MARKETING SCIENCE INSTITUTE

# MSI Reports



2009

WORKING PAPER SERIES

ISSUE FOUR

NO. 09-004

MARKETING SCIENCE INSTITUTE

# MSI Reports



2009

WORKING PAPER SERIES

ISSUE FOUR

NO. 09-004

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

*MSI Reports* (ISSN 1545–5041) is published quarterly by the Marketing Science Institute. 

The views expressed here are those of the authors.



**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 "Apply 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 2008–2010 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.

# Service Productivity Strategy

Roland T. Rust and Ming-Hui Huang

*Should companies always seek to maximize service productivity? This study finds that service quality may decline if firms automate too fast. The pursuit of short-term profits through cost reduction may damage future sales.*

**Roland T. Rust** is Distinguished University Professor and David Bruce Smith Chair in Marketing and Executive Director of the Center for Excellence in Service at the Robert H. Smith School of Business, University of Maryland. **Ming-Hui Huang** is Professor of Electronic Commerce in the Department of Information Management, National Taiwan University.

**Report Summary**

As service becomes an ever-larger part of every developed economy, service productivity has increasingly become the focus of attention. Many companies are using information technology to utilize automation more extensively, reduce the use of labor, and increase service productivity. This course of action may have negative consequences for customer retention, however, if service quality declines.

Here, Roland Rust and Ming-Hui Huang build a rigorous theory of service productivity and use it to derive empirical propositions. The propositions are tested using data from more than 700 service companies in two time periods. The empirical analysis largely supports the theory. Many important managerial implications emerge from the research, including (1) for a given level of technology, firms should seek an optimal level of productivity rather than try to maximize productivity; (2) as technology advances and automation costs decrease over time, the optimal level of productivity increases; (3) as price increases, the optimal level of productivity decreases; (4) as the future is given more weight, the optimal productivity level decreases; (5) as customer retention is more sensitive to the level of service, the optimal level of productivity decreases; (6) as wages rise, the optimal level of productivity increases; and (7) for firms with higher sales and market share, the optimal productivity level increases.

The authors show that too myopic a viewpoint and/or pursuit of too high a level of productivity boosts current sales at the cost of reducing the service level and future sales. Specifically, large service companies tend to be too productive, relative to the optimal level, and should place less emphasis on cost reduction through automation and more emphasis on providing good service. The authors also show that in a recession, relative use of labor should be greater and automation should be used relatively less. ■

## Introduction

Increasing service productivity (output divided by service labor input[1]) is widely believed to be a worthy goal for service companies. Michael Hammer, author of *Reengineering the Corporation*, says that "increasing service productivity through front-office reengineering is the challenge of the decade" (Hammer 2005). In addition, there is research evidence that increasing productivity by having customers perform part of the service themselves (co-production) can be beneficial (Bendapudi and Leone 2003). Yet there is substantial anecdotal evidence that productivity increases can go too far. Consider the following three examples:

"Their check-in counter was SUCH a mess. There was 8 echeck-in kiosks open, only 2 baggage handlers, 1 supervisor, and a long a** line of people waiting … I got the attention of the supervisor and notified her that I had a paper ticket. She waved me off and told me to check-in through a kiosk. I tried that and it's impossible … All the while, she's trying to get all the people who have checked in through a kiosk to form ANOTHER line to have their checked baggage tagged … Someone finally got the supervisor's attention and put their paper ticket about 6 inches in front of her eyes. She says, 'Oh, it IS a paper ticket. You can't check in through the kiosk. You're going to need a rep's help … But they are on break. They'll be back in about 10 minutes.' Only 2 reps show up 10 minutes later. They finally help all of us who have paper tickets … After that fiasco, we still had to wait in the damn bag tagging line."

(Yelp.com 2007)

"In February, I received an e-mail warning me that free online access would end by March 11 if I didn't 'renew now.'… So, I decided to cancel my Sirius service before the cutoff date … The only way to cancel is to contact customer service by the phone number the company publishes on its Web site. When I tried that option, I encountered a Byzantine automated-menu system and lengthy hold delays followed by hang-ups … for my third cancel-by-phone attempt, I tried this alternative number instead. After yet another wait of more than a half-hour, I reached someone."

(Dipert 2009)

"This evening, while doing my regular fund transfers on online HSBC … I gave all the particulars, clicked confirm, ok. Service error. Come back later. Oklah. Few minutes later, try again. Confirm. Ok. Went through. Moved on to the next transaction. Checked balance. The previous transaction went through twice! … Called up to HSBC, they denied everything. Said it's the customer's (aka ME) fault … Service. Go fly kite. Angry angry angry."

(*Terra Cin* 2009)

These examples suggest that companies may sometimes seek too high a level of service productivity, at the expense of service quality, customer satisfaction, customer retention, and future sales. We seek to investigate this idea, using both theoretical and empirical analysis.

Our goal in this paper is to contribute a new way of thinking about service productivity. The traditional way of thinking about service productivity is that it should be maximized, because greater productivity (all other things being equal) produces greater profits on the firm level and expands the economy on the aggregate level (Banker, Chang, and Natarajan 2005; Brown and Dev 2000; Brynjolfsson and Hitt 1994). Our view of service productivity is more nuanced. We build a new theory of service productivity that distinguishes between long-term structural effects, based on level of technology, and short- to medium-term effects, based on decisions about the tradeoffs between the use of service personnel and the use of automation.[2] As the level of technology increases, automation becomes more effective. But what is the right amount of automation at a particular point in time? In the short or medium term, our theory suggests that firms should seek an optimal level of service productivity—that is, service productivity should neither be too low nor too high. Our theory enables us to derive a number of important managerial propositions, and a broad-based empirical analysis shows that these propositions are largely confirmed.

In the last 100 years, the service sector has almost tripled as a percentage of most developed economies, and the importance of service in all facets of the economy has dramatically increased, leading to the current industry-led focus on service science (Maglio and Spohrer 2008; Spohrer and Maglio 2008). As service has transformed business and become an ever greater part of the economy, it has become increasingly apparent that the primary task of marketing is providing service, leading marketing theorists to call for a "service dominant logic" for marketing (Vargo and Lusch 2004). The service marketing area has developed and matured over the last 30 years, with many researchers exploring how to satisfy customers and provide better service (e.g., Berry and Parasuraman 1993; Fisk, Brown, and Bitner 1993; Lovelock and Gummeson 2004; Shostack 1977).

The key customer measure of service, customer satisfaction, has been shown to have a very important impact on many key performance metrics, such as consumer spending growth (Fornell, Rust, and Dekimpe in press), customer retention (Rust and Zahorik 1993), long-term profitability (Mittal et al. 2005), market share (Anderson, Fornell, and Lehmann 1994), share of wallet (Keiningham, Perkins-Munn, and Evans 2003), shareholder value (Anderson, Fornell, and Mazvancheryl 2004; Gruca and Rego 2005), stock prices (Fornell et al. 2006), and willingness to pay (Homburg, Koschate, and Hoyer 2005).

Although the impact of customer satisfaction is impressive, top executives are continually struggling with the tradeoff between improving service to customers and cutting costs. Researchers in marketing have shown that this tradeoff between customer satisfaction and productivity is especially pronounced in the service sector, in contrast to the goods sector, in which increasing customer satisfaction and increasing productivity often go hand in hand (Anderson, Fornell, and Rust 1997).

At the firm level, executives often focus on increasing service productivity to cut costs. For example, many airlines have installed automated kiosks at airports to reduce labor costs by reducing the need for check-in personnel. Many companies have set up automated phone answering systems that compel customers to negotiate seemingly endless menus to get the service they seek. The Internet has provided yet another means of reducing front-line service personnel and cutting costs. In general, the use of self-service technologies to drive service productivity has been one of the most important service consequences of advances in information technology (Baily and Lawrence 2001; Meuter et al. 2000; Meuter et al. 2005).

At the aggregate economy level, policymakers are concerned about service productivity because GDP growth is increasingly dependent upon it. It is of great concern, then, that service productivity growth has lagged behind productivity growth in the goods sector (Blackstone 2007). The implicit conclusion is that service productivity growth is essential to both firm profitability and the growth of the aggregate economy. This would seem to imply that firms should be maximizing their service productivity (Banker, Chang, and Natarajan 2005; Brown and Dev 2000).

We seek to provide a deeper view of service productivity by separating long-term productivity improvement effects, based on technology, from shorter-term effects based on the tradeoff between the use of front-line service personnel and the use of automation and self-service technology. We build a formal theory that suggests that while better technology that improves productivity is always better, in the shorter term there is an optimal level of productivity, based on an optimal level of service personnel and level of automation, beyond which it is not desirable to go. The tradeoff occurs because of the consideration of the value of customers over time, in addition to short-term profit maximization. Our theory also reveals the conditions under which the optimal level of productivity should be higher and lower. The empirical results largely confirm the conclusions from our theory.

We also use our theory to investigate how firms should manage service productivity during a recession. Although it is tempting for a company to increase productivity during an economic downturn, slashing costs by firing large numbers of workers, our analysis shows that it is actually better to allow service productivity to decline somewhat during a recession—keeping more service personnel than the short-term financial environment would tend to suggest.

