirical model that enables the estimation of the effect of complexity on purchase choice and market share. In Section 3.1, I describe the transition from the theoretical space into the empirical one. In Section 3.2 I describe the identification of the complexity effect in the empirical model. In Section 3.3 I derive the empirical market share formulation. In Section 3.4 I discuss the endogeneity of price.

3.1 The Empirical Specification

I move from the theoretical space to the empirical one by assuming that the firm's cost and revenue are linear functions and additively separable. Additionally, since different businesses have different revenue and cost structures, different decision making processes and different preferences, I allow the revenue and cost functions to represent heterogeneous consumers. Thus, customer *i*'s revenue and cost from device *j* can be defined as:

$$
\begin{aligned}
&R_{ij}\left(x_j, y_j, z_j\right) = x_j'\gamma_{r,i} + y_j'\beta_{r,i} + z_j\delta_{r,i} \\
&C\left(P(x_j, y_j, z_j), T(x_j), x_j, y_j, w_j\right) = \alpha_i p_j + T(x_j) + x_j'\gamma_{c,i} + y_j'\beta_{c,i} + w_j\delta_{c,i} \\
&T\left(x_j\right) = T_1\left(x_j\right) + \Delta T\left(G(x_j)\right)
\end{aligned}
\tag{3.1}
$$

Where: x_j, y_j, z_j, p_j, and w_j are the same constructs that are defined in the theoretical model. α_i represents the customer specific price effect, $\gamma_{r,i}$ and $\gamma_{c,i}$ are the customer specific weights of the features on revenue and costs, respectively, $\beta_{r,i}$ and $\beta_{c,i}$ are the customer specific weights of the characteristics on revenue and costs, respectively, and $\delta_{r,i}$ and $\delta_{c,i}$ represent the customer specific effects of the intangible and subjective attributes on revenue and cost.

To model the usage time $T\left(x_j\right)$, I start by explicitly defining the complexity score of device *j*, $G\left(x_j\right)$. To that effect, I adopt Shugan's Cost of Thinking (1980) idea and define device *j*'s complexity score as a weighted sum of the values of the vector x_j. The weights represent actual difficulty levels associated with the activation of each feature:

$$G\left(x_j\right)=\sum_{k=1}^{N_1} d_k x_{kj}$$
$$where \quad d_k \geq 1 \quad \forall k \tag{3.2}$$

In the case where all weights are identical and equal to one, the device complexity is a simple feature count; more complex devices are simply devices that have richer feature sets. However, since not all features are equally complex, even by themselves, the weights are needed, and features which are more difficult to use will have a weight which is higher than 1. Hence two devices may have the same number of features but different complexity scores.

Next, I allow the effect of complexity on profit to have a flexible polynomial form:

$$\Delta T\left(G(x_j)\right)=\sum_{l}\theta_{il}\left(\sum_{k=1}^{N_1} d_k x_{kj}\right)^l \tag{3.3}$$

The case where $l=2$ allows for an increasing, a diminishing returns or a concave effect of complexity on the profit function as can be seen in Figure 2. The case of an increasing function supports the idea that more is always better. In other words, it may reflect the belief the more features a device has, the better it is. Diminishing returns is in line with classic economic theory – more inputs to production function will, at some point, show diminishing returns. A concave effect demonstrates the existence of complexity as I define it: more features, after some point, have a negative effect on profits because, for instance, their existence creates complexity - a costly interaction effect of all features.

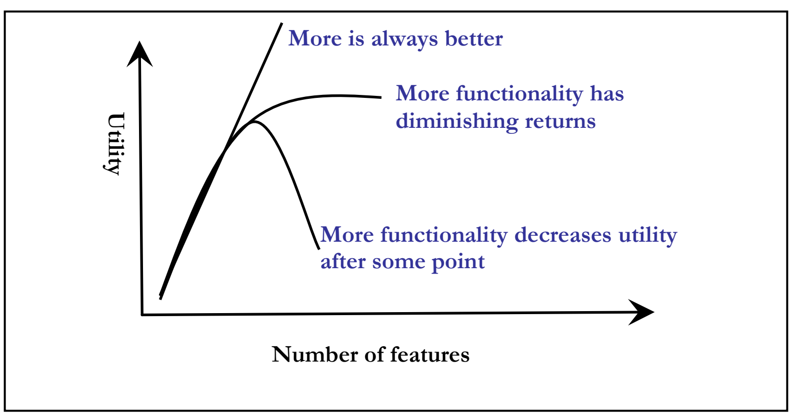

Figure 2 - Possible Effects of the Number of Features on Utility

The last part of the model is the time incurred by activation of each feature excluding the complexity effect. Since it is modeled as a linear function of the features, its effects are already captured in $\gamma_{c,i}$ in

Equation (3.1). I collect the cost and revenue effects (which are not separately identifiable) to get a simplified profit function[3]:

$$\Pi_{ij} = x_j'\gamma_i + y_j'\beta_i - \alpha_i p_j - \theta_{i1}\left(\sum_{k=1}^{N_1} d_k x_{kj}\right) - \theta_{i2}\left(\sum_{k=1}^{N_1} d_k x_{kj}\right)^2 + (z_j, w_j)'\delta_i \tag{3.4}$$

Since the researcher does not have the complete information set (in particular, attributes and subjective preferences), I adopt a random utility approach where stochastic components capture the effects of the unobservable factors as follows:

$$\Pi_{ijm} = x_j'\gamma_i + y_j'\beta_i - \alpha_i p_{jm} - \theta_{i1}\left(\sum_{k=1}^{N_1} d_k x_{kj}\right) - \theta_{i2}\left(\sum_{k=1}^{N_1} d_k x_{kj}\right)^2 + \xi_{jm} + \varepsilon_{ijm} \tag{3.5}$$

In this formulation, ξ_{jm} represents the unobserved (by the researcher) product attributes which are market dependent. These may include specific marketing efforts of manufacturer's that exist in market *m* and influence customer's choice. ε_{ijm} represents the customer specific idiosyncratic tastes that come from the customer-specific effect of subjective attributes, customer specific cost effects, and other intangible factors not observed by the researcher. These errors are assumed to be distributed iid Type I Extreme Value across consumers and devices.

This formulation is the well-known random coefficient, or heterogeneous logit model of demand[4]. I make two additional assumptions necessary for the estimation. First, I assume that there are M different markets, and that different sets of devices are sold in these markets. Second, the customer also has the option of not purchasing any device from the choice set, i.e. choosing the 'outside good', denoted by $j = 0$. This assumption is not only reasonable, but also necessary: without an outside good option, a homogenous increase of the price of all products will not change quantities demanded. Under these assumptions, the theoretical model in Equation (2.6) is empirically specified as:

$$\underset{j \in J_m}{Max} \quad \Pi_{ijm} = x_j'\gamma_i + y_j'\beta_i - \alpha_i p_{jm} - \theta_{i1}\left(\sum_{k=1}^{N_1} d_k x_{kj}\right) - \theta_{i2}\left(\sum_{k=1}^{N_1} d_k x_{kj}\right)^2 + \xi_{jm} + \varepsilon_{ijm} \tag{3.6}$$

where J_m represents the choice set available in market *m*.

[3] In the following specifications I use the quadratic form for complexity (i.e $l = 2$ in Equation 3.3) clarity of exposition. In the empirical investigation, I examine different degrees of the polynomial.
[4] Note that without loss of generality, the terms in this formulation are added. In case where the effect of some terms should be negative to reflect an effect of a cost mechanism, the parameter's value is negative.

3.2 Identification of the Features and the Complexity Effects

The empirical model in Equation (3.6) captures the main effects of features on the profit through the γ_i parameter vector. In addition, a feature may affect the profit by its contribution to the complexity of the device. This additional effect is captured by θ_{i1} and θ_{i1}, which are the same for all features that enter the complexity score. Since θ_{i1} captures a linear effect of each feature, as does the vector γ_i , there is an identification issue that needs to be addressed. Formally, profit function can be written as:

$$\begin{aligned}
\Pi_{ijm} &= y_j{}'\beta_i + x_j{}'\gamma_i + \theta_{i1}\left(\sum\nolimits_{k=1}^{N_1} d_k x_{jk}\right) + \theta_{i2}\left(\sum\nolimits_{k=1}^{N_1} d_k x_{jk}\right)^2 - \alpha_i p_{jm} + \xi_{jm} + \varepsilon_{ijm} \\
&= y_j{}'\beta_i + \sum\nolimits_{k=1}^{N_1}\gamma_{ik}x_{jk} + \theta_{i1}\left(\sum\nolimits_{k=1}^{N_1} d_k x_{jk}\right) + \theta_{i2}\left(\sum\nolimits_{k=1}^{N_1} d_k x_{jk}\right)^2 - \alpha_i p_{jm} + \xi_{jm} + \varepsilon_{ijm} \\
&= y_j{}'\beta_i + \sum\nolimits_{k=1}^{N_1}(\gamma_{ik} + \theta_{i1}d_k)x_{jk} + \theta_{i2}\left(\sum\nolimits_{k=1}^{N_1} d_k x_{jk}\right)^2 - \alpha_i p_{jm} + \xi_{jm} + \varepsilon_{ijm}
\end{aligned} \tag{3.7}$$

The identification of θ_{i2} is achieved by the quadratic functional form of the complexity function. However, without further information, it is not possible to separate the main (γ_{ik}) and the secondary (θ_{i1}) *linear* complexity effect for the individual features. To that effect, I utilize data on heterogeneity in familiarity levels of different features by different customers. I assume that different customers are familiar with different features, no one knows all features (in fact, some features may be new), and the researcher knows which features the customer is familiar with. Under these assumptions, I construct an exclusion criterion that utilizes the heterogeneous familiarity levels. This criterion will allow identification of the linear parameters. The customer, at the time of purchase, is not familiar with all the features a new device offers. In fact, even after purchasing the device, the user may not be familiar with all the functionalities and capabilities a device offers. This can be clearly seen in Figure 3 which reveals differences in familiarity levels of different customers of representative copying device features. Features that are not familiar to the user do not provide the buyer with additional (positive or negative) utility. This is the exclusion criterion.

