University of Pennsylvania
ScholarlyCommons

Publicly Accessible Penn Dissertations

1-1-2012

Services in Manufacturing Industries: Contributions to Quality and Competition

Jose Guajardo
University of Pennsylvania, josegu@wharton.upenn.edu

Follow this and additional works at: http://repository.upenn.edu/edissertations

Part of the Business Administration, Management, and Operations Commons, and the Management Sciences and Quantitative Methods Commons

Recommended Citation

Guajardo, Jose, "Services in Manufacturing Industries: Contributions to Quality and Competition" (2012). *Publicly Accessible Penn Dissertations*. 516.
http://repository.upenn.edu/edissertations/516

This paper is posted at ScholarlyCommons. http://repository.upenn.edu/edissertations/516
For more information, please contact libraryrepository@pobox.upenn.edu.

Services in Manufacturing Industries: Contributions to Quality and Competition

Abstract

Motivated by the increasingly important role of services in manufacturing industries, this dissertation examines implications of this trend for quality management and competition by firms engaged in the production of joint product-service offerings. Broadly defined, we study the following research questions: How do the service contracts offered by manufacturers affect product quality? How does consumer demand respond to product quality and service attributes when manufacturers compete on services? How are consumer intentions influenced by product quality and service quality perceptions, and how does consumer heterogeneity influence this relationship? We empirically study these questions in the aerospace, automobile and consumer electronics industries, respectively. In the first study, we examine the impact of Performance-Based Contracting on product reliability in an application in the aerospace industry (aircraft engines), and show that the incentive alignment induced by performance-based contracts positively influences product reliability by different mechanisms. In the second essay, we formulate and estimate a structural model to analyze the impact of service competition and product quality in the U.S. automobile industry. We show that the impact of service attributes (warranty length, service quality) on consumer demand critically depends on the firm's product quality. Finally, in the third essay (consumer electronics industry), we examine the joint influence of product quality and service quality perceptions on consumer intentions toward a brand, and show that consumer heterogeneity plays a significant role in defining this relationship. Collectively, our results suggest that the joint consideration of product and service is essential for the development of an effective competitive strategy and for the management of quality by manufacturing firms.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Operations & Information Management

First Advisor
Morris A. Cohen

Second Advisor
Gerard Cachon

Keywords
econometrics, manufacturing, operations management, operations strategy, service operations

Subject Categories
Business | Business Administration, Management, and Operations | Management Sciences and Quantitative Methods

SERVICES IN MANUFACTURING INDUSTRIES: CONTRIBUTIONS TO QUALITY AND

COMPETITION

José Andrés Guajardo Andrades

A DISSERTATION

in

Operations and Information Management

For the Graduate Group in Managerial Science and Applied Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the
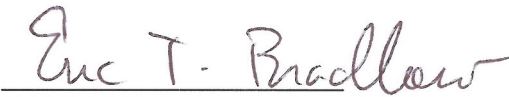
Degree of Doctor of Philosophy

2012

Supervisor of Dissertation

Morris A. Cohen, Professor of Operations and Information Management

Graduate Group Chairperson

Eric T. Bradlow, Professor of Marketing, Statistics, and Education

Dissertation Committee

Gérard P. Cachon, Professor of Operations and Information Management

Morris A. Cohen, Professor of Operations and Information Management

Serguei Netessine, Professor of Global Technology and Innovation

SERVICES IN MANUFACTURING INDUSTRIES: CONTRIBUTIONS TO QUALITY AND

COMPETITION

José Andrés Guajardo Andrades

A mis padres

ABSTRACT

SERVICES IN MANUFACTURING INDUSTRIES: CONTRIBUTIONS TO QUALITY

AND COMPETITION

José Andrés Guajardo Andrades

Morris A. Cohen

Motivated by the increasingly important role of services in manufacturing industries, this

dissertation examines implications of this trend for quality management and competition

by firms engaged in the production of joint product-service offerings. Broadly defined,

we study the following research questions: How do the service contracts offered by

manufacturers affect product quality? How does consumer demand respond to product

quality and service attributes when manufacturers compete on services? How are

consumer intentions influenced by product quality and service quality perceptions, and

how does consumer heterogeneity influence this relationship? We empirically study these

questions in the aerospace, automobile and consumer electronics industries, respectively.

In the first study, we examine the impact of Performance-Based Contracting on product

reliability in an application in the aerospace industry (aircraft engines), and show that the

incentive alignment induced by performance-based contracts positively influences

product reliability by different mechanisms. In the second essay, we formulate and

estimate a structural model to analyze the impact of service competition and product

quality in the U.S. automobile industry. We show that the impact of service attributes

(warranty length, service quality) on consumer demand critically depends on the firm's

iv

product quality. Finally, in the third essay (consumer electronics industry), we examine the joint influence of product quality and service quality perceptions on consumer intentions toward a brand, and show that consumer heterogeneity plays a significant role in defining this relationship. Collectively, our results suggest that the joint consideration of product and service is essential for the development of an effective competitive strategy and for the management of quality by manufacturing firms.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Manufacturers are increasingly providing not only physical goods to their customers, but also supporting services. This trend can be observed in industry sectors as diverse as aerospace and defense, automobiles, computers, consumer electronics, and industrial equipment, among many others. The provision of services in traditional manufacturing industries opens up a number of opportunities and challenges for firms. Two important related aspects of it are service competition and the joint management of quality for products and supporting services. While a number of theories and existing knowledge in Operations Management, Economics, and Marketing, are relevant for different aspects of these problems, little empirical evidence exists for understanding how firm decision-making regarding quality and competition affect consumer decision-making in the presence of joint product-service offerings. This dissertation provides empirical evidence that contributes to a better understanding of this problem, by analyzing three central

1

aspects of it in three important and well established manufacturing industries, i.e., aerospace, automobiles, and consumer electronics. Above all else, we seek to characterize how product and service attributes interact in non-trivial ways, how the resulting interaction affects consumer decision-making, and what implications can be derived for firm decision-making.

This dissertation is organized in the form of three self-contained essays. In each of them, we employ econometric tools and statistical methods to analyze data that reflects actual consumers' and firms' decision-making in these industries. In the first essay (chapter 2), we study how service contracts offered by the manufacturer influence product quality. More specifically, using a proprietary dataset provided by a major manufacturer of aircraft engines, we empirically investigate how product reliability is impacted by use of two different types of after-sales maintenance support contracts: time and material contracts (T&MC) and performance-based contracts (PBC). A number of competing arguments based on the theory of incentives establish why product reliability may increase or decrease under PBC, justifying the need for empirical analysis. We build a two-stage econometric model that explicitly accounts for the endogeneity of contract choices, and find evidence of a positive and significant effect of PBC on product reliability. The estimation of our model indicates that product reliability is higher by 25-40% under PBC compared to under T&MC, once the endogeneity of contract choice is taken into account. Our results are consistent with two mechanisms for reliability improvement under PBC: more frequent scheduled maintenance and better care performed in each maintenance event.

In the second essay (chapter 3), we focus on the more general question of how service attributes and product quality jointly influence consumer demand when manufacturers compete on services. For this purpose, we formulate a structural econometric model to analyze the impact of service attributes (warranty length, after-sales service quality) on consumer demand in the U.S. automobile industry. Our results indicate that service attributes play a compensatory role with respect to product quality, i.e., the impact of warranty length and service quality on demand increases when product quality decreases. Conversely, both service metrics are complementary with respect to demand, i.e., the better the service quality, the higher the marginal effect of longer warranties. Our results estimate a median willingness to pay for one year of warranty of about $850, which is equivalent to 2.5% of the average vehicle price in our sample. We find that, for an average car in our sample, the effect on demand of a 1% price decrease is equivalent to increasing product quality by 3%, which is in turn equivalent to increasing the warranty length by 9%.

In the third essay (chapter 4), we study how product quality and service quality perceptions affect customer loyalty in an application to the consumer electronics industry. Our analysis, based on a survey of consumers, reveals that both attributes are jointly associated with customer intentions. On average, product quality and service quality perceptions act as complements in terms of customer likelihood to recommend the brand. More importantly, we analyze the moderating role of customer characteristics on the relative influence of product and service quality perceptions, and document significant heterogeneity in service-sensitivity in the population. Our results indicate, in particular,

that there are important differences across gender and income. Service quality perceptions play a more important role for women and for high income segments, relative to men and low income groups, for which product quality perceptions are found to be a more important driver of customer loyalty.

Collectively, the empirical evidence provided in this dissertation suggests that the joint consideration of product and service is essential for the development of an effective competitive strategy and for the management of quality by firms. In the concluding chapter, we elaborate on different aspects of this general conclusion, as well as on future research directions that emerge from this dissertation.

Chapter 2

Impact of Performance-Based Contracting on Product Reliability

2.1. Introduction

The movement towards a service-based economy has led many manufacturing firms to recognize the strategic importance of after-sales product support services that enable the availability of properly functioning products. In many industries, such services are a major source of revenue, profit, and growth and thus act as a source of sustainable competitive advantage (Cohen et al. 2006). This is especially true in those industries where products are complex and the consequences of product downtime can be severe. Moreover, when products have relatively long life cycles (e.g., aircraft, engines, semiconductor fab equipment, medical imaging devices, etc.), they present the firms supplying after-sales support with ample opportunities to provide repair and maintenance

services. As many OEM firms in such industries reposition themselves to become service providers, it has become critical for them to evaluate and define contractual relationships with their customers for the provision of after-sales support. Traditionally, after-sales services have been performed under time and material contracts (T&MC), under which the supplier is compensated for the amount of resources consumed (such as spare parts and labor) whenever product maintenance is required. However, a new form of a support contract has emerged in recent years: performance-based contracts (PBC). Under PBC, also referred to as Performance-based Logistics (PBL) in the defense sector and Power by the Hour® in the commercial sector, a supplier is paid based on the realized outcome of customer value. For example, an airline customer pays an engine service provider in proportion to the number of aircraft flying hours, which is affected by engine up-time (i.e., the number of hours the engine was available for use), and which determines the value derived by the customer.

PBC has been popularized in the aerospace industry in particular because it was recognized early on that it brings a potential benefit of aligning incentives among customers and suppliers. PBC compensates the supplier based on the same outcome that the customer cares about (i.e., product utilization), and hence the supplier is motivated to increase product performance, associated with metrics such as product reliability and availability. It has been noted that the risk of supplier moral hazard is high under the traditional T&MC since the provision of services that the customer procures under it, such as the spare parts and repair labor, is typically a high margin and profitable activity

for suppliers. PBC can mitigate this problem, since under this contract the supplier is responsible for these costs.

Past theoretical results support the notion that adoption of PBC can result in increased product performance at a lower cost (Kim et al. 2011, Hypko et al. 2010, Randall et al. 2010). Empirical evidence to support this conclusion, however, is not conclusive. Kirk and DePalma (2005) analyze a Navy PBC program and, based on a review of historical repair frequency data for several programs, offer the following mild observation: "there is some evidence that the PBC contract may have helped to improve availability and reliability." To the question "Do you have evidence of higher performance/lower cost based on past or current [PBC] programs?" included in a recent survey conducted in the aerospace industry (Newsome 2008), the responses were decidedly mixed: "Yes" – 33%, "No" – 36%, "Too early to tell" – 31 %. A recent Government Accountability Office (GAO) report on PBC in the defense industry presents a similar view, and points to the limitation of existing studies: "Many DoD [Department of Defense] program offices that implemented [PBC] arrangements have limited cost data, and various other factors – such as the lack of business case analyses – further limit an evaluation of the costs of this support strategy. Available data from the programs GAO reviewed indicated mixed results" (GAO 2008).[1] A 2009 DoD study (DoD 2009), which reviewed over 30 weapon system programs, reported that it is widely accepted that PBC leads to higher material

---

[1]Reflecting the intense scrutiny over PBC employed in defense and other government services, House of Representatives recently held a hearing on the benefits and costs of PBC. See "Performance-based Acquisitions: Creating Solutions or Causing Problems?" Full Committee Thursday, May 08, 2008. http://homeland.house.gov/hearings/index.asp?ID=136.

availability, but that conclusions concerning cost reduction and other performance implications such as reliability improvement are less clear.

Our review of industry studies indicate that they are primarily based on comparisons of observed average measures or on simple regression analyses that leave out numerous confounding factors, bias adjustments, and statistical robustness checks. As we discuss below, there are several complexities that make the quantification of the impact of PBC on metrics such as product reliability a challenging task. Indeed, the fact that prior evaluations of PBC have ignored the modeling of such complexities helps to explain why the debate concerning the potential for PBC to provide significant improvement of product performance remains unresolved at this time. The main goal of this study is to conduct a thorough examination of the competing arguments that associate PBC with improved product performance in a specific context (i.e., aircraft engine reliability) and empirically evaluate them, thereby offering a new perspective to this issue that has captured practitioners' attention in the past few years since PBC became widely adopted.

Our focus on the research question "Does PBC result in reliability improvement?" is motivated by the intense debate currently underway among practitioners and policy makers concerning the same issue. Indeed, product reliability is regarded as a prime performance metric in the aerospace industry, not only because it is an important driver of product utilization, but also because in this industry product failures due to imperfect reliability lead to direct and large financial losses as well as to potentially disastrous consequences (e.g., loss of life). A good illustration of the role of reliability in this industry is the recent engine failure incident that affected the Qantas Airways Ltd. which

received extensive media coverage worldwide. The mid-air blowout of one of Qantas' Trent 900 engines frightened all airline passengers, and had direct short-term financial effects on both Qantas and Rolls-Royce, the engine manufacturer; the stock prices of the two companies fell by roughly 10% in the days/weeks after this incident (Clark and Mouawad 2010). In the closely-related defense sector, the DoD considers achieving a high level of product reliability as one of the three essential elements of enabling mission capability, along with availability and maintainability (DoD 2005).

In the academic literature, a few studies have proposed models that link PBC with reliability improvement. Kim et al. (2011), based on a game-theoretic model analysis under a limited set of assumptions, suggest that the answer to the reliability improvement question is an unqualified "yes". However, as we demonstrate in this study, there are several real-world and theoretical considerations that are not captured in Kim et al. (2011) which make the reasoning much more nuanced and the outcome less certain. In fact, as we discuss in more detail in Section 2.3, there are a number of competing arguments that may lead to a positive, negative, or neutral answer to the question. In this study, we address this question by presenting the results of an empirical analysis performed on a unique proprietary dataset provided by Rolls-Royce, a leading manufacturer and a service provider for aircraft engines. Rolls-Royce, like many companies in the aerospace industry, offers its customers the two types of contracts we have mentioned: T&MC and PBC.

One of the key challenges in measuring the effect of PBC on product reliability is the presence of heterogeneous customer preferences for contract types that requires us to

model customers' contract choice decisions explicitly. Without controlling for this endogenous contract selection process, as we demonstrate, one could erroneously conclude that there is no statistically significant effect of contract type on product reliability. Instead, we propose a two-stage framework that explicitly deals with the endogeneity inherent in contract choice by a customer, and we provide evidence at the 95% confidence level that product reliability under PBC is in fact higher by about 25-40% compared to that under T&MC. The conclusions from this analysis are robust to a large number of alternate specifications and modeling assumptions. Thus, our findings lend support to the view that adoption of PBC results in product performance enhancement. In addition, our results are consistent with two separate mechanisms by which reliability improvement can be achieved under PBC: more frequent scheduled maintenance and better care performed in each maintenance event. The latter impact can be accomplished by such activities as conducting more thorough checks leading to better identification of defects, preemptive parts replacements, and possible product re-design.

2.2. Related Literature

While there is an abundance of theoretical papers on the subject of supply chain contracting in the Operations Management (OM) literature (e.g., Cachon 2003 provides an extensive review of more than 200 papers in this area), most of them consider simplified settings in which contract choices made by customers with heterogeneous preferences are not taken into account. In contrast, endogenous contract choice is a central feature in our analysis. Heterogeneity in customer preferences can be explained in

many different ways, but especially relevant to our problem context is heterogeneity created by information asymmetry, namely, by adverse selection and moral hazard. In this respect, OM papers such as Corbett and de Groote (2000) and Iyer et al. (2005) that analyze the optimal design of contracts in the presence of information asymmetry are relevant to our analysis. We note, however, that these papers focus mostly on how firms make choices from a menu of contract terms within a single class of contracts (e.g., menu of price-quantity pairs) – as suggested by the standard mechanism design theory – rather than choose one contract type from multiple classes (e.g., T&MC vs. PBC). The latter situation, which we analyze in this study, is not thoroughly investigated in the literature.

Although numerous papers based on game-theoretic analyses have produced an abundance of predictions on how supply chain contracts should be optimally structured, empirical validations in the OM literature are scarce. Indeed, as Cachon (2003) notes, "the literature contains a considerable amount of theory, but an embarrassingly paltry amount of empiricism." Although a small number of empirical papers do exist under the broad theme of supply chain coordination – for example, Novak and Eppinger (2001) and Novak and Stern (2008, 2009) examine the impact of product characteristics on vertical integration decisions in the automobile industry – the issue of contractual incentives has not received comparable empirical scrutiny. This study contributes to the literature by providing an empirical analysis of unique transaction data that sheds light on the role of supply chain contracting in affecting product reliability, a key variable of interest in the OM and quality management areas. There are other empirical OM papers that are related to our study, including Ramdas and Randall (2008; product reliability in the automotive

industry), Deshpande et al. (2003a; after-sales services in the defense industry), and Terwiesch et al. (2005; incentive conflicts in the semiconductor industry). However, none of these papers explicitly consider the incentives created by contract type choices.

In contrast to the current state of empirical research on contracting in the OM literature, progress has been made in other areas. Chiappori and Salanie (2003) detail the development of the contracting literature in economics and point out that, just as in OM, early contracting papers were predominantly theoretical and did not account for the process of contract selection by heterogeneous agents. Only recently have the economists started addressing this shortcoming, focusing mainly on behavior of individuals (e.g., Ackerberg and Botticini 2002). A few very recent exceptions (some still unpublished) have investigated revenue sharing contracts in the video rental industry featuring contract self-selection (Mortimer 2008, Ho et al. 2010, 2011). Not surprisingly, given the different nature of the problems of engine maintenance vs. video rentals, our analysis involves contract types, incentives mechanisms, and performance metrics that are quite distinct from theirs. Except for these examples, there is no empirical paper that considers contract self-selection in supply chains that we are aware of. Similarly, some empirical papers have examined the influence of variants of PBC in applications in service sectors with a focus on the behavior of individuals (e.g., Lazear 2000, Prendergast 2002 in labor economics; Lu et al. 2003, Shen 2003 in health care; Heinrich 2002 in public policy). However, whether the findings from such firm-to-individuals settings extend to firm-to-firm settings is unclear (because of differences in buyer characteristics such as price sensitivity, attitude towards risk, access to capital, number of stakeholders involved, etc.),

and comparable studies in the latter setting are sparse. Finally, in a related but different context of offshore software development, Gopal et al. (2003) analyze the impact of fixed-price vs. T&MC on software vendor profits (Gopal and Sivaramakrishnan 2008 perform a similar analysis). As we do, they use a two-stage modeling approach that includes the determinants of contract choice, but in the context of procurement of an intangible product (software). In addition, their research differs from ours because the contracts they study are not based on product performance, which is the central focus of our analysis.

This research contributes to the operations management (OM) literature by being one of the few studies that empirically examine and test predictions from supply chain contracting models, thereby bridging the gap between theory and empirical evidence in this area. Moreover, we add a new dimension to the literature by showing that customers' product support contract choice decisions are integral in linking a contract type with product performance. This finding counters the arguments found in Kim et al. (2011), Hypko et al. (2010), and Randall et al. (2010) who suggest that it is sufficient to simply count the number of product failures under the two contracts (T&MC and PBC), take averages, and infer that one contract leads to higher reliability (they all suggest PBC does). We show that this is not necessarily the case; without accounting for customers' self-selection of contracts (and thus failing to consider endogeneity), this approach can lead to misleading conclusions.

2.3. Industry Background and Theoretical Motivation

Our research setting is in the maintenance, repair and overhaul (MRO) market for commercial aircraft. According to Standard and Poor's (2011), the MRO sector generated revenues of $111 billion in 2009, of which $62 billion was attributed to military MRO, $42.7 billion to air transport (commercial aircraft) MRO, and $6.2 billion to business and general aviation MRO. More generally, reported statistics (see Cohen et al. 2006 and the references therein) indicate that sales of spare parts and after-sales services in the U.S. represents 8% of annual domestic product, meaning that customers spend approximately $1 trillion every year to maintain assets they already own.

In recent years, customers in the commercial aerospace industry, i.e., the airlines, have increasingly adopted outsourcing strategies for MRO services in order to focus on their core competencies and to reduce costs. This trend has led to expansion of the range of MRO services offered by suppliers of various types of aircraft subsystems (e.g., hydraulic power systems, engines, avionics systems). A unique feature of this market is that it is quite common for the OEM's that manufacture the systems and subsystems to offer support services for their own products. This is due to the highly customized and complex nature of the products, which can make it difficult for a third party to provide the level of product care that customers require. Consequently, the provision of such services has been very profitable for OEM firms such as Pratt & Whitney, General Electric Co., Rolls-Royce, Boeing and Lockheed Martin.

Not surprisingly, customers, who often end up paying far more for after-sales services than for the products themselves over the life of product usage, have started to demand

that suppliers provide contract options that reduce their cost for product support. PBC was introduced as a response to this demand, and it is Rolls-Royce, one of the major aircraft engine manufacturers, that is universally recognized as the company that pioneered this concept. Rolls-Royce offers two different types of contracts to their customers: T&MC and PBC. The main distinction between the two is the basis of compensation to the supplier. Namely, under PBC the customer agrees to pay a fee in proportion to aircraft flying hours, which in turn is affected by the availability of all major aircraft sub-systems (including engines). Flying hours – a key measure of product utilization in the aerospace industry – depends heavily on subsystem reliability as well as other factors such as the stocking levels of spare parts inventory and the speed of repair at maintenance depots. In our dataset, we observe information on product reliability but not on the other factors that may also drive product utilization, explaining our focus on this single metric in our analysis. In defining product reliability, it is important to make a distinction between unplanned (or unscheduled) maintenance and planned (or scheduled) maintenance events. In our analysis we associate product reliability with unplanned maintenance events because they represent unforeseen disruptions that can lead to the loss of customer value and to the costly measures that are required to mitigate their impact. Planned maintenance events, on the other hand, reflect managerial interventions that are scheduled in advance, and the time it takes to complete this type of maintenance has much lower impact on product utilization, relative to unplanned maintenance. In fact, the mean repair time for unplanned maintenance events in our sample is in the order of several weeks.

The main focus of this research is on the question "Does PBC result in reliability improvement?" and in the quantification of such effect (if any). We will hypothesize that the answer to this question is positive, based on the basic theoretical argument that incentive alignment is enhanced under PBC. However, as we will discuss next, the presence of information asymmetry in our setting leads to potential double moral hazard and adverse selection effects, which, besides making the answer to the question not obvious, gives rise to a number of competing mechanisms by which product reliability can be affected throughout the product support process under different contract types.

We start by discussing the supplier side. First, note that the supplier has an opportunity to exert effort to improve reliability of existing products when they are in her possession during repair and maintenance processes. An appropriate framework to study the supplier's incentive to do so under T&MC or PBC is the moral hazard model. Our setting fits this framework because, in most practical cases, the supplier's reliability improvement efforts are discretionary and unobservable to the customer and are influenced by contractual incentives. This is especially true for reliability issues that require engine overhauls, since, in such instances, the customer loses visibility to the supplier's repair capabilities as soon as she sends a defective engine to a repair depot operated by the supplier. Recall that the supplier is compensated in proportion to the product usage under PBC, whereas under T&MC, he is paid each time the product fails and the customer requests repairs. Since product utilization increases with reliability, then, a supplier under PBC is motivated to improve product reliability in order to maximize his revenue. On the other hand, a supplier under T&MC has a skewed

incentive that may lead him not to invest in reliability improvement or even degrade reliability, since each failure incident represents an opportunity for him to generate revenue by selling high-margin spare parts and charge for labor and other resources consumed. Thus, the supplier's service revenue could actually grow with lower reliability. This reasoning was formalized in Kim et al. (2011), who reach a similar conclusion in their analysis based on a game-theoretic model. In addition, this intuitive relationship between a contract type and reliability outcome is supported by some industry reports.[2]

If it was the case that reliability is the only variable that can be influenced by a managerial decision, based on the argument above, it is relatively straightforward to hypothesize that PBC is superior to T&MC in incentivizing the supplier to improve reliability. There are, however, two theoretical constructs that can be hypothesized to moderate the aforementioned discussion. The first issue is the multitasking aspect (Holmström and Milgrom 1991): given that the supplier can invest in not only reliability improvement but also in parts inventory, repair time reduction, and other efficiency gains, is it necessarily true that the level of reliability improvement is significantly higher under PBC than under T&MC? It is possible, for example, that improving reliability is

---

[2] According to Thomas (2005), "… Rolls-Royce had officially won praise from the US Navy for its innovative 'PBtH' support for the F405 engine". According to Captain Win Everett, Program Manager for the US Navy's Undergraduate Flight Training Systems at NATC 'Patuxent River' (Maryland), 'under Rolls-Royce, engine availability has exceeded the current target of 85%, the average time between engine removals has increased from 700 hours to over 900 hours, and expected engine removals have fallen by 15 per cent.'" In addition, Business Wire (2008) cites reliability improvements after introduction of PBC contracts by Rolls-Royce. In PBC environments not directly related to engines or Rolls-Royce, Geary and Vitasek (2008) observe that, "... there were also 90 improvements made to the APU (auxiliary power unit), with 20 of those being reliability improvements", and "the contract with Raytheon... the system design and support concept used in this program have resulted in a 200% improvement in MTBOF (mean time between operations failures) and a 400% improvement in mean time to repair."

17

too costly and therefore the supplier chooses other means to increase the product utilization. Second, an argument can be made that a reputational concern will prevent the supplier from neglecting reliability improvement, regardless of which contract he is subject to. One may further argue that the supplier is required to deliver the highest level of reliability that he can provide, since the customers in the airline industry are under heavy regulatory mandates. If this reasoning is correct and these effects in fact dominate, we would expect to see no significant relationship between a contract type and reliability. Thus, whereas a simple reasoning that focuses solely on the supplier-side moral hazard would lead us to believe that PBC is more effective in incentivizing the supplier to improve reliability than T&MC, arguments based on multitasking and reputation effects make the ultimate outcome of the supplier's actions not obvious.

Next, we turn our attention to the customers. It can be argued that, regardless of the contract type for after-sales services, it is unambiguously in the interest of a customer to make sure that high reliability is maintained given the high opportunity costs of having an aircraft on the ground. However, a careful examination of the customers' incentives under each contract type reveals that the effects of double moral hazard may be significant in this setting. This is because reliability is a function of not only the supplier's actions, but also the level of care that the customer exerts and the pattern of usage that the customer adopts. For customers, PBC can be viewed as a form of insurance that provides them with protection against unforeseen out-of-pocket charges that are incurred when an unexpected product malfunction occurs (unlike T&MC, under which the customer has to pay for spare parts and labor after those events). Therefore, it can be

hypothesized that – relative to T&MC customers who are responsible for all costs associated with product failures – PBC customers may tend to operate their product with less care, contributing to wear-and-tear and thus making the product more prone to failures (Padmanabhan 1995 builds a model based on the same intuition, applying it to the case of extended warranties). If such situations are common, then we expect that the impact of a contract type on reliability to be not significant or may even be reversed, i.e., PBC may in fact degrade reliability (which would happen if the customer's moral hazard is dominant).

In summary, although the basic reasoning suggests that adopting PBC will result in higher product reliability based on an alignment of incentives, other theoretical arguments moderate or counter this view. Earlier theoretical studies (Kim 2011, Kim et al. 2011) point to the former conclusion, but they do not capture many of the confounding factors that we have identified. Thus, this question is best answered empirically. We state the hypothesis of our empirical research as follows.

Hypothesis: Reliability under PBC is significantly higher than reliability under T&MC.

