Three Essays on Product Quality and Pricing

by

Cristina Nistor

B.A. Mathematics and Economics, Bryn Mawr College, 2005
M.A. Mathematics, Bryn Mawr College, 2005

Submitted to the Sloan School of Management
in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Management
at the
Massachusetts Institute of Technology

June 2012

Signature of Author: _____

MIT Sloan School of Management
June 4th, 2012

Certified by: _____

Catherine Tucker
Douglas Drane Career Development Professor in IT and Management
and Associate Professor of Marketing
Thesis Supervisor

Accepted by: _____

Ezra Zuckerman Sivan
Nanyang Technological University Professor
Chair, MIT Sloan PhD Program

Table of Contents

Three Essays on Product Quality and Pricing

By

Cristina Nistor

Submitted to the Sloan School of Management
on June 4, 2012 in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy in Management

Abstract
This dissertation consists of three essays on product quality and pricing.

Essay 1: Pricing and Quality Provision in a Channel: A Model of Efficient Relational Contracts

The first essay analyzes how quality concerns affect relationships in a channel. A firm concerned about uncontractible quality for a customizable good has to pay higher prices to sustain a relationship with the supplier. If the customizable good has very volatile demand, premium payments on this good cannot be sustained. Instead, the downstream firm pays a premium for a good with more stable demand that is correlated with demand for the customizable good. I use a novel dataset containing sales made by a wholesaler to Asian restaurants in the Southeastern United States to test this prediction empirically. As predicted by the proposed model, if customizable goods have very volatile demand, the high end restaurants do not pay a premium on those goods but instead pay a premium for other goods with more stable demand.

Essay 2: Third Party Marketing Approvals

The second essay measures the effect of competition in a certifier market. When customers purchase new products, there is often a degree of uncertainty about their quality. A common solution is to rely on a third-party certifier to provide some form of accreditation that signals quality. However, the incentives of a third-party certifier may not be completely benign. Competitive certification markets may lead the certifiers to provide unduly positive evaluations of quality to gain market share or provide unduly negative evaluations in order to gain credibility with end-users. This paper exploits an unusual natural experiment to evaluate the extent to which third-parties can be relied upon to correctly report product quality. It focuses on the FDA's decision to allow third parties to prepare certifications for certain medical devices, and observes how this decision to introduce competition at the reviewer stage has affected the quality of products allowed to go to market. There is evidence that allowing third party certification leads to significantly lower product quality. However, experience with using a third party reviewer in the past diminishes the negative effect of reviewer competition.

Essay 3: Layaway and the Quasi-Endowment Effect of Installment Payments

The third essay explores the quasi-endowment effect. The paper evaluates how much consumers are willing to prepay for a purchase which will be experienced in the future. In particular, the results indicate that prepaid installment plans allow the consumer to start deriving utility for the purchase from the moment of the first payment. This quasi-endowment effect is felt only for goods that are purchased for own consumption.

Thesis Committee:

Catherine Tucker (Chair), *Douglas Drane Career Development Professor in IT and Management, Associate Professor of Marketing*

Drazen Prelec, *Digital Equipment Corp. Leaders for Global Operations Professor of Management*

Birger Wernerfelt, *J.C. Penney Professor of Management*

Juanjuan Zhang, *Class of 1948 Career Development Professor, Associate Professor of Marketing*

Acknowledgements

I am grateful for the many people who supported my time as a student at MIT Sloan.

I am thankful that my advisor, Catherine Tucker allowed me to find my own research path while devoting great care and thought to my academic attempts. I learned to identify good ideas and great datasets while working with her and I have been happy to benefit from her overall advice about publishing and research. I am also grateful for my committee members, Drazen Prelec, Birger Wernerfelt and Juanjuan Zhang, for working with me over the years and spending a great deal of time caring about my education. My thanks also go to all other members of the MIT faculty, who provided examples and feedback on my work: Sharmila Chatterjee, Duncan Simester, Joshua Ackerman, Michael Braun, Renee Gosline, John Hauser, John Little, Glen Urban.

My thanks go to Sharon Cayley and Hillary Ross for helping me disentangle the academic requirements of the program.

I thank Mami and Adrian for their love and support, and for always believing that I could and should do the best research that I also enjoy.

I also thank my extended family Ngoc, Le and Hannah and Sarah, Brandon, Audrey, Ryden and Evan for being the steady point in my changing life.

Most of all, I am grateful for my husband's support and love. Matt's presence has made me want to be a better researcher and person.

Essay 1:
Pricing and Quality Provision in a Channel: A Model of Efficient Relational Contracts*

Abstract

I model how quality concerns affect relationships in a channel. A firm concerned about uncontractible quality for a customizable good has to pay higher prices to sustain a relationship with the supplier. If the customizable good has very volatile demand, premium payments on this good cannot be sustained. Instead, the downstream firm pays a premium for a good with more stable demand that is correlated with demand for the customizable good. I use a novel dataset containing sales made by a wholesaler to Asian restaurants in the Southeastern United States to test this prediction empirically. As predicted by the proposed model, if customizable goods have very volatile demand, the high end restaurants do not pay a premium on those goods but instead pay a premium for other goods with more stable demand.

*I am indebted to Catherine Tucker and Birger Wernerfelt, for guidance and encouragement. I also thank the restaurant supply company that provided the data for this study. Helpful comments were provided by Sharmila Chatterjee, Bob Gibbons, Drazen Prelec, Matthew Selove, Duncan Simester, Juanjuan Zhang; and also by attendants at the 2012 AMA conference and seminar participants at Berkeley, Chapman, Chicago, Emory, Georgia Tech, MIT, Toronto, and UNC.

1 Introduction

"You would expect that the customer pays us high prices for the value-added services we provide. Well, that doesn't happen. [...] We cross-sell our other products to these customers by offering them significant breaks on the value-add products in return for their commitment to buy the book-and-ship products exclusively from us. In a way, in these relationships the commodity products subsidize the specialty products."

-Stephen Kaufman, Arrow Electronics Inc. CEO[1]

Suppliers and downstream partners face choices over wholesale prices every time they interact in a channel decision. I look at how these choices are affected by different needs for uncontractible quality of the goods provided. My analysis is in the context of repeated interactions, which give firms the opportunity to sustain a level of effort from the supplier that would not be possible in a one-time interaction. I develop a model that shows how channel members can use "relational contracts" (Baker et al. 2002) to provide uncontractible services to the downstream players who value them.

The model is a repeated game played by a supplier and a downstream firm. Both players have the same information about demand, costs and characteristics of the downstream firm. In addition to the physical good, the supplier can add a dimension of quality (such as product customization) that is not contractible through a formal contract.[2] I show that downstream players who care about the uncontractible product customization will pay higher wholesale prices than those who do not. They are willing to pay a premium price as a payment to reward the supplier's effort and sustain the relationship.

If some goods are customizable and some are not, we might expect that the most efficient relationship would set a price premium on the customizable goods to provide stronger effort

[1] Quoted in HBS Case "Arrow Electronics, Inc." (Narayandas 1998)

[2] These services may be uncontractible because they are not yet known at the time the contract is written, or because it would be too difficult to specify, monitor or enforce them.

incentives to the supplier in times of high demand. However, I show that if demand for the customizable goods is very volatile, this creates a problem. When demand is very high, the premium payment required by the supplier may be so high that it exceeds the value of the relationship to the retailer. This could cause the retailer to violate the relational contract by refusing to pay these high premium payments.

Firms face a dilemma between wanting to provide stronger effort incentives to the supplier in times of high demand and, on the other hand, not wanting to require unrealistically high premium payments from the retailer in times of high demand. I show that in some cases, setting a price premium on *non-customizable* goods can help resolve this dilemma. If demand for different types of goods is correlated, but demand for non-customizable goods is less volatile, then a price premium on non-customizable goods can serve the dual role of providing the supplier with stronger incentives in times of high demand while still limiting the "spikes" in these payments to a range that is acceptable to the retailer.

Thus, the model has surprising insights about which types of payments can sustain optimal effort in a relational contract. I show that a price premium on a non-customized good is sometimes the most efficient way to reward a supplier for effort on a customizable good. Alternative payment schemes would result in the relationship breaking down or in less-than-optimal effort.

I test the model on a large novel dataset from a supplier of sushi restaurants from the Southeast region of the United States. I supplement the data with menu prices for each restaurant and use these as a proxy for the restaurant's type as a downstream player.

The results are consistent with the predictions of the model: the supplier customizes raw products to various degrees if the restaurant requests it. Restaurants that require customization pay a premium on *non-customized* products such a rice.[3] This is consistent

[3]The Robinson-Patman Act prohibits price discriminations between purchasers of commodities of like grade and quality which are likely to result in substantial injury to competition. However, it is not applicable in the case of goods that are improved through the service provided by the manufacturer. In effect, a

with the model, as products like fresh fish have volatile demands that are quickly affected by supply shocks or other non-demand related shocks, so any payment scheme linked to the actual product being customized would not be able to sustain the relational contract that provides the customization effort. On the other hand, dry products like rice are relatively stable goods that are perfect for serving as a basis for the premium payment to sustain the relational contract.

My results are consistent with the intuition of managers in the industry. In an interview with the manager for the supply company, he says the main goal for the sales force is to make the high-end restaurants happy with the relationship. In terms of wholesale prices, the supplier believes that high-end restaurants do not care about prices as long as they get a consistent quality of goods and all their service demands are satisfied. On the other hand, the manager is aware that none of the low end restaurants will want to pay for extra services and that they will switch suppliers as soon as they find a vendor with lower prices.

Conversations with a manager from a value-oriented restaurant[4] revealed that this type of downstream player is mainly concerned with obtaining low prices for an adequate quality of the actual good. The manager decided to use several suppliers by comparing prices for each good and choosing to purchase the lowest price good from each supplier. Even after a few years of operation, the restaurant still uses several suppliers and buys certain goods from each of them.[5]

customized good has added service that does warrant a different price from different restaurants. Most importantly, there is no intention of competitor harm (Luchs et al. 2010) in the bargaining over prices and quality of the products delivered.

[4]Personal interview with a manager of a small family owned and operated sushi restaurant in a large city in Latin America. The restaurant is aimed at middle income college students and workweek lunch crowd. They believe their main competitors are fast-food restaurants priced around 4$ to 5$ because their menu has bundled prices of 5$ to 7$.

[5]The manager I interviewed acknowledged that he sometimes calls the sales agent to complain about wholesale prices if the price increase is unexpected (5% or 7% instead of 2%). Moreover, he claims that the supplier will sometimes substitute products if an item becomes unavailable and the restaurant is faced with the difficult choice of returning the products and not having enough material to work with or just keeping the inadequate products. He speculates that the situation arises because suppliers are sometimes hit with limited supply of an item, and they keep the small quantities for their 'bigger' customers. I would argue that this explanation is indeed very plausible, and that it is likely the supplier is making sure their 'best'

This paper analyzes a theoretical model of relational contracts with a specific empirical application to the restaurant industry. However, the problem is more general: relational contracts are used in any channel where suppliers and downstream firms face a problem of uncontractible effort, costly monitoring of product quality, or a classic problem of channel coordination without a way to write enforceable contracts.

Arrow Electronics is another example of a company that uses a pricing structure like the one modeled in this paper to sustain a relationship with its customers. The company is a distributor for two types of micro-chips: standard commodity chips and "value-added" chips they customize for each client. The latter type of products requires effort investment from Arrow. As mentioned by the company's CEO, the company does not charge a price premium for value-added products they customize for each client's specific needs. Instead, the company prices these value-added chips competitively and uses them to attract and retain customers who then pay higher margins on the standard chips.

A third example of a supplier that uses this type of pricing structure is a healthcare consultant who prepares patient satisfaction reports for hospitals in a large city in the South-western United States.[6] The clients are hospitals who have ongoing contracts for five years that can be terminated with one month notice. The clients sometimes have requests that are not written in the initial contract because it is difficult to know ahead of time what will be needed. They ask for added services, such as individual reports, results by area of management, or reports broken out by any variable in the data. The consultant generally does not charge for these extra projects, which he usually accepts, even though they are effort intensive. Instead, the price of the standard contract is set higher to make up for these added requests. If the added project is too big, then a one-time invoice will be sent to the client, but the price will still be lower than what it would be if the client were to pay regular

customers are getting the needed inputs.

[6]This example is based on a personal interview with the consultant.

prices (for example, a large extra project that would usually require a $30,000 budget if done for a different customer will actually be billed as a $8,000 project). The practice of having a price premium on the standard contracts and then offering extra services for free or lower prices is meant to sustain a relationship with the clients and encourage them not to price shop for each service individually, but rather think of all transactions as part of a relationship with the consultancy.

In Section 2 I review related literature. Section 3 introduces the theoretical model. The data are presented in Section 4, with the empirical analysis following in Section 5. Section 6 concludes with implications for channel management and directions for future research.

2 Related Literature

Bargaining as a mechanism of setting wholesale prices has been analyzed in Iyer and Villas-Boas (2003). This paper extends the results by allowing infinitely repeated interactions that make relational contracts possible. Additionally, Iyer and Villas-Boas (2003) show that two-part tarrifs are not useful in a bargaining setting. Ho and Zhang (2008) also find that two part tarrifs are not successful in coordinating a channel. They verify empirically that framing may improve efficiency of using two-part tarrifs. This paper is consistent with their findings: I establish that two-part tarrifs are too inflexible to sustain efficient relational contracts and be used in a channel where wholesale prices are set through a repeated bargaining process.

Recent theoretical models and empirical work show manufacturer's response to a strategic retailer (Bayus and Putsis 1999; Geylani et al. 2007; Dukes et al. 2009). In a related context, I show that there are strategic reasons for the non-dominant retailer to assume low-bargaining power over wholesale prices in order to maintain a relationship with the manufacturer. Under certain conditions, the relationship with the manufacturer is a relational contract that ensures optimal quality provision in the channel. Additionally, the current paper identifies reasons for the weak retailer's inability to dictate low wholesale prices.

A related stream of literature deals with information that can affect profits of channel partners. This information could be about the costs faced by the retailer, or a signal about product quality geared toward end consumers, or information about manufacturing costs or even strategic information sharing in a channel (Jeuland and Shugan 1983; Desiraju and Moorthy 1997; Shaffer and Zettelmeyer 2002; Corbett et al. 2004; Liu and Zhang 2006; Guo and Iyer 2010). Information asymmetry in bargaining is documented in Busse et al. (2006). This paper assumes that all parties in the channel have the same information about demand, costs and each other's types.

Most empirical studies of channel relations to date have focused on how differences in market power of various brands or manufacturers affect the proportion of promotional passthrough (Besanko et al. 2003; Nijs et al. 2009) or have investigated the competitive structure of a channel (Kadiyali et al. 2000; Sudhir 2001; Villas-Boas and Zhao 2005; Draganska et al. 2010). This paper looks for the reasons leading to different market power levels of retailers in a channel and estimates the resulting differences in prices they pay to the upstream supplier.

This paper is related to a stream of literature that studies relationships in a channel or in business-to-business environments and identifies characteristics of how such relationships may influence transactions (Spekman et al. 1998; Jap 1999; Cannon and Perreault 1999; Ghosh et al. 2006; Shervani et al. 2007; Tuli et al. 2007, 2010; Rietz et al. 2011). Additionally, this literature stream has analyzed how the use of punishment strategies and channel conflict can sustain close ties (Kumar et al. 1995, 1998). The present paper contributes to this literature by formulating a theoretical model of repeated interactions that provides insights into the mechanism behind prices and quality in channel relationships.

This paper uses an infinitely repeated game in the style of Rotemberg and Saloner (1986); Abreu (1988); Lal (1990); Dechenaux and Kovenock (2007); Thomadsen and Bhardwaj (2011). Like these earlier papers, I derive conditions in which repeated interactions enable

firms to cooperate in equilibrium. By applying this framework to the problem of pricing and quality provision in a channel with multiple products, I generate new insights about the optimal pricing tactics that make cooperation sustainable. This leads to the surprising result that in some cases more efficient cooperative outcomes are possible when the downstream firm pays a premium on goods that are not being customized, instead of on the more volatile customized goods.

An area often neglected in marketing is that of relational contracts that develop in a channel. There is a growing economic literature on contracts and relational contracts in particular (e.g Baker et al. 2002; Levin 2003; Brown et al. 2004; Gibbons 2005; Plambeck and Taylor 2006; Halac 2011). This paper brings to light the importance of relational contracts in marketing channels. It explains how downstream players may end up with different levels of market power in a channel by allowing the use of relational contracts between retailers and suppliers. This paper also compares different payment schemes that allow efficient relationships to provide optimal effort. The results are consistent with Levin (2003), who finds in a different context, that the discretionary payments to sustain the relationship cannot be too large to prevent the principal from walking away, or too low to prevent the agent from walking away. I extend these results to study pricing mechanisms that make relational contracts sustainable in a channel. One important difference is that I explicitly model channels in which the supplier sells multiple goods and study which types of goods are suitable for the price premiums that sustain relational contracts.

