Bargains-then-ripoffs:
Innovation, pricing and lock-in in enterprise software

Ian Larkin*
Harvard Business School
ilarkin@hbs.edu

November 20, 2007

Abstract

In industries with quick innovation cycles and switching costs, vendor profitability is often driven by the ability to "lock in" customers. Despite a large theoretical literature, there are few empirical studies on the success of vendor pricing and product strategies to achieve lock-in. This paper uses a comprehensive database of transactions in two major product lines for a large enterprise software vendor to examine the use and success of pricing and product strategies to increase lock-in. Regarding pricing, the paper demonstrates that the vendor engages in significant "bargain-then-ripoff" pricing; customers completely new to it receive discounts that are nearly 50% greater than existing customers who are purchasing upgrades of a product. On the product side, one common lock-in strategy – the offering of broad product "suites" which span multiple product lines – appears to be of limited effectiveness. The vendor cannot limit its initial "bargain" to customers new to a product line, even if the customer has bought a different product the vendor previously. This suggests that lock-in is product-, not vendor-specific. Finally, not all customers get locked-in; I find that customers with high IT capabilities, and/or those with strong financial performance (which are very highly correlated), are likely to receive an "initial bargain" and then switch suppliers or use legacy systems, rather than pay locked-in rates. The "best" customers in terms of future revenue potential therefore avoid getting locked-in.

* This paper is based on the second chapter of my dissertation in the Business and Public Policy program at the Haas School of Business, University of California, Berkeley. I would like to thank my committee – David Mowery, Steven Tadelis, Catherine Wolfram and Bronwyn Hall – for their considerable guidance. Funding from the Ford Foundation and the Harry S. Truman Foundation is gratefully acknowledged. All errors are my own.

1 Introduction

Researchers in industrial organization and strategy commonly hold that achieving product lock-in is critical for vendors in information good industries. Lock-in, defined as a state where "switching suppliers [by customers] is virtually unthinkable" (Varian, 2003), allows information good vendors to sell at quasi-monopolistic prices and reap profits that can be plowed back into research and product development, a key part of achieving long-term profitability in industries with quick innovation cycles.

Lock-in occurs when customers incur switching costs if they change vendors, making it less costly for them to pay locked-in rates to their current vendor than to switch to a new one. Switching costs, therefore, become a critical part of the strategic toolkit for information good vendors, as noted by Shapiro and Varian (1998):

> You just cannot compete effectively in the information economy unless you know how to identify, measure, and understand switching costs and map strategy accordingly.

Switching costs and lock-in affect vendor decisions around dynamic pricing, the rate of product innovation, the depth and breadth of product and service offerings, and a host of other strategic issues. There is a plethora of theoretical papers about such markets, and widely accepted conventional wisdom from case studies and the like. However, there is little detailed empirical research around the key determinants of lock-in, both in terms of vendor strategy and customer type, especially in the business-to-business (B2B) procurement environment. Since the B2B procurement market is some $4 trillion annually (Muthoo, 2000), this is clearly an area ripe for research. Furthermore, the few empirical studies that exist use aggregate measures of market share and pricing, not actual contract-by-contract data, to examine switching costs and lock-in. These studies therefore cannot speak to price discrimination strategies used by vendors and heterogeneous customer response to vendor pricing and product strategies.

In this paper, I use a proprietary database of individual customer sales for a leading enterprise software vendor, along with a unique dataset of switching behavior of these customers, to examine vendor strategy and customer response in enterprise software, a market with very high switching costs. Specifically, I examine the extent to which two vendor strategies achieve lock-in: price discrimination strategies between existing and new customers, and strategies to sell multiple products to one customer. I also study how different subsets of customers respond to these vendor attempts to promote lock-in, specifically looking at whether some customer types are more likely to get locked in than others.

Overall, the empirical results demonstrate the difficulty vendors have in successfully locking in their customer bases. I first demonstrate that the vendor uses “bargain-then-ripoff” price discrimination, as predicted by the theoretical literature, in an attempt to attract new customers through low prices. The results show that discounts for new customers are 15 to 25 percentage points larger than discounts for existing customers.. Since average discounts are around 30 percentage points, new customers are getting “bargains” that are at least 50% larger than the average customer.

However, the results also show that a significant share of customers take advantage of initial bargains and subsequently switch to other providers, who are likely also giving them bargains. These customers have, as measured by industry surveys, greater capabilities in information technology, which is highly correlated with financial performance and company growth. Therefore, the “best” potential customers for the vendor in terms of future revenue potential appear to be less likely to get locked in and pay high prices. Finally, the results show that the effectiveness of one technique commonly believed to increase lock-in – offering multiple products – is fairly limited. Lock-in increases when the vendor sells similar products in the same product line, but there is no additional lock-in benefit from selling products outside of the product line the customer already uses[1].

These results have clear implications for vendor strategy in enterprise software. Overall, contrary to existing literature, the results suggest that broad software “suite” providers are not significantly better at increasing lock-in than narrow, focused software firms. This finding is consistent with the industry’s heterogeneous structure, which includes a few large “suite” providers and thousands of small, focused firms. Secondly, the results indicate that the strategy of offering heavy discounts to new customers, which is a common industry phenomenon, may be counterproductive. The customers who get the largest discounts happen to be customers with the highest IT capabilities and the best financial

[1] I will define the terms “product” and “product line” in the context of enterprise software in much more detail later in the paper.

performance, attributes that suggest they are the most likely to switch providers. More nuanced pricing structures might enhance revenue capture for the vendor.

This paper makes several important contributions. It is among the first to study price discrimination arising from lock-in in a private, B2B setting, and is likely the first to directly measure an average price differential paid by existing versus new customers in this setting. Price is the key variable in most theoretical models of switching costs, and difficulty in obtaining data has led it to be under-researched. Similarly, it is among the first studies to look at heterogeneous customer response to vendor pricing and product strategy; other empirical studies only compare customers to non-customers, and cannot identify switchers from non-switchers. Finally, there are several streams of further research evident from the study, most notably in building a deeper understanding to lock-in dynamics from the customer perspective.

The paper is laid out as follows. In the next section, I discuss factors that lead enterprise software customers to incur switching costs, as well as industry purchasing and upgrade dynamics[2]. Section three reviews the current theoretical and applied literatures on vendor-customer dynamics in industries marked by significant switching costs and product lock-in. In section four, I build hypotheses on vendor pricing and product strategy, and response to such strategies by customer type. In section five, I review the data, estimation strategy and empirical results. In the final section I discuss the implications of the results for vendor lock-in strategies, overview the study's limitations, and talk about potential next steps.

2 Institutional background

2.1 A brief overview of the enterprise software product space

To understand switching costs and product lock-in in enterprise software, it is first useful to define the product space. As shown in Figure 1, enterprise software can be broadly divided into "applications," which are programs with which employees interact to examine, manipulate and use data, and "infrastructure," the programs that operate the applications,

[2] Larkin (2007), based on the first chapter of my dissertation and concerned with incentive system gaming by enterprise software salespeople, contains a detailed discussion of price negotiations in enterprise software, which is highly relevant to this chapter as well.

and store data in a secure yet flexible way. Only employees in a company's IT department will interact with infrastructure programs in any meaningful way.

Figure 1 also shows that there are multiple "product lines" within each space; for example, within the applications space there are separate product lines for financial management, human resource management, supply chain management, customer relationship management, and a multitude of others. While the exact definition of a "product line" varies somewhat between industry analysts and vendors, the broad boundaries between the lines are widely accepted.

Within each product line, there are multiple products. As Figure 1 indicates, these may be point products, such as a billing software solution within the financial management product line. However, most "products" are in fact made up of multiple point products within the same line; for example, a "cost management" product might be made up of a vendor's payment management and cost accounting point products, within its financial management product line. Complicating the picture, each individual point product usually has multiple "grades," with increased functionality, speed and/or reliability as the grade increases. A typical vendor SKU book contains literally tens of thousands of individual "products" that are different combinations of point products at different grades. It is critical to re-emphasize that the study defines a "product line" as a set of products meeting the same broad customer need[3], and a "product" as an individual solution*, or bundled set of solutions* (which is in fact much more common in the data used), within a single product line.

Finally, Figure 1 outlines the two most common types of software vendors: product specialists, who concentrate on solutions in a single product line, and industry specialists, who sell industry-specific solutions across product lines. Nearly all (if not all) enterprise software companies start as one of these types; SAP, the largest enterprise software application vendor, started as a financial management software firm, while Oracle, one of the largest application and infrastructure players, started as a database management firm on the infrastructure side. Both of these companies, and a few others, have emerged via organic growth and acquisitions as "suite providers" (not shown in 1), offering products in

[3] To be more precise, I accepted the product line definitions the vendor used for its products. Like most enterprise software companies, this vendor's engineering and marketing departments are organized by product line.

many different product lines. There are not many suite providers with significant market share in multiple product lines, but the few that exist – including SAP and Oracle on the applications side, and IBM, Microsoft, Oracle and Symantec[4] on the infrastructure side – are quite large and influential.

2.2 Product upgrades and switching costs in enterprise software

The enterprise software industry is characterized by very fast innovation cycles, where vendors develop and sell upgrades to their current products that meet the same underlying business need but are improved in terms of functionality, speed, integration with other software platforms, and other key areas[5]. Vendors release small functionality upgrades for major products at least annually; buyers usually receive these upgrades for free as part of their initial contract. However, customers usually must purchase major upgrades to the product, which occur about every five years. While it is not uncommon to skip an upgrade cycle (which is essentially the equivalent of continuing to use Microsoft Office 2000 when Office XP is available), customers are often afraid that skipping cycles will leave them at a competitive disadvantage because their businesses will not have access to software functionality their competitors are using to run their companies more efficiently.

In order to successfully implement a newly purchased software package, a corporate customer incurs a plethora of costs beyond what it pays for the actual software licenses. These costs are typically broken down into three components: hardware, software customization (defined as customer-specific changes to the software code purchased), and training and implementation. A common industry estimate is that these costs are roughly eleven times the cost of the software itself (Varian, 2002).

The first time a customer purchases software for a certain business need, for example when it replaces a non-electronic invoicing system with a server-based, automated package,

[4] Symantec is an interesting example of a company that started in the consumer security software space, but moved into enterprise software organically and through the acquisition of VERITAS, a leading enterprise storage vendor. Microsoft too started as a consumer software company, but has invaded the enterprise space organically and via a large acquisition of Great Plains software, a large enterprise applications provider.