Finally, and as we will discuss further in the following sections, it is important to account for the customers' contract selection mechanism in order to isolate the effect of PBC on product reliability. As we have described in our discussion so far, the setting we study is characterized by information asymmetry. In particular, a customer may keep his/her product usage profile private. Then, because of the insurance role of PBC, a customer who tends to overuse the product or who possesses products that tend to

undergo high levels of stress would prefer PBC. If such an adverse selection motive is significant, then we expect to see a high rate of PBC selection among the customers who tend to use the supplier's maintenance services more often. Since the work of Rothschild and Stiglitz (1976), contract self-selection has been recognized as one of the important determinants of insurance markets, and it is also a feature that we expect to be relevant in our setting given the insurance role of PBC. While selection effects have been studied in the case of insurance markets and empirical testing of the associated theories has grown in the last ten years, empirical evidence indicates that information asymmetry leads to self-selection in some sectors but not in others (Chiappori and Salanie 2003). Additionally, some examples of advantageous selection in the insurance markets have been reported in the recent literature (Einav and Finkelstein 2011). Thus, the precise nature of a selection mechanism is far from obvious and is likely to depend on specific settings. Ultimately, whether the presence of contract self-selection is significant in our setting and which outcome the selection leads to are empirical questions, which we examine in our analysis.

2.4. Data

The dataset which Rolls-Royce made available to us consists of five years of data (July 2002 - July 2007, hereafter the observation period) of maintenance events (engine removals) for different models of aircraft engines produced by Rolls-Royce. A removal of the aircraft engine may be necessary due to a part failure (unplanned removal) or for maintenance purposes (planned removal), resulting in a shop visit to the service provider.

For a better understanding of the data, it is useful to describe more precisely how the data for this research was provided to us. We obtained two different data files: a spreadsheet containing all of the removal events between July 2002 and July 2007 for the engines in our sample (hereafter called the "removals file"), and a spreadsheet containing a list with all engines registered with Rolls-Royce for each customer (hereafter called the "engines file"). For each removal of an engine unit, a list of the relevant information contained in the removals file is as follows: engine unit ID, engine model, date at which the engine entered the repair shop, cumulative flying hours ("time since new", TSN) at the time of a shop visit, cumulative cycles (CSN, defined in Table 2.1) at the time of a shop visit, removal type (planned or unplanned), aircraft tail number in which the engine is installed, aircraft model ID, ID of the customer that owns the product, the contract type (T&MC or PBC) under which the product receives service. The engines file, on the other hand, contains a list with all of the engines registered with Rolls-Royce at the end of the observation period, i.e., July 2007. For each engine in this file we know: engine unit ID, engine model, cumulative aircraft flying hours and cycles at the end of the observation period, and the ID of the customer that owns the aircraft.

After cleaning data by removing inconsistent observations (e.g., for a given unit, reported flying hours at a shop visit in 2006 are less than the flying hours reported in a shop visit in 2005), our sample consisted of 763 engine units for which at least one engine removal is observed in the 5 years observation period. There are essentially two engine models in our sample: for one of these product models there are 3 different versions, and for the other there are 2 different versions. The engine models are installed

in three different types of aircraft. For all types of product, we observe engine units covered by either PBC or T&MC. In the sample of 763 engine units with removals, 21.4% are covered by T&MC, 78.6% are covered by PBC. Among the pool of 763 engines, 305 of them (40%) had at least one unplanned removal during the observation period. These 305 units are associated with 48 different customers.

As mentioned in the previous section, our approach for capturing the impact of a contract type on product reliability focuses on unplanned removal events. Unplanned engine removals are highly undesirable events for aircraft owners since an aircraft on-the-ground generally results in high opportunity costs, with estimates as high as hundreds of thousands of dollars per day for an unplanned removal for a fully loaded wide body commercial aircraft. Furthermore, there is always a possibility that unplanned failure may lead to a catastrophic event. In contrast, in the case of planned removals the shop visit is programmed in advance, and the appropriate replacements can be scheduled to be available to avoid having an aircraft on-the-ground. Indeed, we have observed in practice that a release of a product on the scheduled completion date for a planned maintenance event has a high priority. Therefore, whether a removal is planned or unplanned largely determines the downtime of an engine and hence of an aircraft.

Hence, for the purpose of studying product reliability, our main analysis focuses on the sample of 305 engines with at least one unplanned removal. We note that, in the sample of 305 engines with unplanned removals, 21.3% of the engines are covered by T&MC, and 78.7% are covered by PBC, i.e., exactly the same proportions as observed in the full sample: 39.9% of the T&MC engines had unplanned removals, and 40% of the

PBC engines had unplanned removals. The observed proportions suggest that focusing the analysis on unplanned removals does not generate, a priori, a sample bias (see Section 2.7 for further discussion of sampling issues and robustness checks). We note also that reliability is, by definition, a product level variable, i.e., what fails is an individual product. Consequently, the unit of analysis of interest in this research is an engine unit. As will be noted throughout, focusing on an engine unit allows us to capture several details associated with its reliability (initial condition of the product, product type, etc.). From the two data files described earlier, we are able to obtain and calculate several variables of interest which describe both the product and the customer. Table 2.1 provides definitions and descriptive statistics for the variables used in our analysis.

While there are several possible ways to approach the measurement of product reliability, we have chosen the mean time between unplanned removals (MTBUR) to be the measure employed in our main analysis since (1) it is in fact a key reliability metric that practitioners in the aerospace and other industries constantly monitor and (2) it can be computed from available data. MTBUR represents the average time that a product is used without the need for an unplanned removal for repair and maintenance purposes. Unlike some other metrics (such as mean time between removals, which includes both planned and unplanned removals), MTBUR is a good representation of the physical reliability inherent in the product since an unplanned removal event occurs only when an engine fails randomly, free of managerial interventions, unlike planned removals. In the rest of our discussion we use the terms product reliability and mean time between unplanned removals interchangeably.

Unit	No. obs.	Variable Name	Variable definition	Mean	Median	Std. Dev.	Min	Max
Engine	305	ini_age	Time since new in July 2002 $(TSN(T_B))$	3,231	2,962	2,760	0	12,916
Engine	305	final_age	Time since new in July 2007 $(TSN(T_E))$	10,424	11,094	4,819	363	21,600
Engine	305	nyears_em_intro	Years since engine model introduced	3.1	4	1.7	0	5
Engine	305	eng_avgflighttime	Average flight time $TSN(T_E)/CSN(T_E)$	1.22	1.10	0.27	0.82	2.50
Engine	305	n_unplanned_rem	No. of observed unplanned removals	1.21	1	0.46	1	3
Engine	305	n_planned_rem	No. of observed planned removals	0.52	0	0.64	0	3
Customer	48	fleetsize	No. of engine units registered at Rolls-Royce	40	8	106.2	2	593
Customer	48	fleetmix	No. of engine models registered at Rolls-Royce	1.65	1	1.0	1	5

Table 2.1: Definition of variables and descriptive statistics. TSN = Time since new, measured in flying hours. CSN = Number of cycles since new (a cycle is defined as the interval between a takeoff and a landing). The variables ini_age, final_age, and eng_avgflighttime are all measured in flying hours.

Although MTBUR is an appropriate metric of product reliability, calculating it still poses nontrivial issues for our analysis because unplanned removals – and, in fact, removals of any kind – are quite rare events in our data. In our observation period of 5 years, the majority of the products in the dataset (81.6%) exhibit only one unplanned removal; the remaining units had either two or three unplanned removals (16.1% and 2.3% of the sample, respectively). Additionally, the data suffers from censoring since information on any unplanned removals that occurred before July 2002 or after July 2007 is excluded. Defining a rule to compute the MTBUR, therefore, is a challenging task. We

illustrate the problem and the procedure used to calculate the MTBUR with an example
(see Figure 2.1).
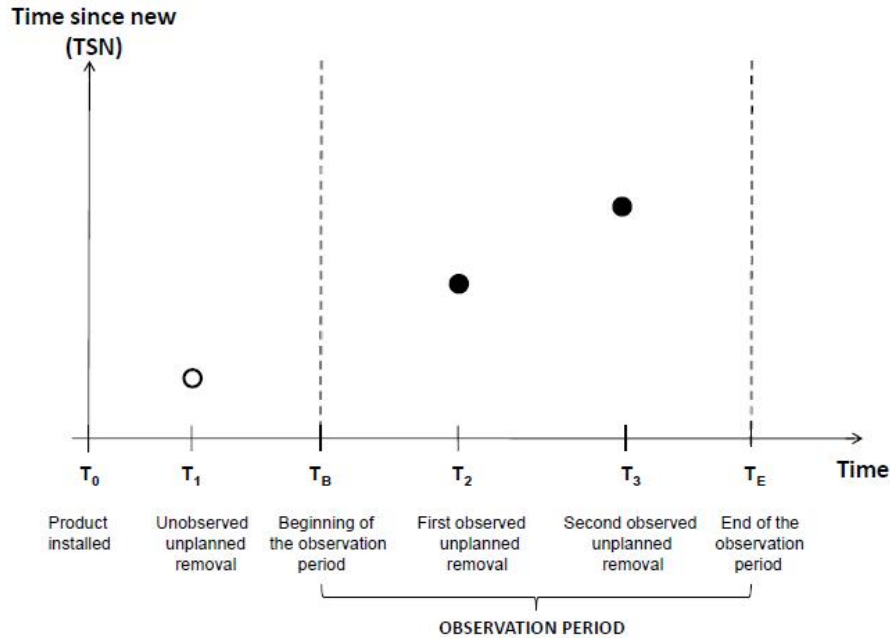
Figure 2.1: Example of an unplanned removal sequence.

Consider a product that was installed in an aircraft at time T_0, before the beginning of
the observation period T_B (July 2002 in our case). Assume that a first unplanned removal
occurred at time $T_1 < T_B$, i.e., this event was unobservable to us. Suppose we observe the
two unplanned removals at times T_2 and T_3, which occurred before the end of the
observation period T_E. Let TSN(T) denote the time since new of a product at time T.
(Note that TSN is measured in flying hours, i.e., hours of actual activity in the air, which
is different from calendar time. In Figure 2.1, the former is shown on the y-axis and the
latter is on the x-axis.) We do not observe the first unplanned removal and we do not
even know if it took place or not. In other words, we only know the values of T_B, T_2, T_3,

25

T_E, and the respective measures TSN(T_2), TSN(T_3), and TSN(T_E), but not the values of T_0 (the time at which the product was installed), T_1 (the time the first unplanned removal occurred), the corresponding flying hours TSN(T_1), and the initial age of the product at the beginning of the observation period TSN(T_B). We build our main proxy for MTBUR as follows (in section 2.7 we discuss other proxies for MTBUR and engine reliability):

$$\text{MTBUR} = \frac{\text{TSN}(T_E) - \text{TSN}(T_B)}{\text{Number of observed unplanned removals}}$$

Our proxy for MTBUR is thus defined by the inverse of the observed failure rate. In the example illustrated in Figure 2.1, it is equal to [TSN(T_E)- TSN(T_B)]/2. However, as we pointed out, the data do not include the value TSN(T_B). We compute an estimate for TSN(T_B), say TSN*(T_B), by assuming that there was a constant rate of usage for the product throughout the observation period. Specifically, we estimate this value as a linear projection of the line defined by the first observed removal and the age of the product measured at the end of the observation period, i.e., we estimate the slope of the line using the first removal and the end of the observation period as the two data points. We then project the line back to T_B in order to obtain an estimate of the initial age of the product defined as max{0, TSN*(T_B)}, which is measured in flying hours. We believe that this approximation provides a reasonable estimate for MTBUR. Note that if we omit subtracting the initial age of the engine, the proxy would overestimate the true MTBUR for all engines that had unplanned removals before the beginning of the observation period. The descriptive statistics for our MTBUR proxy are displayed in Table 2.2; recall that this variable is measured in flying hours.

	Overall sample		T&MC only		PBC only	
Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
MTBUR	6,456	3,278	6,016	3,844	6,575	3,106

Table 2.2: Descriptive statistics for the MTBUR.

Note that the mean MTBUR is slightly higher for PBC than for T&MC engines (by about 10%). An alternative approach to what is described above, i.e., computing the mean time between unplanned removals for each individual product and using it as a dependent variable in a regression model, would be to infer the mean time between unplanned removals (as an output of the analysis) by estimating the underlying distribution of the time between removals using techniques drawn from duration models (see, for example, Cameron and Trivedi 2005, Ch. 17). This approach – the analysis of which is discussed in Section 2.7 – has some econometric challenges on its own. In particular, it is difficult to account for the endogeneity of contract choices, a central feature of our model. Thus, in the main part of this study, we focus our discussion on the results based on the first approach which allows us to use a well-established two-stage econometric framework that explicitly deals with endogeneity of contract choices.

2.5. Econometric Model

Our goal is to build a model that captures the effect of contract type on product reliability, measured by MTBUR. While the previous literature (Kim 2011, Kim et al. 2011) present separate analyses of PBC and T&MC (but under the same sets of conditions), in practice both contracts are offered to the customers simultaneously and therefore assignment of customers to contracts is not random. Instead, a customer may

select a contract based on their private knowledge of engine usage that they possess and/or based on other firm characteristics. Thus, a major challenge associated with isolating the marginal effect of different contract types on product reliability is the inherent endogeneity associated with contract type choice by customers, which has been regarded as a key econometric issue in testing contract design hypotheses (Masten and Saussier 2002). A good illustration of the biases that can be generated by not accounting for self-selection in an application to firm entry and performance can be found in Shaver (1998). General econometric discussions on the importance of accounting for self-selection and related methods can be found elsewhere (e.g., Heckman 1979, Maddala 1983).

To account for endogeneity in contract selection, we utilize a two-stage treatment effects model (see Maddala 1983, p. 120). This control function approach allows us to estimate the effect of a binary treatment (PBC) on a numeric outcome (product reliability), given that the treatment assignment is not random but rather is determined by an endogenous decision process. The approach utilizes a two-stage structure that involves a first stage to explain contract choice (Equation 2.2) and a second stage to explain product reliability (Equation 2.1).

$$y_i = x_i\boldsymbol{\beta} + \delta z_i + \varepsilon_i \tag{2.1}$$

$$z_i = \mathbf{1}(w_i\gamma + v_i > 0) \tag{2.2}$$

The observed reliability of product i, denoted by y_i, is explained by the exogenous covariates x_i and the binary endogenous variable z_i (that in our case is equal to 1 for products covered by PBC contracts and 0 otherwise). As is standard in discrete choice

28

models with latent variable representation, e.g., probit, the binary variable for contract choice (z_i) is modeled as an indicator function, dependent on a set of exogenous covariates w_i, which drive the choice process. The error terms (ε_i, v_i) of the outcome and choice equations, respectively, account for unobservable characteristics which are allowed to be correlated, and are modeled as a bivariate normal random variable with distribution $N_2(0,0,\sigma^2,1,\rho)$; where the variance of v_i is normalized to one for identification purposes. If the correlation between both error terms is equal to zero then the outcome and choice equations can be estimated independently (Equation (2.1) could be estimated by OLS), i.e., the endogeneity of contract type is not relevant for the problem; otherwise, OLS estimation of Eq. (2.1) will produce a biased estimate for δ. For additional information on two-stage models based on control functions, the reader is referred to Maddala (1983, pp. 117-125).

In order to properly capture the main effect of our interest, i.e., significance and the magnitude of the coefficient δ in Equation (2.1), we need to specify the observable covariates influencing product reliability (x_i) and contract choice (w_i). There are characteristics of both the product and the customer (the user) that can play roles. Of those, we need to identify the most salient characteristics in order to avoid collinearity issues in our sample of 305 observations. Since we are not aware of other papers that analyze after-sales repair and maintenance contracts in the aerospace industry at the level required for our model, there is no precedent for many of the variables that we use, but we take clues from the reliability and contract theories, and from our in-depth knowledge of the industry.

An obvious factor that can influence the observed MTBUR is the initial condition of the product. Reliability theory (see Rausand and Høyland 2004) argues that very young and very old engines are more likely to have low reliability: younger engine units typically require adjustments at the beginning of their life, while older engines may fail more often due to part wear and tear caused by usage over time. To check these conjectures, we examined the distribution of the MTBUR for different ranges of initial product age and observed that MTBUR is lower for both new and old products, and is higher for medium age products, which is in line with the reasoning proposed above. In order to account for such nonlinearities, we include both linear and quadratic terms for the initial age of the product in our model specification. A polynomial function has been also used by Hubbard (1998), among others, to capture the effect of initial age in a related setting.

Another product characteristic that is related to engine reliability is the average flight time of an engine. For example, more take-off/landing cycles per flight hour may decrease the reliability of the engine since most of the wear and tear happens during these time periods. In our dataset there are records of the time since new (TSN) and the cycles since new (CSN) for each engine, measured at the end of the observation period. The average flight time for each engine is then the ratio TSN/CSN. We include this control variable in the outcome equation.

While the initial age of the engine and the average flight time capture a relevant part of both the initial conditions and usage patterns of the product, engine model characteristics can also affect reliability and should be controlled for. As mentioned in the

data description, there are 2 engine models (5 versions in total) and 3 different aircraft models in our sample. Not every engine model can be installed in every aircraft model. In our data, there are three possible variables that help account for engine model effects: engine model dummies, aircraft model dummies, and a continuous variable indicating the number of years since the engine model was introduced into the market. They are strongly correlated and therefore cannot be included simultaneously in the model. Out of the three candidates, we choose to include the last variable in the model specification, as it captures the essence of the differences across products more parsimoniously in a single coefficient. This variable is a reasonable way to approach the effect of engine characteristics on product reliability, e.g., less is known about engine models that were introduced more recently to the market, which could influence reliability and how likely are these engines to be covered by PBC. In Section 2.7 we discuss models which use engine model dummies and aircraft model dummies, instead, but our results remain robust to these variations.

As we discussed in Section 2.3, a customer can also affect engine reliability, and thus customer characteristics need to be included in the specification of the outcome equation. For example, geographic location of the owner can be a proxy for closeness and availability of repair shops and for other local market conditions such as weather. We include dummy variables for the geographical region of the owner, categorized as U.S. vs. non-U.S. customers. We collect this information based on the customer ID, which we know from our database. Naturally, we would like greater granularity to capture local effects at the level of countries or even regions within a country. Recall, however, that we

only have 48 customers in our sample, and so defining the geographical variable too narrowly would generate an identification problem due to low frequency data.

Further, we include the variables fleetsize (number of engine units that a customer has registered with Rolls-Royce) and fleetmix (number of different engine models in the portfolio of a customer), which serve as proxies for the customer characteristics that may be correlated with the engine reliability in different ways. For example, one may expect fleetsize to be positively correlated with reliability because a customer with a larger fleet is in a better position to request high quality services provided by Rolls-Royce. On the other hand, a negative correlation will be observed if fleetsize reflects usage patterns, in which case a customer with a larger fleet may use the engines with less care expecting a slack in engine capacity during low utilization periods. Similarly, both positive and negative correlations between fleetmix and reliability are possible. A positive correlation is likely if service priority goes to a customer who demands more attention because of the support needs of his diverse engine portfolio. Alternatively, one can hypothesize a negative correlation on the ground that a less diversified engine portfolio simplifies the engine operation and therefore leads to higher reliability. Whichever may be the outcome, it is clear that the customer characteristics captured in these variables can affect reliability. Hence, it is important to use this information in the model specification to control for customer effects. Note that an alternative way to capture the effect of customer characteristics on reliability is using a random (or fixed) effects model; this approach is discussed in Section 2.7.

Finally, another factor that could influence the time between unplanned removals is the occurrence of planned maintenance events (planned removals). Planned maintenance events have to follow a schedule that is prescribed by the Federal Aviation Administration, i.e., after every so many flying hours the engine must come in for a planned maintenance check. Of course, there is some discretion in following these rules: the airplane owners will often try to minimize schedule disruptions by arranging planned maintenance at the time when the aircraft is close to the repair facility and when its utilization is relatively low. All else being equal, we expect that the engines that have had more frequent planned maintenance to have lower need for unplanned removals, resulting in a higher MTBUR. Thus, we include in the outcome equation a variable indicating the number of observed planned maintenance events for a given engine.

Regarding the choice equation, we need to include covariates that influence the type of contract selected by a customer. Insights from the contract theory (e.g., Bolton and Dewatripont 2005) suggest that customers who intend to overuse engines and take poor care of them will be incentivized to opt for PBC contracts to begin with (i.e., adverse selection). That is, allocation and sharing of the risk induced by different contract types is one possible reason for self-selection by customers. Clearly, the two contract types of interest here have different implications for risk allocation: under PBC, the risk is shifted entirely towards the supplier, while under T&MC the risk is shifted towards the customer. Thus, PBC can be thought of as an insurance policy that creates predictable cash flows for the customer at a cost. Individual customers would then analyze this trade-

off using their internal knowledge about their risk-aversion, and we expect that a customer who is more (less) tolerant to risk will opt for T&MC (PBC).

While it is notoriously difficult to find good proxies for risk-aversion, one of the commonly used proxies is the size of the company which, in the context of our setting, corresponds to fleetsize. According to supplier managers we interviewed, fleetsize is probably the most relevant variable to explain contract selection, as it is observed anecdotally that customers with a larger fleet are more likely to choose PBC for after-sales product support. This is related to the notion of risk we discussed above: the expected volume of cash flows needed for repair and maintenance services increases with the size of the customer fleet, which may cause larger firms to be more likely to sign on for PBC. This conjecture is in line with the data, which show that the median fleet size of T&MC and PBC customers are 2 and 12, respectively. In addition, we include the variable fleetmix as part of the choice equation specification in order to reflect the connection between the complexity of a customer's engine portfolio and his contract choice. For example, a customer may have internal expertise to service some engine types, but ownership of multiple engine types would require a complex mix of internal capabilities and therefore could lead the customer to opt for a comprehensive support arrangement through PBC. Alternatively, a more diversified fleet may be seen as a measure of risk diversification, which may lead fleetmix to be negatively correlated with the choice of PBC. Hence, the mix of the customer product portfolio needs to be controlled for.

Similarly, the location of the owner partially captures local characteristics such as prices, availability of repair shops, competition, and marketing efforts by Rolls-Royce that may influence the decision to choose PBC. Therefore, we have included geographical dummies in w_i, as we did for the outcome equation. Finally, another factor that can influence contract choice for a given engine is how complex it is to deal with maintenance for a given engine model. As discussed earlier, the number of years since an engine model was introduced to the market seems to summarize product characteristics that can be related to the type of customers that prefer either type of contract, for example, engine models for which there is little knowledge of maintenance procedures (i.e., newer models) could be perceived as riskier by customers, and therefore they could be more likely to be covered by PBC as a way to mitigate risk. We thus include this variable in the choice equation.

Note that our modeling approach allows for correlation between the unobservable terms of the outcome and contract choice equations. This feature is useful since we do not observe all variables related to the risk profile of the customer, which can influence both product reliability and contract choice. In particular, in our case we expect this correlation to be negative due to adverse selection and customer-side moral hazard, i.e., customers that are more likely to use engines more intensely (increasing failure risk) have more propensity to sign for PBC, which is hypothesized to increase reliability. Similar arguments have been discussed, for instance, in the case of extended warranties for new car buyers (Padmanabhan 1995, Hollis 1999), where it is argued that heavy users have stronger incentives than light users to sign on for extended warranties, since their

products are more likely to experience failures. We also note that the control function approach we implement does not involve exclusion restrictions, in contrast to estimation based on instrumental variables methods. That is, all variables in w_i can be included in x_i. This is due to the assumption of the nonlinear selectivity term introduced as control in the outcome equation (see Eq. 2.3 below), which allows for consistent estimation of the δ coefficient in our problem under the maintained exogeneity assumption for covariates other than z_i (Maddala 1983 p.121). See Petrin and Train (2010) and Cameron and Trivedi (2005), which contain broader discussions on control function approaches for endogeneity correction.

2.6. Results and Analysis

We now turn to the estimation and results obtained for the two-stage model defined by Equations (2.1) and (2.2). The model can be estimated using the usual two-step procedure, defined e.g., in Maddala (1983, pp. 120-122). First, a probit model is estimated for the choice equation, where the probability that observation i receives "the treatment" (engine unit i covered by PBC) is given by $\Pr(z_i = 1 | w_i) = \Phi(w_i \gamma)$, where Φ is the cdf of the standard normal distribution. Let ϕ be the pdf of the standard normal distribution. From the results of this first stage, a selectivity term is derived as follows:

$$\text{Selectivity term}\_i = \begin{cases} \dfrac{\phi(w_i \gamma)}{\Phi(w_i \gamma)} & \text{if } z_i = 1, \\ \dfrac{-\phi(w_i \gamma)}{1 - \Phi(w_i \gamma)} & \text{if } z_i = 0. \end{cases} \tag{2.3}$$

The selectivity term calculated from the first stage is used as a regressor in the outcome equation, which allows for consistent estimation of the effect of the endogenous

treatment, PBC, on product reliability, by accounting for the endogeneity in contract choice. We estimate the model in STATA using the aforementioned two-step procedure. For a reference, we also report results obtained from estimating the outcome equation using OLS, i.e., ignoring the endogeneity problem of contract choice. Table 2.3 displays the results obtained from the two-stage model and from the OLS model. We report clustered standard errors at the customer level in all cases, which are robust to correlation among the unobservable terms of observations from a given customer. In our problem, clustered standard errors at the customer level are of comparable magnitude to the ones obtained with the usual two-stage correction. In particular, our conclusions on how PBC affects reliability remain unaffected by the use of the common two-stage standard errors correction.

Our estimate for the effect of PBC on MTBUR is positive and indicates that, on average and all else being equal, PBC significantly increases engine reliability (p-value=0.003). Confidence intervals for the effect of PBC on reliability at the 90% and 95% confidence levels are given by [1154,3895] and [882, 4168], respectively. This provides support for the hypothesis of a positive and significant effect of PBC on product reliability. The model also indicates that the number of planned removals has a positive effect on product reliability: as expected, the more frequent the planned maintenance events, the less frequent the need for unplanned maintenance. Other product attributes with significant effects include the initial age of the engine and the number of years since the engine model was introduced. The effects found indicate that both linear and quadratic terms of the initial engine age are significant, in line with our arguments, and

37

that engine models that were introduced earlier to the market are associated with a larger

MTBUR, again in line with our reasoning. Similarly, the model indicates that customer

characteristics also affect reliability: once controlled by contract choice, product

characteristics, and planned maintenance, engines pertaining to U.S. customers with

smaller fleet and greater fleet diversity, have greater MTBUR, relative to their

counterparts.

TWO-STAGE MODEL				OLS MODEL
Choice equation		Outcome equation		
Variable	Coeff.	Variable	Coeff.	Coeff.
Fleetsize	0.024***	PBC	2525***	30
	(0.008)		(816.7)	(364.1)
Fleetmix	-0.436	n_planned_rem	1084***	1108***
	(0.383)		(257.9)	(250.2)
region_usa	-1.43*	ini_age	0.582**	0.629***
	(0.866)		(0.224)	(0.213)
nyears_em_intro	-0.206	ini_age_square	-0.00007***	-0.00007***
	(0.178)		(0.00002)	(0.00002)
Constant	1.57	eng_avgflighttime	861	555
	(1.088)		(965.3)	(973.1)
		region_usa	2562***	1876***
			(360.9)	(327.5)
		fleetsize	-4.32***	-0.967
			(1.557)	(0.965)
		fleetmix	298*	130
			(157.4)	(136.2)
		nyears_em_intro	1461***	1346***
			(166.7)	(177.4)
		selectivity_term	-1674***	
			(463.6)	
		Constant	-3969**	-1064
			(1732)	(1601)
χ^2(df)=9.75(4)		R^2=0.54, F=162.65		R^2=0.53, F=47.96

Table 2.3: Two-stage and OLS models results (N=305). Clustered standard errors (customer-level) are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. df=degrees of freedom.

As discussed, an important feature of our model is that it controls for the endogenous nature of contract choices. Interestingly, as we show in Table 2.3, the results obtained by ignoring this endogeneity through using a simple OLS regression (i.e., assuming that the correlation between the error terms in the choice and outcome equations is zero) would suggest that there is no effect at all of PBC on product reliability. To the contrary, once the endogeneity is taken into account, our two-stage model predicts a positive and significant effect of PBC on reliability. In fact, in our two-stage model, the estimate for the correlation between the unobservable terms in the choice and outcome equations is negative (-0.69), and a likelihood ratio test rejects the null hypothesis of the correlation term being zero at the 5% significance level. Further, the selection correction term is significant at the 1% significance level. This evidence confirms the important role of accounting by the selection process in order to estimate the effect of PBC on product reliability in our application, illustrating that the model cannot be estimated by OLS due to the self-selection in the contract choice decision.

