MARKETING SCIENCE INSTITUTE

MSI Reports

Drivers of Technological Novelty and Superior Customer-Need Fulfillment in New Product Development (04-117)
Stefan Wuyts and Shantanu Dutta

Consumer Preferences for Mass Customization (04-118)
Benedict G. C. Dellaert and Stefan Stremersch

The Effect of Service Experiences over Time on a Supplier's Retention of Business Customers (04-119)

Ruth N. Bolton, Katherine N. Lemon, and Matthew D. Bramlett

A Spatial-Choice Model for Product Recommendations (04-120)
Sangkil Moon and Gary J. Russell

Firm-Sponsored Satisfaction Surveys: Positivity Effects on Customer Purchase Behavior? (04-121)
Utpal M. Dholakia, Vicki G. Morwitz, and Robert A.Westbrook

Analogies and Imaginary Consumers: A Case Study of New Product Development (04-122)
José Antonio Rosa, Steve Hoeffler, William Qualls, and Jonathan Bohlmann

2004
WORKING PAPER SERIES
ISSUE FOUR
NO. 04-004

MSI Reports

2004

WORKING PAPER SERIES

ISSUE FOUR

NO. 04-004

The Marketing Science Institute supports academic research for the development—and practical translation—of leading-edge marketing knowledge on issues of importance to business performance. Topics are identified by the Board of Trustees, which represents MSI member corporations and the academic community. MSI supports academic studies on these issues and disseminates findings through conferences and workshops, as well as through its publications series.

Marketing Science Institute
1000 Massachusetts Avenue
Cambridge, MA
02138-5396

Phone: 617.491.2060
Fax: 617.491.2065
www.msi.org

Working Paper Series
The articles that appear in *MSI Reports* have not undergone a formal academic review. They are released as part of the MSI Working Paper Series, and are distributed for the benefit of MSI corporate and academic members and the general public.

Subscriptions
Annual subscriptions to *MSI Reports* can be placed online at www.msi.org. Questions regarding subscriptions may be directed to pubs@msi.org.

Single reports
Articles in *MSI Reports* are available in downloadable (PDF) format at www.msi.org.

Past reports
MSI working papers published before 2003 are available as individual hard-copy reports; many are also available in downloadable (PDF) format. To order, go to www.msi.org.

Corporate members
MSI member company personnel receive all MSI reports (PDF and print versions) free of charge.

Academic members
Academics may qualify for free access to PDF (downloadable) versions of MSI reports and for special rates on other MSI print publications. For more information and to apply, go to "Qualify for academic membership" on www.msi.org.

Classroom use
Upon written request, MSI working papers may be copied for one-time classroom use free of charge. Please contact MSI to obtain permission.

Search for publications
See the searchable publications database at www.msi.org.

Submissions
MSI will consider a paper for inclusion in *MSI Reports*, even if the research was not originally supported by MSI, if the paper deals with a priority subject, represents a significant advance over existing literature, and has not been widely disseminated elsewhere. Only submissions from faculty members or doctoral students working with faculty advisors will be considered. "MSI Working Paper Guidelines" and "MSI 2004-2006 Research Priorities" are available in the Research section of www.msi.org.

Publication announcements
To sign up to receive MSI's electronic newsletter, go to www.msi.org.

Change of address
Send old and new address to pubs@msi.org.

The Effect of Service Experiences over Time on a Supplier's Retention of Business Customers

Ruth N. Bolton, Katherine N. Lemon, and Matthew D. Bramlett

A firm's decision to renew service contracts can be greatly affected by a few extremely favorable experiences—especially if they occur closer to the time of contract renewal. To retain customers, suppliers should focus on timing of services delivered, sufficient utilization of the service contract, and occasionally providing exceptional service to customers.

Report Summary

What is the link between a supplier's marketing and service operations and its customers' subsequent repatronage behavior? In this study, Bolton, Lemon, and Bramlett develop a model of service contract renewal for an individual firm purchasing multiple contracts from the same supplier.

They model the firm's decision to renew a service contract as a function of service quality and price, where service quality is measured by the supplier's service operations metrics over time. Based on data for high technology support services in business markets in Germany and the U.K., the study investigates how average service levels, variability in service levels (especially extreme outcomes), and timing of service delivery influence firms' service contract renewal decisions.

Findings show that a firm that has a few extremely favorable experiences for a given service contract is more likely to subsequently renew that service contract. Firms are also influenced by favorable extreme outcomes for other contracts from the same supplier when deciding whether to renew the focal contract—but they weigh this information less heavily. Firms weigh recent experiences more heavily than earlier experiences when deciding whether or not to renew, so the timing of service experiences may be critical to the survival of buyer-seller relationships.

Current practice typically focuses on managing variability *across* contracts and customers (for example, meeting targets across all firms served), rather than variability *within* contracts (for example, targeting specific contracts and firms). However, certain business customers may be systematically underserved because conventional quality control mechanisms (across customers, over time) fail to capture, recognize, or create solutions within customers. These study results suggest that suppliers managing relationships with firms that hold multiple service contracts should carefully manage the amount and timing of resources allocated to *each* contract to deliver value within the firm/supplier relationship. ■

Ruth N. Bolton is Professor of Marketing, W.P. Carey School of Business, Arizona State University. **Katherine N. Lemon** is Associate Professor of Marketing, Wallace E. Carroll School of Management, Boston College. **Matthew D. Bramlett** is Statistician, Division of Vital Statistics, National Center for Health Statistics.

Introduction

Consider this scenario: A firm owns expensive capital equipment, such as computing and information technology; engineering, medical, and manufacturing equipment; or financial, health, or energy management software. Because the equipment is critical to the firm's operations, the firm has purchased contracts for maintenance, support, and repair services. Each service contract is uniquely associated with a piece of equipment, so the firm holds many service contracts. A typical contract has a fixed price and it requires the supplier to provide service on a specific item for a specified time period. Over the duration of the contract, the firm utilizes the services as needed. At the end of each contract, the firm decides whether or not to renew the service contract associated with a specific piece of equipment. The firm makes separate (but not necessarily independent) decisions regarding each contract that it holds.

What factors influence the firm's decision to renew service contracts? The firm's decision will depend on the expected value or benefits to be derived from the supplier of the service contract versus the price of the contract. Unlike an initial purchase decision, a firm's renewal decision is much less likely to depend on contract specifications or marketing communications from suppliers (cf. Ganesh, Arnold, and Reynolds 2000; Kalwani and Narayandas 1995). Instead, the firm's assessment of the value of the renewed contract will likely depend on its prior service experiences under the existing contract or on its prior service on similar contracts for other equipment. The benefits of renewing a service contract may seem highest when the firm has previously experienced an equipment failure that was satisfactorily resolved by the incumbent supplier. We believe that a heavier dedication of the supplier's resources to the firm—implying higher usage of the supplier's services—should be associated with a higher expected value for an existing contract. A firm's experiences with a service supplier over time will alter subsequent buying patterns.

These observations highlight a number of research questions about the role of prior service experiences in a firm's service contract renewal decision. First, what will have more influence on the renewal decision, the supplier's service contract specifications, the firm's prior service experiences, or variability in its experiences over time (Rust et al. 1999)? Second, if a firm has experienced average levels of service quality in the past—including a few incidents of extremely high quality service—will an "extra mile" experience disproportionately influence the renewal decision? Third, if the supplier allocates a substantial amount of resources, such as engineer work minutes, to the firm early in the relationship, will this affect the firm's decision to renew a service contract differently than if the supplier allocates a substantial amount of resources late in the contractual relationship, closer to the renewal decision? Last, will the firm's experience with other service contracts from the same supplier spill over to influence its decision regarding the contract up for renewal?

Answers to these questions are critical in today's markets because finely tuned relationships between firms and their suppliers are necessary for total quality management, process reengineering, just-in-time delivery, and other activities that are coordinated across the entire value chain (Deming 1986; Hammer 2001; Levy 1997). They also have strategic implications for suppliers that are attempting to improve service delivery (Anderson, Fornell, and Rust 1997), as well as for buying firms who want to encourage "lean" suppliers (MacDuffie and Helper 1997). Lastly, they can provide guidance to service organizations that are attempting to exploit their knowledge of business customers to increase customer retention and improve business performance—especially in high technology markets (Heide and Weiss 1995).

We have modeled firms' repatronage behavior for service contracts over time. Our approach is different from prior research in three ways.

- Our dynamic model of service contract renewal for an individual firm—at the contract level—recognizes interdependencies among service contract renewal decisions due to the firm's purchase of multiple contracts from the same supplier. In contrast, prior studies of business-to-business (B2B) relationships typically have estimated static models at the firm level that compare different relationship stages (Cannon and Perreault 1999; Dwyer, Schurr, and Oh 1987; Heide and Weiss 1995; Lusch and Brown 1996; Wathne, Biong, and Heide 2001).

- We have modeled a firm's decision to renew a service contract as a function of service quality and price where service quality is measured by the supplier's *service operations metrics over time*, such as engineer work minutes allocated to the firm during a specific time period. Prior studies of B2B relationships have typically relied on key informants' *perceptions* of the supplier and industry.
- We have incorporated longitudinal data about the supplier's service operations to investigate how average service levels, variability in service levels (especially extreme outcomes), and timing of service delivery influence firms' service contract renewal decisions.

