Impact of Customer Web Portals on Call Center: An Empirical Analysis[1]

Anuj Kumar, Rahul Telang

The Heinz College

Carnegie Mellon University

Abstract

Firms are investing millions to deploy web based self-services at their call centers primarily to reduce operating costs. The rationale 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 like telephony. We conduct a field study at the call center of a prominent US health insurance firm to examine the benefits of web based self service channel. On the one hand, the web channel may substitute the telephony channel in some cases. On the other hand, the web also exposes customers to more 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 the diff-in-diff specifications to show that the web based self service usage leads to 14 percent increase in the telephone calls. We conduct several robustness checks to show that our specifications are robust to any potential selection of customers in the web self-service usage. We further show that the impact of web portal is moderated by the portal and individual customer characteristics. We find that if the information is unambiguous and easily retrievable on the web, calls decline by 29 percent. However, for ambiguous information, the calls increase substantially. Our research provides insights into the challenges and opportunities of self service technologies design.

Keywords: Self-service, call center, customer support, value of web portal, multichannel service management

1. Introduction

Call centers and its contemporary contact centers have been firms' preferred medium for communication with their customers. Industry estimates suggest that 70 percent of all customer-business interaction takes place via the call centers. Fortune 500 companies, on average, operate 30 call centers each. 40 percent of the total AT&T's telephone traffic was 1-800 calls directed towards various call centers (Mandelbaum 2006). Firms have increasingly recognized that the call centers offer an opportunity to collect rich customer interaction data which can be analyzed to provide customized goods, services and experience to the customers. This, in turn, may lead to higher customer satisfaction and loyalty. Growth in newer and cheaper technologies and a desire to improve efficiency has 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 telephony, email, SMS. On the other hand, there are self-service channels (popularly referred to as self services technology SSTs channels) like the interactive voice response units (IVRs/VRUs), the web based self-service portals, and the ATM (Automated Teller Machines). The SSTs provide a wide array of choices for the customers to contact the firm.

Since a large proportion of the total call center cost is the direct labor cost, firms are increasingly using SSTs like the IVR and the web based self-services to cut the costly customer service representative (CSR) costs. Yankee group research report (2006) estimates that the average cost of serving a customer via the self-service channels (less than $1) is an order of magnitude cheaper than serving her via the assisted channels (between $5-$10). Within the 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. Due to this compelling economics, the web based self-service has grown from 7.7 percent of the total customer interaction in 2005 to 14.4 percent in 2007 (Yankee 2006). 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 and Sullivan 2005). The web based self-services allow

customers to not only execute transactions at convenience without any physical interface with the firm but also allow the customers to seek relevant information and manage their online account independently. It is expected that once the customers learn and start availing the convenience of this option, it would reduce the demand for the assisted channels. SSTs should also result in higher customer satisfaction and profitability (Anderson et.al. 1994, Xue et al 2007). However, the benefits of SSTs can be realized only when the customers embrace and use this technology.

Prior research has explored the determinants for customer adoption of the SSTs. Meuter et al. (2000) find that the ability of the SSTs to bail customer out of immediate troubling situation, ease of access and use, and fascination with capability of technology are the main sources of customer satisfaction with the SSTs. Meuter et al. (2005) show that the customer readiness variables like role clarity, motivation and ability are key mediators in increasing the customer likelihood of trying the SSTs. Dhabolkar et al. (2001) use attitudinal theories to propose a conceptual framework for predicting a customer's decision to use the SSTs. However, these studies rely on questionnaire or survey tools to elicit customer adoption and use of the SSTs and 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 newspaper cannibalizes sales of the paper format if their contents are similar. However, it may even enhance sales, if the contents in the two formats are different. Biyalogorsky et al. (2003) study how the addition of an Internet channel to the existing brick-and –mortar store channel for a music records firm affects the sale and find a negative but insignificant impact.

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 usage data and show that the efficiency is related to the firm profitability. Campbell et al.(2010) conduct a field study on the impact of online banking channel adoption on various

measures. They show that the users who adopted the online banking channel reduced their dependence on the IVR and ATM (*substitution*) but increased their consumption of the firm call center and local branches (*augmentation*). They use a difference-in-difference estimator 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.

Our work is also directly related to the broader studies that examine the role of IT and productivity. However, much of this literature uses aggregate data at the firm, industry or even country level (Siegel et.al 1994, Brynjolfsson et.al. 1995, Barua et.al 1995). To measure the impact of IT on productivity at more granular level, the IT researchers have focused on intermediate output measures like inventory, quality etc. A few recent papers have explored the IT productivity issue at an individual level. Aral et al. (2007) find a positive effect of IT use and information flow on individual white collar worker performance. Bulkley et al. (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 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 US health insurance firm with about four million customers. The firm offers customer support via telephony through its call center. Around 2004-2005, it introduced a web based self-service portal for its customers. We collect a random sample of 60000 customers from 2005 to 2007. About 10000 customers adopted the web portal. Using the vertical differentiation framework in economics, we show that the web portal usage could potentially substitute 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 which 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 a diff-in-diff specification to estimate the net impact of web portal usage on telephone calls. Our results show that the portal registration leads to 14 percent increase in the number of calls to the call center. Since the portal registration is not completely random, we conduct several robustness checks to show 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 29 percent decrease in calls where the related information is unambiguously provided on the portal. In contrast, the portal usage leads to 66 percent increase in the 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 a survey data. Our data is detailed enough to draw some interesting conclusions not explored in the prior literature. Third, we conduct various robustness checks to ascertain that the diff-in-diff specifications we employ is robust to any possible selection bias. Therefore, our data allows us to draw causal inferences. Last but not the least, we provide an empirical evidence of the value of the portal design in reducing calls.