With respect to the choice equation, the results indicate that fleetsize and the geographical region have explanatory power for the contract type of an engine. In particular, our results show that engines owned by customers with greater fleet sizes are more likely to be covered by PBC than by T&MC, in line with our hypothesis, data, and managerial expectations. Our two-stage model estimates thus indicate that the factors influencing the likelihood of an engine being covered by PBC positively are in turn negatively correlated with engine reliability, which, in addition to the negative estimate of the correlation between the unobservables of the outcome and choice equations, is

consistent with the argument of adverse selection. The first stage probit predictions that drive our modeling of the selection process are fairly reasonable. Comparing the proportions of engines covered by PBC vs. T&MC, the actual data in our sample reflects proportions of 78.7%/21.3% respectively (240/65 engine units), almost identical to our model's predictions of 79.3%/20.7% respectively (242/63 engine units). The model replicates the observed proportions remarkably well, and formal tests of equality of predicted and actual proportions cannot be rejected at any relevant significance level.

In summary, the results in this section provide support for our main hypothesis of a positive effect of PBC on reliability, once the endogeneity of contract choice is accounted for. One intriguing question that we have not addressed is: what is the mechanism by which PBC improves reliability? The results of our model are consistent with two possible mechanisms. First, our estimation results indicate that the frequency of scheduled maintenance is positively correlated with reliability. More frequent planned maintenance could be one mechanism by which PBC improves reliability. Summary statistics in Table 2.4 indicate that the frequency of scheduled removals is indeed higher under PBC.

	Overall sample		T&M only		PBC only	
Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
MTBPR	6,675	3,114	8,034	3,709	6,332	2,849
n_planed_rem	0.518	0.644	0.354	0.543	0.563	0.663

Table 2.4: Descriptive statistics for MTBPR[3] and number of planned removals.

---

[3] The MTBPR (mean time between planned removals) is computed similarly to the MTBUR: the only change is the use of the number of observed *planned* removals in the denominator. The relevant sample for this calculation consists of the engines that had at least one planned removal. The sample for the statistics for the variable n_planed_rem consists of the 305 engines we analyze in our models.

Further, we note that without including the planned removals as a control in our empirical model (not reported in Table 2.3), we obtain a bigger and still highly significant coefficient on PBC, 2968, which suggests that part of the PBC effect is related to higher frequency of planned maintenance. Based on both observations, i.e., the descriptive statistics noted above and the increase in the magnitude of the PBC coefficient when planned removals are excluded from the model specification, our results suggest that more frequent scheduled maintenance under PBC is one mechanism by which reliability improvements may be achieved. However, as the estimation of our main model illustrates (Table 2.3), there is an effect of PBC on reliability even when we control for the number of scheduled maintenance events, which suggests that more frequent planned removals is not the only mechanism by which PBC influences reliability. Indeed, this observation is consistent with an alternative mechanism by which reliability is improved under PBC, namely, reliability improvement resulting from the supplier's provision of better care in each maintenance event. This could be achieved by various actions such as preemptive replacements of parts, more thorough identification of defects, and possible product re-design.

2.7. Robustness

In this section we examine the robustness of our main findings with respect to variations in a number of relevant model constructs, such as model specification, reliability proxies, modeling techniques, and sampling issues. Table 2.6 reports the results of some of these

tests. The remaining tests, mentioned in this section but not reported in the table due to space constraints, are available upon request.

Model specification: The first issue we address is related to the variables used to control for product type. In our main model, we included the number of years since the introduction of an engine model to the market as the variable capturing product effects (nyears_em_intro), as it captures the product effects more parsimoniously while giving more accurate first stage predictions. As we mentioned in earlier sections, there are different ways to control for product characteristics, e.g., dummy variables to characterize aircraft model, engine model, or engine type. They cannot be included simultaneously in the model, as different proxies are almost perfectly collinear. As a robustness check, we estimated our model using alternate aircraft model dummies, dummies for the 5 engine model versions and dummies for the 2 main engine types. Our results remain qualitatively the same with these variations.

A second issue is the use of the initial age of the engine in the outcome regression. As we noted, this variable is not available to us, and we used extrapolation to obtain an estimate of it. A concern that can be raised is that we are using this variable as an input to calculate our dependent variable, as well as to calculate an independent variable. While we do not believe that this is critical in our case – our proxies for product reliability would be less meaningful if we did not take into account the initial age of the engine, to start with – we try a different proxy for capturing the initial conditions of the engine. We obtain the year of production of each aircraft from the Federal Aviation Administration

(FAA) and related websites, matching the aircraft serial numbers in our dataset.[4] Using

this information, we construct a variable reflecting the number of years since the aircraft

was manufactured with respect to 2002 (the beginning of the observation period). While

we believe that the initial age of the engine (measured in flying hours) is a more precise

proxy for the initial conditions of the engine, the age of the aircraft can capture some

relevant initial conditions of the engine in a similar fashion. We use age of the aircraft

instead of initial age in Equation (2.2) to estimate our model. The results obtained,

displayed in column 1 of Table 2.6, show that our results and conclusions are robust to

this variation.

Another issue related to the model specification is the potential role of uncontrolled

unobserved heterogeneity in driving our results. Given the clustered nature of our data

(i.e., engines pertaining to different customers), it could be hypothesized that unobserved

customer effects could have an influence in our estimates. As a way to explore the

robustness of our results with respect to this issue, we included an unobserved

heterogeneity term at the customer level in the contract choice and outcome equations,

which is modeled as a draw from a normal distribution with mean zero and unknown

variance, i.e., a two-stage random effects model. We find that our main results remain

qualitatively the same under this variation (see column 2 in Table 2.6), and that the

---

[4] We obtained data on this variable from aeronautic institutions from different countries. Websites consulted include: USA (FAA database, http://www.faa.gov), France (Direction Générale de l'Aviation Civile, http://www.immat.aviation-civile.gouv.fr/immat/servlet/aeronef_liste.html#), UK (Civil Aviation Authority http://www.caa.co.uk/application.aspx?catid=60&pagetype=65&appid=1), Sweden (Transport Styrelsen, http://www.luftfartsstyrelsen.se/templates/LS_LuftFartyg_Sok____39453.aspx), . For other countries like Brazil, Italy, Belgium, South Africa, Canada, Thailand, etc., we used different industry websites, e.g. http://www.airport-data.com, www.airframes.org, http://www.aerotransport.org/php/querybuilder.php?tab=regn, http://www.planespotters.net/Production_List/.

variance contributed by the unobserved heterogeneity term at the customer level explains only 2.7% of the total variance attributed to unobservable factors. In other words, most of the variance is already captured in the idiosyncratic error term at the engine unit level, which was already considered in Equation (2.2). Furthermore, we also examine whether the PBC effect holds at the customer level. For that purpose, we estimate a two-stage between effects model, i.e., a regression on group means where each costumer defines a different group, which is an extension of the two-stage random effects model described above. We find that the PBC effect remains significant and is of comparable magnitude to the one reported in our main model, which suggests that the PBC effect is also present at the customer level. Altogether, these robustness checks suggest that unobserved heterogeneity is not a major concern for our conclusions.

Overall, for the robustness tests considered here, the PBC effect is significant at the 95% confidence level in all cases, with estimates in the range of [1655,2521] flying hours, which is in line with our main model. This represents a reliability improvement under PBC in the 25-40% range relative to T&MC.

Measurement of product reliability: The first issue here is the proxy we use to characterize product reliability. Consider again the example Figure 2.1. Given the nature of our data, there are several ways in which a proxy for product reliability could be defined, depending on how the right tail of the distribution is accounted in the calculations. In addition to the MTBUR variable defined in section 2.4 and used in our models – which considers the right tail of the distribution in its entirety and therefore is a metric that is less sensitive to our estimation of the initial age of the engine – we explore

different definitions of product reliability that deals with the right tail of the distribution in different ways. Based on the example in Figure 2.1, we define the following alternative proxies for MTBUR:

- MTBUR_proxy2 = $\dfrac{\text{TSN(Latest observed unplanned removal)} - \text{TSN}(T_B)}{\text{Number of observed unplanned removals}}$

- MTBUR_proxy3 = average$\{\text{TSN}(T_2) - \text{TSN}(T_B),\ \text{TSN}(T_3) - \text{TSN}(T_2),$ $\text{TSN}(T_E) - \text{TSN}(T_3)\}$,

- MTBUR_proxy4 = max$\{$MTBUR_proxy2, MTBUR_proxy3$\}$.

Proxy 2 ignores the right tail of the distribution where unplanned removals did not occur and it is therefore also more sensitive to the estimated initial age of the engine, while proxies 3 and 4 incorporate that information in different ways than our main proxy. The descriptive statistics for these alternative MTBUR definitions are displayed in Table 2.5; all four proxies are measured in flying hours.

Variable	Overall sample		T&MC only		PBC only	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
MTBUR_proxy2	3,328	2,458	2,872	2,632	3,451	2,400
MTBUR_proxy3	3,507	1,837	3,147	2,029	3,605	1,773
MTBUR_proxy4	4,079	2,227	3,609	2,536	4,206	2,123

Table 2.5: Mean and standard deviation of alternative reliability proxies.

While the values in Table 2.5 differ across the proxies (reflecting the distinct ways in which they account for the right tail of the distribution), in all cases we see that MTBUR of PBC engines is slightly higher than MTBUR of T&MC engines. We run our two-stage models using these MTBUR proxies as dependent variables, and find that the results that are consistent with the ones from our main MTBUR proxy. In particular, the PBC effect

remains significant at 95% and 99% confidence levels in all cases (see columns 3, 4 and

5 in Table 2.6).

VARIABLES	(1)	(2)	(3)	(4)	(5)
PBC	2521.2***	1817.6***	1476.7**	1257.7***	1391.6***
	(865.3)	(484.9)	(605.0)	(429.2)	(514.6)
n_planned_rem	1002.9***	1070.6***	484	700***	885***
	(256.7)	(260.7)	(295.8)	(150.4)	(216.8)
ini_age		0.567**	0.222	0.224	0.324*
		(0.225)	(0.154)	(0.154)	(0.163)
ini_age_square		-0.00007***	-0.00001	-0.00002	-0.00003
		(0.00002)	(0.00002)	(0.00002)	(0.00002)
aircraft_age	575**				
	(234.1)				
aircraft_age_square	-26**				
	(11.91)				
eng_avgflighttime	1137.8	636.7	1367**	716	1111
	(1,163)	(932.6)	(679.5)	(583.9)	(670.2)
region_usa	2781.5***	2286.3***	958***	1156.7***	1394.3***
	(331.7)	(386.4)	(355.9)	(182.7)	(221.3)
fleetsize	-4.7**	-2.5**	0.5	-1.1	-0.4
	(1.896)	(1.242)	(1.382)	(1.046)	(1.105)
fleetmix	168	212	-22	63	41
	(191.2)	(158.9)	(148.0)	(98.15)	(116.4)
nyears_em_intro	1565.3***	1230.8***	681***	775***	863***
	(163.1)	(172.2)	(113.1)	(100.2)	(112.7)
selectivity_term	-1601.7***	-1054.9***	-1016**	-839***	-953***

Constant	(494.7)	(204.4)	(411.7)	(241.9)	(337.1)
	-6154.1***	-2483.1	-3025.4*	-2219.7**	-2953.1**
	(1,824)	(1,543.4)	(1,547)	(1,077)	(1,351)
	R^2=0.54, F=158.75	R^2=0.54, χ^2(df)=1128(10)	R^2=0.25, F=42.98	R^2=0.54, F= 145.88	R^2=0.49, F=166.54

Table 2.6: Selected robustness checks. All columns display results for the outcome equation of the two-stage model considered in each case. (1): Dependent variable=MTBUR; (2): Dependent variable=MTBUR, the model includes a random effect at customer level; (3): Dependent variable=MTBUR_proxy2; (4): Dependent variable=MTBUR_proxy3; (5): Dependent variable=MTBUR_proxy4. Clustered standard errors (customer-level) are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

A second consideration related to the measurement of product reliability is the cross-sectional modeling approach we have used. Thus far, our analysis has relied on constructing proxies for MTBUR of each engine unit and performing two-stage estimations. As mentioned earlier, a duration model offers an alternative way to analyze our data. In fact, the sample in our dataset represents multiple spells data since some products have more than one unplanned removal. The main advantage of this modeling approach is that it inherently accounts for some data censoring issues, which are common in duration data such as ours. Unfortunately, dealing with the endogeneity issue in the context of duration models is difficult, in the sense there are no sufficiently established techniques that we could employ in our problem to convincingly account for the endogeneity of contract choices (research in this area is ongoing, see Bijwaard 2007 for a recent contribution to this research stream). However, the following informal approach has been used by some researchers: (1) run the probit model (Equation 2.2) to predict contract choices, (2) calculate the selectivity term from this analysis (Equation 2.3), and (3) perform duration analysis using the selectivity term as one of the regressors. A similar

approach in the context of sample selection for duration models was used by Rao et al. (2001), based on the generalization of the Heckman selection model proposed in Lee (1983). Although consistency of this approach is not, to our knowledge, fully established, we use this procedure in this subsection as a robustness check, which can be taken as complementary evidence with the caveat on endogeneity as described above. In order to proceed with this approach, we analyze the data at the removal level (instead of at the engine unit level), and we must examine the influence of contract type on the respective unplanned removal rates. As is standard in duration analysis, we conduct experiments using both semi-parametric (Cox) and parametric (exponential, log-logistic, and log-normal) transition rate models. The explanatory variables are the same as the ones used in the two-stage regression models; the only modifications in explanatory variables we incorporate – in order to take advantage of the multiple spell nature of the data – are to replace the initial age of the product with the age of the product at the beginning of each spell in the outcome equation and to use the number of observed planned removals in each spell, instead of in the full observation period. For example, for the spell associated with the occurrence of the third unplanned removal of a product unit, instead of the initial age (age of the product at the beginning of the sample period) we include the age of the product unit at the moment of the second removal (for both the linear and quadratic terms). Note that this also makes these models much less sensitive to our estimation of the variable ini_age, which may serve as additional evidence of the robustness of our results to that issue. We estimate the models using clustered standard errors at the customer level. We obtain largely the same findings obtained with the two-stage models,

48

in particular, for the effects of PBC on product reliability, regardless of the modeling technique employed (semi-parametric and parametric effects defined above).

Sampling issues: The first issue is the low frequency of unplanned removals in our data. As we indicated earlier, for a significant part of our sample we observe only one unplanned removal. As a robustness check, we run our model for the sample of engines which had at least two unplanned removals (ignoring engines with only one unplanned removal). We obtain coefficients for the PBC effect similar to the ones obtained for the whole sample. Some significance is lost, however; under clustered standard errors at the customer level we obtain a p-value for the PBC effect of 0.124, and a p-value of 0.058 under non-clustered standard errors. It is reasonable to expect some reduction in the significance levels in this case, as the number of engine units with at least two removals is very small (56 engine units). We repeat the exercise for duration models, which have the advantage of accounting for each spell separately (thus alleviating the problem of dealing with such a small sample size to some extent), and again obtain similar PBC effects which are now significant at the 95% confidence level. While we do not attempt to make inferences based on this small sample of 56 engine units, these exercises alleviate potential concerns regarding the role of low frequency of unplanned removals in our analysis.

Perhaps a more fundamental sampling concern is that our definition of MTBUR for the two-stage model is conditional on the occurrence of at least one unplanned removal. This means that engines without observed unplanned removals are not considered in the estimation of our model. Naturally, this is driven by the nature of the dependent variable,

mean time between unplanned removals, which is undefined if unplanned removals did not occur at all. As we pointed out, the concern of a potential bias in our sample due to this issue is alleviated by the fact that the observed proportions of PBC and T&MC engines in the sample of 305 engine units for which we observe unplanned removals is the same as the proportion observed in the full sample of 763 engines. Duration models, however, allow for the possibility of estimating the model including those engine units for which there are no observed events (unplanned removals), by incorporating the associated spells into the likelihood function. We estimate duration models for the full sample of 763 engines, using the same procedure and modeling assumptions described earlier, and obtain largely the same findings as in the case of the 305 engine unit sample. This alleviates concerns regarding the potential role of the conditional nature of the sample used in our analysis.

2.8. Conclusions

Our analysis suggests that product reliability improvement is achieved under PBC. Our estimates indicate that, in comparison to traditional T&MC, there is a positive and statistically significant effect of PBC on the mean time between unplanned removals (MTBUR) of a product and that its magnitude is in the 25-40% range. We also show that endogeneity of contract choice, which has not been either modeled or discussed previously in this context, is clearly an issue here. Indeed, without explicitly accounting for this endogeneity, the significance of the PBC effect disappears. These findings are supported by numerous robustness checks under a number of alternative model

specifications and modeling approaches. The results obtained from our analysis provide a first step towards understanding the overall impact of PBC on product reliability.

Although our analysis shows firm evidence that supports these conclusions, it is not free of limitations. One of them is our model specification that allows us to explicitly deal with the endogeneity of contract choices at the expense of treating the rest of the covariates as exogenous. As a side effect, this approach precludes us from exploring a potential correlation between the frequency of planned removals and the unobservables in the outcome equation. This, however, is not unreasonable since planned removals are subject to strict regulation in the aerospace industry (although there may be some discretion by aircraft owners in this regard). In addition, there are several issues stemming from the nature of the data. First, our ability to accurately measure engine reliability is limited by the low frequency of failures in our sample. Failures of aircraft engines are rare events, even though we observe the system for the relatively long period of 5 years. This problem creates some imprecision in the definition of the dependent variable, MTBUR. We deal with this issue by conducting analysis based on several proxies for MTBUR and using alternative modeling approaches (two-stage models, duration model analysis), and find that the results are consistent overall. Second, while our dataset is rich in terms of characterizing the removal incidents for a given product, we have only limited data to characterize a customer (such as the customer risk profile) that may impact the results. To alleviate this concern, we included in our analysis such variables as fleet size, fleet mix, and the region of the owner, as a way to account for some relevant customer characteristics. Third, our dataset does not provide the detailed

characteristics of the contracts; partly because of this, we only distinguish between T&MC and PBC. This limitation prevents us from exploring, for example, the influence of price and contract length on the customer's contract choice. We also do not observe data before and after the adoption of PBC, which would have made it possible to study how incentives evolve over time. However, we tested whether the influence of unobserved heterogeneity could influence the results, and found that the PBC effect remains remarkably robust, which suggests that the influence of these unobserved factors should not affect our results significantly.

Lastly, it is important to note that the goal of our research is driven in large part by data availability. For instance, our analysis does not address cost implications of PBC because cost data were not made available to us. Undoubtedly, an access to a richer set of data including financial and managerial information, other performance metrics such as product availability, and contract term details would enable a more complete analysis with a larger scope than the current study offers, leading to a deeper understanding of the benefits and downsides of PBC. Despite the limitation, we believe that our study provides valuable insights to practitioners who are continuously striving to achieve the highest level of product reliability.

In this research we present one of the few studies that empirically estimate the impact of contracting on supply chain outcomes. While we cannot claim that the conclusions obtained in this study are applicable to all supply chain settings, our findings are of interest, not only to the aircraft repair and maintenance industry but also to all industries in which the quality of manufactured products is an important driver of firm performance.

We also believe that this study is practically relevant and timely because it is the first empirical investigation that rigorously tests the hypothesis that reliability improvement is achieved under PBC, an issue of considerable interest among practitioners who are currently weighing the costs and benefits of adopting PBC-based relationships for after-sales product support.

Chapter 3

Service Competition and Product Quality in the U.S. Automobile Industry

3.1. Introduction

A fundamental trend in manufacturing industries is the movement from a "pure manufacturing" paradigm to a business model in which a central role is assigned to the service component of products based on the value they provide to consumers (Cohen et al. 2006). The movement towards a service-based economy has coincided with this change and has encouraged many manufacturing firms to put more emphasis on the delivery of services associated with their product offerings (Shankar et al. 2009). It has been reported that the sales of after-sales services and spare parts represent 8% of the annual gross domestic product in the U.S. (Cohen et al. 2006). In the automotive sector, in particular, recent industry reports indicate that for manufacturers in this industry,

service and parts operations are on average 54% more profitable than the main business of producing and selling vehicles, and account for an average of 36% of revenues (Koudal 2008). In the technology sector, companies like IBM that traditionally sold manufactured goods, today derive more than 50% of their revenue from services (Suarez et al. 2008). In short, services have become an important part of an OEM's competitive strategy in traditional manufacturing industries. While existing models of product differentiation in manufacturing industries have provided some insights in explaining the consequences of pricing and other product characteristics on demand, they have mostly ignored the impact of services. This study takes a step in addressing this issue, by formulating an empirical model to analyze the role of services as part of a firm's competitive strategy in the U.S. automobile industry, and the joint effect that service attributes and product quality have on consumer demand.

The automobile industry has served as a preferred setting for empirical studies on product differentiation (e.g. Berry et al. 1995 and 2004, Sudhir 2001, Train and Winston 2007, among many others), and constitutes a natural choice for our research. Indeed, Standard & Poor's (2011) reports that in this industry "product quality and design are becoming less of an issue in differentiating foreign and domestic manufacturers," as a result of the actions taken by Detroit automakers to improve their designs and to streamline their manufacturing processes in recent decades. Services, on the other hand, represent an important differentiating factor for automakers, especially given the high level of competition and low concentration in the U.S. automobile market (Koudal 2008). The fact that auto OEM's have been adjusting their service strategies in recent years, as

we describe shortly, provides some support for this notion.

We focus on services during the in-warranty period. In particular, we measure the service dimension of a brand by the length of its warranty, along with a metric of the after-sales service quality delivered during the in-warranty period. The warranty period covered by OEM's has steadily increased over time, from about three months in the 1930's, to as much as ten years in recent years (Murthy and Blischke 2006). The length of the warranty defines the period in which repair services are provided by the OEM as part of the value that consumers obtain with the purchase of the car, and therefore is a managerial decision that partially reflects the service intensity provided by OEM's. Firms have been active on adjusting the length of their warranties in the last decade: Ford, Chevrolet, Acura, Mazda, Mitsubishi, Audi and Kia are just some examples of brands that have increased the length of their warranties in that period. For example, Ford increased their powertrain warranty from 3 years/36,000 miles to 5 years/60,000 miles in 2007. In the words of a spokeperson from Ford when asked about the reasons behind the warranty length increase: "We think that some people weren't considering Ford products because we didn't have an extended powertrain warranty versus some of our competition. We hope that it will increase our competitiveness... We think that customers do want it, and do care about it" (Warranty Week 2006). On the other hand, Chrysler and Volkswagen both decreased their warranty length at least once during the same period. Indeed, company sources have suggested that the increase in warranty length by Chrysler for the 2008 model year "wasn't as valuable to consumers as we might have hoped" (Automotive News 2009), and as a result the company cut it back in 2009.

Firms face an important trade-off when defining their warranty period: while longer warranties may potentially increase product demand, they also generate significant costs. For U.S.-based automakers, these costs have typically been in the range of $10 billion per year, which represents roughly 2-4% of their yearly revenue (Warranty Week 2011). These data, along with the aforementioned managerial actions and the attention of the trade press to them, reflect the importance of improving our understanding of the role of warranties and service attributes as drivers of demand, the conditions under which they influence demand, and the magnitude of these effects.

In this study we formulate and estimate a structural model to measure the impact of service attributes on consumer demand in the U.S. automobile industry. Combining data from multiple sources, we propose an empirical model using market-level data for new light cars sold in the U.S. between 2001 and 2007, a period in which firms actively adjusted their service and warranty strategies. Our analysis is based on a random coefficients logit demand model that allows for customer heterogeneity in tastes for different car attributes and, unlike most existing models, incorporates the two aforementioned variables to characterize firm service strategies, i.e. warranty length and the quality of after-sales service. Our results provide new evidence to explain the influence of warranty length and service quality on the demand for a given model, relative to the influence of other characteristics such as price and product quality. Most existing empirical models of competition in this and other industries deal with the endogeneity of prices while assuming that all other characteristics in the demand specification are exogenous. Our model is different in that we not only endogenize

pricing but also the warranty length decision. We do so by generating instruments based on the factors driving firm decision-making for the warranty length. Our findings indicate that, when the endogeneity of warranties is ignored, service attributes do not have a significant impact on demand. Once the endogeneity of warranties is considered, however, there is a significant effect of warranty length on demand, while service quality does not have a significant effect when this variable is considered in isolation. Our estimates imply a median willingness to pay for one year of warranty of about $850, and that for a vehicle with average characteristics in our sample, a 1% price decrease has the same effect on consumer demand as an improvement in warranty length of 9% or as a 3% improvement in the vehicle's product quality.

We also investigate complementarities and substitution effects between warranty length, service quality and product quality. Indeed, the following example from a survey by CNW Marketing Research[5] illustrates that there may be important differences in warranty effects across firms. The company conducted a survey (September-November 2006) with shoppers of GM, Hyundai, and Toyota, asking them how important the length of the warranty was in their shopping decision. The results of the survey revealed that 45.1% of all intenders considered the warranty length to be "extremely or very important" in order to have these companies on their shopping list. Breaking down the results at the company level, however, showed important differences, i.e., the percentage of intenders that considered the length of the warranty extremely or very important was 54.6% for Hyundai intenders, 53.4% for GM intenders, and only 28.4% for Toyota

---

[5] Available at http://www.cnwbyweb.net (subscription required to access data)

intenders. To our knowledge, the existing literature has not offered an explanation consistent with these data. While multiple hypotheses can be proposed to explain the difference in these specific cases, our research conducts a systematic analysis to understand the joint influence of service attributes and product quality on consumer demand. We propose that warranty length, service quality and product quality, interact in a non-trivial way in the consumer's value function, and we investigate the nature of these interactions. In particular, we test whether the effect of service attributes on consumer demand is independent of, or is a complement or substitute for product quality. The theory of compensatory effects on consumer decision-making (e.g. Dieckmann et al. 2009) support the hypothesis that both dimensions act as substitutes, i.e., good service serves the main purpose of compensating consumers for poor product quality. Alternatively, service attributes could be complements with product quality if consumers see both dimensions as reinforcing their brand preference, i.e., if the primary effect of offering good product quality and good service is to create better brand image. Our results indicate that the value that consumers derive from warranty length and service quality in the U.S. automobile industry increases when product quality decreases, i.e. service attributes have a bigger impact on demand when product quality is low, providing evidence for a compensatory rather than a complementary role of services relative to product quality. Similarly, we test whether both service attributes have independent, complementary or substitution effects on demand, and find evidence for a complementary relationship in this case, which is contrary to our findings for the case of product quality. This suggests that a firm that increases its warranty length would make the most out of

this investment (in terms of its impact on demand) by simultaneously investing in providing better service quality. The results of our analysis thus indicate that the joint consideration of product and service is essential for the development of an effective competitive strategy.

3.2. Related literature

Service competition is a major topic of interest in operations management (OM). In traditional service industries, theoretical models have examined competition when consumer demand depends on price and service levels (So 2000, Cachon and Harker 2002, Allon and Federgruen 2007 and 2009, Bernstein and Federgruen 2007), and in empirical OM research several studies have tested some of these and related theories in, e.g., the fast food industry (Allon et al. 2011) and the banking industry (Buell et al. 2011). In manufacturing industries, in contrast, service competition has been the subject of theoretical models in OM, e.g., service competition between a manufacturer and a retailer (Cohen and Whang 1997), between retailers that interact strategically with a manufacturer (Tsay and Agrawal 2000), and between manufacturers (Lu et al. 2011). The empirical evidence in the case of manufacturing industries however, is scarce, and indeed we are not aware of any OM papers that analyze the impact of service competition on consumer demand in manufacturing industries empirically. In the economics literature, on the other hand, theoretical models of product differentiation (e.g. Shaked and Sutton 1982, Caplin and Nalebuff 1991), have prompted numerous empirical studies, especially in the automobile industry where researchers have studied different aspects of firm

competition and consumer demand (e.g. Berry et al. 1995 and 2004, Train and Winston 2007, among many others). Similar to the OM literature, these economic models of demand have omitted the role of supporting services by automakers. Our study thus attempts to fill this gap by analyzing the role of service attributes as drivers of consumer demand in the U.S. automobile industry. Moreover, as our results illustrate, considering the interaction between service attributes and product quality is essential in order to disentangle the effects of service attributes on demand in this industry, and thus analyzing service competition in a manufacturing industry offers new evidence that goes beyond what has been done in service industries.

