TECHNOLOGY SUPPORT AND DEMAND FOR CLOUD INFRASTRUCTURE SERVICES: THE ROLE OF SERVICE PROVIDERS

A Dissertation
Presented to
The Academic Faculty

by

German F. Retana

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
Scheller School of Business

Georgia Institute of Technology
December 2013

TECHNOLOGY SUPPORT AND DEMAND FOR CLOUD INFRASTRUCTURE SERVICES: THE ROLE OF SERVICE PROVIDERS

Approved by:

Dr. D.J. Wu, co-chair
Scheller School of Business
Georgia Institute of Technology

Dr. Chris Forman, co-chair
Scheller School of Business
Georgia Institute of Technology

Dr. Sridhar Narasimhan
Scheller School of Business
Georgia Institute of Technology

Dr. Marius Florin Niculescu
Scheller School of Business
Georgia Institute of Technology

Dr. Sandra Slaughter
Scheller School of Business
Georgia Institute of Technology

Dr. Ramnath K. Chellappa
Goizueta Business School
Emory University

Date Approved: August 6, 2013

To Maru, Pa, Ma, la Machita, and Nana.

ACKNOWLEDGEMENTS

This thesis represents far more than what appears in these pages and my own effort to produce them. It represents years of guidance and support from faculty, peers, family, and friends.

I am especially grateful to the co-chairs of my doctoral committee, D.J. Wu and Chris Forman, whose exceptional support and teachings, always thoughtful, honest and assertive, were key pillars in the development of my research program. They gave me the freedom to explore and take risks, and offered me unparalleled mentorship along the way. I am also deeply grateful to Sri Narasimhan and Florin Niculescu, whose guidance and support as advisors were invaluable as well. I finally thank Sandra Slaughter and Ram Chellappa, who as members of my dissertation committee offered me keen feedback that further improved my research.

Beyond the faculty in my committee, I cannot stress how fortunate I was to be part of the IT Management program at the Scheller College of Business. Professors Eric Overby, Han Zhang, Saby Mitra, Jeffrey Hu, Lizhen Xu, and Mike Smith, all offered me guidance and wisdom at different stages in the program. I also had the pleasure of learning from other professors at Scheller and thank Stylianos “Stelios” Kavadias, Ajay Kohli, Jeongsik “Jay” Lee, Frank Rothaermel, Jerry Thursby, Marie Thursby, Debby Turner, and Manpreet Hora for their instruction. The Ph.D. experience was further enriched by my student cohort, and I thank Peng Huang, Wen Wen, Chih-hung Peng, Denny Yin, Tianshi Wu, Mike Frutiger, Hemang Subramanian, Jiao Xu, and Vincenzo Palermo for being such outstanding colleagues.

Furthermore, my research has benefited from the many conferences, workshops and seminars that I attended. I owe my sincere gratitude to the participants of the 2011 INFORMS, 2011 WISE, 2012 Winter Conference on BI, 2012 AACRIS, 2012 INFORMS, 2012 ICIS, and to participants of seminars at the Fox School of Business at Temple University, the Krannert School of Management at Purdue University, the School of Computing at the National University of Singapore, the Sauder School of Business at the University of British Columbia, and INCAE Business School for their helpful comments and suggestions on my research.

I also express my sincere thanks and appreciation to the members of the business analytics team from the provider who kindly granted me access to their data. Not only did they make me feel part of their team each time I visited their premises, but I could have not developed my research without their patience and fanatical collaboration.

Finally, outside of the school's walls, my family's unconditional and loving support was of prime value. I especially thank my wife, María "Maru" Barreto, whose enduring patience, loving encouragement, and faithful support during our adventure in Atlanta, cannot be expressed with words. My parents, German Retana and Olga Solano, have always been my example to follow and from a distance always made me feel their love. Same goes for my sisters, Melissa and Natalia, whose good vibes always reached my heart.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Service providers have long recognized that their customers play a vital role in the service delivery process since they are not only recipients but also producers, or co-producers, of the service delivered. Moreover, in the particular context of self-service technology (SST) offerings, it is widely recognized that customers' knowledge, skills and abilities in co-producing the service are key determinants of the services' adoption and usage. However, despite the importance of customers' capabilities, prior research has not yet paid much attention to the mechanisms by which service providers can influence them and, in turn, how the providers' efforts affect customers' use of the service.