In the next section of the paper, we build a theory of service productivity and use it to derive a number of managerially-relevant propositions. The third section of the paper describes the empirical analysis that we used to test our theory. The fourth section presents results, and we finish with discussion and conclusions. Details of the derivations underlying the theoretical development are provided in an appendix.

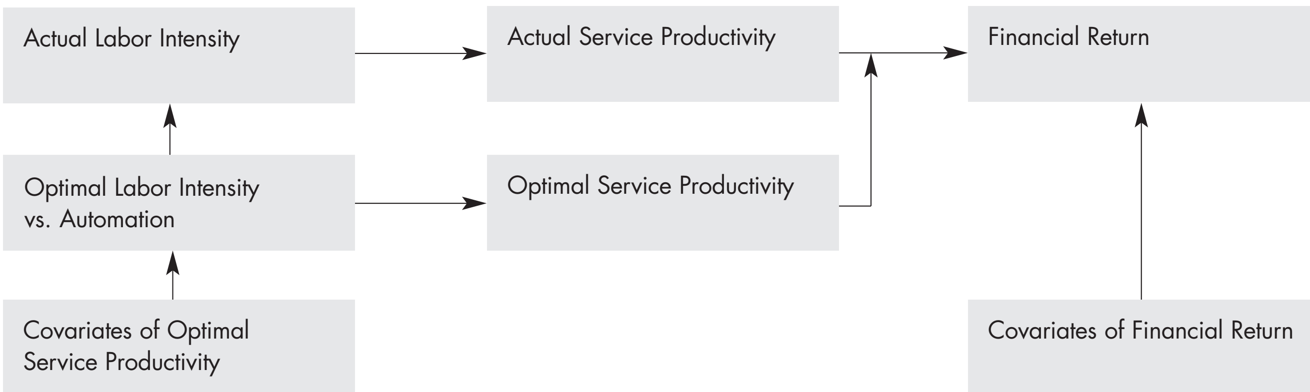
## A Theory of Service Productivity

### Overview

Productivity is typically defined as units of output divided by units of input. This means that for a manufacturing firm, the goal of maximizing productivity is often a reasonable one because the quality of the output can be maintained more or less constant. Traditional manufacturing-based cost-driven productivity concepts suffer insufficiencies in the service economy because they assume all units of output are identical, and they focus on single transactions. The service environment creates quite a different situation because the quality of the service provided often depends critically on the degree to which labor is employed, and the focus is typically on continuing customer relationships. For example, a 100-seat restaurant can achieve very high productivity by having only one waitperson, but at the cost of providing very poor service and losing customers. On the other hand, the same restaurant, if it wishes to have one waitperson for every table, can provide very excellent service, but at high cost. In other words, the service environment provides the opportunity for large variability in both productivity and level of service provided, depending upon decisions made about use of labor. This implies that for service firms, maximizing productivity may not be such a good idea. Instead, the firm should seek to find the balance between productivity and service quality that maximizes financial return, noting that a favorable financial return requires one eye to the future and satisfying and retaining customers (Parasuraman 2002). We wish to build a theoretical framework that incorporates these effects—placing service productivity decisions within a decision framework in which a firm maximizes its financial return.

Figure 1 presents a broad overview of our theoretical framework. We explicitly consider the firm's tradeoff between productivity and level of service provided. We posit that a firm's financial return is affected by the firm's service productivity level and the firm's optimal productivity level, with the best performance being achieved when the service productivity equals the optimal productivity level. Other covariates unrelated to productivity may also impact financial return. The service productivity is driven by the firm's decisions about labor intensity vs. use of automation. The dotted line in Figure 1 between optimal labor intensity and actual labor intensity indicates that the firm's decisions about the labor/automation tradeoff may be affected by its knowledge

Figure 1
**Theoretical Framework**




of the optimal labor intensity level. The optimal labor intensity level is, in turn, determined by other firm-specific variables in the marketplace.

To further explain our general theory, we first briefly summarize our theory in words, list its major assumptions, and then provide a more formal development. The key marketing problem in service productivity is how best to serve the customer—that is, what mix of service personnel (labor) and self-service technology (automation) should be used. We visualize the firm's goal as maximizing its financial return. To incorporate the impact of service level on ongoing customer relationships, our theory incorporates financial return over time (Gupta, Lehmann, and Stuart 2004; Rust, Lemon, and Zeithaml 2004; Srivastava, Shervani, and Fahey 1998), reflecting the idea that financial returns from service and service-based solutions involve long-term relational processes in addition to short-term sales (Tuli, Kohli, and Bharadwaj 2007).

The firm must pay a market wage per unit labor; it also knows the cost per unit for automation. The service level provided to the firm's customers is a function of the degree to which service is automated and the effectiveness of that automation. Less labor intensity in service increases service productivity and lowers costs, but at the price of a lower level of service (Brown and Dev 2000; Oliva and Sterman 2001).

Sales are the sum of new customer sales (attraction) and existing customer sales (retention). Both attraction and retention are affected by the previous-period service level, through such mechanisms as direct experience, word-of-mouth, positive public relations, etc. A higher previous service level makes it easier to attract new customers (Danaher and Rust 1996) and retain existing ones (Rust and Zahorik 1993).

The firm chooses what mix of labor and automation to use in delivering the service, with higher labor intensity resulting in a higher service level and greater value to the customer.[3] The goal should be to choose the mix that maximizes financial return. To reflect both current and future financial return, our formal theory operationalizes financial return as the discounted profit flows from the customer base (the customer equity), which has been shown to be an effective proxy of the value of the firm (Gupta, Lehmann, and Stuart 2004; Kumar and Shah 2009). The relative levels of labor and automation determine the service productivity level. We will show that for a given technology level and automation cost, an optimal service productivity level exists. We will further explore how the optimal service productivity level is affected by a number of important factors, including contribution margin, price, weight given to the future, the importance of service in driving retention, wage rate, automation costs, level of technology, and current sales level.

The term *productivity* refers to output divided by input. In service, the most important kind of productivity is *labor productivity*, which is typically calculated as output divided by labor hours or labor cost. For our purposes, we define *service productivity* as dollar sales divided by number of employees (e.g., Basker 2007; Bertschek and Kaiser 2004; Converse 1939). When wages are held constant (as assumed in our theoretical development), service productivity is proportional to sales divided by labor cost, the operationalization we will use when deriving our propositions, because that formulation is easier to work with.

## Assumptions

We summarize here the key assumptions that we use to drive our theory. We employ a two-period model in which the first period is the current decision period and the second period represents the future. This is a common approach in analytical decision modeling in marketing and economics (e.g., Desiraju and Shugan 1999; Kopalle and Lehmann 2006). Based on the logic and research cited in the previous section, we incorporate the following assumptions into our theory:

***Assumption A1:*** *The wage rate, cost of automation, and level of technology are fixed in the market in the short run and are known to the firm.*

***Assumption A2:*** *The firm chooses its mix of labor and automation in delivering service so as to maximize financial return, operationalized here as the discounted profit flows from its customers.*

***Assumption A3:*** *Attraction and retention are both affected by the previous-period service level.*

***Assumption A4:*** *Automation is cheaper than labor, per unit of service.*

***Assumption A5:*** *Quality increases as the amount of labor per unit of service increases.*

While these assumptions hold in a wide variety of industries (and in fact the propositions are tested across a broad spectrum of service industries) the assumptions may hold better in some industries, geographies, etc., than others. Assumption 1, about the wage rate being known, may not hold in some very new industries. Assumption 2 may not hold in some service organizations (e.g., government) where financial return is not the goal. Assumption 4, that automation is cheaper than labor, may not hold in some highly personalized industries (e.g., hair styling) where the technological level does not yet permit automation to be competitive. Assumption 5, that a higher labor intensity increases quality, may not be true for some industries in which technology is sufficiently advanced. In such industries, a totally automated service may be preferred.

## Formal theory

We consider a firm that seeks to attract and retain customers to maximize its financial return, operationalized here as the discounted profit stream from its customers. This goal is consistent with theory related to the marketing–finance interface (Srivastava, Shervani, and Fahey 1998) and customer equity (Gupta, Lehmann, and Stuart 2004; Rust, Lemon, and Zeithaml 2004). We posit a two-period model, with previous (implied) periods having resulted in a customer base that the firm sells to in Period 1. (Thus, we are considering an established firm, not an initial start-up.) We assume that the firm produces units of a service product, for which it charges a fixed price.[4] The firm sets its proportion of labor per unit (equivalently, it sets the proportion of automation per unit), and the observed service level is a function of the labor per unit and the automation per unit (with automation per unit weighted by the relative effectiveness of automation, which reflects the technological regime). Both attraction in Period 2 and the retention of Period 1's customers in Period 2 are a function of the service level in Period 1.

To formalize, let:
$\theta$ = proportion of labor per unit
$1 - \theta$ = proportion of automation per unit
$W$ = wages per unit labor
$A$ = automation cost, with $A < W$
$\alpha$ = level of technology (relative effectiveness of automation[5]), $0 < \alpha < 1$
$m$ = gross margin
$R$ = revenue per unit
$P$ = service productivity
$Q$ = unit sales

Then the labor cost per unit is $\theta W$, and the automation cost per unit is $(1 - \theta)A$. The service (labor) productivity is sales divided by labor cost, or $p = QR/Q\theta W = R/\theta W$. We assume that the service level, $S$, results from the extent to which automation is used, taking into account the relative effectiveness of automation:

$$S = \theta + \alpha(1 - \theta) = \alpha + (1 - \alpha)\theta \qquad (1)$$

Let $Q$ be the sales in Period 1, resulting from the previous service level. Then the Period 1 profit per unit is the gross contribution per unit, minus the labor costs, minus the automation costs. The total profit for Period 1 is thus:

$$p_1 = (mR - \theta W - (1 - \theta)A)Q = (mR - A - (W - A)\theta)Q \qquad (2)$$

We assume that $mR > A$ (otherwise profit for both periods would be guaranteed to be negative, and the firm would choose not to operate).