This paper is organized as follows. We describe our study setting in Section 2. Section 3 outlines our theoretical framework. We describe our data and econometric specifications in section 4. We also discuss various threat of identification in this section. Finally, in section 5, we conclude with managerial implication of our research and 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 US. The firm sells several different health insurance plans to a customer base of approximately four million. After the plans are sold, firm serves its customers through its operational unit. The operational unit performs three broad activities

1. Client Services – Initial account set up and routine periodic activities like issuing invoices etc.

2. Call Center Services – Resolving customer queries.

3. Claims Services – Processing customer claims 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 about 70% of the total operating cost of $47 million in 2007. The firm received over 3 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, provide network and so on to the customers. Customers first register on the web portal by creating their secure username and password. Thereafter, the customers can access information on the web portal. Since the customers can find detailed information on the portal, the firm hoped that it would reduce the customers' need to call the customer service for their queries.

The web portal provides information in five broad areas. First, it provides information on the 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 account. 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 like location, profile, credentials and quality performance data. This helps the customers make more informed choice of providers. Lastly, the portal provides the

customers access to a health encyclopedia to get information on disease, care management, surgeries and procedures. Thus, the web portal hosts comprehensive information on customers' plan benefits, spending and health records. It also hosts large information on various health providers, general disease and health care management. The customer has to sift through this large information to search for the relevant information and then process it to find the desired answer to her query. This is very different from the situations where a customers can perform simple transaction (account transfer, bill payments etc) through a 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 the customer service representative (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 information search which is complex, uncertain and imprecise. In contrast, the telephony is a high quality and high cost channel. The wait times in call center are typically large and the call center is not open 24/7. However, the 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 (Thisse et al. 1979). [2] In the following, we provide a simple model

---

[2] Unlike a vertical differentiation model, the web and the telephony need not be a perfect substitute. Consumption of telephony need not completely eliminate the need for consumption of the web. For example, sometimes the web may provide some complementary information. As long as, on average, users perceive telephony to be a higher quality than the web, our model insights will go through.

of how the web portal can affect the telephony use. The goal is not a structural model to be estimated but a model that provides insights and testable hypothesis.

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 health event.[3] The higher the severity, the higher is the value of its resolution. Let P_c & P_w be the expected probabilities of query resolution through the telephone call and the portal visits respectively. Given that the portal is a lower quality alternative to the telephony, we expect $P_c > P_w$. Let C_c be the cost (monetized disutility) of a telephone call and C_w be the cost of a portal visit. As per our assumption, we expect $C_c > C_w$. Thus the expected utility from the telephony is $U_c = P_c \bullet V - C_c$, and from the web is $U_w = P_w \bullet V - C_w$.

Before the web portal, customers had only the call center to resolve her 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 telephony, or the web portal, or to do nothing. Given a clear order for the telephony and the portal, customers will choose telephony 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
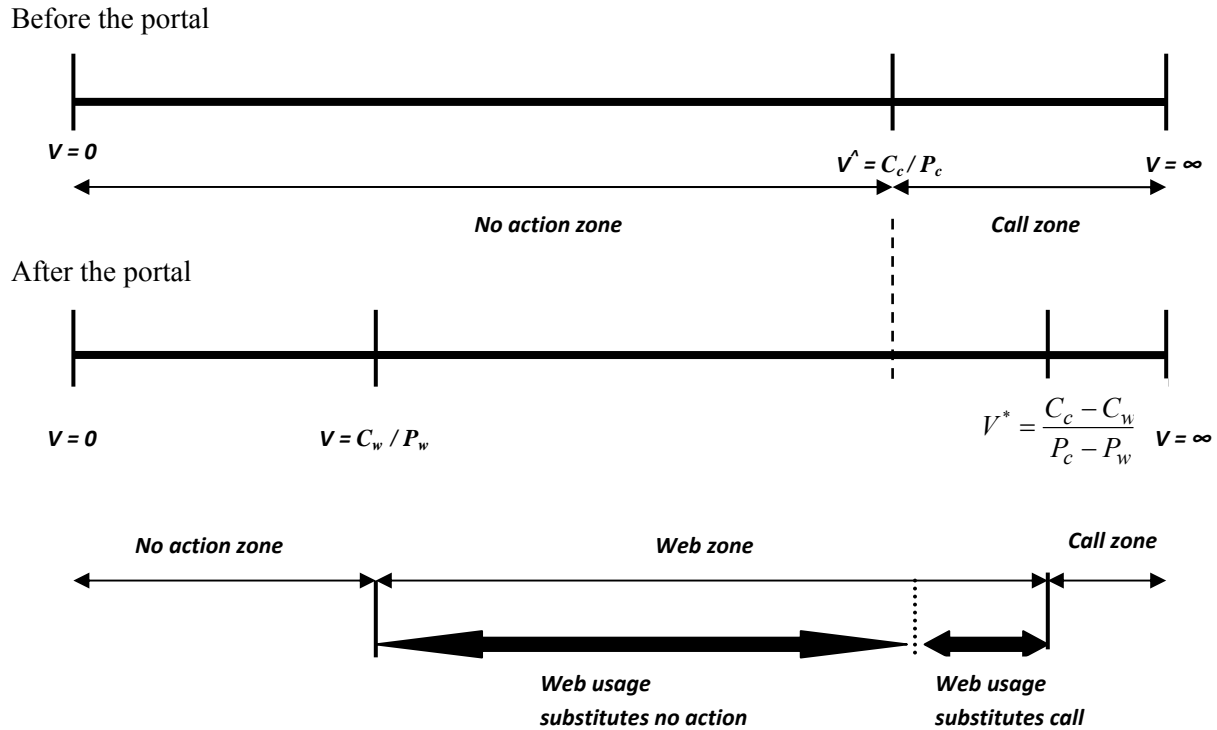
$$U_c(V^*) = U_w(V^*), \qquad \text{Or, } V^* = (C_c - C_w)/(P_c - P_w).$$

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.

---

[3] We will elaborate on this in more detail in our data section.