This dissertation addresses research questions associated with the role of a provider's technology support and education in influencing customer use of an SST, namely public cloud computing infrastructure services. The unique datasets used to answer these research questions were collected from one of the major global providers in the cloud infrastructure services industry. This research context offers an excellent opportunity to study the role of technology support since, when adapting the standardized and commoditized components of the cloud service to their individual needs, customers may face important co-production costs that can be mitigated by the provider's assistance. Specifically, customers must configure their computing servers and deploy their software applications on their own, relying on their own capabilities. Moreover, the cloud's offering of on-demand computing servers through a fully pay-per-use model allows us to directly observe variation in the actual use customers make of the service.

The first study of this dissertation examines how varying levels of technology support, which differ in the level of participation and assistance of the provider in customers' service co-production process, influence the use that customers make of the service. The study matches and compares 20,179 firms that used the service between

March 2009 and August 2012, and who over time accessed one of the two levels of support available: *full* and *basic*. Using fixed effects panel data models and a difference-in-difference identification strategy, we find that customers who have access to full support or accessed it in the past use (i.e., consume) more of the service than customers who have only accessed basic support. Moreover, the provider's involvement in the co-production process is complementary with firm size in the sense that larger firms use more of the service than smaller ones if they upgrade from basic to full support. Finally, the provider's co-participation through full support also has a positive influence on the effectiveness with which buyers make use of the service. Firms that access full support are more likely to deploy computing architectures that leverage on the cloud's advanced features.

The second study examines the value of early proactive education, which is defined as any provider-initiated effort to increase its customers' service co-production related knowledge and skills immediately after service adoption. The study analyzes the outcome of a field experiment executed by the provider between October and November 2011, during which 366 randomly-selected customers out of 2,673 customers that adopted during the field experiment period received early proactive education treatment. The treatment consisted in a short phone call followed up by a support ticket through which the provider offered initial guidance on how to use the basic features of the service. We use survival analysis (i.e., hazard models) to compare the treatment's effect on customer retention, and find that it reduces by half the number of customers who leave the service offering during the first week. We also use count data models to examine the treatment's effect on customers' demand for technology support, and find that the treated customers ask about 19.55% fewer questions during the first week of their lifetimes than the controls.

CHAPTER 1

INTRODUCTION

Service providers have always faced the challenge that the quality of the service they offer and the value their customers derive from it depends to a great extent on the customers' own participation in the service delivery process (Bitner et al. 1997, Mills and Morris 1986). Service customers are both recipients and producers, or co-producers (Xue and Harker 2002), of the service delivered. For instance, in order to receive most business-oriented IT services, customers must be willing and capable to communicate their requirements and other essential information to the provider (e.g., Bettencourt et al. 2002). In the context of online self-service technologies (SSTs), such as online banking portals, customers must manipulate the web sites on their own to find the information and execute the transactions they need (e.g., Xue et al. 2011). Finally, in the context of online learning sites, another SST, students must be able to navigate through the technology in order to access the courses' content and manage their learning process (e.g., Tyler-Smith 2006). In general, the greater the level of participation and responsibilities customers have in the service system, the more the value they receive from the service depends on their own individual abilities (Chase 1978, Frei 2006).

Research in the context of online SSTs has consistently shown that customers' knowledge, skills and abilities are key determinants of their adoption and continued use of the services (e.g., Buell et al. 2010, Frei 2008, Xue and Harker 2002, Xue et al. 2007). However, despite how critical customers' capabilities are, prior literature has not yet explored the role that a provider may play in influencing them. Research has generally considered customers' service co-production skills as given and thus exogenous to the provider (e.g., Xue et al. 2011). Some researchers have suggested that providers should support their customers since the accessibility to external knowledge sources and support are important in determining users' decisions to adopt a new IT product (e.g., Li et al. 2005, Morgan and Finnegan 2007). Others have indicated that offering technical support is a competitive necessity for high technology vendors (Das 2003). Nevertheless, to our knowledge, no prior work has systematically examined the effect that a provider's technology support *post-adoption* may have on customers' use of its service. Our research aims to take a first step in addressing this gap in understanding.