[5] This is also true of consumer software, where vendors like Microsoft are continually putting out new versions of their products. The Microsoft Vista operating system, for example, looks almost nothing like Windows 95, but is only 11 years older and fulfills essentially the exact same role for consumers. It is also a full 4 upgrade cycles removed from Windows 95, so Microsoft averaged an upgrade every 2.5 years during this timeframe.

a customer cannot avoid these additional expenses in hardware, customization and implementation. However, when purchasing a product upgrade with new functionality, some of these expenses are not replicated if a customer stays with its current vendor. Conversely, if a customer switches vendors, these sunk costs will have to be replicated, leading to switching costs between vendors.

The most obvious example is training costs. Employees are trained to use a particular software platform, which is usually quite complex, and business processes – or the way employees interact with data and information flow – have been changed to meet the software's needs. Much of this training and business process redesign will carry over if the buyer simply buys an upgrade of its current product, but will have to be replicated substantially if a customer chooses to switch vendors. This is because products providing the same basic functionality but designed by two vendors differ in the exact method by which the software is used and the way in which it interacts with business processes and information flow. Similarly, when upgrading to a new version of the same product, a customer can often "piggy back" on previous customized code it purchased for the original product. But if a customer switches vendors, this customized code is almost always completely scrapped, since it interacts with code written by the replaced vendor that will no longer be used. Hence, switching enterprise software vendors for a particular product incurs a substantial additional cost when compared with buying an upgrade from the existing vendor of the product[6].

Switching costs also reflect the business risk of changing vendors. Even if additional training and replication costs are minimal, the experience good nature of enterprise software, and uncertainty around how successfully a business will adapt to a new package, can lead executives to be heavily biased towards an existing vendor. This bias is enhanced by "urban legends" of customers' businesses being harmed after switching software providers. For example, in the year 2000 Nike famously (and very publicly) blamed poor implementation of a new package from the supply chain software firm i2 for a quarterly loss that far exceeded industry estimates and led to a one-day, double-digit decline in both i2 and Nike's stock prices.

---

[6] It is interesting to note that new hardware is almost always bought for applications undergoing a major upgrade; customers are loath to use old hardware for the latest applications. Hardware costs, therefore, are usually not a source of switching costs.

2.3 Sales dynamics in enterprise software

Like many large-scale, business-to-business procurement transactions, enterprise software customers tend to purchase from vendors via one-on-one negotiations, leading to a unique price for a certain package for that customer. The results of negotiations are usually kept secret as a contractual matter, and the vendor therefore is able to engage in price discrimination across customers. Of course, vendors do not know a customer's willingness to pay or product utility, nor do they know a customer's switching cost. As a classic experience good, buyers too do not know their utility from using a product, and usually report in surveys and the industry press that they incur switching costs that were much larger than anticipated when implementing a software package.

The price book is scaled to give customers the incentive to buy multiple products within a line, and to switch vendors. The list price of a multi-product solution is almost universally cheaper than the combined list prices of the two point products[7]. Also, it is quite common for a customer using a product defined by a vendor as a direct competitor to get a "competitive switch" discount from list price. Current customers sometimes get an "upgrade credit" for their current product that reduces the list price they face to upgrade; such credits are usually only for the last version of the product, and for major upgrades it is common not to extend any credit to current customers.

The complicated nature of the price book, as well as its secrecy (price books are usually not publicly available), means that both the vendor and customers tend to focus on discounts rather than raw prices. There is not an accepted "market price" for products, while customers are trained to think about the size of the discount they receive (Riciutti, 2004). Also, the vendor itself focuses its sales management practices on discount management.

2.4 Prevalence of product upgrades and switching costs in other settings

The importance of quick product upgrades and customer switching costs, along with first degree price discrimination by vendors as a tool to cope with these factors, is by no means

[7] This is true for "suites," or solutions with products from different lines, as well.

unique to enterprise software. Many business-to-business information good industries, such as enterprise servers and mainframes, semiconductor equipment, telecom equipment and services, and numerous others have similar industry economics, upgrade patterns and pricing practices. Non-information industries such as defense equipment, commercial aircraft and professional consulting and accounting services are also marked by significant switching costs, product and service innovation, one-on-one negotiations, and first degree price discrimination.

3 Existing literature on switching costs and lock-in

Researchers have devoted considerable attention to studying industry dynamics when switching costs and the potential for lock-in are high. I restrict my review to studies on price discrimination (either temporal or across customer groups), product strategies to increase lock-in, and the very small literature on heterogeneous customer response to such strategies[8].

3.1 Price discrimination: "bargains-then-ripoffs"

The main theoretical prediction of the switching cost literature is that when sellers cannot commit to pricing across periods, buyers will be compensated for incurring switching costs when making their initial purchase. Rational vendors will realize that they can charge a price premium to customers who incur a switching cost, and are therefore willing to lower price for the initial purchase. This logic was first modeled explicitly by Klemperer (1987), and Farrell and Klemperer (2006) cite no fewer than 20 theoretical studies of this basic dynamic[9]. Most of these models assume that customers have identical switching costs, and therefore predict switching will not occur. Chen (1997) and Taylor (2003) model situations where switching costs are not known with certainty, and demonstrate that switching will occur when a competing vendor offers a later-stage "bargain" too good to pass up, or when the installed vendor attempts a "ripoff" that is too large. Since buyers are needlessly

[8] For a very thorough review of the larger literature on switching costs, see Farrell and Klemperer (2006).
[9] I show a very basic version of such a model in the next section.

incurring a switching cost on the subsequent purchase, this switching is inefficient. A key facet of all of these models is that new customers receive lower prices than “loyal” ones.

There have been surprisingly few empirical studies on “bargain then ripoff” behavior, probably because pricing data in markets with price discrimination across customers is hard to come across[10]. The study most like this one, Greenstein (1993), shows that an incumbent vendor is much more likely to be selected by the federal government in subsequent purchases of the same product. However, price is not a focus of his study; rather, he looks at the importance of product compatibility with past purchases.

There are other empirical studies with variations of the “bargain-then-ripoff” result, although none look directly at transaction-level behavior. Calem and Mester (1995) show that switching costs help explain why lower prime interest rates, which affect a bank’s cost of capital, are not immediately translated into lower credit card interest rates. Credit card issuers can effectively price discriminate between old and new customers, and the paper shows that interest rates fall more slowly for card issuers with larger installed bases. Elzinga and Mills (1998) use a dataset with 42 observations to show that new wholesale cigarette customers were more likely to benefit from an industry price war than existing customers, although the mechanisms leading to switching costs are very unclear. Borenstein (1991) attempts to show that margins on leaded gasoline tend to be higher than those of unleaded gasoline due to higher search costs for leaded gasoline; again, his data do not directly show repeat versus new purchases, although he does have a number of tests ruling out alternative explanations of the margin difference.

3.2 Product strategy to increase lock-in

A smaller literature examines how vendors use product breadth in an attempt to increase switching costs and the degree of lock in[11]. The key result from this literature is that

[10] A number of studies look at industries where vendors cannot price discriminate across old and new customers; Farrell and Klemperer (2006) contain a review. Most of these studies are for consumer goods and look at the impact of market share on pricing decisions, and do not look at individual transactions. An exception is Viard (2007), which looks at the impact of 1-800 number portability by businesses on switching and pricing behavior in the telecom market; however, in his study vendors charge the same price to all customers.

[11] As noted by Farrell and Klemperer (2006), there are very few studies on the effect of switching costs on product *quality*, although they speculate that high switching costs will lower overall quality due to the dampened competitive effects of having locked-in customers.

vendors will offer complementary products in order to increase the degree of customer lock-in. Most of the theoretical models in this space, such as Klemperer (1992) and Klemperer and Padilla (1995), actually consider shopping costs, where consumers face disutility from searching for products from separate vendors. Shopping costs are often defined as a type of switching cost, if the shopping costs must be repeated when purchasing the same basket of goods from separate vendors. However, there may be other types of switching costs, for example if there are economies of scale in training or implementation costs. These models predict that vendors will inefficiently expand their product lines, in that they are not a low cost provider for all products, because doing so will lead to greater lock-in. On a related note, the tying or bundling literature, such as Whinston (1990) and Nalebuff (2006), consider the anticompetitive effects of requiring purchase of one (often new) product in order to purchase another (often old) product where the customer faces switching or shopping costs[12].

The empirical literature has support for the basic hypothesis that increased product breath increases lock-in. Dranove and White (1996) show that shopping and switching costs across hospitals help explain size and pricing decisions by large hospitals. Chen and Hitt (2002) demonstrate that pricing, and arguably switching costs, are positively influenced by product breadth and quality in the online brokerage business. As with the price discrimination literature, these studies do not directly look at actual transactions, instead using posted market prices and data on market share or overall sales. There are also a significant number of case studies of the consumer software industry demonstrating how Microsoft used lock-in to existing products to prop up sales of new ones (Whinston, 2001; Bresnahan, 2001). Finally, the marketing literature has a number of studies that show that brand loyalty is higher for firms with greater product breadth; however, these surveys largely rely on survey data (Chen and Hitt, 2002), not actual purchase data, and do not attempt to address the underlying endogeneity issue (e.g. firms have higher product breadth because they had the resources to expand after strong initial products attracted loyalty).

3.3 Customer response to vendor lock-in strategies

[12] There are also many models of bundling which consider differences in consumer taste and willingness-to-pay; models considering switching or shopping costs are only a small subset of the bundling literature.

Surprisingly, there are few studies that consider customer strategies for avoiding lock-in. Customers are usually thought of as price takers who do not try to influence their real or perceived switching costs. In fact, nearly all the empirical studies of switching costs implicitly assume that "consumers myopically maximize current utility without considering the future effects of their choices" (Farrell and Klemperer, 2006).

The most comprehensive examination of these issues is in Greenstein (1997), who provides a detailed qualitative study of buyer reaction to switching costs and potential lock-in in the mainframe computer industry. Greenstein convincingly documents that buyers are very cognizant of the risks and costs of lock-in, and actively take steps to reduce these risks. Strategies include deliberately using legacy products for periods longer than the vendor presupposed, changing bidding procedures to reduce incumbent bias, increasing procurement oversight practices and data collection efforts, increasing training programs, and other steps. The study does not, however, explain how such efforts vary by customer type, or directly demonstrate how these behaviors affect pricing or switching behavior.