One of the service variables we focus on is length of warranty. Four main rationales regarding the economic role of warranties have been proposed in the literature (see e.g. Emons 1989 for a comprehensive discussion): protection against product failures (insurance role), provision of product quality information to consumers (signaling role), a mechanism to discriminate consumer risk preferences if customer heterogeneity is not fully observable by the seller (sorting role), and to incentivize the seller to improve product quality (incentives role). These theories would thus be consistent with consumer preferences for longer warranties, all else being equal. Regarding the insurance role (e.g. Heal 1977), warranties provide consumers with some security against poor product quality and are often used by manufacturers as a value-added feature to promote their products (Thomas 2006). The signaling argument, on the other hand, predicts that higher quality products will have longer warranties, and is perhaps the one that has received the most attention. Since Spence's model of perfect competition in which warranties serve as

signals of reliability (Spence 1977), several theoretical models have qualified this finding in alternative settings (e.g. Cooper and Ross 1985, Gal-Or 1989). Empirical tests of the signaling role of warranties are also numerous. While early papers like Wiener (1985) showed a positive association between warranty length and product quality providing support for the signaling argument, others like Douglas et al. (1993) showed that the opposite is possible. More recently -and more broadly- Chu and Chintagunta (2011) empirically tested for the different roles of warranties in the U.S. automobile and PC server industries, finding support for the insurance and sorting role of warranties, but not for the signaling and incentives roles. Given the numerous papers studying the economic role of warranties, we do not address this question and rather build our model upon some of the findings in this literature, to study how service attributes such as warranty length and service quality, along with product quality, jointly influence consumer demand.

Empirical models of demand related to ours that include consumer response to warranties include Menezes and Currim (1992) and Chu and Chintagunta (2009). Menezes and Currim (1992) formulate a theoretical model to define the appropriate warranty length for firms, and they also perform some empirical testing in the automobile industry. Their empirical analysis is based on a sales response model (aggregate demand function), for which OLS analysis is performed, and in which several attributes (including aggregate functions of other firms' actions) are used to explain total sales for a given model. They do not deal with the endogeneity of either the price or the warranty length in the demand specification. In a paper more closely related to our study, Chu and Chintagunta (2009) empirically analyze the value of warranties in the U.S. server market.

Their research in this B2B setting quantifies the value of warranties for manufacturers, intermediaries and customers, and finds a positive value for warranties in all cases. As in our case, their demand model is based on a random coefficients logit model that allows for customer heterogeneity and that is based on market data, but they only account for the endogeneity of the pricing decision.

While past empirical studies are certainly relevant for our analysis, we establish at least three important differences. First, these papers focus on warranties exclusively, while our interest is in services more broadly defined, which includes not only the firm's warranty length but also its service quality in the demand specification. Second, our focus is on understanding the effect on demand of the interaction between service attributes and product quality, in order to enlighten firm decision-making regarding both of them. To our knowledge, our findings in this regard have not been established in previous empirical literature. Third, from a methodological perspective, unlike past papers, note that we explicitly deal with the endogeneity of both pricing and warranties, and our identification strategy for warranty effects could be used in other settings[6]. With respect to the third aspect, most existing models of product differentiation have accounted for the endogeneity of prices in the demand specification, under the assumption of exogeneity of all other product characteristics. This assumption has been recognized as an important shortcoming in this literature (Berry 1994). As a response, a recent and growing stream of research on endogenous product choice has considered models in which some product

---

[6] See the section on model identification for a discussion of the required assumptions under which our identification strategy is valid.

characteristics other than price are treated as endogenous (see Crawford 2012 for a recent review of this research stream; good examples include Draganska et al. 2009 and Fan 2011). Our research thus also relates to the endogenous product choice literature, as we deal with the endogeneity of both pricing and the warranty length decision by firms. Finally, our research is also related to the numerous empirical studies in OM that examine the automobile industry, including Fisher et al. (1999), Ramdas and Randall (2008), Olivares and Cachon (2009), and Gallino et al. (2012), among many others, that as ours attempt to examine some aspect that contribute to an understanding of factors influencing the matching between what firms supply and what consumers demand in this industry.

In short, this study contributes to the aforementioned streams of research by being one of the first to empirically analyze the value of service attributes as drivers of demand in manufacturing industries, and by being (to our knowledge) the first study to empirically analyze complementarities between service attributes and product quality in the context of demand models in a competitive manufacturing setting. Product quality, service quality and warranty length are variables of longstanding importance in OM research. The new empirical evidence of their impact on demand in a competitive setting provided in this study, contributes to a better understanding of the strategic implications of the joint management of these variables by firms.

3.3. Data and Industry Background

Market-level data was collected from different sources for our analysis. We obtained data

on sales and product characteristics from Ward's Automotive, for all new light cars sold in the U.S. in the period 2001-2007. Vehicle specifications include miles-per-gallon, length, width, height, horsepower and weight, among other features. Data about warranty length were obtained from Automotive News for the period 2003-2007; we completed and validated the data for the period 2001-2007 from the manufacturers' websites and the 2009 Official Warranty Guide (J&L Warranty Pros). Data on product quality and service quality at the brand level were obtained from J.D. Power's press releases. Aggregate yearly data on transactional prices were obtained from a secondary source, based on J.D. Power data. Below, we discuss some characteristics of these data sources in more detail, along with aspects of the industry that help to clarify our analysis.

Warranty data: Automakers include manufacturer warranties bundled with the purchase of every new car, to protect consumers against defects for a certain period of time/usage. There are three main types of warranties bundled with a new car: basic, powertrain, and corrosion. The basic warranty (a.k.a. bumper-to-bumper) is the most comprehensive and covers most parts of a vehicle. The powertrain warranty (a.k.a drivetrain) covers the major cost components of the car such as the engine, transmission, etc., usually for an extended period of time (for some brands in some years, the coverage period is the same for basic and powertrain warranties). The corrosion warranty covers the vehicle against rust. For example, the Acura 2007 model year vehicles had basic, powertrain, and corrosion warranties of 4/50,000, 6/70,000, and 5/unlimited [years/miles], respectively. For a given brand, there is a high correlation between the warranty terms for these three types of warranties, and also between the years/miles metrics. Most of the existing

studies on warranties have focused on the duration of the powertrain warranty for several reasons. First, the powertrain warranty covers the most expensive parts of the vehicle. Second, most of the changes in warranty strategies by OEM's refer to powertrain warranty duration and therefore is the richer source of longitudinal variation, e.g., Acura from 4/50,000 to 5/70,000 in 2006, Chevrolet from 3/36,000 to 5/100,000 in 2007, Kia from 5/60,000 to 10/100,000 in 2001, Mazda from 3/50,000 to 4/50,000 in 2003, Mitsubishi from 5/60,000 to 10/100,000 in 2004, among many others. Finally, the powertrain warranty is the warranty that automakers advertise the most. Consistent with these arguments, we use the length of the powertrain warranty in years as our warranty variable.

Quality data: J.D. Power publishes yearly reports on product quality and service satisfaction at the brand level. Here we provide a brief description of the data used in our analysis; further details can be obtained via http://www.jdpower.com/.

Our product quality metric is based on J.D. Power's Initial Quality Study (IQS), which determines the number of problems per 100 vehicles in the first 90 days of ownership. The study examines 217 vehicle attributes, and reports on a broad range of problems reported by owners, including defects/malfunctions (complete breakdown or malfunction of any component, feature, or item) and design problems (components or features that may be functioning as designed, but are perceived to be difficult to use or understand, or are in a poor location). Every year this information is summarized in a brand-level metric. For example, in 2004 the best brand was Lexus with 87 problems per 100 vehicles, the worst was Hummer with 173 problems per 100 vehicles, and the

industry average was 119 problems per 100 vehicles. In 2007, the best brand was Porsche with 91 problems per 100 vehicles, the worst was Land Rover with 170 problems per 100 vehicles, and the industry average was 125 problems per 100 vehicles. We take the negative of the number of problems per vehicle as our product quality metric PQ_{jt}, such that a higher value for PQ_{jt} (smaller number of problems per vehicle) denotes higher product quality.

Similarly, J.D. Power's Customer Service Index (CSI) measures the satisfaction of vehicle owners who visited the dealer service department for maintenance or repair work during the first three years of ownership. According to the J.D. Power's description, the CSI study "provides an overall customer satisfaction index score based on six measures: service initiation, service advisor, in-dealership experience, service delivery, service quality, and user-friendly service." The score is based on a 1000 point scale. For example, in 2004 the best brand was Lincoln with a score of 912, the worst was Daewoo with a score of 754, and the industry average was 862. Similarly, in 2007 the best brand was Jaguar with a score of 925, the worst was Isuzu with a score of 780, and the industry average was 876. We note that this metric refers to after-sales service at dealers during the first three years of ownership, which is coincident with the minimum warranty period observed in the industry. This metric thus reflects services that occurred during the in-warranty period and, in conjunction with the warranty length, defines the variables that we use to characterize the service dimension of a brand. Specifically, our service quality variable SQ_{jt} is a scaled version of the customer service index (CSI score/1000).

Figure 3.1 displays the relationship between the IQS and CSI metrics for 2004 and 2007, for the brands in our sample. Figures 3.2 and 3.3 do the same for the relationship between warranty length and IQS and CSI, respectively. We note several interesting observations. Pooling the data for our period of analysis at the brand level (2001-2007), we obtain a correlation of -0.64 between IQS and CSI, denoting a positive relationship between product quality and service quality (recall that the IQS index reflects negative product quality). As illustrated in Figure 3.1, most brands are located on the diagonal of the graph. The graph from 2004 suggests some exceptions, like Saturn (low product quality, high service quality), and Hyundai and Toyota (high product quality, low service quality). Similarly, we obtain correlations of 0.13 between warranty and IQS, and -0.29 between warranty and CSI. These statistics reflect a negative correlation between warranty length and both product quality and service quality. Note that the negative correlation between warranty and product quality counters the signaling role of warranties.
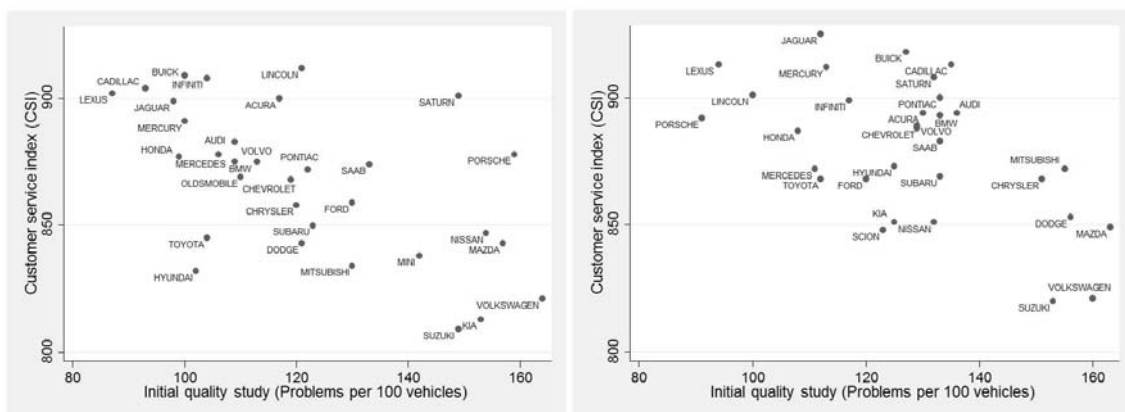
Figure 3.1: CSI vs. IQS (2004: left, 2007: right)
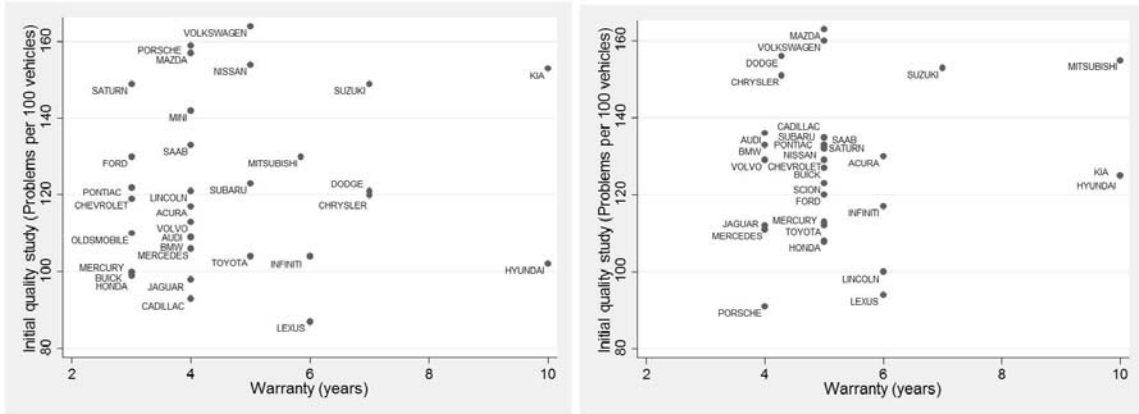
Figure 3.2: IQS vs. warranty length (2004: left, 2007: right)
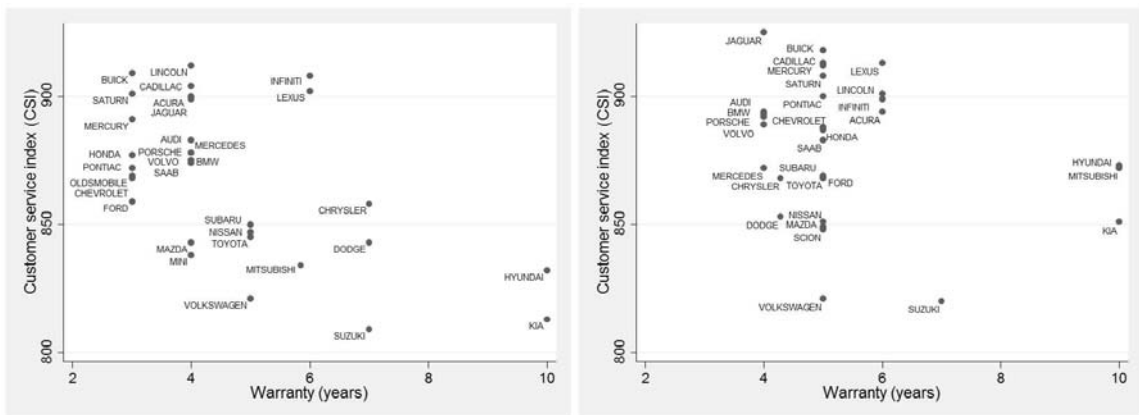
Figure 3.3: CSI vs. warranty length (2004: left, 2007: right)

Finally, we note that, while the service experience at dealers is not fully determined by OEM's, they can and do influence the service process in several ways (see e.g. Cohen et al. 2000). First, OEM's impose guidelines and service standards on their dealers. Second, they can facilitate the quality of service delivered by dealers through a wide range of managerial interventions, e.g., by setting up parts pooling mechanisms, sharing information, using vendor-managed inventory and implementing a generous parts return policy for dealers. Third, OEM's usually set up incentive programs, whereby a dealer's compensation is, in part, based on service performance. Finally, the design of the service

network, for example, the definition of the number of dealers, is ultimately defined by the OEM.

Sales, prices and product characteristics: We obtained data on sales and product characteristics from Ward's Automotive for all new light cars sold in the U.S. between 2001 and 2007. This includes cars belonging to the segments small, middle, large and luxury, as categorized by Ward's. Sales data is available at the make-model level (e.g. Toyota Corolla) monthly. Data on product characteristics (e.g. miles-per-gallon, length) is available for each model year and for each of the versions of a given make-model. As noted in existing studies (e.g. Berry et al. 1995, Sudhir 2001), a certain level of aggregation is required to match sales data with the respective product characteristics. For a given product characteristic, e.g. length, we consider the average length of the options of a given model year as the length associated with that model year (an approach also taken by Balachander et al. 2009).

In addition, we obtained yearly data on transactional prices at the make-model by model year level from a secondary source based on J.D. Power data. These data are collected at the daily level by J.D. Power from a sample of dealers in the U.S. covering about 70% of the geographical areas and 15-20% of total U.S. sales. These data reflect transactional prices paid by consumers after rebates and as such, are more informative of actual consumer expenses than the manufacturer suggested retail price which is usually used in research papers due to the unavailability of information about transactional prices. We only have access to these data at the aggregate yearly level, more precisely, the average across time of the transactional prices paid by consumers for a given make-

model by model year in the period September-August of each year, which is the definition used for calendar year in our analysis. Note that sales data from Ward's do not distinguish between different model years of a given make-model sold in the same calendar year. In practice, however, in a given calendar year different model years of the same make-model are sold simultaneously. The pricing data set contains sales at the model year level for each calendar year for the sample of dealers described above. We use the distribution of sales in this data set and apply it to the market level sales data from Ward's to obtain sales at the make-model by model year level. A similar approach to matching both data sets was taken by Copeland et al. (2011).

Sample: We match all data sources as described above. Our final sample consists of 2122 yearly observations for all new light cars sold in the U.S. in calendar years 2001-2007, which includes model years from 2000 to 2008. Our unit of analysis is a make-model by model year and calendar year, e.g. 2005 Toyota Corolla in calendar year 2005.

3.4. Model

In this section we describe the structural model formulated to study the role of service attributes as drivers of demand in the automobile industry. It considers decision-making by both consumers (demand model) and firms (supply model). In what follows, we describe the demand model in detail, and provide a high level discussion of the underlying supply model. As will be shown, while we do not estimate supply side parameters in our analysis, our empirical formulation does make extensive use of the underlying supply model to derive identification conditions for demand parameters. Thus,

71

the formulation of this supply model and its assumptions allow us to deal with the endogeneity of the warranty length decision by firms in the demand specification. We do discuss the identification strategy in detail, and finalize the section with a brief outline of the estimation procedure.

3.4.1. Demand

In this section we describe the structural model formulated to study the role of service attributes. We consider a random coefficients logit demand model, where the utility that consumer i derives from purchasing vehicle j ($j = 1, ..., J$) in calendar year t ($t = 1, ..., T$) depends on the vehicle price p_{jt}, warranty duration w_{jt}, product quality PQ_{jt}, service quality SQ_{jt}, and a vector of observable vehicle characteristics (size, horsepower to weight ratio, etc.) x_{jt}, as follows:

$$u_{ijt} = \alpha_i p_{jt} + x_{jt} \beta_i + h(w_{jt}, PQ_{jt}, SQ_{jt})\gamma + \xi_{jt} + \varepsilon_{ijt} \tag{3.1}$$

The term ξ_{jt} represents unobserved product attributes common to all consumers, and ε_{ijt} is a type I extreme value idiosyncratic shock. Consumers maximize utility, and purchase vehicle j in calendar year t if and only if $u_{ijt} \geq u_{irt}$ for all $r = 0, 1, ..., J$. Here, $r = 0$ defines the outside good, i.e. the option of not purchasing a new light car in year t, where $u_{i0t} = \varepsilon_{i0t}$. The individual-level coefficients α_i and β_i are decomposed into a mean effect common to all consumers (β's) and individual deviations from that mean (σ's), as is common in the literature (e.g. Berry et al. 1995, Sudhir 2001). The total effect of attribute x_{jt}^k on the utility of consumer i is thus modeled as $(\beta_k + \sigma_k v_{ik}) x_{jt}^k$, where β_k and σ_k are

72

parameters to be estimated, and v_{ik} is a shock from a standard normal distribution; the same holds for α_i. It is useful to note that u_{ijt} can be thus expressed more compactly as a function of the mean utility δ_{jt} common across all consumers, and the heterogeneity terms μ_{ijt} and ε_{ijt} as:

$$u_{ijt}=\delta_{jt}(p_{jt},x_{jt},w_{jt},PQ_{jt},SQ_{jt};\theta_1)+\mu_{ijt}(p_{jt},x_{jt},w_{jt},PQ_{jt},SQ_{jt},v_i;\theta_2)+\varepsilon_{ijt} \tag{3.2}$$

Here, θ_1 is a vector containing all parameters of the mean utility (α, β, and γ), and θ_2 all the heterogeneity parameters (σ). Let d_{jt} contain all M vehicle characteristics. δ_{jt} and μ_{ijt} are thus defined as:

$$\delta_{jt} = \alpha\, p_{jt} + x_{jt}\beta + h(w_{jt},PQ_{jt},SQ_{jt})\gamma + \xi_{jt} \tag{3.3}$$

$$\mu_{ijt} = \sum_{m=i:...M} \sigma_m d^m_{jt} v_{im} \tag{3.4}$$

The function $h(w_{jt},PQ_{jt},SQ_{jt})$ defines the way in which warranty length, product quality and service quality enter into the utility function. Under the linearity assumption for these variables, the utility function would take the following form:

$$u_{ijt} = \alpha_i p_{jt} + x'_{jt}\beta_i + \gamma^1 w_{jt} + \gamma^2 PQ_{jt} + \gamma^3 SQ_{jt} + \xi_{jt} + \varepsilon_{ijt} \tag{3.5}$$

This formulation is useful to capture the main effects of the variables of interest, and is also consistent with the linearity assumption made for the rest of the covariates. We refer to the model derived from the utility function in Eq. (3.5) as the main effects model.

We are also interested in testing, however, whether service attributes and product quality act as complements, substitutes or independently in the demand function. For this purpose, we consider an enhanced formulation in which the function $h(w_{jt},PQ_{jt},SQ_{jt})$ not

only includes the main effects for these three variables -as described in Eq. (3.5)- but also their two-way interaction terms $w_{jt}*PQ_{jt}$, $w_{jt}*SQ_{jt}$, and $PQ_{jt}*SQ_{jt}$. These interactions reflect the non-linearities of interest, i.e., the complementarity/substitution effects between service attributes and product quality. A recent study by Guajardo and Cohen (2012) has shown that service quality and product quality act as complements in terms of how they determine the likelihood to recommend the brand in an application in the consumer electronics industry, i.e., the better the perceptions of product quality of a person, the higher the impact of better perception about service quality on the person's likelihood to recommend the brand. In the context of our demand model, we would expect service attributes and product quality also to be complements if the dominating mechanism by which they affect consumer demand is through their impact on brand image. On the other hand, service attributes could act as substitutes for product quality if the main mechanism by which they affect consumer demand is by compensating consumers for poor product quality, i.e., if the car purchase process can be described as a compensatory process with respect to these attributes. In a compensatory decision process (e.g. Dieckmann et al. 2009), strengths along one or more dimensions of product or service quality can compensate for weaknesses along others. This is in contrast to the case of non-compensatory processes, in which no compensation is possible if certain attributes of a product or service are weak, even if it possesses strengths along other dimensions. The role of compensatory effects on consumer decision-making would thus provide a basis for characterizing our service attributes as substitutes for product quality. Similar arguments can be hypothesized for the interaction between warranty and service

quality, i.e., a negative interaction under the hypothesis of compensatory attributes, or a complementary (positive) relationship not only if they both contribute to better brand image, but also if the main mechanism by which they affect consumer demand is by providing complementary functionality (longer and better service support). In the case of warranties and product quality, their insurance role (Heal 1977, Emons 1989) implies that warranties should be more important for consumers when the product is expected to fail more often, i.e. when product quality is lower, which would provide additional support for the hypothesis of a negative coefficient for the term $w_{jt}*PQ_{jt}$. Alternatively, all three attributes may exhibit independent effects on consumer decision-making, in which case no significance would be obtained for the interaction terms. In this scenario of competing theories, whether service attributes act as complements, substitutes, or independently of product quality in the demand function is ultimately an empirical question, which we test in our analysis.

3.4.2.   Supply

We assume that firms compete on prices and warranties. This assumption is consistent with some prior theoretical models (e.g. Spence 1977, Gal-Or 1989), which have modeled competition based on these two variables, taking other factors such as product quality as given[7]. As noted, by offering warranties, firms incur important warranty costs. We incorporate these costs in our formulation, and define the profit function for firm f in

---

[7] Naturally, firms choose not only prices and warranties, but also vehicle characteristics and positioning with respect to product quality and service quality. The main argument behind our formulation is based on the nature of the model and the timing of firm decisions. While firms can easily adjust prices and warranties, decisions on vehicle characteristics as well as actions influencing product quality and service quality occur over a longer time horizon than the one we consider here. We discuss this issue more extensively in subsequent sections.

period t as follows:

$$\pi_f = \sum_{j \in J_f}(p_{jt} - mc_{jt} - wc_{jt})Ms_{jt}(\boldsymbol{p_t}, \boldsymbol{w_t}, \boldsymbol{PQ_t}, \boldsymbol{SQ_t}, \boldsymbol{x_t}, \boldsymbol{\xi_t}; \boldsymbol{\theta}) \quad\quad (3.6)$$

In Eq. (3.6), J_f represents the set of vehicles produced by firm f, mc_{jt} the marginal

costs of production of vehicle j, wc_{jt} the expected per-unit warranty costs, M is the size

of the market, and s_{jt} the market share of vehicle j (market-level variables are in bold). As

in most existing models, (e.g. Berry et al. 1995, Sudhir 2001), we consider a marginal

cost function g_1 based on the projection of costs onto observable vehicle characteristics

x_{jt}^s (e.g., horsepower, size) and unobservable cost shifters φ_{jt}, i.e.,

$$mc_{jt} = g_1(x_{jt}^s, \varphi_{jt}) \quad\quad (3.7)$$

Next, we consider the warranty cost function. Generically, let $N(t)$ be the stochastic

process for the number of failures of a vehicle by time t, and $Y_n(t)$ the cost of failure n at

time t, independent of $N(t)$. A standard formulation for expected warranty costs (e.g.

Thomas 2006 pp. 67) if the warranty length is set to W is thus wc(W) =

$E[\sum_{n=1,\ldots,N(W)}Y_n(t)]$. If the time between failures is iid, the expected warranty costs are

given by wc(W)= $E[N(W)]E[Y_n(t)]$. The term $E[N(W)]$ represents the expected number

of failures during the warranty period, which depends on the warranty length and the

failure process. For example, if $N(t)$ is assumed to be a homogeneous Poisson process

and λ is the failure rate per time unit, then $E[N(W)]=\lambda W$; if the failure process is more

complex, in general $N(t)$ will not necessarily have a tractable closed-form solution. For

our purposes (and using our notation), however, it suffices to note that $E[N(W)]$ is a

function of product quality PQ_{jt} and the warranty length w_{jt}. With respect to the expected

cost per event, $E[Y_n(t)]$, we must consider heterogeneity across brands. In particular, and

in line with the previous literature (e.g. Cohen and Whang 1997), providing a certain

level of service quality is costly, and thus the cost per event will depend on the quality of

service provided when servicing the vehicle, for which our SQ_{jt} variable can serve as a

proxy. Let xb_{jt} denote other observable characteristics that capture part of the brand

heterogeneity in warranty costs, and ς_{jt} unobservable factors. The warranty costs function

g_2 can be thus represented conceptually as:

$$wc_{jt} = g_2(w_{jt}, PQ_{jt}, SQ_{jt}, xb_{jt}, \varsigma_{jt}) \tag{3.8}$$

Finally, we turn to firm behavior. The vast majority of studies on product

differentiation focus exclusively on firms' pricing behavior. In our case, we assume that

firms compete on both prices and warranties, and make their decisions in order to

maximize profits in each period, according to the profit function (3.6). While we do not

estimate supply side parameters, the supply model presented here, and the drivers of

marginal costs and warranty costs in particular, provide the fundamentals for our

identification strategy for the parameters in the demand model.

