A Proposal for an Econometric Analysis of Switching Costs in the Software Indsutry

Steve Kahl

15.575

Spring 2004

I. Research Problem

The total cost of ownership to implement a software package typically is much higher than the purchase price of software, sometimes as high as twelve times the price. Integrating software with the pre-existing technical architecture, changing business processes, training new users on the system, and maintenance costs for the software make it difficult for customers to replace software products. These high switching costs can lock customers in and increase the value of the installed base. As a result, the size of the installed base is an important performance measure for the software industry. The installed base, in turn, impacts competitive strategy. A firm such as Microsoft with a significant installed base may deter entry from new firms. Or, the value of an installed base may influence product strategy, pricing, and technology adoption.

Despite the importance of switching costs in the software industry, little is understood empirically about its impact on firm strategy, in particular what price. Most work done on switching cost is theoretical in nature, not tied to any particular industry analysis, and is highly generalized. Little empirical work has been done to delineate the different kinds of switching costs and their impact on software pricing. We propose an empirical study of the affect of switching costs on pricing in the supply chain management software market using a hedonic pricing model.

This project has three research goals in mind: 1) Provide a more detailed taxonomy of switching costs in the software industry; 2) Estimate the magnitude of switching costs on the price users will pay; 3) deduce implications for competitive strategies.

The organization of this research proposal is as follows. Section II discusses how this research project will help contribute to the existing literature on switching costs. Section III discusses important aspects of the data required for analysis. Section IV identifies the main hypotheses we intend to test and explains the methodology.

II: Literature Review

There exists a large and mostly theoretical literature on switching costs. We believe the main contribution of this study will be an empirical investigation to test some of these theoretical models and to estimate the effect of each component of switching costs on the price users are willing to pay.

An important contribution of the switching cost literature is an elaboration of the different kinds of switching costs (Klemperer 1987; Nilssen 1992; Klemperer 1995; Farrell and Klemperer Forthcomming). Typically, switching costs have been divided into three categories:

- **Informational:** The training costs associated with learning the new product
- **Transactional:** Costs associated with completing the transaction of switching from one vendor to another. For example, switching equipment suppliers requires returning the old equipment and picking up the new

- **Contractual:** Firm's actions that create switching costs, e.g., frequent flier programs or specific binding elements of a contract.

Although a good start, we believe that this list is not exhaustive and that switching costs vary across industries. The switching costs associated with an airplane differ from those in consumer goods. The switching costs associated with an airplane are more transactional in nature – switching engine suppliers requires significant modification to other parts of the airplane, but the switching costs in consumer products are more contractual and brand related. We believe that a more detailed analysis of switching costs in a particular industry will extend this initial list in a constructive manner. Thus, we intend to provide a detailed empirical analysis of the software industry to determine its specific switching costs. For example, implementing software typically involves changing business processes, so these organizational changes would be another category of switching costs.

Despite initially disaggregating switching costs, in the theoretical models most economists broadly treat switching costs as a single variable (Klemperer 1987a; Klemperer 1987b; Klemperer 1987c; Farrell and Shapiro 1988; Beggs and Klemperer 1992). The disadvantage of this approach is that we would like to know the effect of different kinds of switching costs. This is important for competitive reasons. For example, new entrants in software typically target technology switching costs, arguing that their solution is cheaper to implement. Even though this may be true, technology

switching costs may not be as important as the organizational change switching costs. Entrants may be focusing on the wrong kind of switching costs.

Moving in this direction, Nilssen developed a model which showed that an increase in transaction costs has more of an impact relative learning costs (Nilssen 1992). However, she only considered two kinds of switching costs and held total switching costs constant. We plan to extend this approach using an empirical model to test the effects of all the different kinds of switching costs.

Another important contribution of the literature is the development of theoretical models to explain firm behavior in markets with switching costs. The general conclusion of these models are twofold 1) markets with switching costs have higher prices; 2) the prices that firms may extract change from period to period, so this price advantage is a long term view (Klemperer 1995; Farrell and Klemperer Forthcomming). When competing for customers in the initial period, a firm may lower prices in order to attract customers, but once the customer is locked-in the firm can charge higher prices.

