Information Systems Research
Vol. 23, No. 3, Part 1 of 2, September 2012, pp. 721–737
ISSN 1047-7047 (print) | ISSN 1526-5536 (online)

http://dx.doi.org/10.1287/isre.1110.0390

Does the Web Reduce Customer Service Cost? Empirical Evidence from a Call Center

Anuj Kumar
Warrington College of Business Administration, University of Florida, Gainesville, Florida 32611,
anuj.kumar@warrington.ufl.edu

Rahul Telang
School of Information Systems and Management, Heinz College, Carnegie Mellon University,
Pittsburgh, Pennsylvania 15213, rtelang@andrew.cmu.edu

Firms are investing millions to deploy Web-based self-services at their call centers. The rationale for such investment is that the firm's cost of interacting with its customers through the Web-based channel is an order of magnitude cheaper than the assisted channels such as telephony. We conduct a field study at the call center of a prominent U.S. health insurance firm to examine this cost-saving rationale of the Web-based self-service channel. On the one hand, the Web channel may substitute for the telephony channel in some cases. On the other hand, the Web also exposes customers to a vast amount of information about their health policy, claims, and coverage; this information can create uncertainty leading to customers seeking more information and hence making more telephone calls. We designed a quasi-natural experiment in our field setting and used difference-in-difference specifications to show that the Web-based self-service usage leads to a 14% increase in telephone calls. We conduct several robustness checks to show that our specifications are robust to any potential selection of customers in the Web-based self-service usage. We further find that the impact of Web portal usage is moderated by the Web portal characteristics. We find that if the information is unambiguous and easily retrievable on the Web, calls for such information decline by 29%. However, for ambiguous information, the calls increase substantially. Our research provides insights into the challenges and opportunities of self-service technologies design.

Key words: self-service; call center; customer support; Web portal; multichannel service management; health insurance
History: Anitesh Barua, Senior Editor; Ramnath Chellappa, Associate Editor. This paper was received on December 4, 2009, and was with the authors 7 months for 2 revisions. Published online in *Articles in Advance* December 7, 2011.

1. Introduction

Call centers and their contemporary contact centers have been firms' preferred media for communication with their customers. Industry estimates suggest that 70% of all customer–business interaction takes place via call centers; Fortune 500 companies, on average, operate 30 call centers each (Mandelbaum 2006). Of AT&T's total telephone traffic, 40% was 1-800 calls directed toward various call centers (Mandelbaum 2006). Firms have increasingly recognized that the call centers offer an opportunity to collect rich customer interaction data that can be analyzed to provide customized goods, services, and experiences to the customers, leading to higher customer satisfaction and loyalty. Growth in newer and cheaper technologies and a desire to improve efficiency have led to firms introducing several alternative channels of customer support at the call centers over time. On the one hand, there are assisted channels where firm representatives assist customers via telephone, e-mail, or short message service (SMS). On the other hand, there are self-service channels (popularly referred to as self-services technology, or SST, channels) such as the interactive voice response units (IVRs/VRUs), the Web-based self-service portals, and the ATM (automated teller machine). The SSTs provide a wide array of choices for customers to contact the firm.

Because a large proportion of the total call center cost is direct labor cost, firms are increasingly using SSTs such as IVRs and Web-based self-services to cut costly customer service representative (CSR) costs. A Yankee Group research report (2006) estimates that the average cost of serving a customer via self-service channels (less than \$1) is an order of magnitude cheaper than serving her via assisted channels (between \$5 and \$10). Within self-service channels, the Web-based self-service costs just \$0.24 for an interaction as opposed to \$5.50 for a CSR-assisted interaction via telephony. Because of this compelling economics, the Web-based self-service has grown from 7.7% of the total customer interaction in 2005 to 14.4% in 2007 (Yankee Group 2006). The

market for the Web-based self-service solutions is expected to grow from US$309.3 million in 2005 to US$887.3 million in 2012 at a compound annual growth rate of 16.2% (Frost & Sullivan 2006). The Web-based self-services allow customers to not only execute transactions at their convenience without any physical interface with the firm but also seek relevant information and manage their online accounts independently. It is expected that once customers learn and start availing themselves of the convenience of this option, the demand for the assisted channels would go down. SSTs should also result in higher customer satisfaction and profitability (Anderson and fornell 1994, Xue et al. 2007). However, the benefits of SSTs can be realized only when customers embrace and use this technology.

Prior research has explored the determinants for customer adoption of SSTs. Bitner et al. (2000) and Meuter et al. (2000) find that the ability of SSTs to bail a customer out of an immediate troubling situation, ease of access and use, and fascination with the capability of technology are the main sources of customer satisfaction with SSTs. Meuter et al. (2005) show that the customer readiness variables such as role clarity, motivation, and ability are key mediators in increasing the customer likelihood of trying SSTs. Bobbitt and Dabholkar (2001) use attitudinal theories to propose a conceptual framework for predicting a customer's decision to use SSTs. However, these studies rely on questionnaire or survey tools to elicit the customer adoption and use of SSTs, and they do not inform how the adoption of SSTs affects the demand for other available alternative channels. The interaction among different channels of sale has been studied by marketing scholars. Deleersnyder et al. (2002) study 67 newspapers and find that the Internet format of newspapers cannibalizes sales of the paper format if their contents are similar. However, sales may be enhanced if the contents in the two formats are different. Biyalogorsky and Naik (2003) study how the addition of an Internet channel to the existing brick-and-mortar store channel for a music records firm affects the sales, and they find a negative but insignificant impact.

There is a vast literature on issues surrounding the management of call centers (Gans et al. 2003, Aksin et al. 2007). Most of this research is devoted to modeling query arrival and service time, and thus they provide guidance on forecasting, capacity planning, and personnel scheduling at call centers. Only recently have some academic papers started looking at how customers' demand at call centers is affected by different firms' strategies. For instance, Kumar and Telang (2010) have examined how product customization affects customers' calls at a health insurance firm's call center. In the present paper, we study how customers' demand for the other channels at call centers is affected by their usage of a Web-based self-service channel—a question not empirically studied in this literature.

Two studies are directly relevant to our work. Xue et al. (2007) study the adoption and usage of various service channels (teller, ATM, online banking) offered by a large retail bank and show how demographic characteristics affect channel use. They also compute various measures of consumer efficiency using the channel use data and show that the efficiency is related to the firm profitability. Campbell and Frei (2010) conduct a field study on the impact of online banking channel adoption on local branches, IVR, ATMs, and call centers. They show that the users who adopted the online banking channel reduced their dependence on the IVR and ATM channels (*substitution*) but increased their consumption of the firm's call center and local branches (*augmentation*). They use a difference-in-difference method for identification, which is similar to our approach. Their augmentation result is also similar to our finding that the use of the Web portal is associated with an increase in the call center calls. However, our settings are quite different. Unlike online banking, in our setting, the Web portals are mostly used for information search rather than conducting transactions. Moreover, in this paper we provide a clear mechanism through which the Web portal usage leads to more calls, and we provide a clean identification of the aggravation effect (i.e., using Web portals encourages more calls to a call center).

Our work is also directly related to broader studies that examine the role of information technology (IT) and productivity. However, much of this literature uses aggregate data at the firm, industry, or even country level (Siegel 1997, Brynjolfsson and Hitt 1995, Barua et al. 1995). To measure the impact of IT on productivity at a more granular level, the IT researchers have focused on intermediate output measures such as inventory, quality, etc. A few recent papers have explored the IT productivity issue at an individual level. Aral et al. (2011) find a positive effect of IT use and information flow on individual white-collar worker performance. Bulkley and Van Alstyne (2006) find a positive effect of information flow and network structure on white-collar worker performance. We also have customer-level data, and we examine the effects of IT at an individual level. In particular, we examine how individuals use a widely available IT-enabled self-service tool like the Web portal and how it affects the costs at the call center. However, unlike the productivity studies, we will show that the adoption of an IT tool may have an adverse effect on firms' intended productivity goals.

In the present work, we conduct a field study at a major U.S. health insurance firm with about four million customers. The firm offers customer support via telephony through its call center. Around 2004 to 2005, the firm introduced a Web-based self-service portal for its customers. We collect a random sample of 60,000 customers from 2005 to 2007. About 10,000 customers adopted the Web portal. Using the vertical differentiation framework in economics, we show that the Web portal usage could potentially substitute for the telephone calls, especially for simple information search. We also use human decision-making theories to highlight that the portal usage may expose customers to more information that may aggravate their concerns, leading them to make more telephone calls. This is particularly true when the information search is less directed, and the information is uncertain and complex. We set up a quasi-natural experiment and employ difference-in-difference specifications to estimate the net impact of Web portal usage on telephone calls. Our results show that the portal registration leads to a 14% increase in the number of calls to the call center. Because the portal registration is not completely random, we conduct several robustness checks to show that our specification is robust to any potential selection of customers in the portal registration. We also show that the results are moderated by the portal and customer characteristics. We find that the portal usage leads to a 29% decrease in calls where the related information is unambiguously provided on the portal. In contrast, the portal usage leads to a 66% increase in calls where the information is ambiguous and comprehensive with no interactive features.