3.4.3. Identification and Instruments

In practice, all the observed variables in our demand specification $(p_{jt}, x_{jt}, w_{jt}, PQ_{jt}, SQ_{jt})$ are

determined or influenced by firm decisions. On the other hand, ξ_{jt} reflects characteristics

or shocks not observed in the data, such as style, prestige, and reputation, that affect the

demand for different products. An endogeneity problem for the demand parameters emerges if some of the observed variables are set by firms upon observing the demand shocks ξ_{jt}. As noted, most existing studies have accounted for the endogeneity of prices in the demand specification under the assumption of exogeneity for all other product characteristics. While this assumption has been widely acknowledged as a shortcoming since Berry (1994), the underlying argument for it relies on the fact that, while the prices are easily adjustable by firms according to the market conditions and therefore p_{jt} is likely to be correlated with ξ_{jt} (i.e. price endogeneity), other product characteristics captured by x_{jt} (e.g. horsepower, size) are defined by firms well in advance of the time when a model is sold in the market, and thus are assumed to be uncorrelated with ξ_{jt}. To account for the endogeneity of prices, instrumental variables can be used in the estimation. A well-known example is Berry et al. (1995)'s model for the auto industry involving prices p_{jt} and product characteristics x_{jt} in the demand specification.

Their supply model considers firms competing on prices, and makes a Bertrand-Nash equilibrium assumption. Under these assumptions, they propose a set of instruments to deal with price endogeneity: the sum (or average) of product characteristics x_{jt} for (i) other cars of the same firm, and (ii) cars of other firms. Product characteristics x_{jt} are exogenous by assumption, and are thus also used as instruments. This set of instruments has been widely used to deal with price endogeneity since then (e.g., Sudhir et al. 2001, Train and Winston 2007, Balachander et al. 2009). We use this set of instruments to deal with price endogeneity; as in Sudhir (2001), instead of considering the average

characteristics for cars of all other firms, we compute the average characteristics of other firms' cars in the same market segment (small, middle, large, luxury), which refines the set of instruments by using cars that are *closer* to each other in terms of characteristics.

Our specification of the demand model involves not only prices and vehicle characteristics, but also brand level attributes w_{jt}, PQ_{jt}, and SQ_{jt}. As in the case of prices, firms can easily set the length of the warranty w_{jt} in response to the unobserved factors in ξ_{jt}. A similar observation was made by Menezes and Currim (1992), who noted that in contrast to changes in product quality, changes in warranty length and price could be carried out almost instantaneously. Thus, warranties are expected to be endogenous in the demand specification in the same way (in terms of timing) as prices. Conversely, let us consider firm actions to influence PQ_{jt} and SQ_{jt}. Note that firms can affect product quality by introducing changes in product design, using better parts/components (Ramdas and Randall 2008), redesigning their processes, etc. All of these factors will be reflected over a term longer than our yearly period of analysis. The time-to-market of a vehicle, for example, can take several years from the beginning of the design stage to product launch. Similarly, factors influencing service quality such as the implementation of optimization-based technologies for the management of parts inventories, more investment in spare parts inventory, the design of a more efficient service network, and a higher focus on services more generally, will usually involve long-term efforts and cultural changes by firms (see e.g. Cohen et al. 2000). We thus argue that the observed PQ_{jt} and SQ_{jt} are not easily adjustable contemporaneously by firms upon observing the shocks ξ_{jt}, and therefore will consider PQ_{jt} and SQ_{jt} to be exogenous in the demand specification. We

derive instruments for the warranty length based on the exogeneity assumption for PQ_{jt} and SQ_{jt} and the structure of our supply model as follows. Consider vehicles j and r, produced by different brands. Note that given our model in which firms compete on prices and warranties, w_{jt} and w_{rt} are the result of the strategic interaction of firms and are therefore correlated. If firms set their warranties optimally (or at least, take into account the expected warranty costs), w_{jt} will be correlated with the drivers of warranty costs, e.g. PQ_{jt} (see Eq. 3.8). Similarly, w_{rt} will be correlated with PQ_{rt}. Noting that u_{ijt} does not depend on the attributes of vehicle r, then PQ_{rt} is a valid (source of) instruments for w_{jt}. We thus consider the average of product quality of other brands as an instrument for the warranty of a given brand.

$$w_{jt}, \ldots\ldots\ldots\ldots, w_{rt}$$

$$PQ_{jt}, \ldots\ldots\ldots\ldots, PQ_{rt}$$
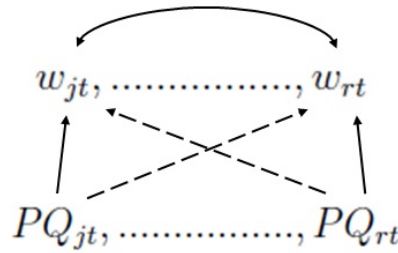
Figure 3.4: Instruments

We apply this same argument to generate instruments using the rest of the drivers of warranty costs, i.e. SQ_{jt} and xb_{jt} (Eq. 3.8). In xb_{jt} we include dummies for the region of the manufacturer (coded into three categories: USA, Europe, Asia) to partially capture heterogeneity across brands. Thus, our set of instruments for warranties includes the average product quality of other brands, the average service quality of other brands, and the proportion of brands belonging to the different geographical regions. Note that

heterogeneity at the vehicle model level is already captured through the x_{jt}-based instruments. Finally, since 2003 it has been mandatory for firms traded in the U.S. to disclose warranty costs in their financial statements, which was private information before that. This modifies the information set under which firms make their warranty decisions and, being unrelated to the demand side, may thus serve as additional source of exogenous variation to instrument for warranties. We thus construct an indicator function (pre vs. post 2004 calendar year) and include it as an additional instrument.

Finally, note that we observe cross-sectional and longitudinal price variation for all models and years, as well as for product quality and service quality for all brands and years. Although the variation in warranties is more limited (e.g. it does not allow us to include brand fixed effects in the demand specification), we do observe cross-sectional variation across brands in our warranty variable each year and longitudinal variation for several brands at some point in our observation period. For some of these brands with longitudinal changes we also observe variation in warranty length for different model years of the same make-model being sold in the same calendar year. Finally, there is variation in the observed warranties in the market (as well as in the rest of the variables) due to changes in the choice set of vehicles available in the market each year. Also note that, along with the warranty length, we include other brand-level variables in the demand specification (product quality, service quality, and manufacturer geographical region), which alleviates concerns about brand fixed effects as potential confounders.

3.4.4. Estimation

The estimation of random coefficients demand models is discussed in detail in Berry (1994), Berry et al. (1995), and elsewhere. Here we briefly review the key aspects of the estimation procedure[8].

Under the Type I extreme value distribution assumption for ϵ_{ijt}, the market share for product j in calendar year t obtained from Eq. (3.2) is given by:

$$s_{jt} = \frac{exp(\delta_{jt}+\mu_{ijt})}{1+\sum_{k=1}^{J} exp(\delta_{kt}+\mu_{ikt})} = \int_{v} \frac{exp(\delta_{jt}+\mu_{ijt}(v_i,...;\theta_2))}{1+\sum_{k=1}^{J} exp(\delta_{kt}+u_{ikt}(v_i,...;\theta_2))} P(v)dv \tag{3.9}$$

where $P(v)$ is the joint distribution over all elements of v_i, which in our case is the product of standard normals. The integral in Eq. (3.9) does not have a closed form, and is evaluated using simulation, drawing values from the distribution of v for a sample of individuals. The estimation of the model proceeds as follows. For a given draw of θ_2, the actual[9] and predicted (Eq. 3.9) market shares are equated by means of a contraction mapping that allows us to obtain a unique solution for δ_{jt}, which is in turn used to compute the value of ξ_{jt}, or more precisely, ξ_{jt} (θ) (inner loop). Let Z denote the available instruments, which includes the exogenous characteristics in the demand specification. The sample analogs to the moment conditions $E[\xi Z] = 0$ can thus be constructed by using $\xi(\theta)$. An outer loop searching for the parameters $\hat{\theta}$ that solve the minimization of the GMM objective function completes the estimation routine

---

[8] For additional details, we refer the interested reader to Knittel and Metaxoglou (2012) and Nevo (2000). Our implementation largely follow theirs.

9 Market shares are obtained by dividing actual sales by the market size. As in most previous studies, we define the market size as being the number of households in the U.S. for a given year. Data on the number of households was obtained from the U.S. Census Bureau (available at http://www.census.gov)

(i.e., $\hat{\theta} = arg_\theta min\xi(\theta)'Z\Phi^{-1}Z'\xi(\theta)$). Here, the weighting matrix Φ is a consistent estimate of E[Z'ξ(θ) ξ(θ)' Z], and is obtained employing the usual two-stage procedure (see Nevo 2000 for more details). Finally, as noted by Knittel and Metaxoglou (2012), the estimation procedure is subject to variability depending on the optimization algorithms and initial values considered. Similarly, Dube et al. (2011) note the dangers of using loose tolerance levels in the estimation procedure. Consistent with best practices recommended in both cases, we use multiple optimization algorithms, 50 different starting values, and best-practice tolerance levels in our implementation[10].

3.5. Empirical Analysis

3.5.1.   Main effects model

Our specification for x_{jt} builds upon existing literature, using variables similar to those used by Berry et al. (1995), Sudhir (2001) and Balachander et al. (2009). We include the size of the car measured as the product between the length and the width (SIZE), the ratio of horsepower to weight (HPWT), and the miles per dollar (MPD) of the vehicle as product characteristics in x_{jt}. The MPD variable is obtained by dividing the miles-per-gallon by the dollars-per-gallon in a given year. We obtained monthly average prices for gasoline from the U.S. Department of Energy (http://www.eia.doe.gov), which are aggregated at the calendar year level to calculate the MPD variable. Fuel prices are expressed in 2007 dollars using the CPI index for the respective year, and the same is

---

[10] We use five optimization methods in our experiments (available in the implementation by Knittel and Metaxoglou 2012): quasi-newton 1 and 2, nelder-mead simplex, solvopt and conjugate gradient. Also, we use tolerance levels of e^{-14} for the inner loop and e^{-6} for the outer loop.

done for the vehicle price p_{jt} (PRICE), i.e., all monetary variables in our analysis are expressed in 2007 dollars (data on CPI's were obtained from the U.S. Department of Labor, Bureau of Labor Statistics, available at http://www.bls.gov/cpi/cpirsdc.htm). We also include in x_{jt} dummy variables to indicate whether a model year is from the previous year (PREVY_MY) or the next year (NEXTY_MY), dummy variables for manufacturer region (MANUF_EUR, MANUF_ASIA), dummy variables to indicate whether the model was launched in the last 2 years (INTRO2Y) or is soon (2 years) to be out of the market (EXIT2Y), and a time trend (TREND). Table 3.1 displays descriptive statistics for the relevant variables in our sample of 2122 observations, including statistics for the warranty length (WARR), product quality (PQ), and service quality (SQ). Table 3.2 displays the correlation matrix for these variables.

Variable	Mean	Std. Dev.	Min	Max
PRICE ($1,000)	34.938	22.161	10.286	170.689
WARR (years)	4.7	1.9	3	10
PQ (-1*problems per vehicle)	-1.289	0.235	-2.670	-0.760
SQ (CSI score/1000)	0.863	0.032	0.781	0.925
MPD ([miles/$]/10)	1.028	0.336	0.439	3.770
HPWT (100 x hp/lb)	0.643	0.171	0.203	1.578
SIZE (sq. inches/10,000)	1.314	0.145	0.792	1.708

Table 3.1: Descriptive statistics.

Variable	PRICE	WARR	PQ	SQ	MPD	HPWT	SIZE
WARR	-0.19***						
PQ	0.29***	-0.10***					
SQ	0.26***	-0.24***	0.66***				
MPD	-0.39***	0.02	-0.18***	-0.41***			
HPWT	0.76***	-0.18***	0.24***	0.30***	-0.56***		
SIZE	0.28***	-0.22***	0.21***	0.32***	-0.44***	0.24***	

Table 3.2: Correlation matrix. Note: *, **, ***, Significant at the 0.1, 0.05, 0.01, confidence levels, respectively.

We start our discussion of the results with the estimation of the main effects model. The random coefficients model, which allows for customer heterogeneity and accounts for the endogeneity of prices and warranties, is obtained by performing the estimation procedure described in the previous section[11]. If customer heterogeneity is ignored (i.e. $\mu_{ijt} = 0$), the model can be estimated by OLS regression (if the endogeneity of prices and warranties is not accounted for) or by using instrumental variable techniques (e.g. 2SLS). We refer to the latter case as IV LOGIT. Table 3.3 displays the results obtained in each of the aforementioned cases.

The results in Table 3.3 show that the coefficient for the price moves in the expected direction, i.e., demand becomes most sensitive to price as the price endogeneity is accounted for. Indeed, sensitivity to price more than doubles, similar to the findings in Berry et al. (1995) and Petrin (2002). Similarly, the coefficient of the warranty variable has a negative sign in the OLS regression (-0.034). However, once the endogeneity of the warranty length is accounted for using the instrumental variables described in the discussion of our identification strategy (IV LOGIT and Random coefficients model), we obtained a positive and significant effect of warranty length on demand[12]. Note that the fact that a positive coefficient for warranty (0.082 in the random coefficients model, 0.132 in the IV LOGIT) is only obtained after correcting for the endogeneity of the warranty variable and that a negative coefficient is obtained otherwise, is consistent with

---

[11] In our implementation, we include random coefficients for variables for which we observe most substantial variation at the make-model level, i.e. PRICE, HPWT, SIZE and MPD.

[12] We also estimated the model accounting only for the endogeneity of price, ignoring the endogeneity of warranty length (not reported in the text). We obtained a negative coefficient of warranty length in that case.

a scenario in which brand reputation (which is part of the unobservable) is negatively correlated with the warranty length, which in turn explains the bias in the warranty coefficient if we do not employ our strategy for endogeneity correction. Other variables with a significant effect on demand are PQ, HPWT, SIZE, and dummy control variables for vehicle model year, manufacturer region (significant effect for European automakers only), model exit, and the time trend. The results displayed for the random coefficients model also include the magnitude of the estimates for the heterogeneity parameters, which indeed reveals significant heterogeneity effects for price, the horsepower to weight ratio, and the vehicle size.

With respect to the estimation of the random coefficients model, as noted in previous sections, we use different optimization algorithms and starting values. The reported solution in Table 3.3 is the one for which the value of the GMM objective function is minimized (equal to 169.1 in this case), and satisfies both first and second order conditions of optimality (i.e. zero gradient and positive-definite Hessian)[13]. Most important, the results of the main effects model illustrate the effect of the instruments used in estimation, which act to adjust the price and warranty coefficients in the expected direction. In the first stage of the 2SLS procedure, the test for excluded instruments leads to rejection of the null hypotheses of excluded instruments having no explanatory power both in the case of PRICE and WARR (p-value<0.0001 in both cases), with R^2 and F statistic of 0.78 and 222.7 in the case of PRICE, and 0.44 and 50.5 in the case of WARR,

---

[13] The estimation procedure arrives at the same optimal solution in 52% of the runs, which is in the order of magnitude of recent reports, e.g. Knittel and Metaxoglou (2012) and Dube et al. (2011), and is aligned with their findings about the need to use multiple starting values, optimization algorithms, and tight tolerance levels.

respectively. The underidentification test also leads to rejecting the null of underidentification (p-value<0.0001, Anderson LM statistic=146.7). Overall, the model has desirable statistical properties and the tests performed indicate that our proposed instruments are appropriate for our application.

					RANDOM COEFFICIENTS MODEL			
Variable	OLS		IV LOGIT		Main effect (β)		Heterogeneity (σ)	
	Estimate	Std.Err.	Estimate	Std.Err.	Estimate	Std.Err.	Estimate	Std.Err.
PRICE	-0.035***	0.003	-0.083***	0.010	-0.096***	0.010	0.028***	0.006
WARR	-0.034*	0.018	0.132***	0.049	0.082*	0.048		
PQ	0.670***	0.168	0.755***	0.186	0.950***	0.183		
SQ	-7.038***	1.449	1.193	2.257	0.069	2.126		
HPWT	0.059	0.298	4.369***	0.905	2.319**	0.941	0.729***	0.249
SIZE	1.625***	0.257	3.245***	0.422	2.497***	0.440	0.436***	0.118
MPD	-0.169	0.138	0.302*	0.180	-0.110	0.227	0.131	0.194
NEXTY_MY	-2.64***	0.075	-2.652***	0.083	-2.654***	0.089		
PREVY_MY	-1.885***	0.063	-1.928***	0.071	-1.916***	0.060		
INTRO2Y	0.089	0.090	0.115	0.100	0.103	0.094		
EXIT2Y	-0.891***	0.123	-0.943***	0.136	-1.015***	0.132		
TREND	-0.042**	0.019	-0.152***	0.030	-0.265***	0.037		
MANUF_EUR	0.057	0.106	1.209***	0.260	0.907***	0.252		
MANUF_ASIA	-0.101	0.077	-0.049	0.110	-0.034	0.103		
CONSTANT	-1.844	1.533	-13.135***	2.765	-8.626***	2.839		

Table 3.3: Estimation of the main effects model. Note: *, **, ***, Significant at the 0.1, 0.05, 0.01, confidence levels, respectively.

3.5.2.   Model with complementarities

The model in the previous section is useful to study the main effects of our variables of interest and to illustrate the way in which our instrumentation strategy works. As noted earlier, however, we are also interested in investigating complementarities/substitution

effects between service attributes and product quality. We now turn to the discussion of the results of the model involving two-way interactions between warranty length, service quality and product quality. We "mean center" the variables involved in interaction terms (WARR, PQ and SQ), i.e. we subtract the mean from each individual observation, such that the individual coefficients for the single terms of these variables reflect the effect when the other two variables are set to their average values. The results of the random coefficients model are displayed in Table 3.4. The GMM function in the optimal solution is 128.6 in this case, the solution satisfies both first and second order conditions for optimality, and the estimation procedure led to the reported optimal solution in 52% of the runs. Similarly to the main effects model, the model has the desirable statistical properties and the tests performed indicate that our proposed instruments are appropriate. In the first stage of the 2SLS procedure, the test for excluded instruments leads to rejection of the null hypotheses of excluded instruments having no explanatory power in all cases of PRICE, WARR, WARRxPQ, WARRxSQ (p-value<0.0001 in all cases), with R^2 and F statistic of 0.78 and 217.8 in the case of PRICE, 0.45 and 50.1 for WARR, 0.27 and 22.6 for WARRxPQ, and 0.21 and 16.7 for WARRxSQ, respectively. The underidentification test leads to rejecting the null of underidentification (p-value<0.0001, Anderson LM statistic=113.9). Again, the instruments exhibit reasonable statistical properties.

	Main effect (β)		Heterogeneity (σ)	
Variable	Estimate	Std.Err.	Estimate	Std.Err.
PRICE	-0.101***	0.011	0.035***	0.006
WARR	0.142*	0.079		
PQ	1.369***	0.326		
SQ	-2.123	3.089		
WARR X PQ	-1.217***	0.331		
WARR X SQ	7.465***	2.871		
PQ X SQ	-29.236***	6.943		
HPWT	1.021	1.071	1.119***	0.291
SIZE	2.204***	0.527	0.353***	0.106
MPD	-0.562*	0.330	0.318	0.233
NEXTY_MY	-2.728***	0.093		
PREVY_MY	-1.889***	0.064		
INTRO2Y	0.058	0.098		
EXIT2Y	-0.982***	0.147		
TREND	-0.341***	0.045		
MANUF_EUR	0.708**	0.310		
MANUF_ASIA	-0.048	0.107		
CONSTANT	-7.281***	1.314		

Table 3.4: Estimation of the model with complementarities (Random coefficients model). Note: *, **, ***, Significant at the 0.1, 0.05, 0.01, confidence levels, respectively.

Next, we concentrate on the results obtained for our variables of interest, i.e., the joint effect of WARR, PQ, and SQ. The negative and significant coefficient for the interaction term WARRxPQ indicates that the marginal effect of an additional year of warranty coverage on demand decreases with product quality, or, conversely, is higher when product quality is lower. A longer warranty acts as a partial substitute for product quality, which is consistent with the insurance role of warranties. Similarly, we obtain a negative and significant coefficient for the term PQxSQ. Noting that the effect of service quality is not significant when treated in isolation in the main effects model, this result suggests

that service quality is of value for consumers only when the product quality is low. Jointly, these results provide support for the compensatory role of service attributes with respect to product quality, ruling out potential complementarities between product and service attributes in the demand function. In contrast, we obtain a positive and significant coefficient for the term WARRxSQ, indicating that the marginal effect on demand of an additional year of warranty coverage increases with service quality, i.e. there is a complementary relationship between the length of the warranty and service quality.

3.5.3.   Discussion

Our results indicate that warranties have a significant effect on consumer demand, and that the marginal value of an additional year of warranty decreases with product quality and increases with service quality. Further analysis of our main effects model reveals that the median implied willingness to pay for an additional year of warranty, obtained as the ratio of the marginal utility of warranty length to the marginal disutility of price, is approximately $850 which is equivalent to about 2.5% of the average vehicle price in our sample. In other words, and all else being equal, for an average vehicle in our sample, increasing the length of the warranty by one year is equivalent to decreasing the vehicle price by about $850, in terms of their effect on consumer demand. This estimate seems reasonable by industry standards. Indeed, a similar number was quoted in a recent industry report which mentioned that "consumers pay about 2 percent of the vehicle price per year of extended service" (Consumer Reports 2008).

Also, if we focus on the mean utility implied by the main effects model, we note that, for a car with average characteristics in our sample, the effect on demand of a 1% price

decrease is equivalent to increasing product quality by 3%, and is in turn equivalent to increasing the warranty length by 9%. These benchmarks are useful for understanding the relative impact of different managerial interventions with respect to consumer demand. Indeed, these demand estimates can inform managerial decision-making by allowing managers to anticipate the effect of alternative interventions on consumer demand, which together with their usually good knowledge about the costs involved for each of these interventions, could be used to quantify trade-offs involved in managerial decision-making regarding these variables.

Further analysis of our model with complementarities also reveals that the value of one year of warranty is on average about two times more important for U.S. manufacturers than for foreign firms during our period of analysis. Our analysis provides an explanation for this observation, i.e., that U.S.-based brands had on average lower product quality and higher service quality than their foreign counterparts in that period. According to the results of our model, both of these factors imply a higher marginal effect of warranty length on consumer demand. Figure 3.5 illustrates the relationship obtained for the marginal effect of warranty length on consumer utility as a function of PQ and SQ (the grid surface), and also displays some of the brands in our sample.
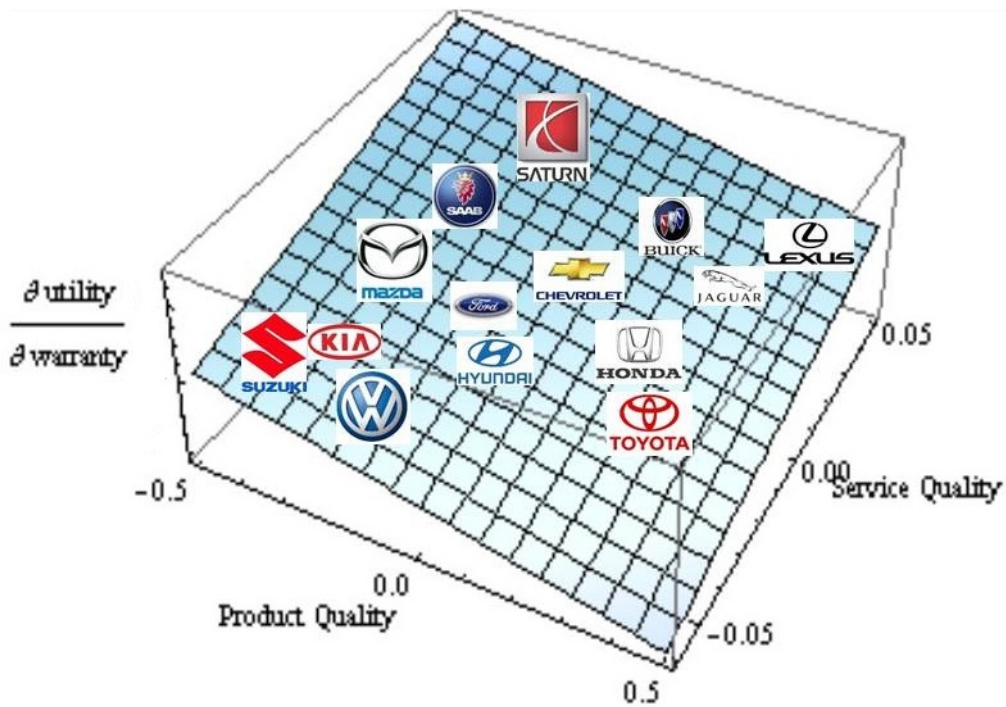
Figure 3.5: Marginal effect of warranty on utility function

As an illustration of the implications of our results, we return to the example in previous sections (survey data Sept-Nov 2006) in which remarkable differences were observed for the importance of the warranty length for customers of Hyundai, GM and Toyota. Our results provide an explanation for these differences which is consistent with the data observed in that case. While the main brands of GM and Toyota both had a 5 year powertrain warranty in the survey period, GM had worse product quality and better service quality than Toyota. The results of our model suggest that both lower product quality and higher service quality increase the effect of warranty length on demand, consistent with the observation from the survey that the warranty length was considered extremely or very important for 53% of GM customers, in contrast to only 28% of Toyota customers, even though the warranty length was 5 years in both cases. Similarly, while

92

Hyundai had slightly better product quality and worse service quality than GM, the fact that Hyundai's warranty was 10 years makes the overall impact of its warranty length on consumer utility higher than in the case of GM and Toyota according to our model, consistent with the observation from the survey that Hyundai customers are those for whom warranty length was of higher importance when considering the brand in the shopping list, among the three brands considered in the survey. Overall, these observations offer face validity of our estimates.

Finally, another interesting implication of our results concerns the complementarity between service attributes. Consider, for example, the case of Kia, a brand that was characterized as having both low product quality and low service quality during our period of analysis. Kia increased the length of their powertrain warranty from 5 years/60,000 miles to 10 years/100,000 miles in 2001, i.e., the brand offered "America's No. 1 warranty" in conjunction with Hyundai. Our results imply that Kia would have benefited the most out of this great warranty coverage (in terms of the effect of this policy on consumer demand), if it had contemporaneously invested in providing better service quality, along with the warranty length increase. In short, being good at one service dimension (service quality), amplifies the effect of being good at another service dimension (warranty length).

3.5.4. Robustness

We briefly discuss some of the relevant modeling choices and examine the robustness of our main findings with respect to variations in some model constructs.

First, in previous sections, we discussed a number of reasons why we use the length of the powertrain warranty in years as our warranty variable. We performed experiments using alternative definitions, and our main findings remain robust if, for example, we consider miles instead of years, or if we consider a weighted average between powertrain and basic warranty as our warranty variable. Perhaps a more sensitive issue is the definition of our product quality variable. Arguably, there is no perfect way to measure quality. We believe, however, that the metric of problems per vehicle based on the initial quality study by J.D. Power is a reasonable choice. Indeed this metric captures a relatively objective metric of product quality, that has been widely available in the past, and that has had lots of visibility for consumers historically. J.D. Power also publishes data on vehicle dependability, which measures quality problems after 3 years of ownership. We collected some of these data and found a high correlation (0.75) between the vehicle dependability metric and our initial quality variable, which suggest some consistency in both product quality metrics. Furthermore, we think that quality problems in the first three months of ownership may be much more disruptive than quality problems after three years, which is one reason to prefer the initial quality variable used in our study (in addition, considering initial quality instead of vehicle dependability allows us to perform the analysis with a larger sample size). Nevertheless, we estimated the model using the vehicle's 3-year dependability metric to construct the product quality variable, and found results that are largely consistent with our main results (with the exception of the coefficient for the WARRxSQ variable that became marginally non-significant).