Background

This study deliberately focuses on service contract renewal rather than initial purchase decisions. Because previous research has shown that competitive offerings are less important for firm repatronage decisions than initial purchases (Ganesh, Arnold, and Reynolds 2000; Heide and Weiss 1995), we have focused on the influence of prior service experiences on future purchase behavior. Like the seller of the service, we do not observe the decision-making process within the organization. We simply observe the firm's choice between two primary alternatives, to renew or not renew a contract. Our study doesn't distinguish among the secondary alternatives available to the firm if the service contract is not renewed. The firm may switch to another supplier, rely on in-house service, or discontinue using service in this category due to changing needs. Prior research has studied self-reported switching behavior by firms that have continued to use a service category (Wathne, Biong, and Heide 2001).

Modeling service contract renewal

Base Model. Following a long tradition of theoretical and empirical work concerning firm behavior (Coughlan 1985; Heide and Weiss 1995), we represent the firm's decision to renew a service contract by a binary choice model. The firm (i) renews a service contract (j), by choosing the alternative (k = 1,2) with the highest expected future value (i,j,k^*). That is,

$$P(i,j,k^*) = Prob\ (U_{ijk^*} > U_{ijk}) \quad (1)$$

where:

$$U_{ijk} = V_{ijk} + e_{ijk} \quad (2)$$

and V_{ijk} represents the deterministic component and e_{ijk} represents the stochastic component of the firm's value function. In our formulation, the firm's decision regarding a service contract from a current supplier depends on whether the service contract's value exceeds a threshold or cutoff point that justifies renewal. Firms in different market segments that have different characteristics and needs will value the service contract differently.

Prior research concerning customers' evaluations and purchases of services indicates that service value (V_{ij}) depends upon quality and price (*Price*), where quality has multiple dimensions and price includes monetary and nonmonetary costs (Heide and Weiss 1995; Zeithaml 1988). In this study, we explicitly distinguish between two dimensions of service quality: quality that meets customer needs and quality that results from freedom from deficiencies (Anderson, Fornell, and Rust 1997; Juran and Godfrey 1998). The first dimension, *design quality* ($DesignQ_{it}$), focuses on the elements of the product or service that the customer expects to receive based upon the services or benefits promised by the supplier or stated in the service contract. The second dimension, *experience quality* $(ExperienceQ_{it-1})$ focuses on the customer's prior experience with each of the elements of the product or service.

Design Quality. In studies of B2B relationships, relational norms have been shown to influence exchange relationships between firms (Heide and John 1992; Stinchcombe 1986). The firm's norms about design quality and price—as stipulated in the service contract—will primarily influence a firm's initial purchase of a service contract. It may also influence repeat purchases. Thus, the base model can account for potential heterogeneity across firms with respect to their need for design quality, for example, depending on how critical the piece of equipment being covered by the service contract is perceived.

Experience Quality. Organizational norms about the quality of service also evolve through ongoing interactions and are products of past interactions (Coleman 1990). The firm has opportunities to assess service quality during its interactions with suppliers, which span the spectrum between the external environment and the firm. Based on these interactions, the firm makes more effective purchase decisions (Sinkula 1991). Specifically, prior experiences of a high level of service quality (more visits, more effective support, etc.) will be associated with a higher likelihood of contract renewal.

Completing the base model, these notions can be expressed algebraically as follows:

$$V_{ij} = v\,(DesignQ_{it},\ ExperienceQ_{it-1},\ Price_{it}) \qquad (3)$$

where higher service value is positively associated with higher quality and lower prices.

Variability in service quality over time

Reliability or consistency in the deployment of resources over time plays a role in the success of B2B relationships (Dwyer, Schurr, and Oh 1987; Parasuraman, Zeithaml, and Berry 1985; Tsikritsis and Heineke 2004). Specifically, increases in variability in service quality over time should *decrease* the value of a service contract and, consequently, the likelihood of the firm renewing a contract with its supplier. Recently, this phenomenon has been partially integrated into a mathematical model of consumer—not organizational—purchase behavior.[1] Rust et al. (1999) developed a Bayesian model that predicts that reducing a consumer's uncertainty regarding perceived product quality increases the likelihood that he or she subsequently chooses a consumer product. However, they were unable to eliminate a natural confound between changes in the mean and variance of perceived service quality over time. In their second experiment, they attempted to resolve this issue by studying cross-sectional variation rather than variation over time. In this study, we examine variability in service levels over time, after controlling for average service levels.

Extreme Outcomes. Bayesian approaches assume that service quality over time is normally distributed, where variability is a surrogate for uncertainty. In contrast, many service operations metrics—employee labor, materials, resources allocated, response time, resolution time—are characterized by a *skewed*, nonnormal distribution. Their distributions are characterized by a lower boundary of zero, a majority of observations within a certain range, and a few extreme outcomes. For example, a supplier can usually deploy a single technician to deliver a Service X—using certain materials and procedures—within 24 hours of the firm's request. However, the supplier may sometimes need to allocate additional resources—such as deploying an expert engineer—to deliver the same Service X. In this paper, we investigate how infrequent but extremely high (or low) levels of delivered service influence subsequent purchase decisions.

Timing. Experiments that manipulate perceived quality in the laboratory potentially mask effects arising from the *timing* of service experiences across natural purchase intervals. Yet, the timing of service experiences may critically influence subsequent purchase decisions. For example, if a supplier allocates a substantial amount of resources, such as engineer work minutes, to a firm early in the relationship, will this affect the firm's decision to renew a service contract differently than if the supplier allocates a substantial amount of resources late in the contractual relationship, closer to the renewal decision? We address this issue by tracking service experiences over a multi-year time frame, and comparing the effects of early versus recent service experiences on service contract renewal decisions.

Model Formulation

We now develop a dynamic model of the firm's service contract renewal decision by considering how two temporal features influence the firm's renewal decision: extreme outcomes in prior service experiences and the timing of service experiences.

Extreme outcomes arising from variability

There is substantial evidence from the judgment and decision-making literature that variability over time creates experiences that can be encoded favorably or unfavorably (Loewenstein 1988), where losses typically loom larger than gains (Kahneman and Tversky 1979; Tversky and Kahneman 1991). Buying firms' decisions (based on organizational norms) are likely to be equally sophisticated, especially in technology-intensive markets (John, Weiss, and Dutta 1999). Thus, firms are likely to consider higher moments of the service quality distribution, those beyond mean service quality levels.

A natural extension of people's loss aversion is that options *with extreme values within an offered set* will be relatively less attractive than options with intermediate values (Tversky and Simonson 1993). People use a weighted average of transactions to form judgments based on past experiences—so extreme positive or negative outcomes can be very influential (Fredrickson and Kahneman 1993).[2] Exposure to an option with a favorable or unfavorable extreme outcome may lead firms—managed by people—to prefer (or become averse to) that option in future service renewal decisions.

How do extreme values arise from variability in prior service experiences? How do they influence the firm's contract renewal decisions? Typically, service contracts specify upper (or lower) bounds on certain aspects of service delivery and service suppliers manage operations to achieve certain targeted levels within these bounds (Holcomb 1994). However, there will be variability in actual experienced service levels over time for a given firm. For example, a support service supplier might offer a fixed-price contract that promises to send a technician within 4 hours of a request. The majority of firms receive a visit from a technician that resolves the problem, but some requests necessitate a visit from a team, including an engineer, to resolve the problem. The visit involving the team may be more favorably evaluated because it indicates that the supplier has provided exceptional service. The supplier is willing to "go the extra mile" by sending extensive resources (an entire team) to fulfill its service promise within the context of a fixed-price contract. The firm makes the same request in both cases, but it is the additional allocation of resources by the service supplier that we hypothesize influences the firm's evaluation of the experience and subsequent contract renewal decision. When service contracts specify upper or lower bounds, we believe a disproportionate frequency of extreme experiences relative to the targeted service level will influence the individual firm's subjective expected value for a service contract. Thus,

H1: Favorable (unfavorable) extreme outcomes experienced over prior time periods ($Extreme_{t-1}$) will positively (negatively) influence firms' renewal decisions for service contracts at time t, after controlling for average service levels.

Timing of service interactions

Recent experimental results indicate that the order or timing of service experiences has an important influence on people's reference points and that *people* shift their reference points after a stimulus is presented (Chase and Dasu 2001). Furthermore, there is empirical evidence that reference points, or expectations, regarding key marketing variables influence *firm* performance (Glazer, Steckel, and Winer 1989; Rajendran and Tellis 1994). Hence, we investigate whether the timing of service experiences influences firms' decisions regarding service contracts. Following Boulding et al.'s (1993) model of perceived service quality, we can represent the firm's predictive expectation of the experience quality associated with a service contract by an averaging model, in which experience quality at time t is a weighted average of past (actual) service experiences (Q_{mt}), summed over service attributes $m = 1 \dots M$. Algebraically,

H2a: $ExperienceQ_t = \Sigma_m Q_{mt-1} + \omega (Q_{mt} - Q_{mt-1})$ $t > 1$ (4a)

Equation 4a is a flexible specification that subsumes three special cases that have an interesting managerial interpretation:

- a prior-based model ($\omega = 0$) in which the firm's assessment of experience quality is completely dependent on information obtained at time $t - 1$ (Q_{mt-1}),
- a recency model ($\omega = 1$) where current information obtained at time t (Q_{mt}) completely supersedes any prior information, and
- an equal weighting model ($\omega = .5$) in which current and prior information are equally important.