We believe our research is unique in many ways. First, there are few studies that have examined the effect self-service channels adoption on a firm call center costs. Second, we estimate the impact of Web-based self-service usage on the actual transactional data in a field study, as opposed to survey data. Our data are detailed enough to draw some interesting conclusions not explored in the prior literature. Third, we conduct various robustness checks to ascertain that the difference-in-difference specification we employ is robust to any possible selection bias. Therefore, our data allow us to draw causal inferences. Last, but not least, we provide empirical evidence of the value of the portal design in affecting calls.

The rest of this paper is organized as follows. We describe our study setting in §2. Section 3 outlines our theoretical framework. We describe our data and econometric specifications in §4. We also discuss various threats of identification in this section. Finally, in §5, we conclude with managerial implication of our research, outline limitations of the current work, and discuss some future research directions.

2. Research Site

Our study setting is a large health insurance firm in the United States. The firm sells several different health insurance plans to a population of approximately four million customers. After the plans are sold, the firm serves its customers through its operational unit. The operational unit performs three broad services:

1. Client services—Setting up accounts and performing routine periodic activities such as issuing invoices, etc.,
2. Call center services—Resolving customer queries, and
3. Claims services—Processing customer claims both through computer systems and manually.

Activities 1 and 3 are predominantly automated through the information systems set up in the firm. The call center services accounts for approximately 70% of the total operating cost of $47 million in 2007. The firm received over three million calls in 2007.

2.1. Web Portal

To contain large call center costs, the firm started a customer Web portal in 2004–2005. The portal provided detailed information regarding health policy coverage, claims, billing, provider network, and so on, to the customers. Customers first register on the Web portal by creating a secure user name and password. Thereafter, customers can access information on the Web portal. Because customers can find detailed information on the portal, the firm hoped that it would reduce the customers' need to call the customer service with their queries.

The Web portal provides information in five broad areas. First, it provides information on plan coverage and membership details. Second, it allows the users to track their health-care costs, monitor their claim status, get useful information about the costs of health services, and manage their spending accounts. Third, it allows the customers to access their personal health records, explore treatment options, and get support in health-care decisions and preventive recommendations. Fourth, it allows customers to access various health providers' information such as location, profile, credentials, and quality performance data. This helps the customers make more informed choices of providers. Finally, the portal provides customers with access to a health encyclopedia for information on diseases, care management, surgeries, and procedures. Thus, the Web portal hosts comprehensive information on customers' plan benefits, spending, and health records. It also hosts a large amount of information on various health providers, general disease, and health-care management. The customer has to sift through this large amount of information to search for the relevant information and then process it to find the desired answer to her query. This is very different

from situations where a customer can perform a simple transaction (make an account transfer, pay a bill, etc.) through Web portals by clicking on specific links.

3. Theoretical Framework

Customers use available channels of inquiry to resolve queries pertaining to their health and health insurance plans. In our setting, before the Web portal, the telephony was the only choice available to customers for query resolution. After the introduction of the portal, for a given query, the customer has a choice to either search and process this information on the portal or call a CSR for information and resolve her query.

The Web portal can be considered a low-quality and low-cost channel in line with the vertical differentiation models. The cost to access the Web (after choosing to register) is relatively low, but the customer has a lower probability of finding answers to her questions, especially if the questions pertain to information search that is complex, uncertain, and imprecise. In contrast, telephony is a high-quality and high-cost channel. The wait times in a call center are typically large, and the call center is not open 24/7. However, customers can find precise answers to their queries from the CSRs, who have superior expertise and resources. Literature on the vertical product differentiation then provides a roadmap for how the demand for both channels realizes (Gabszewicz and Thisse 1979, Telang et al. 2004).[1] In the following, we provide a simple model of how the Web portal can affect telephony use. The goal is not a structural model to be estimated but a model that provides insights and testable hypotheses.

As a rational economic agent, the customer will consider the costs and benefits of choosing a particular channel. Let V be the economic value of query resolution to a customer. The value of query resolution depends on the severity of the health event.[2] The higher the severity, the higher the value of its resolution. Let P_c and P_w be the expected probabilities of query resolution through telephone calls and portal visits, respectively. Given that the portal is a lower-quality alternative to telephony, we expect $P_c > P_w$. Let C_c be the cost (monetized disutility) of a telephone call, and let C_w be the cost of a portal visit. Per our assumption, we expect $C_c > C_w$. Thus the expected utility from the telephony is $U_c = P_c \cdot V - C_c$ and from the Web is $U_w = P_w \cdot V - C_w$.

[1] The Web and telephony need not be perfect substitutes. Use of telephony need not completely eliminate the need for use of the Web. For example, the Web may sometimes provide some complementary information. As long as, on average, users perceive telephony to be of higher quality than the Web, our model insights will go through.

[2] We will elaborate on this in more detail in §4.

Before the Web portal, customers had only the call center to resolve their query. Normalizing the expected utility from doing nothing, U_o to be zero, customers will choose to call if $U_c > U_0$ or if $V^\wedge \geq (C_c/P_c)$. Thus $V^\wedge$ is a threshold value of the query below which the customer does not bother to make a call.

After the introduction of the portal, the customer has another channel available for query resolution. As we argued, the Web is a lower-quality and lower-cost alternative. Therefore, for a given V, the customer has a decision to make: she can choose the telephone, choose the Web portal, or do nothing. Given a clear order for the telephone and the portal, customers will choose the telephone if $U_c > U_w$ for a given V. Thus, the indifference V^* that separates the phone versus the Web portal use is given by

$$U_c(V^*) = U_w(V^*) \quad \text{or} \quad V^* = (C_c - C_w)/(P_c - P_w).$$

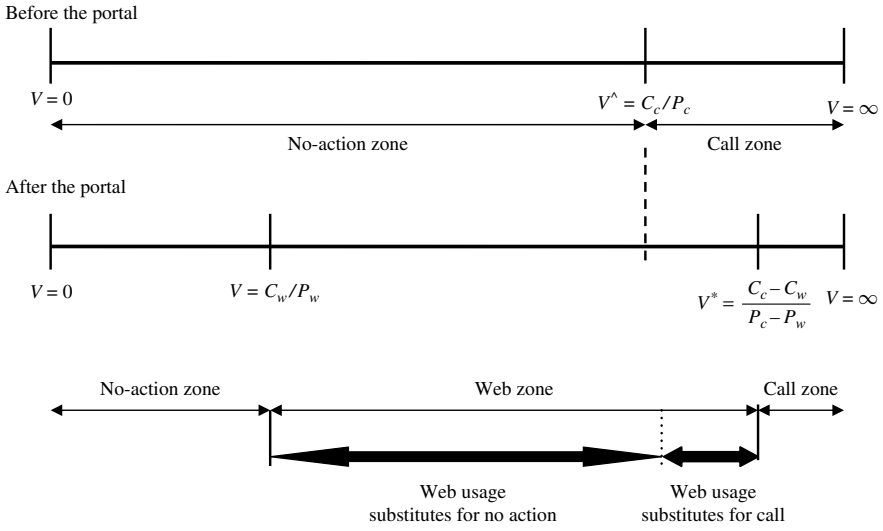
Similarly, the customer will choose the portal over doing nothing if $U_w(V) > U_0$ and $V < V^*$. Thus customers choose the portal when $C_w/P_w < V < V^*$. We can outline these regions pictorially in Figure 1.

Note that it is immediate that $V^* > V^\wedge$. Therefore, after the portal becomes available, the customers use the portal instead of the telephone when $V^\wedge < V < V^*$. This is evident from Figure 1. Therefore, if the portal were to be a substitute for the telephone (the amount of substitution depends on how effective the portal is relative to the telephone after controlling for their costs), we would expect a reduction (or at most no change) in phone calls.

In Figure 1, notice the region from V to $V^\wedge$ where the customers do not use the telephone for query resolution (because of the low value of the query and/or the high cost of telephony). However, the low cost of the portal can facilitate the portal visit for the same query. In other words, the users are willing to use the portal for a query resolution when they chose to do nothing before the portal registration. This is similar to the increase in consumption of online banking services observed by Hitt and Frei (2002). We will show that our data match this aspect quite well. This suggests that the portal provides an opportunity for the users to visit the portal and gather information about their claims, health status, and so on, for even the low- and medium-severity events.