Figure 1- Effect of the Portal

Before the portal

$V = 0$ $V^{\wedge} = C_c / P_c$ $V = \infty$

No action zone   Call zone

After the portal

$V = 0$ $V = C_w / P_w$ $V^* = \dfrac{C_c - C_w}{P_c - P_w}$ $V = \infty$

No action zone   Web zone   Call zone

Web usage substitutes no action   Web usage substitutes call

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

In Figure 1, notice the region from V to $V^{\wedge}$ where the customers do not use the telephony for query resolution (due to the low value of the query and/or the high cost of the 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. We will show that our data matches 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 the medium severity events.
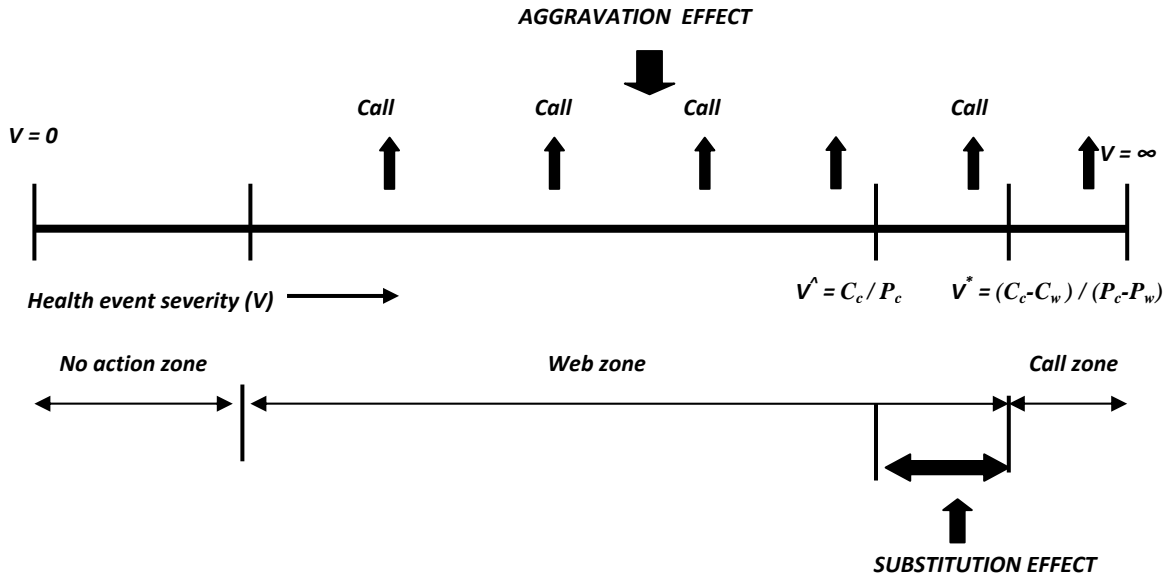
However, this stylized model relies on the fact that the portal is at least a weak substitute. However, in our setting, the web portal provides large and passive information to its customers. It does not actively provide any tool for a particular query resolution. If the customer searches for unstructured and uncertain information then the portal may not only fail to resolve the query but it may also bring additional information to customer notice. There is a large literature in the human decision making that suggests the provision of more information (if it creates ambiguity) may lead to more information seeking behavior (Cox 1967, Murray 1991). Dermer et al. (1973) found in a laboratory experiment that subjects intolerant to ambiguity prefer 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 the visits to a portal can lead to customer frustration and more calls.[4]

Notice from 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. The portal contains comprehensive information on product benefits, spending, health records and health providers. So the customer gets to see all information on a desired topic, some of it is relevant and some irrelevant, depending on the context. In processing this information, it is conceivable that sometimes this information 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). Since 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 below.

---

[4] http://www.easyerlang.com/papers/self-help.htm, accesses on October 14, 2010.

Figure 2- Combined model for call generation with the portal usage

AGGRAVATION EFFECT

Call   Call   Call   Call

V = 0   V = ∞

Health event severity (V) ⟶

$V^\wedge = C_c / P_c$ $V^* = (C_c - C_w) / (P_c - P_w)$

No action zone   Web zone   Call zone

SUBSTITUTION EFFECT

Thus the impact of 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 the telephony usage and highlight when the aggravation 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 05- Dec 07 for a random sample of 60000 customers from the entire population of the firm customers. We also collected some demographic data like age and gender for these customers. The firm started offering web portal services to its customers since 2004 onwards. After data cleaning, we get a useable sample of 59280 customers. 48631 customers did not register on the web portal at all and the remaining 10659 customers registered on web portal up to Dec 07 (5918 customers registered till December 2005, 2715 customers registered in year 2006 and remaining 2626 customers registered in year 2007). Thus our data offers us 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 call like claims

information, plan /drug / network coverage information, web related technical and usage information, membership /ID card related information, provider information, and so on. Since the portal use itself may prompt calls like 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 (hereinafter calls or telephone calls means 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, we gathered that a health event and associated severity of that health event is a major driver of customer calls. [5] We use three separate measures as 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.[6] We also compute the total dollar amount due from the customer in her claims to indicate customer liability. Note that these three variables put together provide a 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.

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

---

[5] One of the authors accomplished these during his three months internship with the firm.
[6] We used negotiated charges that the insurance firm pays to the health provider rather than the billed provider charges (amount the provider bills to the customer) as proxy for health event intensity because different quality providers will bill different charges for same procedure but the insurance firm will negotiate similar charges from such diverse providers for similar procedures.

(for 2006 cohort) or quite similar (for 2007 cohort) to the non registered customers but after the web portal registration their calls increased significantly.