To formally present the exclusion criterion, I define X as the full set of features. I further define the partial set of features that are *familiar* to customer i as X_i. Utility that is not related to complexity will include only features in the set X_i. The complexity function, however, will still include all the features that the device offers, i.e. all the features in the set X. The reason for this is the fact that even if the user is not familiar with a feature, it is still part of the device and contributes to its complexity. The user has to browse through a full menu in order to reach or activate a feature required for a particular operation. Using this new notation, the profit from purchasing a device can be expressed as:

$$\Pi_{ijm} = \beta_{0i} + y_j{}'\beta_i + \sum_{k\in X_i} \gamma_{ik} x_{jk} + \theta_{i1}\left(\sum_{k\in X} d_k x_{jk}\right) + \theta_{i2}\left(\sum_{k\in X} d_k x_{jk}\right)^2 - \alpha_i p_{jm} + \xi_{jm} + \varepsilon_{ijm} \qquad (3.8)$$

Rearranging terms results in the following specification of the profit function:

$$\Pi_{ijm} = \beta_{0i} + y_j{}'\beta_i + \sum_{k\in X_i} \overbrace{(\gamma_{ik} + \theta_{i1} d_k)}^{benefits-costs} x_{jk} + \theta_{i1}\left(\sum_{k\in X\setminus X_i} d_k x_{jk}\right) + \theta_{i2}\left(\sum_{k\in X} d_k x_{jk}\right)^2 - \alpha_i p_{jm} + \xi_{jm} + \varepsilon_{ijm} \qquad (3.9)$$

Equation (3.9) shows that once θ_{i1} is estimated, the main effect of the familiar features, γ_{ik} are separately identified. By estimating $\gamma_{ik} + \theta_{i1} d_k$, and θ_{i1} , I can solve for γ_{ik} .

As a final note, the main effects of the familiar features (captured by the terms $\gamma_{ik} + \theta_{i1} d_k$) include both costs and revenue effects that are not related to complexity. A main effect will be negative when the cost effect outweighs the benefit from this feature. Clearly, the more difficult a feature is to use, the higher the cost effect of this feature. The unfamiliar features have difficulty costs that affect the profit through the complexity tern.

3.3 From Customer's Profit Maximization to Market Shares

Using McFadden (1974) logit formula and under the T1EV distribution of the error terms, ε_{ijm}, the profit maximization in Equation (3.6) yields the individual consumer's choice probabilities for each device in market *m*. Aggregating these probabilities over all consumers in each market is the market shares for these devices in market *m*.

To derive the consumer's choice probability for device *j* in market *m*, I start by specifying the heterogeneity of the parameters. In a logit formulation, it is particularly important to account for heterogeneity because demand elasticities of a homogeneous logit model do not allow for products of similar characteristics to be close substitutes. Since heterogeneity of the taste parameters amounts to modeling random coefficients, I assume that the customer's taste parameters ($\beta_i, \gamma_i, \theta_{i1}, \theta_{i2}, \alpha_i$) are multivariate normally distributed, conditional on demographics, as follows:

$$\begin{aligned}&[\beta_i, \gamma_i, \theta_{i1}, \theta_{i2}, \alpha_i]' = [\beta, \gamma, \theta_1, \theta_2, \alpha]' + \Lambda D_i + \Sigma \nu_i \\ &D_i \sim F_D(D) \\ &\nu_i \sim MVN(0, I)\end{aligned} \tag{3.10}$$

where D_i is a vector of demographic variables, Λ is a matrix of coefficients that measure how the taste characteristics are affected by demographics, $P_D(D)$ is an empirical demographics distribution, ν_i capture the unobserved customer tastes assumed to be multivariate normal, and Σ is a scaling matrix. Using these assumptions, the profit function can be decomposed as:

$$\begin{aligned}&\Pi_{ijm} = \delta_{jm}\left(x_j, y_j, p_{jm}, G(x_j), \xi_{jm}; \beta, \gamma, \alpha, \theta_1, \theta_2\right) + u_{ijm}\left(x_j, y_j, p_{jm}, G(x_j), D_i, \nu_i; \Lambda, \Sigma\right) + \varepsilon_{ijm} \\ &\textit{where} \\ &\delta_{jm}\left(x_j, y_j, p_{jm}, G(x_j), \xi_{jm}; \beta, \gamma, \alpha, \theta_1, \theta_2\right) = x_j'\gamma + y_j'\beta - \alpha p_{jm} - \theta_1\left(G(x_j)\right) - \theta_2\left(G(x_j)\right)^2 + \xi_{jm} \\ &u_{ijm}\left(x_j, y_j, p_{jm}, G(x_j), D_i, \nu_i; \Lambda, \Sigma\right) = \left[x_j, y_j, p_{jm}, G(x_j), G(x_j)^2\right]' * \left(\Lambda D_i + \Sigma \nu_i\right) + \varepsilon_{ijm} \\ &\qquad = x_j'\left(\Lambda_x D_i + \Sigma_x \nu_i\right) + y_j'\left(\Lambda_y D_i + \Sigma_y \nu_i\right) - p_{jm}\left(\tau_p D_i + \sigma_p \nu_i\right) \\ &\qquad - \left(G(x_j)\right)\left(\tau_G D_i + \sigma_G \nu_i\right) - \left(G(x_j)\right)^2\left(\tau_{G^2} D_i + \sigma_{G^2} \nu_i\right) + \varepsilon_{ijm}\end{aligned} \tag{3.11}$$

where $\left[\Lambda, \Sigma\right] = \left[\Lambda_1, \Sigma_1, \Lambda_2, \Sigma_2, \tau_1, \sigma_1, \tau_2, \sigma_2, \tau_3, \sigma_3\right]$, Λ_1, and Σ_1 are matrices of coefficients that measure how the **feature** parameters vary with demographics and unobserved individual characteristics, Λ_2 and Σ_2 are matrices of coefficients that measure how the **characteristic** parameters vary with demographics and unobserved individual characteristics, τ_1 and σ_1 are scalars that measure how price sensitivity varies with demographics and individual characteristics, and τ_2, σ_2, τ_3 and σ_3 are scalars that measure how the complexity variables are effected by demographic variables and unobserved preferences.

The specification is completed by specifying the utility of the outside good option as suggested by Berry (1994) to:

$$\Pi_{i0m} = \xi_{0m} + \tau_0 D_i + \sigma_0 \nu_{i0} + \varepsilon_{i0t} \tag{3.12}$$

Since in the logit models, only differences in utilities matter, I use the outside good as the base good and define $\pi_{ijm} = \Pi_{ijm} - \Pi_{i0m}$. I further use the standard practice (Berry et. al, 1995, Nevo 2000) and normalize the utility of the outsize good to zero.

Utilizing the distributional assumption in the error term ε_{ijm} gives rise to the logit formula (McFadden 1974). Thus, the probability that customer *i* will purchase device *j* in market *m* can, be expressed as:

$$p_{ijm} = \frac{\exp\left(\delta_{jm} + u_{ijm}\right)}{1 + \sum_{k=1}^{J_m} \exp\left(\delta_{km} + u_{ikm}\right)} \quad and \quad p_{i0m} = \frac{1}{1 + \sum_{k=1}^{J_m} \exp\left(\delta_{km} + u_{ikm}\right)} \tag{3.13}$$

The market share of device *j* in period m is the average of these probabilities across all customers:

$$\begin{aligned} s_{jm} &= \int_{D,v} \frac{\exp\left(\delta_{jm} + u_{ijm}\right)}{1 + \sum_{k=1}^{J_m} \exp\left(\delta_{km} + u_{ikm}\right)} dF_D(D) dF_v(v) \\ &= \int_{D,v} \frac{\exp\left(\delta_{jm} + u_{ijm}\left(x_j, y_j, p_{jm}, G(x_j), D_i, v_i; \Lambda, \Sigma\right)\right)}{1 + \sum_{k=1}^{J_m} \exp\left(\delta_{km} + u_{ikm}\left(x_k, y_k, p_{km}, G(x_k), D_i, v_i; \Lambda, \Sigma\right)\right)} dF_D(D) dF_v(v) \end{aligned} \tag{3.14}$$

where $F_D(D)$ is the joint distribution of the demographic variables and $F_v(v)$ is a multivariate standard normal distribution. Since there is no closed form for this integral, a simulation is necessary to approximate the market share as follows:

$$s_{jm} = \frac{1}{ns} \sum_{i=1}^{ns} \frac{\exp\left(\delta_{jm} + u_{ijm}\left(x_j, y_j, p_{jm}, G(x_j), D_i, v_v; \Lambda, \Sigma\right)\right)}{1 + \sum_{k=1}^{J_m} \delta_{km} + u_{ikm}\left(x_k, y_k, p_{km}, G(x_k), D_i, v_i; \Lambda, \Sigma\right)} \tag{3.15}$$

where v_is are drawings from the multivariate normal distribution and D_i are drawn from the empirical demographics distribution.

3.4 The Endogeneity of Price

The estimation of the market share system in Equation (3.15) could have been done using simulated maximum likelihood techniques had it not been for the unobserved device specific error component ξ_{jm}.

This error term is the ground for a price endogeneity issue which requires a different treatment of this non-linear system.

Endogeneity is a phenomenon that exists when an explanatory variable is correlated with the error term in a regression context. It usually arises if one or more of the following situations are present: (1) omitted variables (2) simultaneity (3) measurement error. In this study, I assume that the model characteristics and features of device *j* are exogenous. However, I treat price as an endogenous variable for the following reasons. First, similarly to Berry et. al (1995) I assume that the unobserved variables captured by ξ_{jm} include perceptual attributes (such as reliability and style), marketing activities (such as advertising efforts and channel power) and other physical characteristics and features of the devices not present in the model. The consumers and manufacturers have full information about the components of ξ_{jm}, but the researcher does not and therefore ξ_{jm} is treated as part of the error. Since there is reason to believe that the omitted variables affect the price of the devices (for instance, a more reliable device will have a higher price), there exists a correlation between the price variable and ξ_{jm} causing endogeneity. Second, simultaneity is a classic source of price endogeneity that originates from the fact that equilibrium quantities and prices are determined by simultaneous systems of demand and supply. Dube et. al, for instance, show that in a static Bertrand Oligopoly model, equilibrium prices are in fact functions of costs and share. As share increases, marginal costs of production can decrease (because of learning or returns to scale) and this decrease can bring down the price. Lastly, measurement error may exist if the data on prices does not include market related discounts, service costs etc… Any systematic error in measurement finds its way into the error term and may increase the correlation between the price and the unobserved error term.

Regardless of the source of the price endogeneity, not accounting for it will result in downward biases of inconsistent price estimates (Besanko et. al, 1998). To achieve consistent price estimates, I employ the instrumental variable estimation methodology suggested by Berry et. al (1995). Instrumental variables (IV) techniques require instruments which are additional variables that are correlated with the endogenous variable but are not correlated with the error term in a regression. While IV techniques have been well established methodologies in the context of linear regression models, the non-linear specification in Equation (3.15) requires a different approach. In his paper, Berry (1994) suggests an approach that uses instrumental variables by utilizing the share of the outside good. Berry's technique, however, works for the homogenous and for the nested logit models only. To account for price endogeneity in the heterogeneous model, I use the technique presented in Berry et. al (1995) that builds on Berry's original approach and enhances it to the heterogeneous case. I outline their estimation procedure in the next section.