3 Theoretical Model

I build a model with two risk-neutral players: a manufacturer and a retailer. The upstream firm can sell two goods to the downstream player. The value of good one can be increased if the seller invests (non-contractible) effort into customizing it for the retailer. Good two is a basic good, which cannot be customized. The players trade repeatedly at dates $t = 0, 1, 2....$

For simplicity, let good one, which has an uncontractible quality dimension, be denoted by subscript u while the basic good, good two is denoted by b.

On-time delivery, consistent good customer service, personal help and even customized improvements to the physical good are possible examples of quality dimensions that may be uncontractible or unknown (such as in Iyer and Villas-Boas (2003)) at the time the bargaining process over prices takes place. The model allows for these uncontractible quality dimensions to be quite general: it suffices that a formal contract may be hard to enforce, or hard to specify or costly to monitor in order for the results to be valid.

The firms find it impossible to write completely enforceable contracts because of the complexity of the transactions, but their desire to continue to do business with each other in the future acts as an incentive to maintain a relational contract. Under certain conditions, this relationship is self-enforcing and leads to optimal quality for the buyer because both parties fear the loss of future benefits if they deviate from cooperating.

I do not explicitly model consumer demand, but I assume the retailer needs quantities $q_{t,u}$ and $q_{t,b}$ of goods one and two, respectively, in period t. These amounts can vary randomly from period to period, and I allow for $q_{t,u}$ and $q_{t,b}$ to be correlated.

The downstream firm's utility at time t is given by $U_{R_t} = q_{t,u}(\alpha_1 + e_t - P_{t,u}) + q_{t,b}(\alpha_2 - P_{t,b})$ with e_t the non-contractible effort per unit put in by the seller at time t. $P_{t,u}$ and $P_{t,b}$ are the unit prices at time t, and α_1 and α_2 are constants.

The seller's utility is given by $U_{S_t} = q_{t,u}(P_{t,u} - c(e_t)) + q_{t,b}(P_{t,b})$ with $c(e)$ being the per-unit cost of effort, which I will assume for simplicity is of the form $c(e) = \beta e^2$. Without loss of generality, set marginal cost equal to zero.

The retailer wants to maximize $\sum_{t=0}^{\infty} \delta^t U_{R_t}$ while the supplier wants to maximize $\sum_{t=0}^{\infty} \delta^t U_{S_t}$.

Note that optimal effort for maximizing total channel profits is given by $e^* = \frac{1}{2\beta}$. This is the optimal *per unit* effort which implies that optimal investment costs for the seller are higher when quantity $q_{t,u}$ is high.

The supplier has the outside option to sell the goods at unit prices $\bar{P}_u$ and $\bar{P}_b$. The outside option does not reward the supplier for the effort expended on customization. For example, the outside option may be a low end retailer who does not value quality or it could be that the effort is specific to the retailer who requests it.

3.1 Timing

The parties initially agree on a relational contract $(P_u, P_b, e(q))$ with constant unit prices for each good and effort which can be a function of quantity. The restriction to constant unit prices can be justified on the grounds of simplicity (Schmalensee 1989). However, it is important to allow per-unit effort to vary with quantity because in some cases the optimal effort cannot be sustained if quantity is very large.

At each time t, the game has 3 stages:

1. Nature draws $q_{t,u}$ and $q_{t,b}$ and the buyer and the seller observe this demand for that particular period.

2. The seller decides how much effort to put into providing the service, e_t. The buyer observes this choice.

3. The buyer can either agree to pay the unit prices from the relational contract P_u and P_b or can deviate and go to the outside market, which offers $\bar{P}_u$ and $\bar{P}_b$.

3.2 Results

There is always a bad equilibrium where the supplier invests zero effort and the retailer offers $(\bar{P}_u, \bar{P}_b)$. Assume that if the relationship breaks down then the players revert to this bad equilibrium.

For the relational contract to be sustainable there can never be a time when either firm is better off deviating than staying in the relationship. I derive conditions for both the supplier and the retailer to stay in the relationship.

At any time t the utility for the buyer if it stays in the relationship is given by its utility in the present period added to the discounted stream of utilities it gets in the future while it is in the relationship:

$$q_{t,u}[\alpha_1+e_t(q_{t,u})-P_u]+q_{t,b}[\alpha_2-P_b]+\sum_{T=t+1}^{\infty}\delta^{T-t}\{E[q_{T,u}(\alpha_1+e_T(q_{T,u})-P_u)]+E[q_{T,b}(\alpha_2-P_b)]\}. \tag{1}$$

Note that I allow effort e_T to be a function of quantity $q_{T,u}$ as specified in the relational contract mentioned before.

On the other hand, if the retailer deviates and offers lower prices, the relationship will end. In this case, the prices in this period and all future periods will be $\bar{P}_u$ and $\bar{P}_b$, and effort will be zero in all future periods. Thus, the retailer's utility if it deviates at any time t is given by:

$$q_{t,u}[\alpha_1+e_t(q_{t,u})-\bar{P}_u]+q_{t,b}[\alpha_2-\bar{P}_b]+\sum_{T=t+1}^{\infty}\delta^{T-t}\{E[q_{T,u}(\alpha_1-\bar{P}_u)]+E[q_{T,b}(\alpha_2-\bar{P}_b)]\}. \tag{2}$$

Note that given the assumptions about timing of the model, the supplier has already invested a sunk effort at time t, so effort is not affected until the next period $t+1$.

Comparing the conditions (1) and (2) and rearranging terms gives the boundary under which the retailer will always want to stay in the relationship:

$$(P_u-\bar{P}_u)\Big[q_{t,u}+\frac{\delta}{1-\delta}E[q_{T,u}]\Big]+(P_b-\bar{P}_b)\Big[q_{t,b}+\frac{\delta}{1-\delta}E[q_{T,b}]\Big]\leq\frac{\delta}{1-\delta}E[q_{T,u}e_T] \tag{3}$$

Intuitively, the present value of all the premium payments the retailer makes must be less than the present value of the benefits it receives due to the supplier's higher effort in order for the relationship to be sustainable.

The utility for the *supplier* if it stays in the relationship at any time t is:

$$q_{t,u}\Big[P_u - \beta[e_t(q_{t,u})]^2\Big] + q_{t,b}[P_b] + \sum_{T=t+1}^{\infty} \delta^{T-t} E\Big[q_{T,u}P_u + q_{T,b}P_b - q_{T,u}\beta[e_T(q_{T,u})]^2\Big]. \quad (4)$$

On the other hand, if the supplier deviates and provides lower effort than specified in the relational contract, the retailer will only agree to pay $\bar{P}_u$ and $\bar{P}_b$ (the supplier's reservation prices) in the current periods and in all future periods as well. Thus, at any time t, the supplier's utility if it deviates and provides zero effort is:

$$q_{t,u}[\bar{P}_u] + q_{t,b}[\bar{P}_b] + \sum_{T=t+1}^{\infty} \delta^{T-t} E[q_{T,u}\bar{P}_u + q_{T,s}\bar{P}_b]. \quad (5)$$

By comparing conditions (4) and (5) for the supplier and rearranging terms, the supplier will always want to stay in the relationship only if at each time t:

$$\begin{aligned}(P_u - \bar{P}_u)\Big[q_{t,u} + \frac{\delta}{1-\delta}E[q_{T,u}]\Big] \quad + \quad (P_b - \bar{P}_b)\Big[q_{t,b} + \frac{\delta}{1-\delta}E[q_{T,b}]\Big] \geq \\ \geq q_{t,u}\beta[e_t(q_{t,u})]^2 + \frac{\delta}{1-\delta}\beta E\Big[q_{T,u}[e_T(q_{T,u})]^2\Big] \quad (6)\end{aligned}$$

Intuitively, the present value of the premium payments the supplier receives in this period and all future periods must be greater than the present value of its cost of effort in order for the supplier to stay in the relationship.

I first study the case in which optimal effort is always sustainable. In this case, I prove that it is always possible to create an efficient relational contract that sets a price premium only on the good with uncontractible quality. However, if demand for the good with uncontractible quality is very volatile, it is not possible to sustain the optimal effort in all periods. In this case, I derive conditions in which a contract that places a premium on the *non-customizable* good leads to a more efficient outcome than a contract that places a premium on the *customizable* product.

By combining the constraints for the supplier and the retailer, a necessary condition for a contract to be sustainable is that:

$$\beta\left[q_{t,u}[e_t(q_{t,u})]^2 + \frac{\delta}{1-\delta}E[q_T e_T(q_T)]^2\right] \leq \frac{\delta}{1-\delta}E[q_T e_T(q_T)]. \tag{7}$$

Intuitively, the present value of the cost of effort (to the supplier) in the current and all future periods must be less than the present value of the benefit of effort (to the retailer) in all future periods.

If I substitute in the expression for optimal effort $e^* = \frac{1}{2\beta}$, condition (7) becomes:

$$\beta\left(\frac{1}{2\beta}\right)^2\left[q_{t,u} + \frac{\delta}{1-\delta}E[q_{T,u}]\right] \leq \left(\frac{1}{2\beta}\right)\left(\frac{\delta}{1-\delta}\right)E[q_{T,u}].$$

Rearranging terms and simplifying gives:

$$q_{t,u} < \left(\frac{\delta}{1-\delta}\right)E[q_{T,u}]$$

The relationship is always sustainable (for all possible $q_{T,u}$) only if:

$$\max(q_{t,u}) < \left(\frac{\delta}{1-\delta}\right)E[q_{T,u}] \tag{8}$$

This condition states that optimal effort can only be sustained if the maximum quantity is small enough relative to the expected quantity, and if firms place enough weight on the future. Intuitively, the expected value of the relationship must be large enough to compensate for the effort required in times of greatest demand.

Proposition 1. *If it is possible to sustain optimal effort for all levels of demand (inequality (8) holds) then there is an efficient relational contract that sets a premium only on the good with uncontractible quality* ($P_u > \bar{P}_u, P_b = \bar{P}_b$).

Proof: Set $e(q) = e^* = \frac{1}{2\beta}$, and set $P_u - \bar{P}_u = \beta(\frac{1}{2\beta})^2$. The supplier's constraint (6) always holds with equality. The assumption that condition (8) holds (and therefore (7) holds as well) implies that the retailer's constraint (3) always holds as well. QED

Intuitively, in cases where optimal effort is sustainable, a premium on the good with uncontractible quality can be used to exactly offset the supplier's cost of effort in each period. The retailer is always willing to make this premium payment if condition (8) holds.

The next step in the analysis is to identify an arrangement that gives both the seller and the buyer the appropriate incentives to sustain an efficient relational contract when optimal effort (first best) is not sustainable.

I will now derive conditions that guarantee a more efficient contract is possible when premium payments are made on the non-customizable good instead of the customizable good. As an example, to provide intuition for how this can happen, I assume the quantity traded for each product has the following distribution:[7]

$$\begin{aligned} Prob(q_u = L) &= 1 - \omega \\ Prob(q_u = H) &= \omega \\ Prob(q_b = L) &= 1 - \omega \\ Prob(q_b = M) &= \omega \end{aligned}$$

where $L < M < H$ and:

$$M < \left(\frac{\delta}{1-\delta}\right) L$$

$$H > \left(\frac{\delta}{1-\delta}\right) L.$$

[7]For simplicity I assume demand for the two goods is the same in the low condition. However, this assumption could be relaxed and the prices for the good that does not have uncontractible quality would need to be adjusted for it.

I assume demand for the goods is perfectly correlated, although any sufficiently strong positive correlation would be enough for the results to hold.

Proposition 2. *When ω is small enough, a more efficient relational contract is possible when premium payments are made on the non-customizable good (the basic good) than when they are made on the good with uncontractible quality. A premium on the non-customizable good is also more efficient than a fixed premium payment.*

Proof:

When ω becomes small enough, the retailer's constraint (3) converges to:

$$(P_u - \bar{P}_u)\Big(q_{t,u} + \frac{\delta}{1-\delta}L\Big) + (P_b - \bar{P}_b)\Big(q_{t,b} + \frac{\delta}{1-\delta}L\Big) \leq \frac{\delta}{1-\delta}Le(L).$$

The supplier's constraint (6) converges to:

$$(P_u - \bar{P}_u)\Big(q_{t,u} + \frac{\delta}{1-\delta}L\Big) + (P_b - \bar{P}_b)\Big(q_{t,b} + \frac{\delta}{1-\delta}L\Big) \geq q_t\beta E[e(q_t)]^2 + \frac{\delta}{1-\delta}\beta L[e(L)]^2.$$

Because the probability of low demand L approaches 1, any contract that achieves optimal effort at times of low demand will dominate a contract that does not. Therefore, I will explore contracts such that both of these conditions hold:

- optimal effort occurs in times of low demand
- the contract does not break down in times of high demand.

The first property implies that $e(L) = \frac{1}{2\beta}$. Substituting this term into the supplier's constraint, we see that in order for the supplier to maintain optimal effort in times of low demand requires:

$$\frac{1}{4\beta}\left[L + \frac{\delta}{1-\delta}L\right] \leq [P_u - \bar{P}_u + P_b - \bar{P}_b]\left[L + \frac{\delta}{1-\delta}L\right] \tag{9}$$

This implies $\frac{1}{4\beta} \leq P_u - \bar{P}_u + P_b - \bar{P}_b$.

We also need to ensure that the retailer is willing to make the required premium payment in times of *high* demand, which requires:

$$(P_u - \bar{P}_u)\left(H + \frac{\delta}{1-\delta}L\right) + (P_b - \bar{P}_b)\left(M + \frac{\delta}{1-\delta}L\right) \leq \left(\frac{\delta}{1-\delta}\right) L \left(\frac{1}{2\beta}\right) \qquad (10)$$

Consider a contract that places a premium only on the good with uncontractible quality. By the assumption that $H > \frac{\delta}{1-\delta}L$, the inequality in (10) implies that $P_u - \bar{P}_u < \frac{1}{4\beta}$. However, this implies that the supplier's constraint (9) cannot be satisfied.

Intuitively, because the good with uncontractible quality has such large demand spikes, the contract has to limit the premium payment on this good to prevent the retailer from deviating in times of high demand. However, this constraint on the price premium will be so low that the supplier does not have enough incentives for optimal effort in times of normal (low) demand.

By contrast, by the assumption that $M < \frac{\delta}{1-\delta}L$, the simple, non-customizable good (good b) does not face this problem. Setting $P_b - \bar{P}_b = \frac{1}{4\beta}$, the retailer's constraint (10) still holds in times of high demand, and the seller has incentives for optimal effort in times of low demand because (9) holds.

Finally, a premium on the non-customizable good is superior to a fixed premium payment. Because demand for this good is correlated with demand for the good with uncontractible quality, the supplier has stronger incentives for effort when demand for the uncontractible quality good is high. Although it is not possible to sustain optimal effort in times of high demand, this additional premium payment makes it possible to sustain greater effort than would be possible with a fixed premium payment. QED

I now present a numerical example in which optimal effort is *not* sustainable, and in which a relational contract with a premium payment on the simple, non-customizable good

is more efficient than one with a premium payment on the good with uncontractible quality.

Assume the distribution of $q_{t,u}$ is given by: $P[q_{t,u} = 5] = 0.9$ and $P[q_{t,u} = 50] = 0.1$. Let $\delta = \frac{2}{3}$. This implies that $\left(\frac{\delta}{1-\delta}\right) E[q_{t,u}] = 2(9.5) = 19$. Because $\max[q_{t,u}] > 19$, optimal effort is not sustainable. Intuitively, this good has a distribution that occasionally has such large spikes in the quantity traded (due to either exogenous supply or demand shocks) that the cost of optimal effort in these periods exceeds the value of the ongoing relationship, so a relational contract cannot be used to sustain optimal effort.

However, there is a "near optimal" relational contract that is sustainable using a premium payment on the simple, non-customizable good, under the assumption that quantity of the two goods is perfectly correlated but the distribution of good b is less volatile.