Figure 3 – Trends in telephone calls
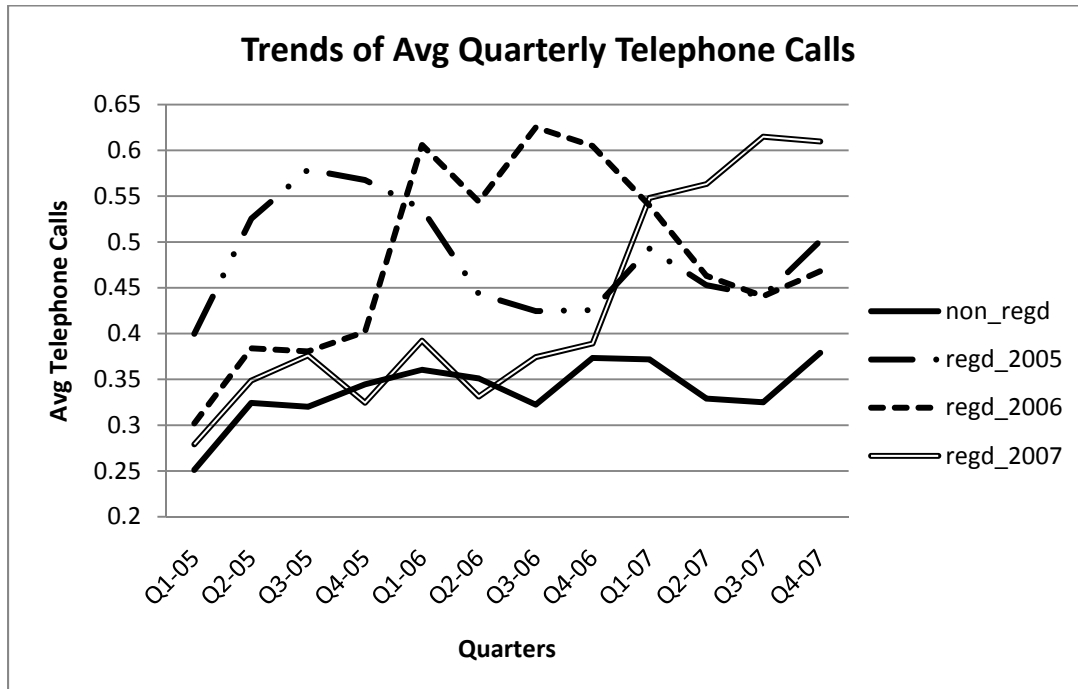
While we have data for the customers registered on the web portal in 2005 and in 2007, for this study, we focus on only those customers who registered in 2006. A simple rationale is, for this cohort, we have 10 months of data before the portal registration (period 2005 or pre-period) and same calendar 10 months of data after the registration (period 2007 or post period). For other cohorts (2005 or 2007 registration), we have unbalanced data. Focusing on the 2006 registered cohort allows us a clean identification as we can control for 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 the period 2007. We provide the summary statistics for the active portal users separately from the all registered customers in the table below.

Table 1 – Summary statistics

Variables	Observations	Period 05		Period 07	
		Mean	Std. Dev.	Mean	Std. Dev.
Non-web portal registered customers					
No of Telephone calls	48631	1.10	2.02	1.13	2.15
Customer Liability ($)	48631	1569.65	10797.57	1276.75	9470.65
Negotiated charges ($)	48631	5333.64	15765.86	6010.51	17281.91
Number of claims	48631	21.76	24.56	24.59	27.25
Number of web visits	48631	0	0	0	0
Age	48631	56.59	17.48	-	-
Proportion of females	48631	0.42	0.49	-	-
2006 web portal registered customers					
No of Telephone calls	2715	1.32	2.17	1.55	2.48
Customer Liability ($)	2715	864.95	4807.14	991.57	3738.43
Negotiated charges ($)	2715	5220.99	12156.81	6837.23	20108.34
Number of claims	2715	24.96	24.99	28.51	26.95
Number of web visits	2715	0	0	1.66	6.63
Age	2715	48.11	12.54	-	-
Proportion of females	2715	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 ($)	838	687.46	2795.86	1238.4	4933.09
Negotiated charges ($)	838	4812.5	10414.69	7260.25	12733.68
Number of claims	838	24.54	23.5	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	-	-

From Table 1, the customers who did not register for the portal made fewer calls (1.10 as opposed to 1.32) and had higher customer liability (1599 as opposed to 864). From Figure1, it is clear that the call trend for the 2006 cohort in the pre period (2005) matches quite well with the non registered users. For our identification, it is not necessary that the registered users and the non-registered users have the same

call intensity. What is needed is they have the same calling trend. We will discuss the potential threat to identification in more detail in subsequent sections.

Since we have a treatment group (portal registered customers), and a control group (no registration), 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.

Figure 4 – Quasi-natural experimental setup

	Pre-portal period 05	Post-portal period 07
2006 Web regd. customers (Treatment Gr)	• CALLS • CONTROL VARIABLES	• CALLS • CONTROL VARIABLES
Non-web regd. Customers (Control Gr)	• CALLS • CONTROL VARIABLES	• CALLS • CONTROL VARIABLES

This is a preferred specification in these settings, as it weeds out the confounding effect of unobserved factors from treatment effect (Meyer 1995, Angrist and Krueger 1999, Gruber 1994, Card 1992). The specification requires that the control group and the treatment group exhibit similar trends before the treatment. We utilize the fixed effect difference-in-difference specification (A) to examine how web portal usage impacts customers' telephone calls. We control for other observable variables that influence telephone calls: number and severity of health event and the financial liability of the customer.

$$tel_{it} = \beta_i + \beta_1 X_{it} + \beta_2 reg_i + \beta_3 post_t + \beta_4 reg.post_{it} + \varepsilon_{it} \qquad \text{--------- (Specification A)}[7]$$

where i index the customers and t indexes period (period 05 and period 07);

tel_{it} = Number of telephone calls by customer i in period t;

reg_i = 1 if the customer i is registered on the web portal and 0 otherwise;

$post_t$ = 1 for period 07 and 0 otherwise;

---

[7] We can also deploy a Poisson regression method given the count data. The results are unchanged. Since we plan to deploy a 2-stage least squares method later, we present the results with the OLS specification.

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 estimates (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 non registered customer in the post period. The fixed effect specification controls for the time invariant unobserved factors that may influence telephone calls. We scaled negotiated provider charges values in ten thousand dollars and customer liability values in thousand dollars to make the estimate β_1 more readable. We found similar results with taking logarithm of these two variables. Notice that we use a dummy variable (*reg*) to indicate the portal registration or not. 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 as sometimes the CSRs encourage the customers to visit the web after the phone call. The results of specification (A) with robust standard errors are given in Table 2.