3.5 The Estimation Procedure

The estimation procedure of the heterogeneous demand system suggested by Berry et. al (1995) is described by the following algorithm:

Step 1 – For given values of Λ and Σ, calculate the market share for device *j* in market *m* according to Equation (3.15). This requires knowledge of δ_{jm}. Berry et. al (1994) use the following contraction mapping to solve for δ_{jm}:

$$\delta_{jm}^{h+1} = \delta_{jm}^{h} + \ln\left(S_{jm}\right) - \ln\left(s(x_j, y_j, p_{jm}, G(x_j), \delta_{jm}^{h}; \Lambda, \Sigma)\right) \tag{3.16}$$

where S_{jm} is the observed market share of device *j* in market *m,* and $s(x_j, y_j, p_{jm}, G(x_j), \delta_{jm}^{h}; \Lambda, \Sigma)$ is the computed market share using Equation (3.15). Iterating on this equation until the difference between δ_{jm}^{h+1} and δ_{jm}^{h} is less than a given threshold yields the estimator $\hat{\delta}_{jm}\left(\Lambda, \Sigma\right)$.

Step 2 – Solve the following linear regression system:

$$\hat{\delta}_{jm}\left(\Lambda, \Sigma\right) = x_j'\gamma + y_j'\beta - \alpha p_{jm} - \theta_1\left(G(x_j)\right) - \theta_2\left(G(x_j)\right)^2 + \xi_{jm} \tag{3.17}$$

This is a linear regression system and the classic linear instrumental variables estimation method such as GMM can be used, given proper instruments can be found. This step produced the estimates: $\hat{\beta}, \hat{\gamma}, \hat{\alpha}, \hat{\theta}_1, \hat{\theta}_2$ that are, in fact, functions of the values of Λ and Σ from step 1.

Step 3 – Perform a GMM procedure to estimate Λ and Σ by Searching for the values of Λ and Σ that minimize the following objective function:

$$\min_{\Lambda,\Sigma} \xi_{jt}' Z\left(Z'\Psi Z\right)^{-1} Z'\xi_{jt} \tag{3.18}$$

where Z is a set of instruments to be used, Ψ is a weighting matrix and ξ_{jt} is given by:

$$\xi_{jt} = \hat{\delta}_{jm} - x_j'\hat{\gamma} + y_j'\hat{\beta} - \hat{\alpha} p_{jm} - \hat{\theta}_1\left(G(x_j)\right) - \hat{\theta}_2\left(G(x_j)\right)^2 \tag{3.19}$$

In practice, optimizing over Λ and Σ is done, as suggested by Nevo (2000), by using the Nelder and Mead (1965) non-derivative "simplex" search method. The optimal weighting matrix Ψ is a function of the true or consistent estimates of the parameters. Nevo suggests using a naïve weighting matrix, such as $\Psi = Z'Z$ to obtain consistent estimates of the parameters. These parameters can then be used to construct an optimal weighting matrix which in turn can be used to obtain the parameters estimates.

4 Data

The effect of complexity on choice in business environments will be examined by analyzing the factors that influence businesses' purchase decisions of industrial devices. The Hardcopy Peripheral (HCP) market includes all the printers, faxes, scanners, single function digital copiers and multifunctional devices. The particular sub-market of the HCP market at the focus of this investigation includes single and multifunctional copier devices targeted at corporate work groups. In what follows, I describe the single and multifunctional copier market and the unique datasets that will be used for the empirical analysis.

4.1 The Single Functional and Multifunctional Copier Market

The single functional and multifunctional copier market includes either single functional copiers (henceforth referred to as SFC) or multifunctional devices (referred to as MFC) that have a copier function and perhaps other capabilities from the print/fax/scan capability set. The evolution of this market is represented by the two devices depicted in Figure 4. On the left is a modern multifunctional device that can copy, fax, scan, print, create booklets, staple, and adjust image size among other things. The device on the right is one of the first copiers that were introduced by Xerox in the late 1940s. It has a single function - copy, and its only feature is the number of copies it can produce.

The particular set of devices I am investigating is the set of all SFC/MFC devices that have a copying speed of 30-90 ppm[5] and that were targeted to businesses of ten or more employees between 1998 and 2006. These devices make up 13% of the total unit sales but 72% of the total revenues in the SFC/MFC market. Such devices are sold mostly to businesses (rather than to households, home offices or small offices) and are therefore a suitable exemplar of an industrial market. Additionally, these devices are, on average, more powerful and versatile in their feature sets than the slower devices (1-30 ppm). Since I am examining the sensitivity of business customers to a consequence of feature richness, the variation in the feature set in these devices makes this market particularly appropriate for this investigation. The average number of features in these devices has increased over time, as can be seen in Figure 5, and represents well the feature creep phenomenon of the digital era. The devices in this study range in price from $1050 to $200,000 with an average of $24,500. The price of these devices reinforces why they are usually not sold to home or small offices. The average selling price in this market steadily decreases over time, which can be partially explained by the steady increase in the number of competing unique devices in this market (Figure 6) as well as the falling costs of computational capacity necessary for the production of these devices.

4.2 Datasets

A comprehensive and unique data set set was assembled from three different datasets and was used in the empirical analysis. The first dataset (henceforth Dataset 1) contains SFC/MFC model-specific information on characteristics, features, quarterly unit sales and quarterly average prices. The second dataset (henceforth Dataset 2) contains survey responses of ratings of 'ease or difficulty of use' and 'familiarity' for different features of SFC/MFC devices. The third dataset (Dataset 3) contains information on yearly distribution of business demographics in the US between 1998 and 2006.

Dataset 1 was assembled from two main data sources. *Buyers Laboratory Inc.* (*BLI*) supplied model specific characteristics and features dataset. *International Data Corp. (IDC)* supplied model specific quarterly sales and prices. I merged the two datasets by model name. In the merge, I excluded devices that had

[5] Pages Per Minute

incomplete characteristics or sales data[6]. Additionally, *IDC* sometimes reports total sales of families of devices (for instance, the number of shipment units of Xerox 123 and Xerox 123a are reported together under model Xerox 123 with the average selling price of these devices). In those cases where *BLI* reports a separate characteristic set for each family member but *IDC* only reports sales of the base model – I include only the base model in my sample. In these cases, I calculate the base device's price as a weighted average (weighted by sales) of all devices that constitute a family of devices. The merged dataset contained 333 unique devices sold between 1998 and mid 2006. To check the representativeness of this sample, I first examined how each vendor in the market is represented in the sample. Figure 7 shows the brand sales distribution of the top vendors in the market, as well as the sales volumes of the vendors in the sample. In the sample, the sales of each of the 4 biggest vendors in the market (HP, Xerox, Canon and Ricoh – yearly sales of 10000+ units) are more than 86% of their actual market sales. The medium sized (yearly sales of 2500–5500 units) vendors' sales are above 88% of their actual sales, with the exceptions of Savin and Toshiba. These two manufacturers are not represented in my sample due to the lack of features data for their devices. Small vendors (less than 2000 units per year) were not included in the sample as well. Overall, devices from 11 vendors are represented in the sample, and represent 74.4% of the actual unit sales of the investigated market.

The second dataset I use summarizes the results of a survey I conducted in order to collect specific familiarity and ease (or difficulty) of use ratings for *each* copier feature. In this survey, I listed 32 features and functions of multifunctional devices, and asked the respondents to choose their familiarity level with each feature (on a scale of 1 – 'not familiar' to 4 – 'familiar and using often'). Additionally, I asked the respondent to rate their ease (or difficulty) of using each feature (on a scale of 1 – 'very easy to use' to 5 – 'very difficult to use'). I surveyed 101 respondents with different business roles (managers, engineers, copier experts, administrators and starting level business employees) and from different organizations. The descriptive statistics of the survey results are shown in Table 3. The average of each feature's 'ease of use' ratings was used as weight in calculating the empirical complexity level of a device (Equation (3.2)). The means of the

[6] Some characteristics and features were hand collected from vendors' websites, device manuals and various internet sources, and these devices were not omitted.

familiarity ratings provide the additional information that is necessary to achieve identification in the empirical model, as described in Section 3.2.

The third dataset contains information on the distribution of business units demographic available from the Country Business Patterns section of the US Census. Data on the number of firms of different sizes was obtained for the years 1998-2006. This data will be used to investigate how firm demographics impact the effect of complexity on choice in the heterogeneous model.

4.3 Variables

The variables contained in Dataset 1 are prices, sales, characteristics and features. The **characteristics** vector (y_j in Equation (3.5)) used in the analysis consists of variables that were pointed out in discussions with industry experts and field professionals as ones that may have considerable effects purchase decisions. The characteristics I use in the analysis are: price (average selling price), copying speed (pages per minutes), size (cubic inches), warm-up time (seconds), paper capacity (the sum of the maximum amount of paper each drawer, tray or cassette holds) and main platform (device's primary function – copier or printer). Table 1 presents the descriptive statistics of these variables. In the analysis, I use log of characteristics as it allows for diminishing returns effects and also demonstrates a better fit. I use manufacturer dummies to account for brand effects and to control for interface similarities across devices of the same manufacturer. I also include a fourth quarter dummy to control for seasonality (September-December usually demonstrate peeks in sales in this market). Since the number of unique devices per quarter has an increasing trend (see Figure 5), I follow Ackerberg and Rysman (2005) and include the number of unique devices per quarter as an additional variable. These authors show that in the presence of product congestion, when there are significant changes in the size of the choice sets, standard estimation may produce biased elasticities. To correct for these biases, they propose incorporating a function of the number of products in the market into the demand equation. This way, the price elasticities will be identified by the price variations and not by the change in the number of products.

The **features** that were used in the analysis are reported in Table 2 along with their definitions and descriptive statistics. In this investigation, I use the 34 main copier features as reported by *Buyers Lab*. For the main effect, I used only features that demonstrate sufficient variation. For example, the feature 'copier book' and 'copier preset' are offered by all the devices (99%), and were therefore excluded. Other features that I used were general paper-handling related features such as duplexing, sorting and stapling. The capabilities print, fax and scan were also included as features since specific data on the features offered by each of these capabilities was not available.

The **complexity** score for each device was constructed as a weighted sum of the features this device offers (see Equation (3.2)). The weight associated with each feature was the average (across respondents) of the difficulty or ease of use ratings of this feature. These averages can be seen in Table 3, in the leftmost column. The complexity score used in this analysis is constructed from copier features (book, booklet, copy control, covers, editing, erase, image insert, image repeat, image rotate, job programs, job build, xy zoom, margin shift, reversal, OHP interleaving, photo, poster, program ahead, sheet insertion, stamping, timer, 2 in 1, language), paper features (feeder, duplex, sort and staple) and function features (print, fax, scan). It should be noted that the complexity score is calculated from all the features mentioned above, including those which do not appear in the main effects of the model. The reason for this is that not all devices have the same set of features, and therefore the total number of features has a greater variation than the variation of each feature by itself. The total number of features for the devices range from 9 to 32, with the mean number of features being 22, and the standard deviation being 4.2. The complexity level for each device ranges from 17.8 to 60.77, with mean of 44.18 and standard deviation of 8.5.