Let the distribution be $P[q_{t,b} = 5] = 0.9$ and $P[q_{t,b} = 18] = 0.1$.[8] Consider the following relational contract: $P_u - \bar{P}_u = 0$, $P_b - \bar{P}_b = \frac{1}{4\beta}$, $e_u(q_{t,u} = 5) = \frac{1}{2\beta}$, and $e_u(q_{t,u} = 50) = \left(\frac{1}{2\beta}\right)\left(\frac{3}{5}\right)$. This contract calls for zero premium payment on the customizable good but positive premium payment on the non-customizable good. It also calls for optimal effort when demand is low and $\frac{3}{5}$ optimal effort when demand is high. By substituting these values into the constraints for the retailer and the supplier, it can be seen that each firm always wants to stay in this contract, both in periods of low and high demand.

Intuitively, setting a price premium on the more stable, non-customized good provides stronger incentives to the supplier in times of high demand without letting the total premium payment grow so large that the retailer would want to exit the relationship.

This same contract is not sustainable if the same premium payment is placed on the good with uncontractible quality because the premiums become more than the retailer is willing to pay during the large spikes in quantity traded of this good.

[8] It is not necessary that demand for the good be perfectly correlated, but this helps keep the example simple.

4 Data

The analysis uses a novel dataset from a supplier of Asian food products to restaurants in the Southeast United States. The supplier is a mid-size supplier who has multiple competitors in the area. The sales agents work with a few restaurants throughout the entire period that the restaurant is a client but there are instances in which the agents will take over someone else's account.[9]

The sales agents take orders from clients and negotiate prices. Restaurants' order frequency varies from once a day to a few times a week. The sales agents encourage the restaurants to try new goods and explain how items can be used if the items are new. The agents have some latitude over the prices they can negotiate with the restaurants. In this industry, discounts are not used very much, but the agents are known to give either small discounts for bulk orders or other small discounts. The most frequent type of discount given is a 2% discount that sales force awards periodically to their customers. This discount is a significant one for a typical sushi (or similar) restaurant. The National Restaurant Association claims that restaurant margins are between 4 and 6% before taxes.[10]

The dataset contains invoice level data on prices, products details and quantities. I supplement this information with characteristics for the restaurant who orders the products. I used Urbanspoon.com and yelp.com to find average menu prices for the restaurants. The menu prices are in four categories: "Under $10", "$10 to $15", "$15 to $25", and "Over $25". The restaurants which had no menu price available online (either from customer reviews or from the restaurant) were excluded from the analysis. The price levels are used as a proxy for the restaurants type: High End restaurants will generally be considered those with prices over $10.

[9] In the dataset, there are a few instances when a switch happened because an agent resigned and some other similarly random events. It is very rare for a salesperson to be taken off an account for performance problems.

[10] www.restaurant.org

Table 1: Sales Agent Orders Percent By Item Perishability

	Dry	Fresh	Frozen	Other
Agent 1	18	40	36	6
Agent 2	26	23	45	7
Agent 3	25	22	47	6
House	21	24	48	7
Overall	22	28	43	6

The dataset spans 49 weeks[11] and several states: Alabama, Florida, Georgia, and South Carolina. The goods are either from domestic sources, or imported from China, Japan, Scotland, and Vietnam. There are around 600 unique product codes being transacted, with 178 Dry goods such as Rice, 33 Fresh goods such as Big Eye Tuna, 305 Frozen items such as Smoked Shrimp and 83 Other items like Chopsticks. The Dry, Fresh and Frozen products represent 22%, 28%, and 3% of the transactions respectively.

Table 1 reports the percent distribution of goods (overall times ordered) for each agent, by type of good.

The dataset contains specific notes about the restaurants' requests for customization of the products. The customizable products correspond to goods with uncontractible quality from the theoretical model. While not all products can be customized, there are certain products (such a fresh fish) that are amenable to customization. I classify these products based on the requests available in the dataset: if a product code is ever associated with a request for specific service, then I include it in the set of products that can be potentially customized. This classification does not imply that the product will always be customized or that all restaurants ask for special dimensions of the product. Table 2 shows the distribution of unique product codes that can be customized with the distribution of the customizable products for overall number of orders (with intuitive ordering weights).

The restaurant supply industry is affected by supply shocks due to uncertainty about

[11]Data span April 2010 to February 2011.

Table 2: Customizable Items

Product name	% of Total Orders	% of Customized Orders
Frozen Scottish Whole Salmon	9.09	27.64
Y/F Tuna Loin	3.68	11.20
Fluke(Hirame)	2.64	8.03
Farm Rock	2.53	7.69
Fz. Escolar Block	2.20	6.70
California Uni	1.83	5.56
Fz. Smoked Salmon Chunk	1.81	5.51
Frozen Hamachi Fillet	1.66	5.06
(Frozen) Scottish Salmon Fillet for Sushi	0.85	2.60
Big Eye Tuna Loin	0.78	2.38
Fz Hamachi Loin Farm Japan	0.55	1.69
Asi Beff Gyoza	0.51	1.55
Mushidako Octopus	0.49	1.48
Fresh Hamachi Fillet(Japan)	0.43	1.30
Aji	0.43	1.29
Fz. Albacore Tuna Loin	0.42	1.28
Fz. Escolar Block	0.41	1.25
Live Mirugai(Geoduck)	0.41	1.23
Madai (Japan)	0.29	0.87
Fresh Kanpachi Fillet	0.24	0.73
Unagi	0.24	0.72
Spanish Mackeral (USA)	0.23	0.69
OO-Toro Southern Blufin	0.22	0.68
Atlantic Whole Salmon (Farm Raised)	0.21	0.62
Tuna Ground	0.20	0.61
Apex Y/F Tuna Saku AAA	0.14	0.43
Blue Fin O-Toro	0.13	0.40
YF Tuna Loin	0.10	0.29
Bluefin Tuna Loin	0.08	0.25
BE Super Frozen Tuna Saku	0.03	0.10
Y.F. Tuna Loin	0.02	0.07
Chillian Sea Bass	0.02	0.06
BF Frozen O-Toro(Saku)	0.01	0.03
Awabi (Abalone)	0.00	0.01
YF Super Frozen Tuna Saku	0.00	0.01
Overall Customized	32.87	100.00

raw, fresh materials available at any time. Uncertain seafood catch affects the quantity of items being transacted: if there is historically volatile supply of a particular type of fish, the restaurants may include it only in a few menu items or keep a list of special dishes when that particular fish is in season. However, to a certain degree, most fresh goods are volatile in terms of supply: weather patterns or demands in other countries[12] may affect availability for restaurants. For the empirical analysis, having a volatile transacted quantity for raw products influences how the restaurants pay for customization of these products. The model suggests that volatile items cannot sustain relational contracts that will provide adequate levels of quality so the payments for customization will be made in the context of a bundle of goods and will be tied to more stable goods.

5 Analysis

5.1 Main Effect

Table 3 shows that the supplier charges higher prices to high-end restaurants. I consider that a High End Restaurant is one whose menu prices are over $10. These are the restaurants who customize the products most, and thus would be willing to pay more for the customization. I estimate a simple OLS specification with fixed effects for each product code, and perform the estimation with the negotiated transaction prices. Table 3 includes robustness checkes for several other specifications. The coefficient for High End restaurants is positive, indicating that higher end restaurants pay more for their inputs on average. In fact, further analysis shows that demand volatility and customization are moderating factors for the payment premiums.

[12]For example, the sushi industry in the US is currently suffering from a low supply of eel because the main exporter, China, has recently seen an increase in domestic demand for it.

Table 3: High End Restaurants with Price Over $10 Pay Higher Prices

	(1) Price	(2) Log Price	(3) Price	(4) Log Price	(5) Price	(6) Log Price	(7) Price	(8) Log Price	(9) Price	(10) Log Price
High End Restaurant	0.703***	0.0280***	0.820***	0.0264***			0.878***	0.0366***	0.528***	0.0233***
	(0.161)	(0.00476)	(0.166)	(0.00484)			(0.168)	(0.00480)	(0.162)	(0.00478)
High End Over $15					0.947***	0.00975***				
					(0.106)	(0.00371)				
Item fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Agent Fixed Effects	No	No	Yes	Yes	No	No	No	No	No	No
Observations	44972	44972	44972	44972	44972	44972	42957	42957	43047	43047
R^2	0.8862	0.8515	0.8863	0.8516	0.8864	0.8514	0.8880	0.8564	0.8847	0.8506

Models (5) and (6) define High End to be any restaurant over $15. Models (7) and (8) exclude Other products. Models (9) and (10) exclude restaurants over $25. Robust standard errors are in parentheses. Coefficients that are significantly different from zero are denoted by the following system: *10%, **5%, and ***1%.

Table 4: Payment for Customization is Not Tied to the Customized Product if Demand is Volatile

	(1) Volatility		(2) Volatility		(3) Volatility		(4) Volatility	
	Low	High	Low	High	Low	High	Low	High
High End Restaurant	-2.325***	1.845***	-2.214***	1.972***	-2.277***	2.206***	-0.958	0.716***
	(0.424)	(0.299)	(0.425)	(0.306)	(0.447)	(0.327)	(0.630)	(0.268)
High End x Customized	2.971***	-0.763**	2.943***	-0.739**	2.924***	-1.124***	4.312***	0.103
	(0.483)	(0.341)	(0.481)	(0.339)	(0.503)	(0.365)	(0.845)	(0.301)
Item fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Agent Fixed Effects	No	No	Yes	Yes	No	No	No	No
Observations	15295	29677	15295	29677	14552	28405	3005	41967
R^2	0.9207	0.8688	0.9209	0.8690	0.9219	0.8708	0.9504	0.8825

Volatility in (1) through (3) is measured by demand for the top customizable products. Model (4) uses volatility of all customizable products. Model (3) excludes "Other" products. Robust standard errors are in parentheses. Coefficients that are significantly different from zero are denoted by the following system: *10%, **5%, and ***1%.

5.2 Payments that Sustain Relational Contracts are Not Tied to Volatile Customized Services

The High End Restaurants pay higher prices in general to sustain the relational contract that allows them to receive more service from the supplier. The requests for customization that I observe in this supplier dataset are detailed notes on fish weight and parts. For example, a restaurant may ask for the fish to be "Clean and White please" or for a specific size of fish ("15LBS back loin head part please !!!!!"), or even request how frozen the items should be ("BE Super Frozen Tuna Saku -1Bag BIGG" or "2lb not super frozen"). There is no general sense that any of these requests are improving the quality of the fish in a vertical sense (i.e. big tuna is not alway better than baby tuna, and super frozen fish is not always better than slightly frozen one) but rather the restaurants are requesting customizations based on their needs for the menu, for the special offerings, or maybe based on some customer preference for one type of dish over another. Based on these requests, I create a set of products that are customizable. Not all transactions of these goods are customized, but the variable "customized" is defined to be 1 if the good is ever customized.

The theoretical model predicts that to sustain the relational contract, payments for customization may be tied to non-volatile products. Volatility of the customized product is an important moderator in the model: the theoretical results predict that for stable demand of customized products, there is no restriction on placing the extra payment for effort directly on the price of the customized product. However, if the customized product has volatile demand, then the model predicts that the payment may not be tied to its price (as these prices would be too volatile to sustain a relationship and provide adequate effort levels). Thus, if the customized product is stable, then the payment could be tied to its price, but if the customized product is volatile, then the payment will not be tied to it.

In the dataset for restaurant supplies, I define a volatility index for each product-

restaurant pair by computing the standard deviation divided by the mean of the quantity ordered. Furthermore, at the restaurant level, I separate the restaurants into volatility tiers based on how volatile their demand for customized goods is.[13] The restaurants with the most volatile demand for customizable goods are considered "High Volatility" while restaurants that have a relatively stable volatility for customizable products are classified as "Low Volatility". For item i being sold to restaurant r, I estimate the empirical specification:

$$Price_{ir} = \gamma HighEnd_r \times Customizable_i + \alpha HighEnd_r + \beta Customizable_i + ItemGroup_i + \epsilon_{ir}.$$

Consistent with the theoretical model, Table 4 shows that the payments that sustain the relational contracts which assure customization for some products cannot be tied to the demand of customizable volatile products. Instead, these are linked to the payments of stable goods.

5.3 More Frequent Orders Make It Possible to Sustain Larger Payments on Customized Goods

Further analysis of the High End restaurants reveals that not all of these customers need to pay the premium payments on stable products in order to sustain the relationship with the supplier. This is not surprising: customers who interact more often with their suppliers are able to sustain a relationship more easily than those who do not. The theoretical model predicts that downstream customers who interact frequently with their supplier (those with high δ discount factor, who care about the future more) will be able to sustain the relationship more easily by paying a premium on the items they customize than those who have only ocasional interactions with the supplier. Intuitively, comparing two high end restaurants

[13]The Volatility Tiers are based on which quartile of volatility the restaurant's demand for customized goods is in: a restaurant who orders a very volatile customized good will be in Volatility Tier 4 while a restaurant whose customized goods demands are all in the lowest quartile of demand volatility will be in Volatility Tier 1.

who care about the relationship, but with one buying only once a week while the other one makes purchases every day, it becomes evident that the restaurant with the more frequent interaction will have more to gain from staying in the relationship. Thus, the restaurant with more frequent interactions will be able to pay some part of the price premium on the customizable good because the future value of its relationship with the supplier will be higher than the payment. The same is not true for the restaurant with once-a-week delivery, whose premium payment on the volatile, customizable good will be lower because the future value of its relationship is lower. Table 5 shows that restaurants who have more frequent delivery (and therefore a higher discount factor δ) are able to pay a premium for the customizable items and sustain the relationship with that premium.

Table 5: High End Restaurants: Relationship Strength is a Moderator

	All	Stable	Volatile	All	Stable	Volatile
Customized x Orders	1.468***	2.365***	1.309***	1.464***	2.175***	1.347***
	(0.196)	(0.366)	(0.249)	(0.201)	(0.358)	(0.257)
Customized	4.397***	2.112***	5.305***	4.485***	2.320***	5.322***
	(0.407)	(0.586)	(0.586)	(0.422)	(0.581)	(0.604)
Orders	-0.974***	-2.825***	-0.782***	-1.148***	-2.504***	-1.109***
	(0.191)	(0.336)	(0.245)	(0.190)	(0.325)	(0.245)
Item fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Agent Fixed Effects	No	No	No	Yes	Yes	Yes
Observations	31753	10331	21422	31753	10331	21422
R^2	0.8670	0.9250	0.8411	0.8675	0.9254	0.8417

Volatility is measured by demand for the top customizable products. Robust standard errors are in parentheses. Coefficients that are significantly different from zero are denoted by the following system: *10%, **5%, and ***1%.

I use the numbers of orders per week as a proxy for relationship strength in the specification for item i being sold to restaurant r:

$$Price_{ir} = \gamma Orders_r \times Customizable_i + \alpha Orders_r + \beta Customizable_i + ItemGroup_i + \epsilon_{ir}.$$

All restaurants order on average at least once a week, with some ordering more often. The

High End volatile restaurants are able to sustain the relationship by putting the price premium on the customizable items as their number of orders per week increases. This is consistent with the prediction that a higher δ will allow stronger relationships to be sustained despite volatile demand.

6 Conclusion

I develop a model of relational contracts to explain why high end downstream firms pay higher prices on the goods they purchase. I assume that these firms require a higher level of personalization for some of the items they purchase and that the added services are not ex-ante contractible through a formal contract. I prove that the payments that can sustain a relationship are linked to goods with stable demands, even if these goods are not customizable. Moreover, the relational contract provides uncontractible services for the customizable goods.

Empirically, I show that High End restaurants pay more on the products they purchase and make more requests for customization. Thus, I infer that the higher prices paid by High End restaurants are made to compensate and incentivize the supplier for providing the added services the restaurants require. If the customizable goods have volatile demands, then the payments sustaining the relational contract are based on goods with stable demand which cannot be customized.

I believe the findings of the empirical application show that the theoretical model of relational contracts is plausible and used in practice. More generally, I find that relational contracts are a useful tool to ensure appropriate levels of uncontractible quality to players that value this added quality dimension.

There are several limitations to my study. I do not explicitly model customer demand and retail prices, which could lead to double marginalization. Futher research is needed to determine if relational contracts can have a role in coordinating the channel to solve the

problem of double marginalization. It would also be interesting to extend the model to include multiple competing retailers who need relational contracts.

References

Abreu, D. (1988). On the theory of infinitely repeated games with discounting. *Econometrica 56*(2), 383–396.

Baker, G., R. Gibbons, and K. J. Murphy (2002). Relational contracts and the theory of the firm. *The Quarterly Journal of Economics 117*(1), pp. 39–84.

Bayus, B. L. and W. P. Putsis (1999). Product proliferation: An empirical analysis of product line determinants and market outcomes. *Marketing Science 18*(2), pp. 137–153.

Besanko, D., J.-P. Dubé, and S. Gupta (2003). Competitive price discrimination strategies in a vertical channel using aggregate retail data. *Management Science 49*(9), pp. 1121–1138.