The overarching research question of this dissertation is: *Does a provider's technology support and education influence its customers' service use?* We examine this question in the context of an emerging SST, namely public cloud infrastructure services. These services, also known as public Infrastructure-as-a-Service (IaaS), are a business-to-business (B2B) SST in which on-demand computing and storage resources (i.e., servers) are offered on a pay-as-you-go basis (Mell and Grance 2011). Our research context offers us an excellent opportunity to examine how a provider's technology support and education influence its customers' behavior for at least three reasons. First, cloud customers have a great level of responsibility in the service outcome. When using these services, "the consumer does not manage or control the underlying cloud infrastructure but has control over operating systems, storage, and deployed applications; and possibly limited control of select networking components" (Mell and Grance 2011). Second, the novelty of the service and some of its technical features represent potential challenges to customers, as is suggested by industry insight. A 2011 survey found that only 25% of IT staff in global organizations had cloud experience with public infrastructure or platform-as-a-service, and 50% of the organizations claimed that their staff was "less than somewhat prepared to handle" these services (Symantec 2011). Together, the high degree of involvement of customers in the service delivery process and the technical challenges posed by the service require customers to engage in significant service co-production efforts. Finally, the service is offered entirely on-demand and there are no contracts or subscriptions that lock in customers in any way, require minimum spending levels, or charge penalty fees for surpassing some level of consumption. Therefore, our context allows us to directly observe the variance in customers' demand for the service.

In the second chapter of this dissertation we start addressing our main research question by examining *how varying levels of involvement by the provider in its customers' service co-production processes influence the latters' use of the service*. In this chapter, the studied customers can choose (and switch between) two levels of technology support, *basic* and *full*. Full support differs from basic support in that when offering it the provider educates buyers and helps them in their service co-production processes, whereas basic support only deals with simple quality of service issues. We first develop a parsimonious service co-production model that examines a customer's tradeoffs when choosing between basic/no support and full support, and the corresponding optimal use. Then, the insights from the analytical model are used to motivate our hypotheses.

To test our hypotheses we collect unique data from a major global public cloud infrastructure services provider. Our rich data consist of 22,179 firms that used the provider's public cloud infrastructure service at some point between March 2009 and August 2012. Our econometric approach uses fixed effects panel data models and a difference-in-difference identification strategy to compare customers' use of the service before adopting full support, during their continued access to full support, and after switching from full to basic support. Given that our identification strategy assumes that unobserved factors are changing in the same way for customers who adopt full support (treated) as for those that never do (controls), and that if this assumption is violated then our estimates of the causal effects of full support adoption on service use become inconsistent, we take several additional measures in our econometric strategy. First, we conduct several falsification tests and include additional controls to address reverse causality concerns. We also run our models employing matched subsamples that are constructed based on pre-treatment behavior and using a coarsened exact matching

(CEM) procedure (Blackwell et al. 2010). The procedure supports the premise that customers do not exhibit differential behavior before the treatment. Additionally, we use instruments for the decision to upgrade to full support construct based on the occurrence of unexpected service failures; the underlying assumption for the use of these instruments is that the occurrence of failures, and more importantly the support interactions that take place between customers and the provider when working to overcome the problems they cause, can serve as a signal to customers for the value of full support and will increase the likelihood of upgrading. Finally, we include lagged values of our potentially endogenous variables as instruments (Arellano and Bover 1995, Blundell and Bond 1998) in dynamic panel models using a generalized method of moments (GMM) estimation approach; we augment our instrument matrix in the GMM estimation with the support interaction-based instruments.

We find that customers who adopt and continue having access to the provider's full support use, on average, 188% more of the service relative to customers who have only had access to basic support. Additionally, when full support customers downgrade to basic support, they continue to use, on average, 77% more of the service compared to those who have never accessed full support. We also investigate whether firm size, a measure that has been shown to be correlated with technical sophistication (e.g., Rogers 1995), is complementary with the adoption of full support. We show that larger firms exhibit a greater marginal increase in their use of the service from adopting and having access to full support. They also continue to use more of the service than smaller firms after they opt to switch to basic support. Lastly, we show that technology support helps customers make better and more efficient use of the service. Firms that access full support are more likely to deploy computing architectures that leverage on the cloud's advanced features.