Other work looks at this issue in a more tangential way. One part of the theoretical model in Taylor (2003) considers customers who incur costly switching, only in order to lower prices in subsequent "ripoff" stages from later vendors, who update priors on the customer's true switching costs and lower "ripoff" stage prices to minimize the risk of another switch. Similarly, in an empirical paper, Cabral and Greenstein (1990) examine the effect of a specific strategy by the federal government procurement office to ignore some or all switching costs when considering procurement bids, and showed that it potentially decreased the price premiums incumbent computer vendors charged the government.

4 Theory development and hypotheses

In this section, I briefly outline the theory behind the empirical tests of the next section. Since the basic hypotheses are relatively simple and well established in the existing theoretical literature, no attempt at a formal model is made. However, I do outline the basic logic underlying the empirical tests using simple equations. The three areas of interest are "bargain then ripoff" price discrimination, the degree to which product breadth increases lock-in, and heterogeneous customer response to the vendor's pricing and product strategies.

I first review the basic results of Klemperer (1995), which provides a simple and tractable model of "bargain-then-ripoff" behavior in a market with switching costs. His two-period model has two symmetric suppliers, with costs c_t in periods t=1 and 2. Note that c_1 does not necessarily equal c_2; if c_2 is thought of as after-sales service, spare parts, or other such costs, c_2 could be significantly smaller than c_1. However, since costs are not central to the story in this paper (as marginal costs in enterprise software are close to zero), I will assume constant costs and drop the subscripts.

On the demand side, a single customer has reservation utility of r_t for consumption of the relevant good in period t, and switching cost s should she switch providers in period 2. It is assumed that $r_t > c$, so there are gains from trade in each period, and that each supplier gives a "take-it-or-leave-it" offer to the customer. The customer then accepts one offer (or no offers) in each period.

Backwards induction yields the "bargain-then-ripoff" result. With perfect information, customers will purchase the second good from the same supplier it purchases from in period 1. The incumbent supplier, with full information about s, will price (just below) c + s in Period 2, to take advantage of the market power provided by the switching cost. The non-selected supplier will not price below c, and the customer will not switch at this price. Foreseeing this market power, both suppliers are willing to price at *less* than c in period 1, in order to reap profits in period 2. Competition results in first period pricing of (slightly more than) c - s.

Assuming multiple customers but separate games with each customers, a vendor will always sell at a price below cost to a customer new to it, and use the pricing power arising from the switching costs to price higher than cost in subsequent periods. In the model's simple notation:

$$P_1 = c - s \text{ for new customers} \quad (1)$$

$$\text{and} \quad P_2 = c + s \text{ for existing customers} \quad (2)$$

Therefore, customers new to a vendor pay less than existing customers. This basic intuition yields the paper's first hypothesis:

H1: For the same product, vendors will charge existing customers of a product's previous version more than they charge new customers to the product.

There are a couple of other points worth noting. First, this simple model predicts customers will never switch. Secondly, the degree of the initial "bargain" (and the later "ripoff") is directly related to the switching cost of the customer. In this case the difference in price between new and locked-in customers is twice the switching cost, since subtracting (2) from (1) yields $P_2 - P_1 = 2s$.

The next hypothesis reviews the effect of multiple products sales when switching costs across products are shared. The basic idea is that the difference between "bargain" and "ripoff" pricing will be reduced for customers who purchase several products from a vendor when switching costs are shared between these products.

To see the intuition behind this hypothesis, assume three vendors exist for products m and n. Vendor A sells both products, while the other vendors are specialists, Vendor B of product m and Vendor C of product n. Assume some switching costs can be shared across products purchased from the same vendor, so that $s_m + s_n > s_{m+n}$. Also assume costs are symmetric, so the multi-product firm's costs are exactly double each of the single-product firms' costs.

This simple model cannot determine whether customers purchase from Vendor A or from Vendors B and C; in both cases net pricing and customer welfare is the same. However, it does predict that customers who choose to purchase both products from Vendor A will pay a higher initial price per product than customers who choose to purchase only one product from vendor A, and one product from Vendor B or C.

From equation (1), we know that a customer purchasing the products separately will pay in period 1:

$$P_{1m} = c - s_m \text{ to Vendor B} \tag{2}$$

$$P_{1n} = c - s_n \text{ to Vendor C} \tag{3}$$

$$P_{1m,n} = 2c - (s_m + s_n) \tag{4}$$

Customers who purchase both products from Vendor A pay:

$$P_{1m,n} = 2c - (s_{m+n}) \tag{5}$$

Since $s_m + s_n > s_{m+n}$, we know that (4) < (5), or customers purchasing from a single vendor pay more per product in period 1 than customers purchasing from separate vendors. Similar intuition would show that the price in period 2 is lower for customers purchasing from a single vendor. Essentially, purchasing multiple products with shared switching costs from lessens the need for bargain-then-ripoff behavior. This leads to a second hypothesis:

> **H2: Customers who purchase more than one product from a vendor, or who have previously purchased another of the vendor's products, are given smaller initial "bargains" for these products than customers completely new to the vendor; they also pay lower "locked-in" rates for both products than existing single-product users.**

I now introduce customer heterogeneity in switching costs to the equation. With perfect information and the ability to price discriminate, heterogeneous switching costs do not change the basic result around price discrimination and lack of switching. However, the story changes when vendors are uncertain about exact switching costs. First, as demonstrated by Chen (1997) and Taylor (2003), some customers will switch providers in equilibrium. In both models, only customers with low switching costs ever switch vendors.

It is impossible to observe customer switching costs; however, studies in the Information Systems literature have shown that there is a strong correlation between the strength of a company's overall IT department, as measured by vendor and analyst surveys, and how cost effective the company is in working with systems integrators and third-party service providers. These studies use cross sectional regressions to show that companies that are judged by analysts as effective IT users spend less than predicted on outside professional services than companies who are judged to be poor at IT, controlling for industry, company size and the like. As noted in the previous section, professional customization, training and business process redesign are a significant portion of the switching costs companies incur when implementing a software package. A related study shows that firms judged capable at IT are more comfortable with the risks of starting projects with new providers, so they arguably have lower psychological switching costs as well. Furthermore, these costs and risk tolerances are not identifiable ex-ante, so a vendor will not know with certainty which customers will have low switching costs[13].

Therefore, existing theory predicts only customers with low switching costs actually switch providers when vendors are uncertain about true switching costs, and the applied IS literature contains evidence that survey results on IT capabilities predict lower-than-average spending in a major source of switching costs. A clear hypothesis emerges from this line of thinking:

H3: Customers who have stronger IT capabilities are more likely to switch vendors.

Customers with stronger IT capabilities, therefore, are more likely to take advantage of "bargain then bargain" pricing across vendors instead of "bargain-then-ripoff" pricing from a single vendor.

There are two competing explanations for this result; I note them here for interest, but I am agnostic as to which is the causal mechanism leading to switching. The first is a "reverse lemons" or adverse selection result, where only "poor" customers find it profitable to participate in the "ripoff" stage of the market. The second is Taylor's signaling result, where customers strategically switch, even when it is unprofitable in the short term to do so, to affect vendor belief about their true switching costs and reduce later-stage "ripoff" behavior.

Finally, I develop hypotheses on the effect of product breadth on switching behavior. As shown above, customers who have purchased multiple products with shared switching costs should pay a lower price in the "ripoff" period than those who purchase single products

[13] Vendors can and do read the same set of surveys used by these studies, but cannot estimate the difference in switching costs between customers strong and weak at IT management. Also confounding this result is, as discussed in the next section, the fact that customers with strong IT capabilities are much more likely to be customers with strong revenue and profit growth, who are therefore more attractive to vendors, leading to a higher willingness to discount in the hopes of winning future business.

from the vendor. With uncertainty around true switching costs, the reduced prices paid by these customers should reduce their likelihood of switching to single-product vendors This is because single-product vendors, selling products with higher average switching costs, will have a smaller scope to induce switching since customers will know their ripoff-stage prices will go up. This leads to the paper's final hypothesis:

> **H4: Customers who have purchased multiple products from a vendor are less likely to switch vendors for any product**

It is interesting to note that I do not test a natural extension of hypotheses 3 and 4, that customers with strong IT capabilities are less likely to purchase multiple products from a single vendor. However, this result is a longstanding qualitative finding in the Information Systems literature, and the combination of these two hypotheses may provide insight into the mechanisms behind this widely-accepted conventional wisdom.

5 Data, estimation strategy and results

5.1 Data

The data for this study were provided by a leading enterprise software vendor, representing individual transactions for all sales in North America for two of its major product lines between 1997 and 2002. These product lines were selected because the vendor is perceived to have strength in both, reducing concerns about unobserved differences in quality biasing the results[14]. In total the dataset contains 1,852 purchases made by 756 customers over the course of 22 financial quarters. The database excludes two types of purchases for the selected products: deals under $50,000, which are usually small add-on purchases sold at or near list price, and "site license" deals, which give the customer the right to use as many licenses as it wishes for a particular product. Site license contracts were not available, and much of the data used later in identification is not relevant to them, since, for example, there is no notion of the level of discount granted. Still, deals in the database account for approximately 80% of the revenue for these two product lines.

[14] I do attempt to control for sub-product quality, as discussed below.

It is important to note that I focus only on license revenue in the empirical analysis. Maintenance services are priced at a fixed percentage of list price for all products, and salespeople have no control over their discounts[15]. It is therefore not meaningful to investigate pricing behavior for maintenance services. (From an econometric point of view there is no variation in these data, making them unusable.) The purchase of professional services may impact both pricing and switching costs, but the vendor explicitly avoids pricing these services below marginal cost, even to win a new account, and competitive pressures keep prices for these services from spiraling too high[16]. I therefore only use professional service spending as a control.

I augmented the transaction database with two other sets of data. First, I added publicly available information on customers, including, where available, analyst estimates of a customer's IT capabilities, taken from a proprietary meta survey collated by Archstone Consulting[17]. Second, I conducted follow-up interviews in 2006 to determine whether each transaction in the database led to the sale of an upgrade in the period between the end of the database in 2002 and the follow-up interview. For each transaction, I determined whether the customer bought an upgrade in that timeframe, switched to a competitor's product, or was still using the legacy product[18]. The final dataset contains five classes of information:

1. Sale information, which includes products bought (licenses, maintenance and services), list price, price paid, and whether the customer bought products from multiple lines on the same purchase order[19].