Another concern is related to the potential role of brand fixed effects as confounders. As noted earlier, the level of variation in our warranty variable does not allow us to control for brand fixed effects in the model specification, and thus the role of omitted brand fixed effects is certainly a valid concern. We noted, however, that the warranty is not the only brand-level variable in our model, and indeed, we are including product quality, service quality and an indicator of the manufacturer region as brand level variables. We make the following two observations in this regard. First, note that omitted brand-level factors are part of the unobservable term, and accordingly this is part of the reason behind our identification strategy for warranties, i.e. in our formulation we are implicitly accounting for them in the estimation of warranty effects. Second, we collected some additional brand-level variables, like the number of dealers of each brand (obtained from Automotive News's market data books) and the brand age, and estimated our models including these variables in the model specification as a way to partially control for some brand effects not captured in the main formulation, and our main findings remained robust to these variations.

Finally, one could postulate a three-way interaction effect between product quality, service quality and warranty, meaning that the two-way interaction effects postulated by our model could in turn be moderated by the remaining variable as a third factor. We extended our model specification by adding the three-way interaction term WARRxPQxSQ and our main findings remain robust to this variation.

Table 3.5 displays the results obtained for some of the robustness checks discussed in this section. For ease of display, we only show results for the mean utility estimates in each case.

Variable	(1) Estimate	Std.Err.	(2) Estimate	Std.Err.	(3) Estimate	Std.Err.	(4) Estimate	Std.Err.
PRICE	-0.101***	0.012	-0.084***	0.011	-0.077***	0.013	-0.102***	0.012
WARR	0.125*	0.066	0.135*	0.080	0.116	0.073	0.154*	0.083
PQ	1.805***	0.369	0.477***	0.146	0.971***	0.327	1.415***	0.346
SQ	-4.416	3.048	-0.507	3.120	-0.585	2.811	-1.373	3.242
WARR X PQ	-1.856***	0.415	-0.572***	0.171	-0.956***	0.308	-2.052***	0.437
WARR X SQ	9.788***	2.368	6.467	4.098	4.875*	2.701	6.597**	2.916
PQ X SQ	-35.018***	8.100	-20.725***	4.203	-26.718***	6.593	-23.633***	7.015
NDEALERS					0.201***	0.057		
WARRxPQxSQ							-18.446**	7.197

Table 3.5: Selected robustness checks. Notes: All columns display results for the mean utility estimates of the random coefficients model, for the most relevant variables. (1): Powertrain warranty in (10,000's) miles instead of years; (2): PQ in demand specification measured using the problems per vehicle metric of J.D. Power's dependability study instead of the initial quality study; (3): Including NDEALERS (number of franchised dealers, measured in 1000's) in the model specification; (4): Including the 3-way interaction WARRxPQxSQ. *, **, ***, significant at the 0.1, 0.05, 0.01, confidence levels, respectively.

3.6. Conclusions

In this study we formulate and estimate a model to study the impact of service attributes on demand, and the moderating role of product quality in that relationship. We focus on services for the in-warranty period, and characterize the service strategy of a firm by both its warranty length and its after-sales service quality. Our results indicate that both service metrics are complementary, i.e., the better the service quality of a brand, the

higher the marginal effect of offering longer warranties on demand, and vice versa. Thus, these two service attributes reinforce each other. In contrast, no complementarities are observed for service attributes and product quality, and services play, rather, a compensatory role with respect to product quality, i.e., the impact of both service variables on demand increases when product quality decreases. Collectively, our results suggest that competing on services is more effective (in terms of its effect on demand) for firms that have lower product quality, and that a firm that increases its warranty length would benefit most by simultaneously investing in improving its service quality. These findings illustrate that firms would benefit by defining their product and service strategies jointly rather than independently, i.e. they show that the joint consideration of product and service is essential for the development of an effective competitive strategy. In particular, the positioning of a firm with respect to product/service quality dimensions directly influences the marginal effect of its warranty length on consumer demand (Figure 3.5).

Our model thus also provides a tool for managers to evaluate the impact of offering different warranty lengths on consumer demand (for a given positioning in product quality and service quality), which if complemented with actual warranty cost data (which are internally available), would help companies to define optimal warranty levels. More generally, we illustrated how our model can inform decision-making regarding how alternative managerial interventions (e.g. price decrease, quality increase, warranty increase) impact consumer demand, e.g. for an average vehicle our model suggests that reducing the price by 1% is equivalent (in terms of its impact on consumer demand) to

improving product quality by 3%, which is in turn equivalent to increasing the length of the warranty by 9%. We believe that the demand-side estimates derived from our analysis constitute a critical missing component for managerial decision-making in practice, as managers are usually very good at estimating the implied costs of different managerial interventions, while demand effects are much more difficult to isolate.

Our analysis of the service strategy of firms focused exclusively on the in-warranty period. In practice, firm service strategies also include the out-of-warranty period, and thus the definition of the warranty length could have implications in terms of the profits that firms derive from selling extended warranties. If appropriate data becomes available, modeling the interaction between in-warranty and out-of-warranty policies offers a promising avenue for future research. Also, our identification strategy requires the exogeneity of product quality and service quality in the demand specification, which, as noted, is justified by the timing of our model, in which firms and consumers make decisions within a one-year horizon. While appropriate to capture short-term effects, our model does not capture longer-term dynamic aspects involved in firm decision-making and consumer demand. The formulation of a dynamic model that endogenizes long-term effects of firm investment in product quality and service quality is thus a natural -and much more complex-extension of our analysis. Finally, as noted throughout the document, to our knowledge this is the first study to empirically analyze how firm service strategies interact with product quality in a demand model of firm competition. The formulation of similar studies in other manufacturing industries would allow for a

broader understanding of the value of services as part of firms' competitive strategies. We

hope to conduct further research in these areas.

Chapter 4

Product Quality or Service Quality? The Impact of Customer Heterogeneity

4.1. Introduction

The movement toward a service-oriented model in traditional manufacturing industries has opened a number of important challenges and opportunities for firms (Shankar et al. 2009). One of these challenges is the management of quality for products and for those services bundled with the product that support the creation of value through product use. Indeed, when such services are added to the core product offering, it has been noted that improving manufacturing quality is a necessary but not sufficient condition for the attainment of increased, overall, perceived quality (Vargo and Lusch 2004). To the extent that consumers' overall assessment of a brand depends not only on product quality but also on associated service quality, firms can benefit by managing product quality and

service quality jointly, as opposed to independently (Gummesson 1988). Along the same lines, Shankar et al. (2009) noted that if a commoditized product can be paired with consistent and reliable quality of service, then a company can better differentiate its offering to its customers. While the potential benefits may be important, the joint management of product quality and service quality can be challenging, since doing so can add significant complexity to a firm's operations. The challenge, in particular, involves the potential interaction between product and service quality as perceived by consumers. A good illustration of this is discussed by Kindström (2010), who interviewed managers from several manufacturers with well-known premium brands, reporting that "many managers expressed concerns about the possible undermining of their strong brands when moving their product bundle mix towards offering services. Their main concern was that customers would find it difficult to ascertain the quality of the service (especially early on), and that poor-quality service could reflect negatively on the product, alienating customers from both service and product alike". It is therefore critical for firms to understand how consumers perceive quality for their products and supporting services, and the influence of these perceptions on the overall level of customer satisfaction and loyalty. Our study takes a step in the direction of understanding the joint effects of product quality and service quality on consumer decision-making, by studying the impact of customer perceptions of product and service quality on their intentions to recommend the brand, and the moderating role of customer heterogeneity on that relationship.

While several studies addressing consumer intentions and satisfaction have highlighted the role of service quality in service industries (e.g., Parasuraman et al. 1988,

Rust et al. 2002, Mittal et al. 2005), as well as the role of product quality in manufacturing industries (e.g., Churchill and Suprenant 1982, Mitra and Golder 2006, Tellis and Johnson 2007), little research has been done to study the effects of product quality and service quality in a unified manner, and moreover, to study how these effects vary for different groups of consumers. We formulate the product-service quality perceptions map in Figure 4.1 to illustrate potential interaction between each quality dimension. The upper-right (bottom-left) quadrant represents consumers who perceive that both service and product quality are good (bad). Consumers with a high perception of product quality but a low perception of service quality will be located in the bottom-right part of the graph, and those where product quality is perceived to be low but service quality is high, will be located in the upper-left area of the graph.
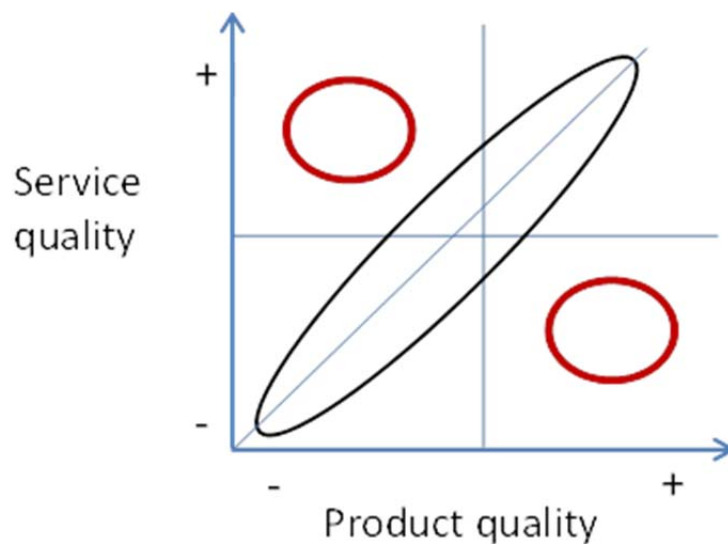

Figure 4.1: Product-service quality perceptions map.

Several questions of interest emerge from this representation of the problem. First, what is the distribution of consumers in the product-service quality perceptions map?

Second, how are perceptions of product quality and service quality jointly associated with consumers' intentions toward the brand? Third, are the effects of product quality and service quality complementary or compensatory? Fourth, how do consumer characteristics (gender, age, education, etc.) moderate these effects? Finally, how can the answers to these questions inform managerial decision-making with regard to both strategic questions concerning product-service design and positioning, and operational/tactical questions associated with customer prioritization?

In this chapter, we provide an empirical analysis that sheds some light on these questions. Our analysis is based on the results of a survey-based study we conducted for a major manufacturer in the consumer electronics industry. The goal of the company was to support a customer centric strategy that used services associated with their high quality products as a means to achieve competitive differentiation, and more precisely, to develop key performance indicators (KPI's) that would support managerial interventions to improve service quality as perceived by their customers. The focus of that study was on the high definition TV's product segment (HDTV). Our goal in this research is to measure how consumer perceptions of quality for both the product and service affect their intentions to recommend the brand. Here, the product is a HDTV, while services refer to the call center for customer service and in-home repair services. Our analysis, which was motivated by our direct collaboration with this company, is based on a sample of consumers that received product support service recently (i.e., 2-5 months previous to the survey date). The results of the analysis suggest that in our application: (i) more than 30% of the consumers had asymmetric perceptions for the quality of products and

103

support services, (ii) product quality and service quality perceptions are jointly associated with consumers' likelihood to recommend the brand, (iii) on average, there is a complementary relationship between product quality and service quality perceptions, i.e., the higher the consumer perceptions of product quality, the higher the marginal impact of higher service quality perceptions on the person's likelihood to recommend the brand, (iv) customer heterogeneity plays an important moderating role. In particular, service quality is more important for women and for higher income segments, relative to men and lower income segments. In short, our analysis shows that:

- Customers' perceptions of both product quality and service quality are important drivers of their likelihood to recommend the brand.

- There is a complementary relationship between product quality and service quality perceptions in terms of customers' likelihood to recommend the brand.

- Different segments of customers are more sensitive to product quality while others are more sensitive to service quality.

We elaborate on the consequences of our findings for managerial decision-making, specifically for product-service differentiation.

4.2. Related Literature

The topics of product quality and service quality management and measurement have been analyzed extensively in the operations management and marketing literatures. There is a considerable history of research in the operations management field devoted to the measurement and management of product quality (e.g., Garvin 1987, Juran 1992), while

there has been extensive research, primarily in marketing, on the development of tools for the measurement of service quality (e.g., Parasuraman et al. 1985, Cronin and Taylor 1992). However, to date, both topics have for the most part been treated independently, and little empirical evidence has considered product quality and service quality jointly. Exceptions include some studies in the automobile industry (Oliver and Swan 1989, Archer and Wesolowsky 1996, Mittal et al. 1999, Devaraj et al. 2001, Guajardo et al. 2012), the mobile phone industry (Chai et al. 2009), the IPod/ITunes bundle (Chan et al. 2008), and the lumber industry (Hansen and Bush 1999). Here we briefly discuss some of these studies, and point out the distinctive features of our research.

Oliver and Swan (1989) find that both satisfaction with the dealer and satisfaction with the product were related to overall customer satisfaction. Acher and Wesolowsky (1996) employ a critical incidents technique to conclude that negative service incidents can affect consumer intentions toward both the dealer and the OEM, and that the effect is less drastic in the case of negative incidents with the product. Mittal et al. (1999) shows that service quality is initially more important than product quality in terms of consumer intentions toward the OEM, and that later in the ownership cycle that relationship is reversed (product quality becomes more relevant than service quality both for intentions toward the OEM and the dealer). Devaraj et al. (2001) find that there is a strong association between product quality and customer loyalty, but that there is only partial support for the effect of service quality on loyalty. Chai et al. (2009) study the Chinese mobile industry, and find spillover effects for customer satisfaction with the OEM and the service provider. Chan et al. (2008) found no evidence of impact of service

satisfaction (ITunes) on overall customer loyalty, and no evidence for an effect of product quality on customer satisfaction; their sample size of about 40 students who experienced after-sales services, however, may be a limitation associated with these unintuitive results. Hansen and Bush (1999) found that product quality was more relevant than service quality in explaining overall satisfaction in an application in the lumber industry. Collectively, for the most part, these studies show that there is indeed a potential benefit for a manufacturer to jointly consider the effects of both product quality and service quality on consumers' intentions. This conclusion is also supported by the results of our analysis.

Our research, however, differs from these studies in several aspects. First, unlike existing studies, we analyze the moderating role of customer characteristics on the relative association between product-service quality and customer intentions, and show that the effects of product and service quality are significantly heterogeneous in the population. This study is, to our knowledge, the first systematic empirical study that highlights the role of customer heterogeneity in the product-service quality context. This, in turn, has important consequences for managerial decision-making regarding product-service differentiation, as we discuss in detail in a later section. Second, we discuss and document complementarities between product and service quality perceptions. This is a feature that has been mostly omitted in existing studies. Indeed, to our knowledge, the only other study that considers complementary/substitution effects between product quality and service quality is our study of warranties and service quality in the automobile industry (Guajardo et al. 2012), which deals with a different problem, i.e.,

investigating the impact of product quality and service attributes on new car sales. As our analysis will show, we arrive at different results in this study of customer loyalty in the consumer electronics industry, i.e., the conclusions from our previous research do not extend to this new setting. Finally, we analyze a domain where, to our knowledge, this problem has not been studied before, i.e., HDTV's in the consumer electronics industry. This domain differs from the existing studies in that the OEM has high influence over the service delivery outcome. Indeed, in our domain the OEM directly controls the call center for customer service, and also has incentive programs in place in order to assure that repair services performed by third parties are completed according to desired target service levels. Managerial action (e.g., differentiated service design) based on our results is thus suggested, and can be implemented directly by the OEM.

4.3. Research Setting and Data

Our research is motivated by our direct interaction with a major manufacturer in the consumer electronics industry. Manufacturers in this industry offer products such as televisions, audio and video equipment for home and cars (CD players, DVD players, etc.), image equipment (cameras), and amplification equipment (musical instruments, public address systems, etc.). Industry-wide market shares, as a percentage of total sales, among these product segments are as follows (IBISWorld 2010): Television 57.1%, audio equipment 15.6%, car audio and video equipment 12.7%, image recording/playback 4.6%, and imaging 10.0%. According to recent statistics, the biggest players in the industry are Sony, Samsung, Panasonic and LG, which have a combined

market share of about 37% (IBISWorld 2010). The low concentration and high level of competition in this industry has led to a high degree of parity with respect to product price and performance and thus it is natural that firms look for ways to differentiate themselves from their competition. Services that add value to core product offerings are one possible way to achieve that goal (Grönroos 1990). Indeed, the company we worked with has recently adopted this approach as a part of their overall competitive strategy.

Our study focused on the company's U.S. market for the high definition television segment. As indicated above, the television segment is the largest segment in the industry in terms of market share, i.e., our study focuses on the most important product line in terms of sales. Supporting services for customers under warranty consist of interactions with the call center for customer service (owned and operated by the OEM) and the in-home repair services (provided by exclusive service providers or by authorized service centers). Thus, in our context, service quality refers to customers' perceptions of overall quality for both call center interactions and in-home repair services.

Collaborating with the company, we designed and analyzed a survey, intended to capture consumer perceptions of different variables of interest[14]. The survey was run by a third-party market research company in November 2010, using a web-based interface. Due to the nature of the phenomenon of interest (how consumer perceptions of service quality and product quality are associated with their intentions toward the brand), the target group of customers for the survey was defined as HDTV owners of the brand, who

---

[14] Another independent survey was run to capture consumer expectations for a sample of consumers who did not experience service interaction with the brand in the last 2 years. We make reference to that survey where appropriate.

recently (July-September 2010) experienced a service interaction with the company. The survey obtained data for a random sample of such customers. Among other features, the survey captured consumer perceptions about product quality and service quality, and their likelihood to recommend the brand. A 7-point scale was used, and consumers were asked to rank their evaluation of the quality of the product and supporting services (1: Poor,…, 7: Excellent), as well as their likelihood to recommend the brand based on their overall experience (1:Not at all likely,…,7:Extremely likely). The survey also collected information on customer characteristics such as gender, age, education and income (household income before taxes). Overall, the sample consisted of 345 HDTV owners of the brand who recently experienced a service interaction with the company, with valid answers for their perceptions of product quality, service quality, and likelihood to recommend the brand. Table 4.1 provides some descriptive statistics for various sub-groups in the sample, as well as descriptive statistics and pairwise correlations for the variables of interest.

The statistics in Table 4.1 indicate some apparent differences for distinct groups in the sample. For example, average scores for the quality perceptions and the likelihood to recommend the brand are consistently greater for men than for women, as well as for older rather than for younger customers. Interestingly, the relative magnitude of the association between product-service quality and likelihood to recommend the brand seems to also vary across the different groups, i.e., for women, the correlation between service quality and likelihood to recommend is bigger than the correlation between product quality and likelihood to recommend, while for men, we observe exactly the

opposite. A similar observation can be made for income. In particular, for the high income segment, the correlation between service quality and likelihood to recommend is bigger than the correlation that was observed for product quality. These results however are reversed for the lower income segment. Naturally, these observations are based solely on univariate statistics and pairwise correlations, i.e., they do not consider any confounding factors and are thus only an indication that differences across customer groups do exist. To more accurately disentangle the relevant associations, a multivariate approach that simultaneously evaluates these characteristics is needed. This analysis is the focus of the next section.

Customer characteristic	Frequency (%)[15]	Average Scores (1-7 scale)			Pairwise correlations		
		Product Quality (PQ)	Service Quality (SQ)	Likelihood to Recommend (LtR)	PQ,LtR	SQ,LtR	PQ,SQ
Gender							
Men	80.3%	5.78	5.35	5.17	0.82	0.78	0.60
Women	19.7%	4.94	4.76	4.46	0.75	0.84	0.58
Age							
Age1 (15-34)	18.2%	4.95	4.74	4.87	0.68	0.71	0.46
Age 2 (35-54)	36.7%	4.78	4.87	4.61	0.78	0.78	0.56
Age 3(55 or more)	45.1%	5.48	5.73	5.46	0.85	0.83	0.63
Education							
No college	33.7%	5.40	5.30	5.25	0.82	0.83	0.64
College or more	66.3%	5.00	5.25	5.03	0.81	0.78	0.58
Income							
less than $100,000	63.3%	5.16	5.19	5.01	0.85	0.78	0.67
$100,000 or more	36.7%	4.98	5.31	5.06	0.74	0.83	0.52
Full Sample	100%	5.13	5.23	5.03	0.80	0.80	0.59

Table 4.1: Sample description - statistics and correlations

---

[15] For each customer characteristic, only valid responses are considered in the calculations. Sample size is each case is as follows: Gender=345 obs., Age=335 obs., Education=312 obs., Income=259 obs.

The graph in Figure 4.2 illustrates the allocation of consumers in our sample in the product-service quality perceptions map. For a better visual display, we categorize the answers in the 7-point scale for the perceptions of product and service quality as low (1,2,3), moderate (4,5), and high (6,7). The numbers inside the bars represent the fraction of consumers in each group as a percentage of respondents conditional on their ratings of product quality (without brackets), and as percentages relative to the full sample (in brackets). For each group, the average score for the likelihood to recommend the brand question is displayed in square brackets, to the right of the group. For example, 60.5% of the customers who rated product quality as low also rated service quality as low (this group represents 13.3% of the full sample), and the average score for the likelihood to recommend the brand for that group is 1.5. As can be observed in the graph, an important group of consumers (68.4%) are located on the "diagonal" of the graph. However, 31.6% of the consumers have different perceptions of product and service quality, even with the aggregation of the data into categories that we used. A second observation is that, as expected, important differences are noted in the average score for the likelihood to recommend the brand for each group. In particular, the scores obtained for the extreme groups ({PQ,SQ}={LOW,LOW},{HIGH,HIGH}) of 1.5 and 6.8, respectively, suggest that consumers' perceptions of product and service quality may capture an important part of their brand loyalty scores.
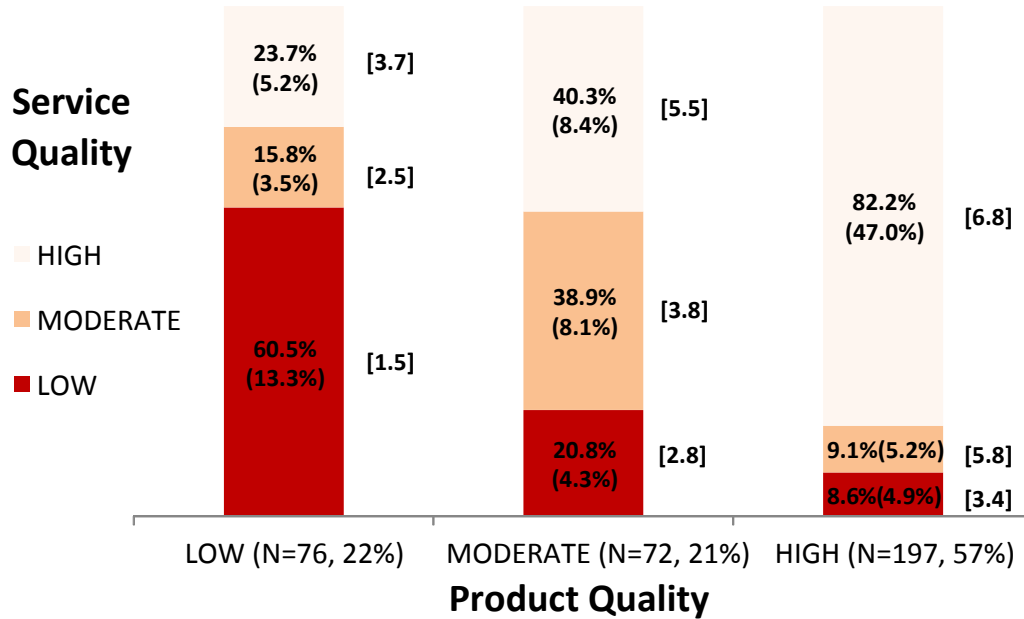
Figure 4.2: Product-service quality perceptions map - Application in HDTV's

4.4. Model and Results

4.4.1. Hypotheses and modeling framework

Recall that our main goal is to analyze the joint association between the consumers' perceptions for product and service quality with their loyalty toward the brand, the moderating role of customer heterogeneity on that association, and the existence of complementary/substitution effects. We use the likelihood to recommend the brand as a proxy for consumer loyalty. The likelihood to recommend is recognized as one of the key metrics to measure consumer intentions (Farris et al. 2006), and together with repurchase intent and customer satisfaction, have been used widely to capture consumer loyalty in marketing research. Moreover, existing research has shown that the likelihood to recommend is highly correlated with actual repurchase behavior, providing empirical

support for the use of this metric as a dependent variable (Zeithaml et al. 1996). As noted in the previous section, in our case this variable is measured using a 7-point scale, and is the dependent variable in our analysis.

As also noted previously, several studies have shown that consumer intentions and satisfaction are associated with service quality in service industries (e.g., Parasuraman et al. 1988, Rust et al. 2002, Mittal et al. 2005), as well as with product quality in manufacturing industries (e.g., Churchill and Suprenant 1982, Mitra and Golder 2006, Tellis and Johnson 2007). There are not many studies that integrate both aspects, but for the most part, the ones that do (Oliver and Swan 1989, Archer and Wesolowsky 1996, Mittal et al. 1999, Hansen and Bush 1999, Devaraj et al. 2001, Chan et al. 2008, and Chai et al. 2009) have found support for the relationship between both product quality perceptions and service quality perceptions with customer intentions. We would expect the same in our case, and based on the existing evidence we thus formulate hypothesis 1.

H1: Product quality perceptions and service quality perceptions are positively associated with consumers' likelihood to recommend the brand

While research in marketing has explored the role of customer heterogeneity, relevant applications in operations management have rarely integrated customer heterogeneity with operations-related variables, such as product quality and service quality. In particular, none of the papers analyzing product-service quality in a unified framework noted earlier[16] have studied the moderating role of customer heterogeneity in influencing

---

16 Oliver and Swan 1989, Archer and Wesolowsky 1996, Mittal et al. 1999, Hansen and Bush 1999, Devaraj et al. 2001, Chan et al. 2008, and Chai et al. 2009.

the relationship between product-service quality perceptions and consumer intentions toward a brand. Previous research, however, has illustrated the role of customer heterogeneity in different settings. Studies on consumer satisfaction, in particular, have found systematic differences in the way that different groups of consumers respond to satisfaction ratings (Bryant and Cha 1996) and behave in terms of the satisfaction-repurchase link (Mittal and Kamakura 2001). Thus, it is important to control for customer characteristics in the model specification. While in a different context than ours, a few recent studies in operations have highlighted the role of heterogeneous effects of service time in consumer decision-making, including Lu et al. (2012), who report heterogeneous effects of service time on consumer demand in a retail store (ham purchase at the deli section of a super-center), and Campbell and Frei (2011), who report significant variation in consumers' sensitivity to service wait times in the banking industry. More generally, in product-service settings, the notion of a "context-specific" value creation process (Ng et al. 2012), in which customer value depends on the context of usage, which by definition may vary for different customers, has been acknowledged. Thus, in our case, we would expect significant heterogeneity for product quality and service quality effects across the population, which we reflect in the general formulation of hypothesis 2. There are, however, some theories and existing research that would help in formulating a more detailed version of the general statement in hypothesis 2, if we look individually at consumer characteristics such as gender, age, education, and income. We discuss some of them next.