Despite these general conclusions, these models are inconclusive about how a firm should behave when responding to potential entrants. A firm may want to continue to increase market share but sacrifice price by competing with entrants for new customers, or a firm may chose to preserve price by focusing on repeat customers but sacrifice market share (Klemperer 1989; Klemperer 1995). We plan to design our study so that we can be more explicit about firm behavior during periods of firm entry.

A final shortcoming of these models is that they generally assume that the customer is aware of these switching costs and that it impacts their willingness to pay. Fortunately, Greenstein has justified this assumption by documenting how customers invest based upon analysis of technological switching costs (Greenstein 1997). In our study, the customer's understanding of the switching costs are endogenous to our model as we derive our switching costs from a combination of vendor information and the customer's estimated budget. We also extend Greenstein's analysis to include intangible variables such as organizational change beyond the technical compatibility issues.

III. Data and Data sources – supply chain management software market

In order to more explicitly analyze the effects of switching costs on pricing we propose to study a software market in which we expect significant variation in switching costs – the supply chain management software market. These applications help organizations plan their production processes within a plant, between plants, and coordinate supply and logistic activities with trading partners. We plan to source this data at the project level – capturing budget estimates of customers as well as contractual information from the vendor, SAP.

Supply chain management software is a good market to study because typically these products have large switching costs. These applications need to be integrated with existing hardware configurations, other business applications, and often new business

processes must be designed and implemented in order to take advantage of the technical advances of these products. As firms begin to coordinate supply chain activities with trading partners, logistics service providers, customers, and other service providers, these cost have increased to cover the inter-firm integration requirements. Although the firms may share a supply chain software package, their internal operating environments may differ increasing the complexity of the technical integration on top of the business process changes. Lastly, there are significant training costs as well as many of these packages introduce new planning and scheduling algorithms. For example, supply chain planning packages in the mid 1990's enabled silmultaneous constraint based planning – a technique planners could not previously consider with existing software packages. Therefore, the users not only need to be trained on the new software package but the new techniques as well. These costs combine to create large switching costs.

The business models of these software vendors express the magnitude of the switching costs. Over the lifetime of the project, product sales may only represent 30% of the total revenue the firm receives from the client (Cusumano 2004). The other 70% comes from service and maintenance revenues. And, this is only part of the organizational change cost that the customer pays because outside of the software firm the customer also typically pays for third party consultants such as Accenture or IBM or must pull employees from current projects to implement the change.

Studying the supply chain software market has several other advantages. The origins of the market trace back to the beginning of package business applications, providing a rich

data set. The first applications called MRP and MRPII (MRP stands for material requirements planning) date back to the 70's and 80's. These applications enabled companies to plan the material requirements to build a final product as well as schedule capacity and when the parts were required. The introduction of the enterprise resource planning (ERP) applications integrated these material requirements modules with procurement, financial, and human resource modules to help manage the procurement of component parts to meet production plans. In the early 90's, software firms' took advantage of lowering cost of computing and advancements in operations research to develop more powerful production planning applications that could develop plans quicker and more accurately using constraint-based techniques. In addition, at the planning level these packages enable broader functional integration with logistics and demand planning as well as integrating with trading partners in the supply chain.

Thus, there are three distinct innovation periods in the supply chain market – the introduction of automated material requirements planning, integrating planning with other applications through ERP software, and the introduction of advanced supply chain planning. These three periods provide enable us to study Klemperer's claims about incumbent firm behavior in a market with high switching costs and the threat of firm entry.