Table 2 – Estimates for specification (A)

Dependent variable: Telephone calls	Coefficient estimates (Standard errors in parenthesis)
post dummy	-0.021[*] (0.011)
reg.post dummy	**0.180[***] (0.005)**
Customer liability (in 1000 \$)	0.006[***] (0.001)
Negotiated charges (in 10000 \$)	0.021[**] (0.012)
Number of claims	0.018[***] (0.001)
Constant	0.701[***] (0.012)
N	51346
R-sq (adj. R-sq)	0.69 (0.39)

Note - [***],[**],[*] = statistically significant at the 1%, 5% and 10% levels (two-sided test), respectively

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 web portal registration *(reg.post)* indicating that even after controlling for other

observable factors, and the unobserved customer fixed effects, customers make higher telephone calls after the web portal registration. The magnitude of coefficient estimate for *reg.post* suggests that after the registration, the telephone calls, on average, increase by 0.18, which translates into 14 percent 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 section 4.4. Before that, we provide some robustness checks on our findings.

4.1   Threat to Identification

Given that the web portal assignment is not random, there are various alternative explanations for our results. In the following, we will rule of 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 the trend in calls for the treatment and the control group to match before the treatment. Our data seems to satisfies this requirement as seen from Figure 3. We now formally show the call pattern of the control and the treatment users before the portal registration is similar. We test this assumption by comparing the trends in telephone calls for the treatment and control group over multiple pre-treatment periods (Meyer 1995). So we run specification (A) for the treatment group and the control group in two pre-treatment periods, March-July 05 and Aug- Dec 05. A significant coefficient estimate for the dummy variable *reg.post* in such specification would indicate a differential trend in calls for the treatment and the control group. The results are given in column (A) of Table 3. We find an insignificant coefficient estimate for variable *reg.post* which indicates that the treatment and control users had similar call trends in the pre-period.

Table 3 –Estimates for trends of calls in pre-registration period

Dependent variable : Telephone calls	(A) 2006 regd customers as treatment group	(B) 2007 regd customers as treatment group
reg.post dummy	**0.010 (0.031)**	**-0.022 (0.053)**
Customer liability (in 1000 \$)	-0.004 (0.004)	0.003 (0.004)
Negotiated charges (in 10000 \$)	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	51348	51259
R-sq (adj. R-sq)	0.452 (0.343)	0.608 (0.361)

Note - $^{***, **, *}$ = statistically significant at the 1%, 5% and 10% levels (two-sided test), respectively
(Standard errors are in parentheses).

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 on pre-treatment periods, period 05 and period 06 (March-Dec 06). Note that in these periods, the 2007 web registered customers did not have access to the portal. The results for the specification are given 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 trend 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 (square and cube) 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 health severity window. 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 window.

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 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 are the ones who register for the online banking channel but do not use it. Thus the customers in our treatment group and the control group are both 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.

Table 4 –Estimates for actual web portal usage

Dependent variable: Telephone calls	(A) web users versus registered non-web users	(B) web users versus non-web registered 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 1000 $)	0.021** (0.012)	0.006*** (0.001)
Negotiated charges (in 10000$)	-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	5430	51346
R-sq (adj. R-sq)	0.701 (0.395)	0.692 (0.398)

Note - ***, **, * = statistically significant at the 1%, 5% and 10% levels (two-sided test), respectively

We also considered the actual web *users* as the treatment group and the non-web registered customers as the control group. The results are given 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 form 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 the increase in calls due to the web registration is not restricted to the 2006 cohort alone.

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 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). [8]

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 out 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 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 the age is correlated with the number of calls (for example the older users are more likely to call). However, our outcome variable is not the "number of calls" but the "difference in the number of calls". We do not expect that the age drive the change in calls over such a

---

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

short time period. We use a falsification test to show that the age is not related to the change in calls (see Angrist and Kruger 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 that the age coefficient is insignificant. (See Table 12, Appendix A). We also use the non-registered customer and use one year and two year difference (2005 to 2006 and 2005 to 2007) in phone calls and regress it with the age and other control variables. We again find that the age is uncorrelated with the change in calls.

We provide more details on our instrument and results with the instrumental variable analysis in Appendix A. 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.

4.3   Web portal and Aggravation effect

So far we used the aggregated data (pre and post period) to show the 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 more granular time dimension. In particular, we show that we can identify the aggravation effect of the web portal use established in the model section of the paper. Note from Figure 1, 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 event 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 2715 customers who registered on the web portal in year 2006. We included in the control group, a random sample of 3000 customers from the total 48631 non-web portal registered customers. We divided the entire 05 and 07 periods for each of these selected customers in 15 days intervals to signify their individual health events (We found similar results by making 30 days health events). For each of these windows, we compute the total customer

liability.[9] 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 $50-150 and, (iii) *high severity* for liability more than $150 (These health event categories were created with highly skewed distribution of health severity variables. 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 days windows is given in Table 5.

Table 5- Summary statistics for different severity health events

Variable	Period 05		Period 07	
	Observations	Mean	Observations	Mean
Non-web registered customers				
Low liability health event				
Telephony use	14605	0.052	17582	0.044
Web usage	14605	0	17582	0
Medium liability health event				
Telephony use	8415	0.069	9348	0.052
Web usage	8415	0	9348	0
High liability health event				
Telephony use	9544	0.076	11361	0.069
Web usage	9544	0	11361	0
Web registered customers				
Low liability health event				
Telephony use	13914	0.048	14214	0.047
Web usage	13914	0	14214	0.043
Medium liability health event				
Telephony use	8934	0.066	9835	0.063
Web usage	8934	0	9835	0.061
High liability health event				
Telephony use	9222	0.103	12026	0.092
Web usage	9222	0	12026	0.076

---

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

We now formally explore how the web portal and the telephony usage vary with the health severity. As per our theoretical model, we expect that the users are less likely to use the telephony for medium and low severity health events. Similarly, we expect that the users are likely to use the portals for both the medium and the high severity health events. We estimate the following fixed effect Logit specification .

$$P(wb)_{it} = \beta_i + \beta_1. M_{it} + \beta_2. H_{it} + \varepsilon_{it} \qquad \text{--------- (Specification C)}$$

where i specify customer and t specify the window;

wb_{it} = 1 if customer i uses web in window t and 0 otherwise;

M_{it} = 1 if health event is medium severity during window t and 0 otherwise;

H_{it} = 1 if health event is high severity during window t and 0 otherwise; and

β_i = Group fixed effect for customer i.