However, this stylized model relies on the fact that the portal is at least a weak substitute. However, in our setting, the Web portal provides large and passive information to its customers. It does not actively provide any tool for a particular query resolution. If the customer searches for unstructured and uncertain information, then the portal may not only fail to resolve the query but may also bring additional information to the customer. There is a large

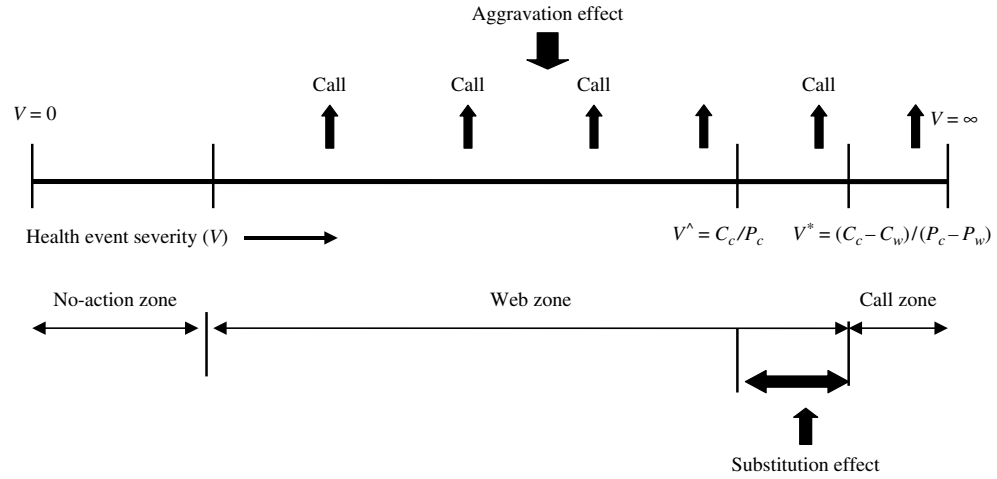
Figure 1 Effect of the Portal



amount of literature in human decision making that suggests that the provision of more information (if it creates ambiguity) may lead to more information-seeking behavior (Cox 1967, Murray 1991). Dermer (1973) found in a laboratory experiment that subjects intolerant to ambiguity prefer a greater amount of information and more readily interpretable signals. Similarly, Miller (1972) found in his experimental study that financial investors seek more data when faced with uncertainty. Some industry studies have also highlighted similar phenomenon where visits to a portal can lead to customer frustration and more calls (Yankee Group 2006).

Notice that in Figure 1 the customers are more likely to visit the Web for moderate economic value query resolution, or they may visit the Web with no particular query in mind. This is also different from the telephony channel, where the customers almost always have a specific query to resolve. Put another way, in the absence of the portal, customers would have done nothing in such cases. Recall that the portal contains comprehensive information on product benefits, health spending, health records, and health providers. The customer gets to see all information on a desired topic; some of it is relevant and some irrelevant, depending on the context. It is conceivable that processing this information sometimes raises additional questions (for example, a customer, who has been prescribed a spinal manipulation, finds four different types of spinal manipulations and associated coverage options listed in her health plan benefits on the portal). Because the customer has the alternative telephony channel available in the present setup, she can possibly call the CSRs and mitigate her uncertainty. In our setup, this suggests that a visit to the Web portal may lead to an *aggravation* effect, necessitating a phone call. This is explained in Figure 2.

Figure 2 Combined Model for Call Generation with the Portal Usage



Thus the impact of the portal on telephony will depend on the strength of the substitution or the aggravation effect. In the following, we empirically estimate the relationship between the portal and telephone usage, and we highlight when the aggravation effect or the substitution effect is more dominant.

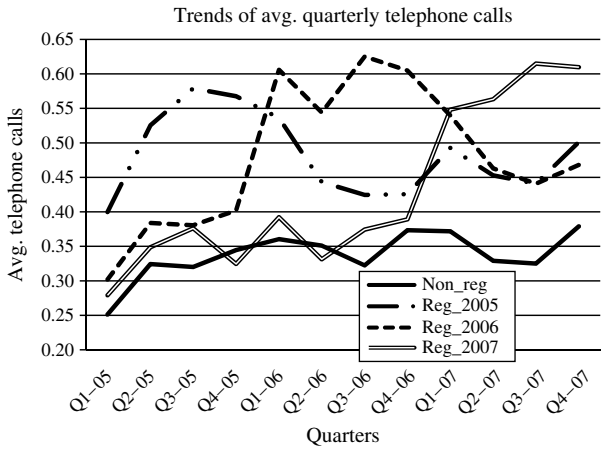
4. Data and Econometric Specifications

We collected the telephone usage data, claims data, and the Web portal usage data for the period March 2005 through December 2007 for a random sample of 60,000 customers from the entire customer population of the firm. We also collected some demographic data such as age and gender for these customers. The firm started offering Web portal services to its customers in 2004. After data cleaning, we get a useable sample of 59,280 customers. Of these, 48,631 customers did not register on the Web portal at all, and the remaining 10,649 customers registered on the Web portal up to December 2007 (5,308 customers registered till December 2005, 2,715 customers registered in the year 2006, and the remaining 2,626 customers registered in 2007). Thus our data offer a quasi-natural experiment setting where we can compare the customer calls before and after the portal registration.

We collected call data for each customer in our sample from the automatic call distributor (ACD) of the call center. The firm classifies the customer calls based on the reason for the calls such as claims information, plan/drug/network coverage information, Web-related technical and usage information, membership/ID card-related information, provider information, and so on. Because the portal use itself may prompt calls such as password reset calls, technical queries, Web search-related queries, and so on, it is important to distinguish and exclude such calls. Because the firm classifies each call, we exclude all calls pertaining to issues surrounding Web portal usage (hereafter, calls or telephone calls means that the Web-related calls are excluded).

After listening to over 100 live calls and holding extensive discussions with the CSRs and the call center managers,[3] we concluded that a health event and associated severity of that health event are major drivers of customer calls. We use three separate measures as a proxy for the health severity. We use the number of claims for a customer to indicate the number of her health events. We use the amount paid (negotiated charges) by the insurance firm to the health provider to capture the intensity of health procedures.[4] We also compute the total dollar amount due from the customer in her claim to indicate customer liability. Note that these three variables put together provide good control for the customer calls in different health situations. We compute these three variables for each customer in each period and use them as control variables in our analysis.

Figure 3 Trends in Telephone Calls



We plot the average quarterly telephone calls for the non-Web-registered customers and the registered customers for 2005, 2006, and 2007 in Figure 3. Figure 3 suggests that (1) the call trend is similar for the Web-registered customers and the non-Web-registered customers in the pre-Web registration period, (2) the calls increase substantially for the Web portal-registered customers during the registration period, and (3) telephone calls stabilize and stay at a higher level for the Web portal-registered customers compared with the non-Web-registered customers after the registration year. The figure also indicates that prior to registration, registered customer calls were a bit higher (for the 2006 cohort) or quite similar (for the 2007 cohort) to the nonregistered customers, but after the Web portal registration, their calls increased significantly.

Although we have data for the customers registered on the Web portal in 2005 and in 2007, for this study, we focus on only those customers who registered in 2006. A simple rationale is, for this cohort, we have 10 months of data before the portal registration (2005, or pre-period) and 10 months of data after the registration (2007, or post-period). For other cohorts (2005 or 2007 registration), we have unbalanced data. Focusing on the 2006 registered cohort allows us to have a clean

[3] One of the authors accomplished these during his three-month internship with the firm.

[4] We used negotiated charges that the insurance firm pays to the health provider rather than the billed provider charges (the amount the provider bills to the customer) as a proxy for health event intensity because different quality providers will bill different charges for the same procedure, but the insurance firm will negotiate similar charges from such diverse providers for similar procedures.

Table 1 Summary Statistics

		2005		2007	
Variables	Observations	Mean	Std. dev.	Mean	Std. dev.
Non-Web portal-registered customers					
No of telephone calls	48,631	1.10	2.02	1.13	2.15
Customer liability ($)	48,631	1,569.65	10,797.57	1,276.75	9,470.65
Negotiated charges ($)	48,631	5,333.64	15,765.86	6,010.51	17,281.91
Number of claims	48,631	21.76	24.56	24.59	27.25
Number of Web visits	48,631	0	0	0	0
Age	48,631	56.59	17.48	—	—
Proportion of females	48,631	0.42	0.49	—	—
2006 Web portal-registered customers					
No of telephone calls	2,715	1.32	2.17	1.55	2.48
Customer liability ($)	2,715	864.95	4,807.14	991.57	3,738.43
Negotiated charges ($)	2,715	5,220.99	12,156.81	6,837.23	20,108.34
Number of claims	2,715	24.96	24.99	28.51	26.95
Number of Web visits	2,715	0	0	1.66	6.63
Age	2,715	48.11	12.54	—	—
Proportion of females	2,715	0.45	0.49	—	—
2006 Web portal-registered customers who used Web portal in 2007					
No of telephone calls	838	1.29	2.11	1.99	3.01
Customer liability ($)	838	687.46	2,795.86	1,238.40	4,933.09
Negotiated charges ($)	838	4,812.50	10,414.69	7,260.25	12,733.68
Number of claims	838	24.54	23.50	30.92	27.59
Number of Web visits	838	0	0	5.39	11.06
Age	838	47.12	11.74	—	—
Proportion of females	838	0.49	0.50	—	—

identification as we can control for the same calendar months for the whole sample. Moreover, with this setup, we can drop the 2006 year data altogether to avoid the transient call increase/decrease during the year of registration for the 2006 cohort. In subsequent sections, we show that we get similar results when we use all registered customers as well. Along with the users who registered for the Web, we also have a control group of non-Web-registered customers who did not register in the entire period.

The summary statistics of all variables discussed above for the 2006 Web portal-registered customers and non-Web portal-registered customers is given in Table 1. Not all the Web portal-registered customers actually used the portal in 2007. We provide the summary statistics for the active portal users separately from all registered customers in the table.

From Table 1, we see a higher increase in calls (1.32 to 1.55) for the 2006 registered customers compared with the non-Web-registered customers (1.10 to 1.13). This suggests that customers make more calls after Web registration. However, we also see differences in non-Web-registered customers and Web-registered customers before Web registration in their call intensity (1.10 versus 1.32) and their liability (1,599 versus 864). From Figure 1, it is clear that the call trend for the 2006 cohort in the pre-period (2005) matches quite well with the nonregistered users. For our identification, it is not necessary that the registered and nonregistered users have the same call intensity. What is needed is that they have the same calling trends. We will discuss the potential threats to identification in more detail in subsequent sections.