---

[15] More correctly, a salesperson can offer an informal discount on maintenance, but must make up for that discount in the form of a lower license discount, since sales managers could approve the latter but could not approve a discount on maintenance. Therefore, all maintenance sales are at list price in the database. Maintenance is a constant percentage of list price for all products. Hence it may be more proper to think of discount from license list price as a blended discount from combined license and maintenance list prices.
[16] Professional services contracts always give a FTE estimate, meaning outrageous pricing would quickly be noticed by the customer, which is almost surely using other professional IT service firms on other projects.
[17] These data were available for 421 of the 756 customers in the database.
[18] All products in the database had a major upgrade between 2002 and 2006. However, there will be a censoring problem with this set of information, since some customers who are observed to have not upgraded may have subsequently upgraded or switched vendors, or may do so in the near future. This is discussed at greater length below.
[19] Customers usually buy multiple products from within the same product line on the same purchase order; these are not counted as multi-product sales. I extended the definition of multi-product sales to include purchases from any other product line within one year of the purchase order in question, and the empirical results did not change.

2. Product information, which includes the product's age (the number of quarters since version 1.1 was introduced), and two measures of product quality. The first is the vendor's own estimate of the product's market share, compiled by its product marketing teams. The second is a dummy variable indicating whether the product in question was rated a leading product by the widely-used Gartner MagicQuadrant reports.
3. Customer status, which is a series of variables indicating a customer's history with the vendor, product line and specific product in question, and some information about the customer's use of other vendors. This category, which will act as the key explanatory variable in the empirical strategy, is discussed in more detail below.
4. External customer information, including name, number of employees, revenues, market capitalization, and the aggregate survey data on customer IT capabilities, where available.
5. Internal vendor factors that could affect price, including salesperson compensation concerns (as defined in the previous chapter of the dissertation), vendor financial performance in the quarter in question, and so on. In addition, all the empirical analyses use quarter fixed effects.

Before summarizing the overall data, it is worthwhile to clearly explain the "customer status" data, which measure for each transaction the depth of previous customer interaction with the vendor's products. Figure 2 provides an overview of these data. As noted, a majority of transactions involved customers who had previously purchased at least one product from the vendor; only 19% of transactions were from completely new customers. Nearly all completely new customers were classified in the price book as being a "competitive switch," which in most (but not all) cases meant the customer received an immediate list price reduction on the product they purchased.

For transactions involving customers with a previous purchase history, 42% included an upgrade of an existing product. (Note that many of these purchases involved other products as well, as the vendor was adding many new product options during the time in question.) However, for only about a quarter of "upgrade" transactions did the customer receive an "upgrade credit" to list price. As noted, upgrade credits were usually given only if a customer had purchased the previous version of the product quite recently; essentially they are used to motivate customers to buy older products for which upgrades are imminent, and many customers prefer to wait until the new version is launched.

The remaining 58% of transactions were with "existing customers" for the purchase of new products only. These are split approximately equally between transactions where the customer bought a product new to it within the same product line as a previous purchase, and transactions involving a completely new product line for an existing customer. Note very few customers buying in the same product line received a "competitive switch" credit, while nearly a third of existing customers buying a new product line received this credit. Overall, 27% of transactions were classified as a "competitive switch," but over one-third of these purchases were made by the vendor's existing customers.

Table 1 shows summary statistics on these figures and the rest of the dataset. As noted, the average deal is large, nearly $800,000, and heavily discounted, at over 34%. Few purchase orders – only 13% – contain products from multiple product lines, although 76% of orders contained purchases of multiple products within the same line[20]. As noted, I deliberately selected the vendor's strongest product lines, so the average product age at the time of sale is relatively old at over five years, considering the age of the sector [21]. The vendor has a very high average market share of 36%[22].

Finally, by 2006, customers had upgraded at least one of the main products involved in 46% of these transactions[23]. For 40% of the transactions in the database, the customer was still using legacy systems. For a full 18% of transactions, the customer had switched to another provider.

Overall, the data suggest that customers use multiple enterprise software providers, and at least some of them switch among them relatively frequently. The vendor extended

[20] Almost always these are coded by the vendor as a single product purchase (because the price book lists almost every conceivable product combination as a single product), but the product in question is in fact a bundle of individual products.
[21] The age is calculated according to the *oldest* product bought in the purchase.
[22] Again, this number is calculated by using the product with the highest market share in the purchase. This is the somewhat optimistic methodology the vendor used to measure its market share.
[23] These data were collected by interview with senior sales managers, and involved some qualitative judgment on their part as to what constituted an upgrade of a "main product."

“switch” credits for over one-quarter of the transactions in the database, and customers for 18% of transactions switched away from the vendor subsequent to purchase. And even if they aren’t switching, most customers are only using this vendor for one or two product lines, meaning they are using other vendors for lines that this vendor could supply. This is somewhat less surprising in view of the average size of customers: customer revenue averages over $20 billion annually and the average customer had over 38,000 employees in 2006.

5.2 Estimation strategy and results

This section reports the empirical strategy and results for my tests on price discrimination to customers of different types, and determinants of the likelihood of switching. For the first category (H1 and H2), I use Ordinary Least Squares (OLS) regressions to predict pricing. For the second category (H3 and H4), I use Maximum Likelihood techniques to predict the probability of switching.

5.2.1 Price discrimination and lock-in (H1 and H2)

The data in Figure 2 and Table 1 suggest several straightforward empirical tests of the effects of lock-in on pricing decisions. Does the vendor engage in “bargain-then-ripoff” behavior according to customer status? Does it need to give large initial “bargains” to existing customers that are purchasing products they have not used before?

Formally, I model the determinants of transaction price as based on customer status and controls:

$$Y_i = \beta_1 + \varphi * CS_i + \beta_2 * \Omega_i + \varepsilon_i \qquad (6)$$

where the subscript *i* refers to the transaction observation, Y refers to the price measure, CS refers to measures of customer status, Ω_i is a vector of deal controls, and ε_i is the error term.

However, there are two challenges in this exercise. The first is to come up with a useable definition of price. As noted, the same sub-product can carry hundreds or even thousands of list prices depending on the specifics of purchase, making direct comparisons of price paid per license very problematic. The second challenge is that customer utility, which undeniably influences price, is not observable and therefore is part of the error term in my regression. This gives rise to the "Heckman problem" where it is unclear whether the empirical results are due to the factors modeled or due to unobservable customer heterogeneity that is correlated with the variables of interest. I first address the problem of defining price and, once I have derived results, return to the problem of unobservable utility.

As noted in the previous chapter of this dissertation, the most logical measure for price is the discount granted to the customer. This normalizes all sales by list price and provides a unit which is directly comparable across operating environments, hardware selected, product add-ons selected, the product "grade" selected, and other key factors.

However, the vendor sometimes changes list price based on customer status, which could bias my results. For 9% of the transactions in the database, an existing customer received an "upgrade credit," lowering its list price and decreasing the discount the vendor can give to achieve the same dollar in license revenue. Similarly, 19% of transactions involved a "competitive switch" credit[24], which also lowered the list price of the transaction.

In terms of identifying "bargain then ripoff" behavior, the bias introduced by these credits goes both ways. In the case of upgrade credits, customers with small observed discounts may in fact be getting a good deal, if the credit was large, yet their discount would suggest they were getting "ripped-off." In the case of competitive switch credits, customers may also receive a small discount, indicating they did not get a "bargain," when in fact the credit meant they received a very good price.

Table 2 shows some basic statistics on median upgrade credits and average discounts by customer status. As noted, the median upgrade credit for the 9% of transactions involving an upgrade credit within the same product line was 30 points off of the "standard" list price, and transactions of this type received, on average, a discount of 21%. For 25% of the transactions, a customer upgraded yet did not receive a credit, and the average discount rises to 28%. This is high-level evidence that ignoring the effects of upgrade credits on discount will introduce bias to the results. Similarly, discounts are lower in all cases for transactions involving competitive switch credits when compared to transactions with customers of the same status but where no credit was given.

[24] 27% of transactions were classified as a "competitive switch," but in some cases list price did not change. The vendor has asked that I not disclose the reason for this discrepancy, although I did take it into account in the econometrics.

I use three separate techniques to deal with this problem. In the first, I use discount as the dependent variable, and interact a dummy variable for whether the customer received an upgrade credit with the customer type dummies. If the model is correct, the resulting coefficients will be an unbiased indicator of *discount* levels, even for transactions involving list price credits. However, they will not be indicative of bargain-then-ripoff behavior for transactions involving credits, because the upgrades bias the translation of those to meaningful price figures.

In the second test, I run the same regression but exclude all observations involving list price credits; therefore, list price is "at book" for all transactions. In the third test, I construct an "adjusted discount" figure by adding back the credit to the product in question and re-calculating the discount. The second test discards data and lessens the power of the empirical tests, while the third test adds noise to the data. It is hoped that the three techniques can be triangulated as a check for robustness.

Referring again to the estimating equation:

$$Y_i = \beta_1 + \varphi_1 * CS_i + \beta_2 * \Omega_i + \varepsilon_i \qquad (6)$$

I define CS_i as a series of dummies identifying a customer's status in a mutually exclusive way. The excluded status is "existing customer, upgrade, no upgrade credit." The variables receiving dummies are:

1 0 0 0 0 0 0	for existing customer, upgrade, upgrade credit
0 1 0 0 0 0 0	for existing customer, same product line
0 1 1 0 0 0 0	for existing customer, same product line, switch credit
0 0 0 1 0 0 0	for existing customer, different product line
0 0 0 1 1 0 0	for existing customer, different product line, switch credit
0 0 0 0 0 1 0	for new customer
0 0 0 0 0 1 1	for new customer, switch credit

Note that the structure of the dummy variables essentially interacts the presence of a list price credit with the appropriate customer status dummy, so the effect of the credit is additive.

The controls include product line[25], operating system, customer industry, and salesperson region dummies; the deal's total purchase price; other elements of the deal such as spending on services and product quality; salesperson tenure; internal vendor concerns such salesperson compensation differences (per the previous chapter of this dissertation) and whether the deal comes in the vendor's final quarter of the fiscal year; and finally customer characteristics. The need for product, customer industry and region dummies is clear; all of these could be related to a deal's discount, regardless of lock-in. The vendor may discount more if a buyer orders more professional services, and may discount more for products of lower quality or with a lower market share. I control for purchase price as the vendor is very likely to offer bigger discounts for larger deals, due to the zero marginal cost environment. I control for salesperson tenure since there may be differences in the propensity and ability to discount. I control for salesperson compensation concerns because they can critically affect a salesperson's willingness to discount. I control for the final quarter of the vendor's fiscal year because executive pay largely depends on fiscal year-end stock price, and executives may be more willing to discount heavily in the fourth quarter. Finally, I control for customer characteristics, most notably different variables indicating size, growth[26], and IT capabilities, since the vendor may be willing to discount more heavily to customers who are larger or growing more quickly. As noted above, I also control for customer industry.