We similarly estimate specification for telephony usage of web registered customers the in pre-web registration period

$$P(tel)_{it} = \beta_i + \beta_1. M_{it} + \beta_2. H_{it} + \varepsilon_{it} \qquad \text{--------- (Specification D)}$$

where, tel_{it} = 1 if customer i uses web in window t and 0 otherwise,

The results of specification (C) and (D) are given in Table 6. We have cluster corrected the standard errors to account for both correlation and heteroskedasticity of error terms due to the multiple health events for the same customer.[10] Notice that the low liability event is the baseline here. For the web, the coefficient estimate for the medium severity health event and the high severity health event are positive and significant. This indicates that the probability of web usage is high for both the medium and the high severity health events compared to the low severity health event. For telephony usage, the coefficient estimate for the high severity health event is significant but the coefficient for the medium severity health event is insignificant. This suggests that in the pre-web registration period, customers are

---

[10] We also used lagged dummy variables for similar severity health event in immediate preceding window to capture lower probability of calls in similar health events in later periods for repeated health events for a customer like dialysis, chemotherapy etc. and find qualitatively similar results.

not very likely to use the telephony for the medium severity events. In short, users use the telephony for high severity events but are willing to use the web portal for the medium severity event as well. This indicates the potential low access cost of the web portal.

Table 6 - Estimates for Specification (C) and (D)

Fixed effect logit model	(A) Web usage	(B) Telephony usage
Medium liability health event (β_1)	0.214*** (0.101)	0.038 (0.076)
High liability health event (β_2)	0.626*** (0.136)	0.358*** (0.084)
Observations	36075	32070

Note - ***, **, * = statistically significant at the 1%, 5% and 10% levels (two-sided test), respectively (Cluster corrected standard errors are in parenthesis).

If the web portal usage actually aggravates customer concerns leading her to use more telephony, we should see a higher increase in the calls in the medium severity health event for the web portal registered customers. Since users do nothing in case of low severity health event, the web is unlikely to lead to the aggravation. Similarly, since the users are anyway likely to use the telephony for high severity events, web is not likely to cause significant aggravation there as well.[11] However, the customers use the web portal significantly during the medium severity events even though few telephone calls occur in this window before the web portal use. Thus if the portals causes the aggravation, then the medium severity health events provide an opportunity for the identification of the aggravation effect.

We test this intuition by a difference-in-difference design with a fixed effect Logit specification

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

where, tel_{it} = 1 if customer i uses telephone in window t and 0 otherwise;

$post_t$ = 1 for period 07 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 due to the web registration. We run the specification (E) separately for the high, medium and low severity

---

[11] At least it is not readily identified.

health events. The estimation results with cluster correction at the customer level are summarized in the Table 7

Table-7: Differential impact of web treatment on low, medium and high severity health event

Dependent variable : Probability of phone call	Coefficient estimates (Cluster corrected standard errors in parentheses)		
	Low severity health event	Medium severity health event	High severity health event
post dummy	-0.144** (0.068)	-0.38*** (0.088)	-0.237*** (0.074)
reg.post dummy	**0.142 (0.100)**	**0.323*** (0.122)**	**0.127 (0.102)**
N	43668	36532	42153

Note - ***, **, * = statistically significant at the 1%, 5% and 10% levels (two-sided test) respectively

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. However, visits to the portals cause the consumers to seek more information by calling the call centers (absent the web portal, they would not have made as many phone calls for these events).

4.4   Types of Calls affected by the Portals and implications for the Portal Design

As noted earlier, the web portal in general provides vast amount of passive information with no interactive tools. This may cause customer uncertainty causing leading them to use the telephony. We now analyze the effect of web portal design on telephone calls by examining the two extreme categories of calls.

The first category of calls (requesting insurance card, downloading insurance forms, and searching for healthcare 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 her 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, (3) customer can easily download all insurance

25

forms (organized in an alphabetical order) from the web portal. 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 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 70-80 page product benefit booklet with little interactive search tools / features – (ii) The claims related information on the portal includes customers' medical and pharmacy claims, health spending, and a comprehensive booklet on health care procedure. In both of these categories, a customer has to sift through voluminous document to search 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 are my generic drug coinsurance rate / co-pay?".* We expect that the product and the 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 period 05 to period 07 for both the web registered and the 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 period 05 to period 07 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).

Table 8 – Summary statistics for specific categories of calls

Telephone calls	Observations	Period 05		Period 07	
		Mean	Std. Dev.	Mean	Std. Dev.
2006 regd. Customers					
Product benefits and claims related calls	2715	0.373	0.989	0.811	1.627
(Provider participation, ID card and form related calls)	2715	0.055	0.263	0.029	0.194
Non-web registered customers					
Product benefits and claims related calls	48631	0.239	0.842	0.427	1.297
Provider participation, ID card and form related calls	48631	0.041	0.232	0.034	0.198

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.

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 1000 $)	0.000 (0.000)	-0.007 (0.016)
Negotiated charges (in 10000$)	-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	51346	51346
R-sq (adj R-sq)	0.50 (0.19)	0.63 (0.25)

Note - ***,**,* = statistically significant at the 1%, 5% and 10% levels (two-sided test), respectively
(Standard errors in parentheses).

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 29 percent decrease in such calls due to the web registration for the web registered group. Likewise an estimated value of +0.246 for the second

category calls signifies 66 percent increase in such calls due to 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 substantial increase in the calls if the information is provided in a comprehensive fashion without adequate interactive/search tools.