Brown, M., A. Falk, and E. Fehr (2004). Relational contracts and the nature of market interactions. *Econometrica 72*(3), pp. 747–780.

Busse, M., J. Silva-Risso, and F. Zettelmeyer (2006). $1,000 cash back: The pass-through of auto manufacturer promotions. *The American Economic Review 96*(4), 1253–1270.

Cannon, J. P. and W. D. Perreault (1999). Buyer-seller relationships in business markets. *Journal of Marketing Research 36*(4), pp. 439–460.

Corbett, C. J., D. Zhou, and C. S. Tang (2004). Designing supply contracts: Contract type and information asymmetry. *Management Science 50*(4), pp. 550–559.

Dechenaux, E. and D. Kovenock (2007). Tacit collusion and capacity withholding in repeated uniform price auctions. *The RAND Journal of Economics 38*(4), pp. 1044–1069.

Desiraju, R. and S. Moorthy (1997). Managing a distribution channel under asymmetric information with performance requirements. *Management Science 43*(12), pp. 1628–1644.

Draganska, M., D. Klapper, and S. B. Villas-Boas (2010). A larger slice or a larger pie? an empirical investigation of bargaining power in the distribution channel. *Marketing Science 29*, 57–74.

Dukes, Geylani, and Srinivasan (2009). Strategic assortment reduction by a dominant retailer. *Marketing Science 28(2)*, 309–319.

Geylani, T., A. J. Dukes, and K. Srinivasan (2007, March). Strategic manufacturer response to a dominant retailer. *Marketing Science 26*, 164–178.

Ghosh, M., S. Dutta, and S. Stremersch (2006). Customizing complex products: When should the vendor take control? *Journal of Marketing Research 43*(4), 664–679.

Gibbons, R. (2005). Incentives between firms (and within). *Management Science 51*(1), pp. 2–17.

Guo, L. and G. Iyer (2010, May). Information acquisition and sharing in a vertical relationship. *Marketing Science 29*, 483–506.

Halac, M. (2011). Relational contracts and the value of relationships. *American Economic Review*. Forthcoming.

Ho, T.-H. and J. Zhang (2008). Designing pricing contracts for boundedly rational customers: Does the framing of the fixed fee matter? *Management Science 54*(4), 686–700.

Iyer, G. and J. M. Villas-Boas (2003). A bargaining theory of distribution channels. *Journal of Marketing Research 40*(1), pp. 80–100.

Jap, S. D. (1999). Pie-expansion efforts: Collaboration processes in buyer-supplier relationships. *Journal of Marketing Research 36*(4), 461–475.

Jeuland, A. P. and S. M. Shugan (1983). Managing channel profits. *Marketing Science 2*(3), pp. 239–272.

Kadiyali, V., P. Chintagunta, and N. Vilcassim (2000). Manufacturer-retailer channel interactions and implications for channel power: An empirical investigation of pricing in a local market. *Marketing Science 19*(2), pp. 127–148.

Kumar, N., L. K. Scheer, and J.-B. E. M. Steenkamp (1995). The effects of perceived interdependence on dealer attitudes. *Journal of Marketing Research 32*(3), pp. 348–356.

Kumar, N., L. K. Scheer, and J.-B. E. M. Steenkamp (1998). Interdependence, punitive capability, and the reciprocation of punitive actions in channel relationships. *Journal of Marketing Research 35*(2), pp. 225–235.

Lal, R. (1990). Price promotions: Limiting competitive encroachment. *Marketing Science 9*(3), pp. 247–262.

Levin, J. (2003). Relational incentive contracts. *The American Economic Review 93*(3), pp. 835–857.

Liu, Y. and Z. J. Zhang (2006). The benefits of personalized pricing in a channel. *Marketing Science 25*(1), pp. 97–105.

Luchs, R., T. Geylani, A. Dukes, and K. Srinivasan (2010, December). The end of the robinson - patman act? evidence from legal case data. *Management Science 56*, 2123–2133.

Narayandas, D. (1998). Arrow electronics, inc. *Harvard Business School Case* (598-022), 461–475.

Nijs, V., K. Misra, E. T. Anderson, K. Hansen, and L. Krishnamurthi (2009). Channel pass-through of trade promotions. *Marketing Science*, mksc.1090.0509.

Plambeck, E. L. and T. A. Taylor (2006). Partnership in a dynamic production system with unobservable actions and noncontractible output. *Management Science 52*(10), 1509–1527.

Rietz, T. A., R. M. Sheremeta, T. W. Shields, and V. L. Smith (2011). Transparency, efficiency and the distribution of economic welfare in pass-through investment trust games. Technical report.

Rotemberg, J. J. and G. Saloner (1986). A supergame-theoretic model of price wars during booms. *The American Economic Review 76*(3), pp. 390–407.

Schmalensee, R. (1989, Autumn). Good regulatory regimes. *RAND Journal of Economics 20*(3), 417–436.

Shaffer, G. and F. Zettelmeyer (2002). When good news about your rival is good for you: The effect of third-party information on the division of channel profits. *Marketing Science 21*(3), 273–293.

Shervani, T. A., G. Frazier, and G. Challagalla (2007). The moderating influence of firm market power on the transaction cost economics model: An empirical test in a forward channel integration context. *Strategic Management Journal 28*(6), 635–652.

Spekman, R. E., J. W. Kamauff, and N. Myhr (1998). An empirical investigation into supply chain management: A perspective on partnerships. *International Journal of Physical Distribution 38; Logistics Management 28*(8), 630–650.

Sudhir, K. (2001). Structural analysis of manufacturer pricing in the presence of a strategic retailer. *Marketing Science 20*(3), pp. 244–264.

Thomadsen, R. and P. Bhardwaj (2011). Cooperation in games with forgetfulness. *Management Science 57*(2), 363–375.

Tuli, K. R., S. G. Bharadwaj, and A. K. Kohli (2010). Ties that bind: The impact of multiple types of ties with a customer on sales growth and sales volatility. *Journal of Marketing Research 47*(1), 36–50.

Tuli, K. R., A. K. Kohli, and S. G. Bharadwaj (2007). Rethinking customer solutions: From product bundles to relational processes. *Journal of Marketing 71*(3), 1–17.

Villas-Boas, J. M. and Y. Zhao (2005). Retailer, manufacturers, and individual consumers: Modeling the supply side in the ketchup marketplace. *Journal of Marketing Research 42*(1), pp. 83–95.

Essay 2:
Third Party Marketing Approvals*

Abstract

When customers purchase new products, there is often a degree of uncertainty about their quality. A common solution is to rely on a third-party certifier to provide some form of accreditation that signals quality. However, the incentives of a third-party certifier may not be completely benign. Competitive certification markets may lead the certifiers to provide unduly positive evaluations of quality to gain market share or provide unduly negative evaluations in order to gain credibility with end-users. Based on competition models of certifier markets, this paper analyzes the consequences of allowing non-governmental parties to enter into the certification market for medical devices. We exploit an unusual natural experiment to evaluate the extent to which third-parties can be relied upon to correctly report product quality. We focus on the FDA's decision to allow third parties to prepare certifications for certain medical devices, and observe how this decision to introduce competition at the reviewer stage has affected the quality of products allowed to go to market. We find evidence that allowing third party certification leads to significantly lower product quality. However, we find that experience with using a third party reviewer in the past diminishes the negative effect of reviewer competition.

*I am especially indebted to my advisor, Catherine Tucker for comments, guidance and encouragement. Helpful comments were provided by seminar participants at MIT.

1 Introduction

When customers purchase new products there is a always a degree of uncertainty about their quality. One common option is to rely on a third-party 'certifier' or some form of accreditation to signal or disclose quality. However, one issue is that the incentives of a third-party certifier may not be completely benign. The certifiers may have to weight the benefit from reporting quality truthfully against losing market share to their competitors. We aim to determine empirically if competition at the certifier level leads to quality distortions in the end market.

In order for many innovations to be brought to market, they have to demonstrate that they meet certain quality thresholds. Certification is a review process that allows new products in the medical, bio and life-sciences to become available to the market. In such fields, certification is a key step in preparing to develop a product for the market, as years of work and research rely on proving the quality of the product and how it will benefit consumers. A common way of achieving this, especially in the fields of medicine, medical devices and biotechnology is to ensure the product is certified. Usually, in both the US and Europe this certification process is managed by the government. However, long delays has led the US to experiment with permitting some devices to be certified by private firms.

This paper analyzes the consequences of allowing competition in the certification market for medical devices. We exploit the FDA's decision to allow third parties to prepare certifications for certain medical devices and observe how this decision affects the quality of the new products that are ultimately used in the market. If we find that third party reviews are as high quality as the government's, this could speed up review processes and make it easier for new products to reach consumers in need. On the other hand, if we find that third party reviews are of lower quality, this could suggest that the government needs to continue playing a key gate-keeper role, and it is the only way for new firms to credibly prove their

quality.

The market for medical devices is regulated by the FDA, which acts as a testing and certifying entity before clearing medical devices for marketing approval. Due to the large number of medical devices and costly process involved in testing these devices, in 1997, the FDA Modernization Act allowed the FDA to start a program that lets accredited companies conduct their independent certification process and submit the results for final marketing approval to the FDA. In doing so, the FDA created a competitive certification market in which the FDA acts both as a player and as an overseer of the final decision for marketing of medical devices. The third party review program includes Class I and Class II devices, which are devices of lowest risk to the patient. On May 20, 1998, FDA issued a list of 147 types of low-to-intermediate risk devices eligible for third party review.[1] After 2001, when the pilot testing program was expanded, more than 670 types of devices became eligible to be certified by third parties. While some devices are still low risk, the expanded list includes devices such as surgical lasers, and medical imaging devices. The test program has proved successful in relieving some of the financial burden of the FDA, but concerns about the quality of the certification process have emerged at the same time.

On the one hand, third party certifiers have the advantage of short turn-around time for reviews, and a friendly review process. For example, Intertek, an accredited third party certifier, introduced in March 2010 a 10-day reviewing program and advertised "smoothest possible 510(k)submissions for its clients".[2] This program would allow the company producing the medical device to get to the market in 50 days instead of the 90 days necessary for FDA-led processing. The companies who develop the products can ask for certain tests to continue while others are still pending and can generally control the timeline of the certifi-

[1]Testimony on the Implementation of the FDA Modernization Act of 1997 by Michael A. Friedman, M.D., October 7, 1998.

[2]http://www.businesswire.com/news/home/20100325005045/en/Intertek-Introduces-10-Day-510-Review-Program

cation process, something that they are not able in a FDA-lead certification investigation.

On the other hand, it is unknown if the quality of the third-party reviews actually gets checked by the FDA after these are submitted. Reputation effects should take care of simple mistakes, but examples of final FDA approvals for third party reports that mirror exactly language used by the company in the original submission lead to concerns over how thorough the overseeing process really is.[3]

Despite an abundance of data collected by the FDA, there are no large scale analyses of the quality of the medical devices that received non-governmental forms of pre-market reviews. Most of the existing reports and analyses focus on the financial benefits of relieving the FDA of the certification reviews. Some reports analyze the number of third party prepared reviews to point to the test program as a success, as it has enabled more medical devices to be marketed in a shorter time.

We intend to fill the gap by looking at the outcomes of a natural experiment for the quality of the medical devices approved. Our work has direct regulatory impact: the test program is under question at the moment, following a few very public and tragic failures of radiology devices.[4] In the long term, established companies can rely on their reputation to weather spells of bad products. Thus, accurate certificate results benefit new companies, who have to prove that their products are high quality. For entrepreneurs, an unbiased and efficient certification market leads to easier entry and a fair chance at success. However, as competition among the certifiers increases, the quality of the reviews is affected by the incentive to be truthful and the drive to gain market share. Our work will quantify the result of these conflicting effects.

Section 2 discusses related literature. Section 3 presents the regulatory context of marketing medical devices. Section 4 describes the datasets used in this paper. Section 5 shows

[3]WSJ.com, Third-party reviews of devices come under scrutiny at the FDA, March 15, 2010
[4]NY Times Articles Series: Radiation Boom

the empirical results of our analysis. Section 6 proposes a mechanism that could explain the results. Section 7 concludes.

2 Related Literature

As far as we know, this is the first paper to test empirically the effects of allowing competition at the certifying stage for a product on the ultimate quality of the product. It is related to a growing stream of theoretical work on standards, forum shopping and certification markets.

Previous theoretical work in markets with information asymetries has analyzed the strategic behavior of certification intermediaries. Lizzeri (1999) considers markets in which certifiers are dependent on sellers for payment and are able to extract the full premium of providing information to the buyer market. In this model, quality is determined exogenously and the certifier's role is to provide a signal about quality to the buyer. He also shows that perfect competition at the intermediate stage of certification will lead to full disclosure and no profits for the certifiers, as the intermediaries compete the information rents away. In subsequent work, Albano and Lizzeri (2001) allow quality to be endogenous and analyze the certifier's strategic response: the intermediary's presence ensures that quality is non-zero but also manages to extract rent for providing quality information so the manufacturer has incentives to provide less than optimal quality because it does not reap the full benefits of providing optimal quality. This paper does not include certifier competition, which is an important part of the empirical results on the current paper.

Competition at the certifier level and its effects on incentive to disclose accurate information is modeled in Pesendorfer and Wolinsky (2003). They find that in the presence of price competition for reviewer services, there is an information externality which leads to an equilibrium in which certifiers rely on other's services and compete on low prices, thus generating low quality reviews. Unlike the present paper, Pesendorfer and Wolinsky (2003) does not consider the effects of certifier reputation on their incentives, but rather focuses on

a central planner intervention of setting minimum prices to achieve a second-best equilibrium of optimal quality disclosure.

Recent work on forum shopping and certifier competition (Lerner and Tirole 2006; Farhi et al. 2011) analyzes the case of competition at the third party certifyer stage and allows manufacturers to choose among different certifiers. This stream of research focuses on the tension between needing to be credible to the end-user market (and reputation concerns) with catering to the sellers of the product who pay for certification. Lerner and Tirole (2006) allow certifiers to differentiate in how friendly they are to the sellers and allow sellers to shop around strategically for the most suitable certifier. This is similar to the medical devices market, where the product sponsor has a choice of firms to use for eligible products. Farhi et al. (2011) focuss on tiered certification and transparency. For medical devices, the transparency level is set: if a review is not successful at the third party stage, it will not become public, which may lead manufacturers to shop around for the certifier who will give them a positive rating, even if it is not founded in true quality level. This model implies that competition at the certifier stage, followed by a rubber-stamp process at the FDA approval stage, might theoretically lead to lower quality products for the end user. Our estimates indicate this may be happening in the medical devices certification market. Moreover, Farhi et al. (2011) find that in markets where certifiers compete with each other over market share, they will decide to implement a shorter time review as a way to attract sponsors. Interestingly, we find that this race to yield reviews in a short time characterizes the market for certifying medical devices. Since review time seems to be a strong selling point for the certifiers in the Accredited Persons program, it is an indication that the third party reviewers are competing with each other over market share by decreasing certification time, rather than through fees.

A separate stream of research related to our paper includes empirical models of financial ratings or education evaluations. Empirical papers on financial certification of risk like

(Bolton et al. 2012) focus on multiple ratings for the same product and the incetive to shop around for seller in a market with both naive and experienced buyers. Becker and Milbourn (2010) find that the credit market becomes less accurate at certifying quality, and distorts quality upwards, when the number of certifiers increases. On the other hand, Doherty et al. (2011) claim that entry in the insurance ratings market leads to different stardards for ratings between the incumbent and the entrant: the entrant has higher standards for the same quality insurer and only higher-than-average quality insurers choose to receive a second rating from the entrant. Unlike these papers, we focus on how a change in certifier competition is reflected in product quality. In the natural experiment we exploit, the coarseness of the ratings, the transparency of the process and the standards of review are all determined before the competition is introduced in the market and we will consider these to be set throughout the period for our estimation. Instead, we focus on the endogenous quality of the products that get certified as a result of competition introduced at the certifier stage. Medical devices are particularly suited for this analysis, as their quality can be objectively measured by the number of adverse effects they generate.

Unlike education outcomes or financial ratings, whose expectations may influence the ultimate quality perception of the product at the consumer end, medical products will function properly in expectation unless their actual quality is less than what it was certified to be. Thus, the results from the natural experiment for medical device may be cautiosly extended to other products that need certification.