The findings of our second chapter, which show the value for the provider in our study of assisting customers in their co-production processes, motivated the provider to proactively engage and offer assistance to customers immediately after service adoption. The provider conducted a field experiment that is the research context of our third chapter, in which we study *what are effects of early proactive education on customers' retention and demand for technology support during the early stage of their co-production processes?* We define early proactive education, or EPE, as any provider-initiated effort to increase customers' service co-production-related knowledge and skills immediately after service adoption. Such effort may, in turn, enable customers to derive a greater utility from the service. To our knowledge, this is the first study to examine how such proactive engagements can be used to aid customers in their service co-production processes (at any stage in their lifetimes). Additionally, our focus on the period immediately following adoption is motivated by practice: more customers abandon the service during the first week than in any other week of their lifetimes, which makes retention during this period critical, and customers' demand for technology support is frontloaded in the sense that they ask most of their questions soon after adoption.

Our work addresses two tensions in the literature. First, in regards to the influence of EPE on customer retention, we note that EPE may foster retention by increasing customers' perceived service quality (Eisingerich and Bell 2008, Sharma and Patterson 1999), setting realistic expectations (Bhattacherjee 2001, McKinney et al. 2002), and aiding customers surpass the initial ramp-up stage (Xue et al. 2007), but it can also make them quickly realize the limitations of the service (Fodness et al. 1993, Nayyar 1990) and consider defecting. We hypothesize the former effect will prevail because (i) customers will derive more value and will be more satisfied from using a service they understand better due to the treatment and, additionally, (ii) even just becoming familiar with the service and learning how to use its basic functionalities already constitutes a co-production skill learned that would be lost if they left. Second, with respect to EPE's influence on customer's demand for (reactive) technology support, EPE may reduce customers' demand for support by making them self-sufficient (Xue and Harker 2002), but it can also increase it by making them provider-dependent (Challagalla et al. 2009). We hypothesize that customers' demand for technology support (i.e., the number of questions they ask through reactive support channels) immediately after adoption will be reduced by EPE because of the provider's ability to preempt the questions customers generally have during this stage of their lifetimes (e.g., the frequently asked questions, or FAQs). Moreover, at this stage customers will not have developed any dependency habits that will lead them to increase their demand for technology support.

To test our hypotheses we collected unique data from a field experiment conducted by the same major public cloud computing infrastructure services provider during October and November 2011. Upon signup, 366 customers selected at random out of 2,673 customers that opened an account during this period received the field experiment's treatment: EPE. The treatment consisted of a short phone call followed up by a support ticket through which the provider offered initial guidance on how to use the basic features of the service. Our empirical strategy leverages the random assignment of the treatment and employs survival analysis and count data models to examine the differences in retention and demand for reactive technology support, respectively, between the two customer groups immediately after service adoption. Our robustness checks thoroughly examine and validate the random assignment assumption.

Regarding customer retention, we find that treated customers' hazard rate (i.e., number of customers who leave the service per unit of time) is about 49.60% lower than that of controls during the first week after adoption. This has a strong managerial implication for the provider since, by improving retention early on when the risk of customer churn is highest, EPE has a positive long term effect on the growth and overall size of the customer base. With respect to customers' early demand for technology support, as measured by the number of questions they ask to the provider through online live chat sessions and support tickets in the weeks following adoption, we find that EPE reduces the average number of questions asked during customers' first week after adoption by 19.55%. This is an important drop in one of the provider's major operational costs: the human labor-intensive offering of reactive technology support.

In this dissertation, in addition to finding answers to our research questions regarding how a provider's technology support and education influence its customers' service use, our focus on the cloud B2B context allows us to contribute to the service co-production literature in other ways as well. For instance, we examine whether service providers help businesses to overcome knowledge barriers at the organizational level, a proposition that has been argued in prior work but which has not been empirically tested (Attewell 1992, Fichman and Kemerer 1999). Furthermore, different from the more commonly studied business-to-consumer (B2C) SST settings (e.g., online banking or retail) that are generally turn-key or ready-to-use solutions (e.g., online banking services), in B2B, service providers face the complex challenge of offering a service amenable to a wide variety of use cases and business needs for a very heterogeneous customer base (Venters and Whitley 2012). Thus, the customers and the provider must invest a significant effort in understanding each other in order to best co-produce the service, as is the case in other B2B services contexts (Bettencourt et al. 2002, Ko et al. 2005).