Because we have a treatment group (portal-registered customers) and a control group (nonregistered customers), we can employ the difference-in-difference specification, widely used in the literature with natural and quasi-natural experimental settings. The setting is pictorially shown in Figure 4.

This is a preferred specification in these settings as it weeds out the confounding effects of unobserved factors from treatment effects (Meyer 1995, Angrist and Krueger 1999, Kumar and Telang 2010, Danaher et al. 2010). The specification requires that the control and treatment groups exhibit similar trends before the treatment. We utilize the fixed effect difference-in-difference specification (A) to examine how Web

Figure 4 Quasi-Natural Experimental Setup

	Pre-portal period 2005	Post-portal period 2007
2006 Web-reg. customers (Treatment group)	• Calls • Control variables	• Calls • Control variables
Non-Web-reg. customers (Control group)	• Calls • Control variables	• Calls • Control variables

portal usage impacts customers' telephone calls. We control for other observable variables that influence telephone calls: the number and severity of health events and the financial liability of the customer:

$$tel_{it} = \beta_i + \beta_1 X_{it} + \beta_2 reg_i + \beta_3 post_t + \beta_4 reg.post_{it} + \varepsilon_{it} \text{ (Specification (A))},^5$$

where i indexes the customers and t indexes the period (2005 and 2007),

tel_{it} = Number of telephone calls by customer i in period t,
reg_i = 1 if customer i is registered on the Web portal and 0 otherwise,
$post_t$ = 1 for 2007 and 0 otherwise,
X_{it} = Control variables (number of claims, negotiated provider charges in the claims, and total customer liability in the claims) for customer i in period t, and
β_i = fixed effect for customer i.

The coefficient estimate (β_4) for *reg.post* is of interest. This estimate captures the increase or decrease in the telephone calls for the registered customers relative to the nonregistered customers in the post-period. The fixed effect specification controls for the timeinvariant customers' unobserved factors that may influence telephone calls. We scaled negotiated provider charges values in tens of thousands of dollars and customer liability values in thousands of dollars to make the estimate β_1 more readable. We found similar results when taking the logarithm of these two variables. Notice that we use a dummy variable (*reg*) to indicate the portal registration. It may be tempting to use the number of actual Web portal visits in the post-period, but that measure may likely suffer from reverse causality because sometimes the CSRs encourage the customers to visit the Web after the phone call. The results of specification (A) with robust standard errors are reported in Table 2.

First, note that the coefficient estimates for the control variables have the expected signs. Higher health severity (liability, frequency of claims, etc.) leads to more phone calls. We find a positive and significant estimate for the interaction term (*reg.post*), indicating that even after controlling for other observable factors, and the unobserved customer fixed effects, customers make more telephone calls after the Web portal registration. The magnitude of the coefficient estimate for *reg.post* suggests that after the registration, the telephone calls, on average, increase by 0.18,

[5] We can also deploy a Poisson regression method given the count data. The results are unchanged. Because we plan to deploy a two-stage least squares method in the appendix, we present the results with the ordinary least squares (OLS) specification.

Table 2 Estimates for Specification (A)

Dependent variable: *Telephone calls*	Coefficient estimates	
post dummy	−0.021*	(0.011)
reg.post dummy	**0.180*** **	**(0.005)**
Customer liability (in \$1,000)	0.006***	(0.001)
Negotiated charges (in \$10,000)	0.021**	(0.012)
Number of claims	0.018***	(0.001)
Constant	0.701***	(0.012)
N	102,692	
R^2 (adj. R^2)	0.69	(0.39)

Notes. Standard errors are in parentheses. Bold values indicate main variable of interest.
***Statistically significant at the 1% level; **statistically significant at the 5%; *statistically significant at the 10% level (two-sided test).

which translates into a 14% increase in the telephone calls for the Web-registered customers. This is a significant increase, both statistically and economically. This also indicates that, in our sample, the Web portal seems to have a stronger aggravating effect as opposed to the substitution effect. We will discuss the identification of the aggravation effect in more detail in §4.4. Before that, we provide some robustness checks on our findings.

4.1. Threats to Identification

Given that the Web portal assignment is not random, there are various alternative explanations for our results. In this section, we will rule out these explanations.

An important consideration is why certain users choose to register on the portal. Usually, the CSRs encourage the callers to consider registering on the portal. In that regard, the users who call more often are more likely to know about the portal and hence have a higher probability of registering on the portal. This is consistent with the summary statistics. However, this alone is not a major threat to the identification. As we noted earlier, the difference-in-difference specification simply requires that the trends in calls for the treatment and control groups match before the treatment. Our data seem to satisfy this requirement, as seen in Figure 3. We now formally show that the call patterns of the control and treatment users before the portal registration are similar. We test this assumption by comparing the trends in telephone calls for the treatment and control groups over multiple pretreatment periods (Meyer 1995). We run specification (A) for the treatment and control groups in two pretreatment periods, March–July 2005 and August–December 2005. A significant coefficient estimate for the variable *reg.post* in such a specification would indicate a differential trend in calls for the treatment and control groups. The results are reported in column A of Table 3. We find an insignificant coefficient estimate for the variable *reg.post*, which indicates

Table 3 Estimates for Trends of Calls in Preregistration Period

Dependent variable: *Telephone calls*	A 2006 reg. customers as treatment group		B 2007 reg. customers as treatment group	
reg.post dummy	**0.010**	**(0.031)**	**−0.022**	**(0.053)**
Customer liability (in $1,000)	−0.004	(0.004)	0.003	(0.004)
Negotiated charges (in $10,000)	0.007	(0.023)	0.009	(0.022)
Number of claims	0.015***	(0.001)	0.018***	(0.001)
Constant	0.514***	(0.117)	0.417***	(0.193)
N	102,696		102,518	
R^2 (adj. R^2)	0.452	(0.343)	0.608	(0.361)

Notes. Standard errors are in parentheses. Bold values indicate main variable of interest.

***Statistically significant at the 1% level; **statistically significant at the 5%; *statistically significant at the 10% level (two-sided test).

that the treatment and control users had similar call trends in the pre-period.

We also ran the same test but with the 2007 Web-registered customers as the treatment group and the non-Web-registered customers as the control group in pretreatment periods, 2005 and 2006 (March–December 2006). Note that in these periods, the 2007 Web-registered customers did not have access to the portal. The results for the specification are reported in column B of Table 3. We again find an insignificant estimate for the variable *reg.post*. In all, the results in Table 3 indicate the call trends for the Web-registered customers and the non-Web-registered customers in our sample are similar in the pre-Web registration period.

A bigger threat to the identification is that a particular health event (say, sudden illness) may increase a customer's desire to find more information and drive her to register on the Web and also make more phone calls leading to a positive estimate for β_4. However, we have detailed information on the health events for a customer. We also include a higher-order polynomial form (squared and cubed) for the health event controls (number of claims, provider charges, and customer liability) in the regression to account for possible nonlinear effects of the health severity on the customer calling behavior. Our estimates remain unchanged. Usually, a sudden change in the health event would lead to an unusual increase in the calling propensity even before the portal registration. We do not find any such evidence in our data. In the next section, we will provide more evidence that the unobserved health events cannot explain an increase in the calls. We will show that the increase in the calls occur disproportionately in the medium-severity health events. This is consistent with our theory that the low cost of accessing the portal provides an opportunity for a customer to visit the portal even for the low- and medium-severity events. If the portal causes aggravation, we should see a higher increase in this region. However, if the unobserved health events were causing the call increase, we should not see a disproportionate increase in the medium-severity health events.

We can further allay the selection concerns by running specification (A) with the treatment group as the registered customers who actually used the portal and the control group as the registered customers who did not visit the portal. This is similar to the Campbell and Frei (2010) paper where they distinguish between active and passive users. The passive users here are the ones who register for the Web portal but do not use it. Thus the customers in both treatment and control groups are registered on the Web portal. The results are given in column A of Table 4. We find a positive and significant coefficient estimate for variable *wb.post*. This indicates that the customers who actually used the portal made more phone calls relative to the customers who merely registered on the portal.

We also considered the actual Web users as the treatment group and the non-Web-registered customers as the control group. The results are reported in column B of Table 4. It is evident that we get similar sign and significance but a higher magnitude for the coefficient estimate.

Note that we dropped customers other than the 2006 Web-registered customers from our total sample of Web-registered customers in our analysis so far. We took several different samples from the Web-registered customers from the 2005 and 2007 cohort and used them as the treatment group. We find qualitatively similar results indicating that the increase in calls as a result of Web registration is not restricted to the 2006 cohort alone.

Table 4 Estimates for Actual Web Portal Usage

Dependent variable: *Telephone calls*	A Web users vs. registered non-Web users		B Web users vs. registered non-Web users	
post dummy	−0.041	(0.062)	−0.021*	(0.011)
wb.post dummy	**0.454*****	**(0.112)**	**0.594*****	**(0.081)**
Customer liability (in $1,000)	0.021**	(0.012)	0.006***	(0.001)
Negotiated charges (in $10,000)	−0.032	(0.036)	0.021**	(0.012)
Number of claims	0.028***	(0.002)	0.018***	(0.001)
Constant	0.628***	(0.062)	0.701***	(0.015)
N	5,430		98,338	
R^2 (adj. R^2)	0.701	(0.395)	0.692	(0.398)

Notes. Standard errors are in parentheses. Bold values indicate main variable of interest.