4.5 Additional Robustness Checks

If the web portal usage actually causes the 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 specification (A) with the dependent variables as (1) total calls minus the calls made within 3 days of the portal visit, (2) calls of category one (where related information is clear on web) minus the calls of this category made within 3 days of the portal visit, and (3) calls of category two (where related information is ambiguous on web) minus the calls of this category made within 3 days of the portal visit. If the web causes aggravation then the calls are made immediately after the portal visit. Therefore, once the calls made immediately after the portal visits are removed, the estimate for the call increase would be insignificant for the total calls and the category two calls. Since, the web helps in resolving the category one calls, the result of a decrease in the calls due to the portal usage for this category should remain unchanged. The estimated coefficients for *reg.post* dummy in specification (A) are given in Table 10. For brevity, we suppress the estimate of all other control variables (health severity variables). We also provide the corresponding coefficient estimates when the calls made within 3 days of web visit are not removed in Table 10 for quick comparison.

Table 10 –Estimates with calls made immediately after web visit removed

Coefficient estimate for *reg.post* dummy	Total calls	Calls for ambiguous information	Calls for unambiguous information
With 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 - ***,**,* = statistically significant at the 1%, 5% and 10% levels (two-sided test) respectively
(Standard errors in parenthesis).

Note that the coefficient estimates for *reg.post* dummy are as per our intuition and supports our claim of the causal increase in the telephone calls due to the portal usage where web provides ambiguous information.

From Figure 3 on page 12, we observe that the customer calls increase sharply with the web registration but then decline and stabilize at a higher level. We excluded 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 increase in the calls may decrease with time. We conduct additional analysis to reassure the readers that the learning on the web does not alter our result qualitatively. We divide the 2006 registered customers into two cohorts, old registered customers (registered on web portal between March– July 2006) and new registered customers (registered on web portal between August–Dec 2006). We find 1317 old registered customers (48.5%) and 1398 new registered customers (51.5%) in our sample of 2006 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 as 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 the customer i is registered on web portal btw August–Dec06 and 0 otherwise; and

$regold_i = 1$ if the customer i is registered on web portal btw March– July06 and 0 otherwise

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

Table 11 – Estimates investigating learning in 2006 registered customers

Dependent variable: Telephone calls	Coefficient estimates (Standard errors in parentheses)
post dummy	-0.021* (0.013)
regnew.post dummy	**0.201*** (0.062)**
regold.post dummy	**0.159** (0.071)**
Customer liability (in 1000 $)	0.006*** (0.001)
Negotiated charges (in 10000 $)	0.025** (0.008)
Number of claims	0.017*** (0.000)
Constant	0.737*** (0.015)
N	51346
R-sq (adj. R-sq)	0.694 (0.398)

Note - ***, **, * = statistically significant at the 1%, 5% and 10% levels (two-sided test), respectively

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 learning to use the web. We intend to examine this aspect in the future work.

## 5.	Conclusions, Managerial Implications and Limitations

We conduct a field study at a large US 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 estimate the net effect of the portal usage on the telephone calls using a diff-in-diff regression design. We find that the web based self-service usage leads to about 14% increase in the telephone calls. A deeper inspection shows that calls go 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 web portal treatment in our sample. Our result is counter-intuitive and contrary to the prevalent rationale of cost savings for investments in SSTs. Our result has large managerial implications for the new generation contact center operations where number of different

channels of customer services like telephony, internet chat, web portal based self-service, interactive voice response unit, email and SMS are offered simultaneously.

Our results provide several broad prescriptions to practicing managers. First, in a multi-channel service set up, 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 usage of the other available service channels is essential for designing, deploying and effectively managing multi-channel service setup. Second, 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 which 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, self-service channel, often times, does not resolve customer query completely and thus should be provided with options to escalate the query to the other possible assisted channels like email, 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 issue in the self-services channel, which is found to be detrimental for customer loyalty (Areily et.al. 2002). Finally, low-value high-volume standard transactions which require least customer effort and knowledge are the ones most suitable for self-services (Roth 2001, Roth 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 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 web portal design.

Present work has several limitations, which also offer opportunity for extensions in future. One limitation is that we don't have data on call durations and thus we cannot rule out the possibility that calls

made after the web 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 for CSR is comprised of actual talking time with customer and the time taken in wrap up activity. Normally, the wrap up activity takes around 20-30% of the total time. So even if the talking time is reduced in the calls after the web self service visit, the wrap up activity still takes the same 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 will be of great value for the firms to segregate customers who use web self service efficiently from those who don't. A random coefficient model / Hierarchical Bayesian model on panel data with multiple realizations for individual customers would allow estimation of individual level parameter and thus distribution of the parameter estimates. This could be an interesting future extension of the present work. We also find some evidence of customer learning to use web with time. An extended dataset with more time period will allow us to explore and estimate the impact of such learning on customer calls.

References

Anderson. E. W., C. Fornell. 1994. "A Customer Satisfaction Research Prospectus" in service quality: New directions in Theory and Practice. Roland T Rust and Richard E Oliver eds. Thousand Oaks, CA. Sage Publication. 241-268.

Angrist J. D., A. B. Kruger. 1991. Does Compulsory School Attendance Affects Schooling and Earnings? Quarterly Journal of Economics. 106. 979-1014

Angrist J. D., A. B. Krueger. 1999. Empirical strategies in labor economics. In O. Ashenfelter and D. Card (Eds). Handbook of Labor Economics.

Aral S., E. Brynjolfsson, M. W. Van Alstyne. 2007. Information, Technology and Information Worker Productivity: Task Level Evidence. NBER Working Paper No. W13172

Ariely D., J. Lynch, Y Moon.2002. Taking Advice from Smart Agents: The Advice Likelihood Model. Paper presented at Marketing Science Institute/Duke joint conference on Customer Relationship Management. Durham NC January.

Barua A., C. Kriebel, T. Mukhopadhyay. 1995. Information Technology and Business Value: An Analytic and Empirical Investigation. Information System Research. 6 (1) 3-23.