Table 3 presents an overview of the results. In model (A) I use the entire sample, dropping the survey data on customer IT capabilities, which is only a control in these tests. It shows significant evidence of bargain-then-ripoff behavior; new customers who do not receive upgrade credits receive discounts of nearly 10 percentage points higher than customers upgrading their products, if the new customer does not receive a switch list price credit. Hypothesis 1 is clearly supported, as new customers are getting bargains.

Hypothesis 2 gains only partial support; the vendor appears to avoid having to offer bargains to customers purchasing new products in the same line, as that coefficient is not significantly different than zero. However, existing customers purchasing outside their current product line receive discounts of over 7 percentage points more than upgrading

[25] I actually control for the 4 major pieces of each product line, so the empirical tests only compare discounts within these sub-lines.
[26] Several different measures of size and growth were used as controls; the tables only report a few of them.

customers, provided they did not receive an upgrade credit. It appears that switching costs do not operate across product lines.

Credits to list price clearly act as a damper on discount behavior; three of the four credit types have statistically significant negative effects on discounts. However, even taking the effects of list price credits into account, new customers, and customers new to a product line, still get discounts that are larger than upgrading customers; remember that the effects of customer type and the customer type interacted with the credit dummy are cumulative. In both cases the sum of the coefficients is still positive, indicating that despite the credits to list price (which are 40 percentage points at the median), new-to-vendor or new-to-product line customers still are getting higher discounts than upgrading customers.

There are other interesting results in the data. First, the scope for "ripoffs" does not appear to go up as a customer upgrades more; there is a negative and non-significant effect on the cumulative number of upgrades. Also, customers purchasing multiple product lines in the same transaction are not given smaller discounts, all else held equal; this again suggests that these customers face independent pricing decisions, and that switching costs do not carry over across product lines. The vendor predictably has more power to control discounts on older products, which are probably more successful and perhaps of higher quality. This may explain why neither of the quality measures appears significant. Finally, bigger, better performing customers achieve bigger discounts; the vendor appears willing to attempt to attract "better" customers by discounting more heavily.

Columns (B) and (C) report the same tests for the sub-sample of transactions where the customer's IT capabilities are measured in the survey. The results are largely the same; the (positive) effect of IT capabilities on discounts goes away when financial data are inserted as well. In fact, the IT capabilities data are highly collinear with the performance data, which is a classic result in the information systems literature. Model (C) suggests that the vendor looks at performance, not IT capability, when making the discount decision. Finally, model (D) strips out some of the controls that had no effect across models; again, the results stay essentially the same.

Table 4 reports results from the same test using the second technique for handling list price credits: dropping observations where these credits were used. I therefore also drop the variables that interacted customer type with receipt of a credit, and I use the additional room in the table to report a few more controls of interest which did not change dramatically from the first specification.

Although measurement is less precise due to the smaller sample size, the effects of interest are of comparable size and are mostly statistically significant. Customers buying new products within the same product line do not receive bargain pricing, consistent with the hypothesis that some switching costs are shared within product lines. But customers buying in a new product line, or that are new to the vendor, receive discounts that are approximately 5 to 10 percentage points bigger. (The existing customer, new product line result is not significant when looking only at the IT capabilities sample, but again this is likely due to the smaller sample size in this specification.)

In the final specification, reported in Table 5, I attempt to adjust the discount on transactions where a list price credit was used, so that the discount given is an accurate reflection of dollar capture by the vendor. For competitive switch customers this is relatively non-controversial, because the credit given to these customers is a pure "bargain" to induce them to ignore their switching costs and buy from the vendor. However, for upgrade customers the story is less clear; customers only receive an upgrade credit when they have purchased the last version of the product, and usually only when they do so relatively late in the old product's lifespan, or very early in the lifespan of the new product. The rationale for this credit, therefore, is arguably not around bargain-then-ripoff behavior, but around increasing incentives to purchase very old and very new products. Nevertheless, I adjust discounts on these transactions as well to be consistent.

To calculate the adjusted discount, I recalculate a "fair" list price by adding back to the list price the amount of the credit. For products where the list price credit matches exactly the product being bought, this is straightforward. However, quite often a buyer would receive a "customized" list price credit based on a salesperson's view of the customer's current product functionality and what the customer was buying from the vendor[27]. The salesperson then got the proposed credit and discount approved by her superiors. The vendor's order system, however, would require a valid SKU (which

[27] This usually occurred because products made up of similar elements had a competitive switch credit, while the exact product the customer wanted did not.

automatically nets out a credit), and would not accept exceptions[28]. Salespeople would therefore look up a SKU that had a list price resulting in the desired post-credit discount (and therefore total order value) they were trying to achieve, even though the products being ordered did not match. These errors in the sales database, from which I received my data, were corrected at the invoicing stage, and unfortunately I do not have access to invoicing data. I do know, however, when a credit was "fudged," as it is clearly indicated in one transaction field. I therefore must rely on incomplete salesperson notes and my own knowledge of the price book to try to "correct" these problems[29].

The "adjusted discount" affects the 17% of deals with list price credits, in each case raising the discount on those transactions. Table 5 shows the results of the specification with these changes. Of course, the interaction terms are no longer necessary, since the list price corrections were designed to put transactions by customer status on the same terms. Again, the results are broadly similar; what is interesting is the discount difference for new customers approaches 15 percentage points in the full models, and is still statistically significant. Customers new to the vendor were most likely to receive competitive switch credits, and it appears when these are taken into account, these customers are getting even bigger "bargains." According to these results, customers with no previous vendor experience receive discounts that are 15 percentage points larger than average, which is nearly half the average discount in the sample. Therefore, these customers' discounts go up by a factor of 50%!

The discount difference between upgrade customers and customers new to a product who are purchasing in the same line becomes negative (remember that almost none of these orders received a credit, so nearly all discounts were the same), but is still not statistically significant. The discount difference between upgrade customers and customers new to a product line is smaller than in the original model, but the difference is still statistically significant, and it appears there are not economies to switching costs across product lines.

[28] This was the vendor's own order system – I love how this is a perfect example of how terrible enterprise software is.

[29] This is why these data are so hard to clean – I didn't worry about this problem before because I was accepting discount numbers as given, and not worrying about credits. Luckily this isn't a gigantic problem – it adds error to maybe 5% of transactions, and the salespeople tell me they always picked products that were "close" (e.g. in the same product sub-field), so it doesn't screw up my dummies. I've also cleaned out a lot of it via the salesperson notes field, where they (sometimes) say exactly what they are doing. Maybe I should just leave this long-winded explanation out and say there's some error in these re-calculations?

The basic results appear robust and consistent across specifications: the vendor is using bargain-then-ripoff behavior, but buying multiple products in different lines does not appear to reduce switching costs for customers. Switching costs do appear to be shared by customers using products within the same product line.

The main alternative explanation for these results, as mentioned previously, is the "Heckman problem" that arises due to unobserved utility. Specifically, customer utility may be correlated with prior purchase history. Customers with high valuations of a product would arguably buy a product earlier in its life-cycle, and thus would enter my dataset as existing customers. Since these customers have higher utility, they are naturally willing to pay more for a product. The data may not represent "bargain-then-ripoff" behavior, but instead demonstrate a textbook example of third-degree price discrimination; indeed, it is a common strategy in technology industries to sell early versions of a product at a high price to those with high valuation of the product, and sell slightly different products to the "mass market" as the product wins acceptance.

A classic method for dealing with such problems is the use of instrumental variables; however, it is difficult to think of a variable uncorrelated with utility but correlated with a customer's prior product choice. I therefore turn to a more indirect method to try and test the basic results in environments where the price discrimination story is less plausible.

The approach is to restrict the sample to products which have only one or two upgrade cycles. Any intertemporal utility difference would be lessened by the relative newness of the product. Also, quite conveniently, the vendor introduced a number of completely new applications during the time period covered by these data that were designed to take advantage of the rise of the Internet, such as software designed to aid business-to-business procurement or collaborate in real-time with suppliers and other partners. These products were purchased by a significant number of existing customers, but also by a large set of new (and newly-formed) customers. In total, 456 of the 1,852 transactions are for products that have had 2 or fewer upgrades.

If intertemporal price discrimination is leading to the low initial price for new customers, we should not see a difference in discount by customer status when focusing on these new products, since all customers would arguably have high utility[30]. Table 6 shows

[30] Thanks to Scott Stern for coming up with this idea.

the same specification run only on the sub-sample of new products. I use adjusted discount as the dependent variable, although very few of the transactions received any upgrade; the results were almost exactly the same using original discount.

The results are consistent with the original story; customers who are upgrading receive comparable discounts to those extended to first-time product purchasers who have purchased within the same line, but customers new to the product line or new to the vendor receive discounts which are much larger. Significance again drops, but stays below the critical 10% significance threshold. Interestingly, receiving a high mark from Gartner predicts a much lower discount; it could be that for early products, getting favorable analyst reviews helps pricing significantly. Overall, it appears that a systematic difference in customer utility correlated with product status is not driving the discount behavior; rather, product status alone is driving it.

5.2.2 Who switches? Modeling switching likelihood by customer type (H3 and H4)

Having demonstrated that the vendor is engaging in "bargain-then-ripoff" pricing behavior, but that the lock-in achieved via this strategy only appears robust for customers upgrading their product or purchasing within the same product line, I now turn to hypotheses concerning customer switching behavior. I test whether high customer IT capabilities increases propensity to switch, and whether the use of multiple products decreases this propensity.

To test these hypotheses, I turn to the data collected via interviews, which indicated whether customers for every transaction chose to upgrade their product line, continued to use a legacy system as of 2006, or switched to another provider. Each customer has a choice of three discrete, non-ordinal states, so the multinomial logit is the logical specification from which to run this test[31]. Formally, I model:

$$\Pr(C_i = J) = f(IT_i, CS_i, \Omega_i, \varepsilon_i) \qquad (7)$$

$$J \in \{U, L, S\}$$

[31] See the previous chapter of this dissertation for a discussion of the assumptions behind the multinomial logit specification.

where C represents the observed choice of the customer; U, L and S refer to upgrade, use the legacy application, or switch, respectively; IT refers to the customer's IT capabilities (H3); CS refers to the customer state with the product and vendor (H4); Ω represents a vector of controls; and ε represents the error term.