This paper is also relevant to a growing stream of research in marketing which analyzes information asymmetries using differences in differences methods (Busse et al. 2006). The literature on disclosure of payment or gifts (Dana and Loewenstein 2003; Malmedier and Schmidt 2011) sheds some light on the tension between maintaining a reputation for truthful reviews and payment maximization, feelings of reciprocity or fairness. Our paper assumes that a product cannot be marketed without certification of an appropriate level of quality

and does not focus on shading the process of reviewing the product or the fees associated with it. Instead, we focus on competition at the reviewer stage and strategic reviewer's response to it.

Recent literature has focused on heathcare IT in order to analyze how adoption of new technology can save lives (Miller and Tucker 2011). Finally, this paper is related to Thirumalai and Sinha (2011) which shows the role of learning and how recalls for medical devices are incorporated into stock prices. Thirumalai and Sinha (2011) does not address the certification part of marketing medical devices but does analyze the consequences of bad quality (as reflected in recalls) on stock prices for the manufacturer. They find that stock prices do not react enough to incorporate the new recall information, which may indicate that the market does not penalize bad quality enough to deter manufacturers from marketing unsafe products. However, the authors find evidence of learning, with firms that had previous recalls being less likely to have future recalls. Our paper includes more granular data on adverse effects in order to study unsafe products: we look at adverse effects that happens in a few years after the product is marketed. Moreover, we are interested on whether introducing competition at the certifier stage makes products less safe, regardless of the financial outcomes of these adverse effects. We find that certifier competition does not lead to better product quality in this market.

3 Medical Devices Certification Process

3.1 FDA Approves Medical Devices for Marketing

This section summarizes the history of medical devices and their certification in the US. We used the FDA website and FDA reports as our main source of information[5] to describe the evolution of the certification process for medical devices as it relates to our analysis.

Medical devices used in the United States are approved for marketing by the Food and

[5]http://www.fda.gov/MedicalDevices/default.htm

Table 1: Device class and Regulatory Controls

Device Class	
Class I General Controls	With Exemptions
	Without Exemptions
Class II General Controls and Special Controls	With Exemptions
	Without Exemptions
Class III General Controls and Premarket Approval	

Drug Administration (FDA). Medical devices range in complexity and use, from the simplest like tongue depressors to very complex machines like x-rays or lasers. These devices are classified in three groups (regulatory classes) shown in Table 1 based on the level of control needed to ensure that the device is safe for use. The classification depends on how the device will be used which is usually described in the labeling of the product. The classification is also risk-based, with devices that are riskiest for the patient or the users being classified as higher class.

Regulatory controls for premarketing submissions and marketing clearing depend on the class of the device. Some devices, usually Class I devices, do not require a premarketing approval. However, if a device is not exempt from premarketing approval, then it requires submission of a 510(k) or PMA application, which the FDA reviews and approves or denies. We will focus on devices that require a 510(k) approval in order to be marketed.

Most Class I devices do not require a 510(k) submission (they are exempt from premarketing approval though they must comply with other regulatory standards of registration/listing, production and labeling), though devices without exemptions will require one. Class II devices without exemptions require a 510k for approval. Almost all of 510(k) eligible devices are Class II devices. Class III devices typically require a PMA unless the device was approved before 1976 or is equivalent to a device approved before then (these are known as *preamendament* devices). If the device is a preamendament device, then a 510(k) is required. Our analysis focuses on Class II devices because these are ones that make up the majority

of third party certification devices.

A *510(k)* is a type of premarket submission made to the FDA to prove that the device is at least as safe and effective as an existing legally marketed device (that is not subject to a PMA).[6] This is know as showing that the device is *substantially equivalent (SE)* to an existing device that is already approved, called a *predicate.* The device can be marketed only after being declared to be SE. The FDA claims that SE determination is done in less than 90 days and is based on information given by the submitter of the device. If the FDA determines that the device is not SE, then the submitter can resubmit another 510(k) with new data, submit a PMA or try to reclassify the device. In 1997 the FDA started a program to allow Third Party Review for conduction the primary review of the 510(k). We will describe the evolution of this program in more detail in Section 3.3. In this paper, we focus on devices which are elligible for 510(k) approvals because those are the only types of devices that can be certified by third parties. We will not consider PMA approvals.

There are three types of Premarket Notification 510(k)'s that may be submitted to FDA: Traditional, Special, and Abbreviated. The Traditional review may be used for any eligible device. The Special and Abbreviated 510(k) methods were developed under the "New 510(k) Paradigm" to help streamline the 510(k) review process. The Special review may be used for a modification to a device that has already been approved through the 510(k) process, if the modification does not affect the intended use of the device or affect the fundamental scientific technology of the device. The Special 510(k) allows the manufacturer to declare

[6] A *PMA* is a Premarket Approval through which the FDA determines if a Class III device is safe and effective. Class III devices are riskier than other devices because they "support or sustain human life, are of substantial importance in preventing impairment of human health, or which present a potential, unreasonable risk of illness or injury". Thus additional checks are needed to approve these devices, and the PMA is the strictest type of regulatory approval process the FDA uses. It includes technical sections with non-clinical laboratory studies and clinical investigations. The PMA is held to high standards of scientific research, as these are meant to approve risky devices, many of which are completely new or have new elements. The FDA has a goal to review and make a determination on PMA applications under 180 days but even in 2012 the actual time needed for these is longer. The Third Party Program does not apply to PMA applications, and the FDA is the only authority for these, so we will not use PMA applications in our analysis.

conformance to design controls without providing the data. Special reviews have a shorter timeline than Traditional reviews: the FDA aims to review and make a decision on Special reviews in 30 days. The Abbreviated review is used when: a guidance document exists, a special control has been established or FDA has recognized a relevant consensus standard. Very few devices that are eligible for a 510(k) review are Special or Abbreviated. Most 510(k) are Traditional reviews. We will use both Traditional and Special types of devices in our analysis. We will focus on Traditional reviews because these are the ones that are used mostly, and that benefit most from third party certification in terms of time savings before being able to be marketed. However, we also use Special reviews as a falsification group: we check whether there is an effect of third party certification for this class of devices.

3.2 Reporting Adverse Effects and Malfunctions

The devices are used by hospitals, nursing homes etc and research facilities. These are required to report malfunctions of the device, but reporting rules vary depending on the seriousness of the incident. If a death occurs, then user facilities report the event to the FDA. However, if a serious injury occurs, then the rules state that the report needs to be done only to the manufacturer. If the manufacturer is unknown, then the facility makes the report to the FDA. In addition, the user facilities submit a semi-annual report to the FDA containg a short list of all the adverse reports they made during that period. This report standard for user facilities is known as the Safe Medical Devices Act of 1990, which started being enforced in November 1991. In 1998, the Food and Drug Administration Modernization Act (FDAMA) made a few changes to the reporting rules, including an annual report of to the FDA summarizing all adverse event reports hospitals filed during that year (instead of the previous semi-annual reports).

In addition to user facilities, there are other parties that are required to file adverse event reports: distributor and manufacturer. Voluntary submitters are also allowed to report

adverse events. All these types of reports are compiled by the FDA and made available to the public on the FDA website. We use all types of reports in our analysis. However, because a single event may be reported by several agents (for example, in a manufacturer report, a user facility report and a voluntary report), we aggregate the reports and use only one instance for each single event. The process of matching the reports to a device is quite difficult, and we tried to keep the most amount of information about the report and the device involved, regardless of the source of the report (for example, if the year of the event was missing in one version of the report but not the device type, then we combined the two reports into a complete one and used it for our analysis).

3.3 Third Party Certification for Medical Devices

In 1997, the FDA implemented FDA Modernization Act of 1997 (FDAMA), which included the *Accredited Persons Program.* This program was based on a pilot run previous by the FDA in which other organizations were allowed to prepare the certification reviews for 510(k) applications. The short pilot ended and was replaced with a longer, more permanent test program in which third parties were allowed to become accredited persons to participate in the certification of 510(k) applications. The accredited person's program allowed third parties to certify eligible devices only, by preparing the review for the 510(k) and then submitting it, along with its recommendation, for direct approval to the FDA. The law also specifies that the FDA must then make a decision on the review in 30 days, which is quite short compared to the usual 90 days required for a direct submission to the FDA.

The Accredited Persons Program also specifies the process through which a company can become accredited for preparing 510(k).

The program does not make it mandatory for all manufacturers of eligible devices to use Accredited Persons: manufacturers can submit 510(k) applications through the regular process to the FDA as well. Thus, comparing failure rates of devices certified by third parties

to FDA devices could lead to biased results because of selection effects at the manufacturer level. Section 5 describes in detail how we control for selection effects.

There are no additional fees for reviews submitted by Accredited Persons, but third parties are free to negociate and set their own fees with the manufacturer as there is no FDA guidance about those fees. The FDA charges a small fee for reviews that use FDA certification: in 2012 the standard fee is $4,049 (small businesses with less than $100 million in gross receipts or sales have a reduced fee of $2,024).

The FDMA was initially quite restrictive in specifying eligible devices. In 2001, the FDA launched an Expansion Pilot, which allowed Accredited persons to review many more Class II devices that they could previously. Broadly, many Class II devices that had no device-specific guidance documents became eligible for Third Party Review in 2001. This created a natural experiment for competition at the certifier stage. Table2 lists the changes in FDA policies over time. We exploit this fortunate law change in our analysis: we use the eligibility for the pilot program, the post policy period and the third party eligibility to study whether introducing competition among certifiers leads to products of different quality levels than those certified by the FDA absent the policy change.

510(k) reviews are meant to show that a device is safe, effective and equivalent to an existing and already approved device (predicate). The Third Party Review is supposed to be similar in quality to an FDA review. The FDA receives two copies of the review, along with additional information, review documentation and recommendation from the third party reviewer. The third party thus takes over the role of performing the time-consuming primary review of the product and the FDA is meant to get an accurate review without actually performing it. The FDA, after reviewing the report, can either approve (clear the device for marketing) or request additional information. The third party then passes along the request for information to the device sponsor, who can then provide it or explain why it thinks it is not needed. If additional information is requested, the 510(k) is

Table 2: Change in FDA Policies over time (Source: www.fda.gov)

Date	Policy Change
August 1996	The third-party review pilot program began with just 15 device types eligible
November 1997	The President signed FDAMA into law. The number of Eligible Devices is 154 (GAO)
September 1998	FDA publishes list of Accredited Persons eligible for the FDAMA.
November 1998	The agency began accepting 510(k) reviews from Accredited Persons, and terminated the Third Party Review Pilot Program that began on August 1, 1996.
March 2001	Devices otherwise eligible for third-party review, but for which a specific guidance document does not exist, are designated as Expansion Pilot devices. For these products, the third party is required to contact FDA prior to initiating its review, to discuss specific areas that the reviewer should address. This contact is not required before a third party reviews subsequent submissions for the same product category. (GAO: FDA expanded the program to include more than 670 class I and class II device types to be eligible for 510(k) review by a third party)
October 2003	FDA increased review fee from $2187 to $3480 ($2784 for small businesses). Third-party submissions remain exempt from the fee.

placed under a 30-day hold (this is true even if the primary review is done by the FDA). For devices using third party reviews, after information is provided, these devices are still considered on the fast timeline of a third party review, which can mean significant time savings for the manufacturer.

In addition to being on a fast-track to approval once it reaches the FDA, the third party certification process has other benefits for the sponsors. The accredited firms claim to have a friendlier process, with the manufacturer having input and feedback into the review process. This is very different from a traditional FDA review where there is no feedback until the end of the process. In addition, third party certifiers attract manufacturers (and compete with each other) on the time it takes to prepare a review. For example, in 2010 Intertek launched a "10-Day 510(k) Review Program - Real-Time Review" for the primary review and submission to the FDA. They claimed that the service will considerably shorten the time to market and raise the probability of a successful application on a first pass. This is indeed a significant time saving for the sponsor: the device will take 40 total days to get be marketed compared to the usual 90 days that would be needed if the manufacturer used the FDA for the primary review.[7]

The companies approved by the FDA as Accredited Persons have specialized expertise in some area of device testing, standards, or foreign regulatory requirements which may help the sponsors during the primary review process, as there is more information sharing and collaboration that would usually occur in a FDA-led primary review.

There are 10 companies approved as Accredited Organizations list in Table 3 (most current list as of 2012). These are approved for different types of devices: not all devices may be reviewed by all companies. The third party companies are located in the US and internationally, which may also be a consideration for foreign manufacturers and importers.

[7]http://www.intertek.com/news/2010/03-25-10-day-510k-review-program

Table 3: Third Parties: Accredited Organizations[8]

Name	Location
British Standards Institution	United Kingdom
Center for Measurement Standards, ITRI	Republic Of China (Taiwan)
Dekra Certification B.V.	Netherlands
Intertek Testing Services	USA
NIOM Scand. Inst. of Dental Materials	Norway
Pharmalink Technical Group, LLC	USA
Regulatory Technology Services, LLC	USA
TUV Rheinland of North America, Inc.	USA
TUV SUD America, Inc.	USA
Underwriters Laboratories, Inc.	USA

4 Data On Medical Device Certification and Post-Approval Performance

4.1 FDA Databases are used for analysis

We use two separate databases that span 12 years to perform an in-depth analysis of the third party certification program and its results.

- The PreMarket Notifications (510(k)) Database records what medical devices were certified and whether or not they were certified by the government or a private party, along with other specification about the device and the manufacturer.

- The MAUDE Database records all instances of adverse medical events connected with the use of a medical device. These reports are either voluntary reports, user facility reports, distributor reports, and manufacturer reports, spanning a large universe of details and seriousness of the averge events.

We restrict the analysis to devices that need 510(k) approvals in ordered to be marketed. We use a differences-in-differences approach to analyze whether the introduction of the program allowing non-governmental pre-market approval of medical devices has changed the overall

quality of the products in this market.

4.2 Which products choose third party review

Third party review is available for a list of eligible devices. However, not all devices eligible benefit in the same way from choosing a third party review over a standard FDA review. In particular, products that are more complex or expensive, and for which potential sales are higher will benefit more from going to the market earlier. Additionally, there are types of products for which FDA review times are especially long, depending on which FDA branch reviews these products.

Not all eligible devices benefit from third party certification. Mass-produced simple devices usually do not have enough sales to recouperate the third party fee, in which case the manufacturer is more likely to submit an FDA review, for which fees are lower. Moreover, devices eligible for Special 510(k) submissions are guaranteed a 30-day response from the FDA, which negates most of the time saving from using third parties. We use these Special submission as a control group, since we do not expect them to benefit from third party certification.

Table 4 shows the percent of products eligible for Third Party Review before the Pilot Program while Table 5 shows the percent of products eligible after the introduction of the Pilot Program. It seems that it takes time for the manufacturers to start taking advantage of the new eligible categories of products, so there is a slow start to the number of products that are actually reviewed by Third Parties in 2001 and 2002. In effect, we will use 2003 as the year when implementation of the pilot becomes widespread in the certification stage. Thus our difference-in-difference estimation will focus on changes in product quality before and after 2003, which we take to be exogenous and the source of our exogenous variation.

Table 6 shows the summary statistics for the period 1997 to 2005, which is used in the main analysis. We focuse on five years of adverse effects after each device is marketed. Thus,

Table 4: Proportion of Third Party Reviews for products already eligible before Pilot Program

Year	FDA Review	Third Party Review	Total
1996	100.00	0.00	100.00
1997	99.07	0.93	100.00
1998	99.18	0.82	100.00
1999	97.46	2.54	100.00
2000	96.70	3.30	100.00
2001	93.07	6.93	100.00
2002	93.80	6.20	100.00
2003	89.70	10.30	100.00
2004	85.36	14.64	100.00
2005	85.91	14.09	100.00
2006	84.89	15.11	100.00
2007	83.99	16.01	100.00
2008	79.04	20.96	100.00
2009	83.09	16.91	100.00
2010	87.89	12.11	100.00
Total	91.71	8.29	100.00

Table 5: Proportion of Third Party Reviews for products eligible in the Pilot Program

Year	FDA Review	Third Party Review	Total
1996	100.00	0.00	100.00
1997	100.00	0.00	100.00
1998	100.00	0.00	100.00
1999	100.00	0.00	100.00
2000	100.00	0.00	100.00
2001	98.76	1.24	100.00
2002	97.90	2.10	100.00
2003	95.23	4.77	100.00
2004	93.27	6.73	100.00
2005	91.54	8.46	100.00
2006	91.85	8.15	100.00
2007	91.81	8.19	100.00
2008	90.80	9.20	100.00
2009	91.48	8.52	100.00
2010	94.55	5.45	100.00
Total	96.17	3.83	100.00

Table 6: Summary Statistics: Devices from 1998 to 2005

	Mean	Std.Dev.	Min	Max	Observations
Adverse Events	0.17	0.71	0	5	17200
Indicator	0.07	0.26	0	1	17200
Third Party Review	0.05	0.21	0	1	17200
Post Policy	0.37	0.48	0	1	17200
Eligible for Pilot	0.29	0.45	0	1	17200
Average Time By Committee	108.65	17.15	76	318	17200

Table 7: Differences in Differences: Adverse Events for 1998 to 2005

		FDA Review	Third Party	
	Difference	Mean	Mean	T-Test
Adverse Effects Before Pilot	-.0105	.1925	.2029	-0.2282
Adverse Effects After Pilot	.04595	.1335	.0875	1.5827

devices approved starting in 2006 would have fewer adverse effects as a result of having fewer years of observation in the dataset.[9]

As some initial evidence of the effect of the policy implemented in the Pilot program, Table 7 indicates that there is a decrease in the incidence of adverse effects for both FDA reviews and Third Party reviews after the policy change; furthermore, this decrease is larger for third party reviews than for FDA reviews. This result seems promising, but we worry about selection into the pilot program: even though some devices are eligible for it, it may be that the manufacturer decides not to send the devices to a third party.