We do not model trial—that is, the situation when a firm does not have past (actual) service experiences—so we do not consider initial purchase conditions (where $t = 1$).[3]

Alternatively, firms might be loss averse or respond to upward or downward trends. Firms might make comparisons so that *deviations* from past experience directly influence the service contract renewal decision. Prospect theory predicts that the deviation from a reference point (positive or negative) influences people's decision making (Kahneman and Tversky 1979; 1984). Loss aversion has been shown for consumer decisions regarding services (Bolton 1998; Bolton and Lemon 1999), but prior research has not investigated whether deviations in service operations over time influence firm behavior. Hence, we investigate whether deviations influence the firm's decision to renew a service contract.

Hypothesis 2b: $ExperienceQ_t = \Sigma_m \alpha (Q_{mt} - Q_{mt-1})$ $t > 1$ (4b)

In both H2a and H2b, the firm's service contract renewal at time t depends on its assessment of experience quality at time $t - 1$.

Summary

We expand our base model by rewriting Equation 3 to include extreme values, as well as covariates to reflect differences across firms and service contracts.

$V_{ijt} = v\,(DesignQ_{t,}\ ExperienceQ_{t-1,}\ Price_{t,}\ Extreme_{t-1,}\ Covariates)$ (5)

Equations 1, 2, and 5 describe a dynamic model of service contract renewal, whereby firms adjust the expected value of a service contract to recognize the uncertainty of post-choice experience quality *($ExperienceQ_{t-1,}$ $Extreme_{t-1}$)*. Equations 4a and 4b describes how expectations are formed for experience quality.

Method

The model is estimated with data describing very large firms who purchase system support services from a global supplier. System support contracts consist of a complex bundle of services. For example, telecommunications companies provide voice and data system support. Firms buy multiple service contracts if they own multiple systems. They typically purchase a separate contract for each system. However, they may decide not to purchase contracts for some systems—instead providing support internally or doing without support—or they may purchase system support contracts from several different suppliers.

Firms obtain system support services by purchasing contracts that can range in price from $15,000 to $300,000, depending on the nature of the contract. The service promised by the support contract can be low, medium, or high, where a higher level corresponds to an incremental increase in bundled services. Furthermore, support is provided for two aspects of systems, hereafter called Technology A and Technology B. An illustrative example of different technologies within a single system might include voice and data lines within telecommunications systems. Each technology has different incidences of support requests and is supported through different service delivery mechanisms. All system support contracts promise 24/7 support with guaranteed response within 2 hours. However, the contracts do not promise to resolve the support request within a certain timeframe; instead, they promise to escalate the handling of highly critical support requests through devotion of additional resources.

An important feature of this supplier's support contracts is that a firm pays a fixed amount for support over a specified time period. The contract price is not dependent on usage levels. A firm's usage of support services is triggered by a system request, but not necessarily a system failure. Typically, the incidence of system requests is not within the control of the firm or the support supplier. (It may be partially within the control of the company that designed and manufactured the system.) As discussed later in the paper, we conducted statistical tests regarding potential covariates to control for the incidence and nature of support requests.

Prior research by the supplier

To gain an understanding of which service operations factors might influence customer renewal, we worked with the supplier firm to conduct some preliminary research, described below as phases 1–3.

Phase 1. A market research company specializing in high technology products and services conducted in-depth interviews with 55 respondents. Respondents included chief information officers, management information system managers, and service technicians who were identified from the supplier's records, and screened by the market research company to ensure that they were involved in the decision-making process for service contracts (either recommending or making the final decision). These respondents worked for firms that purchased between 5 and 300 system support contracts—not all necessarily purchased from the cooperating supplier. The market research company's report recommended that the supplier should deliver consistent, immediate access to an engineer that knows the customer's system and their environment to increase customer loyalty. For example, a German respondent remarked, "I want to talk to the most knowledgeable person who knows my environment."

Phase 2. The cooperating supplier commissioned a customer satisfaction and loyalty survey of the people who make recommendations or decisions regarding system-support contracts for its large business customers. There was a response rate of 30%, yielding 263 observations. Statistical models were developed that linked respondents' system support satisfaction to service attribute perceptions. For example, European decision-makers' satisfaction was positively related to the supplier's ability to meet commitments for technology A requests and respond quickly to technology B requests ($p < .05$).

Phase 3. We translated customers' perceptions of system support quality into objective and concrete measures derived from the service operations database (Acosta-Mejia 1998;

Table 1a
Study Context*

	Descriptive Statistic
Average years as customer of the cooperating supplier	6.8
Estimated share of customer	36%
Mean total support dollars (all suppliers)	704K (U.K.), 870K (Germany)
Number of competing suppliers	9.1
Self-reported intention to renew all contracts (5 = very likely to renew)	4.3
Self-reported % low support contracts (of total contracts with cooperating supplier)	50%

* These data were extracted from preliminary survey research as described in the text. As elsewhere, dollar values are scaled to preserve the confidentiality of the cooperating supplier's data.

Table 1b
Sample Statistics

Variable	Descriptive Statistic
Number of customers with complete data for 24 months	143
Number of contracts	2,442
Average duration of firm's relationship with service supplier	7 years
Number of customers included in model estimation (>5 contracts in Europe)	120
Renewal rates	88%
Contracts with A incidents	23%
Contracts with B incidents	18%
Average A engineer work minutes / contract	28
Number of extreme values for A in past two years	Range 0-5
Average B engineer work minutes / contract	15
Number of extreme values for B in past two years	Range 0-7

Bolton and Drew 1994), following well-established procedures (cf., Kordupleski, Rust, and Zahorik 1993; Zeithaml and Bitner 2000). Our analyses indicated that *average engineer work minutes per contract* for technology A and technology B were two managerially actionable service operations measures that represented system support quality—specifically, experience quality—and could potentially predict service contract renewal. There are separate measures of engineer work minutes per contract for technology A and technology B, respectively.

Engineer work time per contract is not equivalent to response time or resolution time per contract (averaged across requests for a given contract). A customer service representative typically logs a supplier's response to a request for system support and then forwards the request to a service technician or occasionally to an engineer. Hence, engineer work time is not related to response time, resolution time, or usage of total resources—in the same way that the time a patient spends with a doctor is not related to the time he or she spends waiting to be seen, total time elapsed to obtain a diagnosis (including time spent with medical technicians), or usage of total resources (such as x-rays or lab tests). Engineer time (or engineer minutes) represents the time spent by an engineer trained specifically on the technology (A or B) and does not reflect time spent by non-engineers on system support or waiting time (Kumar, Kalwani, and Dada 1997). The distribution of engineer work minutes per contract for technology A and technology B can also be considered to represent the buying firm's experience regarding consistent and immediate access to an engineer who knows the customer and their environment. For example, average engineer work minutes per contract for technology A are positively correlated with percentage of response time commitments met (.64) and percentage of resolutions that are defect free (.50) calculated across contracts ($p < .0001$).[4] Note that, for a given contract, more engineer time is preferred by the buyer (*ceteris paribus*) because it increases defect-free resolutions without affecting the fixed contract price.

The study database

Our dataset was constructed by drawing a stratified probability sample (where the strata are countries) from the cooperating supplier's list of large business customers. Very large business customers are a market segment composed of firms that operate certain enterprise-level systems. The dataset describes 143 firms from Germany and the United Kingdom that purchase system support services. Within Europe, firms face the same set of competing