***Statistically significant at the 1% level; **statistically significant at the 5%; *statistically significant at the 10% level (two-sided test).

We also use the propensity score method to match the Web portal-registered users with the control users (Rosenbaum and Rubin 1983). We calculate the propensity score using the observables (health event controls as well as age and gender). We then use a variety of matching algorithms (one-to-one matching, one-to-many matching) to estimate the effect of the treatment (the portal registration) on the telephone calls. The results are similar to what we find in the main specification (A).[6]

4.2. Instrumental Variable Specifications

Although all robustness specifications suggest our results are robust, we now use an instrumental variable specification to rule out an outside possibility of an omitted variable contaminating our estimate. We need an instrumental variable (IV) that provides an exogenous source of variation for our endogenous variable (portal registration). Such an IV should be uncorrelated to the structural error—ε (or change in calls) in the specification (A) but be correlated to the endogenous Web portal registration variable. We use the age of the users as an instrument. Intuitively, we expect that the younger users are more likely to register for the portal. However, one may worry that age is correlated with the number of calls (for example, the older users are more likely to call). We propose a first-difference form of regression specification to overcome this problem (see specification (B) in the appendix). In specification (B), our outcome variable is not the *number of calls* but the *difference in the number of calls between the two periods*. We do not expect that age drives the change in calls over such a short time period. We use a falsification test to show that age is not related to the change in calls (see Angrist and Krueger 1991 for similar exogeneity tests for the instruments). For the 2007 registered cohort, we have the call data for 2005 and 2006. We use the difference in calls and regress it with age and other control variables. We find insignificant estimates for the age coefficient (see Table A.1 in the appendix). We also use the nonregistered customers and use the one-year and two-year differences (2005 to 2006 and 2005 to 2007) in phone calls and regress it with age and other control variables. We again find that age is uncorrelated with the change in calls.

We provide more details on our instrument and results with the instrumental variable analysis in the appendix. In short, we find that even with the instrumental variable, our results remain robust and that the Web portal registration leads to more phone calls.

[6] The results are available upon request from the authors. The estimate for the average treatment effect is about 0.21, which is similar to the OLS estimate of 0.18.

4.3. Distribution of Telephony and Web Portal Use

So far, we have used the aggregated data (pre- and post-period) to show that Web portal registration and usage leads to more phone calls. However, during this time period, consumers experience different health events, and some are more severe than others. Data aggregation precludes us from exploring these in more detail. We now disaggregate our data in a more granular time dimension. In particular, we show that we can identify the *aggravation* effect of the Web portal use established in §3. Note from Figure 1 that once the customers register on the portal, they are likely to use the portal for the medium-severity events. Absent the portal, they are not likely to call for medium-severity events because of the higher associated costs of the telephony. Therefore, if the Web causes the aggravation effect, our model predicts that we should see an increase in the calls related to medium-severity health events.

To test this prediction, we used the same treatment group of 2,715 customers who registered on the Web portal in the year 2006. We included in the control group, a random sample of 3,000 customers from the total of 48,631 non-Web portal-registered customers. We divided the entire 2005 and 2007 periods for each of these selected customers into 15-day intervals to signify their individual health events (we found similar results by taking 30-day health events). For each of these windows, we compute the total customer liability.[7] To keep the analysis more intuitive, we classify these 15-day windows into three separate health severity categories: (i) *low severity* if the aggregated customer liability for a window is less than \$50, (ii) *medium severity* for liability between \$50 and \$150, and (iii) *high severity* for liability more than \$150. (These health event categories were created for highly skewed distribution of health severity variables in our sample. However, we get similar results with different numerical categorization of the health events.) The summary statistics of health severity variables, number of telephone calls, and number of Web visits for these 15-day windows are given in Table 5.

We now formally explore how the Web portal and telephony usage vary with health severity. Per our theoretical model, we expect that, because of the high cost of telephony use, users are likely to make phone calls mostly during the high-severity event. In short, we expect fewer phone calls during the medium-severity event and especially during the low-severity event.

[7] We can perform the analysis by including provider-negotiated charges or by taking a combination of provider-negotiated charges and customer liability. We focus on the customer liability as our results suggest that it is a significant predictor of calls.

Table 5 Summary Statistics for Different Severity Health Events

	2005		2007	
Variable	Observations	Mean	Observations	Mean
Non-Web-registered customers				
Low liability health event				
Telephony use	14,605	0.052	17,582	0.044
Web usage	14,605	0	17,582	0
Medium liability health event				
Telephony use	8,415	0.069	9,348	0.052
Web usage	8,415	0	9,348	0
High liability health event				
Telephony use	9,544	0.076	11,361	0.069
Web usage	9,544	0	11,361	0
Web registered customers				
Low liability health event				
Telephony use	13,914	0.048	14,214	0.047
Web usage	13,914	0	14,214	0.043
Medium liability health event				
Telephony use	8,934	0.066	9,835	0.063
Web usage	8,934	0	9,835	0.061
High liability health event				
Telephony use	9,222	0.103	12,026	0.092
Web usage	9,222	0	12,026	0.076

To test this, we use 2005 data for Web-registered customers (pre-Web portal registration) and estimate the following fixed effect logit specification:

$$P(tel)_{it} = \beta_i + \beta_1.M_{it} + \beta_2.H_{it} + \varepsilon_{it} \quad \text{(Specification (C))},$$

where

$tel_{it} = 1$ if customer i uses the telephone in window t and 0 otherwise,
$M_{it} = 1$ if the health event is of medium-severity during window t and 0 otherwise,
$H_{it} = 1$ if the health event is of high-severity during window t and 0 otherwise, and
$\beta_i =$ Group fixed effect for customer i.

Similarly, we expect that because of the relatively low cost of the portal, the users are likely to use the portal even for the medium (and possibly low)-severity event. Thus, whereas the telephone is used for the high-severity events, portals are used for the medium- and low-severity health events. In other words, we should see a change in the distribution of usage when portals are used as opposed to when the telephone is being used. To test this, we use 2007 data for Web-registered customers (postregistration) to estimate similar specification as in (C) but for the Web portal use:

$$P(wb)_{it} = \beta_i + \beta_1.M_{it} + \beta_2.H_{it} + \varepsilon_{it} \quad \text{(Specification (D))},$$

where i specifies the customer and t specifies the window; $wb_{it} = 1$ if customer i uses the Web in window t and 0 otherwise.

Table 6 Estimates for Specifications (C) and (D)

	(C) Preregistration telephony usage	(D) Postregistration Web usage
Medium liability health event (β_1)	0.038 (0.076)	0.214*** (0.101)
High liability health event (β_2)	0.358*** (0.084)	0.626*** (0.136)
Observations	32,070	36,075

Note. Cluster-corrected standard errors are in parentheses.
***Statistically significant at the 1% level; **statistically significant at the 5%; *statistically significant at the 10% level (two-sided test).

The results of specifications (C) and (D) are given in Table 6. We have used a cluster robust variance matrix estimator for the standard errors to account for correlation and heteroskedasticity of the error terms because of multiple health events for the same customer.[8] Notice that the low liability event is the baseline here.

For telephony usage, the coefficient estimate for the high-severity health events is significant, but the coefficient for the medium-severity health events is insignificant. This suggests that in the pre-Web registration period, customers are not very likely to use telephony for the medium-severity events. However, for the Web, the coefficient estimate for the medium- and high-severity health events are positive and significant. Even the magnitude of the medium event dummy estimate is large compared with the telephony use estimate. Thus users are more likely to use the Web even for the medium-severity event than telephony. In summary, our preliminary analysis shows that consistent with our theoretical model, users use the telephone for the high-severity events but are willing to use the Web portal for the medium-severity events as well. This indicates the potential low access cost of the Web portal.

4.4. Web Portal and the Aggravation Effect

If the Web portal usage actually aggravates customer concerns, leading her to use the telephone more, we should see a higher increase in the calls in the medium-severity health events for the Web portal-registered customers. Because users do nothing in cases of low-severity health events, the Web is unlikely to lead to the aggravation. Similarly, because the users are anyway likely to use the telephony for high-severity events, the Web is not likely to cause significant aggravation there as well.[9] However, customers use the Web portal significantly during the

[8] We also included lagged dummy variables for similar medium- and high-severity health events in preceding windows in our specifications to allow for lower probability of calls in repeated health events for a customer such as dialysis, chemotherapy, etc., and find qualitatively similar results.

[9] At least it is not readily identified.

medium-severity events even though few telephone calls occur in this window before the Web portal use. Thus, if the portal causes aggravation, then the medium-severity health events provide an opportunity for the identification of the aggravation effect.