Period 2 customer retention, $r$, is a function of the service level:

$$r = \beta S = \beta\alpha + \beta(1 - \alpha)\theta \qquad (3)$$

where $\beta$ is the extent to which the service level drives retention.

Period 2 attraction, $a$, is also a function of the service level[6]:

$$a = \gamma S = \gamma\alpha + \gamma(1 - \alpha)\theta \qquad (4)$$

where $\gamma$ is the extent to which service drives customer attraction. This results in a second-period profit, $p_2$, of:

$$p_2 = (mR - A - (W - A)\theta)(\beta\alpha Q + \beta(1 - \alpha)\theta Q + \gamma\alpha + \gamma(1 - \alpha)\theta) \qquad (5)$$

Then the financial return of the firm is:

$$FR = p_1 + kp_2 \qquad (6)$$

where $k$ is the discount factor.

The firm's profit maximization problem is thus to maximize the financial return with respect to the degree of labor per unit, $\theta$. The maximum discounted profit occurs when[7]:

$\theta = 1$, if $\theta^* \geq 1$
$\theta = 0$, if $\theta^* \leq 0$
$\theta = \theta^*$ otherwise

where

$$\theta^* = ((mR - A)/2(W - A)) - (Q/(2k(1 - \alpha)(\beta Q + \gamma))) - (\alpha/(2(1 - \alpha))). \qquad (7)$$

If we restrict ourselves to the most common and most interesting case, in which neither all labor nor all automation is the best strategy, comparative statics on $\theta^*$ yield a set of profit maximization guidelines for managers with respect to the productivity strategy. Let us denote the optimal service productivity level as *OPT*, defined as the level of service productivity that maximizes financial return. *OPT* may be derived as follows (obtained by inserting the optimal labor intensity into the definition of productivity):

$$OPT = R/W\theta^* \qquad (8)$$

We note that $\partial OPT/\partial\theta^* < 0$, implying that as the optimal labor effort per unit decreases, optimal productivity increases.

The following eight propositions regarding the covariates of optimal productivity emerge from the preceding analysis (proofs are given in the Appendix):

***Proposition 1:*** *As margin increases, optimal productivity decreases.* As profit margins are higher, the firm should use more labor and less automation. The intuition here is that satisfying a customer is worth a lot, because of the value of the customer. High profit margins incentivize the firm to "pull out all the stops" to satisfy the customer by providing better service.

***Proposition 2:*** *As price increases, optimal productivity decreases.* As the price level is higher, the firm should use more labor and less automation. The reasoning is similar to the reasoning for Proposition 1: higher prices mean the customer becomes more valuable, providing incentive for better service.

***Proposition 3:*** *As the future is weighted more, optimal productivity decreases.* When the future is more important, satisfying customers today becomes more important, to increase customer retention. This implies a higher service level and less attention to productivity.

***Proposition 4:*** *As the extent to which service drives retention increases, optimal productivity decreases.* When service is more important to drive customer retention, having better service is more important, suggesting that productivity should be emphasized less.

***Proposition 5:*** *As wages rise, optimal productivity increases.* Automation becomes more attractive as labor gets more expensive. This is because there is more benefit (cost savings) from substituting cheaper automation for more expensive labor.

***Proposition 6:*** *If wages are low enough, as automation costs decrease, optimal productivity increases.* For typical parameter values, as automation becomes cheaper, it is better to use more of it. Exceptions to this, although seemingly counterintuitive, may occur in extreme cases if service is exceptionally effective in driving future customer attraction. In that case, if wage levels are high, cost savings from reduced automation costs can be used to "buy" more labor, to increase future customer attraction, which decreases productivity.

***Proposition 7:*** *As the level of technology increases, optimal productivity increases.* Technological advance makes automation more effective at providing service, leading to more use of automation.

***Proposition 8:*** *As initial sales increase, optimal productivity increases.* The intuition is that if the firm already has a lot of customers, then making money from them in the current period is more important. If the firm has few customers, but needs to get more, then the service level needs to be emphasized more than current profits, to grow the firm's customer base.

The theory also enables us to derive (see the Appendix) the following additional propositions:

***Proposition 9:*** *For a given level of technology, there is an optimal productivity level.* Driving productivity to a level that cannot be supported by the current level of technology boosts short-term profit at the expense of overall financial return, because the resulting decline in quality hurts future sales.

***Proposition 10:*** *Too myopic a viewpoint (putting not enough weight on future profits) boosts short-term productivity and profitability at the expense of overall financial return.* Too much attention to the present at the expense of the future leads to reducing service levels to cut costs in the short run.

### Service productivity in a recession

To explore this issue, we use the above theoretical model, but note that first-period sales will be depressed in the current period because of reduced demand. That is, the market's response to previous-period service level will be less than what it otherwise would be. This means that the current-period sales, $Q$, will be lower. The managerial result of this temporarily reduced demand is the following (see the Appendix for proof):

***Proposition 11:*** *In a recession, the firm should make relatively more use of labor (as opposed to automation) than usual, and the optimal level of service productivity will be lower than usual.* During a recession, the future is more important than the present[8] because it is harder than usual to obtain current sales and profits. This leads to greater emphasis on service to increase future customer attraction and retention, leading to a higher use of labor (relatively speaking) and less emphasis on productivity.

In spite of this result, one notes that many or even most companies cut their labor force in downturns. This is not necessarily inconsistent with the above result. If the labor force is cut *proportionally less* than sales drop, then productivity will actually decrease even as the labor force is being trimmed. The above result does suggest, however, that it is possible to lay off too many workers, with the result being that current service and future sales and profits are harmed.

It should be noted that the above result does not depend on the wage rate being lower in a recession. To the extent that the wage rate drops during a recession, the above result will become even stronger.

## Method

### Empirical model

We build an empirical model that can test most of the propositions obtained in the previous section. We first propose the empirical model in general terms, and then later show how the model is operationalized with respect to data and specific measures and how the propositions from the previous section may be tested. The general empirical model may be expressed as:

$$FR_j = \delta_0 - \delta_1(P_j - OPT_j)^2 + \Sigma_c D_c Y_{cj} + \varepsilon_j \quad (9)$$

where $FR_j$ is the financial return for company $j$; $P_j$ is firm $j$'s labor productivity; $OPT_j$ is the optimal level of productivity for company $j$; $Y_{cj}$ is covariate $c$ of financial return for firm $j$; $\delta_0$, $\delta_1$ and the $D_c$'s are model parameters to be estimated; and $\varepsilon_j$ is the error term, assumed normally distributed. This formulation explores whether or not an optimal productivity level exists. It is general enough to include the case where there is an optimal value ($\delta_1 > 0$) as well as the case where productivity should be maximized ($\delta_1 < 0$).

Most of the propositions from the theory in the previous section involve statements about how various variables (e.g., profit margin, prices, etc.) impact the optimal productivity level. The empirical model captures these by expressing the optimal productivity level, $OPT_j$, as a function of the levels of these covariates, $X_{ij}$, of the optimal productivity level.[9] If $X_{ij}$ is covariate $i$ for firm $j$, then the expression for $OPT_j$ is:

$$OPT_j = \Sigma_i B_i X_{ij} \quad (10)$$

where the $B_i$'s are coefficients to be estimated. Combining equations 9 and 10, we obtain a nonlinear equation that enables us to estimate all of the model parameters at once, using nonlinear estimation methods:

$$FR_j = \delta_0 - \delta_1(P_j - \Sigma_i B_i X_{ij})^2 + \Sigma_c D_c Y_{cj} + \varepsilon_j \quad (11)$$

In our empirical model, we test many of the variables that the theory tells us should have an impact on the optimal level of productivity. Specifically, we define the following covariates for optimal productivity:

$$\begin{aligned} X_{1j} &= \text{profit margin} \\ X_{2j} &= \text{price level} \\ X_{3j} &= \text{growth rate} \\ X_{4j} &= \text{industry concentration} \\ X_{5j} &= \text{wage rate} \\ X_{6j} &= \text{firm age} \end{aligned} \tag{12}$$

We show later how these covariates are operationalized using our data. We use a nonlinear least squares estimation with the Gauss-Newton iterative method to estimate the model parameters. This estimation regresses the residuals onto the partial derivatives of the model with respect to the parameters until the estimates converge.

### Testing the propositions

The empirical model provides empirical evidence for or against many of the propositions put forward in the theory of the previous section. We will now explain the correspondence between the propositions resulting from our theory and the empirical model.

The coefficient for $X_{1j}$, $B_{1j}$, provides a direct test of Proposition 1. Based on Proposition 1, $B_{1j}$ should be negative. Thus, we can test the null hypothesis, $B_{1j} \geq 0$, against the alternative hypothesis, $B_{1j} < 0$. Likewise, based on Proposition 2, $B_{2j}$ should be negative, and based on Proposition 5, $B_{5j}$ should be positive.