Table 2

Constructs, Measures, and Descriptive Statistics*

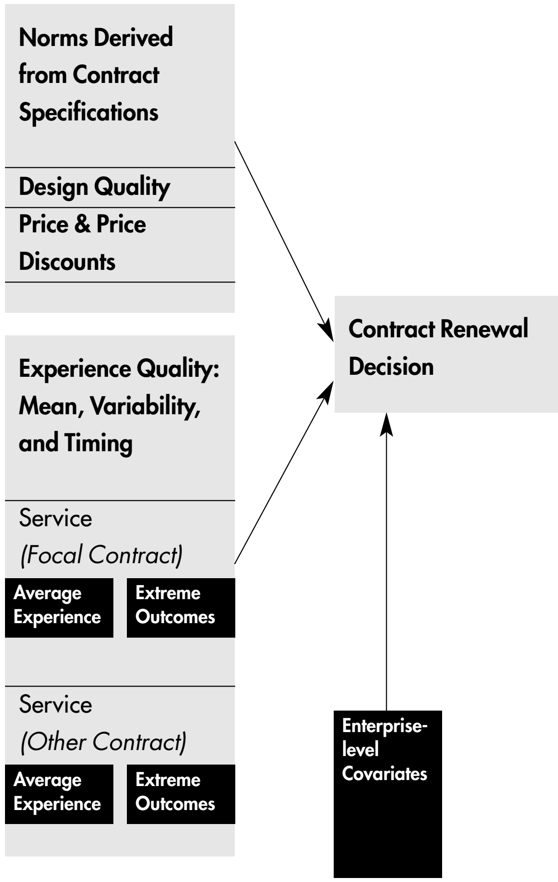
Construct	Measure*	Value
Extreme outcomes within the focal contract (H1)	Number of incidents in which tech. A engineer work minutes exceeds 240	.09 (.40)
	Number of incidents in which tech. B engineer work minutes exceeds 120	.11 (.46)
Average extreme outcomes across other contracts (H1)**	Number of incidents in which tech. A engineer work minutes exceeds 240 minutes, averaged across other contracts	.11 (.14)
	Number of incidents in which tech. B engineer work minutes exceeds 120 minutes, averaged across other contracts	.25 (.28)
Contract-level Covariates		
Experience quality for the focal contract	Avg. tech. A engineer work minutes for a support request	19.46 (40.58)
	Avg. tech. B engineer work minutes for a support request	6.0 (18.69)
Experience quality across other contracts**	Avg. tech. A engineer work minutes for a support request, averaged across other contracts	7.9 (9.65)
	Avg. tech. B engineer work minutes for a support request, averaged across other contracts	5.8 (6.49)
Design quality	Contract type–dummy variables for medium and high	Med: .15 (.35) High: .03 (.18)
Price norm	List price (divided by 1,000)	5.28 (6.62)
Deviation from price norm	Discount off list price (percent) Dummy variable for missing discount information	.20 (.20) .15 (.35)
Enterprise-level Covariates		
System characteristics	Avg. number of tech. A requests across contracts	2.6 (35.76)
	Avg. number of tech. B requests across contracts	.16 (.16)
	Number of contracts: 0-10	46%
	11-20	27
	>20	28

*Two righthand columns show mean and standard deviation across all observations for 1997 and 1998. Model operationalized with either 1997 values (prior-based model) or 1998 values (recency model).
** The measures of these constructs are subject to one of the four transformations described in the text.

suppliers. Firms simultaneously hold service contracts from nine (on average) different service suppliers. Descriptive statistics are shown in Table 1a and Table 1b.

We obtained firms' annual billing records describing the contracts they held in 1998, and then tracked whether they renewed these contracts in 1999. We also obtained monthly

Figure 1



service operations records for 1997-98 describing interactions associated with each contract between each firm and the supplier. These time series data were reported at the contract level (see Table 1b). Although all the information was drawn from the service supplier's information systems (that is, only the seller side of the dyad), the database provides an unusually complete description of the B2B relationship because it incorporates cross-sectional and time-series observations at the contract level and enterprise level. Thus, we are able to model the firm's service contract renewal decisions in 1999 (yes/no *vis-à-vis* 1998) as a function of its experiences in the previous two years (1997-98).

Measurement of model constructs

The measures of each model construct and their descriptive statistics are summarized in Table 2. The operationalized model is depicted in Figure 1. We determined whether or not a firm renewed a contract by comparing the billing records for January 1999 versus January 1998. A service contract *for a given system* is considered to be renewed if: (1) the firm purchases a new contract (for the same system) at the same service level or (2) the firm upgrades by purchasing a new contract (for the same system) at a higher service level. We include upgrades within renewals because they entail the renewal of the same set of bundled features (that is, the same contract), plus the purchase of additional features. We do not observe any instances in which firms downgrade (that is, purchase a lower quality contract).

Design Quality, Experience Quality, and Price. The design quality (*DesignQ*) specified in the service contract can be jointly represented by dummy variables that indicate whether the contract provides low, medium, or high levels of support. The contract's price (*Price*) is the list price, plus the discount (if any). The service supplier offers price discounts to certain customers—where discounts are typically higher when list prices are higher (and competition more intense)—so larger discounts are offered on contracts that firms are less likely to renew. Recall that prior research indicated that we could measure firms' experience quality with two process metrics: average engineer work minutes for technology A support incidents per contract, and average engineer work minutes for technology B support incidents per contract. Firms who are allocated, on average, more engineer work minutes to support a given contract should be more likely to renew their service contract (*ceteris paribus*). Thus, firms' service experiences ($ExperienceQ_{t-1}$) are represented by two variables: average engineer work minutes per contract allocated to system support for technology A and technology B. Instead of estimating ω, we measured $ExperienceQ_{t-1}$ using the average value for 1997 (prior-based model), 1998 (recency-based model), an average of 1997 and 1998 (equal weight model), or the difference between 1997 and 1998 (deviations model). See Table 2 for descriptions of each measure.

Extreme Outcomes. In qualitative research, managers and technical end users in customer organizations reported favorable responses to support incidents in which the supplier exceeded industry norms in its allocation of resources. Preliminary research suggested industry norms such that an engineer could resolve a technology A request in 4 hours and a technology B request in 2 hours. Hence, we derived measures of extremely favorable outcomes by identifying incidents in which engineer work minutes allocated for a technology A support request exceeded 240 and by identifying incidents in which engineer work minutes allocated for a technology B support request exceeded 120. The supplier's service operations records indicate that 13% of technology A support incidents entailed a resource allocation of more than 240 engineer work minutes, and 11% of technology B support incidents entailed an allocation of more than 120 engineer work minutes. Hence, we measured favorable extreme outcomes ($Extreme_{t-1}$) for a particular service contract by counting (separately) the number of support incidents above these cutoff values for technology A and technology B over the past two years.[5]

Treatment of Multiple Contracts. The preceding paragraphs describe two measures of experience quality and two measures of extreme outcomes—namely, average engineer work minutes per contract and counts of extreme values of engineer work minutes/incident for technology A and technology B. These four measures describe service *for a focal contract* (j). Since any firm may hold multiple contracts, firms' assessments of experience quality and favorable extreme outcomes *for other contracts provided by the same service supplier* may also influence their focal contract renewal decisions. Thus, it is both useful and relevant to calculate four similar measures to characterize the firms' average experience quality and extreme outcomes for each of the other contracts. For each measure, we average across the other contracts, $p = 1 \ldots p_i$ (where p is not equal to j)—thereby creating an additional four measures of experience quality and extreme outcomes. For example, if a firm has four contracts, each contract is, in turn, the focal contract with its own unique variables, with average values calculated across the other three contracts held by the firm.

Covariates. We incorporate the frequencies of technology A and technology B support incidents as covariates into Equation 5, measured as the average number of incidents calculated across all support contracts (that is, the focal contract and other contracts) between a particular firm and this service supplier. Both covariates are statistically significant ($p < .05$). Following Bolton and Drew (1994), we also conducted analyses to investigate whether specific characteristics of the service request—such as the type of support requests, the severity of the problem, or the average amount of downtime minutes per contract—should be included as covariates in the service contract renewal model. Statistical tests indicated that none of the aforementioned variables had a statistically significant effect ($p > .05$), after incorporating the frequencies of technology A and technology B support incidents as covariates.

Estimation Procedure

The model is summarized by equations 1, 2, and 5. Recall that, in this model, a firm renews a contract if the unobserved latent variable representing the value of support services exceeds a threshold value. Since firms may hold multiple contracts, a firm may make repeated (that is, dependent) renewal decisions—creating dependent observations for each firm. If the firm i makes k_i contract renewal decisions, the residual term e_{ijk} in Equation 2 can be expressed as follows:

$$e_{ijk} = \beta_i + \varepsilon_{ijk}, \quad (6)$$

where β_i is the unknown random effect for firm i and the ε_{ijk} are the independent residuals. It is assumed e_{ijk} can be decomposed into fixed and

Table 3

Hypothesis Two: Comparison of Models with Alternative Reference Points*

Reference Point	Log Likelihood	Akaike Information Criterion	BIC (Schwartz Criterion)
H2a: Prior-based model: Firm relies on 1997 experiences only	–746.75	1,525.5	1,512.89
H2a: Recency model: Firm relies on 1998 experiences only	–679.43	1,390.85	1,413.05
H2a: Equal weight model: Firm relies on both 1997 and 1998 experiences	–755.67	1,543.34	1,503.73
H2b: Deviations only model: Firm compares 1998 with 1997 experiences	–764.89	1,561.78	1,549.16
Model with maximum likelihood value	H2a supported (Recency Model, $\omega = 1$)		

*All models have the same number of exogenous variables, so log likelihood values are directly comparable. These results are for a model with a reciprocal transformation of $Extreme_{t\text{-}1}$, but a recency model dominates for the three other transformations.

random effects, where the distribution of random effects is assumed to be multivariate normal. Thus, our model can be specified and estimated as a binary logistic regression with mixed effects (Gibbons and Hedeker 1997; Hedeker, Gibbons, and Flay 1994). In a mixed effects model, the degree of dependency is jointly estimated with the usual model parameters, adjusting for the dependence resulting from the multiple contracts per firm. We can estimate this model with maximum likelihood techniques using MIXOR (Gibbons and Hedeker 1997). MIXOR uses marginal maximum likelihood estimation, utilizing a Fisher-scoring solution. We chose a complementary log-log response function such that $P(i,j,k^*) = 1\text{-}exp[\text{-}exp(V_{ijk^*} - V_{ijk})]$.