We test this intuition by estimating the change in telephone calls across these three health events for users who registered on the portal versus those who did not. We expect that users who registered on the portal are more likely to increase their telephone usage in the *medium-severity* events. We again use a difference-in-difference design with probability of a telephone call as our dependent variable:

$$P(tel)_{it} = \beta_i + \beta_1.post_{it} + \beta_2.reg_{it} + \beta_3.reg.post_{it} + \varepsilon_{it}, \quad \text{(Specification (E))}$$

where

$tel_{it} = 1$ if customer i uses the telephone in window t and 0 otherwise,
$post_t = 1$ for 2007 and 0 otherwise, and
$reg_i = 1$ if customer i is Web portal registered and 0 otherwise.

The *reg.post* dummy in the specification (E) indicates a net increase in the probability of calls because of the Web registration. We run the specification (E) separately for the high-, medium-, and low-severity health events. The estimation results with cluster correction at the customer level are summarized in the Table 7.

Note that the *reg.post* dummy is significant (economically and statistically) only in the medium-severity health events. Thus, consistent with our model of the aggravation effect, we find that the availability of the Web portal encourages users to visit the portals for the medium-severity events (see specification (D)). However, visits to the portals cause consumers to seek more information by contacting the call centers (absent the Web portal, they would not have made as many phone calls for these events), leading to an increase in phone calls during the medium-severity health event.

Table 7 Differential Impact of Web Treatment on Low-, Medium-, and High-Severity Health Events

Dependent variable: *Probability of phone call*	Coefficient estimates		
	Low-severity health event	Medium-severity health event	High-severity health event
post dummy	−0.144** (0.068)	−0.381*** (0.088)	−0.237*** (0.074)
reg.post dummy	0.142 (0.100)	0.323*** (0.122)	0.127 (0.102)
N	43,668	36,532	42,153

Note. Cluster-corrected standard errors are in parentheses.
***Statistically significant at the 1% level; **statistically significant at the 5%; *statistically significant at the 10% level (two-sided test).

4.5. Types of Calls Affected by the Portals and Implications for the Portal Design

As noted earlier, the Web portal in general provides a vast amount of passive information with insufficient interactive tools. This may cause customer uncertainty, leading them to use the telephone. We now analyze the effect of Web portal design on telephone calls by examining two extreme categories of calls.

The first category of calls (requesting an insurance card, downloading insurance forms, and searching for health-care providers) pertains to events for which the Web portal provides simple tools to easily retrieve the relevant information unambiguously. The firm's Web portal provides three such functionalities: (1) customers can click a button on the portal that automatically sends an ID card request to the firm whose status can be readily tracked; (2) customers can fill in their zip code and the type (specialty) of health care required, and the Web portal provides details of all the participating health-care providers in the vicinity along with their quality scores; and (3) customers can easily download all insurance forms (organized in alphabetical order). We expect that the consumer can readily find this information and perform simple transactions with minimal effort. Therefore, we expect that the Web portal should substitute for these types of calls.

For the second category of calls (product- and claims-related information), the related information on the portal is provided in a comprehensive fashion with little interactive tools. We identify two such areas on the portal: (i) the product-related information is given in a 70- to 80-page product benefit booklet with little interactive search tools or features, and (ii) the claims related information on the portal includes customers' medical and pharmacy claims, health spending, and a comprehensive booklet on health-care procedures. In both of these categories, a customer has to sift through voluminous documents to search for the relevant information. Based on listening to hundreds of live calls, examples of some such calls are

- "My doctor has prescribed __ but I see four different types of coverage on my benefits pages on Web. I cannot locate or understand my coverage on my benefit page on Web. Please tell me whether it is covered under my plan?"
- "I thought my plan allowed for __ specialist visits, but I see three different conditions for it on my benefits page __ please clarify."
- "I need __ procedure for my treatment. I am confused about the cost of this procedure as appearing in health care/procedure cost page on portal."
- "What is my generic drug coinsurance rate/co-pay?"

Table 8 Summary Statistics for Specific Categories of Calls

		2005		2007	
Telephone calls	Observations	Mean	Std. dev.	Mean	Std. dev.
2006 reg. customers					
Product benefits and claims related calls	2,715	0.373	0.989	0.811	1.627
Provider participation, ID card and form related calls	2,715	0.055	0.263	0.029	0.194
NonWeb registered customers					
Product benefits and claims related calls	48,631	0.239	0.842	0.427	1.297
Provider participation, ID card, and form-related calls	48,631	0.041	0.232	0.034	0.198

We expect that the product- and claims-related calls should increase with the portal usage.

We test this intuition with our data. The mean values for these categories of calls for the 2006 Web-registered and non-Web-registered customers are given in Table 8. We see a reduction in the mean value for the first category of calls from 2005 to 2007 for both Web-registered and non-Web-registered groups, but this reduction is much higher for the Web-registered group. We also see an increase in the second category of calls from 2005 to 2007 for both groups, but the increase is much higher in the Web-registered customers. These trends support our intuition, but we now test it more rigorously via the regression specification (A).

We estimate specification (A) with the 2006 Web-registered customers as the treatment group and the non-Web-registered customers as the control group. The only difference is we now use the two call categories, as discussed above, as the dependent variable. The results are given in Table 9.

We find a negative and significant coefficient estimate for the *reg.post* dummy for the first category of calls and a positive and significant coefficient for the second category of calls. The coefficient estimate value −0.016 for the first category of calls signifies a 29% decrease in such calls because of the Web registration. Likewise, an estimated value of 0.246 for the second category of calls signifies a 66% increase in such calls because of the Web registration. These results clearly indicate that the Web portal is useful in reducing calls where it provides tools to customers to readily retrieve relevant information unambiguously. But at the same time, it leads to a substantial increase in the number of calls if the information is ambiguous and comprehensive without adequate interactive/search tools.

Table 9 Impact of Web Portal Use on Type of Calls

Dependent variable: *Telephone calls*	Calls for unambiguous information	Calls for ambiguous information
post dummy	−0.011*** (0.000)	0.173*** (0.006)
reg.post dummy	**−0.016**** **(0.011)**	**0.246***** **(0.032)**
Customer liability (in $1,000)	0.000 (0.000)	−0.007 (0.016)
Negotiated charges (in $10,000)	−0.002 (0.001)	0.008 (0.012)
Number of claims	0.000*** (0.000)	0.007*** (0.001)
Constant	0.036*** (0.001)	0.104*** (0.072)
N	102,692	102,692
R^2 (adj. R^2)	0.50 (0.19)	0.63 (0.25)

Note. Bold values indicate main variable of interest.
***Statistically significant at the 1% level; **statistically significant at the 5%; *statistically significant at the 10% level (two-sided test).

4.6. Additional Robustness Checks

If Web portal usage actually causes telephone calls, such calls should happen within a reasonable period immediately after the Web portal usage. We now show evidence of this. We ran three separate specifications (A) with the dependent variables as (1) total calls minus the calls made within three days of the portal visit, (2) calls of category 1 (where related information is clear on the Web) minus the calls in this category made within three days of the portal visit, and (3) calls of category 2 (where related information is ambiguous on Web) minus the calls of this category made within three days of the portal visit. If the Web causes aggravation, then the calls are made in the short time period following the portal visit. Therefore, once these are removed, the estimates for the call increase would be insignificant for the total number of calls and category 2 calls. Because, the Web helps in resolving the category 1 calls, the result of a decrease in the calls because of the portal usage for this category should remain unchanged. The estimated coefficients are given in Table 10 (we estimate specification (A) after eliminating calls three days after the portal visit).

Note that in the first row of Table 10, we provide the corresponding coefficient estimates with total number of calls for quick comparison (for brevity, we suppress the estimate of all other control variables and health severity variables). Note that the coefficient estimates for the *reg.post* dummy are as per our

Table 10 Estimates with Calls Made Immediately After the Web Visit Was Removed

Coefficient estimate for *reg.post* dummy	Total number of calls	Calls for ambiguous information	Calls for unambiguous information
All calls	0.180*** (0.051)	0.246*** (0.032)	−0.016** (0.015)
Calls within 3 days from the portal visit removed	0.022 (0.053)	0.083 (0.637)	−0.019*** (0.006)

Note. Standard errors are in parentheses.
***Statistically significant at the 1% level; **statistically significant at the 5%; *statistically significant at the 10% level (two-sided test).

intuition and support our claim of the causal increase in the telephone calls because of the portal usage where the Web provides ambiguous information.

From Figure 3, we observe that the customer calls increase sharply with the Web registration but then decline and stabilize at a higher level. We exclude the year of registration from our analysis to avoid picking up any transient aggravation and/or learning effect of the Web. It may be possible that the customers learn to use the portal with time and the magnitude of the increase in calls may decrease with time. We conduct additional analysis to reassure readers that learning on the Web does not alter our result qualitatively. We divide the 2006 Web-registered customers into two cohorts, old registered customers (registered on the Web portal between March and July 2006) and new registered customers (registered on the Web portal between August and December 2006). We find 1,317 old registered customers (48.5%) and 1,398 new registered customers (51.5%) in our sample of 2006 Web-registered customers. Now we run revised specification (A) with these old and new registered customers as two treatment groups, and the non-Web-registered customers are the control group:

$$tel_{it} = \beta_i + \beta_1 X_{it} + \beta_2 regnew_i + \beta_3 regold_i + \beta_4 post_t + \beta_5 regnew.post_{it} + \beta_6 regold.post_{it} + \varepsilon_{it},$$

where

$regnew_i = 1$ if customer i is registered on the Web portal between August and December 2006 and 0 otherwise, and

$regold_i = 1$ if customer i is registered on the Web portal between March and July 2006 and 0 otherwise.