4.4 Market Size and Instruments

The estimation of the taste parameters in the demand equation (Equation (3.15)) requires information on the potential market size of SFC/MFC devices with 30-90 ppm, in order to find the share of the outside good. In this research I model the choice decision conditional on a preliminary decision to purchase a copying device. Therefore, the market size can be taken as the total US SFC/MFC sales of 1-90 ppm MFC/SFC devices to businesses of ten or more employees during 1998-2006. *IDC* reports sales of *all*

SFC/MFC devices, by speed, and I use these measures to calculate the quarterly market size for MFC/SFC 1-90 ppm. To calculate the outside good, I assume that when a firm decides not to buy a device from the sample choice set, it buys a different SFC/MFC device available in the market. The outside good in this case includes all the single and multifunctional devices of speeds 1-29 ppm, as well as all the devices with speeds 30-90 ppm that were manufactured by vendors which were not in the sample[7]. The share of the inside good sales constitutes 13% of the total SFC/MFC market between 1998 and 2006.

The estimation methodology also requires instruments to account for the endogeneity of price, discussed in Section 3.4. Any factor that is correlated with the price of device *j* at time *m* and is not correlated with the demand error term ξ_{jm} can be an appropriate candidate. In the analysis, I use several instruments that can be considered proxies for manufacturing costs. The first is constructed, according to an industry conversion function, from the weight of the device. According to industry specialists, the variable manufacturing costs is a linear function of the weight of a device. I use log of this constructed variable and several powers of it as instruments. Another instrument that I use is a technology price index. Between 1998 and 2006, technology prices dropped significantly. Since technology prices are correlated with manufacturing costs of digital SFC/MFC devices, they are correlated with the prices of these devices. As a technology index, I used the quarterly average price per gigabyte of storage[8]. Additionally, I use the traditional lagged prices as instruments. In this market, a customer buys a device every 3-5 years, so the price in a previous period is not expected to affect current demand.

5 Results and Discussion

In this section I present four sets of results – a benchmark homogeneous model, with and without complexity, and a heterogeneous model, with and without a complexity term. The results of the analysis support the notion that there exists a complexity effect, which is reflected in actual purchase behavior, and that this effect is in fact concave. In other words, as more features are added into a device, the profit or

---

[7] I exclude production devices (90+ ppm) since they are specialized printing and copying devices for the printing industry.

[8] These figures were acquired from Gartner Inc and The World Bank (http://www.worldbank.org/)

utility from this device increases and then decreases after an optimal point. Furthermore, I find evidence that not accounting for complexity may generate biases in the estimates of features and price. These findings illustrate the importance of incorporating a complexity measure into choice models.

5.1 Results of a Homogenous Model

In the homogenous model I assume that all consumers have the same preferences or tastes for price, features, and characteristics. In other words, I set

$$[\beta_i, \gamma_i, \theta_{i1}, \theta_{i2}, \alpha_i] = [\beta, \gamma, \theta_1, \theta_2, \alpha] \tag{5.1}$$

in Equation (3.10). This model is estimated using the technique developed in Berry (1994) to account for price endogeneity in a homogeneous logit formulation. To achieve identification (as discussed in Section 3.2), I construct X_i (in Equation (3.9)) as the set of features that are familiar (to some level) to more than 30% of the respondents[9]. I test two homogeneous models – Model A incorporates the complexity terms and Model B which does not include them. The results are reported in Table 4.

Comparing the two models reveals noteworthy differences in the values and significance of the coefficients. The main difference between the models is the significance of features estimates. Some features (such as *image insert, image repeat, image rotate,* and others) appear to have a significant effect on choice in Model A but show up as insignificant in Model B. Other features' coefficients (such as *fax* and *covers*) that have a negative and significant sign in Model B show up as insignificant in Model A. The later results seem more realistic because, for instance, the negative and significant *fax* coefficient in Model B (-0.16) cannot be intuitively explained. The insignificance of the *fax* coefficient in Model A can be explained by the fact that most firms already had standalone fax devices in the time of this investigation, so this addition to the copier may be redundant. Another important difference between the models is demonstrated in the price effect. In model A, when complexity is accounted for, the price sensitivity is higher. This result is consistent with the fact that price is negatively (and significantly) correlated with complexity. I will elaborate on this issue in the next section, where I discuss the results of the heterogeneous model. The higher price coefficient in Model A

[9] Results are robust to other familiarity thresholds.

is reasonable – people are more sensitive to an increase in price when they explicitly take into account the additional costs incurred by complexity. Both models have very close R^2 measures, with Model A demonstrating a slightly better fit than Model B. Below I discuss the results of Model A where complexity is accounted for.

The coefficients of the complexity function are both significant and negative: $\theta_1 = -1.13$ and $\theta_2 = -0.22$. These coefficients of the quadratic function (Equation (3.5)) clearly support a concave complexity effect: as more features are added into the device, the utility from these devices increases and then decreases. In the main effects of the features, most features demonstrate a positive effect on utility, which means that their benefits outweigh their costs (net of the complexity costs) to the firm. The features that demonstrate a negative and significant effect require further discussion. According to the model a negative effect reflects a situation where the benefits from these features are exceeded by the costs they impose on the firm. *Copier photo*, for instance has a negative effect, and this result is intuitive in industrial settings. Photo capabilities are not only associated with more costly supplies (special paper for instance), but are also providing opportunities for employees to spend more personal time in their workplace. It may therefore be an undesirable feature to a significant fraction of firms. The feature *job programming* is also negative and significant. This feature is, on average, the second most difficult to use among all features (with mean difficulty level of 2.65, see Table 3). The results may reflect the fact this feature imposes a high activation, operating or support costs arising from the difficulty of actually using or learning the *job programming* feature. Similarly, the activation of *energy save* mode incurs a cost of waiting time (until the device warms up again) which may not offset the energy cost savings this feature provides. Alternatively the person who derives the potential benefit from this feature is not the same as the person who incurs the time costs. Finally, the results of the homogeneous model reflect average taste parameters in the population. It could be that some consumer segments do in fact find the *energy save* or *job programming* features desirable. In other words, it may be possible that manufacturers include these features in their devices so that they can appeal to these segments.

The effects of the characteristics are significant and are consistent with a cost minimization objective of the consumer. The *copy speed* (0.19) and the *paper capacity* (0.26) are significant and positive; *device size* (-0.15) and *warm-up time* (-0.31) are significant and negative. These results indicate that faster devices that require less maintenance and take up less space are preferred.

The positive and significant coefficient of *copier base* (3.33) indicates that in this market, devices whose primary function is copying are preferred. The evolution of multifunctional devices began when manufacturers of single functional copiers and single functional printers started to use convergence strategies and added print, copy, fax and scan capabilities into their products. The result was a market of multifunctional devices that generally have two different base configurations. Multifunctional devices whose base configuration is a copier have a stronger preference among consumers. This result is consistent with market behavior. For instance, Hewlett Packard (initially a printers company) has famously started from their strong Small Office/Home Office multifunctional device and printer base and has failed to gain significant market share in the SFC/MFC market, even after a number of attempts. The strong preference for devices with a copier base may also reflect a category loyalty present in this market: multifunctional devices were first introduced as copiers with enhancements and consumers are used to the idea of multifunctional copiers and standalone printers.

The fourth quarter dummy is positive and significant (0.14) demonstrating the peak sales in the fourth quarter.[10] I will elaborate on this in the next section. The number of unique devices in the market has a negative and significant effect (-0.02), demonstrating that crowding causes market shares to shrink, as expected. The brand effects are consistent with the brand name values in this market. The Xerox brand is the base brand, and is inferior only to Canon and Konica.

Overall, these results are intuitive and indicate that devices that are simpler to use, are faster and cheaper to use are preferred in the MFC/SFC market. More importantly, the effect of complexity at the time of purchase is concave pointing to the existence of an optimal complexity level that maximizes the buyer's utility.

[10] Other quarters were not significant.

5.2 Heterogeneous Model Results

In the heterogeneous model, I implement the methodology discussed in Section (3.4) and estimate the taste parameters of the demand model specified in Equation (3.15). In this estimation, I assume random coefficients of price, complexity, squared complexity, and a constant term as follows:

$$\begin{aligned} [\theta_{i1}, \theta_{i2}, \alpha_i]' &= [\theta_1, \theta_2, \alpha]' + \Lambda D_i + \Sigma \nu_i \\ [\beta_i, \gamma_i]' &= [\beta, \gamma]' \end{aligned} \tag{5.2}$$

The firm demographics D_i include firm sizes in the US in the years 1998 and 2006. For each quarter, I draw a random sample of 200 firm sizes from the empirical distribution of sizes of businesses in the U.S. I include the square of the firm size as an additional demographics variable in order to capture non-linear effects, as suggested by Nevo (2000).

The results of the heterogeneous model are shown in Table 5. Model C presents the results of the model with complexity, and Model D presents the results of the model estimated without complexity terms. The means of the distribution of the marginal utilities of the heterogeneous parameters are shown in the leftmost column of Models C and D. For the other variables of the model, the values in this column are the homogeneous point estimates of their coefficients. Estimates of heterogeneity around the means of price and the complexity variables are presented in the next three columns. The column labeled "Std Deviation" captures the effects of the unobserved demographic effects. The columns labeled "Firm Size" and "Firm Size^2" show the demographic effect on the slope coefficients. Due to the large number of parameters estimated in these models, the standard errors have not yet been computed and are work in progress. The subsequent discussion focuses on the magnitudes of the taste estimates.

The significant price effect (-1.10) in the heterogeneous model is stronger than in the homogeneous model (-1.01), as expected (Besanko et. al, 1998). This magnitude of the price coefficient reveals that price elasticity in this market is low. In fact, because the shares of each device in the market are very small (as there are many unique devices competing each quarter, see Figure 6), the own price elasticity can be approximated by the price coefficient[11]: -1.10. This means that this product category exhibits low own price elasticities. This result is not surprising, though, since the SFC/MFC devices of 30-90 ppm are usually leased to

[11] Exact calculations of elasticities will be supplied later.

businesses, not sold. Firms, when their lease expires (usually every 3-5 years) need to buy (lease) a new device under the new terms. In many instances, the price of the device is subsumed in a contract which focuses on price per page and includes services and consumables. The device price, in these situations, may be subsidized by higher margins from consumables or services, so sensitivity to net price may appear low. More information on prices of consumables and services may shed light on this issue. Additionally, the dramatic decline in prices over time, along with the increase in the number of unique devices, may suggest that consumers are less price sensitive and are looking for differentiations in other attributes. The top graph in Figure 8 presents the frequency distribution of the price coefficient generated using a sample of 3400 firm sizes drawn from the empirical distribution of firm sizes in the US. From the graph, it is clear that heterogeneity in price sensitivity exists and is substantial as some firms demonstrate elastic demand and others have an inelastic price effects.

The positive interaction coefficient of price and firm size (1.34) reveals that larger firms have lower price sensitivities, and since these are fairly expensive devices this is a rational finding. Further, amplifying the preveious observation, large firms are more likely to both lease and lease multiple devices. Indeed a well documented trend in the industry is the move towards full "fleet management" contracts where the price of any one device is of marginal significance.