In this market, 120 firms (84%) purchase more than 5 contracts. Hence, we are able estimate a binary logistic regression with mixed effects based on 2,442 observations or contracts—that is, about 20 contracts/firm. Rodriguez and Goldman's (1995) simulation results have demonstrated that binary response models estimated with fixed effects, rather than random effects, can be biased when the underlying random parameter values are large. A fixed-effects model also produces slightly less conservative estimates of the coefficients when intraclass correlations are high. Therefore, by using a random effects model, we provide a stronger test of the hypotheses.

Results

Model comparisons and fit

Table 3 shows test statistics for models that incorporate alternative time frames as described in H2a and H2b—that is, reference points based on prior experiences, recent experiences, equal weighting of experiences, or deviations from prior experiences. The model incorporating recent service experiences dominates. We discuss this finding in more detail later in this section. We initially specified the subjective expected value, Equation 5, to be linear additive. In the linear additive specification, counting the number of extremely favorable outcomes implicitly assumes that the firm gives equal weight to all extreme outcomes associated with the same contract. The firm's subjective expected value for the service contract might be nonlinear with respect to extreme outcomes. For example, there might be diminishing marginal returns from favorable extreme outcomes. Hence, we estimated four models that incorporated alternative transformations of the $Extreme_{t\text{-}1}$ variables: (1) a conventional linear additive term ($Extreme_{t\text{-}1}$), (2) a quadratic term only ($Extreme_{t\text{-}1}^2$), (3) a natural logarithm transformation ($\ln(Extreme_{t\text{-}1})$), and (4) reciprocal transformation ($1/Extreme_{t\text{-}1}$). To be conceptually consistent, the same transformation was performed on all four $Extreme_{t\text{-}1}$ variables. Note that, for the logarithmic and reciprocal transformations, when $Extreme_{t\text{-}1} = 0$, the trans-

Table 4

Logistic Regression Results Based on Recent Experiences, Reciprocal Transformation of *Extreme*$_{t-1}$

Construct	Measure		Exp. Sign	Coefficient (Std. Error)	Explanatory Power[1]
H1: Extreme outcomes *within contract*	Number of incidents in which tech. A engineer work minutes exceeds 240 minutes, in 1997-98		+	+ 1.254* (.765)	4.1%
	Number of incidents in which tech. B engineer work minutes exceeds 120 minutes, in 1997-98		+	− .889 (.617)	
H1: Extreme outcomes *across other contracts*	Number of incidents in which tech. A engineer work minutes exceeds 240 minutes, in 1997-98		+	+ .018*** (.006)	
	Number of incidents in which tech. B engineer work minutes exceeds 120 minutes, in 1997-98		+	+ .016 (.015)	
Experience quality *for focal contract*	Tech. A engineer work minutes per request in 1998		+	.007 (.007)	77.1%
	Tech. B engineer work minutes per request in 1998		+	+ .926*** (.154)	
Experience quality *for other contracts*	Tech. A engineer work minutes per request in 1998		+	+ .069*** (.016)	
	Tech. B engineer work minutes per request in 1998		+	+ .184*** (.031)	
Design quality	Contract type:	Medium	+	.165 (.266)	.7%
		High	+	− .389 (.281)	
Price norms	List price		−	− .050*** (.018)	
Discount from normative (list) price	Discount off list price		−	− 2.489*** (.705)	5.4%
	Discount missing		n/a	+ 2.280*** (.304)	
Enterprise-level covariates: system characteristics	Avg. number of tech. A requests across contracts		n/a	+ .059** (.025)	12.7%
	Avg. number of tech. B requests across contracts		n/a	− 5.704*** (.869)	

Constant			.128
Log Likelihood	−679.43***	Hit Rate	88%
Pseudo R^2	17%	Mean Absolute Deviation	.12

* $p < .10$, ** $p < .05$, *** $p < .01$

[1]Based on the coefficient multiplied by the standard deviation of the variable, scaled to sum to 100.

formed values is undefined so (instead) we set it equal to zero.

Log likelihood function values for the recency model with these four transformations are: linear (–688), quadratic (–704), natural logarithm (–696), and reciprocal (–679). These values are directly comparable because the number of variables in the model does not vary. Based on the log likelihood function value (as well as the Akaike Information Criterion and the Schwartz Criterion), the model incorporating reciprocal transformations of the four recent measures of $Extreme_{t-1}$ dominates. The order of the two tests—alternative reference points or alternative transformations—does not affect any of our results. We discuss results regarding functional form in more detail later in this section.

Table 4 displays the final model that incorporates recent experiences and reciprocal transformations of the four measures of $Extreme_{t-1}$. This model is discussed in the remainder of the paper. The model fits the data reasonably well, with a pseudo-R^2 value of 17%. Unlike ordinary least squares, the pseudo R^2 for a logistic regression model is calculated by comparing the estimated model with an equal probability model. Hence, the pseudo R^2 value for the model is satisfactory. The hit rate (88%) compares favorably with Morrison's (1969) proportional chance criterion (78%). Since the aforementioned hit rate was calculated when the model was estimated on the entire sample, we also calculated the hit rates using split sample methods. Specifically, the model was estimated on a random sample of 75% of the observations and predictions were made for the corresponding holdout sample—the remaining 25% of the observations. This procedure was repeated three times to evaluate the model's predictive ability. The hit rates for the three holdout samples were 87%, 84%, and 86%. These values also compare favorably with the proportional chance criterion. In the remainder of this section, we discuss the results of the hypothesis tests, in the order in which the hypotheses were first presented.

Extreme outcomes

H1 predicts that favorable extreme outcomes will increase the likelihood of contract renewal. In this study, favorable extreme outcomes are measured by the count of incidents with extremely high engineer work minutes over a two-year period. This hypothesis is supported: favorable extreme outcomes in technology A, both within the focal contract and across other contracts, significantly increases the likelihood of renewal ($p < .05$, $p < .01$). Thus, the results indicate that exceptional efforts by the supplier (positive extreme outcomes) are recognized and valued by firms, and there is a spillover effect across contracts. Conversely, a lack of effort (unfavorable variability or no positive extreme outcomes) within and/or across contracts is associated with a lower likelihood of renewal. Recall that we speculated that the firm's subjective expected value for the service contract might be nonlinear with respect to extreme outcomes. As described earlier, we estimated models that incorporated four alternative transformations of the four $Extreme_{t-1}$ measures and found that the reciprocal transformation dominated. Hence, a single "extra mile" experience on a contract is weighed more heavily than multiple experiences on the same contract.

Favorable extreme outcomes account for 4% of the explained variance. The significant role of extreme outcomes in service contract renewal decisions is consistent with prior cross-sectional research concerning managerial decision making. In a B2B context, the mere labeling of performance as positive or negative, with or without a reference point, has been found to affect perceptions of risk by top management (Sitkin and Pablo 1992) and to affect organizational action (Neale et al. 1986). The incidents with extremely high engineer work minutes, and consequently effort, are beyond industry norms for support service—as well as reference points set by firms' own recent contract experience—and are viewed as positive or favorable performances.

Assessments of experience quality

H2a and H2b are competing hypotheses about how firms assess experience quality. H2a with $\omega = 0$ states that a firm's service contract renewal decision at time t depends entirely on prior experiences (1997 operations data); H2a with $\omega = 1$ states that it depends entirely on recent experiences (1998 operations data); H2a with $\omega = .5$ states that it depend equally on prior and recent experiences (measured by averaging

1997 and 1998 operations data); and H2b states that deviations or trends directly influence the firm's decision (measured by subtracting the 1997 operations data from the 1998 operations data). The four alternatives were assessed by comparing the test statistics shown in Table 3.

The recency model fits the data better than the other three models. This result implies that average service operations experiences are evaluated directly, rather than as a deviation from earlier experiences. Moreover, the timing of experiences is important to firms—recent experiences are weighed more heavily than early experiences. The dependence of the renewal decision on *recent* service operations is consistent with the results of psychological experiments in which people give more weight to experiences at the end of a series (Loewenstein 1988; Loewenstein and Prelec 1993; Varey and Kahneman 1992). It is also consistent with Hansen and Danaher's (1999) finding that judgments of service quality and purchase intentions are driven more by the performance of the final event than the initial event—regardless of the trend in service levels. In summary, we find support for H2a, with assessments of experience quality based upon recent interactions with the service supplier.

Covariates: Design quality, experience quality, and price

Experience quality accounts for the majority of the explained variance: 77.1% (see Table 4, right column). The variable measuring the average engineer work minutes for technology B incidents for the focal contract shows the strongest effect, accounting for 68.5% of the explained variance in the model, and experience quality on other contracts accounts for 8.6%. In contrast, the variables controlling for design quality and price—including any price discount—account for 6.1% of the explained variance. This finding is consistent with a Mittal, Kumar, and Tsiros (1999) study of purchase intentions in the automobile industry that showed that the importance of service attributes increases over time, whereas the importance of design attributes decreases. Variables indicating medium or high support levels are not statistically significant (see Table 4, $p > .05$), so the repatronage behavior of firms does not appear to depend on contractual terms regarding service, although the initial purchases may. Repatronage behavior does depend on price; higher-priced contracts are less likely to be renewed (see Table 4, $p < .01$). The effect of a discount off list price is significant and negative (see Table 4, $p < .01$), so that a larger discount is significantly associated with a smaller likelihood of renewal. We expected this result because discounts are offered to firms on contracts that are subject to more intense competition, and list price and discount off list price are correlated at .33 ($p < .01$).