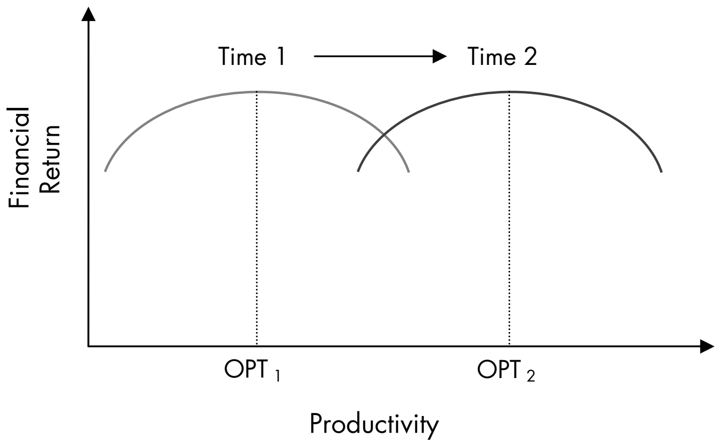
Proposition 3 cannot be tested directly, because we do not have a direct measure of how much the future is weighted. We can, however, construct a proxy for this variable. Presumably a higher sales growth rate will lead to weighing the future more. (For example, if a company multiplied its sales by 100 each year, it would be obvious that the future would be much more important than the present.) Thus we test Proposition 3 using the growth rate. Based on Proposition 3, we would expect $B_{3j}$ to be negative.

Similarly, we need to build a proxy to test Proposition 4 because we do not have a direct measure, from publicly available sources, of the degree to which service drives retention. However, we can argue that retention becomes more important when there are many competitors, because the customer has more choices. Thus, our proxy for the degree to which service drives retention is market concentration. If the market is more concentrated, that means there are fewer competitors to which customers can flee, meaning the impact of the service level on retention is reduced. Thus, based on Proposition 4, we would then expect that $B_{4j}$ would be positive.

Propositions 6 and 7 require variables for automation costs and level of technology. While it is not possible to obtain direct measures of these variables that are comparable across different industries, we can test these propositions indirectly. It is well known that technology does not go backwards over time ("you can't put the genie back in the bottle"), and automation costs are known to decrease over time as technology advances. Thus, if we estimate Equation 9 at two points in time, we would expect from Propositions 6 and 7 that the average optimal productivity level across companies would increase over time (see Figure 2). Thus a test of whether the average $OPT_j$ increases over time provides indirect evidence for or against propositions 6 and 7.

For the initial sales level, which is needed for Proposition 8, we use firm age as a proxy, making the assumption that mature firms will have higher levels of initial sales than start-up firms. Proposition 9, that an optimal level of productivity exists, can be tested directly using the empirical model. If $\delta_1 > 0$, the firm's financial performance will be best when its productivity is equal to the optimal level of productivity. If $\delta_1 < 0$, then financial return will be maximized when the firm's productivity is as large as possible.

Figure 2
**Optimal Productivity over Time**




### Data

We tested the model empirically using the Compustat North America database, which includes more than 30,000 active and inactive publicly held U.S. and Canadian companies. We focused our analysis on service firms with North American Industry Classification System (NAICS) codes of 42–92 for the years 2002 and 2007 as those choices made it possible to investigate both robustness of the results over time and trends in optimal productivity. NAICS was adopted by the U.S. Census Bureau in April 1997 to replace the Standard Industrial Classification (SIC) system. NAICS better reflects the current economy structure, including as it does new service sectors such as information and professional, scientific, and technical services. Our universe of service firms for this study is comprehensive, in that it includes all service sectors: wholesale and retail trade, transportation and warehousing, information, finance and insurance, real estate and rental and leasing, management of companies and enterprises, administrative and support services, educational services, health care, entertainment and recreation, accommodation and food services, and other services. After removing firms with missing values, we have 741 firms in 2002 and 751 firms in 2007 for hypothesis testing. Table 1 summarizes the company characteristics for the two data years.

### Measures

In our empirical model (Equation 11), we operationalize financial return as return on assets (ROA). ROA is one of the most frequently used indicators for assessing firm financial performance in the marketing literature (e.g., Aksoy et al. 2008; Anderson, Fornell, and Mazvancheryl 2004; Fornell et al. 2006; Narver and Slater 1990; Noble, Sinha, and Kumar 2002). The relationship between productivity and optimal productivity is used as a predictor of ROA, along with two additional widely used covariates, firm size and selling, general, and administrative expenses (SG&A). The predictors of optimal service productivity include profit margin, selling price, sales growth rate, Herfindahl–Hirschmann index, wage rate, and firm age. Table 2 summarizes the descriptive statistics, correlations, and variance inflation factors (VIFs) for all measures for the two data years.

The specific operationalization of the variables in the empirical model is as follows:

**Return on Assets (ROA).** ROA is calculated as net income divided by total assets. It measures the utilization of the assets of a firm, a performance indicator that reflects the role of technology and labor. ROA is a popular gauge of profitability because it is relatively more stable than other return indices such as return on equity (ROE); can be calculated for companies with negative shareholder's equity; is useful for analyzing competing companies in the same industry (Anderson, Fornell, and Mazvancheryl 2004); and is useful for gauging the profitability of a company on an absolute basis. High ROA firms are more profitable than low ROA firms.

**Labor Productivity.** Following prior research (e.g., Basker 2007; Bertschek and Kaiser 2004; Converse 1939; Datta, Guthrie, and Wright 2005; Guthrie 2001; Huselid 1995), we calcu-

Table 1
**Summary of Company Characteristics**

| | 2002 | | 2007 | |
|---|---|---|---|---|
| | Number of firms | Percentage | Number of firms | Percentage |
| **Industry** | | | | |
| Wholesale and retail trades and logistics | 46 | 6.21% | 47 | 6.26% |
| Information and technical services | 98 | 13.23% | 67 | 8.92% |
| Finance and insurance | 530 | 71.52% | 574 | 76.43% |
| Education and health care | 12 | 1.62% | 12 | 1.60% |
| Recreation, food, and accommodation | 55 | 7.42% | 51 | 6.79% |
| **Sales (in millions of dollars)** | | | | |
| < 10 | 45 | 6.07% | 22 | 2.93% |
| 10–100 | 326 | 43.99% | 316 | 42.08% |
| 100–500 | 189 | 25.51% | 196 | 26.10% |
| 500–1,000 | 68 | 9.18% | 61 | 8.12% |
| > 1,000 | 113 | 15.25% | 156 | 20.77% |
| **Number of Employees** | | | | |
| < 50 | 35 | 4.72% | 27 | 3.60% |
| 50–100 | 59 | 7.96% | 52 | 6.92% |
| 101–400 | 227 | 30.63% | 270 | 35.95% |
| > 400 | 420 | 56.68% | 402 | 53.53% |

late labor productivity as the log of sales per employee where dollar sales are used to capture total output and number of employees is used to measure labor input.[10] The sales per employee metric is considered to be a good measure of labor productivity, with its greatest use being to compare industry competitors and to examine the historical performance of the company.

**Covariates of Optimal Productivity.** There are six variables included as the covariates of optimal productivity. *Profit margin* is calculated as the proportion of a firm's net sales to its gross sales. It shows how much of a firm's sales dollars are profit. *Price* is calculated as the ratio of selling costs to one minus the proportion profit margin. In this formula, price is the total price summed over all services provided by a firm, not the unit or average price of services. The wide variations in a firm's service offerings would make the unit price, even if it were available, not easily comparable across firms and industries. *Sales growth rate* is calculated as the difference between a firm's current year's sales and the previous year's sales divided by the previous year's sales. *Market concentration* index (HHI) is the sum of the square of market shares (Schmalensee 1977) at the four-digit NAICS level. *Wage rate* is defined as labor expenses per employee. *Firm age* is operationalized as the years since a firm's financial data become available in the Compustat data set (a proxy for years since the firm was established).