Bitner M.J.,S. W. Brown, M. L. Meuter. 2000. Technology infusion in service encounters. Journal of the Academy of Marketing Sciences. 28(1) 138-149

Biyalogorsky E., P. Naik. 2003. Clicks and Mortar: The Effect of Online Activities on Offline Sales. Marketing Letters 14 21-32

Bonde A., Can self-service deliver better service? The top 10 mistakes companies make in introducing self-service-and how to avoid them. June 27, 2006 CIO Magazine

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. Journal of American Statistical Association. 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. Journal of Operations Management. 20 175-188
Burrows P., The Era of Efficiency. June 18 2001. Business Week.

Brynjolfsson E., L. Hitt. 1995. Information Technology as a Factor of Production: The role of Differences among Firms. Economics of Innovation and New Technology. 3 183-199

Bulkley N., M. W. Van Alstyne. 2006. An Empirical Analysis of Strategies and Efficiencies in Social Networks. MIT Sloan Research Paper No. 4682-08

Card D. 1992. Do Minimum Wages Reduce Employment? A Case Study of California 1987-1989. Industrial and Labor Relations Review. (46) 38-38

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

Dabholkar P. A., L. M. Bobbitt. 2001. Integrating attitudinal theories to understand and predict use of technology based self-service. International Journal of Service Industry Management. 12(5) 423-450.

Danaher P. J., I. W. Wilson, R. A. Davis. 2003. A comparison of Online and Offline Consumer Brand Loyalty. Marketing Science. 22. 460-76

Deleersnyder B., I. Geyskens, K. Gielens, M. G. Dekimpe. 2002. How cannibalistic is the internet channel? International Journal of Research in Marketing. 19(4) 337-48

Dermer J. D. 1973. Cognitive Characteristics and the Perceived Importance of Information. The Accounting Review. 48(3). 511-519

Frost and Sullivan. 2006. North American IVR and Web-Based Self-Service Markets. FA24-72. www.frost.com

Gabszewicz J. J., J. F. Thisse. 1979. Price competition, Quality and Income Disparities. Journal of Economic Theory. 20. 340-359.

Gruber J. 1994. The Incidence of Mandated Maternity Benefits. The American Economic Review. 622-641

Mandebaum A. 2006. Call Center (Centers): Research Bibliography with Abstracts, Version7. Downloadable from ie.technion.ac.il/serving/Reference/cclib.pdf.

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

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. Journal of Marketing. 69(April2005) 61-83

Meyer B. D. 1995. Natural and Quasi Experiment in Economics. Journal of Business and Economic Statistics. 13(2) 151-161

Miller H.1972. Environmental Complexity and Financial Report. The Accounting Review. January 1972, 31-37.

Murray K. B. 1991. A test of service marketing theory: consumer information acquisition activities. Journal of Marketing. 55(1) 10-25

Neslin S. A., D. Grewal, R Leghorn, V. Shankar, M. L. Teerling, J. S. Thomas, P. C. Verhoef. 2006. Challenges and Opportunities in Multichannel Customer Management. Journal of Service Research. 9(2) 95-112.

Parthasarathy M.,A. Bhattacherjee. 1998. Understanding post-adoption behavior in the context of online services. Information System Research. 9(4) 362-379

Siegel, D. 1997. The impact of Computer on Manufacturing Productivity Growth: A Multiple Indicators, Multiple-Causes Approach. The Review of Economics and Statistics. 79(1) 68-78.

Wallace D.W., J. L. Giese, J. L. Johnson. 2004. Customer retailer loyalty in the context of multiple channel strategies. Journal of Retailing. 80. 249-263

Yankee Group Research Inc. 2006. Great Expectations: Self-Service success can happen. www.yankeegroup.com

Appendix A: Instrumental Variable Approach

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

We find a strong non-linear relationship between the variable web portal registration and age – age, square of age and cube of age (herein after referred to as age variables) bears strong correlation with variable reg. The F statistics 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)[12].

Table 12: Change in call as a function of Age

Dependent variable: Change in telephone calls	(A) For 2007 regd. Customers	(B) For non-web regd. customers
Age	-0.119 (0.107)	0.005 (0.008)
Square of age	0.002 (0.002)	-0.000 (0.000)
Cube of age	-0.000 (0.000)	0.000 (0.000)
Change in customer liability ($1000)	0.001 (0.006)	0.006^{***} (0.001)
Change in negotiated charges ($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	2626	48631
R-sq (adj. R-sq)	0.053 (0.052)	0.043 (0.043)

Note - ***, **, * = statistically significant at the 1%, 5% and 10% levels (two-sided test), respectively

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 don't get the treatment of web registration. We use (1) non-web registered customers, and (2) 2007 web registered customers, as these customers did not get the treatment of web registration in year 2005 and 2006. If age variables are valid IV, we should see insignificant

---

[12] Bound et.al. 1995 indicated that the partial R sq and the F statistic for exclusion of the identifying instruments in first stage regression are useful indicators of the quality of the IV

coefficient estimate for age variables in regression. Table 12 reports the coefficient estimates. We find insignificant coefficients for all age related variables for such subsample of customers in our data. This indicates that age variable satisfy the exogeneity requirement of instrumental variable in our setup (Angrist et.al. 1991)[13].

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

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

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

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

Table 13 – Estimates for IV specification (B)

Dependent variable: Change in telephone calls	2SLS coefficient estimates	First stage coefficient estimates
Change in customer liability (in 1000 $)	0.006^{**} (0.003)	-0.000 (0.000)
Change in negotiated charges (in 10000$)	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)
Square of age		-0.000^{***} (0.000)
Cube of age		0.000^{***} (0.000)
Constant	-0.046^{**} (0.021)	-0.098^{***} (0.018)
N	51346	51346
R-sq (adj R-sq)	0.043 (0.042)	0.026 (0.023)

Note - $^{***}, ^{**}, ^{*}$ = statistically significant at the 1%, 5% and 10% levels (two-sided test), respectively
(Standard errors in parenthesis)

---

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

The results of specification B with age variables are given in Table 13. We see that web portal registration (*reg*) bears a strong partial correlation with age variables and the dummy variables for age bands. The 2SLS regression with instrumental variables show a positive and significant estimate for coefficient of variable *reg* indicating that web portal registration leads to increase in telephone calls. Note that these estimates are higher than the one we obtained with the simple OLS in specification (A).