The effect of average engineer work minutes per request for technology A for the focal contract is not significant ($p > .05$). However, the effect of the average engineer work minutes per request for technology B for the focal contract is significant and positive ($p < .01$). After controlling for the average number of service requests for technologies A and B, when engineer work minutes are high for technology B, it is likely the contract will be renewed. Moreover, average engineer work minutes for technology A and technology B service requests on other contracts with the same supplier also influence the renewal decision. Both coefficients are positive and statistically significant ($p < .01$). When engineer work minutes are high for technology A or technology B for other contracts held by the firm (thereby ensuring efficient and effective service), it is likely the contract will be renewed. Firms prefer higher utilization levels of engineering services because they have paid a fixed price, rather than a variable rate based on usage (Bolton and Lemon 1999).

Discussion

Since the dependent variable is binary (renew, not renew), it is useful to interpret the logistic

Table 5
Managerial Interpretation of Logistic Regression Results

Scenario	Probability of Renewal	Change in Probability vs. Base Scenario	Change in Number of Contracts (Revenue**) per Firm	Change in Revenue** across 143 Customers
Base Scenario: Firm holds a single high support contract, list price = $5,300,* engineer work minutes for tech. A and B for focal contract = sample averages, no extreme outcomes, covariates = sample averages	.84	–	–	–
Scenario 1: Base Scenario, plus firm holds other contracts, where engineer work minutes for tech. A and B for other contracts = sample averages	.96	.12	2.09 ($11.10)	$1,586
Scenario 2: Scenario 1, except engineer work minutes for tech. A and B for focal contract = zero (i.e., no experiences on focal contract)	.72	−.12	−2.09 (−$11.12)	−$1,590
Scenario 3a: Base Scenario, plus extreme outcomes for tech. A for focal contract = sample average	.85	.01	.16 (.85)	$123
Scenario 3b: Base Scenario, plus one "extra-mile" experience for tech. A for focal contract	.95	.11	1.84 ($9.74)	$1,392
Scenario 4a: Scenario 1, plus extreme outcomes for tech. A for other contract = sample average	.97	.13	2.23 (11.80)	$1,687
Scenario 4b: Scenario 1, plus one "extra-mile" experience for tech. A for other contract	.97	.13	2.10 ($11.15)	$1,594

*Throughout the analyses reported in the paper, price has been adjusted by a constant scale factor to preserve the confidentiality of the supplier's data.
**Revenue expressed in 1,000s of dollars.

regression results in terms of the effect of changes in the predictor variables on the probability that the firm renews the contract, as well as on revenue. Table 5 shows some sensitivity analyses for alternative scenarios. We begin with a Base Scenario, in which the firm holds a single contract, purchased at the average list price for the sample, with typical experience quality levels (engineer work minutes for technologies A and B for the focal contract equal to the sample averages), and typical system characteristics (covariates are set equal to the sample averages). There are no extreme outcomes—no "extra mile" experiences. In the Base Scenario, the probability that the firm will renew a single typical service contract is .84. Scenario 1 is similar to the Base Scenario except that the firm holds other contracts with typical experience quality levels (engineer work minutes for technologies A and B for other contracts equal to sample averages). In Scenario 1, the probability that the firm will renew the focal contract, when it is one of many similar service contracts provided by the supplier, is .96. Experiences with other contracts from the same supplier spill over and dramatically affect the firm's probability of renewing the focal contract. Scenario 2 is similar to Scenario 1: the firm holds other contracts with typical experience quality levels, but has *no prior service experiences* for the focal contract. The probability that the firm will renew the focal contract, given no service experience, is much lower: .72.

Scenario 3a is similar to the Base Scenario (no other contracts), except that the firm has experienced a *typical number of extreme outcomes* for technology A on its focal contract. The reciprocal of the number of extreme outcomes for the focal contract is set equal to the sample average. In Scenario 3a, the probability that the firm will renew the focal contract is .85—only slightly higher (.01) than the Base Scenario. In Scenario 3b, when the firm experiences a single "extra mile" experience on the focal contract, the probability that it will renew the focal contract is .95—an increase of .11. The magnitude of the influence of *a single "extra mile" experience* on the probability of the firm's renewal of the service contract is comparable to the influence of holding many other similar contracts. Thus, comparing these four scenarios demonstrates the powerful effect of an "extra mile" experience, and also how its effect can be diluted if there are recurring extreme outcomes on the same service contract.

Scenario 4a is similar to Scenario 1 (the firm holds multiple contracts), except that the firm has also experienced extreme outcomes for technology A *for other contracts* (the reciprocal of the number of extreme outcomes for other contracts is set equal to the sample average). In Scenario 4a, the probability that the firm will renew the focal contract is .97—only slightly higher (.01) than Scenario 1—though substantially higher than the Base Scenario (.13). Scenario 4b is similar to Scenario 1 (the firm holds multiple contracts), except that the firm has also experienced *a single "extra mile" experience* on each of its *other contracts*. In Scenario 4b, the probability that the firm will renew the focal contract is also .97—only slightly higher (.01) than Scenario 1—comparable to Scenario 4a. Comparing scenarios 1, 4a, and 4b demonstrates that the effect of an "extra mile" experience spills over from other contracts to the focal contract, but the magnitude of the effect is much smaller than if the extra experience had been for the focal contract.

The two right-hand columns of Table 5 show the revenue implications for each scenario. We calculate the *change* in the number of contracts held by the firm by multiplying the change in the probability of service contract renewal by the average number of contracts held by a firm. We calculate the change in revenue by multiplying the change in the number of contracts by the average list price ($5,300, where list price has been arbitrarily scaled to preserve the confidentiality of the results). We calculate the change in revenue for the supplier (derived from the 143 firms in the European dataset) by multiplying by 143. These scenarios are necessarily somewhat artificial: the Base Scenario describes a firm with a single contract—all firms do not have this profile, and the supplier has many more customers than 143. However, the projected dollar values are substantial, and give some notion of revenue implications of the scenarios we prepared for managers of the cooperating supplier. For example, the difference in revenues between Scenario 1 (average utilization levels) and Scenario 2 (no utilization of the focal contract) is over $22,000. These projections provide powerful evidence that the supplier should consider reallocating resources—especially engineer work minutes—across contracts and firms over time. In particular, some firms and contracts are receiving multiple "extra mile" experiences whereas other firms and contracts are receiving none. We believe that the supplier should consider proactively creating a single "extra mile" experience for each of the neglected contracts and firms, while reducing the average number of extreme outcomes across contracts.

Contribution to marketing theory

We believe that this study is the first attempt to model the influence of firm/supplier interactions over time on the firm's repatronage decision, and the first attempt to examine the effects of extreme outcomes in service delivery over time on this decision. Our study contributes to our understanding of customer retention in several ways. First, we find that modeling the renewal decision at the individual contract, rather than the overall firm, level provides new insights into the firm's decision. Second, we show that firms attend both to their normative expectations of the service contract (list price) and their experiences with the service contract over time (experience quality) when making the decision to renew a contract. Third, we show that extreme outcomes over time—within and across the contracts that comprise a firm/

supplier relationship—have a significant effect on the repatronage decision. A few instances of delivering exceptional quality to the firm can have a significant, positive effect on renewal. In our study context, recent experiences (in the past year) are particularly important. Overall, this research suggests that firms determine whether or not to renew a contract by evaluating the extent that their experiences with the service supplier—on several dimensions, considering distinct reference points—deliver value.

Enterprise-level versus Contract-level Analysis. Lusch and Brown (1996) observed that many studies have shown that the length of the channel relationship has little effect on a number of important channel constructs (such as trust) and that relationship length may not be useful for explaining channel phenomena. For example, Reinartz and Kumar (2000) recently reported that relationship length has a small correlation with future customer lifetime value. In conjunction with our results, these observations suggest that measures of general relationship constructs may not explain firm behavior as well as specific, experience-based constructs measured at the contract or product level. This conclusion is consistent with Oliver's (1999) argument that satisfaction does not completely explain customer loyalty. Specifically, we find that firms utilize information regarding both the focal contract and the other contracts held by the firm in their decision of whether to retain the focal contract. This finding suggests that it is important to analyze B2B relationships at the contract, product, and/or site level, rather than at the enterprise level. Models that fail to take the richness of the individual-level information into account may lead to inappropriate or less-than-optimal models of resource allocation for the firm.

Experience Quality: Extreme Outcomes and Timing. Service contract renewal is directly influenced by past service experiences—not contract specifications or deviations from norms or earlier experiences. Firms were more likely to renew contracts for which there were higher levels of resources allocated to support services (high average engineer work minutes) and a few incidents of extremely high levels of resources (favorable extreme outcomes). Surprisingly, a few favorable extreme outcomes per contract positively influence business customer retention. It suggests that understanding the firm's decision context is critical to successfully managing service request responses and the firm/supplier relationship. Furthermore, firms gave more weight to recent experiences when deciding whether or not to renew a contract, suggesting that the *timing* of these experiences is critical. This implies that *identical support incidents* may be evaluated very differently, depending upon whether they occur early in the contract relationship or closer to the renewal decision.