**Covariates of ROA.** We include two variables as the covariates of ROA: firm size and SG&A. Firm size is calculated as the log of a firm's number of employees (e.g., Huselid 1995; Koch and McGrath 1996) with the expectation that the larger the firm, the better the firm can combat competition, regardless of

Table 2

**Descriptive Statistics of Variables**

| Variables | M | SD | VIF | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **2002 (N = 741)** | | | | | | | | | | | | | |
| 1. ROA | –.03 | .24 | | 1.00 | | | | | | | | | |
| 2. Log labor productivity | 5.15 | .90 | 4.23 | .06 | 1.00 | | | | | | | | |
| 3. Gross margin (%) | .52 | .48 | 2.73 | .42** | .22** | 1.00 | | | | | | | |
| 4. Selling price | 1,470.67 | 5,212.81 | 1.66 | .02 | .14** | .02 | 1.00 | | | | | | |
| 5. Sales growth rate (%) | –.01 | .23 | 1.04 | –.01 | .07 | .10** | .07 | 1.00 | | | | | |
| 6. Market concentration | .02 | .02 | 1.22 | –.20** | –.33** | –.19** | .00 | .04 | 1.00 | | | | |
| 7. Log wage rate | 3.60 | .97 | 3.75 | –.09* | .81** | .03 | .12** | –.02 | –.23** | 1.00 | | | |
| 8. Firm age (year) | 11.03 | 8.68 | 1.41 | .04 | .02 | .03 | .35** | –.05 | .07* | .09** | 1.00 | | |
| 9. SG&A expense (in millions of dollars) | .29 | .42 | 3.14 | –.55** | –.14** | –.76** | –.02 | –.13** | .02 | .13** | –.01 | 1.00 | |
| 10. Log firm size | –.19 | 2.03 | 2.17 | .17** | –.27** | –.02 | .54** | .03 | .15** | –.24** | .48** | –.14** | 1.00 |
| **2007 (N = 751)** | | | | | | | | | | | | | |
| 1. ROA | –.00 | .21 | | 1.00 | | | | | | | | | |
| 2. Log labor productivity | 5.52 | .92 | 3.42 | .11** | 1.00 | | | | | | | | |
| 3. Gross margin (%) | .48 | .23 | 1.12 | .41** | .06 | 1.00 | | | | | | | |
| 4. Selling price | 3,579.58 | 13,300.45 | 1.77 | .03 | .25** | –.03 | 1.00 | | | | | | |
| 5. Sales growth rate (%) | .15 | .24 | 1.19 | .12** | .15** | .12** | .13** | 1.00 | | | | | |
| 6. Market concentration | .03 | .03 | 1.26 | –.11** | –.28** | –.30** | .06 | –.00 | 1.00 | | | | |
| 7. Log wage rate | 3.92 | .83 | 3.04 | –.07* | .79** | .04 | .21** | .03 | –.27** | 1.00 | | | |
| 8. Firm age (year) | 13.25 | 9.37 | 1.41 | .01 | .00 | –.07* | .23** | –.12** | .16** | .04 | 1.00 | | |
| 9. SG&A expense (in millions of dollars) | .26 | .26 | 1.26 | –.72** | –.14** | .03 | –.08* | –.06 | –.01 | .08* | –.06 | 1.00 | |
| 10. Log firm size | –.17 | 2.14 | 2.47 | .20** | –.17** | –.06 | .54** | .18** | .25** | –.19** | .46** | –.25** | 1.00 |

Note. * $p < .05$, ** $p < .01$. All variables are firm-level variables, except for HHI, which measures market concentration in a firm's industry. VIF was obtained using OLS regression with ROA as the dependent variable. Log firm size is the log transformation per 1,000 employees. Selling price is the total price summed over all services provided by a firm, not the unit or average price of services.

service quality, and the higher the ROA will be. SG&A is a standard reported item in a firm's financial statement that includes all salaries, indirect production, marketing, and general corporate expenses. By normalizing SG&A as a percentage of SG&A per dollar of sales, we expect a negative relationship between SG&A and ROA because higher costs directly reduce a firm's return.

**Industry Effects.** We include a set of industry dummies in the estimation in order to explicitly model industry heterogeneity. The industry dummies are created based on the broad one-digit NAICS categories that include a contrast of five service sectors, with the finance and insurance sector as the reference sector.

## Estimation

We tested the first nine propositions using the nonlinear regression equation specified in Equation 11. We first explored possible multicollinearity using ordinary least squares (OLS)

Table 3
**Results of Model Estimation**

| Parameter | 2002 ($N$ = 741) Estimate (S.E.) | 2007 ($N$ = 751) Estimate (S.E.) |
|---|---|---|
| $\delta_0$ | .236 (.046)**** | .205 (.041)**** |
| $\delta_1$ | .178 (.025)**** | .085 (.019)**** |
| **OPT covariates** | | |
| $\beta_0$: intercept | –.068 (.185) | .826 (.244)**** |
| $\beta_1$: margin | –.312 (.088)**** | –.203 (.132)* |
| $\beta_2$: price | –.231 (.064)**** | –.607 (.127)**** |
| $\beta_3$: growth | –.290 (.048)**** | –.425 (.095)**** |
| $\beta_4$: HHI | .235 (.062)**** | .468 (.081)**** |
| $\beta_5$: wage rate | .918 (.075)**** | 1.056 (.124)**** |
| $\beta_6$: firm age | –.009 (.076) | .255 (.096)*** |
| **ROA covariates** | | |
| $D_1$: SG&A | –.285 (.046)**** | –.467 (.032)**** |
| $D_2$: firm size | –.170 (.034)**** | –.157 (.031)**** |
| **Industry effects** | | |
| $\eta_1$: wholesale and retail trade and logistics | –.048 (.037) | .113 (.029) **** |
| $\eta_2$: information and technical services | –.276 (.053) **** | –.140 (.032) **** |
| $\eta_3$: education and health care | –.079 (.033) *** | .008 (.030) |
| $\eta_4$: recreation, food, and accommodation | .099 (.052) | .157 (.041)**** |
| **Technology effect** | mean (S.D.) | mean (S.D.) |
| OPT ($t_{2002 \text{ vs. } 2007}$ = –15.27)**** | –.068(1.017) | .826 (1.232) |

* $p < .10$
** $p < .05$.
*** $p < .01$.
**** $p < .001$.

regression analyses with ROA as the dependent variable and with all covariates in the equation as the predictors to calculate VIFs. The VIF statistics (see Table 2) are reasonable, with a majority of them below 3.00, indicating multicollinearity is not a concern. We then adjusted all covariates to their industry means by dividing their mean centering scores by their respective two-digit NAICS industry average.[11] The 1% and 99% outliers of each variable were winsorized to reduce the impact of extreme values. All variables in the equations were standardized to have a mean of zero and a standard deviation of one to ensure direct comparability.

## Results

All but one of our propositions received support for both the 2002 and 2007 data years. The results from the empirical analysis are shown in Table 3, with the findings summarized in Table 4.

**Profit Margin.** Proposition 1 predicts that as margin increases, optimal productivity decreases. This prediction received support for 2002 ($\beta$ = –.311, $p$ < .001) and marginal support for 2007 ($\beta$ = –.203, $p$ < .10).

Table 4

**Summary of Findings for Predicting Optimal Productivity**

| Covariates | Descriptions | Expected Effects | Hypothesis Testing Results 2002 | Hypothesis Testing Results 2007 |
|---|---|---|---|---|
| Profit margin | Net sales/gross sales | Negative | Supported | Supported |
| Price | Selling costs/(1 – profit margin) | Negative | Supported | Supported |
| Growth rate | (Current year's sales – last year's sales)/ last year's sales | Negative | Supported | Supported |
| Market concentration | HHI = the sum of the squared market shares of the firms in the industry | Positive | Supported | Supported |
| Wage rate | Wages/per 1,000 employees | Positive | Supported | Supported |
| Firm age | The years since a firm's financial data becomes available in the data set | Positive | Not supported | Supported |
| Technology level | The predicted optimal productivity difference between 2002 and 2007 | Positive | Supported | Supported |

**Price Level.** Proposition 2 argues for a negative relationship between price and optimal productivity. This argument received consistent support for the two data years ($\beta = -.231$, $p < .001$ for 2002; $\beta = -.607$, $p < .001$ for 2007). A service firm's selling price appears to be a robust predictor of optimal productivity.

**Growth Rate.** Proposition 3 states that as the future is weighted more, optimal productivity decreases. Using the current year's sales growth rate as a proxy for a firm's future weight (high growth rate indicates a firm weights future profits more), we found consistent support for both data years ($\beta = -.290$, $p < .001$ for 2002; $\beta = -.425$, $p < .001$ for 2007).

**Market Concentration.** Proposition 4 predicts a positive relationship between the extent to which service impacts customer retention and optimal productivity. Using the degree of market concentration as the proxy for service driving retention, we expect that in cases where the market concentration is high, service will have less impact on customer retention. As predicted, the market concentration index (HHI) was a significant positive covariate for optimal productivity ($\beta = .235$, $p < .001$ for 2002; $\beta = .468$, $p < .001$ for 2007), indicating that if a firm relies more on service to drive retention, then its optimal productivity decreases.

**Wage Level.** Proposition 5 predicts a positive relationship between wage rate and optimal productivity. We received strong and consistent support for the prediction ($\beta = .918$, $p < .001$ for 2002; $\beta = 1.056$, $p < .001$ for 2007), suggesting that as wage rate goes up, optimal productivity goes up.

**Optimal Productivity over Time.** Proposition 6 says that (under most typical conditions) as automation costs decrease, optimal productivity increases. Proposition 7 says that as the level of technology increases, optimal productivity increases. Based on the fact that technology does not go backwards, we can confidently predict that automation costs decrease and the level of technology increases over time. This implies that the optimal productivity level should increase over time. We carried out a *t*-test on the predicted optimal productivity levels of the two data years as a test of this proposition and found that the predicted optimal productivity in 2007 (mean = .826, S.D. = 1.232) was significantly higher than the predicted optimal productivity in 2002 (mean = −.068, S.D. = 1.154), $t = 15.27$, $p < .001$. This gives indirect support for propositions 6 and 7.

**Firm Age.** Proposition 8 predicts that as initial sales are greater (e.g., higher market share or a more mature market), optimal productivity increases, and the firm should use less labor and more automation. Assuming mature firms having higher initial sales than start-up firms, this prediction received support for 2007 ($\beta = .255$, $p < .01$), but not for 2002.