A comparison between Model C and Model D reveals, as in the homogeneous case, that the price coefficient is underestimated when complexity is not accounted for: when complexity is accounted for, the sensitivity to price is slightly higher and this is consistent with the fact that people who take complexity into account will react more to a price increase.

The significant and negative complexity terms are consistent with a concave complexity effect. I calculate the optimal level of complexity by solving for the maximum of the quadratic polynomial presented in Equation (3.3), where I set $l = 2$. In other words, the complexity level that maximizes the utility can be expressed as:

$$\underset{G(x_j)}{\arg max}\left\{\theta_1 G(x_j) + \theta_2 G(x_j)^2\right\} = {-\theta_1}/{2\theta_2} \tag{5.3}$$

In the homogeneous model, this maximum is -2.57 and in the heterogeneous model it is -3.83. Recall that the complexity variables that were used in the estimation were standardized[12], and therefore the optimal values above need to be transformed to the same scale as the original complexity variables. The mean and standard deviation of the original complexity variables (weighted sums of features) are 44.18 and 8.5, respectively. Using these measures, the optimal complexity level computed for these devices, everything else held constant, is 22 in the homogeneous model and 12 in the heterogeneous model. This difference reveals that the homogeneous model overestimates the optimal level of complexity. Not only is it important to account for complexity in these choice models, but also it is important to control for heterogeneous preferences so that the optimal level of complexity will not be biased, causing decreased utility for the consumer.

Further investigation of the source of heterogeneity in the preference for complexity focuses on the interactions between firm sizes and complexity variables. The results show that these interaction effects are positive for the quadratic term and negative for the linear term. This means that larger firms have a stronger preference for devices with a higher complexity level than smaller firms. Larger firms usually have a larger user population and more diverse requirements and needs from these devices. In other words, bigger firms tolerate more complexity perhaps because the users of these devices are more heterogeneous in their needs, and the buyers want to satisfy a bigger range of these needs with one device. The heterogeneous distribution of the complexity term is shown in the bottom graphs of Figure 8. While the average consumer shows a concave complexity response, heterogeneity reveals that that there is a (small) consumer segment who seem to have a positive coefficient for the quadratic term. These consumers may infact exhibit a purchase behavior that supports the "more is better" paradigm. However, the majority of consumers appear to have an optimal complexity level that maximizes their utility. These results stress the importance of examining the sources of the heterogeneity of complexity.

The estimates of the characteristics coefficients in Models C and D are consistent with the results of the homogenous models. A positive *copy speed* coefficient and a negative *warm-up time* coefficient suggest a stronger preference for faster devices. A negative *device size* coefficient and a positive *paper capacity* suggest a

[12] This is due to the fact that the complexity levels show multi-collinearity between the linear and the squared terms.

stronger preference for lower costs manifested by rent and maintenance. The high positive coefficient of the *copier base* characteristic points to the preference for devices whose main role is a copier (vide supra). The seasonality effect is revealed in the positive fourth quarter coefficient demonstrating the increase in demand for these devices in the months Sept-December. There may be several reasons for this finding. First, from the supply perspective, sales people are inclined to achieve their quotas and therefore exert more efforts towards the end of the year. Second, from the demand perspective, firms, in particular procurement units, usually want to exhaust their budgets before the year ends.

The features effects estimates in the heterogeneous model are of similar signs but of different magnitudes than the features effects in the homogeneous model. The preference for almost all features is underestimated by the homogeneous model, with the exception of *interrupt* and *print*. Twelve features (*reversal, 2-in-1, duplex, image insert, poster, image rotate, erase, timer, OHP interleaving, copy control and image repeat*) are underestimated by more than 20%. It seems that controlling for heterogeneity in price and in complexity affects the preference coefficients for different features. This product category exhibits low price elasticity but at the same time high sensitivity to complexity as well as features effects. It may be that the demand in this market is driven more by features differentiation than by price.

In conclusion, the heterogeneous results are consistent with cost savings and profit maximizing objectives of the firms. The complexity effect as demonstrated by its effect on market shares is significantly concave, depends on the number of employees of the consumer, and reveals that there exists an optimal number of features for the SFC/MFC 30-90 ppm product category.

6 Conclusion and Future Research

The technological advancements of the last decade brought with them digital products with a large number of features and options. This enhanced functionality, however, turned out to be a double-edged sword: using the sophisticated set of functions and options requires customers to invest significant time and effort. The tradeoff between enhanced functionality and the increase in the complexity of a device is at the heart of this research. The increase in the complexity of technological and digital products is not a new topic.

In his 1988 book, "The psychology of Everyday Things", Don Norman gives many examples of products, such as the office telephone, that are so complicated to use that even highly educated engineers find then too difficult to use. In fact, Phillips CEO Paul Zevens says that only 13% of all Americans believe that technology products in general are easy to use.

This study was motivated by the need to address this issue and by the desire to find ways to quantify the effect complexity (induced by feature richness) has on actual purchase behavior. In this article, I propose a theoretical framework and develop an empirical methodology to analyze the effect of complexity on consumers' purchases as measured by sales and market shares of industrial devices. Specifically, I use the random coefficients logit to model demand for multi and single functional copiers in business environments. The results reveal that the complexity levels of these devices have a concave effect on demand. In other words, holding everything else constant, there is an optimal complexity level that maximizes the customer's perceived utility from multifunctional copiers. Additionally, preliminary results of this research suggest that there are potential biases in price and features estimates when a customer preference model does not account for complexity.

The methodology proposed in this study can be used by manufacturers of business products that wish to examine the effect of an increasing number of features on their potential market shares. If the behavior manifested in a specific market is indeed concave, an optimal number of features can be found for different consumers. This optimum should then be used (in advanced conjoint methods for instance) in the product planning stages of the devices.

In the long run, knowledge of desired complexity levels will allow manufacturers to reallocate their R&D resources and redirect their innovative efforts to other venues. Additionally, segmentation based on desired complexity levels and specialized products may prove to be profitable strategies. On the demand side, businesses customers that use digital products that fit their preference for complexity will be able to focus on their core business rather than invest extra time and efforts learning and operating complex devices.

While this research is the first to incorporate complexity into quantitative choice analyses, further research on this topic is necessary. A thorough examination of the sources that drive the heterogeneity in the preferences for different complexity levels is important, and could be done with more firm demographics data. In particular, individual firm level usage and purchase data can be helpful in identifying different complexity preferences. For example, the relationship between preference for complexity and average user's age, organizational roles in decision making processes, firm structures and other firm characteristics can be explored. Moreover, examining if complexity effect is different for different industrial product categories is another aspect that needs to be explored.

It would also be interesting to expand the methodology presented in this paper to a dynamic setting. This will allow an examination of how preferences for complexity evolve over time, and can be a useful prediction tool utilizing dynamic models of demand, where consumers have expectations on complexity as well as on price.

In sum, the key takeaway from this research is the importance of including a complexity measure in preference models used to analyze purchase behavior. Ignoring the complexity effect may produce biased results that lead to non-optimal product designs that do not maximize the user's utility. Additionally, since the concave complexity affect reveals itself in purchase behavior (rather than stated preferences), it is important to examine new techniques that will allow for the updating of stated preferences according to actual market behavior.

References

Ackerberg, A. D., and M. Rysman (2005): "Unobserved Product Differentiation in Discrete-Choice Models: Estimating Price Elasticities and Welfare Effects," RAND Journal of Economics, 36(4) 1-19.

Agarwal, M.K. and B. T. Ratchford (1980), "Estimating Demand Functions for Product Characteristics: The case of the automobiles", Journal of Consumer Research, 7, 249-262.

Bajari, P. and C. L. Benkard, (2005), "Demand Estimation with Heterogenous Consumers and Unobserved Product Characteristics: A Hedonic Approach", Journal of Political Economy, 113(6), 1239-1276.

Berry, S. (1994), "Estimating Discrete Choice Models of Product Differentiation". Rand Journal of Economics. Summer 1994: 242-263

Berry, S., Levinsohn., J. and A. Pakes (1995): "Automobile Prices in Market Equilibrium," Econometrica,63.

Besanko, D., S. Gupta, D. Jain (1998): "Logit Demand Estimation under Competitive Pricing Behavior: An Equilibrium Framework", Management Science, 44(11), 1533-1547.

Dube, J.P., B. Bronnenberg, R. Goettler, P.B. Seetharaman, K. Sudhir, R. Thomadsen, and Y. Zhao, (2002), "Structural Applications of the Discrete Choice Model", Marketing Letters, 13(3), 207-220.

Dube, J.P., J. Fox, and C. Su (2009): "Improving Numerical Performace of BLP Static and Dynamic Discrete Choice Random Coefficients Demand Estimation", working paper, SSRN abstract 1338152.

Hans, J. K., S.W. Chung and Y.S. Sohn (2009): "Technology Convergence: When Do Consumers Prefer Converged Products to Dedicated Products?", Journal of Marketing, 73(4), 97-108.

Hauser, J. R. and S.M. Shugan (1980), "Intensity Measures of Consumer Preferences", Operation Research, 28(2), 278-320.

Hauser, J. R. and D. Clausing (1988), "The House of Quality", Harvard Business Review, 66(3), 63-73.

Horsky, D., and S. K. Mohanty (2009): "The Diffusion of New Industrial Products", working paper, University of Rochester.

Horsky, D., S. Misra and P. Nelson (2006): "Observed and Unobserved Preference Heterogenity in Brand-Choice Models", Marketing Science, 25(4), 322-335.

Horsky, D., P. Nelson and S. S. Posavac (2004): "Stating Preference for the Ethereal but Choosing the Concrete: How the Tangibility of Attributes Affects Attribute Weighting in Value Elicitation and Choice", Journal of Consumer Psuchology 14(1&2), 132-140.

Kano, Noriaki; Nobuhiku Seraku, Fumio Takahashi, Shinichi Tsuji (April 1984). "Attractive quality and must-be quality" (in Japanese). Journal of the Japanese Society for Quality Control, 14(2), 39-48

Luo, L., P.K. Kannan, B. T. Ratchford (2008), "Imcorporating Subjective Characteristics in Product Design and Evaluations", Journal of Marketing Research, 45(2), 182-194.

Meyer, R.J., Zhao S. H. and J.K. Hans (2008): "Biases in Valuation vs. Usage of Innovative Product Features", Marketing Science, 27(6), 1083-1096.

McFadden, D., (1974), "Conditional Logit Analysis of Qualitative Choice Behavior", Frontiers in

Econometrics, ed. P. Zarembka , New York, Academic Press.

McFadden, D., and K. Train (2000), "Mixed MNL Models for Discrete Response", Journal of Applied Econometrics, 15, 447-470.

Nelder, J. A., and R. Mead, 1965, "A Simplex Method for Function Minimization", Comput. J., 7, 308-313.