Table 8 shows the differences between devices not elligible for the pilot program and those in the pilot, before and after the policy is implemented. The incidence or adverse effects decreases for both groups, but the decrease is smaller for products eligible for the pilot program. Although the incidence differences between devices in the pilot program and those not in the pilot program are not statistically significant for these aggregate numbers, the fact that those eligible for the pilot show less improvement might suggests the pilot program

[9]The results of the analysis are robust if instead we limit the adverse events to the first two years of adverse effects and include devices marketed before 2009.

Table 8: Differences in Differences: Adverse Events for 1998 to 2005

		Not In Pilot	In Pilot	
	Difference	Mean	Mean	T-Test
Adverse Effects Before Pilot	.0149	.1971	.1822	0.9473
Adverse Effects After Pilot	-.0084	.1272	.1357	-0.4719

is detrimental to the overall safety level of devices that are approved. This comparison is a very raw estimate, and we pursue the analysis further by including various fixed effects for each product type in the next section.

5 Empirical Analysis

In this section we deepen the analysis to control for fixed effects for the product categories and years in the sample. More importantly, we try to see if there is a selection problem with the devices that manufacturers choose to send to third parties for review.

We use the number of years of adverse effects in the first five years after the device is approved as the dependent variable for the main empirical analysis. Table 13 shows our main results. We use the type of approval to restrict the analysis to "Traditional" devices, because we expect that the effect is strongest for the devices which take a long time to get approved under FDA review. These devices are classified by the FDA as "Traditional" devices.

As a first start, we look at a simple regression of adverse events on the third party status of the review shown in Table 9. Model (1) indicates that we cannot attribute the change in quality to third party reviews, and Model (2) adds approval year and product category fixed effects. However, these initial results are not enough to conclude that third party reviews are just as safe as FDA reviews. In particular, it could be that manufacturers could pick and choose the products to send to the third party review, either because they are trying to game the system and sneak bad products on the market (if they are picking the bad products)

Table 9: Raw Effects May Have Selection Bias

	(1) Adv. Events	(2) Adv. Events
Third Party Review	-0.0209	0.0175
	(0.0246)	(0.0256)
Approval Year Fixed Effects	No	Yes
Type of Product Fixed Effects	No	Yes
Observations	13556	13556

Dependent variable is the number of years with a bad event in the first 5 years after approval for a device category.

or that they trying to get their safest products to the market faster in order to get higher profits from sales faster (if they are choosing the safest products). In either case, we worry about selection and want to use the natural experiment given by the expansion of the pilot program.

Models (1) through (4) in Table 10 show the results of difference-in-difference specification to determine if products certified by third parties are more or less safe after the pilot program is introduced compared to the period before the pilot and to other products certified by the FDA. The introduction of the pilot program created two types of products: devices which are eligible for the pilot and those that are not. We use the devices that are not eligible for the pilot as a control group. We compute the "reduced form effect of intention to treat" as in Angrist and Lavy (2009) as this is the effect that captures the introduction of the pilot program without being affected by the endogenous selection of manufacturers choosing to use third party reviews. The policy change can be considered exogenous: the FDA implemented it because of lack of funds and limited man-power, which is not related to final product quality outcome. Moreover, while the FDA did choose the products which would be eligible for the pilot, these are comparable to other Class II products which are not eligible and there is no sense that the FDA picked safer products to be eligible[10].

The analysis is restricted from 1998 to 2008 in order to account for the fact that devices marketed later than 2008 would have fewer than two years of adverse events reported, which could skew the results. *Pilot Eligible* is a dummy variable with 1 for a product that is eligible

[10]Note that if that were the case, then our estimates would be biased towards showing that the pilot produced safer products, which is not the case in our findings.

to be part of the expansion program. This does not include devices that were alredy eligible for third party review before the pilot program was implemented. *Post Policy* is a dummy variable with value 1 for years after the pilot program became implemented: after 2003 third parties started reviewing products eligible in the pilot program. For product i which may be part of the Pilot Program j and was approved in year t (with year 2003 being after the pilot is implemented), the number of adverse effects reflects:

$$AdverseEffect_{ijt} = \alpha + \alpha_1 PilotEligible_j + \alpha_2 PostPolicy_t + \beta_1 \times Pilot_{ij} \times PostPolicy_t + \theta_j + \tau_t + \epsilon_{ijt} \quad (1)$$

Thus, β_1 represents the core coefficient we focus on: whether the pilot program (which made more devices eligible for Third Party certification) had an effect on the adverse effects after the devices were marketed. Coefficient α_1 captures the main effect of a device being Pilot Eligible while coefficient α_2 indicates the effect of the change in Policy as a shift in time. α is a constant term, while θ is a vector of product category fixed effects and τ represents year fixed effects.

Tables 10 indicates that third party certification increases the potential bad events after the device is marketed. Columns (1) through (4) use an indicator dependent variable: it indicates at least one adverse effect in the two observation years after the device is approved. Since the dependent variable is a 0-1 indicator, we use a linear probability model in our main specification. Column (1) indicates the estimates with no fixed effects, column (2) adds year fixed effects while column (3) product category fixed effects and column (4) adds both year and product category fixed effects.

Column (1) shows the main specification in Equation (1). It suggests that devices which are part of Pilot program and are approved after the implementation of the Policy have a higher rate of adverse events than FDA devices. This is a raw difference and does not account for specifics about the product or the time of approval. Here, we can imagine that different products may have different acceptable rates of malfunctions and thus adding

Table 10: Differences in Differences: At Least One Adverse Event in First Two Years

	(1) LPM	(2) LPM	(3) LPM	(4) LPM
Pilot x Post Policy	0.0132**	0.0132**	0.0122**	0.0122**
	(0.00530)	(0.00530)	(0.00529)	(0.00529)
Eligible For Pilot	-0.00792*	-0.00788*	-0.0120***	-0.0119***
	(0.00417)	(0.00417)	(0.00413)	(0.00413)
Post Policy	-0.0226***	-0.0371***	-0.0232***	-0.0387***
	(0.00290)	(0.00476)	(0.00290)	(0.00477)
Approval Year Fixed Effects	No	Yes	No	Yes
Type of Product Fixed Effects	No	No	Yes	Yes
Observations	17782	17782	17782	17782

product category fixed effects is important to control for this counfound. Similarly, there could be a time trend in our data that could make all products safer or less safe in each approval year, if there are unrelated quality standards being implemented. Thus we will also use year of approval fixed effects. Column (2) ads Approval Year fixed effects, column (3) adds Product Type fixed effects and Column (4) has both types of fixed effects. Together, these models indicate that the quality of products decreases after the implementation of the pilot program.

Table 11 checks for robustness of Table 10 to different specifications. Columns (1) through (5) use the same variable as Table 10: it indicates at least one adverse effect in the two observation years after the device is approved. Since the dependent variable is a 0-1 indicator, we use a logit model in our main specification. Column (1) indicates the estimates of a logit model with no fixed effects, column (2) adds year fixed effects while column (3) adds both year and product category fixed effects. The data fits a Poisson distribution: the mean and variance are equal and thus we try to see whether a zero inflated poisson model may be more appropriate to account for some of the zero observations in adverse effects. Column (4) fits a zero inflated poisson model with no fixed effects while column (5) adds approval year and product category fixed effects.

Column (1) suggests that devices which are part of Pilot program and are approved after the implementation of the Policy have a higher rate of adverse events than FDA devices. This is a raw difference and does not account for specifics about the product or the time of

Table 11: 1998 to 2008: At Least One Adverse Event in First Two Years

	(1) Glm Logit	(2) Glm Logit	(3) Glm Logit	(4) Poisson Zero Inf.	(5) Poisson Zero Inf.
Pilot x Post Policy	0.530**	0.533**	0.501**	1.059***	0.901***
	(0.215)	(0.215)	(0.217)	(0.299)	(0.296)
Pilot Eligible	-0.237*	-0.236*	-0.415***	-0.409***	-0.430***
	(0.131)	(0.131)	(0.142)	(0.143)	(0.143)
Post Policy	-0.903***	-2.395***	-2.466***	-0.851***	-2.318***
	(0.119)	(0.430)	(0.431)	(0.120)	(0.390)
inflate					
PilotEligible after Policy				0.649**	0.460*
				(0.259)	(0.259)
Approval Year Fixed Effects	No	Yes	Yes	No	Yes
Product Category Fixed Effects	No	No	Yes	Yes	Yes
N	17782	17782	17782	17782	17782
Log-likelihood	-2170.0	-2139.2	-2001.5	-2355.9	-2326.9
Wald χ^2	62.98	86.96	256.2	307.4	365.4

The dependent variable for Models (1)-(5) indicates at least one bad event in the observation years.

approval. Column (2) adds Approval Year fixed effects, column (3) adds Product Type fixed effects as well. Together, these models indicate that the quality of products decreases after the implementation of the pilot program.

Columns (4) and (5) cosider the possibility that a logistic model may not be the best discription of our data process. Our dataset has a large number of products which have zero adverse effects after introduction. This could be because they are similar to other products but actually safer to use, or it could be that their adverse effects are underreported because they are not as severe as those for other products. To account for the option of two processes creating the zero observations, we use a zero-inflated poisson specification. The results are listed in columns (4) and (5). For the inflation factor, we use a dummy variable which is 1 if the product is in the pilot program after the introduction of the pilot, and zero otherwise. Column (4) includes year fixed effects while column (5) adds product category fixed effects.

The results in Table 11 suggest that products become less safe for consumers after the pilot program is implemented. We check the robustness of this finding by using a different measure of adverse events, by adding several fixed effects and by using a different time frame for the adverse events. Tables 12 and Table 13 show robustness of the results in Table 11.

Table 12: 1998 to 2008: Number of Years with Adverse Effects Two Years After Marketing

	(1) Glm Neg. Binomial	(2) Glm Neg. Binomial	(3) Neg. Binomial	(4) Glm Var. from Nbreg	(5) Neg. Binomial Zero Inf.
Pilot x Post Policy	0.539***	0.513***	0.569**	0.569**	0.755**
	(0.190)	(0.192)	(0.225)	(0.225)	(0.300)
Pilot Eligible	-0.203*	-0.411***	-0.457***	-0.457***	-0.168
	(0.116)	(0.125)	(0.151)	(0.151)	(0.135)
Post Policy	-2.311***	-2.364***	-2.397***	-2.397***	-2.254***
	(0.374)	(0.375)	(0.395)	(0.395)	(0.391)
inflate					
PilotEligible after Policy					0.298
					(0.256)
Approval Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
Product Category Fixed Effects	No	Yes	Yes	Yes	No
N	17782	17782	17782	17782	17782
Log-likelihood	-2606.0	-2431.2	-2342.0	-2342.0	-2458.8
Wald χ^2	101.9	310.4	379.3	252.5	101.4

Dependent variable is the number of years with a bad event in the first 2 years after approval for a device category.

Table 12 checks for robustness of using a different measure of adverse effects: the dependent variable is the number of years with a malfunction in the first two years a device is marketed.

Column (1) fits a negative binomial model with approval year fixed effects, with model (2) adding product fixed effects to that specification. Columns (3) and (4) are equivalent because model (4) uses the variance from model (3) to fit the GLM specification of a negative binomial. Column (5) checks whether the estimates change if we fit a Zero Inflated Negative Binomial model (the data is overdispersed, with the variance almost three times the mean, which indicates that a negative binomial distribution is more appropriate than a poisson distribution).

Table 13 checks for robustness of specifying the timeframe of adverse effects: the dependent variable is the number of years with a malfunction in the first five years a device is marketed. Note that this timeframe for measuring adverse effects implies we can include fewer devices, those approved between between 1998 and 2005 into our analysis, because any device approved after 2005 will have fewer than five years when we can observe malfunctions.

Table 13 indicates that third party certification increases the potential bad events after

the device is marketed. Columns (1) fits a negative binomial model, with models (2) and (3) adding fixed effects to that specification. Columns (4) and (5) are equivalent because model (5) uses the variance from model (4) to fit the GLM specification of a negative binomial. Column (6) uses the GLM model adjusted for Pearson's $\chi 2$. Columns (7) and (8) check whether the estimates change if we fit a Zero Inflated Negative Binomial model.

Column (1) shows the main specification in Equation (1). It suggests that devices which are part of Pilot program and are approved after the implementation of the Policy have a higher rate of adverse events than FDA devices. This is a raw difference and does not account for specifics about the product or the time of approval. Here, we can imagine that different products may have different acceptable rates of malfunctions and thus adding product category fixed effects is important to control for this counfound. Similarly, there could be a time trend in our data that could make all products safer or less safe in each approval year, if there are unrelated quality standards being implemented. Thus we will also use year of approval fixed effects. Column (2) ads Approval Year fixed effects, column (3) ads Product Type fixed effects and column (4) adds both types of fixed effects. Together, these models indicate that the quality of products decreases after the implementation of the pilot program.

Finally, columns (7) and (8) cosider the possibility that a binomial distribution may not be the best discription of our data process. Our dataset has a large number of products which have zero adverse effects after introduction. This could be because they are similar to other products but actually safer to use, or it could be that their adverse effects are underreported because they are not as severe as those for other products. To account for the option of two processes creating the zero observations, we use a zero-inflated negative binomial specification. The results are listed in columns (7) and (8). For the inflation factor, we use a dummy variable which is 1 if the product is in the pilot program after the introduction of the pilot, and zero otherwise. Column (8) adds fixed effects to the bare-bones

Table 13: 1998 to 2005: Number of Years with Adverse Effects Five Years After Marketing

	(1) Glm Neg. Bin.	(2) Glm Neg. Bin.	(3) Glm Neg. Bin.	(4) Neg. Bin. Neg. Bin.	(5) Glm Var. Nbreg	(6) Glm Var. Nbreg x2	(7) Neg. Bin. Zero Inf.	(8) Neg. Bin. Zero Inf.
Pilot x Post Policy	0.347***	0.355***	0.330***	0.483**	0.483**	0.483**	0.343**	0.586***
	(0.124)	(0.124)	(0.117)	(0.214)	(0.214)	(0.207)	(0.163)	(0.177)
Pilot Eligible	-0.312***	-0.312***	-0.453***	-0.518***	-0.518***	-0.518***	-0.152*	-0.241***
	(0.0650)	(0.0650)	(0.0650)	(0.129)	(0.129)	(0.126)	(0.0797)	(0.0931)
Post Policy	-0.587***	-1.108***	-1.157***	-1.298***	-1.298***	-1.298***	-0.131	-0.796***
	(0.0651)	(0.125)	(0.120)	(0.205)	(0.205)	(0.199)	(0.0901)	(0.170)
inflate								
PilotEligible after Policy							0.505***	0.639***
							(0.148)	(0.153)
Approval Year Fixed Effects	No	Yes	Yes	Yes	Yes	Yes	No	Yes
Product Category Fixed Effects	No	No	Yes	Yes	Yes	Yes	No	Yes
N	13556	13556	13556	13556	13556	13556	13556	13556
Log-likelihood	-5925.7	-5909.0	-6404.1	-4443.4	-4443.4	-4443.4	-4533.9	-4383.8
Wald χ^2	105.6	127.3	748.1	369.7	300.1	300.1	5.741	306.0

Dependent variable is the number of years with a bad event in the first 5 years after approval for a device category.