**Optimal Productivity Level.** The $\delta_1$ parameter is positive and significant ($p < .001$) for both data years, supporting Proposition 9. For a given level of technology, there is an optimal productivity level. Based on the estimation results, we calculate the ratio of a firm's labor productivity to its optimal productivity in the two data years to see whether firms tend to have productivity levels that are too high or too low in the two data years. A ratio greater than one is considered too productive, whereas a ratio less than one is not productive enough. A one-sample *t*-test showed that in 2002, firms tended to be slightly overproductive (mean = 1.017, S.D. = .125), which was significantly higher than 1.0 ($t = 3.73$, $p < .001$). In 2007, firms (with the exception of large firms, as we will see below) tended to be underproductive (mean = .905, S.D. = .173), which was significantly lower than 1.0 ($t = -15.06$, $p < .001$). This result suggests that firms were belt tightening slightly too much, on average, in 2002, when the economy was still strong, but in 2007, when the economy was teetering on the brink of recession, companies (with the exception of large firms) were not belt tightening enough.

A further inspection of the distribution pattern of the productivity ratio leads us to the conclusion that big firms (in terms of annual sales) tend to be consistently too productive, when comparing their actual productivity to their optimal productivity. This finding occurs in both data years. In 2002, the average productivity (productivity divided by optimal productivity) of big firms (mean = 1.086, S.D. = .147) was on average 8.6% too high, significantly higher than optimum ($t = 5.01$, $p < .001$) and significantly higher than that of smaller firms (means = 1.010, S.D. = .120; comparison of means test $1.086_{\text{big firms}}$ vs. $1.010_{\text{small firms}}$, $t = 5.05$, $p < .001$). In 2007, big firms (mean = 1.088, S.D. = .281) were 8.8% too productive, while small firms (mean = .884, S.D. = .143) were on average only 88% as productive as they should be. Large firms were on average significantly too productive ($t = 2.74$, $p < .01$) and significantly more productive than small firms ($1.088_{\text{big firms}}$ vs. $.884_{\text{small firms}}$, $t = 10.41$, $p < .001$).

## Discussion

### Theoretical implications

We provide a rigorous theory of the tradeoff between using labor to provide service and using automation and self-service technologies. Our theory is complete enough and specific enough to result in many important conclusions, most of which we have successfully confirmed through empirical research using data from over 700 firms in each of two data years, five years apart.

Our paper makes several important theoretical contributions. The most important is to move beyond the implicit assumption that service productivity should be maximized. By distinguishing between a technology effect (increasing technology and decreased automation costs over time) and a labor-automation substitution effect, we provide new insight into how service productivity works. In the short or medium term, we show that there exists an optimal level of productivity and that the optimal productivity level increases over time. These findings are borne out by the empirical analysis.

Another key contribution is the insight that service productivity needs to be thought about dynamically, in the context of customer retention and its impact over time, rather than just in the context of static profit maximization. This brings service productivity theory in line with modern theory about the interface between marketing and finance. The purpose of service productivity should be to maximize financial return, viewed as discounted cash flows.

Our work also advances theory on the relationship between customer satisfaction and productivity. We show how decisions about the labor-automation tradeoff affect both customer satisfaction (through the service level) and productivity and how increasing productivity can sometimes increase current profits at the expense of level of service (and hence customer satisfaction), future sales and profits, and customer equity.

## Managerial implications

The theoretical model, supported by the empirical analysis, results in many useful managerial implications. These implications challenge the current implicit focus on maximizing service productivity and instead provide solid guidance as to the conditions under which service productivity should be highest. Some of the most important managerial implications that arise from our theory and empirical analysis are the following:

**Optimal Productivity Level.** For a given level of technology there is an optimal productivity level. Automating too fast—driving productivity to a level that cannot be supported by the current level of technology—boosts short-term profits at the expense of overall financial return, because the resulting decline in service level hurts future sales.

**Large Company Productivity.** Our empirical analysis reveals a systematic tendency for large companies to be too productive. This may result from a short-term viewpoint on the part of top managers, many of whom find their compensation levels tied to short-term financial performance (Edmans et al. 2009; Sethi and Namiki 1987). Our study suggests that large companies should focus more on providing good service, automate more slowly, and put less emphasis on short-term productivity and cost reduction.

**Avoiding Myopia.** Too myopic a viewpoint (attaching too low a weight to future profits) will lead the firm to boost short-term productivity and profitability at the expense of future sales and overall financial return. This finding is especially of concern when executives have a more short-term orientation (because of stock option plans or other short-term incentives) than is healthy for the firm.

**Profit Margin Effect.** As profit margins are higher, optimal productivity decreases, and the firm should use more labor and less automation. The intuition here is that satisfying a customer is worth a lot, because of the value of the customer. High profit margins incentivize the firm to “pull out all the stops” to satisfy the customer by providing better service.

**Price Level Effect.** As the price level is higher, optimal productivity decreases, and the firm should use more labor and less automation. As with the profit margin effect, higher prices mean the customer becomes more valuable, providing incentive for better service.

**Future Effect.** As the future assumes more weight (lower discount factor) optimal productivity decreases, and the firm should use more labor and less automation. If the future is worth more, then a retained customer is worth more, which makes providing better service a sensible strategy.

**Sensitivity to Service Level Effect.** As the impact of the service level on customer retention increases, optimal productivity decreases, and the firm should use more labor and less automation. If the customer base is “captive” due to high switching costs, inertia, or lack of decent alternatives, the firm loses the incentive to keep the customer satisfied.

**Wage Level Effect.** As wages rise, optimal productivity increases, and the firm should use less labor and more automation. Higher wages make labor less attractive.

**Automation Cost Effect.** Under certain typical conditions (wages low enough), as automation costs decrease, optimal productivity increases, and the firm should use less labor and more automation. Lower automation costs generally mean automation becomes more attractive. However there are conditions under which we get a seemingly counterintuitive result. Under some conditions, lower automation costs mean that the firm can "afford" more labor (if wages are high enough), to satisfy the customer better, thus shifting the company toward more labor and decreasing productivity.

**Technology Effect.** As the level of technology increases (technology becomes a more effective substitute for labor) optimal productivity increases, and the firm should use less labor and more automation. Because the level of technology only goes up over time, this also suggests a time effect. As time goes by, the optimal productivity increases, and the firm should make more use of automation.

**Initial Sales Effect.** As initial sales are greater (e.g., higher market share or a more mature market) optimal productivity increases, and the firm should use less labor and more automation. The intuition is that if the firm already has a lot of customers, then making money from them in the current period is more important. If the firm has few customers, but needs to get more, then the service level needs to be emphasized more than current profits, to grow the firm's customer base.

### Managing service productivity in a recession

In a recession, managers should use relatively more labor and relatively less automation, resulting in a lower optimal level of productivity. Interestingly, service companies appear to be doing exactly the opposite in the current recession, as service productivity has grown dramatically as service workforces have been slashed (Cooper 2009). Our analysis suggests that too much labor force reduction may be counterproductive for service firms and may have a deleterious effect on service quality, customer satisfaction, and postrecession sales and profitability.

Taken as a group, these results provide managers with insight into when it is better for service productivity to be higher and when it is better for it to be lower. Armed with these insights, managers will be better able to make strategic decisions about when to substitute automation for labor in service provision. Although it is the case that consistent and predictable trends in level of technology and automation costs suggest the wisdom of increasing automation, self-service, and service productivity over time, it is possible to automate too much, too fast. For any given level of technology, an optimal level of service productivity should be sought—that which will maximize the financial returns of the firm.

## Limitations

As with any study, there are limitations to be kept in mind. Our theory, although formal and rigorous, represents a stylized view of the world. Any theory is merely suggestive of reality and is useful only to the extent that it provides insights. One might build alternative theories based on somewhat different assumptions. One might also complicate the theory by considering such factors as (a) multiple levels of labor at different wage rates, (b) competition (perhaps asymmetric), (c) nonlinear relationships where there are linear ones, or (d) considering the use of automation whose primary purpose is raising the level of service rather than reducing costs. All of the above complications would make the theory more realistic, but also less tractable (perhaps intractable) with perhaps limited ability to produce additional insight.

Although we expect the results of this analysis to apply broadly across the service sector (and in many cases in the goods sector as well), these results may not hold universally. In particular, industries in which the main model assumptions do not hold may prove to be an exception to these findings. Such industries tend to be extreme cases (e.g., very new industries, industries in which it is exceptionally difficult to automate, industries in which technology is so far along that a totally automated service is best, etc.).

We should also consider the limitations of our empirical analysis. Although we analyze a very large sample of service companies (more than 700 companies at each of two different points in time), the sample has its limitations. It is limited to one continent (North America), and it is possible that service productivity might operate differently in other economic environments. We do not have direct measures of some of the constructs in the theoretical model (automation costs and level of technology) and must use time as a proxy for them. The two samples (2002 and 2007) do not include the same companies, so the nature of the samples could be different across the two time periods. (The alternative, including only those companies that survive through the 10 data years, might result in a sample selection bias that could miss out on newer and less successful companies.) Also, although it would create severe difficulties for data collection, a more extensive time range would be desirable.