Another advantage of this market is that it is a fairly closed market in the sense that three players, i2, Manugistics, and SAP, have emerged as the market leaders. [GET GARTNER DATA TO SUPPORT THIS CLAIM]. SAP, the largest ERP vendor today,

entered the supply chain management market with an integrated MRP and MRPII solutions as part of its R/2 (mainframe solution) and R/3 (client/server solution) ERP package. Both i2 and Manugistics both entered the market during the third phase, introducing advanced supply chain management technology. Therefore, in period three SAP would be recognized as the incumbent and i2 and Manugistics as the entrants. Other players in the market include Oracle and PeopleSoft but neither has gained significant market share. Unlike the spreadsheet market, though, internally developed custom solutions are a significant factor in the market. Rather than implementing a package solution, some firms opt to build their own using in-house programmers and consultants. However, with the cost of these solutions dropping this has become less prevalent.

With multiple options available, firms have implemented a wide range of solutions. Some firms just implemented supply chain solutions for just one product line or a particular plant; whereas, others implement across the entire company. Some have mixed and matched solutions, for example, using SAP for core production scheduling and i2 for advanced planning. The result is that we would expect wide variation in our independent variables that measure switching costs.

The main disadvantage with this market is the difficulty of capturing data on product features, switching costs, and prices. Getting actual prices paid and list price is complicated by the fact that supply chain software is sold through an iterative direct sales process or is bundled with other products. Determining the switching costs is

complicated by the fact that the costs come from multiple parties – software firm, consultants, hardware firms, communication firms, and the customer itself.

To help resolve these issues, we propose to gather specific project level data for supply chain management implementations for SAP. Following a similar data methodology as Ichniowski in his study on the impact of human resource practices on productivity (Ichniowski, Shaw et al. 1996), I plan to empirically derive the switching costs by collecting project level data. This requires budget information from the customer as well as selling information from SAP. In most cases, purchase of this software required budgetary approval meaning that estimated cost breakdowns were required for all components of the project, including hardware, software, consultants, and organizational change requirements. This budgetary information becomes our estimates for the different kinds of switching costs. It is important that we capture the estimated costs as opposed to the actual costs because the estimated costs reflect what we want to capture: the customer's conception of the switching costs when making a purchase.

However, the customer data is incomplete for our purposes. For competitive reasons, SAP may bundle supply chain applications with other products or offer additional seats for a highly discounted price. These discounts are a form of what Klemperer calls artificial or contractual switching costs (Klemperer 1987a) in that a customer receives a discount but it is only redeemable with SAP. It is important to capture some contractual and product information from SAP as well.

Therefore, the source of our data will include contractual information from SAP as well as detailed budget estimations from customers who purchases SAP's supply chain management applications. We plan to collect data from 1991-2003 in the U.S. market. We begin with 1991 because that coincides with SAP's introduction of R/3 its client/server version of its core ERP software and its entry into the U.S. market in a concentrated fashion. By collecting this data from 1991 forward, we capture two time periods in the market evolution of supply chain software – the introduction of ERP and the introduction of advanced supply chain management applications. Consequently, we will observe SAP both as an entrant into the market and as an incumbent which will provide further insight into the competitive dynamics created by switching costs.

The primary disadvantage with this approach is that is limits the external validity of our results. There could be something about SAP which makes them unique to the marketplace with respect to switching costs. However, a primary goal of the study is to develop a detailed empirical understanding of switching costs and its impact on pricing and competitive behavior. Future studies can expand industries and companies to verify these results.

Related to this issue is a concern with selection bias. We are not capturing data from two types of firms: those that decided not to purchase after thorough review of the costs, and those that decided to purchase from a competitor such as i2 or Manugistics. In order for there to be selection bias, the cases that we did not include must affect both the dependent variable and the independent variables. Even though the missing observations do seem

correlated with dependent variable, it is not clear that they affect the independent variables. These firms behave in a similar fashion, calculating switching costs as part of their purchase decision. The primary driver for not including i2 and Manugistics is a practical one. The supply chain management software sector is highly competitive so trying to collect data from multiple competitors probably would preclude any from participating. Besides, the main motivation for including these companies would be to get the entrant perspective which we can capture with the beginning of SAP's R/3 release.