Netzer, O., O. Toubia, E. T Bradlow, E. Dahan, T. Evgeniou, F. M. Feinberg, E. M. Feit, S. K. Hui, J. Johnson, J. C. Liechty, J. B. Orlin, V. R. Rao (2008), "Beyond Conjoint Analysis: Advances in Preference Measurements", Marketing Letters, 3, 337-354.

Nevo, A. (2000): "A practitioner's Guide to Estimation of Random Coefficients Logit Models of Demand", Journal of Economics and Management Strategy, 9(4) 513-548.

Pogue, D., (2009), "Apple's Sleek Upgrade", NY Times August 26, 2009.

Rosen, S. (1974), "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition", Journal of Political Economy, 82, 43-55.

Shugan, S. M. (1980), "The Cost of Thinking", Journal of Consumer Research, 7(2), 99-111.

Thompson, D. V., R. Hamilton and R.T. Rust (2005): "Feature Fatigue: When Product Capabilities Become Too Much of a Good Thing", Journal of Marketing Research, 42, 431-442.

Wingfield, N (2006), "Must Consumer Electronic Be So Complicated? If Only Home Entertainment Products Could Be As Simple As… Computers", The Wall Street Journal, New York, NY, Nov. 27, 2006, pg. R.4

Zhang, Y., A. Fishbach and A. W. Kruglanski (2007): "The Dilution Model: How Additional Goals Undermine The Perceived Instrumentality of a Shared Path", Journal of Personality and Social Psychology, 92(3), 389-401.

Zhao, S., R. J. Meyer and J. Han (2005): "The Rationality of Consumer Decisions to Adopt and Utilize Product-Attribute Enhancements: Why Are We Lured by Product Features We Never Use?", Experimental Business Research, Marketing, Accounting and Cognitive Perspectives Volume III, ed. Zwick R. and A. Rapoport

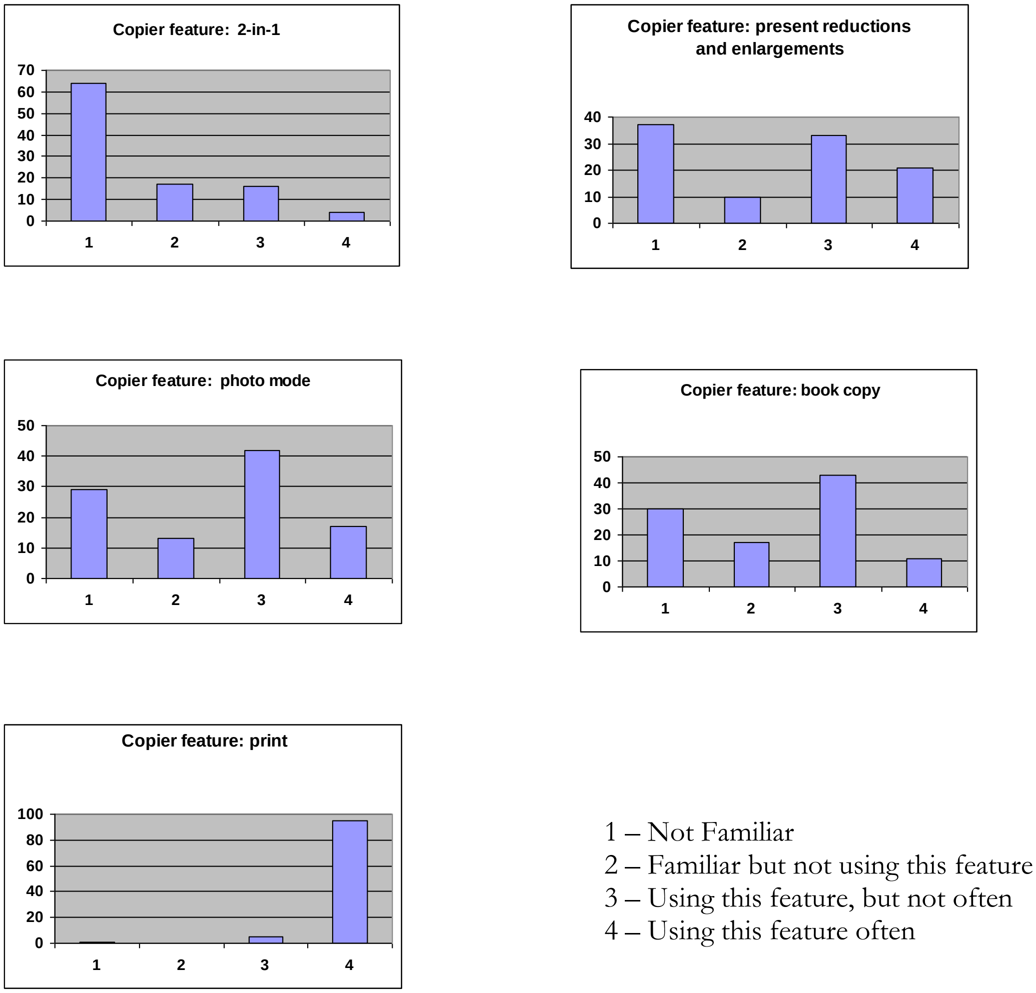

Figure 3: Distribution of familiarity levels of different features

Figure 4 - An MFC and an SFC device

In 1949, Xerox Corporation introduced the first and successful xerographic copier called the Model A (right) – it had a single button and a quantity dial. In 2006, copiers are larger, provide hundreds of features and are controlled by a digital user interfaces (left).