## Directions for Future Research

As our theory and empirical analysis suggest, the optimal service productivity level will only increase over time, so it is essential that we learn more about how to increase service productivity in a way that increases the firm's financial return. We need to know more about the effect of automation efforts not only on current revenues and costs, but also on future revenues and costs. That is, the effect of automation on the perceived service level is of vital importance. For example, what kinds of self-service technology build or maintain customer satisfaction and customer retention? Which kinds result in damage to those future-oriented measures? How does customer lifetime value relate to receptiveness to automation and self-service? Are there ways to usefully segment the market to provide the right level of automation and self-service to different segments? How can incentive plans for the executives of large companies be designed to curb the short-term viewpoint that leads to excessive cost cutting and inadequate attention to level of service?

It would be useful to have longitudinal studies of firms that have sought to increase service productivity. How quickly can service productivity be increased without damaging customer satisfaction and long-term profits? What kinds of service productivity efforts have the greatest impact, and which ones are implemented the fastest? What are the biggest problems that firms have when trying to increase service productivity?

Our formal theory assumes that the effectiveness of labor in providing service is constant, as is the wage level. Future research could instead assume that there are different levels of labor quality available, at different wage levels. The result would complicate the analytical derivations, but could provide an additional degree of realism.

## Conclusions

We have presented a rigorous theory of service productivity and tested it empirically with data from over 700 service firms for each of two data years. Our theory and empirical analysis produced a number of important conclusions about how service productivity should be managed, the most important being that for a given level of technology, there is an optimal level of service productivity. Either too low a level of productivity or too high a level damages the firm's long-term performance, and hence the value of the firm. Our theory and empirical analysis also provide guidance as to when service productivity should be higher or lower as well as cautionary advice for large firms and firms operating in a recession. These findings have very important implications for the firm's decision regarding how it should provide service to its customers—to what extent the firm should attempt to substitute automation and self-service technology for labor in service provision.

## Appendix: Proofs and Derivations

### Optimal Productivity

From Equation 8 we see that the optimal productivity level arises from the optimal level of $\theta$, the proportion of labor per unit. Thus we need to determine $\theta^*$, the optimal level for $\theta$. From equations 2, 5, and 6 we get:

$$FR = (mR - A - (W - A)\theta)(Q + \beta\alpha Q + \beta(1 - \alpha)\theta Q + \gamma\alpha + \gamma(1 - \alpha)\theta) \quad \text{(A1)}$$

Taking $\partial FR/\partial\theta$ yields the expression in Equation 7. There are three cases. For two of the cases ($\theta^* \leq 0$ or $\theta^* \geq 1$), the optimal strategy is all automation or all labor, respectively. For the more interesting and typical third case ($0 \leq \theta^* \leq 1$), we check the second-order conditions:

$$\partial^2 FR/\partial\theta^2 = -2(W - A)(k\beta(1 - \alpha)Q + k\gamma(1 - \alpha)) < 0, \quad \text{(A2)}$$

which confirms that $\theta^*$ does, in fact, maximize financial return.

*Proposition 1: As margin increases, optimal productivity decreases.*

$$\partial\theta^*/\partial m = R/(2(W - A)) > 0 \quad \text{(A3)}$$

$$\partial OPT/\partial m = (\partial OPT/\partial\theta^*)(\partial\theta^*/\partial m) < 0 \quad \text{(A4)}$$

*Proposition 2: As price increases, optimal productivity decreases.*

$$\partial\theta^*/\partial R = m/(2(W - A)) > 0 \quad \text{(A5)}$$

$$\partial OPT/\partial R = (\partial OPT/\partial\theta^*)(\partial\theta^*/\partial R) < 0 \quad \text{(A6)}$$

*Proposition 3: As the future is weighted more, optimal productivity decreases.*

$$\partial\theta^*/\partial k = Q/(2k^2(1 - \alpha)(\beta Q + \gamma)) > 0 \quad \text{(A7)}$$

$$\partial OPT/\partial k = (\partial OPT/\partial\theta^*)(\partial\theta^*/\partial k) < 0 \quad \text{(A8)}$$

*Proposition 4: As the extent to which service drives retention increases, optimal productivity decreases.*

$$\partial\theta^*/\partial\beta = Q^2/(2k(1 - \alpha)(\beta Q + \gamma)^2) > 0 \quad \text{(A9)}$$

$$\partial OPT/\partial\beta = (\partial OPT/\partial\theta^*)(\partial\theta^*/\partial\beta) < 0 \quad \text{(A10)}$$

*Proposition 5: As wages rise, optimal productivity increases.*

$$\partial\theta^*/\partial W = -(mR - A)/(2(W - A)^2) < 0 \quad \text{(A11)}$$

$$\partial OPT/\partial W = (\partial OPT/\partial\theta^*)(\partial\theta^*/\partial W) > 0 \quad \text{(A12)}$$

*Proposition 6: If wages are low enough, as automation costs decrease, optimal productivity increases.*

$$\partial\theta^*/\partial A = (mR - W)/(2(W - A)^2) > 0 \text{ if } W < mR \quad \text{(A13)}$$

$$\partial OPT/\partial A = (\partial OPT/\partial\theta^*)(\partial\theta^*/\partial A) < 0 \text{ if } W < mR \quad \text{(A14)}$$

*Proposition 7: As the level of technology increases, optimal productivity increases.*

$$\partial\theta^*/\partial\alpha = -Q/(2k(1 - \alpha)^2(\beta Q + \gamma)^2) - 1/(2(1 - \alpha)^2) < 0 \quad \text{(A15)}$$

$$\partial OPT/\partial\alpha = (\partial OPT/\partial\theta^*)(\partial\theta^*/\partial\alpha) > 0 \quad \text{(A16)}$$

*Proposition 8: As initial sales increase, optimal productivity increases.*

$$\partial\theta^*/\partial Q = -1/(2k(1 - \alpha)(\beta S + \gamma)) < 0 \quad \text{(A17)}$$

$$\partial OPT/\partial Q = (\partial OPT/\partial\theta^*)(\partial\theta^*/\partial Q) > 0 \quad \text{(A18)}$$

The theory also enables us to derive the following key results:

*Proposition 9: For a given level of technology, there is an optimal productivity level.* This is seen from equations 7 and A2, which provide the first-order and second-order conditions for optimality.

*Proposition 10: Too myopic a viewpoint (putting not enough weight on future profits) boosts short-term productivity and profitability at the expense of overall financial return.*

Let $\theta_m$ be the myopic solution for optimal labor, given myopic discount factor $k_m < k$. From Proposition 3 we know that $\partial\theta^*/\partial k > 0$, which implies $\theta_m < \theta^*$ and therefore that long-term profits and customer equity are harmed. Further:

$$\partial p_1/\partial\,\theta = -(W - A)Q < 0, \tag{A20}$$

which says that too little labor will increase profits in the current period at the expense of total discounted financial return.

*Proposition 11: In a recession, the firm should make relatively more use of labor (as opposed to automation) than usual, and the optimal level of service productivity will be lower than usual.* Let us suppose that Period 1 is a period of temporarily reduced demand (a recession). This implies a lower level of current period sales, $Q$. From A17 and A18, we see that a lower level of $Q$ calls for a higher labor intensity and lower optimal productivity.

## Notes

1. Many operationalizations of output per input exist, such as sales per employee, sales per labor cost, units per employee hour, and others. Which operationalization one chooses is often a matter of data availability and comparability.

2. We use the term "automation" to refer to the use of technology to increase input efficiency and thus enhance service productivity and the term "technology level" to refer to the relative effectiveness of automation in providing quality service.

3. We thus differentiate between long-term automation effectiveness from an advancing technological level (which can, with proper design, result in provision of a higher service level) and short-term automation effects, which generally substitute imperfectly for labor (e.g., the interminable phone menu systems).

4. Having a fixed price simplifies the exposition. Alternatively, we could specify a demand function and make pricing part of the decision problem, but the added complexity would contribute little additional insight with respect to the productivity issue.

5. For simplicity and tractability, we assume that the effectiveness of labor and the effectiveness of automation (at the current level of technology) are constant.

6. Customer attraction is also generally a function of other variables, such as advertising. Again, including additional predictors would complicate the exposition without adding insight into the productivity issue.

7. All derivations are shown in the appendix.

8. This is true as long as current profits are sufficient to keep the company in business, something our formal theory assumes. If that is not the case, then the company should cut costs (increase productivity) in the current period to whatever extent is necessary to maintain the firm's viability.

9. Thus we distinguish between the covariates of optimal productivity, $X_{ij}$, in Equation 10 and the covariates of financial return, $Y_{cj}$, in Equation 9.

10. The log of sales per employee is monotonically related to sales per employee (the other typical operationalization), which means that the optimal value with respect to either definition will result in the same sales per employee as optimum.

11. HHI did not need to be industrially normalized because the predictor was calculated based on industrial-specific sum and sum of squares.

## References


Aksoy, Lerzan, Bruce Cooil, Christopher Groening, Timothy L. Keiningham, and Atakan Yalçın (2008), "The Long-term Stock Market Valuation of Customer Satisfaction." *Journal of Marketing* 72 (July), 105–22.

Anderson, Eugene W., Claes Fornell, and Donald R. Lehmann (1994), "Customer Satisfaction, Market Share and Profitability: Findings from Sweden." *Journal of Marketing* 58 (July), 53–66.

———, ———, and Sanal Mazvancheryl (2004), "Customer Satisfaction and Shareholder Value." *Journal of Marketing* 68 (October), 172–85.