IV. Hypotheses and Model

4.1 Hypotheses

Thus far, we have identified the following hypotheses:

1. We expect that the price customers will pay for supply chain management products will be positively correlated with switching costs and that it will be of similar magnitude as other product attributes. The higher the switching costs the higher the price the customer is willing to pay. This hypothesis is consistent with Klemperer's claim that markets with switching costs lead to higher prices (Beggs and Klemperer 1992).

1a. We will disaggregate switching costs into its various components and we believe that the magnitude of the effect will vary. Thus far, we have identified four categories of switching costs – technical integration, organizational change, training and learning, and contractual. We expect the organizational change switching costs

to have the largest magnitude because these changes require the most significant reorganization across the broadest set of assets.

2. We expect whether the customer is a repeat buyer to mediate the magnitude of this relationship. Klemperer argues that a firm has incentives to exploit repeat customers because they are locked in (Klemperer 1995). Therefore, we would expect two outcomes: 1) a repeat buyer would pay higher price and 2) When interacting switching costs with repeat buyer, we would expect the magnitude to be higher than when the customer is a new to SAP.
3. We also expect whether SAP is an entrant or incumbent to mediate the magnitude of the relationship between switching costs and product price. Because of the benefits of customer lock-in, researchers believe that firms may charge lower prices in order to persuade more customers to buy-in. Therefore, we would expect that when SAP is an entrant it will charge lower prices.

4.2 Hedonic Regression Model

We believe that a hedonic model is the best method to test these hypotheses. This section discusses the advantages of this model as well as our preliminary base model.

Hedonic models assume that a commodity can be considered as consisting of several bundles of smaller characteristics (Berndt 1991). Therefore, we plan to use a hedonic regression model to estimate the magnitude of the effects of switching costs on software prices.

Technology products are particularly strong candidates for hedonic regression models because of their component and modular architecture. Several other studies have analyzed hardware products such as mainframe computer prices (Chow 1967)or microcomputers (Berndt and Griliches 1990). Most relevant to this study is Brynjolffson's and Kemerer's (1996) use of the hedonic model to study the effects of standards and network externalities of pricing in the spreadsheet software package market. The key independent variables in their analysis measured degree of conforming to an industry standard as well a separate measures for installed base and market share. After controlling for time and other product feature effects, they found that network effects and compatibility with a dominant design are positive and are just as important as any other product features (Brynjolffson and Kemerer 1996)

One of Brynjolffson's and Kemerer's concerns with this approach is that other factors beyond the scope of the hedonic model may be important. One area for concern is that because software has low marginal costs, firms have more discretion in their pricing and may strategically price their products. As discussed, Klemperer argues that high switching cost factors into how a firm may strategically price their products in response to potential entry and in order to gain market share. Therefore, we plan to address the strategic pricing issue by including specific measures for switching costs.

Incorporating measures for switching costs extends the hedonic regression model in other important ways. Traditionally, the considered bundled characteristics that determine price are tangible features of the products – essentially their component parts.

Brynjolffson and Kemerer have shown that other intangible features such as conformity to a standard or size of installed base also impact price. Including switching costs suggests another set of intangibles, in particular, organizational changes, training costs, and contract obligations.

Thus, our preliminary general model is:

Log_price = β_0+ β_i* Switching Cost attribute i + β_1*market share + β_j*Product Feature attribute j + β_2*Time + β_3*Repeat Customer + β_4*Entrant + β_5*Project size + ε.

Dependent Variable

We will capture the actual price paid for each product. We expect this information to come from SAP and can validate it with the customer's budget information. We will deflate product prices to 1991 dollars using a GDP deflator and then take the natural log of this number. This transformation will better distribute the data. The main question with the dependent variable is price of what? This is important because reliable price information is critical to this study. Making matters difficult, supply chain management is a product line for SAP, consisting of many different modules that evolved throughout the 12 year period we plan to study. But, capturing data at this level is impractical because customers typically bought several modules at a time. We propose to look through the purchase data to determine logical categories, e.g. production planning or logistics execution, to capture price information.