Table 14: Differences in Differences Specification for Falsification Group: Special Type

	(1) LPM	(2) LPM	(3) LPM	(4) LPM
Pilot x Post Policy	-0.0444*	-0.0415*	-0.0395	-0.0358
	(0.0244)	(0.0244)	(0.0241)	(0.0241)
Eligible For Pilot	0.0665***	0.0637***	0.0580***	0.0540***
	(0.0168)	(0.0169)	(0.0173)	(0.0173)
Post Policy	-0.0360***	-0.159***	-0.0342**	-0.169***
	(0.0136)	(0.0385)	(0.0135)	(0.0382)
Approval Year Fixed Effects	No	Yes	No	Yes
Type of Product Fixed Effects	No	No	Yes	Yes
Observations	3006	3006	3006	3006

Dependent variable: at least one adverse event in the first 5 years after approval for a device category.

specification in column (7).

5.1 Falsification Tests

We use special 510(k)'s as a control, they are seen in 30 days so the benefit from using third parties for the primary review is limited. These are devices that are considered by the FDA to be simpler than the usual Traditional devices, and we expect to find that they are at least as safe as those Special devices approved by the FDA. This is because with a simple (and possibly safer device) there is less room for distortions of quality, and we expect that third parties will be just as good as the FDA in preparing the certification review for these devices. The specifications are similar to the ones used to Equation (1), except we restrict the analysis to Special type approvals.

The results in Table 14 indicate indeed that Special devices are marginally safer after the

pilot is implemented. The result is not significant when using a whether the device had any adverse effects in the first five years as a dependent variable, indicating that indeed devices that allow Special reviews are not affected by the pilot program implementation.

6 Mechanism for the Effect of Third Party Certification

In this section we attempt to analyze how third parties affect the process of certification. In particular, we classify manufacturers by their experience with working with third parties for the initial review process.

Thus, we construct a dummy variable called "Used Third Party Before" for each manufacturer - year of approval group which will be positive if the manufacturer had previously used a third party and zero otherwise. The interaction of this variable with the effect of the pilot program indicates that companies which have previously worked with third parties tend to market safer products. This is an interesting moderator, because it has a few possible interpretations, but ultimately implies that the third party pilot program can be a safe alternative to the FDA reviews, after companies start using third parties regularly.

Columns (1) to (3) estimate an GLM model (for a negative binomial family distribution), and add product category and year fixed effects. Columns (4) and (5) estimate a Zero Inflated Negative Binomial model.

One possible explanation for the negative interaction term is that competition among the third parties (which drives the quality down because they are more lax in standards in order to steal market share from other third party companies) is not actually a problem once a firm establishes a relationship with a particular company. We interview a director for regulatory affairs of one of the manufacturers of medical devices. His opinion was that there is significant habit formation in the industry: companies choose whether to send their products for review to the FDA or to a third party based on company culture instead of more strategic reasons. He believed that once a manufacturer started working with a thirdparty

Table 15: Experience with Third Parties Leads to Safer Products

	(1) Glm	(2) Glm	(3) Glm	(4) Neg. Binomial Zero Inf.	(5) Neg. Binomial Zero Inf.
Pilot x Post Policy x UsedThirdParty	-1.652***	-1.560***	-1.484**	-1.444***	-1.025*
	(0.597)	(0.598)	(0.612)	(0.541)	(0.584)
Pilot Eligible	-0.321***	-0.321***	-0.475***	-0.174**	-0.268***
	(0.0657)	(0.0658)	(0.0722)	(0.0821)	(0.0964)
Post Policy	-0.796***	-1.344***	-1.409***	-0.329**	-0.979***
	(0.0737)	(0.131)	(0.134)	(0.146)	(0.225)
Pilot x Post Policy	0.565***	0.571***	0.578***	0.558***	0.784***
	(0.131)	(0.131)	(0.134)	(0.204)	(0.215)
Used ThirdParty Before	-0.0392	-0.0109	0.327	-0.144	0.0498
	(0.192)	(0.193)	(0.204)	(0.207)	(0.204)
Pilot x UsedThirdParty	0.514	0.511	0.255	0.943***	0.739*
	(0.438)	(0.439)	(0.453)	(0.314)	(0.418)
Post Policy x UsedThirdParty	1.163***	1.165***	0.829***	0.719***	0.563**
	(0.233)	(0.234)	(0.244)	(0.270)	(0.256)
inflate					
PilotEligible after Policy				0.523***	0.663***
				(0.148)	(0.156)
Approval Year Fixed Effects	No	Yes	Yes	No	Yes
Product Category Fixed Effects	No	No	Yes	No	Yes
N	13556	13556	13556	13556	13556
Log-likelihood	-5893.5	-5875.6	-5437.4	-4526.5	-4376.0
Wald χ^2	162.0	185.4	718.0	24.08	6560.9

Dependent variable is the number of years with a bad event in the first 5 years after approval for a device category.

and is satisfied with the outcome, that company is likely to send more products to be certified by the thirdparty.

This story leads us to believe that thirdparties may have an incentive to attract new customers and be too relaxed in their certification when they acquire the customer, but then lose that perverse incentive and their certification standards improve. This could be because manufacturers will be unlikely to switch to another thirdparty after developing a relationship with one company they like already.

However, this is not the only possible explanation for the negative interaction term.

Another interpretation is that once a manufacturer starts working with a third party reviewer, they have more insight into how the review process goes and learn how to improve future products. These improvements could be in the form of a more conservative usage label or more tests to determine safety. This interpretation is supported by anecdotal evidence that the third party review process is much more open than the FDA process. In addition to

the claims third parties make on their websites in order to promote their services, the FDA itself lists a friendly process with open communication as one of the benefits of the third party program.

We cannot distinguish between these two alternative explanations, and there may be more ways to interpret the effect. However, on a larger scale, this effect indicates that the pilot program is not as damaging as it would seem at first look. At this time, few companies actually participate in the pilot program, despite the benefits they could get from getting their products to the market faster. This means that a lot of them are working with third parties for the first time in our data but will likely continue using third party reviews for future products. The effect of this experience with third party reviews should imply that future products will be less biased towards bad quality. It may be that the pilot program will actually yield unbiased reviews once manufacturers work with third parties regularly. This is an empirical question that could be answered in the future.

7 Implications

This paper is a step towards understanding whether third party certification programs affect the quality of the certified products in the end market. We draw on a growing theoretical literature on certification and try to characterize the pull between two conflicting effects: the need to provide safe products for reputation concerns and the desire to gain market share in the certifier market. We use a natural experiment based on the FDA's decision to allow third party firms to prepare certification reviews for certain types of medical devices.

The findings from the analyses above suggest that the introduction of the pilot program lead to a drop in the safety of devices being marketed. This suggests that reputation concerns may be secondary to the desire to acquire market share in this context. However, the results also indicate that manufacturers who use a third party reviewer tend to have safer products in the future. This may mean that as manufacturers start to work with a particular third party

reviewer, there is a lock-in effect that mitigates the certifier's need to fight for market share and allows the certifier to truthfully review the product. An alternative explanation could be that manufacturers learn from the experience of working with third party certifiers and use this knowledge to produce better products subsequently. Future research can investigate which one of these alternatives is a better explanation for the effect we find in this market.

Further research could also use the intensity of the adverse effects, for example, comparing recalls to innocuous malfunctions, to see whether there are differences at the product level depending of the type of certification the products go through before they are marketed. This could be important because previous research about medical devices (Thirumalai and Sinha 2011) has found that recalls are not punished enough by the capital market. Since recalls tend to be very expensive in other categories of products, Thirumalai and Sinha (2011) find an puzzling problem and it would be interesting to see whether the certification mechanism in this market could be a factor into the lack capital market penalties.

References

Albano, G. L. and A. Lizzeri (2001). Strategic certification and provision of quality. *International Economic Review 42*(1), pp. 267–283.

Angrist, J. and V. Lavy (2009). The effects of high stakes high school achievement awards: Evidence from a randomized trial. *The American Economic Review 99*(4), pp. 1384–1414.

Becker, B. and T. Milbourn (2010). Reputation and competition: evidence from the credit rating industry. Harvard Business School Working Papers 09-051, Harvard Business School.

Bolton, P., X. Freixas, and J. Shapiro (2012, 02). The credit ratings game. *Journal of Finance 67*(1), 85–112.

Busse, M., J. Silva-Risso, and F. Zettelmeyer (2006). $1,000 cash back: The pass-through of auto manufacturer promotions. *The American Economic Review 96*(4), 1253–1270.

Dana, J. and G. Loewenstein (2003). A social science perspective on gifts to physicians from industry. *JAMA: The Journal of the American Medical Association 290*(2), 252–255.

Doherty, N. A., A. V. Kartasheva, and R. D. Phillips (2011). Information Effect of Entry into Credit Ratings Market: The Case of Insurers' Ratings. *SSRN eLibrary*.

Farhi, E., J. Lerner, and J. Tirole (2011). Fear of rejection? Tiered certification and transparency. IDEI Working Papers 534, Institut d'conomie Industrielle (IDEI), Toulouse.

Lerner, J. and J. Tirole (2006). A model of forum shopping. *The American Economic Review 96*(4), pp. 1091–1113.

Lizzeri, A. (1999). Information revelation and certification intermediaries. *The RAND Journal of Economics 30*(2), pp. 214–231.

Malmedier, U. and K. Schmidt (2011). You owe me. Manuscript.

Miller, A. R. and C. E. Tucker (2011). Can health care information technology save babies? *Journal of Political Economy 119*(2), 289 - 324.

Pesendorfer, W. and A. Wolinsky (2003). Second opinions and price competition: Inefficiency in the market for expert advice. *The Review of Economic Studies 70*(2), pp. 417–437.

Thirumalai, S. and K. K. Sinha (2011). Product recalls in the medical device industry: An empirical exploration of the sources and financial consequences. *Management Science*, 376–392.

Essay 3:
Layaway and the Quasi-Endowment Effect of Installment Payments*

Abstract

The paper explores the quasi-endowment effect. We evaluate how much consumers are willing to prepay for a purchase which will be experienced in the future. In particular, our results indicate that prepaid installment plans allow the consumer to start deriving utility for the purchase from the moment of the first payment. This quasi-endowment effect is felt only for goods that are purchased for own consumption.

*I am especially indebted to Drazen Prelec for comments and support. Helpful comments were provided by Nathan Fong and Catherine Tucker and seminar participants at MIT.

1 Introduction

A person who decides to buy a home goes through a long and sometimes tedious process of research. Once he finds a suitable candidate and places the offer for the house, the potential buyer starts to experience dread over having the offer rejected and "losing the house". This is inexplicable from the point of view of rational theory: if the offer is rejected, the customer should move on and put in an appropriate offer on the second-favorite house. Even accounting for sunk research costs, which should not count in theory, but do affect consumers, it should be that the same research process has yielded a ranking of favorite houses and not being able to get the top choice because the selling price is too high should not matter as long as some house in the ranking is still a match with the price the consumer is willing to pay. Nevertheless, consumers feel true agony over losing the house they wanted to bid on, and in some cases, increase their bids over what their true willingness to pay, just to be able to buy the house. The struggle to "let go" of a bid that reflects true willingness to pay or increase the bid makes for great reality tv [1] but is baffling from a rational point of view, unless the customer has already felt ownership of the product that he will not ultimately be able to purchase.

Earlier literature implies that firms can dupe naive consumers into developing feelings of endowment. We show that far from being naive, consumers anticipate the endowment effect, and are actually willing to pay for it.

Previous papers (Heyman et al. 2004) that establish and use the term "quasi-endowment effect" see it as an irrational manifestation of attachment that cannot be foreseen by the customer. In particular, the effect is what causes "snipping bidding" towards the end of an online auction by forcing the customer holding the winning bid to increase his bid more than his true willingness to pay just as a response to the fear of "losing" the auction. In our

[1] See "Property Virgins"

extended version of the definition of quasi-endowment, the customer is aware, on some level, that he is deriving utility from the quasi-endowment effect. As such, we are able to show that customers can increase their willingness to pay in order to get the quasi-endowment utility, even in cases where they are not actually feeling the effect. Lab experiments allow us to quantify how much a consumer is willing to pay ahead of time for quasi-endowment. We use anticipated willingness to pay as a proxy for how much the consumer is aware of the effect he will be experiencing.

We explore the possibility that customers get attached to objects from the moment they decide to purchase them and even more so when they make a first downpayment or monetary commitment to purchasing the objects. Quasi-endowment, in this paper, is an extension of the endowment effect. We define the effect broadly: it refers to a sense of ownership that consumers develop for a product, despite the fact that they are not owners or users of that product. This quasi-endowment thus explains the sense of loss that some home buyers feel when their bid is rejected or the deal falls through. The object that inspires the effect need not be as valuable as a house: the decision to buy a book online can also cause the quasi-endowment effect, from the moment the book is placed in the cart. In fact, it is intuitive that consumers feel this effect even more after having paid for the book, even though the physical object is still not available for use.

There is a large literature showing that people quickly become attached to goods they have only briefly owned, or not even started to own yet. This earlier literature has focused on what happens after people are already endowed with a good. In contrast, we show that people are not naive about these effects, and are actually able to anticipate them and willing to pay for them.

An important consideration for installment schemes or layaway plans is the time discounted value of the payments. In that respect, our paper relates to research on time discounting. Prelec and Loewenstein (1998) show that customers prefer to prepay for luxury

or pleasurable items but would want to defer payment and spread it out over the the consumption period for utilitarians items. Our results indicate that customers anticipate the quasi-endowment effect and are willing to pay for it across categories of products, including utilitarian ones.

Layaway plans have become popular recently for all types of products, and are being promoted as the easy and frugal way to make purchases while being on a budget. Considering stores have layaway fees and there is the cost of no interest, layaway plans are actually not a way to save money. However, they are a way to control spending impulses, and thus could be used as a self control mechanism. We do not attempt to prove that quasi-endowment is not related to self-control. Moreover, we don't restrict our analysis to layaway plans. Our results show that the subjects may be using layaway as a way to get to experience the pleasure of quasi-endowment, while we can't rule out that they are also benefiting from the added discipline of making a simple plan to save or of having a mental account for each item they put on layaway (see Thaler (1999)).

Mental accounting may be especially important in the face of low budget constraints. Our research thus relates to a stream of literature on low income households and their behavior. Layaway in itself is associated, wrongly or not, with low status goods, as Prelec and Simester (2001) find when they look at what the introduction of an installment payment plan signals to the customer base of a catalogue retailer for luxury items. We do not look into the signaling aspect of installment plans.

Our analysis uses a pool of subjects that is more diverse than just lower income households, so we anticipate that our results are not limited to a certain demographic group. Thus, unlike Mullainathan and Shafir (2009), we do not anticipate that our findings are limited to low income households or that the effects we find stem from low budget constraints.

2 The Experiments

We use three experiments to test the quasi-endowment effect and find its boundaries.

Experiment 1

The experiment was done in the MIT lab, on a computer, as a package of other marketing experiments that reveal customer preferences. There is no deception, and customers are encouraged to leave comments at the end of the battery of tests. We collect response time, IQ measures, and have access to demographic information for the subjects. The survey places participants in 3 conditions: installment plan with 2 payments, cash at the end of the period, credit card at the end of the period. They are asked to imagine they are purchasing a product for themselves in a month and state how much they are willing to pay for it. Reference prices are provided and the products are from diverse categories (used in previous papers: washer and drier, Lego set, camera, vacation cruise etc).

Respondents

A total of 96 respondents were part of the experiment. We drop responses from 4 participants who seem to ignore the questions and give nonsensical answers. The remaining sample is well balanced, as the assignment is random across conditions. There are 46 males and 45 females and 1 subject who did not answer the gender query. Additional statistics about the sample of subjects are presented in tables A-1, A-2, A-3, and A-4 in Appendix A. Subjects received compensation for their time in the lab, as this experiment was part of a battery of other tests.

Design and Procedure

The experiment elicited willingness to pay from subjects in 3 conditions: cash, installment payment and credit card. The items were selected because they have been used in previous literature and reference prices were realistic at the time of the experiment. Subject did not

have access to the Internet to check prices, but we found that providing reference prices made our estimates conservative. In general, reference prices act as anchors: subjects who see a reference price will state their willingness to pay around that reference price. Thus, providing reference prices can make a weak effect dissapear and the size of any effect that is still significant in the presence of reference prices is conservative. We chose to include reference prices in order to give our subjects some information about the quality and availability of the items they were seeing, without burdening them with specific information about the product. For example, digital cameras vary in price from $50 to $1500, and their price also depends on the store where they are purchased. However, telling participants "Kodak Easyshare M863 Camera (Sells for $119.99 at store CircuitCity)" is completely informative yet quick and easy.