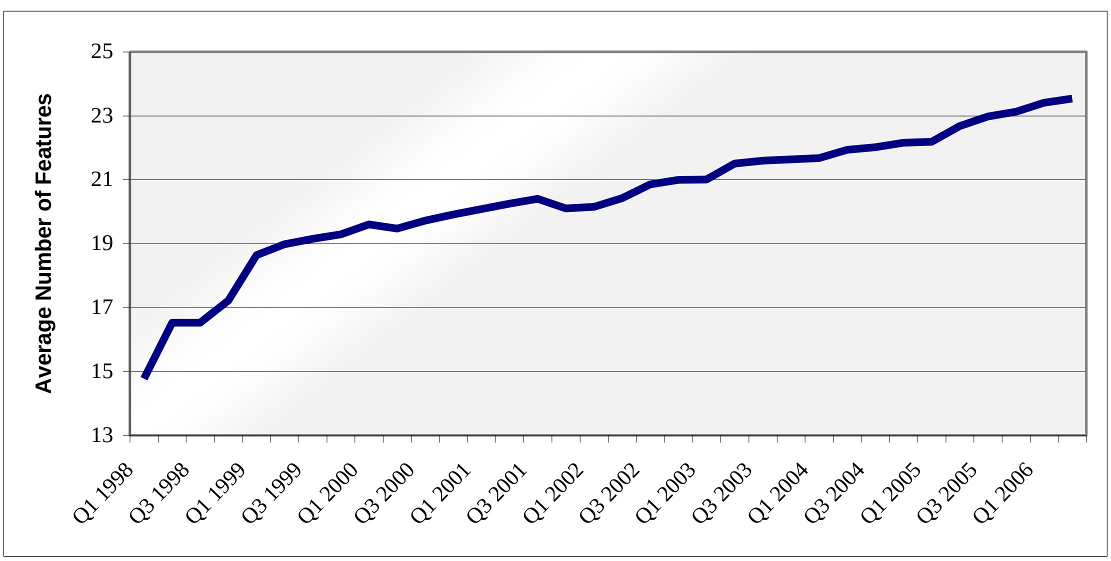

Figure 5 – The Average Number of Features of MFC/SFC 30-90 ppm

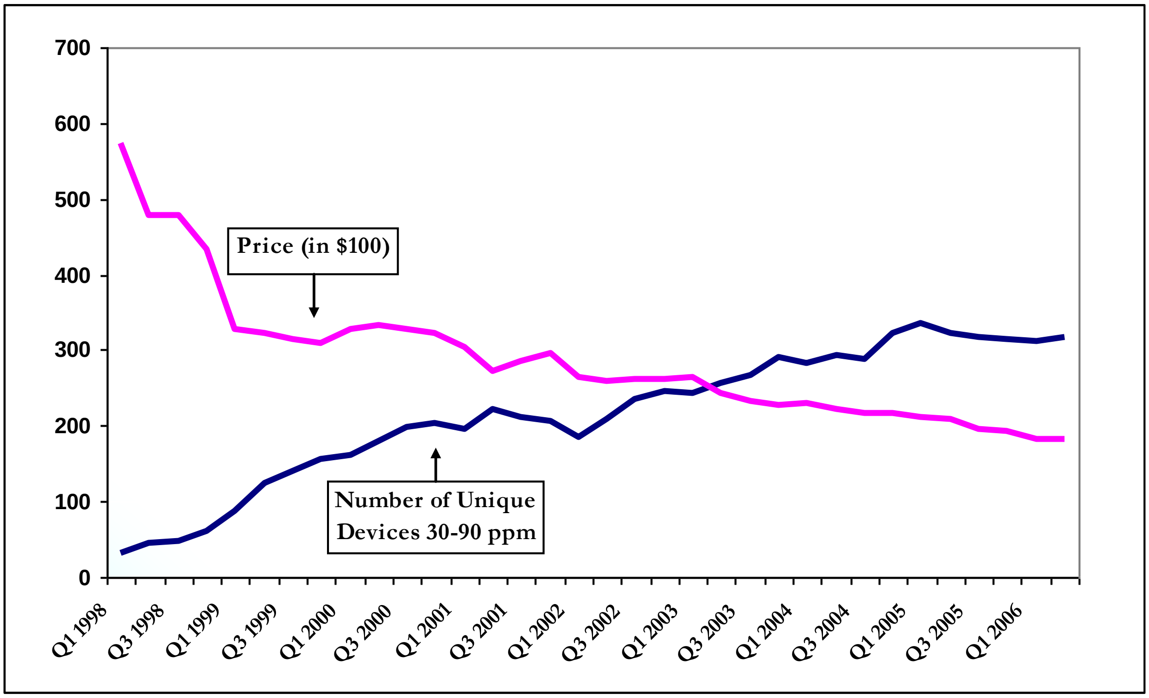

Figure 6 – SFC/MFC Market Trends

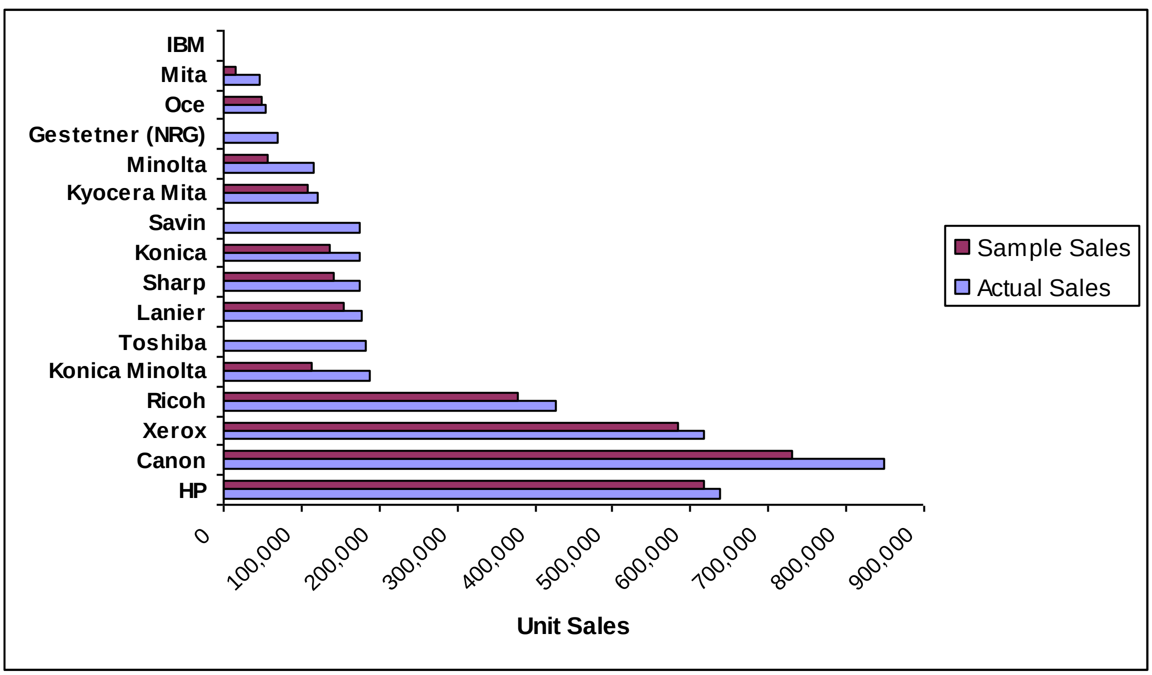

Figure 7 - Sample and Total Sales Volumes of MFC/SFC Devices of 30-90 ppm Sold To Business Units, 1998-2006

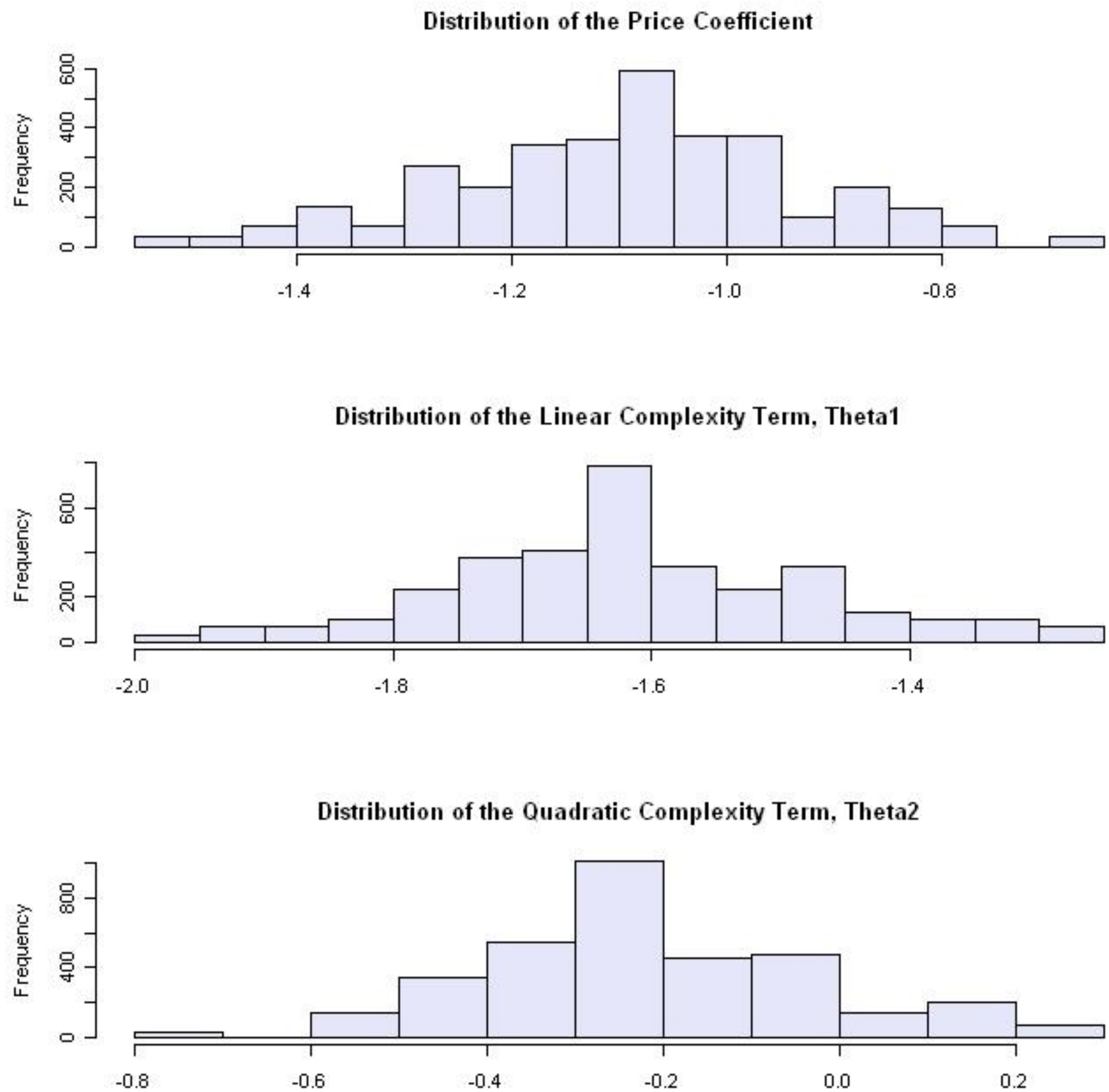

Figure 8 - Frequency Distribution of Heterogeneous Parameters

Table 1 – Characteristics Descriptive Statistics

Variable	Mean	Std Dev	Minimum	Maximum
Copy Speed - letter monochrome pages, per minute	50.76	16.26	31	90
Device Size, cubic inches	52,393.05	68,877.29	3,503	1,081,080
Paper capacity, number of pages	2,356.34	1,570.68	200	7,100
Warm-up time, seconds	213.71	168.39	12	780
Copier base	0.97	0.16	0	1

Table 2 – Features, Definitions and Descriptive Statistics

Feature	Mean	Stdev	Description
2 in 1	0.95	0.22	A feature that allows a user to combine and/or separate up to 4 or more different images onto one or both sides of a single page by reducing or enlarging each original if necessary. Also referred to as image combination or image separation.
Book Copy	0.99	0.08	A feature (key stroke) that is used to produce two single copies (one of each page) from a book that is spread open on a platen.
Booklet Mode	0.83	0.37	After the originals have been scanned into the memory, the copier automatically positions each image on the correct page, thus creating a signature. Copies can be folded to create a booklet.
Copy control	0.98	0.13	The number of access codes for tracking copy usage. The key operator can set copy limits and or reset the usage data for each code. Some copiers also feature a printer that prints usage summaries and or an interface so that this data can be exported to as ID mode, auditron or copy monitor.
Covers	0.92	0.27	The will automatically insert from and or back covers into the completed sets. The draws the cover from alternate paper sources, so thick or color stock may be used. Some copiers can copy on the front and or back of the cover.
Editing	0.27	0.45	Some machines have the ability to mask an area of an original and copy only that area or omit it from the copy.
Energy save	0.96	0.19	This is an energy conserving feature found on many copiers. Once a copy is complete and a preset period of time has elapsed, the copier automatically goes into a standby mode in which partial fusing heat is maintained so that a full warm-up period is not necessary for the next copy.
Erase	0.95	0.23	A feature that is used to erase a portion of the copy. There are several erase models (e.g.. to remove an area around the outer edge of a document or to remove the marks that commonly result when copying from books). Some copiers also allow the user to program the amount of the erased image area.
Image insert	0.38	0.49	Allows a user to merge images scanned from the platen glass, such as pages from a book, magazine or photos, with a document that has been scanned through the document feeder.
Image overlay	0.36	0.48	This feature allows the user to overlay the image of one original onto another to produce a single composite copy.
Image repeat	0.66	0.47	When image repeat is selected, the image on the original is repeatedly imaged onto the copy. The number of times the image is repeated on a copy is programmable by the user.

Image rotate	0.81	0.39	This feature automatically rotates the orientation of an original that is scanned into memory if the paper source that was selected becomes empty during the copy job, provided that the correct size is loaded in another drawer. some units have the ability to rotate the orientation of the original under other circumstances as well.
Interrupt	0.98	0.15	An interrupt key is used to halt copying at any point during a copy run. With some copiers, the status of the original multi-copy operation will be displayed after an interruption. Others will display the status of the run being made while storing data from the interrupted run.
Job build	0.77	0.42	When the number of originals exceeds the document feeder's capacity, a unit that offers job build will allow a user to batch feed originals that will be merged automatically once all the originals have been scanned. The number of originals that can be scanned is dependent on the amount of available memory.
Job programming	0.97	0.18	Job programming, a feature that allows the user to store preprogrammed jobs (e.g.. duplex, margin shift, finish 20 sets and recall them at the touch of a key or keys, is indicated. Some copiers can have preprogrammed jobs downloaded from memory cards.
Job time	0.33	0.47	Once a copy job has been input, the approximate time it will take to produce the job is displayed on the control panel. While copying is in progress, the display indicates the time remaining to complete the job.
Language	0.76	0.43	The language used in the messages on the control panel display can be changed to a language other than English.
Margin shift	1.00	0.05	The ability of a machine to shift the image to the right or left, on the front, back or both sides of a copy, to allow for binding or three hole punching. Most margin shift features are adjustable from 0-20 cm or up to 3/4 inch.
OHP interleaving	0.92	0.27	This feature enables the copier to insert a blank or copied sheet between each transparency when copying a set of transparencies.
Photo Mode	0.99	0.11	A feature that enhances the ability of a copier to copy halftones or photographs. The copier's magnetic roller, containing the toner and enveloper mixture, is slowed down for a more precise coating of the latent image on the photoconductor. This results in a higher quality reproduction of lines and halftones.
Poster Mode	0.23	0.42	A feature that enables the composition of a poster from a few prints.
Preset reduce/enlarge	1.00	0.06	The number of reset reduction and enlargement modes that are factory/service set and are accesses from a key on the control panel.
Program ahead	0.78	0.42	Allows the user to program the next copy job while a job is running.
Reversal	0.78	0.42	This feature automatically reverses an image on while on black or black on white.