———, ———, and Roland T. Rust (1997), "Customer Satisfaction, Productivity, and Profitability: Differences between Goods and Services." *Marketing Science* 16 (2), 129–45.

Baily, Martin H., and Robert Z. Lawrence (2001), "Do We Have a New E-conomy?" *American Economic Review* 91 (2), 308–12.

Banker, Rajiv D., Hsihui Chang, and Ram Natarajan (2005), "Productivity Change, Technical Progress, and Relative Efficiency Change in the Public Accounting Industry." *Management Science* 51 (2), 291–304.

Basker, Emek (2007), "The Causes and Consequences of Wal-Mart's Growth." *Journal of Economic Perspectives* 21 (3), 177–98.

Bendapudi, Neeli, and Robert P. Leone (2003), "Psychological Implications of Customer Participation in Co-Production." *Journal of Marketing* 67 (January), 14–28.

Berry, Leonard L., and A. Parasuraman (1993), "Building a New Academic Field: The Case of Services Marketing." *Journal of Retailing* 69 (1), 13–60.

Bertschek, Irene, and Ulrich Kaiser (2004), "Productivity Effects of Organizational Change: Microeconometric Evidence." *Management Science* 50 (3), 394–404.

Blackstone, Brian (2007), "Is Productivity Growth Back in the Grips of Baumol's Disease?" *Wall Street Journal* (Eastern Edition; August 13), A2.

Brown, James R., and Chekitan S. Dev (2000), "Improving Productivity in a Service Business: Evidence from the Hotel Industry." *Journal of Service Research* 2 (4), 339–54.

Brynjolfsson, Erik, and Lorin Hitt (1994), "New Evidence on the Returns of Information Systems." Cambridge, Mass.: Sloan School of Management, Working Paper.

Converse, P. D. (1939), "Labor Saving Devices in Marketing." *Journal of Marketing* 4 (October), 150–6.

Cooper, James B. (2009), "The Mixed Blessing of Soaring Productivity." *BusinessWeek* (February 23), 10.

Danaher, Peter J., and Roland T. Rust (1996), "Indirect Financial Benefits from Service Quality." *Quality Management Journal* 3 (2), 63–75.

Datta, Deepak K., James P. Guthrie, and Patrick M. Wright (2005), "Human Resource Management and Labor Productivity: Does Industry Matter?" *Academy of Management Journal* 48 (1), 135–45.

Desiraju, Ramarao, and Steven M. Shugan (1999), "Strategic Service Pricing and Yield Management." *Journal of Marketing* 63 (January), 44–56.

Dipert, Brian (2009), "Canceling Sirius: So Frustrating, It Made Me Furious." *EDN* (April 23), http://www.edn.com/article/CA6651597.html?text=Sirius+Satellite+Radio.

Edmans, Alex, Xavier Gabaix, Tomasz Sadzik, and Yuliy Sannikov (2009), "Dynamic Incentive Accounts." Cambridge, Mass.: National Bureau of Economic Research, Working Paper.

Fisk, Raymond P., Stephen W. Brown, and Mary Jo Bitner (1993), "Tracking the Evolution of the Services Marketing Literature." *Journal of Retailing* 69 (1), 61–103.

Fornell, Claes, Sunil Mithas, Forrest Morgeson, and M. S. Krishnan (2006), "Customer Satisfaction and Stock Prices: High Returns, Low Risk." *Journal of Marketing* 70 (January), 3–14.

———, Roland T. Rust, and Marnik G. Dekimpe (in press), "The Effect of Customer Satisfaction on Consumer Spending Growth." *Journal of Marketing Research*.

Gruca, Thomas S., and Lopo L. Rego (2005), "Customer Satisfaction, Cash Flow, and Shareholder Value." *Journal of Marketing* 69 (July), 115–30.

Gupta, Sunil, Donald R. Lehmann, and Jennifer A. Stuart (2004), "Valuing Customers." *Journal of Marketing Research* 41 (1), 7–18.

Guthrie, J.P. (2001), "High-Involvement Work Practices, Turnover and Productivity: Evidence from New Zealand." *Academy of Management Journal* 44, 180–90.

Hammer, Michael (2005), book jacket quote on Rayport, Jeffrey F. and Bernard J. Jaworski, *Why Companies Must Improve Their Service Interfaces with Customers.* Boston, Mass.: Harvard Business School Press.

Homburg, Christian, Nicole Koschate, and Wayne D. Hoyer (2005), "Do Satisfied Customers Really Pay More? A Study of the Relationship between Customer Satisfaction and Willingness to Pay." *Journal of Marketing* 69 (April), 84–96.

Huselid, Mark A. (1995), "The Impact of Human Resource Management Practices on Turnover, Productivity, and Corporate Financial Performance." *Academy of Management Journal* 38 (3), 635–72.

Keiningham, Timothy L., Tiffany Perkins-Munn, and Heather Evans (2003), "The Impact of Customer Satisfaction on Share-of-Wallet in a Business-to-Business Environment." *Journal of Service Research* 6 (1), 37–50.

Koch, Marianne J., and Rita Gunther McGrath (1996), "Improving Labor Productivity: Human Resource Management Policies Do Matter." *Strategic Management Journal* 17 (5), 335–54.

Kopalle, Praveen, and Donald R. Lehmann (2006), "Setting Quality Expectations When Entering a Market: What Should the Promise Be?" *Marketing Science* 25 (1), 8–24.

Kumar, V., and Denish Shah (2009), "Expanding the Role of Marketing: From Customer Equity to Market Capitalization." *Journal of Marketing* 73 (November), 119–3.

Lovelock, Christopher, and Evert Gummeson (2004), "Whither Services Marketing? In Search of a New Paradigm and Fresh Perspectives." *Journal of Service Research* 7 (1), 20–41.

Maglio, Paul, and Jim Spohrer (2008), "Fundamentals of Service Science." *Journal of the Academy of Marketing Science* 36 (1), 18–20.

Meuter, Matthew L., Mary Jo Bitner, Amy L. Ostrom, and Stephen W. Brown (2005), "Choosing Among Alternative Service Delivery Modes: An Investigation of Customer Trial of Self-service Technologies." *Journal of Marketing* 69 (April), 61–83.

———, Amy L. Ostrom, Robert I. Roundtree, and Mary Jo Bitner (2000), "Self-service Technologies: Understanding Customer Satisfaction with Technology-based Service Encounters." *Journal of Marketing* 64 (July), 50–64.

Mittal, Vikas, Eugene W. Anderson, Akin Sayrak, and Pandu Tadikamalla (2005), "Dual Emphasis and the Long-term Financial Impact of Customer Satisfaction." *Marketing Science* 24 (4), 544–58.

Narver, John C., and Stanley F. Slater (1990), "The Effect of a Market Orientation on Business Profitability." *Journal of Marketing* 54 (October), 20–35.

Noble, Charles H., Rajiv K. Sinha, and Ajith Kumar (2002), "Market Orientation and Alternative Strategic Orientations: A Longitudinal Assessment of Performance Implications." *Journal of Marketing* 66 (October), 25–39.

Oliva, Rogelio, and John D. Sterman (2001), "Cutting Corners and Working Overtime: Quality Erosion in the Service Industry." *Management Science* 47 (7), 894–914.

Parasuraman, A. (2002), "Service Quality and Productivity: A Synergistic Perspective." *Managing Service Quality* 12 (1), 6–9.

———, Katherine N. Lemon, and Valarie A. Zeithaml (2004), "Return on Marketing: Using Customer Equity to Focus Marketing Strategy." *Journal of Marketing* 68 (January), 109–127.

———, and Anthony J. Zahorik (1993), "Customer Satisfaction, Customer Retention, and Market Share." *Journal of Retailing* 69 (2), 193–215.

Schmalensee, Richard (1977), "Using the H-Index of Concentration with Published Data." *Review of Economics and Statistics* 59 (May), 186–93.

Sethi, S. Prakash, and Nobuaki Namiki (1987), "Top Management Compensation and Corporate Performance." *Journal of Business Strategy* 7 (4), 37–43.

Shostack, Lynn G. (1977), "Breaking Free from Product Marketing." *Journal of Marketing* 41 (April), 73–80.

Spohrer, Jim, and Paul Maglio (2008), "The Emergence of Service Science: Toward Systematic Service Innovations to Accelerate Co-Creation of Value." *Production & Operations Management* 17 (3), 238–46.

Srivastava, Rajendra K., Tassaduq A. Shervani, and Liam Fahey (1998), "Market-Based Assets and Shareholder Value: A Framework for Analysis." *Journal of Marketing* 62 (January), 2–18.

Terra Cin (2009), "HSBC Online Banking Sucks!" (July 23), available at http://terra-cin.blogspot.com/2009/07/hsbc-online-banking-sucks.html.

Tuli, Kapil R., Ajay K. Kohli, and Sundar G. Bharadwaj (2007), "Rethinking Customer Solutions: From Product Bundles to Relational Processes." *Journal of Marketing* 71 (July), 1–17.

Vargo, Stephen L., and Robert F. Lusch (2004), "Evolving to a New Dominant Logic for Marketing," *Journal of Marketing* 68 (January), 1–17.

Yelp.com (2007), "Three Reviews for United Airlines," second review (August 29), available at http://www.yelp.com/biz/united-airlines-santa-ana.

**Report No. 09-120**