Other terms have similar meaning as in specification (A). The coefficient estimate for β_5 and β_6 indicates the treatment effects of Web usage for the new and old Web portal-registered customers, respectively. We show the estimated coefficient in Table 11. We find positive, significant, and similar treatment coefficient estimates for the old and the new registered customers. The estimate for the old registered customer cohort is a bit smaller but statistically indistinguishable from the estimate for the new registered customers. So overall, our results show a similar magnitude in the call increase for the old and the new Web-registered customers.

However, in the year of registration, we see a large increase in call volume and then some decline (see Figure 3). It may be due to learning to use the Web. We intend to examine this aspect in future work.

5. Conclusions, Managerial Implications, and Limitations

We conduct a field study at a large U.S. health insurance firm and examine how the availability of the Web portal affects the demand for the telephony-based assisted call center services. We first provide a model for the customers' call-generation process that indicates the possibility of both a substitution and an aggravation effect of the Web portal on the telephone calls. We then collect rich individual-level data to estimate the net effect of portal usage on the telephone calls using a difference-in-difference regression design. Our result is counterintuitive and contrary to the prevalent rationale of cost savings for investments in SSTs. We find that Web-based self-service usage leads to about a 14% increase in telephone calls. A deeper inspection shows that the number of calls goes down when the information sought by the users on the Web is unambiguous. However, the aggravation effect is very large, as expected, when the information is ambiguous and likely to cause more questions, leading to more telephone calls. We run several falsification and specification tests to ensure that our results are robust to any potential selection of customers in the Web portal treatment in our sample. More importantly, we believe our analysis yields to causal inferences; i.e., Web portal usage is the reason for the observed changes in telephone usage.

Our result has significant managerial implications for new-generation contact center operations, where the number of different channels of customer services such as telephony, Internet chat, Web portal-based self-service, interactive voice response unit, e-mail, and SMS are offered simultaneously. In particular, our results point to a complex interaction between multiple channels of service provision. Whereas the Web portals are becoming a popular choice as a way to reduce costs for customer service, our data show a

Table 11 Estimates Investigating Learning in 2006 Web-Registered Customers

Dependent variable: *Telephone calls*	Coefficient estimates
post dummy	−0.021* (0.013)
regnew.post dummy	**0.201*** (0.062)**
regold.post dummy	**0.159** (0.071)**
Customer liability (in $1,000)	0.006*** (0.001)
Negotiated charges (in $10,000)	0.025** (0.008)
Number of claims	0.017*** (0.000)
Constant	0.737*** (0.015)
N	102,692
R^2 (adj. R^2)	0.694 (0.398)

Notes. Standard errors are in parentheses. Bold values indicate main variable of interest.

***Statistically significant at the 1% level; **statistically significant at the 5%; *statistically significant at the 10% level (two-sided test).

more subtle and complex effect. Web portals are effective for simple, unambiguous tasks but not so for comprehensive and ambiguous tasks. In fact, in such cases, the portals can actually increase incoming calls, increasing the cost for the call center. Thus, the design of the Web portal is an important dimension that we discuss later.

We have collected a rich sample, but it is from one large firm. However, we believe our findings can be generalized. First, we provide a model for call generation and Web portal usage that is applicable to any call center setting. Second, our main result, that the Web portal causes an aggravation effect, should be applicable in settings where a firm provides comprehensive but unstructured information on Web portals regarding its products and services. A few examples of such settings are other insurance services (such as automobile insurance), IT and IT-related services (such as cable/Internet services), and banking and financial services (such as investment advisory services). In such settings, customers are expected to find answers to their nonstandard queries by searching through a large amount of and unstructured information on firms' Web portals. A paper by Cambell and Frei (2010) points to the generality of our results. Using data from a bank, they show that online banking customers are *more*, not less, likely to go to local branches or call center. This suggests that sometimes introducing one channel does not necessarily reduce the consumption in another channel.

Our results provide several broad recommendations for practicing managers. First, in a multichannel service setup, a channel or a group of channels cannot be managed in isolation. Understanding how customers choose among the available channels and how usage of one channel affects the demand of the other available service channels is essential for designing, deploying, and effectively managing multichannel service setup. Second, a self-service channel should not be seen only as a means of reducing operational cost. Firms often differentiate by offering multiple channels to its customers to choose the one that they find most convenient. Prior research indicates that this may lead to higher customer satisfaction, retention, and loyalty (Wallace et al. 2004, Danaher et al. 2003). Third, a self-service channel often does not resolve a customer query completely, and thus customers should be provided with options to escalate the query to the other possible assisted channels such as e-mail, Web chat, telephony, etc. (Bonde 2006). Management practitioners caution that without such escalation options, the customer may feel frustrated and could churn. Moreover, the escalation option also addresses the lack of human contact in the self-services channel, which is found to be detrimental to customer loyalty (Ariely et al. 2002). Finally, low-value, high-volume standard transactions that require the least customer effort and knowledge are the ones most suitable for self-services (Boyer et al. 2002). Customers use self-service options for the standardized services and thus impose lesser demand on the other available alternative channels. However, for the Web-based self-services options where customers have to first search for relevant information from a comprehensive but passive information source and then process it to find desired answers, the complete resolution of the customer query through the self-service channel alone does not seem so straightforward. The potential aggravation effect of such Web usage can be mitigated by possibly incorporating interactive query resolution tools in the Web portal design.

The present work has several limitations, which also offer opportunities for extensions in the future. One limitation is that we do not have data on call durations, and thus we cannot rule out the possibility that calls made after Web-based self-service may be shorter and thus less costly. We interviewed several managers and CSRs, and they did not feel that calls followed by Web usage are shorter. Moreover, total call handling time is composed of actual talking time with the customer and the time taken for wrap-up activity. Normally, the wrap up activity takes around 20% to 30% of the total time. So even if the talking time is reduced in the calls after the Web-based self-service visit, the wrap-up activity still takes the same amount of time, and thus the net effect on total call handling time may not be much. Moreover, we have only estimated the mean impact of Web portal registration on telephone calls for the population of customers. It would be valuable for firms to segregate customers who use Web-based self-service efficiently from those who do not. A random coefficient model/hierarchical Bayesian model on panel data with multiple realizations for individual customers would allow for estimation of individual-level parameters and thus distribution of the parameter estimates. This could be an interesting future extension of the present work. We also find some evidence of customers learning to use the Web with time. An extended data set with more time periods would allow us to explore and estimate the impact of such learning on customer calls.

Acknowledgments
The authors are thankful to the seminar participants in Statistical Challenges in E-Commerce Research 2009 and the Conference on Information Systems and Technology 2009, the University of Minnesota, and Carnegie Mellon University for many beneficial comments. They also thank

Lowell Taylor, Gérard Cachon, and Kartik Hosanagar for many useful discussions.

Appendix. Instrumental Variable Approach

We use age and its polynomials as instruments for Web registration. We expect that younger users are more likely to use the Web. We also expect that although age may be correlated with calls, it is unlikely to be correlated with the difference in calls, which is our dependent variable.

We find a strong nonlinear relationship between the variable Web portal registration and age—age, the square of age, and the cube of age (hereafter referred to as the age variables) bear a strong correlation with the variable *reg*. The F statistic for exclusion of age variables in the regression is 416. Thus, the age variables satisfy the basic requirement of relevance of instrumental variable (Bound et al. 1995).[10]

Next, we test the exogeneity of our instrumental variables. We test this by regressing the change in telephone calls from 2005 to 2006 (outcome variables) on age variables and other control variables for a subsample of customers who do not get the treatment of Web registration. We use (1) non-Web-registered customers and (2) 2007 Web-registered customers for this purpose, as both of these samples of customers did not get the treatment of Web registration in the years 2005 and 2006. If age variables are the valid IV, we should see an insignificant coefficient estimate for age variables in this regression. Table A.1 reports the coefficient estimates. We find insignificant coefficients for all age-related variables for such a subsample of customers in our data. This indicates that the age variables satisfy the exogeneity requirement of the instrumental variable in our setup (Angrist et al. 1991).[11]

After ensuring that our age variables satisfy the basic requirements of instrumental variables, we propose IVs regression specification (B) with age variables as IV for endogenous variable *reg*:

$$\Delta tel_i = \beta_0 + \beta_1 \Delta X_i + \beta_2 reg_i + \varepsilon_i,$$

$$reg_i = \alpha_0 + \alpha_1 \Delta X_i + \alpha_2 age_i + \zeta_i \quad \text{(Specification (B))},$$

where $E(\varepsilon/\Delta X, age) = 0$ and $E(\zeta/\Delta X, age) = 0$. Note that age in the above specification signifies all age variables, i.e., age, the square of age, and the cube of age.

The results of specification (B) with age variables are given in Table A.2. We see that Web portal registration (*reg*) bears a strong partial correlation with age variables. The two-stage lease squares (2SLS) regression with instrumental variables shows a positive and significant estimate for the coefficient of variable *reg*, indicating that Web portal registration leads to an increase in telephone calls. Note that these estimates are higher than the one we obtained with the simple OLS in specification (A).