Sheet insertion	0.71	0.45	Indication of the maximum number of insert sheets that the copier can automatically insert into completed sets at a certain preprogrammed positions is indicated. The copier draws the slip sheets stock from alternate paper sources, so thick or color stock may be used. Some copiers can copy on the front and/or back of the insert sheets.
Stamping	0.55	0.50	A feature that allows a user to place a stamp, such as a page or distribution number, date, watermark (confidential, draft, etc.) or an image previously stored into memory, onto a copy.
Timer	0.42	0.49	Copiers with this feature can be programmed to automatically turn on and off at user specified times. Some copiers may require an access code in order to operate during "off" hours.
XY zoom	0.92	0.27	This feature allows a user to select vertical and or horizontal reduction and/or enlargement rations independent of each other.
Copy	1.00	0.00	A Copy capability
Fax	0.33	0.47	A Fax capability
Print	0.94	0.24	A print capability
Scan	0.53	0.50	A Scan cpability
Sorter	0.03	0.17	A device used during the production of sets of copies. It consists of bins into which sets of copies are automatically inserted one sheet at a time. Usually this is an optional device
Stapler	0.99	0.08	A sorting device that features an online stapler.
Feeder	0.70	0.46	Document feeders.
Duplex	0.91	0.29	Two sided copies. Auto - the device can automatically produce two sided copies, without having the user to the sides of the paper. 1:2, 2:2, 2:1 option. Manual - manually duplexing small amounts, the user reinserts the paper.

Table 3 – Survey Results

Feature	Difficulty or Ease of Use Ratings				Familiarity Ratings			
	Mean	Stdev	Min	Max	Mean	Stdev	Min	Max
2 in 1	2.14	0.91	1	4	1.60	0.90	1	4
Feeder	1.37	0.79	1	4	3.37	1.11	1	4
Book Copy	2.12	0.99	1	5	2.35	1.02	1	4
Booklet Mode	2.50	1.29	1	5	1.75	0.89	1	4
Copy	1.15	0.55	1	4	3.86	0.35	3	4
Copy Control	2.44	1.40	1	5	1.78	1.02	1	4
Covers	2.19	0.94	1	4	1.80	0.87	1	4
Duplex Automatic	1.54	0.96	1	4	3.38	1.03	1	4
Editing	2.58	1.12	1	5	1.61	0.91	1	4
Energy-Save	1.58	1.45	1	12	3.11	1.09	1	4
Erase	2.41	1.14	1	5	1.67	0.98	1	4
Fax	1.85	1.01	1	5	3.28	0.84	1	4
Image Insert	2.73	1.10	1	5	1.43	0.75	1	4
Image Overlay	2.75	0.97	1	4	1.37	0.66	1	3
Image Repeat	2.21	1.13	1	5	1.59	0.98	1	4
Image Rotate	2.45	1.08	1	5	1.94	0.99	1	4
Interrupt	1.73	0.86	1	4	2.70	0.93	1	4
Job Build	2.43	1.16	1	5	1.67	0.90	1	4
Job Programming	2.65	1.30	1	5	1.68	0.88	1	4
Job Time	1.64	0.90	1	4	2.05	1.16	1	4
Language	1.84	0.88	1	5	2.00	0.88	1	4
Margin Shift	2.09	0.88	1	4	1.91	0.96	1	4
OHP interleaving	2.32	1.06	1	4	1.59	0.81	1	4
Photo Mode	1.72	0.92	1	4	2.47	1.08	1	4
Poster Mode	2.34	1.00	1	4	1.88	0.89	1	4
Preset Reduce/Enlarge	1.50	0.76	1	4	2.38	1.18	1	4
Print	1.11	0.47	1	5	3.92	0.37	1	4
Program Ahead	1.94	0.92	1	4	2.00	1.09	1	4
Reversal	1.95	0.83	1	4	1.61	0.75	1	4
Scan	1.78	1.04	1	5	3.43	0.74	1	4
Sheet Insertion	2.07	1.07	1	5	1.90	0.98	1	4
Sorter	1.27	0.65	1	4	2.58	1.20	1	4
Stamping	2.52	0.81	1	4	1.61	0.80	1	4
Stapler	1.34	0.62	1	4	3.07	1.13	1	4
Timer	2.64	1.22	1	5	1.50	0.61	1	3
XY zoom	1.95	1.04	1	5	2.05	1.08	1	4

Gender Distribution: 44% females, 56% males

Age Distribution: 34% 20-30 years old, 31% 30-50 years old, 35% above 50 years old

Role Distribution: 12% copier experts, 9% office administrators, 32% engineers, 21% managers, 26% starting business roles.

Table 4 - Homogeneous Model

	Model A - With Complexity $R^2 = 0.70$			**Model B - Without Complexity** $R^2 = 0.69$		
	Estimate	Std. Error	Significance Level	Estimate	Std. Error	Significance Level
Constant	-1.56	1.53	0.31	1.98	1.00	0.05
Price	**-1.01**	**0.10**	**<.0001**	**-0.95**	**0.10**	**<.0001**
Characteristics						
Canon	0.89	0.17	<.0001	0.73	0.16	<.0001
HP	-1.52	0.23	<.0001	-1.67	0.23	<.0001
IBM	-1.35	0.65	0.04	-1.57	0.65	0.02
Konica	0.22	0.22	0.31	0.12	0.21	0.58
Konica Minolta	-0.04	0.18	0.81	-0.14	0.18	0.44
Kyocera	-0.49	0.18	0.01	-0.61	0.18	0.00
Kyocera Mita	-0.53	0.24	0.03	-0.62	0.24	0.01
Lanier	-0.84	0.15	<.0001	-0.77	0.15	<.0001
Oce	-2.31	0.18	<.0001	-2.36	0.18	<.0001
Ricoh	-0.27	0.15	0.07	-0.25	0.15	0.10
Sharp	-0.58	0.14	<.0001	-0.63	0.14	<.0001
Print	0.91	0.13	<.0001	0.79	0.13	<.0001
Fax	0.04	0.09	0.69	-0.16	0.08	0.04
Scan	0.41	0.08	0.00	0.23	0.06	0.00
Copy speed	0.42	0.15	0.01	0.45	0.15	0.00
Device size	-0.15	0.07	0.04	-0.20	0.07	0.01
Paper capacity	0.26	0.06	0.00	0.19	0.06	0.00
base	3.33	0.44	<.0001	2.99	0.44	<.0001
Warm Up Time	-0.25	0.04	<.0001	-0.21	0.04	<.0001
N_Devices	-0.02	0.0005	<.0001	-0.02	0.00	<.0001
Fourth quarter	0.14	0.05	0.00	0.14	0.05	0.00
Complexity Variables						
θ_1	**-1.13**	**0.08**	**0.00**			
θ_2	**-0.22**	**0.33**	**0.01**			
Features						
2 in 1	0.65	0.22	0	0.42	0.19	0.03
copy control	0.96	0.26	0	0.91	0.23	<.0001
covers	-0.13	0.13	0.32	-0.31	0.13	0.02
editing	0.13	0.12	0.29	-0.25	0.09	0.01
energy save	-0.88	0.16	<.0001	-0.77	0.15	<.0001
erase	0.51	0.2	0.01	0.41	0.2	0.04
image insert	0.28	0.12	0.02	-0.12	0.08	0.13
image repeat	0.24	0.08	0	0.1	0.08	0.19
image rotate	0.3	0.11	0.01	0.03	0.09	0.7
interrupt	0.85	0.26	0	1.12	0.25	<.0001
job build	0.37	0.12	0	0.17	0.11	0.11
job programmming	-1.26	0.22	<.0001	-1.15	0.22	<.0001
job time	-0.08	0.08	0.33	-0.07	0.08	0.42
language	0.38	0.08	<.0001	0.41	0.09	<.0001
margin shift	0.29	0.58	0.61	-0.66	0.55	0.23
OHP interleaving	0.46	0.17	0.01	0.08	0.14	0.56
photo mode	-2.02	0.3	<.0001	-2.07	0.3	<.0001
poster mode	0.2	0.12	0.09	-0.08	0.09	0.4
program ahead	0.3	0.15	0.04	0.06	0.12	0.61
reversal	0.18	0.11	0.1	-0.11	0.08	0.19
sheet insertion	0.29	0.11	0.01	0.13	0.1	0.23
stamping	0.22	0.1	0.04	-0.1	0.06	0.1
timer	0.51	0.15	0	0.09	0.11	0.38
xy zoom	0.3	0.16	0.07	0.27	0.16	0.08
duplex	-0.29	0.13	0.03	-0.40	0.12	0.00
sort	0.03	0.17	0.87	-0.10	0.17	0.53
staple	2.03	0.32	<.0001	2.14	0.31	<.0001
Feeder	0.45	0.07	<.0001	0.45	0.08	<.0001

Table 5 - Heterogeneous Model

	Model C - With Complexity				Model D -Without Complexity			
	Means	Std Deviation	Firm Size	Firm Size^2	Means	Std Deviation	Firm Size	Firm Size^2
Constant	-3.86	-0.10	0.00	-0.03	0.72	-0.01	0.00	-0.15
Price	-1.10	-0.16	1.34	0.00	-0.94	0.03	1.03	0.00
Characteristics								
Canon	0.90				0.79			
HP	-1.36				-1.57			
IBM	-1.03				-1.28			
Konica	0.35				0.23			
Konica Minolta	-0.02				-0.07			
Kyocera	-0.62				-0.64			
Kyocera Mita	-0.51				-0.67			
Lanier	-0.82				-0.73			
Oce	-2.27				-2.23			
Ricoh	-0.30				-0.29			
Sharp	-0.66				-0.67			
Print	0.88				0.81			
Fax	0.16				-0.15			
Scan	0.44				0.21			
Copy speed	0.42				0.55			
Device size	-0.15				-0.13			
Paper capacity	0.32				0.20			
base	3.61				3.15			
Warm Up Time	-0.34				-0.24			
N_Devices	-0.02				-0.02			
Fourth quarter	0.20				0.14			
Complexity Variables								
θ_1	**-1.61**	**0.14**	**-0.15**	**0.00**				
θ_2	**-0.21**	**0.18**	**1.41**	**0.00**				
Features								
2 in 1	0.97				0.49			
copy control	1.2				0.91			
covers	-0.07				-0.32			
editing	0.26				-0.29			
energy save	-1.00				-0.77			
erase	0.68				0.49			
image insert	0.41				-0.19			
image repeat	0.3				0.12			
image rotate	0.43				0.07			
interrupt	0.71				1.13			
job build	0.39				0.16			
job programming	-1.22				-1.13			
job time	-0.06				-0.1			
language	0.4				0.42			
margin shift	0.77				-0.81			
OHP interleaving	0.6				0.14			
photo	-1.83				-2.02			
poster	0.29				-0.13			
program ahead	0.34				0.04			
reversal	0.27				-0.14			
sheet insertion	0.48				0.14			
stamping	0.27				-0.1			
timer	0.67				0.12			
xy zoom	0.31				0.24			
duplex	-0.22				-0.34			
sort	-0.05				-0.09			
staple	2.07				2.19			
Feeder	0.49				0.43			
Objective Function	4.69				27.51			