Independent Variables

One of the main goals of this study is to more precisely define switching costs for the software industry. In addition, we want to determine if any type of switching costs has a greater effect on price. To accomplish this goal we plan to empirically build the different types of switching costs by analyzing project data. From past experience working on these projects, we have identified four potential categories. All cost data will be deflated to 1991 dollars and then transformed with the natural log.

Technical Integration

SAP only sells the business application that automates supply chain practices. These applications must integrate with other infrastructure software such as database management, application servers, operating systems, transaction management, as well as different kinds of hardware and communication protocols. We can typically isolate these costs in the project budgets because they require more technical personnel or additional hardware and communications purchases.

Organizational Integration

In addition to technical integration, these products also generate substantial amount of organizational change, in particular work processes. For example, leveraging faster and more optimal supply chain planning tools potentially could change how often a manager plans which in turn affects production scheduling, line operations, and interaction with suppliers. It is important to recognize that the firm does not necessarily have to implement SAP's recommended version of supply chain planning in order for switching

costs to exist. Rather, as in Milgrom and Robert's terminology, implementing the application in general has complementarities which requires changes in other areas of the organization (Milgrom and Roberts 1990). The extent to which a firm changes its organizational work flow creates switching costs.

An advantage of capturing data at the project level is projects must estimate these organizational changes independently of the technical integration costs. Typically, independent consultants provide the role of identifying and implementing the changes.

Training and learning

Two distinct cost areas capture total training costs. First, customers must learn the new software application. Specific training initiatives and part of the maintenance package capture this training. The other training costs include teaching the users the new business processes. These costs typically are captured under "change management" activities.

Contractual Costs

Contractual agreements may create a form of switching costs as well. Typically, customers sign up for a maintenance contract which costs 15-20% of the total license over a three year period. These maintenance contracts enable the customer to get free upgrades on purchased products as well as a certain level of customer service. These maintenance contracts in turn create switching costs.

Control Variables

To estimate the magnitude of the effects of switching costs it is important to control for other factors. Thus far, we have identified six control variables: market share, product attributes, time, repeat customer, entrant, and project size. One notable exception from the Brynjolffson and Kemener model is that we do not include a measure for standardization. Unlike the spreadsheet market, we do not anticipate any standards in the marketplace.

Market Share

Similar to Brynjolffson and Kemerer, market share measures the network externality effects of the supply chain product. We would expect this to be positive indicating that when SAP has larger market share they will be able to extract higher prices.

It is calculated by dividing the installed base by total market share. Consistent with past studies, installed base of each product is calculated by summing its sales in all prior years. Since custom development is an important player in the market, total market share cannot be calculated by summing the sales of all other competitors. We will rely upon IDC estimates for total market size.

Total market share may be too broad of a measure. The network effects of purchasing supply chain management do not appear to cross industries. Because of significant industry differences, a retail firm purchasing supply chain management applications does not necessarily increase the value of the product for a chemicals company. Training and

complementary products differ significantly. A more accurate measure of network externalities would be at the industry level; however, market size information is not captured at this level.

Product Attributes

Product attributes and functionality clearly play an important role in customer purchase decision. Consequently, these attributes impact price. At this time, we do not know what the salient product features are but through analyzing SAP and product data we expect to be able to determine the key features. Third party assessments such as AMR, Garnter, and trade magazines may also help indicate the important product features.

Time

Temporal effects are of particular concern for this data set because the time period includes the Year 2000 effect, the Internet bubble, and economic recession. To help account for these factors, we will create a dummy variable for each year. Also, consistent with Brynjolffson and Kemerer's model, we will construct a time trend variable in order to determine any quality-adjusted prices due to temporal factors of technological progress.

Repeat Customer

Repeat customer is a dummy variable coded as 1 if the customer had previously bought any product from SAP. This variable is important because switching cost theory makes the distinction between the prices a firm can charged selling back into the installed base

as opposed to a new customer. As explained in hypothesis 2, we would expect this coefficient to be positive. In addition, when we interact repeat customer with the various switching costs, we expect $\beta_3 + \beta_i$ to be greater in magnitude than β_i in the base model, indicating the effect of switching costs have a greater effect on repeat customers.