In terms of gender, it has been documented by some researchers that men tend to be more technology-oriented than women (e.g., Frankel 1990, Schumacher and Morahan-Martin 2001), and that women are more anxious than men when dealing with technology (e.g., Franz and Robey 1986, Colley et al. 1994). While their studies involve IT and computers, these arguments should extend to our setting of HDTV's, given the technological nature of the product. If women feel more "uncomfortable" than men when facing a HDTV malfunction, they may be more prone to attempt to fix the problem by calling the company. In that case, they may be more likely to appreciate a good resolution of the problem than men, which in turn may cause women to value good service quality higher. Our data contains some evidence aligned with this explanation. The survey asked how frequently a respondent had a service interaction with the company in the last two years, and the data show in fact that women had more frequent interactions. In particular the average number of times a person interacted with the call center is 4.1 times for women, and 2.7 times for men (the difference is significant, p-value=0.0018 for the mean comparison test). Similarly, the average number of times a person reports to have had in-home repair for their HDTV is 1.6 for women and 1.3 for men (p-value=0.0271). It is very unlikely that the HDTV's purchased by women are consistently worse in terms of product reliability than the ones purchased by men, so in light of previous studies it seems more plausible to attribute this evidence to behavioral differences by gender. A related argument, also supported by our data on frequency of interaction, is that, regardless of whether or not they are less technology oriented than men, women are more directly involved in the process of calling the company for service and are present at the

115

moment of in-home repairs, i.e., role theory (e.g., Lips 1994). Finally, an alternative argument to hypothesize gender effects is differences in risk aversion. Indeed, several empirical studies have found that women are more risk averse than men, for products such as insurance (Feldman et al. 1989, Halek and Eisenhauer 2001), retirement savings plans (Sunden and Surette 1998), and the purchase of extended warranties (Chen et al. 2009). If women are more risk averse than men, they would care more about the "peace of mind" aspect of choosing a particular brand, which is partially related to the ability of the company to promptly and effectively resolve product issues, which in turn are drivers of service quality. In other words, even if the product fails, the fact that a company acts adequately to provide supporting services to resolve the problem may be more valuable for women. We have some indication that suggests that women are more risk averse than men in our application. Similarly to the SERVQUAL framework (Parasuraman et al.1988), in our survey we asked consumers to allocate 100 points according to the relevance they assign to seven different dimensions of the quality of service: reliability, responsiveness, assurance, empathy, tangibles, convenience, and warranty. We note that the warranty dimension was defined in the survey instrument as "the company offers you the product warranty terms and coverage options you need"[17]. On average, women allocated 47% more points than men to the importance of the warranty dimension, and the hypothesis that the mean for women is greater than for men for that question is supported (p-value=0.0021). Similarly, we asked about the length of the warranty that

---

[17] Convenience was defined as "the company offers you convenient hours of operations and appointment scheduling times (if necessary)". For the definition of the rest of service quality dimensions, see Parasuraman et al.1988.

consumers expect for their HDTV's and found that expected warranty duration for women is roughly 10% greater than for men on average, and the difference in the expected length of warranty for parts is significant (p-value=0.0248), and partially significant in the case of warranty length labor coverage (p-value=0.0505). This evidence is consistent with previous studies of gender differences in risk aversion in other settings, and suggests that women are indeed more risk averse than men in our case. This in turn may relate to their emphasis on service quality in comparison to men, as discussed above. This discussion supports the formulation of hypothesis 2a.

With respect to age, previous studies have reported that older groups exhibit greater brand loyalty due to greater experience in using a brand's products and due to their tendency to do less search (Ratchford 2001). Devaraj et al. (2001) also find that service satisfaction is positively correlated with age, and Morris et al. (2005) reported that women and men are more similar with respect to technology orientation and knowledge for younger segments of the population, but important differences emerge for older population segments. In other words, the technology gap across gender increases with age. This is consistent with expecting higher service sensitivity for older segments of the population. In the case of education, Propper (1995) reports that employed customers are more sensitive to service time even after controlling for income effects, which partially reflects an education effect if employment and education are correlated. Similar to these studies, in our case we hypothesize that the impact of service quality perceptions is thus more relevant for older and for more educated segments. This is captured in hypotheses 2b and 2c.

117

Finally, with respect to income, studies in the economics literature have shown that the disutility for waiting time is increasing in consumers' income (Propper 1995). This finding is consistent with higher income segments having a higher opportunity cost for time. Indeed, in line with this idea, some studies have reported that higher income customers are more sensitive to service time/capacity in e.g. gasoline retailing (Png and Reitman 1994) and banking (Campbell and Frei 2011). Using this logic, in our context it may be hypothesized that the cost of a "disruption" (HDTV malfunction, based on the time and actions involved in the service and repair process) to the customer is higher for higher income segments. As a result, if the product support issues are resolved effectively and promptly, higher income customers would place more value on such performance (i.e., provision of good service quality) relative to lower income segments, due to heterogeneity in the opportunity costs of customer time for dealing with a repair. Based on these arguments, we formulate hypothesis 2d.

H2: Customer characteristics moderate the relative association between product quality and service quality perceptions with the likelihood to recommend the brand. In particular, service quality perceptions are more important for:

a). Women than for men,

b). Older than for younger segments,

c). More educated than for less educated segments, and

d). Higher income than for lower income segments

Finally, service quality and product quality perceptions can act as complements, substitutes or independently with respect to the likelihood to recommend the brand. In a compensatory decision process (e.g. Dieckmann et al. 2009), strengths along one or more dimensions of product or service quality can compensate for weaknesses along others. This is in contrast to the case of non-compensatory processes, in which no compensation is possible if certain attributes of a product or service are weak, even if it possesses strengths along other dimensions. The role of compensatory effects on consumer decision-making would provide a basis for characterizing perceptions of service quality as substitutes for product quality perceptions. Indeed, consistent with this argument, our recent study in the automobile industry (Guajardo et al. 2012) showed that service quality act as substitute for product quality in the consumer demand function. On the other hand, we would expect service quality and product quality perceptions to be complements if the dominating mechanism by which they affect consumer intention is through their impact on brand image. Indeed, given that the dependent variable explicitly considers brand, this is a plausible scenario in this application in the consumer electronics industry. If this is the case (complementarity hypothesis), we would observe that, for persons with higher perceptions of product quality, there would be a higher the impact of better perception of service quality on their likelihood to recommend the brand. Alternatively, these attributes may exhibit independent effects on consumer decision-making. In this scenario of competing theories, whether service quality perceptions act as a complement, a substitute, or independently of product quality perceptions for the likelihood to recommend the

119

brand, is ultimately an empirical question, and we thus formulate hypothesis 3 in the null form.

H3: The effects of product quality and service quality perceptions on the likelihood to recommend the brand are independent

We investigate these three hypotheses by formulating a (nonparametric bootstrap) regression framework, as described in Eq. (4.1). The subscript i indexes consumers, LtR_i denotes the likelihood to recommend the brand, Q_i the perceptions of quality (for both product and service), X_i customer characteristics (gender, age, education, income), and $Q_i * X_i$ element-by-element interactions between quality and consumer characteristics.

$$LtR_i = \alpha + Q_i\beta + X_i\gamma + Q_i * X_i\delta + \varepsilon_i \qquad (4.1)$$

The model in equation 1 allows for testing of hypotheses 1 and 2. To test hypothesis 3, we enhance the model specification by including interaction between the elements of Q_i, i.e., we include the two-way interaction term between service quality and product quality perceptions as another regressor. We note that the inclusion of the variables X_i and interactions $Q_i * X_i$ account for observed customer heterogeneity. With respect to unobserved heterogeneity, the term ε_i reflects factors such as the previous experience of the consumer with HDTV's. Given the nature and scale of the data, the usual normality assumption is implausible and thus we do not rely on it for testing purposes. For this reason, and taking advantage of the independence of our observations, we employ non-

parametric bootstrap[18], where hypothesis testing is based on bootstrap standard errors and confidence intervals, and thus depends on the empirical distribution obtained by sampling with replacement from our data. As a result, we do not need to make any parametric assumptions for hypothesis testing purposes. Finally, we "mean center" the variables PQ and SQ in our regressions, i.e., we subtract the mean from each individual observation.

4.4.2. Results

Table 4.2 displays the obtained results. We include results for model specifications that gradually incorporate the interaction effects, as well as for a model with the full set of interactions. First, as expected, the hypothesis that product quality and service quality perceptions are associated with the likelihood to recommend the brand is supported (all specifications). The results in Table 4.2 suggest that gender and income moderate the relative importance of the associations between product-service quality and the likelihood to recommend the brand, providing support for hypotheses 2a and 2d. Service quality exhibits a greater association with the likelihood to recommend the brand for women than for men, and for high income customers than for low income customers. Conversely, the association between product quality and likelihood to recommend the brand is higher for men than it is for women, and for low income customers than for high income customers. These effects are present in both the single-variable interaction models (2 and 5) as well as in the model with the full set of interactions (6 and 8). We also note that the models

---

[18] Horowitz 2001, and Cameron and Trivedi 2009 (Ch. 13), contain informative discussions on the use of bootstrapping methods for standard errors, hypothesis testing, and other applications.

suggest that there are no important differences according to age and education, which

rules out hypotheses 2b and 2c.

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PQ	0.602***	0.661***	0.540***	0.632***	0.707***	0.642***	0.638***	0.733***
	(0.064)	(0.073)	(0.149)	(0.119)	(0.078)	(0.167)	(0.064)	(0.166)
SQ	0.511***	0.454***	0.483***	0.532***	0.396***	0.434***	0.553***	0.429***
	(0.059)	(0.067)	(0.109)	(0.108)	(0.073)	(0.120)	(0.058)	(0.120)
Women	-0.102	-0.009	-0.101	-0.101	-0.112	-0.027	-0.128	-0.069
	(0.153)	(0.145)	(0.157)	(0.155)	(0.153)	(0.148)	(0.149)	(0.144)
Age2	-0.340*	-0.337	-0.310	-0.337	-0.271	-0.244	-0.410**	-0.307
	(0.204)	(0.198)	(0.195)	(0.203)	(0.202)	(0.190)	(0.199)	(0.185)
Age3	-0.215	-0.220	-0.191	-0.213	-0.130	-0.126	-0.324*	-0.214
	(0.190)	(0.185)	(0.192)	(0.192)	(0.187)	(0.192)	(0.186)	(0.183)
College	0.0472	0.0621	0.035	0.0543	0.0315	0.049	0.058	0.069
	(0.128)	(0.123)	(0.124)	(0.132)	(0.123)	(0.130)	(0.122)	(0.124)
Income_high	0.116	0.130	0.113	0.114	0.103	0.099	0.150	0.118
	(0.135)	(0.130)	(0.135)	(0.133)	(0.128)	(0.128)	(0.129)	(0.126)
PQ*Women		-0.289**				-0.304**		-0.191
		(0.133)				(0.143)		(0.135)
SQ*Women		0.261**				0.281**		0.230*
		(0.120)				(0.130)		(0.126)
PQ*Age2			0.044			0.180		0.168
			(0.179)			(0.165)		(0.157)
PQ*Age3			0.098			0.138		0.053
			(0.185)			(0.171)		(0.166)
SQ*Age2			0.069			-0.065		-0.053
			(0.144)			(0.132)		(0.127)
SQ*Age3			0.004			-0.046		0.003
			(0.151)			(0.137)		(0.132)
PQ*College				-0.039		0.016		-0.031
				(0.141)		(0.119)		(0.113)
SQ*College				-0.036		-0.118		-0.074
				(0.127)		(0.103)		(0.102)
PQ*Income_high					-0.214*	-0.247**		-0.261**
					(0.125)	(0.132)		(0.126)
SQ*Income_high					0.260**	0.303***		0.307***
					(0.111)	(0.112)		(0.106)
PQ*SQ							0.057***	0.0558***
							(0.020)	(0.018)
Constant	5.206***	5.198***	5.193***	5.196***	5.143***	5.133***	5.118***	5.031***
	(0.192)	(0.190)	(0.192)	(0.199)	(0.190)	(0.197)	(0.197)	(0.197)
Observations	258	258	258	258	258	258	258	258
Adj. R-squared	0.804	0.810	0.802	0.803	0.812	0.818	0.814	0.825

Table 4.2: Regression of LtR on PQ, SQ, and customer characteristics.
Note: Bootstrap standard errors in parentheses, hypothesis testing based on bootstrap confidence intervals: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

The results displayed in columns 7-8 in Table 4.2, consider enhanced model

specifications including an interaction term between product and service quality

perceptions (PQ*SQ). These results suggest that, on average, there is a complementary

relationship between product and service quality, i.e., the greater the product quality, the greater the effect of greater service quality. In other words, a good perception of service quality will only strengthen the likelihood to recommend the brand by a person that already had a good perception of product quality, relatively more than it would do for a person with a bad perception of product quality. And a bad perception of service quality will make it even more unlikely for a person who had a bad perception of product quality to recommend the brand, relative to a person that had a good perception of product quality.

While the results in Table 4.2 suggest that the influence of product and service quality vary by gender and income, we can further analyze these results by testing a linear hypothesis which actually compares the relative influence of product quality and service quality for each segment. Table 4.3 shows results of hypothesis tests that compared the coefficients obtained for the difference (SQ-PQ) for each group. As an example, in the case of model 2, for women the test compares the coefficients for SQ-PQ=(0.454+0.261)-(0.661+-0.289)=0.343, and for men the difference is SQ-PQ=(0.454)-(0.661)=-0.207.

Segment #	Model	Segment Gender	Age	College	Income	(SQ - PQ) Estimate	Std. Error
1	2	Woman	n.a.	n.a.	n.a.	0.343*	0.201
2	2	Man	n.a.	n.a.	n.a.	-0.207	0.135
3	5	n.a	n.a.	n.a.	>$100,000	0.163	0.171
4	5	n.a	n.a.	n.a.	<$100,000	-0.310**	0.145
5	6	Woman	35-54	No college	<$100,000	0.132	0.260
6	6	Woman	55 or more	No college	<$100,000	0.193	0.325
7	6	Woman	35-54	College	<$100,000	-0.002	0.273
8	6	Woman	55 or more	College	<$100,000	0.059	0.321
9	6	Woman	35-54	College	>$100,000	0.548**	0.285
10	6	Man	15-34	No college	<$100,000	-0.208	0.270

11	6	Man	35-54	No college	<$100,000	-0.453**	0.225
12	6	Man	55 or more	No college	<$100,000	-0.392*	0.224
13	6	Man	35-54	No college	>$100,000	0.097	0.273
14	6	Man	55 or more	No college	>$100,000	0.158	0.278
15	6	Man	15-34	College	<$100,000	-0.341	0.258
16	6	Man	35-54	College	<$100,000	-0.586**	0.247
17	6	Man	55 or more	College	<$100,000	-0.525**	0.227
18	6	Man	15-34	College	>$100,000	0.208	0.269
19	6	Man	35-54	College	>$100,000	-0.037	0.250
20	6	Man	55 or more	College	>$100,000	0.024	0.236

Table 4.3: Test for the coefficient of (SQ-PQ).
Note: Non-parametric hypothesis testing based on bootstrap standard errors and confidence intervals. n.a.=not applicable. In the case of model 6, we report results only for segments for which we have at least 5 observations in the sample. Significance codes based on bootstrap confidence intervals: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

The results for models 2 and 5 show not only that the influence of PQ and SQ varies by gender and income, but also that the relative influence of both quality attributes varies significantly. Surprisingly, our results indicate that women value service quality more than product quality (model 2). Similarly, our results indicate that in the low income segment, product quality is valued more than service quality (model 5). The results from model 6 allow a further break down of the results for different groups. This greater granularity, however, comes at the cost of having fewer observations for each group, which in particular is expected to have an impact on significance levels obtained for the women segment (due to the smaller sample size in that case). Table 4.3 reports results for the tests of groups for which we have at least five observations. We note that in almost all cases involving women, the coefficient of service quality is greater than the coefficient of product quality, although the tests reflect that the difference is significant only for the group of women in the 35-54 age segment, with college degree, and household income greater than $100,000. In the case of men, results reflect that the weight that men place

on product quality is greater than the weight on service quality when income is less than $100,000. For the higher income segment, there are no important differences in the value that men place on product quality and service quality. Remarkably, differences for men are thus only significant when income is lower than $100,000 (groups 11, 12, 16 and 17). The results of models 2 and 5 are thus consistent with the results obtained with model 6. This can be considered as complementary evidence to support hypothesis 2, i.e., that sensitivity to service quality perceptions versus product quality perceptions vary significantly with customer characteristics, in particular, our results support important differences by gender and income.

4.4.3. Limitations of the study

The study presented here is subject to several limitations. First, the analysis is based on observational data from a survey, and our results do not imply causality. Rather, we document observed associations in the data. Given the limited work that has been done in this research stream, we think that the results suggested by our research are relevant and should stimulate further work in the area, which could shed new light in terms of the causal effects that may be present. In particular, this study can serve as motivation for the development of controlled experiments designed to gain insights about causal effects. Second, as we have noted, the analysis is conditional on the fact that some product malfunction has occurred. Recall that the sample considered only customers that have had a recent service interaction with the company. As we discussed, however, given that we want to capture consumers' perceptions about service quality, this sample selection issue is inevitably related to the nature of the analysis we are pursuing. It is nonetheless, a

caveat in terms of interpreting our results, i.e., all reported effects are for the population of consumers that experienced at least one service encounter recently. For example, our results indicate that product quality and service quality perceptions have, on average, similar effects on the likelihood to recommend in the full sample. One would expect that for a person that did not experience any product issue, the association between product quality and likelihood to recommend would be higher than the one in our analysis. Perceptions about service quality in that case, though, would be meaningless since those customers would not have had a service encounter. Finally, there are unobserved variables that can moderate the observed effects, e.g., how experienced is the customer with the product, price differences, and factors other than perception of product quality and service quality that could affect brand reputation and the likelihood to recommend (e.g., exposure to advertising). We noted, however, that product quality and service quality have a significant association with the consumers' likelihood to recommend, as reflected in the high correlations (Table 4.1, Figure 4.2) and the levels of explained variance obtained in the analysis (Table 4.2). To the extent that the unobserved variables are not systematically related with the included variables, they should not affect our results.

4.5. Managerial Insights

Our analysis suggests that perceptions of service quality and product quality are jointly associated with the customers' likelihood to recommend the brand, and that some groups of customers are much more sensitive to services (e.g., women, high income) than

others. This suggests that a differentiation strategy, that provides different segments with different value propositions in terms of product-service, could offer an effective way of efficiently addressing customers' needs for OEM's. We discuss managerial implications of our findings in terms of this product-service differentiation strategy, and also implications for operating such a strategy in practice.

4.5.1. Product-service differentiation strategy

Our results show that there is significant heterogeneity across the population of customers regarding the importance of product quality and service quality perceptions. In particular, our results show that women and high-income segments are much more sensitive to service quality than men and low-income segments (customer loyalty for the latter group is mostly driven by product quality perceptions exclusively). These results suggest that OEM's could benefit by offering differentiated product-service bundles to their customer base. Many industry examples exist where a company specializes its product-service offerings to different segments of the population (e.g., auto repair services orientated to women, exclusive restaurants orientated to high income customers, financial services targeted to high wealth individuals, etc.), or where a company offer different service levels to different groups at different prices (e.g., first-class vs. economy-class airline tickets, extended service plans for computers, etc.). The challenge of designing differentiated product-service bundles for OEM's in the consumer electronics industry is similar, but is more difficult to implement given the huge customer base and wide range of products offered by the companies. How then can an OEM achieve, or take advantage of, product-service differentiation? We discuss some approaches below.

A first way to achieve product-service differentiation is through the creation of multiple brands that target different customer segments. In the automobile industry, this is the strategy that General Motors pioneered historically and more recently pursued by creating Saturn. Indeed, Saturn was launched with the goal of having a brand that excels at customer service, and consequently its service-supply-chain was very carefully designed in order to achieve that goal (Cohen et al. 2000). Moreover, while the brand was characterized as having best-in-class service quality, its product quality was not at that level, achieving a positioning in terms of product quality and service quality that was not pursued by many other brands in the industry (Guajardo et al. 2012). While operating multiple brands certainly presents additional challenges at different levels (operational, marketing, organizational, etc.), it may offer an effective way to offer product-service bundles that attract different segments of the population, thus allowing an OEM to increase market share.

A second way is through selling extended service contracts (also referred as extended warranties). OEM's often offer a menu of extended service plans at different prices, beyond the basic manufacturer warranty that comes bundled with the product, which allows consumers to self-select into the service plan, according to their willingness to pay and depending on how sensitive to services they are. While customary in many industries (e.g., recent examples are Apple Care and Toyota Care plans), this approach is more difficult to implement by an OEM in the consumer electronics industry, because extended service contracts provide a very profitable business for retailers (see Hesse 2012 for a recent related discussion). Indeed, some reports have indicated that profits from service

contracts accounted for 50% of Best Buy's operating income, and that the margins they provide are around 18 times greater than the margins derived from the goods themselves (Berner 2004). A potential supply-chain conflict between OEM's and retailers would thus emerge if OEM's would sell their own extended service plans, as it would directly affect a profitable business for retailers. In such a scenario it is possible that retailers would. prefer to sell HDTV's from OEM's that do not compete in the extended service market. A further obstacle is the fact that a large portion of the after-sales support is outsourced to third party providers in this industry. Therefore, while intuitively appealing, this approach to product-service differentiation by an OEM would require constructing appropriate agreements throughout the supply-chain, e.g. by defining acceptable incentive plans or compensations mechanisms for retailers (e.g., promise of exclusivity for new product launches), in order to be embraced and implemented in the retail channel. Of course, use of direct sales channels by OEM's, e.g. by selling products to final consumers through their own webpages, provides a mechanism for OEM's to offer extended service contracts much more easily.

Another way of achieving product-service differentiation is through the definition of varying service-level targets for different product and customer combinations, i.e., by setting differentiated service-level standards depending on the service sensitivity of customers for particular products. In general terms, the set of product-customers covered by an OEM could be divided into different groups, where each of these groups is subject to different service-level targets that are set depending on the criticality of the service dimension for each group. Similar strategies have been analyzed elsewhere, e.g. in the

context of parts inventory management in military applications by Deshpande et al. (2003a), who show that a service differentiation strategy can be an effective way for allocating inventory investment by providing higher service for critical parts, at the expense of accepting lower service levels for parts with less importance. This is a product-based priority service parts system. In our case, our results illustrate the importance of accounting for customer heterogeneity in preferences for service. For the sake of illustration, in the case of HDTV's examined in this research, a differentiated service strategy could be reflected if three different groups were defined, e.g. a platinum segment consisting of women purchasing HDTV's costing $1,500 and above, a gold segment for men purchasing HDTV's costing $1,500 and above, and a silver segment that considers all customers that bought HDTV's costing less than $1,500. Different service-level targets could be defined for each of these segments for both call center and in-home repair services, e.g., 99% first time resolution rate target for the platinum segment, 95% for the gold segment, and 90% for the silver segment. In short, based on our results, we propose that a customer-product priority system for the delivery of support services could provide a significant benefit. Finally, we note that different segments potentially will care differently about various aspects of the service delivery process. Thus, defining service standards for those attributes of the service process that are targeted specifically to each segment could be more effective. To provide a concrete example, as mentioned earlier, in the context of this specific project in the consumer electronics industry, we developed a system of KPI's for the OEM we worked with. These KPI's essentially compared perceived versus expected quality of various specific

130

service dimensions. By analyzing such metrics obtained separately by gender, we found that the service dimension in which the company was best evaluated by men ("Able to effectively resolve your product issues during the first phone call") was precisely the dimension that women evaluated as being the worst. In short, men were very satisfied with what the company was doing in that service dimension (the company exceeded men's expectations by 16%), but women were very unsatisfied with the very same dimension (the company fell short by 28% in terms of women's expectations for that attribute). This example illustrates that by treating segments differently, an OEM can gain further understanding of the managerial action that needs to be taken in order to address the particular needs of each segment.

4.5.2.   Operating a differentiated system

Implementation of a product-service differentiation strategy opens up a number of issues and challenges. In this section, we briefly discuss some of them. Our focus is on the operationalization of a differentiated service-level target strategy introduced in the previous section[19]. We discuss three simple potential operational policies: separation, "round-up", and integration. See Deshpande et al. (2003b) and Kranenburg and van Houtum (2008) for related discussions in the context of spare parts inventory management, and Gurvich et al. (2008) and Bassamboo and Zeevi (2009) for discussions of call centers operations under multiple customer classes.

---

[19] A detailed discussion of the branding and extended service strategies goes beyond the scope of this document.

A first alternative is to operate separate service systems for different segments. In the context of the illustrative example of the previous section, this would imply having different service systems for the platinum, gold and silver segments. We have observed that some airlines indeed use this strategy, where different telephone numbers for customer service are used for different segments. In this context, e.g. the best agents could be assigned to the platinum segment, or more expensive, (higher service capacity) infrastructures, could be used for the platinum and gold segments. In the case we analyze in this research, one recently tested example of such improvements is a policy a.k.a. "concierge service", which consists of having all customer issues handled by the same agent, with the goal of delivering personalized service. Naturally, this policy would be expensive to operate, but our proposed differentiation strategy would allow focusing the intervention on the appropriate segments in a cost effective or profitable way. This alternative, i.e. separate service systems, could be more expensive to operate in terms of resources (in comparison to the integrated operation in which resources are pooled), but it also could be simpler to manage as the resources required to service each specific segment could be managed independently.

A second alternative operating policy is a "round-up" system, that simply sets a single service-level target equal to the one appropriate for the most service-sensitive customer-product class, and applies that standard to all segments. While perhaps such a strategy would be less efficient (more costly), it also represents an alternative that would be easy to implement in practice, as it does not require upfront customer screening or real-time resource rationing. This policy also benefits from complete pooling of the resources. As

noted, different segments potentially care differently about specific service attributes, and therefore a natural thing to do under this policy of offering a single service-level to all would be to design and manage the service support system in a manner that leads to customer satisfaction for the most service-sensitive group.

A third alternative is to have an integrated service system to cover all products and customers. This policy could be efficient in terms of capacity utilization (e.g., through the pooling of resources), but opens up the challenge of how to screen customers and prioritize the allocation of resources in a manner that is consistent with differentiated service target standards. Conditional on being able to screen customers (or customer-products), an integrated system could work by having different service protocols operated by the same, (common) resources. For example, in the context of a call-center, the same agents could use different protocols for responding to calls if the customer is a woman rather than a man. Services related to in-home repair also could easily make use of such call center customer screening, as they involve interaction through the call-center first, i.e., resources for in-home repair services could be assigned depending on the segment.

Finally, we note that more complex rules of operation can of course be proposed, e.g., Gurvich et al. (2008) propose a rationing-type of policy in the spirit of Deshpande et al. (2003a), in which the call center accounts for service differentiation by using an idle-server-based threshold policy. In practice, the decision of what policy to operate requires a careful analysis that considers the costs, benefits and risks involved in each case. Field experiments or simulations could be used to quantify the specific outcomes in practice.

4.6. Conclusions

We have analyzed the association between product and service quality perceptions with consumer intentions to recommend a brand, and the moderating role of customer heterogeneity in that relationship. Our analysis of a case in the consumer electronics industry indicates that perceptions of product quality and service quality jointly influence consumer intentions toward the brand, and that customer heterogeneity greatly affects the relative value of product quality and service quality. In particular, service quality has a bigger impact for women and for high income segments, relative to men and low income segments. We illustrated that while the study is descriptive in nature, the insights derived from it potentially could have important managerial implications. In particular, we discussed a number of strategic implications in terms of product-service differentiation as a potential strategy that the OEM's could pursue. We also discussed a number of operating policies that could be used to implement differentiated service-levels in practice.

From a broader perspective, our analysis illustrates that the joint consideration of product and service by an OEM is important to understand consumer decision-making. Our results suggest e.g. that a company could benefit from managing and monitoring product quality and service quality in a joint fashion. Indeed, a company could lose important information by analyzing these attributes independently. We have observed that a usual practice in industry is for manufacturing companies to create Customer Service Departments with goals solely related to service delivery and support, which are not integrated with product quality. Our results suggest that an OEM should have a

Customer Satisfaction Department which looks at the value creation for customers by considering both product and service quality. This suggests that organizational integration of the functions relevant to both product and service quality within a company and the development of KPI's involving both product and service quality attributes should be considered. Second, as noted by Anderson et al. (1997), "simultaneously managing the product and service subsystem also helps firms manage profitability by allocating resources more effectively between product and service subsystems". Our results have implications concerning the moderating role of customer characteristics such as gender and income, as they could help to identify the most effective allocation of resources to various customer segments targeted for managerial intervention. One specific example is advertising, which could be oriented to influence perceptions regarding product quality and/or service quality depending on the type of customers that are more likely to buy a product category. Third, in this research we introduced the concept of the product-service quality perceptions map, which as our analysis illustrates, could be useful for detecting opportunities for a firm to jointly manage the quality of its products and services. In particular, the perceptions map could be used as a tool for segmentation of the customer base and positioning of product-service bundles. Finally, it is interesting to note that our findings regarding the joint effect of product quality and service quality on consumer decision-making where also present in our recent research in the U.S. automobile industry (Guajardo et al. 2012). That analysis showed that consumer demand in that industry is also jointly influenced by product and service attributes offered by different brands. While in the exposition in this section we do not perform

competitive analysis (as we have only single-firm data), we would expect that the joint consideration of product and service in terms of a firm's competitive strategy is essential in the consumer electronics industry as well.