Managerial implications

This study has implications for suppliers seeking to maximize the likelihood that customers are retained, and it suggests specific strategies for allocating marketing and operations resources over the duration of the firm/supplier relationship. Current practice typically focuses on managing variability *across* contracts and customers (for example, meeting targets across all firms served), rather than variability *within* contracts (for example, targeting specific contracts and firms). However, certain business customers may be systematically underserved because conventional quality control mechanisms (across customers, over time) fail to capture, recognize, or create solutions for this within-customer issue. Hence, the results suggest that suppliers managing relationships with firms that hold multiple service contracts should carefully manage the amount and timing of resources allocated to each contract to deliver value with-in the firm/supplier relationship. As suppliers continue to coordinate marketing and operations activities across the entire value chain (TQM, Six Sigma, process reengineering), understanding the effects of these elements on the firm/supplier relationship is especially critical.

Sufficient Utilization of Contract. Failure to ensure adequate utilization of support services—within and across the portfolio of contracts held by a firm—is likely to lead to decreases in share of customer. It may also ultimately lead to the termination of the firm/supplier relationship. When fixed-price contracts for support services are underutilized, firms are averse to repurchasing them. In our sensitivity analysis, the probability that a firm renewed a contract when it did not utilize the supplier's services for the focal contract was 72%. This is substantially lower (12%) than if it had typical utilization levels for the focal contract. The revenue implications are substantial, as well. The probability of renewal would be even lower if the firm did not hold other contracts. Currently, the supplier reactively allocates resources in response to customer requests. To increase customer retention, suppliers should consider proactively allocating resources within and across all contracts relevant to the firm/supplier relationship, rather than simply reacting to requests for service or support.

At Least One "Extra-Mile" Experience. In this study, favorable extreme outcomes increased contract renewal, supporting the idea that organizational memory—like human memory—recalls service in terms of snapshots of extreme service experiences over time. This observation suggests that suppliers should look for opportunities to deliver exceptional service to each firm—occasionally exceeding upper bounds for the expected level of service. The current rule of thumb in service delivery is typically underpromise, overdeliver. However, this heuristic fails to identify the *firm* to whom the supplier should deliver exceptional service and *when* the service should be delivered and *how often*. It also fails to recognize that firms adapt to over-delivery (multiple extreme outcomes on the same contract), thus decreasing impact on firm behavior. To address this issue, the supplier described in this study might ensure that its engineers spend sufficient time on each contract to effectively handle each request—rather than, say, delegating some responsibilities to less efficient or effective service technicians. In the short run, additional resources are allocated to each request. However, in the long run, it is likely that multiple extreme outcomes would become less frequent. The supplier could also make proactive support calls or visits to firms who do not request support over a specified time period (for a particular contract) to provide a favorable experience with system support services.

Customer Retention. Finally, the model suggests that it is critical to incorporate service operations metrics into models of repatronage decisions. The results from this study suggest that it may ultimately be possible to predict retention behavior solely from internal records (such as CRM systems, operational databases), without utilizing perceptual measures as mediators. Suppliers may be able to implement successful customer retention programs if they keep accurate records of individual firm experiences (at the contract level), and update relationship strategies for individual business customers as their circumstances change.

Limitations, conclusions, and directions for future research

In this research, we have examined antecedents of the service contract renewal decision utilizing a longitudinal, multicountry, cross-sectional database. We have only examined this model in one product category—high technology support services. In future research, it will be important to extend the research to other product categories. For example, we believe that this type of dynamic model could be very useful in computer-mediated environments that provide customer-specific service experiences.

Causal Attributions. In this research, we suggest that extreme outcomes are important in the decision. Firms may not follow a strict temporal integration model in making decisions about service contracts, but a weighted average of prior transactions in which extreme outcomes, or "snapshots" are highly influential. Another explanation for the influence of extreme outcomes on the decision is provided by attribution theory (Folkes, Koletsky, and Graham 1987). Within the customer organization, managers' attributions regarding extreme values—such as the service supplier is making an extra effort (or that it is inept)—may lead them to value the contract more (or less). This notion is consistent with Narayanan and Lehmann (1998, p. 309) who report that managers tend to exhibit control bias. They ask: what event that I can control could have caused this event?

Thus, extreme outcomes in service experiences provide opportunities for managers within firms to make causal attributions toward service suppliers that may lead to favorable (or unfavorable) decisions regarding contract renewal. Consequently, we might expect that, in other study contexts, the decision-maker's attributions about locus (whether the service experience is attributed to the supplier or the customer) and controllability (whether the service experience was preventable or due to circumstances beyond its control) might vary. Attribution theory and intertemporal decision-making research underscore the importance of extreme outcomes in understanding firms'—and people's—behavior. Additional research is needed to understand the process by which these extreme values affect managerial decision making.

Dynamic Reference Effects. We believe that we have only scratched the surface in researching and understanding the role of dynamic reference effects on firm decision making (see also Rajendran and Tellis 1994). Future research in this area could focus on understanding, in more depth, how managers within organizations share information and create shared beliefs, assumptions, and norms that guide service contract renewal decisions. In addition, future research could examine the specific mechanism by which extreme outcomes (both favorable and unfavorable) affect the managers' decisions—as well as firms' decisions—perhaps in a laboratory context.

Conclusion

Over the past decade, firms have invested extensive resources into customer relationship management systems to track interactions and transactions at the customer level over time. We believe that our model and findings provide some insights into how to exploit this information. Our study highlights the importance of developing dynamic models of customer decision making. If, as this research suggests, the extent and timing of the supplier's interaction with a firm influences decision making, ignoring such dynamic effects will result in incorrectly specified models of firm behavior—and, consequently, incorrect allocation of marketing resources. As the field of marketing seeks to deepen its understanding of buyer/seller relationships, understanding the effects of marketing, operations, and service decisions over time will be imperative. ■

Acknowledgements

The authors gratefully acknowledge the assistance of the anonymous cooperating supplier. The opinions expressed in this paper are solely those of the authors and do not reflect the views or policy of the United States government or the National Center for Health Statistics.

Notes

1. An alternative approach is to incorporate uncertainty through a Bayesian updating model. If we assume that a customer's uncertainty about a measurable characteristic of an alternative can be characterized by a normal distribution, then the expected utility that he or she will derive from the alternative depends on the mean and variance of the characteristic's distribution. If we assume an exponential utility function, then expected utility can be determined by a tradeoff between expected value and risk (Jia and Dyer 1996).

2. Fredrickson and Kahneman (1993, p. 46) note, "Extra moments of misery could make the overall experience less aversive if the added moments are less miserable than others and are given substantial weight. In an averaging model, the weights assigned to individual moments are

constrained to add to unity for any episode….We favor a special case of the averaging model, in which most moments of an episode are assigned zero weight in the evaluation, and a few select 'snapshots' receive larger weights."

3. We do not consider values of ω other than 0, .5 and 1. Instead of estimating ω, we measure experience quality using either data from 1997 (early), data from 1998 (recent), or both (i.e., equal weighting or deviations), as shown in Table 2. In preliminary analyses, we also investigated whether quarterly (rather than annual) measures were appropriate; annual measures performed better.

4. Percentages are calculated using the total number of requests for the contract as the denominator. The company's definition of defect free is that no re-work was required for the system within 90 days. Correlations were calculated for the most recent year of operations.

5. It is inappropriate to measure favorable extreme variability by calculating the standard deviation of engineer work minutes per contract because positive deviations have a different meaning than negative deviations, and negative deviations are truncated at zero. Extreme favorable variability can be measured by skewness, but such statistics are not easily understood by company managers. Hence, we simply counted extremely favorable outcomes that lie above the cutoff values. Since extreme outcomes are, by definition, rare, we must count their frequency over a two-year time period to observe a sufficient number of nonzero observations for model estimation.

References

Acosta-Mejia, C.A. (1998), "Monitoring Reduction in Variability with the Range." *IIE Transactions* 30, 515–23.

Anderson, Eugene W., Claus Fornell, and Roland T. Rust (1997), "Customer Satisfaction, Productivity, and Profitability: Differences Between Goods and Services." *Marketing Science* 16 (2), 129–45.

Bolton, Ruth N. (1998), "A Dynamic Model of the Duration of the Customer's Relationship with a Continuous Service Provider: The Role of Satisfaction." *Marketing Science* 17 (1), 45–65.

__________, and J.H. Drew (1994), "Linking Customer Satisfaction to Service Operations and Outcomes." In *Service Quality: New Directions in Theory and Practice*, eds. R.T. Rust and R.L. Oliver, 173–200. Newbury Park, Calif: Sage Publications.

__________, and Katherine N. Lemon (1999), "A Dynamic Model of Customers' Usage of Services: Usage as an Antecedent and Consequence of Satisfaction." *Journal of Marketing Research* 36 (May), 171–86.

Boulding, William, A. Kalra, Richard Staelin, and Valarie A. Zeithaml (1993), "A Dynamic Model of Service Quality: From Expectations to Behavioral Intentions." *Journal of Marketing Research* 30 (February), 7–27.