Entrant

Entrant is a dummy variable, coded 1 for the years 1991-1994. We define SAP as an entrant into the market during the initial roll out of R/3. 1995 marks the year in which i2 and Manugistics emerged as viable competitors in the emerging advanced supply chain management market; therefore, 1995 and beyond SAP should be considered as the incumbent. As explained in hypothesis 3, we would expect this coefficient to be negative because entrants are rapidly trying to build market share with the expectation that they can extract higher rents after lock in.

Project Size

We control for project size because different sized projects will have different price expectations. The product price would be significantly larger in a large scale project in which a customer implements a full ERP solution with supply chain management than a smaller project in which the customer implement supply chain management at one location.

To calculate product size, we propose to divide total project costs adjusted to 1991 dollars, weighted by total number of expected users. The total project costs and the total

number of users can be calculated from the budgetary information provided from the customers. We weight by total number of expected users because firms get different rates for consulting and other complementary IT products such as hardware depending upon the size of the organization.

V. Conclusion

From this study, we hope to not only better understand the influence of switching costs on software prices, but also how this impacts competition and firm strategy. Noted switching cost theorists have argued that switching costs can impact adoption of standards and the size and breadth of the product portfolio (Klemperer 1995). This is an important question in software particularly as firms debate product strategies - the best-of-breed (or narrow product line) versus the suite of products (wide product line) as well as adopting industry standards versus a proprietary solution. We believe that we can use the knowledge gained about the different component and effects of the price software firms can charge to help provide some answers to these important strategic debates.

Bibliography

Beggs, A. and P. Klemperer (1992). "Multiperiod Competition with Switching Costs." Econometrica **60**: 651-666.

Berndt, E. (1991). The Practices of Econometrics, Classic and Contemporary. Reading, MA, John Wiley and Sons.

Berndt, E. and Z. Griliches (1990). Price Indexes for Microcomputers: An exploratory study. Conference on Research in Income and Wealth Workshop on Price Measurement and Their Uses, Washington D.C.

Brynjolffson, E. and C. F. Kemerer (1996). "Network Externalities in Microcomputer Software: An Economic Analysis of the Spreadsheet Market." Management Science **42**(12): 1627-1647.

Chow, G. C. (1967). "Technological Change and the Demand for Computers." American Economic Review **57**(7): 1117-1130.

Cusumano, M. (2004). The Software Business: What Every Manager, Programmer, and Entrepreneur Must Know to Thrive and Survive in Good Times and Bad. New York, Free Press.

Farrell, J. and P. Klemperer (Forthcomming). "Coordination and Lock-In: Competition with Switching Costs and Network Effects."

Farrell, J. and C. Shapiro (1988). "Dynamic Competition with Switching Costs." RAND Journal of Economics **19**: 123-137.

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

Ichniowski, C., K. Shaw, et al. (1996). "The Effects of Human Resource Management Practices on Productivity." American Economic Review **87**(3): 291-313.

Klemperer, P. (1987). "The Competitiveness of Markets with Switching Costs." RAND Journal of Economica **18**: 138-150.

Klemperer, P. (1987a). "Markets with Consumer Switching Costs." The Quarterly Journal of Economics **102**(2): 375-394.

Klemperer, P. (1987b). "The Competitiveness of Markets with Switching Costs." RAND Journal of Economica **18**: 138-150.

Klemperer, P. (1987c). "Entry Detterence in Markets with Consumer Switching Costs." Economic Journal **97**: 99-117.

Klemperer, P. (1989). "Price Wars Caused by Switching Costs." Review of Economic Studies **56**: 405-420.

Klemperer, P. (1995). "Competition when Consumers have Switching Costs." Review of Economic Studies **62**: 515-539.

Milgrom, P. and J. Roberts (1990). "The Economics of Modern Manufacturing: Technology, Strategy, and Organization." American Economic Review **80**(3): 511-528.

Nilssen, T. (1992). "Two Kinds of Consumer Switching Costs." RAND Journal of Economics **23**: 579-589.