We believe that our analysis and the results illustrate issues that have not been addressed in the existing literature. We thus believe that ongoing research could serve to motivate the development of new theories and models for the joint management of product and service quality in service-intensive manufacturing industries. Existing studies seeking to understand the joint association of product-service quality with consumer intentions and/or loyalty (Oliver and Swan 1989, Archer and Wesolowsky 1996, Mittal et al. 1999, Hansen and Bush 1999, Devaraj et al. 2001, Chan et al. 2008, Chai et al. 2009) have not studied the moderating effect of consumer characteristics in their analysis, and therefore, existing evidence does not allow us to contrast our findings in this regard with similar analyses conducted in other industries. We believe, however, that new research in this area would be beneficial to understand how the moderating role of customer characteristics potentially differs, depending on the type of products and services that are encountered in different domains.

Chapter 5

Conclusions

Each of the three essays of this dissertation provides an empirical analysis related to the interaction between product quality and service quality in manufacturing industries, based on actual data from consumers' and firms' actions in the aerospace, automobile, and consumer electronics industries, respectively. In the first essay, we analyzed whether and how the customer-service relationship reflected in different service contracts affect product quality. We concluded that indeed the type of repair and maintenance services contracts offered by an OEM have a significant impact on product quality, as explained by different incentive mechanisms. In the second essay, we showed that product quality and service attributes jointly affect consumer demand. Indeed, we showed that service attributes are more important for consumers when product quality is lower, as revealed by actual customer purchases in the U.S. automobile market. In the third essay, we study how product quality and service quality perceptions affect consumer intentions,

specifically, with respect to the likelihood to recommend the brand. Again, we find evidence of a joint effect of product and service quality on consumer intentions, but in this case the relationship reverses. In this industry, product quality and service quality perceptions act as complements. Taken together, the insights derived from this dissertation help to deepen our understanding of the consequences of jointly managing products and services in manufacturing industries, and have direct implications for the development of joint product-service strategies by firms. Indeed, the applications analyzed in this dissertation provide three concrete examples of such joint product-service strategies by manufacturers: service contracts for product support, product-service competitive strategy, and product-service differentiation for a population with heterogeneous valuations of product and service attributes.

From a more general perspective, the results in this dissertation shed some light on a more general question: When is it more effective for manufacturers to compete on services? While certainly a full answer to this question goes beyond what we have considered in this dissertation, our results do provide several insights that are relevant to start answering that question. First, in chapter 3 we show that for a model that considers short-term behavior by consumers and firms, i.e., a static model, the evidence from our analysis of the automobile industry suggests that competing on services is more important, from the perspective of the effect on consumer demand, for companies that are not leaders in product quality. That is, in the short term, services acts mainly as a compensation for low product quality. Our model, however, does not consider long-term effects of firm investment in services, which are more related to the creation of brand

image. In chapter 4, in contrast, we analyze a setting closer to the brand value creation in the consumer electronics industry, and indeed we find that in this case service quality perceptions act as complements to product quality perceptions in terms of consumers' likelihood to recommend a brand, i.e., better service quality perception help to amplify customer loyalty when the product quality perceptions are higher. These seemingly contradictory findings are explained by contextual differences. In our demand model in the automobile industry, our results reveal that the primary mechanism by which services affect consumer decision-making when it comes to a short-term model of purchases is functionality, i.e., services provide compensatory attributes for low product quality that affect customer value creation from product use. However, when it comes to consumers intentions toward a brand, the main mechanism by which service quality perceptions affect consumer intentions is not functionality but rather through affecting brand image. This latter effect would be consistent with a situation where services are more helpful if firms provide better product quality, if we look at long-term effects in terms of brand image. Thus, to continue deepening our understanding of when and how manufacturers should compete on services, future research needs to integrate short-term and long-term effects of services on consumer decision-making, as well as short-term and long-term trade-offs for firms.

Finally, we discuss several avenues for future research that could build upon the insights derived in this dissertation. This research would contribute to expanding our understanding of service competition in manufacturing industries and, in particular, would consider the role of product quality and service quality in a scenario in which

manufacturers compete on services. First, the development of a dynamic economic model of firm competition that considers short-term and long-term effects of a firm's investments in product quality and service quality, coupled with a demand model that allows for heterogeneity in customer preferences for product and service, would be a natural way to integrate some of the most important insights derived from this dissertation. It would also be a good illustration of how empirical research could lead to the development of more complete and realistic theory. Second, as also noted in chapter 3, the overall manufacturer's service strategy should consider both in-warranty and out-of-warranty periods. As discussed in chapter 4, the existence of this out-of-warranty period generates an inherent supply chain conflict between the OEM's and retailers. Indeed, in industries in which retailers represent the most important sales channel, the provision of extended warranties is mostly a retailer's business. Some theoretical models have analyzed some variants of this problem, but little empirical evidence that integrates both aspects of the problem exists. The formulation of an empirical model in this area is another avenue for future research that would contribute to having a broader understanding of service competition in a supply chain context. Finally, given the inherent context-specific nature of the evidence generated by empirical research, the development of new empirical studies in different manufacturing industries is necessary to potentially arrive at conclusions that could be claimed to be more general across industries.

Bibliography

Ackerberg, D., M. Botticini. 2002. Endogenous matching and the empirical determinants of contract form. Journal of Political Economy, 110(3), 564-591.

Allon, G., A. Federgruen. 2007. Competition in service industries. Operations Research, 55(1) 37–55.

Allon, G., A. Federgruen. 2009. Competition in service industries with segmented markets. Management Science, 55(4) 619–634.

Allon, G., A. Federgruen, M. Pierson. 2011. How much is a reduction of your customers' wait worth? An empirical study of the fast-food drive-thru industry based on structural estimation methods. MSOM, 13(4) 489-507.

Anderson, E.W., C. Fornell, R.T. Rust. 1997. Customer satisfaction, productivity, and profitability: Differences between Goods and Services. Marketing Science, 16(2) 129-145.

Archer, N. P., G. O. Wesolowsky. 1996. Consumer response to service and product quality: a study of motor vehicle owners. Journal of Operations Management, 14(2) 103-118.

Automotive News. 2009. Chrysler drops lifetime warranty, (August 24), (accessed July 22, 2011), [available at http://www.autonews.com/article/20090824/RETAIL03/308249890]

Balachander, S., Y. Liu, A. Stock. 2009. An empirical analysis of scarcity strategies in the automobile industry. Management Science, 55(10) 1623-1637.

Bassamboo, A., A. Zeevi. 2009. On a data-driven method for staffing large call centers. Operations Research, 57(3) 714-726.

Berner, R. 2004. The warranty windfall: Service contracts are cash cows—but retailers are mum about their importance. Business Week, December 20, 84–86.

Bernstein, F., A. Federgruen. 2007. Coordination mechanisms for supply chains under price and service competition. MSOM, 9(3) 242–262.

Berry, S. 1994. Estimating discrete choice models of product differentiation. RAND Journal of Economics, 25(2) 242-262.

Berry, S., J. Levinsohn, A. Pakes. 1995. Automobile prices in market equilibrium. Econometrica, 63(4) 841-890.

Berry, S., J. Levinsohn, A. Pakes. 2004. Differentiated products demand systems from a combination of micro and macro data: The new vehicle market. Journal of Political Economy, 112(1) 68-104.

Bolton, P., M. Dewatripont. 2005. Contract Theory. The MIT Press, Cambridge, MA.

Bryant, B. E., J. Cha. 1996. Crossing the threshold: Some customers are harder to please than others, so analyze satisfaction scores carefully. Marketing Research, 8(4) 21-28.

Buell, R.W., D. Campbell, F.X. Frei. 2011. How do incumbents fare in the face of increased service competition? Working paper.

Cachon, P.C., P.T. Harker. 2002. Competition and outsourcing with scale economies. Management Science, 48(10) 1314–1333.

Cachon, G. 2003. Supply chain coordination with contracts. In: S. Graves and T. de Kok, Editors, Handbooks in Operations Research and Management Science: Supply Chain Management, North Holland, Amsterdam, The Netherlands.

Cameron, A.C., P.K. Trivedi. 2005. Microeconometrics: Methods and Applications. Cambridge University Press, New York.

Cameron, A.C., P.K. Trivedi. 2009. Microeconometrics using Stata, College Station: Stata Press.

Campbell, D., F. Frei. 2011. Market heterogeneity and local capacity decisions in services. MSOM, 13(1) 2-19.

Caplin, A., B. Nalebuff. 1991. Aggregation and imperfect competition: On the existence of equilibrium. Econometrica, 59(1) 25-59.

Chai, K-H.,Y. Ding, Y. Xing. 2009. Quality and customer satisfaction spillovers in the mobile phone industry. Service Science, 1(2) 93-106.

Chan, M.C., Y. Ding, K.H. Chai. 2008. The collective effects of product and service quality on customer satisfaction – An empirical study on iPod and iTunes. In Proceedings of the 2008 IEEE ICMIT, 934-939.

Chen, T., A. Kalra, B. Sun. 2009. Why do consumers buy extended service contracts? Journal of Consumer Research, 36(12) 611–623.

Chiappori, P.-A., B. Salanie. 2003. Testing contract theory: a survey of some recent work. In Advances in Economics and Econometrics, M. Dewatripont, L.P. Hansen and S.J. Turnovsky, editors. Volume 1, Cambridge University Press.

Chu, J., P. Chintagunta. 2009. Quantifying the economic value of warranties in the U.S. server market. Marketing Science, 28(1) 99-121.

Chu, J., P. Chintagunta. 2011. An empirical test of warranty theories in the U.S. computer server and automobile markets. Journal of Marketing, 75(3) 75-92.

Churchill, G. A., C. Suprenant. 1982. An investigation into the determinants of customer satisfaction. Journal of Marketing Research, 19(11) 491-504.

Clark, N., J. Mouawad. 2010. In Qantas blowout, concern for jet engine maker. The New York Times, (November 5), (Accessed July 8, 2011), [available at http://www.nytimes.com/2010/11/06/business/global/06engine.html]

Cohen, M.A., S. Whang. 1997. Competing in product and service: A product life-cycle model. Management Science, 43(4) 535-545.

Cohen, M.A., H. Lee, C. Cull, D. Willen. 2000. Saturn's Supply Chain innovation: High value in after-sales service. Sloan Management Review, 41(4) 93-101.

Cohen, M. A., N. Agrawal, V. Agrawal. 2006. Winning in the Aftermarket. Harvard Business Review, 84(5) 129-138.

Colley, A.M., M.T. Gale, T.A. Harris. 1994. Effects of gender role identity and experience on computer attitude components. Journal of Educational Computing Research, 10(2) 129–137.

Consumer Reports. 2008. Extended warranties: A high-priced gamble, (April 2008), (accessed February 15, 2012), [available at http://www.consumerreports.org/cro/2012/05/extended-warranties-a-high-priced-gamble/index.htm]

Cooper, R., T. Ross. 1985. Warranties and double moral hazard. The RAND Journal of Economics, 16(1) 103-113.

Copeland, A., W. Dunn, G. Hall. 2011. Inventories and the automobile market. The RAND Journal of Economics, 42(1) 121-149.

Corbett, C.J. and X. de Groote. 2000. A supplier's optimal quantity discount policy under asymmetric Information. Management Science, 46(3) 444-450.

Crawford, G.S. 2012. Endogenous product choice: A progress report. International Journal of Industrial Organization, 30(3) 315-320.

Cronin, J.J., S.A. Taylor. 1992. Measuring service quality: A reexamination and extension. Journal of Marketing, 56(2) 55-68.

Department of Defense. 2005. DoD guide for achieving reliability, availability, and maintainability. http://www.acq.osd.mil/dte/docs/RAM_Guide_080305.pdf.

Department of Defense. 2009. DoD weapon system acquisition reform: Product support assessment. https://acc.dau.mil/CommunityBrowser.aspx?id=328610.

Deshpande, V., M. A. Cohen, K. Donohue. 2003a. An empirical study of service differentiation for weapon system service parts. Operations Research, 51(4) 518–530.

Deshpande, V., M.A. Cohen, K. Donohue. 2003b. A threshold inventory rationing policy for service-differentiated demand classes. Management Science, 49(6) 683-703.

Devaraj, S., K. F. Matta, E. Conlon. 2001. Product and service quality: The antecedents of customer loyalty in the automotive industry. Production and Operations Management, 10(4) 424-439.

Dieckmann, A., K. Dippold, H. Dietrich. 2009. Compensatory versus non compensatory models for predicting consumer preferences. Judgment and decision-making , 4(3) 200–213.

Douglas, E.J., D.C. Gelnnon, J.I. Lane. 1993. Warranty, quality and price in the US automobile market. Applied Economics, 25(1) 135-141.

Draganska, M., M. Mazzeo, K. Seim. 2009. Beyond plain vanilla: Modeling joint product assortment decisions. Quantitative Marketing and Economics, 7(2) 105-146.

Dube, J-P., J.T. Fox, C-L. Su. 2011. Improving the numerical performance of BLP static and dynamic discrete choice random coefficients demand estimation. Econometrica, forthcoming.

Einav, L., A. Finkelstein. 2011. Selection in insurance markets: theory and empirics in pictures. Journal of Economic Perspectives, 25(1) 115-138.

Emons, W. 1989. The theory of warranty contracts. Journal of Economics Surveys, 3(1) 43-57.

Fan, Y. 2011. Ownership consolidation and product characteristics: A study of the U.S. daily newspaper market. Working paper.

Farris, P.W., N.T. Bendle, P.E. Pfeifer, D.J. Reibstein. 2006. Marketing metrics: 50 + metrics every executive should master. New Jersey: Wharton School Publishing.

Feldman, R., M. Finch, B. Dowd, S. Cassou. 1989. The demand for employment-based health insurance plans. Journal of Human Resources, 24(1) 115–42.

Fisher, M.L., K. Ramdas, K.T. Ulrich. 1999. Components sharing in the management of product variety: A study of automotive braking systems. Management Science, 45(5) 297-315.

Frankel, K. 1990. Woman and computing. Communications of the ACM, 33(11) 34-45.

Franz, C.R., F.Robey. 1986. Organizational context, user involvement, and the usefulness of information systems. Decision Sciences, 17(2) 329-356.

Gallino, S., G.P. Cachon, M. Olivares. 2012. Does inventory increase sales? The billboard and scarcity effect in U.S. automobile dealerships. Working paper.

Gal-Or, E. 1989. Warranties as a signal of quality. The Canadian Journal of Economics, 22(1) 50-61.

Garvin, D.A. 1987. Competing on eight dimensions of quality. Harvard Business Review, 65(6) 107-119.

Geary, S., K. Vitasek. 2008. A Contractor's Guide to Life Cycle Product Support Management. Supply Chain Visions.

Gopal, A., K. Sivaramakrishnan, M. S. Krishnan, T. Mukhopadhyay. 2003. Contracts in offshore software development: An empirical analysis. Management Science, 49(12) 1671–1683.

Gopal, A., K. Sivaramakrishnan. 2008. On vendor preferences for contract types in offshore software projects: the case of fixed price vs. time and materials contracts. Information Systems Research, 19(2) 202-220.

Government Accountability Office. 2008. Improved analysis and cost data needed to evaluate the cost-effectiveness of Performance Based Logistics. GAO-09-41.

Grönroos, C. 1990. Service management: A management focus for service competition. International Journal of Service Industry Management, 1(1) 6-14.

Guajardo, J.G., M.A. Cohen. 2012. Product quality or service quality? The impact of customer heterogeneity. Working paper.

Guajardo, J.A., M.A. Cohen, S. Netessine. 2012. Service competition and product quality in the U.S. automobile industry. Working paper.

Gummesson, E. 1988. Service quality and product quality combined. Review of Business, 9(3) 14-19.

Gurvich, I., M. Armony, A. Mandelbaum. 2008. Service level differentiation in call centers with fully flexible servers. Management Science, 54(2) 279–294.

Halek, M., J.G. Eisenhauer. 2001. Demography of risk aversion. Journal of Risk and Insurance, 68(1) 1–24.

Hansen, E., R.J. Bush. 1999. Understanding customer quality requirements: Model and application. Industrial Marketing Management, 28(2) 199-130.

Heal, G. 1977. Guarantees and risk-sharing. The Review of Economic Studies, 44(3) 549-560.

Heckman, J. 1979. Sample selection bias as a specification error. Econometrica, 47(1) 153-161.

Hesse, H.S. 2012. Retail strategies for extended warranty sales and impact on manufacturer base warranties. Decision Sciences, 43(2) 341-367.

Ho, J., K. Ho, J. Mortimer. 2011. Welfare effects of full-line forcing contracts in the video rental industry. Working paper.

Ho, K., J. Ho, J. Mortimer. 2011. The use of full-line forcing contracts in the video rental industry. Working paper.

147

Hollis, A. 1999. Extended warranties, adverse selection, and aftermarkets. The Journal of Risk and Insurance, 66(3) 321-343.

Holmström, B., P. Milgrom. 1991. Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. Journal of Law, Economics, and Organization, 7, 24–52.

Horowitz, J. 2001. The Bootstrap, in James J. Heckman and E. Leamer (Eds.), Handbook of Econometrics, vol. 5 New York: Elsevier Science.

Hubbard, T. N. 1998. An empirical examination of moral hazard in the vehicle inspection market. The RAND Journal of Economics, 29(2) 406-426.

Hypko, P., M. Tilebein and R. Gleich. 2010. Benefits and uncertainties of performance-based contracting in manufacturing industries. Journal of Service Management, 21(4) 460-489.

IBISWorld. 2010. Global consumer electronics manufacturing - Global industry report. 27th October, 2010.

Iyer, A.V., L.B. Schwarz, S.A. Zenios. 2005. A principal-agent model for product specification and production. Management Science, 51(1) 106-119.

Juran, J.M. 1992. Juran on quality by design: The new steps for planning quality into goods and services. The Free Press, New York, NY.

Kim, S.-H. 2011. Strategic reliability investments in multi-indenture supply chains. Working paper, Yale University.

Kim, S.-H., M. A. Cohen, S. Netessine. 2007. Performance contracting in after-sales service supply chains. Management Science, 53(12) 1843-1858.

Kim, S.-H., M. A. Cohen, S. Netessine. 2011. Reliability or inventory? Analysis of product support contracts in the defense industry. Working paper, Yale University.

Kindström, D. 2010. Toward a service-based business model – Key aspects for future competitive advantage. European Management Journal, 28(6) 479-490.

Kirk, R. L., T. J. DePalma. 2005. Performance-based logistics contracts: a basic overview. CNA Corporation, (November), (accessed April 23 2009), [Available at http://www.cna.org/documents/D0012881.A2.PDF]

Knittel, C.R., K. Metaxoglou. 2012. Estimation of random coefficient demand models: Two empiricists' perspective. Working paper.

Koudal, P. 2008. Ladies and gentlemen, start your service engines: Competing on service excellence in the automotive industry. Deloitte Research, (accessed July 22, 2011), [available at http://www.deloitte.com]

Kranenburg, A.A., G.J. van Houtum. 2008. Service differentiation in spare parts inventory management. Journal of the Operational Research Society, 59 946-955.

Lazear, E. P. 2000. Performance pay and productivity. American Economic Review, 90(5) 1346-1361.

Lee, L. F. 1983. Generalized economic models with selectivity. Econometrica, 51(2) 507-512.

Lips, H.M. 1994. Gender role socialization: Lessons in femininity. In Jo Freeman (ed.), Women: A Feminist Perspective. Mountain View, CA: Mayfield, 128-148.

Lu, M., C. Ma, L. Yuan. 2003. Risk selection and matching in performance-based contracting. Health Economics, 12(5) 339-354.

Lu, J-C., Y-C. Tsao, C. Charoensiriwath. 2011. Competition under manufacturer service and retail price. Economic Modelling, 28(3) 1256-1264.

Lu, Y., M. Olivares, A. Musalem, A. Schilkrut. 2012. Measuring the effect of queues on customer purchases. Working paper.

Maddala, G. S. 1983. Limited-Dependent and Qualitative Variables in Econometrics. Cambridge University Press. Cambridge, UK.

Masten, S., S. Saussier. 2002. Econometrics of contracts: An assessment of developments in the empirical Literature on contracting. Ch.16, In: The Economics of Contracts-

Theories and Applications (E. Brousseau and J-M Glachant, eds.), Cambridge University Press.

Menezes, M.A.J., I.S. Currim. 1992. An approach for determination of warranty length. International Journal of Research in Marketing, 9(2) 177-195.

Mitra, D., P.N. Golder. 2006. How does objective quality affect perceived quality? Short-term effects, long-term effects, and asymmetries. Marketing Science, 25(3) 230–247.

Mittal, V., P. Kumar, M. Tsiros. 1999. Attribute-level performance, satisfaction, and behavioral intentions over time: a consumption-system approach. Journal of Marketing, 63(2) 88-101.

Mittal, V., W.A. Kamakura. 2001. Satisfaction, repurchase intent, and repurchase behavior: Investigating the moderating effect of customer characteristics. Journal of Marketing Research, 38(2) 131–42.

Mittal, V., E.W. Anderson, A. Sayrak, P. Tadikamalla. 2005. Dual emphasis and the long-term financial impact of customer satisfaction. Marketing Science, 24(4) 544–555.

Morris, M.G., V. Venkatesh, P.L. Ackerman. 2005. Gender and age differences in employee decisions about new technology: An extension to the theory of planned behavior. IEEE Transactions on Engineering Management, 52(1) 69-84.

Mortimer, J. H. 2008. Vertical contracts in the video rental industry. Review of Economic Studies, 75 165-199.

Murthy, D.N.P., W.R. Blischke. 2006. Warranty management and product manufacture. Springer-Verlag, London.

Nevo, A. 2000. A practitioner's guide to estimation of random-coefficients logit models of demand. Journal of Economics & Management Strategy, 9(4) 513-548.

Newsome, Jon. 2008. Aviation Week Performance-Based Logistics Survey. (accessed November 15, 2010), [Available at http://thecenter.utk.edu/images/Users/1/PBL/WhitePapers/AviationWeekPBLSurveyConductedforSAPAmericas.pdf]

Ng, I.C.L., F. Ponsignon, R.S. Maull, S.L. Vargo. 2012. The shifting boundaries of marketing and operations under Service Dominant Logic. Working paper.

Novak, S., S. D. Eppinger. 2001. Sourcing by design: Product complexity and the supply chain. Management Science, 47(1) 189-204.

Novak, S., S. Stern. 2008. How does outsourcing affect performance dynamics? Evidence from the automobile industry. Management Science, 54(12) 1963-1979.

Novak, S., S. Stern. 2009. Complementarity among vertical integration decisions: evidence from automobile product development. Management Science, 55(2) 311-332.

Olivares, M., G.P. Cachon. 2009. Competing retailers and inventory: An empirical investigation of General Motors' dealerships in isolated markets. Management Science, 55(9) 1586-1604.

Oliver, R.L., J.E. Swan. 1989. Equity and disconfirmation perceptions as influences on merchant and product satisfaction. Journal of Consumer Research, 16(12) 372-383.

Padmanabhan, V. 1995. Usage heterogeneity and extended warranties. Journal of Economics & Management Strategy, 4(1) 33-53.

Parasuraman, A., V.A. Zeithaml, L.L. Berry. 1985. A conceptual model of service quality and its implications for future research. Journal of Marketing, 49(4) 41-50.

Parasuraman, A., V.A. Zeithaml, L.L. Berry. 1988. SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. Journal of Retailing, 64(1) 12-40.

Petrin, A. 2002. Quantifying the benefits of new products: The case of the minivan. The Journal of Political Economy, 110(4) 705-729.

Petrin, A., K. Train. 2010. A Control Function Approach to Endogeneity in Consumer Choice Models. Journal of Marketing Research, 47(2) 3–13.

Png, I.P.L., D. Reitman. 1994. Service time competition. The RAND Journal of Economics, 25(4) 619–634.

Prendergast, C. 2002. Uncertainty and incentives. Journal of Labor Economics, 20(2) 1071-1102.

Propper, C. 1995. The disutility of time spent on the United Kingdom's national health service waiting lists. The Journal of Human Resources 30(4) 677–700.

Ramdas, K., T. Randall. 2008. Does component sharing help or hurt reliability? An empirical study in the automotive industry. Management Science, 54(5) 922-938.

Randall, W.S., T.L. Pohlen, J.B. Hanna. 2010. Evolving a theory of performance-based logistics using insights from service dominant logic. Journal of Business Logistics, 31(2) 35-61.

Rao, H., H. R. Greve, G.F. Davis. 2001. Fool's gold: Social proof in the initiation and abandonment of coverage by Wall Street analysts. Administrative Science Quarterly, 46(3) 502-526.

Ratchford, B.T. 2001. The economics of consumer knowledge. Journal of Consumer Research, 27(4) 397-411.

Rausand, M., A. Høyland. 2004. System Reliability Theory: Models, Statistical Methods, and Applications. Wiley, New Jersey.

Rothschild, M., J. Stiglitz. 1976. Equilibrium in competitive insurance markets: an essay on the economics of imperfect information. Quarterly Journal of Economics, 90(4) 629-649.

Rust, R.T., C. Moorman, P.R. Dickson. 2002. Getting returns from service quality: Revenue expansion, cost reduction, or both. Journal of Marketing, 66(10) 7–24.

Schumacher, P., J. Morahan-Martin. 2001. Gender, Internet and computer attitudes and experiences. Computers in Human Behavior, 17(1) 95-110.

Shaked, A., J. Sutton. 1982. Relaxing price competition through product differentiation. The Review of Economic Studies, 49(1) 3-13.

Shankar, V., L. Berry, T. Dotzel. 2009. A practical guide to combining products and services. Harvard Business Review, 87(11) 94-99.

Shaver, J. M. 1998. Accounting for endogeneity when assessing strategy performance: does entry mode choice affect FDI survival? Management Science, 44(4) 571–585.

Shen, Y. 2003. Selection incentives in a performance-based contracting system. Health Services Research, 38(2) 535-552.

So, K.C. 2000. Price and time competition for service delivery. MSOM, 2(4) 329-409.

Spence, M. 1977. Consumer misperceptions, product failure and producer liability. The Review of Economic Studies, 44(3) 561-572.

Standard & Poor's. 2011. Industry Surveys: Aerospace and Defense (February 10).

Standard & Poor's. 2011. Industry surveys: Autos & auto parts (June 23).

Suarez, F.F., M.A. Cusumano, S. Kahl. 2008. Services and the business models of product firms: An empirical analysis of the software industry. Working paper.

Sudhir, K. 2001. Competitive pricing behavior in the auto market: a structural analysis. Marketing Science, 20(1) 42-60.

Sunden, A.E., B.J. Surette. 1998. Gender differences in the allocation of assets in retirement savings plans. American Economic Review, 88(2) 207–11.

Tellis, J.G., J. Johnson. 2007. The value of quality. Marketing Science, 26(6) 758–773.

Terwiesch, C., J. Ren, T. H. Ho, M. A. Cohen. 2005. An empirical analysis of forecast sharing in the semiconductor equipment supply chain. Management Science, 51(2) 208-220.

Thomas, T. J. 2005. Performance logistics now coming to Australia. Australian Defense Business Review, 24(7) 23-24.

Thomas, M.U. 2006. Reliability and warranties: Methods for product development and quality improvement. CRC Press, Florida, USA.

Train, K.E., C. Winston. 2007. Vehicle choice behavior and the declining market share of U.S. automakers. International Economic Review, 48(4) 1469-1496.

Tsay, A.A., N. Agrawal. 2000. Channel dynamics under price and service competition. MSOM, 2(4) 372-391.

Vargo, S.L., R.L. Lusch. 2004. Evolving to a new dominant logic for marketing. Journal of Marketing, 68(1) 1-17.

Warranty Week. 2006. Ford's powertrain warranties, (July 25), (accessed July 22, 2011), [available at http://www.warrantyweek.com/archive/ww20060725.html]

Warranty Week. 2011. Automotive warranty report, (April 7), (accessed July 22, 2011), [available at http://www.warrantyweek.com/archive/ww20110407.html]

Wiener, J. L. 1985. Are warranties accurate signals of product reliability? The Journal of Consumer Research, 12(2) 245-250.

Zeithaml, V.A., L. L. Berry, A. Parasuraman. 1996. The behavioral consequences of service quality. Journal of Marketing, 60(4) 31-46.