The subjects were seated in front of a computer screen and started by doing unrelated marketing tests. At some point in the sequence of experiments, they read the following if they were in the Credit Card condition:

- "Imagine that, one month from now, you are planning to purchase a Product from the list below for yourself. You will pay for this product using your credit card at the time you receive the product. The store you plan to purchase from accepts all major credit cards."

or if they were in the Cash condition:

- "Imagine that, one month from now, you are planning to purchase a Product from the list below for yourself. You will pay for this product using cash at the time you receive the product. "

or if they were in the Installment condition:

- "Imagine that, one month from now, you are planning to purchase a Product from the list below for yourself. You will pre-pay for this product using 2 cash installments.

The first installment is to be paid tomorrow and the last one at the time you receive the product, a month from now. The product becomes yours when you make the first payment but you can take it home only after you have finished paying for it. "

They were then presented with a list of items (held constant across conditions) and asked to name the total price they would be willing to pay. Additionally, in the installment condition, there was space for them to write in two installments. To minimize the danger that faulty adition influence our results, we also provided them with a space to first name the total price and it was this total number that is used in all the analysis.

Results and Discussion

The results indicate that subject anticipate the quasi-endowment effect for most items purchased in installments. Notably, the effect is not valid for the Amazon gift card, which is consistent with previous research. Since a gift card (especially for amazon, with its wide variety of merchandise) is a money equivalent, we do not expect subjects to feel quasi-endowment for it. The results show that subjects who pay in cash or by credit card value the gift card just as much as those whose split the payment into installments.

It is interesting to note that the Wii system is also only marginally more desirable in the installment condition: at the time of the experiment, the Wii was still in limited supply and we believe that the participants did not anticipate being able to purchase one for any price, so their willingness to pay reflects the price of the Wii but not necessarily their willingness to pay for it.

Overall, the subjects were willing to pay more than an overall premium of $62.33 (significant at 5% level) if allowed to pay in installments right away. The result indicates not only that the subjects anticipate the pleasure of feeling quasi-ownership for these items, but that they are willing to pay at least 20% of the total price in order to secure this pleasure.

Interestingly, we do not find any difference between the cash payments and the credit

Table 1: Installment and Cash Payment

	Washer/Drier	Ipod Nano	Wii	Lego Set	Kodak Camera	Cruise	Amazon Card	Pen
Installment	129.6***	25.84**	20.41	4.652**	13.24*	239.3**	1.352	48.68**
Premium	(2.67)	(2.44)	(1.38)	(2.25)	(1.85)	(2.39)	(0.58)	(2.23)
Cash	704.7	144.0	183.6	21.17	91.30	1254.1	44.06	131.5
	(21.91)	(20.73)	(18.56)	(15.41)	(19.69)	(19.12)	(28.49)	(9.15)
Observations	57	56	58	52	57	56	55	53

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2: Credit Card and Cash Payment

	Washer/Drier	Ipod Nano	Wii	Lego Set	Kodak Camera	Cruise	Amazon Card	Pen
Credit Card	2.889	7.250	3.093	-0.887	2.848	145.1	1.686	-18.91
Difference	(0.06)	(0.72)	(0.24)	(-0.43)	(0.44)	(1.57)	(0.62)	(-0.94)
Cash	704.7	144.0	183.6	21.17	91.30	1254.1	44.06	131.5
	(18.86)	(20.13)	(19.92)	(14.74)	(19.75)	(19.21)	(22.85)	(9.13)
Observations	65	64	64	57	66	64	63	62

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

card payments. Both are supposed to happen in a 1 time payment a month from when the experiment takes place, and there is no difference between reported willingness to pay across these conditions. This result is consistent to results from Prelec and Simester (2001) who find that subjects are willing to pay more when using a credit card, but that they are unaware of these tendencies.

Experiment 2

Study 1 was performed in the laboratory and was free from unwanted outside influences, but has the downside of not being incentive compatible. Experiment 2 is an incentive compatible field experiment for a small sample size.

Respondents

The experiment is a second price auction for a pair of Red Sox tickets. The 43 Mba students of a summer class in MIT Sloan submitted bids. Those who did not win the auction were not compensated. However, those who won the auction were allowed to purchase the tickets at face value (around $100).

Design and Procedure

The experiment was a second price auction for a pair of sold out Red Sox tickets. The students received bid sheets with an identifier to annonimize the participants and instructions about how the auction would unfold. The bid sheets were split in two payment conditions and randomized:

- payment in full when the winner received the tickets a week later after the auction
- payment of 10% of the winning bid at the time of the auction and the rest when they received the tickets a week after the auction

Results and Discussion

The quasi-endowment effect should make the students who pay the 10% downpayment willing to pay higher prices for the tickets. Indeed, we find that the mean for the Cash condition is \$101 while for the Installment condition it is \$134, a 33% increase (significant at 6%, significance based on log transformation of prices). For the median, the difference is even higher: \$40 compared to \$100.

We repeated the same experiment with a separate group of students the following week, but the results are not conclusive because there were some students who looked up the prices of the tickets. In the face of an outside option, the results of the auction were closely related to the actual market value of the tickets and not to the private valuation of each participant.[2]

[2]The means for the Cash and Installment conditions were \$104 and \$109 respectively and the medians were \$110 and \$90, while the face value for the tickets was \$100.

Experiment 3

This experiment tests whether all purchases can bring about the quasi-endowment effect. In particular, we test whether items purchased as gifts have the same property as items purchased for the self. If the quasi-endowment exists for gifts, then we can generalize that all purchases are anticipated and that subjects are willing to pay more if they anticipate purchasing anything. However, if gifts are not associated with quasi-endowment, then we identify a stronger variant of the effect: subjects anticipate the quasi-endowment only when they are the ones who will ultimately directly benefit from the purchase.

Respondents

The 84 respondents were participants in the MIT lab in a battery of marketing tests. They were presented with a similar questionnaire to Study 1, with slight variations in the instructions. Subjects who had participated in the earlier experiments were not allowed to be part of this experiment.

Design and Procedure

The subjects were seated in front of a computer screen and started by doing unrelated marketing tests. At some point in the sequence of experiments, they read the following if they were in the Credit Card condition:

- "Imagine that, one month from now, you are planning to purchase a Product from the list below as a gift for a family member. You will pay for this product using your credit card at the time you receive the product. The store you plan to purchase from accepts all major credit cards. Please write the price you are willing to pay for each product in the list."

or if they were in the Cash condition:

- "Imagine that, one month from now, you are planning to purchase a Product from the

list below as a gift for a family member. You will pay for this product using cash at the time you receive the product. "

or if they were in the Installment condition:

- "Imagine that, one month from now, you are planning to purchase a Product from the list below as a gift for a family member. You will pre-pay for this product using 2 cash installments. The first installment is to be paid tomorrow and the last one at the time you receive the product, a month from now. The product becomes yours when you make the first payment but you can take it home only after you have finished paying for it. "

The manipulation *as a gift for a family member* ensures that the subjects consider these items to be desirable, but are not planning to use them themselves.

Table 3: Installment Premium vs Cash Payment For Gifts

	Washer/Drier	Ipod Nano	Wii	Lego Set	Kodak Camera	Cruise	Amazon Card	Pen
Installment	94.28*	17.34	9.071	-0.683	3.905	-3.507	-0.259	30.70
Premium	(1.68)	(1.59)	(0.82)	(-0.17)	(0.49)	(-0.04)	(-0.19)	(1.33)
Cash	715.5	148.5	200.3	26.93	100.8	1439.1	47.91	128.7
	(21.08)	(23.42)	(30.63)	(11.62)	(21.34)	(27.21)	(60.45)	(9.07)
Observations	52	50	49	46	51	50	49	45

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Results and Discussion

The analysis reveals gifts do not seem to cause the quasi-endowment effect. We cannot reject the null hypothesis that the total of the installment payments is the same as the willingness to pay in cash.

We further check whether the null effect of the quasi-endowment for gifts is caused by a difference in our samples, or whether there is some systematic difference in valuations of the items for gifts as opposed to the same items if purchased for self. Table B-2 compares the Cash condition across Study 1 and Study 3, while table B-3 compares the studies in the Installment condition. We find no significant difference in how much the items are valued.

Pooling the data from Study 1 and Study 3 indicates that the quasi-endowment effect is marginally significant still, though not for all the items included in the study. Tables B-4 and B-5 show that for gifts for others and items purchased for personal use, the quasi-endowment has the expected sign.

3 Conclusion and Implications

We have showed that the quasi-endowment effect is a powerful force in driving payment decisions and increasing willingness to pay for consumers. We have shown that consumers are not naive about the quasi-endowment effect, they anticipate the feelings of ownership that will occur, and are actually willing to pay for the pleasure that comes from this feeling of ownership. Moreover, we have analyzed how robust the effect is in a variety of situations, and quantified how its size may differ based on several factors. As far as we know, this is the first paper to explore and measure the quasi-endowment effect in an experimental setting. However, there is evidence that merchants and retailers are aware of the existence of the effect and are using it to improve their profits (and arguably to create happier consumers).

The quasi-endowment effect is intuitively used by retailers, both in brick and mortar store and in online settings. The helpful sales assistant who takes away your apparel items and sets up a cabin to try them on is not only freeing your hands to shop for more items, but also reserving the current items for you and allowing you to enjoy the quasi-endowment pleasure of these items. Shopping baskets in stores where the average purchase is a few items are meant to encourage larger purchases, but also to make sure the consumer feels that what is in the basket is valuable and it's his. Online merchants have developed shopping carts that will save items in between browsing sessions, to increase the chances of a purchase by lower repeat search costs, but also to remind the consumer that these items are his and it would painful to delete them. In the opposite direction, some retailers allow consumers to "steal" each other's shopping cart items, to create a sense of frenzy and increase bidding on items.[3]

[3]One such example is RueLaLa.com, a popular online reseller of high fashion items with limited time sales.

References

Heyman, J. E., Y. Orhun, and D. Ariely (2004). Auction fever: The effect of opponents and quasi-endowment on product valuations. *Journal of Interactive Marketing 18*(4), 7–21.

Mullainathan, S. and E. Shafir (2009). *Savings Policy and Decisionmaking in Low-Income Households*, pp. 121–145. Russell Sage Foundation Press. ¡p¿September 2010¡/p¿.

Prelec, D. and G. Loewenstein (1998). The red and the black: Mental accounting of savings and debt. *Marketing Science 17*(1), pp. 4–28.

Prelec, D. and D. Simester (2001). Always leave home without it: A further investigation of the credit-card effect on willingness to pay. *Marketing Letters 12*(1), 5–12.

Thaler, R. H. (1999). Mental accounting matters. *Journal of Behavioral Decision Making 12*(3), 183–206.

Table A-1: Experiment 1: Payment Condition and Gender Summary

Male	0	1	Total
Credit Card	45.45	54.55	100.00
Cash	42.42	57.58	100.00
Installments	64.00	36.00	100.00
Total	49.45	50.55	100.00

Note that 1 observation is missing.

Table A-2: Experiment 1: Payment Condition and Income Levels Summary

Income Level	1	2	3	4	5	Total
Credit Card	62.50	12.50	9.38	0.00	15.63	100.00
Cash	57.58	18.18	12.12	3.03	9.09	100.00
Installments	57.69	7.69	7.69	11.54	15.38	100.00
Total	59.34	13.19	9.89	4.40	13.19	100.00

Note that 1 observation is missing.

Table A-3: Experiment 1: Payment Condition and Education Summary

Education Level	1	2	3	4	5	Total
Credit Card	0.00	6.06	27.27	24.24	42.42	100.00
Cash	3.03	9.09	27.27	36.36	24.24	100.00
Installments	0.00	12.00	24.00	32.00	32.00	100.00
Total	1.10	8.79	26.37	30.77	32.97	100.00

Note that 1 observation is missing.

Table A-4: Experiment 1: Payment Condition and Employment Summary

Employment	1	2	3	4	Total
Credit Card	15.15	12.12	54.55	18.18	100.00
Cash	18.18	15.15	39.39	27.27	100.00
Installments	19.23	7.69	53.85	19.23	100.00
Total	17.39	11.96	48.91	21.74	100.00

Table A-5: Experiment 3: Payment Condition and Gender Summary

Male	0	1	Total
Credit Card	56.67	43.33	100.00
Cash	54.55	45.45	100.00
Installments	52.63	47.37	100.00
Total	54.88	45.12	100.00

Table A-6: Experiment 3: Payment Condition and Income Levels Summary

Income Level	1	2	3	4	5	Total
Credit Card	41.38	20.69	13.79	6.90	17.24	100.00
Cash	65.63	12.50	9.38	3.13	9.38	100.00
Installments	55.56	27.78	5.56	0.00	11.11	100.00
Total	54.43	18.99	10.13	3.80	12.66	100.00

Table A-7: Experiment 3: Payment Condition and Education Summary

Education Level	2	3	4	5	Total
Credit Card	3.33	36.67	23.33	36.67	100.00
Cash	12.12	24.24	27.27	36.36	100.00
Installments	5.00	25.00	25.00	45.00	100.00
Total	7.23	28.92	25.30	38.55	100.00

Table A-8: Experiment 3: Payment Condition and Employment Summary

Employment	1	2	3	4	Total
Credit Card	26.67	6.67	50.00	16.67	100.00
Cash	15.15	15.15	45.45	24.24	100.00
Installments	15.79	5.26	57.89	21.05	100.00
Total	19.51	9.76	50.00	20.73	100.00

Table B-1: Credit Cards and Cash Payment for Gifts

	Washer/Drier	Ipod Nano	Wii	Lego Set	Kodak Camera	Cruise	Amazon Card	Pen
Credit Card	37.05	0.337	0.127	-4.554	-6.197	-110.9	-0.906	-4.801
Difference	(0.64)	(0.03)	(0.01)	(-1.41)	(-0.83)	(-1.18)	(-0.64)	(-0.22)
Cash	715.5	148.5	200.3	26.93	100.8	1439.1	47.91	128.7
	(18.33)	(20.50)	(24.30)	(11.88)	(19.65)	(22.63)	(48.56)	(8.73)
Observations	61	61	59	59	62	59	62	51

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B-2: Cash Payment For Gifts vs Self: Comparison of Means across Studies

	Washer/Drier	Ipod Nano	Wii	Lego Set	Kodak Camera	Cruise	Amazon Card	Pen
Self	-10.77	-4.453	-16.66	-5.760*	-9.515	-185.0**	-3.842**	2.818
	(-0.20)	(-0.45)	(-1.59)	(-1.80)	(-1.41)	(-2.11)	(-2.17)	(0.14)
Gift	715.5	148.5	200.3	26.93	100.8	1439.1	47.91	128.7
	(19.02)	(21.58)	(27.12)	(12.04)	(21.06)	(23.25)	(38.50)	(8.76)
Observations	65	65	64	59	66	64	63	58

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B-3: Installment Payments For Gifts vs Self: Comparison of Means across Studies

	Washer/Drier	Ipod Nano	Wii	Lego Set	Kodak Camera	Cruise	Amazon Card	Pen
Self Difference	24.58	4.049	-5.316	-0.424	-0.182	57.78	-2.230	20.81
	(0.59)	(0.36)	(-0.30)	(-0.22)	(-0.02)	(0.58)	(-1.02)	(0.85)
Gift	809.7	165.8	209.4	26.25	104.7	1435.6	47.65	159.4
	(25.92)	(19.35)	(15.24)	(17.96)	(17.20)	(18.95)	(28.51)	(8.64)
Observations	44	41	43	39	42	42	41	40

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B-4: Installment Premium vs Cash Payment For All Purchase Types

	Washer/Drier	Ipod Nano	Wii	Lego Set	Kodak Camera	Cruise	Amazon Card	Pen
Installment	113.6***	21.90***	14.18	1.898	8.558	122.0*	0.326	41.20**
Premium	(3.12)	(2.92)	(1.49)	(0.89)	(1.61)	(1.79)	(0.23)	(2.62)
Cash	710.2	146.3	192.0	24.10	96.06	1346.6	46.02	130.2
	(30.67)	(31.31)	(31.83)	(17.86)	(29.03)	(31.30)	(51.37)	(12.95)
Observations	109	106	107	98	108	106	104	98

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B-5: Credit Cards and Cash Payment For All Purchase Types

	Washer/Drier	Ipod Nano	Wii	Lego Set	Kodak Camera	Cruise	Amazon Card	Pen
Credit Card	18.04	3.841	1.030	-2.751	-1.690	20.11	0.339	-12.83
Difference	(0.47)	(0.53)	(0.11)	(-1.41)	(-0.34)	(0.30)	(0.22)	(-0.87)
Cash	710.2	146.3	192.0	24.10	96.06	1346.6	46.02	130.2
	(26.44)	(28.95)	(30.86)	(17.59)	(27.88)	(29.21)	(42.33)	(12.72)
Observations	126	125	123	116	128	123	125	113

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$