[10] Bound et al. (1995) indicate that the partial R^2 and the F statistic for exclusion of the identifying instruments in the first stage regression are useful indicators of the quality of the IV.

[11] Angrist et al. (1991) similarly show that their instrumental variable, *quarter of birth*, did not have any significant relationship with *earnings* (dependent variable) for the subsample of college graduates who were not affected by compulsory schooling laws (treatment).

Table A.1 Change in Call as a Function of Age

Dependent variable: *Change in telephone calls*	A For 2007 reg. customers		B For non-Web reg. customers	
Age	−0.119	(0.107)	0.005	(0.008)
Age^2	0.002	(0.002)	−0.000	(0.000)
Age^3	−0.000	(0.000)	0.000	(0.000)
Change in customer liability (in $1,000)	0.001	(0.006)	0.006***	(0.001)
Change in negotiated charges (in $10,000)	0.024	(0.038)	0.021***	(0.000)
Change in number of claims	0.023***	(0.002)	0.017***	(0.001)
Constant	1.865	(1.173)	−0.112	(0.124)
N	2,626		48,631	
R^2 (adj. R^2)	0.053	(0.052)	0.043	(0.043)

***Statistically significant at the 1% level; **statistically significant at the 5%; *statistically significant at the 10% level.

Table A.2 Estimates for IV Specification (B)

Dependent variable: *Change in telephone calls*	2SLS coefficient estimates		First-stage coefficient estimates	
Change in customer liability (in $1,000)	0.006**	(0.003)	−0.000	(0.000)
Change in negotiated charges (in $10,000)	0.021	(0.011)	0.001	(0.001)
Change in number of claims	0.018***	(0.000)	0.000	(0.000)
reg dummy	**0.670*** **	**(0.309)**		
Age			0.013***	(0.000)
Age^2			−0.000***	(0.000)
Age^3			0.000***	(0.000)
Constant	−0.046**	(0.021)	−0.098***	(0.018)
N	51,346		51,346	
R^2 (adj. R^2)	0.043	(0.042)	0.026	(0.023)

Notes. Standard errors are in parentheses. Bold values indicate main variable of interest.

***Statistically significant at the 1% level; **statistically significant at the 5%; *statistically significant at the 10% level (two-sided test).

References

Aksin, Z., M. Armony, V. Mehrotra. 2007. The modern call center: A multi-disciplinary perspective on operations management research. *Production Oper. Management* **16**(6) 665–688.

Anderson, E. W., C. Fornell. 1994. A customer satisfaction research prospectus. R. T Rust, R. E. Oliver, eds. *Service Quality: New Directions in Theory and Practice*. Sage, Thousand Oaks, CA, 241–268.

Angrist, J. D., A. B. Krueger. 1991. Does compulsory school attendance affect schooling and earnings? *Quart. J. Econom.* **106**(4) 979–1014.

Angrist, J. D., A. B. Krueger. 1999. Empirical strategies in labor economics. O. Ashenfelter, D. Card, eds. *Handbook of Labor Economics*, Vol. 3A. Elsevier, Amsterdam, 1277–1366.

Aral, S., E. Brynjolfsson, M. W. Van Alstyne. 2011. Information, technology and information worker productivity: Task level evidence. *Inform. Systems Res.* Forthcoming.

Ariely, D., J. Lynch, Y. Moon. 2002. Taking advice from smart agents: The advice likelihood model. Paper presented at the Marketing Science Institute/Duke Joint Conference on Customer Relationship Management, January, Duke University, Durham, NC.

Barua, A., C. Kriebel, T. Mukhopadhyay. 1995. Information technology and business value: An analytic and empirical investigation. *Inform. System Res.* **6**(1) 3–23.

Bitner, M. J., S. W. Brown, M. L. Meuter. 2000. Technology infusion in service encounters. *J. Acad. Marketing Sci.* **28**(1) 138–149.

Biyalogorsky, E., P. Naik. 2003. Clicks and mortar: The effect of on-line activities on off-line sales. *Marketing Lett.* **14**(1) 21–32.

Bobbitt, L. M., P. A. Dabholkar. 2001. Integrating attitudinal theories to understand and predict use of technology based self-service: The Internet as an illustration. *Internat. J. Service Indust. Management* **12**(5) 423–450.

Bonde, A. 2006. Can self-service deliver better service? The top 10 mistakes companies make in introducing self-service—And how to avoid them. *CIO Magazine* (June 27) http://www.cio.com/article/22496/Can_Self_Service_Deliver_Better_Service_.

Bound, J., D. A. Jaeger, R. M. Baker. 1995. Problem with instrument variable estimation when the correlation between instruments and the endogenous explanatory variable is weak. *J. Amer. Statist. Assoc.* **90**(430) 443–450.

Boyer, K. K., R. Hallowell, A. V. Roth. 2002. E-services: Operating strategy—A case study and a method for analyzing operational benefits. *J. Oper. Management* **20**(2) 175–188.

Brynjolfsson, E., L. Hitt. 1995. Information technology as a factor of production: The role of differences among firms. *Econom. Innovation New Tech.* **3**(3–4) 183–199.

Bulkley, N., M. W. Van Alstyne. 2006. An empirical analysis of strategies and efficiencies in social networks. MIT Sloan Research Paper 4682-08, Massachusetts Institute of Technology, Cambridge.

Campbell, D., F. Frei. 2010. The cost structure, customer profitability, and retention implications of self-service distribution channels: Evidence from customer behavior in an online banking channel. *Management Sci.* **56**(1) 4–24.

Cox, D. F. 1967. *Risk-Taking and Information Handling in Consumer Behavior*. Harvard University Press, Boston.

Danaher, P. J., I. W. Wilson, R. A. Davis. 2003. A comparison of online and offline consumer brand loyalty. *Marketing Sci.* **22**(4) 460–476.

Danaher, B., S. Dhanasobhan, M. D. Smith, R. Telang. 2010. Converting pirates without cannibalizing purchasers: The impact of digital distribution on physical sales and Internet piracy. *Marketing Sci.* **29**(6) 1138–1151.

Deleersnyder, B., I. Geyskens, K. Gielens, M. G. Dekimpe. 2002. How cannibalistic is the Internet channel? A study of the newspaper industry in the United Kingdom and the Netherlands. *Internat. J. Res. Marketing* **19**(4) 337–348.

Dermer, J. D. 1973. Cognitive characteristics and the perceived importance of information. *Accounting Rev.* **48**(3) 511–519.

Frost & Sullivan. 2006. North American IVR and Web-based self-service markets. Report FA24-72, Frost & Sullivan, San Antonio, TX. http://bp.avaya-news.com/bpnf/2006/support/docs/082306MarketShareReport.pdf.

Gabszewicz, J. J., J. F. Thisse. 1979. Price competition, quality and income disparities. *J. Econom. Theory* **20**(3) 340–359.

Gans, N., G. Koole, A. Mandelbaum. 2003. Telephone call centers: Tutorial, review, and research prospects. *Manufacturing Service Oper. Management* **5**(2) 79–141.

Hitt, L.M., F. X. Frei. 2002. Do better customers utilize electronic distribution channels? The case of PC banking. *Management Sci.* **48**(6) 732–748.

Kumar, A., R. Telang. 2010. Product customization and customer support costs: An empirical analysis. *Manufacturing Service Oper. Management* **13**(3) 347–360.

Mandelbaum, A. 2006. Call center (centres): Research bibliography with abstracts, Version 7. Retrieved October 28, 2011, http://ie.technion.ac.il/serveng/References/ccbib.pdf.

Meuter, M. L., M. J. Bitner, A. L. Ostrom, S. W. Brown. 2005. Choosing among alternative service delivery modes: An investigation of customer trial of self service technologies. *J. Marketing* **69**(2) 61–83.

Meuter, M. L., M. J. Bitner, A. L. Ostrom, R. I. Roundtree. 2000. Self-service technologies: Understanding customer satisfaction with technology-based service encounters. *J. Marketing* **64**(3) 50–64.

Meyer, B. D. 1995. Natural and quasi-experiments in economics. *J. Bus. Econom. Statist.* **13**(2) 151–161.

Miller, H. 1972. Environmental complexity and financial report. *Accounting Rev.* **47**(1) 31–37.

Murray, K. B. 1991. A test of services marketing theory: Consumer information acquisition activities. *J. Marketing* **55**(1) 10–25.

Rosenbaum, P. R., D. B. Rubin. 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* **70**(1) 41–55.

Siegel, D. 1997. The impact of computers on manufacturing productivity growth: A multiple-indicators, multiple-causes approach. *Rev. Econom. Statist.* **79**(1) 68–78.

Telang, R., U. Rajan. T. Mukhopadhyay. 2004. The market structure for Internet search engines. *J. Management Inform. Systems* **21**(2) 137–160.

Wallace, D. W., J. L. Giese, J. L. Johnson. 2004. Customer retailer loyalty in the context of multiple channel strategies. *J. Retailing* **80**(4) 249–263.

Xue, M., L. M. Hitt, P. T. Harker. 2007. Customer efficiency, channel usage, and firm performance in retail banking. *Manufacturing Service Oper. Management* **9**(4) 535–558.

Yankee Group. 2006. Great expectations: Self-service success can happen. Report, Yankee Group, Boston. http://www.ccng.com/files/public/Yankee_SelfService.pdf.