Under my hypotheses, a customer with strong IT capabilities should be more likely to switch, while a customer for whom the transaction in the database was not the first would be more likely to upgrade or use the legacy application. I use a full set of controls, very similar to those used earlier, to ensure that I capture the effects of all relevant data.

Table 7 reports the estimation results. The estimated coefficients in the multinomial logit specification do not carry economic meaning (Horowitz and Savin, 2001), and only the signs and significance of the coefficients, and the significance of the overall regression, are of interest. To estimate coefficients with economic meaning, it is necessary to calculate the marginal effects of the explanatory variable on the outcome variable at the average value of the explanatory variable. The results of these calculations are reported in Table 8 These values can be interpreted as the change in probability of choice *j* due to a one-unit change in the value of the independent variable, evaluated at that variable's mean. These changes in probabilities are most usefully reported in comparison to a "baseline" choice, which in this case is to continue to use the legacy application.

Interestingly, H3 is not supported; customers with strong IT capabilities are, if anything, more likely to use legacy products than to switch or upgrade. A one unit increase in the survey's measurement of IT capabilities increased the probability of using the legacy application by 9% versus switching, and by 12% versus upgrading. This result is consistent with the IS literature, which suggests that customers with greater IT capabilities are more comfortable using older products, and products from more vendors.

Similarly, H4 is not well supported. There are few meaningful differences in probability of switching or upgrading by customer type, except that brand new customers are much more likely to switch, and existing customers new to a product line are slightly more likely to switch or use legacy systems than to upgrade. Since the variables in question are dummies, the coefficients can be interpreted as the change in probability of the two outcomes from moving from "0" to "1." A customer new to the vendor is 24% more likely to switch than to use a legacy product, which is a gigantic effect. Customers purchasing a new product within the same product line are not statistically more likely to upgrade, use legacy products, or switch. Overall, these results reinforce the negative message around product breadth and switching costs in the earlier section; customers using multiple products in the same line appear no less likely to switch than single-product customers.

One other interesting result is the strong effect that financial performance appears to have on the likelihood of switching. For every percentage point increase in the customer's 5-year cash flow change, it has a 7% greater probability of switching than using a legacy product, and a 3% greater probability of using a legacy product than upgrading. While this study did not form hypotheses around these results, they are nevertheless fascinating from a vendor strategy point of view, as they suggest the best performing customers are most likely to switch and the least likely to upgrade their products[32].

Of course, these results do not definitively demonstrate a causal link among product experience, IT capabilities, customer growth and switching behavior. Some unobserved factor could be influencing these results. However, at least two of the above factors – financial performance and IT capability – seem exogenous to a firm's approach to to dealing with the upgrade of a purchased application. Similarly, the original purchase was made several years before the subsequent choice was made. Therefore, endogeneity does not appear to be an issue with these regressions, and it is difficult to think of easily-identifiable variables that could be added to the story.

6 Discussion

This research evaluates vendor strategy and customer response in light of the significant switching costs in enterprise software. It provides relatively strong evidence that, as one analyst put it, "buyers are in control" in this industry (Riciutti, 2004). Over one-quarter of the customers in the database exhibit switching behavior, either to this vendor in the purchase observed, or from the vendor after making the observed purchase.

There are two aspects of this switching behavior which are particularly worrying from a vendor perspective. The first is that the customers who are most likely to switch –

[32] I tried multiple measures of size and financial performance, and when run separately they come out with similar effects. Also, when all financial and size metrics are dropped, the IT capabilities variable switches sign, and predicts switching over legacy use.

those with strong financial performance – are also customers receiving the best initial deals. As noted, they appear to receive discounts that are 50% greater than average, which represents a truly staggering amount of foregone revenue. (Of course, whether these customers would buy from the vendor in the first place without such high discounts is unknown.) Although marginal cost is zero, the industry's significant fixed product development costs probably mean that these customers are being served at less than average cost[33]. In short, current pricing norms in enterprise software almost surely rely on vendors being able to take advantage of a "ripoff" stage in order to reach long-term profitability. Customers who pay "ripoff" prices are in effect subsidizing those who switch; with the rise of the Internet and more scalable computing platforms, it will be interesting to see how long these customers remain content to do so.

Second, and just as worrying, this study directly contradicts findings from studies in the consumer market on switching costs. Varian, for example, marks the findings of Chen and Hitt (2002) (in the online brokerage market) as highly significant, "since breadth of product offerings is under control of the firm; if a variety of products can be offered at a reasonable cost, then it should help in reducing the likelihood of customer switching." However, the strongest factors affecting the likelihood of switching in these results are largely outside of a vendor's control, such as financial performance and customer industry[34]. Product breadth is not a predictor of lock-in, and providers seem to generate no ability to increase price or reduce switching by offering multiple product lines. Again, this should not be surprising when industry structure is examined; this is a very diffuse, specialized industry with numerous single-product providers and many large players. However, recent acquisitions sprees by several medium and large players deserve scrutiny (at least from the perspective of achieving lock-in, which is one commonly cited rationale) in light of these results.

The results suggest that vendors should think about two broad strategies in an attempt to increase revenue capture, in light of the significant switching costs arising from using enterprise software. The first is improved pricing practices. It is clear that vendors need to more carefully control discount behavior, particularly to customers purchasing

[33] EBITDA for a successful enterprise software vendor is typically 10-20%, reinforcing this viewpoint.
[34] The customer industry dummies were quite significant in several cases; however, I did not report these in the results since they were not part of the hypotheses generated.

products lines additional to those that they already are using. "Competitive switch" credits only appear to worsen the problems of rampant "bargain then bargain" behavior.

Just as importantly, the results suggest that vendors should focus on developing strength within product lines in the first instance, rather than expanding product breadth. The results do suggest that lock-in increases as customers use multiple point products within a particular product line. If there are true implementation, training and product use synergies across products, then pricing should improve; the study suggests that these synergies do not exist outside of product lines, despite vendors' efforts to make product functionality and use similar across lines.

Of course, this is a quickly innovating industry. It is yet to be seen whether the scaleable, "one size fits all" push by some vendors will rationalize software pricing; it could also increase switching behavior by reducing switching costs. Similarly, many customers are pushing for entirely new pricing models, for example based on "on demand" rentals of software, rather than purchases.

From the customer's viewpoint, the study suggests that bargains can be had by being willing to switch vendors, and using multiple vendors within and across product lines. Again, this matches existing conventional wisdom in the IS literature, and the corroboration provided by these findings speaks to the validity of the results.

This study has some obvious weaknesses; probably the strongest is the relatively short time period covered by the data, and the lack of robust data on outside options for customers. Without these additional data, it is very hard to construct tests that rule out the "Heckman problem" or other alternative explanations. Still, this study is an important first step, as pricing in B2B environments is understudied. Similarly, increased customer data, such as costs paid for implementation, customization, training and the like would deepen the analysis of customer behavior and the factors leading to switching. Finally, data from several other vendors, hopefully of different sizes and product foci, would address questions about the validity of the results.

Finally, the study suggests several interesting next steps. One clear area for future research is to analyze the reasons for heterogeneity in customer switching costs, and whether and why these differences persist in spite of competition. One interesting avenue suggested by the results is the link between a firm's financial performance and its willingness to switch providers. The causal mechanism, if any, behind this correlation is unclear. Firms could be sending signals about having low switching costs; vendors may be trying to extract "ripoffs" that are too large; something else may be going on; or the correlation could be spurious. However, the strong correlation between financial success and pricing on the one hand and switching on the other makes this an interesting research avenue. Other potential reasons for low switching costs, such as best practices in training, careful customization, regimented product use, and so on, all bear further investigation. Greenstein (1997) talked qualitatively, but persuasively, about many of these factors, but his article has not generated more quantitative research.

Finally, from both the vendor and customer points of view, understanding the "shared" switching cost story would be of considerable benefit. It would help vendors design and sell products that increase lock-in and involve smaller up-front costs for customers. As noted in the introduction, customers report with near-unanimity that switching costs end up higher than they project, and they would clearly be interested in methods to share these costs across products. The theory discussion predicted that shared switching costs would lessen the need for "bargain-then-ripoff" behavior, and the overall results suggest that this result would likely benefit all parties in enterprise software.

References

Borenstein, S. (1991). "Selling Costs and Switching Costs: Explaining Retail Gasoline Margins," *RAND Journal of Economics*, 22: 354-369.

Bresnahan, T. (2001). "The economics of the Microsoft case", mimeo, Stanford University, Department of Economics, retrieved at: http://www.stanford.edu/~tbres/Microsoft/The_Economics_of_The_Microsoft_Case.pdf

Cabral, L. and S. Greenstein (1990). "Switching costs and bidding parity in government procurement of computer systems," *Journal of Law, Economics, and Organization*, 6: 463-469.

Calem, P. and L. Mester (1995). "Consumer behavior and the stickiness of credit card interest rates," *American Economic Review*, 85: 1327-1336.

Chen, P.Y. and L. Hitt (2002). "Measuring switching costs and their determinants in internet enabled businesses: a study of the on-line brokerage industry," *Information Systems Research*, 13: 255-274.

Chen, Y. (1997). "Paying customers to switch," *Journal of Economics and Management Strategy*, 6: 877-897.

Dranove, D. and W.D. White (1996). "Specialization, option demand, and the pricing of medical specialists," *Journal of Economics and Management Strategy*, 5: 277-306.

Elzinga, G. and D. Mills. (1998), "Switching costs in the wholesale distribution of cigarettes," *Southern Economic Journal*, 65: 282-293.

Greenstein, S. (1993). "Did Installed Base Given an Incumbent Any (Measurable) Advantages in Federal Computer Procurement?," *RAND Journal of Economics*, 24 (1): 19-39.

Greenstein, S. (1997). "Lock-in and the Costs of Switching Mainframe Computer Vendors: What Do Buyers See?" *Industrial and Corporate Change*, 6(2): 247-273.

Farrell, J. and P. Klemperer (2006). "Coordination and Lock-In: Competition with Switching Costs and Network Effects," forthcoming in M. Armstrong and R. Porter (eds.), *Handbook of Industrial Organization, Vol. 3*, retrieved at: http://www.paulklemperer.org/

Horowitz, J. and N. Savin (2001). "Binary Response Models: Logits, Probits and Semiparametrics," *Journal of Economic Perspectives*, 15 (4), 43-56.