Cannon, J. P., and W. D. Perreault, Jr. (1999), "Buyer-Seller Relationships in Business Markets." *Journal of Marketing Research* 26(4), 439–62.

Chase, Richard B., and Sriram Dasu (2001), "Want to Perfect Your Company's Service? Use Behavioral Science." *Harvard Business Review* 79 (June), 78–84.

Coleman, J.S. (1990), *Foundations of Social Theory*. Cambridge, Mass.: Harvard University Press.

Coughlan, Anne T. (1985), "Competition and Cooperation in Marketing Channel Choice: Theory and Application." *Marketing Science* 4 (2), 110–29.

Deming, W.E. (1986), *Out of the Crisis*. Cambridge, Mass.: MIT Center for Advanced Engineering Study.

Dwyer, F. Robert, Paul H. Schurr, and Sejo Oh (1987), "Developing Buyer-Seller Relationships." *Journal of Marketing* 51 (2), 11–27.

Folkes, Valerie S., Susan Koletsky, and John L. Graham (1987), "A Field Study of Causal Inferences and Consumer Reaction: The View from the Airport." *Journal of Consumer Research* 13 (March), 534–9.

Fredrickson, B.L., and Daniel Kahneman (1993), "Duration Neglect in Retrospective Evaluations of Affective Episodes." *Journal of Personality and Social Psychology* 65 (1), 45–55.

Ganesh, Jaisahkar, Mark J. Arnold, and Kristy E. Reynolds (2000), "Understanding the Customer Base of Service Providers: An Examination of the Differences Between Switchers and Stayers." *Journal of Marketing* 64 (3), 65–87.

Gibbons, R.D., and D. Hedeker (1997), "Random-effects Probit and Logistic Regression Models for Three-level." *Biometrics* 53 , 1527–37.

Gibbons, R.D., and D. Hedeker (1997), "A Random-effects Ordinal Regression Model for Multilevel Analysis." *Biometrics* 50 , 933–44.

Glazer, Rashi, Joel H. Steckel, and Russell S. Winer (1989), "The Formation of Key Marketing Variable Expectations and Their Impact on Firm Performance: Some Experimental Evidence." *Marketing Science* 8 (1), 18–34.

Hammer, M. (2001), "The Superefficient Company." *Harvard Business Review* (September), 82–91.

Hansen, D.E., and P.J. Danaher (1999), "Inconsistent Performance during the Service Encounter: What's a Good Start Worth?" *Journal of Service Research* 1 (February), 227–35.

Hedeker, D., R. D. Gibbons, and B. R. Flay (1994), "Random-Effects Regression Models for Clustered Data: With an Example from Smoking Prevention Research." *Journal of Consulting and Clinical Psychology* 62, 757–65.

Heide, Jan B., and George John (1992), "Do Norms Matter in Marketing Relationships?" *Journal of Marketing* 56 (2), 32–44.

__________, and Allen M. Weiss (1995), "Vendor Consideration and Switching Behavior for Buyers in High-technology Markets." *Journal of Marketing* 59 (3), 30–43.

Holcomb, M.C. (1994), "Customer Service Measurement: A Methodology for Increasing Customer Value through Utilization of the Taguchi Strategy." *Journal of Business Logistics* 15 (1), 29–52.

Jia, J., and J.S. Dyer (1996), "A Standard Measure of Risk and Risk Value Models." *Management Science* 42 (December), 1691–1705.

John, George, Allen M. Weiss, and Shantanu Dutta (1999), "Marketing in Technology-Intensive Markets: Toward a Conceptual Framework." *Journal of Marketing* 63 (Special Issue), 78–91.

Juran, J.M., and A.B. Godfrey, Eds. (1998), *Juran's Quality Handbook*, 5th ed. New York, N.Y.: McGraw-Hill.

Kahneman, Daniel, and Amos Tversky (1979), "Prospect Theory: An Analysis of Decisions Under Risk." *Econometrica* 47 (March), 263–91.

__________, and __________ (1984), "Choices, Values and Frames." *American Psychologist* 39 (April), 341–50.

Kalwani, Manohar U., and Narakesari Narayandas (1995), "Long-Term Manufacturer-Supplier Relationships: Do They Pay Off for the Supplier Firm?" *Journal of Marketing* 59 (1), 1–16.

Kordupleski, R., Roland T. Rust, and Anthony J. Zahorik (1993), "Why Improving Quality Doesn't Improve Quality (Or Whatever Happened to Marketing?)" *California Management Review* 35 (Spring), 82–95.

Kumar, P., M.U. Kalwani, and M. Dada (1997), "The Impact of Waiting Time Guarantees on Customers' Waiting Experiences." *Marketing Science* 16 (4), 295–314.

Levy, David (1997), "Lean Production in an International Supply Chain." *Sloan Management Review* 38 (Winter), 94–102.

Loewenstein, G.E. (1988), "Frames of Mind in Intertemporal Choice." *Management Science* 34 (2), 200–14.

__________, and D. Prelec (1993), "Preferences for Sequences of Outcomes." *Psychological Review* 100 January, 91–108.

Lusch, Robert F., and James R. Brown (1996), "Interdependency, Contracting, and Relational Behavior in Marketing Channels." *Journal of Marketing* 60 (4), 19–38.

MacDuffie, J.P., and S. Helper (1997), "Creating Lean Suppliers: Diffusing Lean Production Through the Supply Chain." *California Management Review* 39 (4) Summer, 118–51.

Mittal, Vikas, Pankaj Kumar, and Michael Tsiros (1999), "Attribute-Level Performance, Satisfaction and Behavioral Intentions over Time." *Journal of Marketing* 63 (2), 88–101.

Morrison, Donald G. (1969), "On the Interpretation of Discriminant Analysis." *Journal of Marketing Research* 6 (1), 156–63.

Narayanan, Sunder, and Donald R. Lehmann (1998), "An Investigation of Factors Influencing Causal Attributions in Managerial Decision Making." *Marketing Letters* 9 (3), 301–12.

Neale, M.A., M.H. Bazerman, G.G. Northcraft, and C. Alperson (1986), "'Choice Shift' Effects in Group Decisions: A Decision Bias Perspective." *International Journal of Small Group Research* 2 , 33–42.

Oliver, Richard L. (1999), "Whence Consumer Loyalty?" *Journal of Marketing* 63 (Special Issue), 33–44.

Parasuraman, A., Valarie A. Zeithaml, and Leonard L. Berry (1985) "A Conceptual Model of Service Quality and Its Implications for Future Research." *Journal of Marketing* 49 (4), 41–50.

Rajendran, K.N., and Gerard J. Tellis (1994), "Contextual and Temporal Components of Reference Price." *Journal of Marketing* 58 (1), 22–34.

Reinartz, Werner J., and V. Kumar (2000), "On the Profitability of Long-Life Customers in a Noncontractual Setting: An Empirical Investigation and Implications for Marketing." *Journal of Marketing* 64 (4), 17–35.

Rodriguez, G., and N. Goldman (1995), "An Assessment of Estimation Procedures for Multi-level Models with Binary Responses." *Journal of the Royal Statistical Society* A158, 73–89.

Rust, Roland T., J. Jeffrey Inman, Jianmin Jia, and Anthony Zahorik (1999), "What You Don't Know about Customer-perceived Quality: The Role of Customer Expectation Distributions." *Marketing Science* 18 (1), 77–92.

Sinkula, James M. (1991), "Market Information Processing and Organizational Learning." *Journal of Marketing* 58 (1), 35–45.

Sitkin, S.B., and A. Pablo (1992), "Reconceptualizing the Determinants of Risk Behavior." *Academy of Management* 17 (1), 9–38.

Stinchcombe, Arthur L. (1986), "Norms of Exchange." In *Stratification and Organization: Selected Papers*, ed. A.L. Stinchcombe, 231–67. Cambridge, U.K.: Cambridge University Press.

Tsikritsis, N., and J. Heineke (2004), "The Impact of Process Variation on Customer Dissatisfaction: Evidence from the U. S. Domestic Airline Industry." *Decision Sciences* 35 (1) Winter, 129–42.

Tversky, Amos, and Daniel Kahneman (1991), "Loss Aversion in Riskless Choice: A Reference-Dependent Model." *The Quarterly Journal of Economics* 106 (4), 1039–61.

_________, and Itamar Simonson (1993), "Context-Dependent Preferences." *Management Science* 39 (10), 1179–89.

Varey, C., and Daniel Kahneman (1992), "Experiences Extended over Time: Evaluations of Moments and Episodes." *Journal of Behavioral Decision Making* 5, 169–85.

Wathne, Kenneth H., Harald Biong, and Jan B. Heide (2001), "Choice of Supplier in Embedded Markets: Relationship and Marketing Program Effects." *Journal of Marketing* 65 (April), 54-66.

Zeithaml, Valarie A. (1988), "Consumer Perceptions of Price, Quality, and Value: A Means-End Model and Synthesis of Evidence." *Journal of Marketing* 52 (July) 2-22.

__________, and Mary Jo Bitner (2000), *Services Marketing: Integrating Customer Focus Across the Firm*, 2nd ed., 234-8. New York, N.Y.: Irwin-McGraw Hill.