Klemperer, P. (1987). "Markets with Consumer Switching Costs," *Quarterly Journal of Economics*, 102: 375-94.

Klemperer, P. (1992). "Equilibrium Product Lines: Competing Head-to-Head May be Less Efficient," *American Economic Review*, 82: 740-55.

Klemperer, P. (1995). "Competition when Consumers have Switching Costs," *Review of Economic Studies*, 62: 515-39.

Muthoo, A. (2000). "B2B: A Primer," *Business 2.0 (online)*, 6/17/2000.

Nalebuff, B. (2006). "Bundling as an Entry Deterrent," *Quarterly Journal of Economics*, forthcoming.

Padilla, A.J. (1995). "Revisiting dynamic duopoly with consumer switching costs," *Journal of Economic Theory*, 67: 520-530.

Ricciuti, M. (2004). "Eyes on Enterprise: SAP-Microsoft Talks Underscore Harsh Market Reality," CNET News, July 6, 2004: retrieved from http://news.com.com/Eyes+on+enterprise/2009-1001_3-5250319.html.

Shapiro, C, and H. Varian (1998). *Information Rules: A Strategic Guide to the Network Economy*. Cambridge, MA: Harvard Business School Press.

Taylor, C. (2003). "Supplier surfing: competition and consumer behavior in subscription markets," *RAND Journal of Economics*, 34: 223-246.

Varian, H. (2003). "Economics of Information Technology," mimeo, retrieved from http://people.ischool.berkeley.edu/~hal/Papers/mattioli/mattioli.html

Whinston, M.D. (1990). "Tying, foreclosure and exclusion," *American Economic Review*, 80: 837-859.

Whinston, M.D. (2001). "Exclusivity and tying in US v. Microsoft: what we know, and don't know," *Journal of Economic Perspectives*, 15: 63-80.

Figure 1: Representative enterprise software product taxonomy

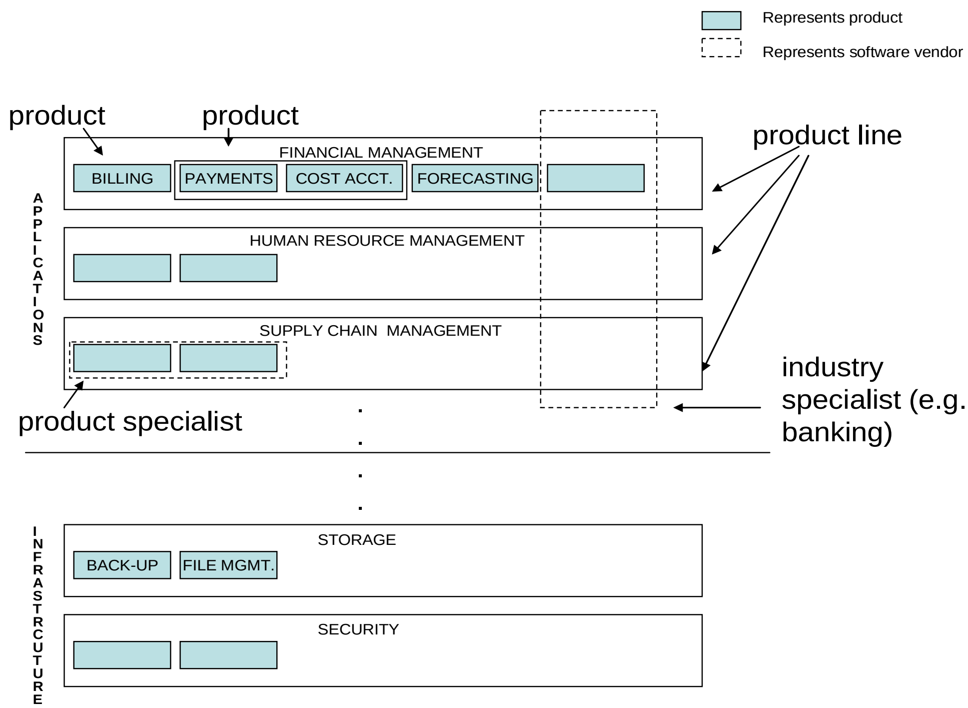
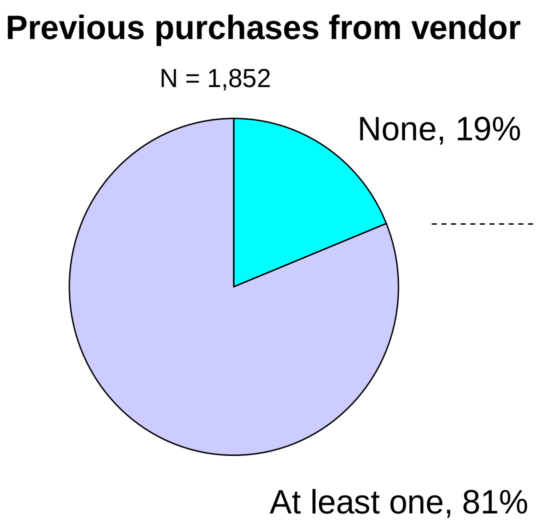


Figure 2: Transaction breakdown by customer's previous purchase history



Of transactions that are first-time purchases:

– 86% of transactions classified as "competitive switch" (16% of total transactions)

– 14% not defined as "competitive switch" (3% of total)

Of transactions that are not first-time:

– 42% include upgrade of existing product* (34% of total)

  – 26% of these received an "upgrade credit" (9% of total)

– 30% are only for different product(s) within same product line (24% of total)

  – 4% are "competitive switch" (1% of total)

– 28% are only from new product line (23% of total)

  – 43% are "competitive switch" (10% of total)

* includes any product purchase where customer owned previous version of some part of product, even if customer purchased additional products at the same time

Table 1: Deal dataset; Summary statistics for key variables, N=1,852 (unless noted)

Variable	Unit	Mean	Std. Dev	Minimum	Maximum
Basic sale information					
Total list price	$1,000	1,190	1,308	50	15,500
Total price paid	$1,000	783	711	50	6,125
Total discount given	%	34.2	11.9	10	90
Multi-product sale	1=yes	0.13	0.33	0	1
Services spend (as % of license spend, not including maintenance)	%	0.24	0.13	0	0.65
Salesperson tenure	# of quarters	13.8	7.8	1	**
Product information					
Product age	# of quarters since Version 1.1	>20**	**	1	**
Market share	%	36.0	14.7	3	**
Gartner MagicQuadrant	1=yes	0.31	0.46	0	1
Customer status					
Cumulative number of upgrades***	#	0.3	0.8	0	**
New to product line	1=yes	0.42	0.45	0	1
New to vendor	1=yes	0.19	0.17	0	1
Competitive switch	1=yes	0.27	0.21	0	1
Upgraded product by 2006	1=yes	0.42	0.49	0	1
Switched by 2006	1=yes	0.18	0.39	0	1
Using legacy product in 2006	1=yes	0.40	0.41	0	1
External customer information					
Annual revenue of customer	$ bn	20.5	19.8	**	**
Customer employee size (2006)	1,000	38.1	17.5	**	**
Five-year cash flow change of customer	%	10.3	6.8	**	**
IT capabilities (5=highest)****	1 to 5 scale	2.8	0.8	1	5
Vendor-specific factors					
% by which vendor exceeded street EPS estimate	%	12.3	**	0	**
Salesperson commission on deal	$1,000	71.4	112.1	1	**
Salesperson commission had the deal closed a quarter earlier	$1,000	64.7	108.3	1	**
Salesperson commission had the deal closed a quarter later	$1,000	61.8	107.8	1	**

Note: ** represents that the data is not reported per agreement with the provider of the dataset (to protect its identity or identity of customers).
*** Equals 0 if this is the first purchase of the product, and "new to product" variable is also 0
**** N = 1,419; data unavailable for customers of the remaining purchases

Table 2: Median upgrade credit and average discount by customer type

Customer type	% of transactions	Median upgrade credit*	Average discount
Existing customer			
Upgraded same product			
- Received upgrade credit	.09	30	21
- Did not receive credit	.25	-	31
Within same product line			
- Received competitive switch credit	.01	40	23
- Did not receive credit	.23	-	37
New to product line			
- Received competitive switch credit	.10	40	28
- Did not receive credit	.13	-	42
New customer			
- Received competitive switch credit	.16	40	30
- Did not receive credit	.03	-	46

* Average number of percentage points subtracted from the "standard" list price to determine list price for customer of that status

Table 3: OLS results using full sample and reported discount

	(A)	(B)	(C)	(D)
Sample	Full sample	IT capabilities sample	IT capabilities sample	Full sample
Number of observations	1,852	1,419	1,419	1,852
Dependent variable	Discount	Discount	Discount	Discount
Constant	7.02 (2.35)***	7.35 (2.65)***	7.56 (2.79)***	8.03 (2.84)***
Customer status variables				
Cumulative number of upgrades	-0.56 (0.60)	-0.31 (0.43)	-0.61 (0.55)	
Existing customer, same line	1.67 (1.32)	1.81 (1.41)	1.41 (0.98)	1.97 (1.55)
Existing customer, new line	7.21 (3.98)*	6.78 (3.60)*	6.45 (3.49)*	6.50 (3.56)*
New customer	9.97 (4.75)**	9.76 (4.70)**	10.08 (4.81)**	9.54 (4.56)**
List price credit interactions				
Existing product*upgrade credit	-5.45 (2.61)**	-4.58 (2.40)*	-4.32 (2.35)*	-5.01 (2.78)*
Same line*CS credit	-4.92 (3.45)	-3.45 (3.21)	-4.39 (3.50)	-4.78 (3.33)
New line*CS credit	-4.61 (2.99)*	-4.87 (3.09)*	-4.97 (3.14)*	-4.50 (2.66)*
New Customer*CS credit	-6.04 (2.62)**	-6.44 (2.70)**	-6.21 (2.67)**	-5.67 (2.51)**
Sale and product variables				
Multi-product line sale	-0.31 (0.40)	-0.29 (0.38)	-0.67 (0.51)	
Log contract size	1.93 (0.71)***	1.87 (0.69)***	1.89 (0.70)***	1.80 (0.67)***
Log product age	-2.31 (1.40)*	-2.56 (1.31)*	-2.67 (1.40)*	-3.41 (1.69)**
Log market share (%)	0.67 (0.50)	1.31 (0.87)	1.09 (0.79)	
Gartner MagicQuadrant	-0.31 (0.48)	-0.09 (0.12)	-0.18 (0.20)	
Customer variables				
Log customer employees	0.30 (0.16)*		0.41 (0.20)**	0.23 (0.14)*
5-year cash flow change (%)	1.67 (0.56)***		2.09 (0.67)***	1.88 (0.60)***
IT capabilities		3.41 (1.09)***	-1.67 (1.31)	
Controls for salesperson compensation concerns	**Y**	**Y**	**Y**	**Y**
Quarter fixed effects	**Y**	**Y**	**Y**	**Y**
Customer industry fixed effects	**Y**	**Y**	**Y**	**Y**
Region fixed effects	**Y**	**Y**	**Y**	**Y**
Operating system fixed effects	**Y**	**Y**	**Y**	**Y**
R-squared	**0.291**	**0.267**	**0.279**	**0.255**

SEs in parentheses; *, **, * represent statistical significance at the 1%, 5% and 10% levels, respectively**

Table 4: OLS results discarding deals with list price credits

	(A)	(B)	(C)	(D)
Sample	Full sample	IT capabilities sample	IT capabilities sample	Full sample
Number of observations	1,501	1,178	1,178	1,501
Dependent variable	Discount	Discount	Discount	Discount
Constant	6.51 (3.01)**	7.32 (3.39)**	7.44 (3.44)**	7.78 (3.67)**
Customer status variables				
Cumulative number of upgrades	-0.25 (0.30)	-0.54 (0.41)	-0.80 (0.59)	
Existing customer, same line	0.90 (0.67)	1.22 (0.96)	1.05 (0.87)	0.81 (0.61)
Existing customer, new line	5.66 (3.01)*	5.09 (3.53)	5.25 (3.22)	5.39 (3.17)*
New customer	11.14 (5.41)**	9.56 (5.09)*	8.98 (4.94)*	10.45 (5.10)**
Sale and product variables				
Multi-product sale	-1.45 (1.29)	-1.09 (1.00)	-1.21 (1.07)	
Log contract size	0.98 (0.51)*	0.94 (0.50)*	0.88 (0.49)*	0.78 (0.40)*
Log product age	-4.70 (2.26)**	-4.55 (2.25)**	-4.60 (2.30)	-5.12 (2.30)**
Log market share (%)	0.23 (0.31)	0.41 (0.42)	0.33 (0.36)	
Gartner MagicQuadrant	-0.69 (0.56)	-0.77 (0.59)	-0.45 (0.50)	
% services spend	0.56 (0.40)	0.67 (0.45)	0.76 (0.52)	
Log salesperson tenure	1.41 (0.77)*	1.31 (0.72)*	1.33 (0.72)*	1.20 (0.71)*
Customer variables				
Log customer employees	0.18 (0.12)		0.33 (0.22)	0.25 (0.18)
Log customer revenue	-0.09 (0.10)		-0.04 (0.08)	
5-year cash flow change (%)	1.01 (0.45)**		1.45 (0.79)*	1.67 (0.78)**
IT capabilities		2.76 (1.32)**	-0.78 (0.50)	
Controls for salesperson compensation concerns	**Y**	**Y**	**Y**	**Y**
Quarter fixed effects	**Y**	**Y**	**Y**	**Y**
Customer industry fixed effects	**Y**	**Y**	**Y**	**Y**
Region fixed effects	**Y**	**Y**	**Y**	**Y**
Operating system fixed effects	**Y**	**Y**	**Y**	**Y**
R-squared	**0.260**	**0.233**	**0.279**	**0.229**

SEs in parentheses; *, **, * represent statistical significance at the 1%, 5% and 10% levels, respectively**

Table 5: OLS results using adjusted discount

	(A)	(B)	(C)	(D)
Sample	Full sample	IT capabilities sample	IT capabilities sample	Full sample
Number of observations	1,852	1,419	1,419	1,852
Dependent variable	Adjusted discount	Adjusted discount	Adjusted discount	Adjusted discount
Constant	6.51 (3.01)**	7.32 (3.39)**	7.44 (3.44)**	7.78 (3.67)**
Customer status variables				
Cumulative number of upgrades	-0.25 (0.30)	-0.54 (0.41)	-0.80 (0.59)	
Existing customer, same line	-0.66 (0.40)	-0.98 (0.67)	-0.76 (0.57)	-0.41 (0.33)
Existing customer, new line	5.98 (3.01)*	5.55 (2.95)*	5.49 (3.01)*	5.77 (3.03)*
New customer	14.26 (6.78)**	15.67 (8.51)*	14.91 (8.13)*	13.67 (6.50)**
Sale and product variables				
Multi-product sale	-1.45 (1.29)	-1.09 (1.00)	-1.21 (1.07)	
Log contract size	0.98 (0.51)*	0.94 (0.50)*	0.88 (0.49)*	0.78 (0.40)*
Log product age	-4.70 (2.26)**	-4.55 (2.25)**	-4.60 (2.30)	-5.12 (2.30)**
Log market share (%)	0.23 (0.31)	0.41 (0.42)	0.33 (0.36)	
Gartner MagicQuadrant	-0.69 (0.56)	-0.77 (0.59)	-0.45 (0.50)	
% services spend	0.56 (0.40)	0.67 (0.45)	0.76 (0.52)	0.34 (0.26)
Log salesperson tenure	1.41 (0.77)*	1.31 (0.72)*	1.33 (0.72)*	1.20 (0.71)*
Customer variables				
Log customer employees	0.18 (0.12)		0.33 (0.22)	0.25 (0.18)
Log customer revenue	-0.09 (0.10)		-0.04 (0.08)	
5-year cash flow change (%)	1.01 (0.45)**		1.45 (0.79)*	1.67 (0.78)**
IT capabilities		2.76 (1.32)**	-0.78 (0.50)	
Controls for salesperson compensation concerns	**Y**	**Y**	**Y**	**Y**
Quarter fixed effects	**Y**	**Y**	**Y**	**Y**
Customer industry fixed effects	**Y**	**Y**	**Y**	**Y**
Region fixed effects	**Y**	**Y**	**Y**	**Y**
Operating system fixed effects	**Y**	**Y**	**Y**	**Y**
R-squared	**0.260**	**0.233**	**0.279**	**0.229**

SEs in parentheses; *, **, * represent statistical significance at the 1%, 5% and 10% levels, respectively**

Table 6: OLS results restricted to new products

	(A)	(B)
Sample	Full sample	Full sample
Number of observations	456	456
Dependent variable	Adjusted discount	Adjusted discount
Constant	8.39 (4.01)**	8.02 (3.86)**
Customer status variables		
Version 2.0 of product (excluded is version 1.0)	-0.03 (0.03)	-0.05 (0.03)
Existing customer, same line	-1.42 (1.07)	-0.97 (0.89)
Existing customer, new line	8.75 (5.03)*	7.45 (4.75)*
New customer	13.29 (7.95)*	14.04 (8.23)*
Sale and product variables		
Multi-product sale	-0.65 (0.41)	
Log contract size	1.45 (0.67)**	1.61 (0.78)**
Log product age	-2.06 (1.56)	-1.78 (1.58)
Log market share (%)	0.01 (0.01)	
Gartner MagicQuadrant	-6.71 (4.00)*	
% services spend	0.41 (0.33)	
Log salesperson tenure	3.07 (2.01)	3.61 (2.42)*
Customer variables		
Log customer employees	0.37 (0.22)	0.31 (0.22)
Log customer revenue	-0.10 (0.12)	
5-year cash flow change (%)	3.41 (1.75)*	3.67 (1.98)*
IT capabilities		
Controls for salesperson compensation concerns	**Y**	**Y**
Quarter fixed effects	**Y**	**Y**
Customer industry fixed effects	**Y**	**Y**
Region fixed effects	**Y**	**Y**
Operating system fixed effects	**Y**	**Y**
R-squared	**0.239**	**0.198**

SEs in parentheses; *, **, * represent statistical significance at the 1%, 5% and 10% levels, respectively**

Table 7: Deal timing model, results after multinomial logit estimating probability of switching and upgrading

Dependent variable = 2006 product use status; N=1,419; robust standard errors clustered by product in parentheses
(note, legacy is the base outcome).

	(A) Pr (switch)	(B) Pr (upgrade)
IT capabilities	-0.03 (0.02)*	-0.02 (0.01)*
Cumulative number of upgrades	-0.09 (0.08)	-0.07 (0.04)*
Existing customer, same line	-0.02 (0.02)	-0.05 (0.04)
Existing customer, new line	0.09 (0.05)*	-0.06 (0.04)*
New customer	0.12 (0.06)**	0.03 (0.04)
Multi-product sale	-0.04 (0.03)	-0.02 (0.02)
Product age	-0.07 (0.04)*	0.04 (0.02)*
Log deal size	-0.10 (0.04)**	0.06 (0.04)
Customer 5 year cash flow change	0.15 (0.07)**	-0.08 (0.05)*
Controls not reported	Product, operating system, industry, sales region, salesperson tenure, length of customer time with salesperson, quarter of original transaction	

*****, **, * represent statistical significance at the 1%, 5% and 10% levels, respectively**

Table 8: Deal timing model, marginal effects after multinomial logit
Dependent variable = 2006 product use status; N=1,419; robust standard errors clustered by product in parentheses
Columns (A) and (B) report the difference in marginal effects after multinomial logit, compared to Pr (legacy); standard error of comparison in parentheses

	(A) Pr (switch) – Pr (legacy)	(B) Pr (upgrade) – Pr (legacy)	(C) X (average variable value)
IT capabilities	-0.087 (0.048)*	0.122 (0.070)*	2.8
Cumulative number of upgrades	-0.090 (0.010)	-0.021 (0.014)*	0.3
Existing customer, same line	-0.067 (0.050)	-0.088 (0.060)	0.24
Existing customer, new line	0.098 (0.057)*	-0.045 (0.032)	0.23
New customer	0.238 (0.111)**	0.15 (0.077)*	0.19
Multi-product sale	-0.030 (0.027)	-0.027 (0.021)	0.13
Product age	-0.017 (0.009)*	0.009 (0.005)*	>20 (hidden for confidentiality)
Log deal size	0.322 (0.141)**	0.217 (0.155)*	6.56
Customer 5 year cash flow change (%)	0.067 (0.039)*	-0.034 (0.020)*	10.3
Controls not reported	Product, operating system, industry, sales region, salesperson tenure, length of customer time with salesperson, quarter of original transaction		

*****, **, * represent statistical significance at the 1%, 5% and 10% levels, respectively**