From the managerial standpoint, this dissertation provides a framework that can be employed by SST providers in a broad range of industries to analyze the data they collect on customer behavior, enabling them to measure how a customers' use of the service is influenced by its access to technology support. This is particular relevant for emerging SSTs, such as some cloud offerings, that are far from a ready-to-use, turn-key solution, but, rather, are more akin to a general purpose technology (Bresnahan and Trajtenberg 1995) that demands considerable adaptation or co-production efforts from customers.

CHAPTER 2

TECHNOLOGY SUPPORT AND IT USE:

EVIDENCE FROM THE CLOUD

2.1 Introduction

Service customers frequently perform actions that are essential to the value they receive from the service. For example, online banking customers must manipulate a web site to obtain the information they need, while in many business IT services contexts the customer must transfer essential information to the provider. This buyer role as both a recipient and producer of services, known as service co-production, plays a key role in determining the quality of service output and use of the service. For instance, in the context of business-to-consumer (B2C) online *self-service technologies* (SSTs) such as those associated with online banking, retailing, or auctions, among others, research has consistently shown that customers' capabilities in co-producing the service are a key determinant of their adoption and continued use (e.g., Buell et al. 2010, Frei 2008, Xue and Harker 2002, Xue et al. 2007).

Nonetheless, the factors associated with customers' capabilities, such as their *knowledge, skills and abilities* (KSAs), have traditionally been considered as given and thus exogenous to the provider (e.g., Xue et al. 2011). With the exception of Field et al. (2012), who highlight the role of face-to-face interactions with a service provider in the B2C setting, little is known on how providers can influence end users' KSAs. Prior work has suggested that the availability of technical support (Morgan and Finnegan 2007) as well as firms' accessibility to external knowledge sources (Li et al. 2005) are important in determining organizations' decisions to adopt a new IT product. However, *post-adoption,* less is known regarding how different types or levels of technology support may influence firms' actual use of a service or product. To our knowledge, no prior work has demonstrated whether a provider's technology support can influence buyers' use of its service.

We aim to take a first step in narrowing this gap in understanding by examining whether provider technology support influences customer use of a particular commodity SST, namely public cloud infrastructure services.[1] Our central research question is: *Does a provider's technology support influence its customers' IT use?* We address these questions in a particular business-to-business (B2B) context. In our setting, provider customers are firms who choose levels of technology support and intensity of use of public cloud computing infrastructure services.

In addition to further probing the impact of provider support on service use, our focus on a B2B context allows us to advance the literature on service co-production in other ways. In particular, we examine whether service providers help businesses to overcome knowledge barriers at the organizational level, a proposition that has been argued in prior work but which has not been empirically tested (Attewell 1992, Fichman and Kemerer 1999). B2B differs fundamentally from B2C in the level of involvement the provider has in co-participating in the service delivery process and the adaptation process that firms must engage in to best use the service. In the B2C context, and in particular in the online banking context examined by Field et al. (2012), there is little heterogeneity in the uses individual customers can give to the service. For example, an online banking portal will allow visitors to perform core banking transactions such as deposits, payments or transfers, and all visitors perform these actions in a mostly standard fashion. Therefore, the provider needs to make a relatively low effort in understanding users' individual needs in order to train them on how to use the service, as the methods used will not differ significantly from one person to another. Very differently, business

[1] Public cloud infrastructure services, or public Infrastructure-as-a-Service (IaaS), are a B2B SST in which on-demand computing and storage resources (i.e., servers) are offered on a pay-as-you-go basis (Mell and Grance 2011).

service providers such as those offering cloud services face the complex challenge of offering a service amenable to a wide variety of use cases and business needs for a very heterogeneous customer base (Venters and Whitley 2012). Firms will vary in terms of their industries, projects, sizes, geographical location of their own customers, and their internal IT capabilities, to name a few variables. Similar to how it has been documented in other B2B services contexts (Bettencourt et al. 2002, Ko et al. 2005), the buyers and the provider must invest a significant effort in understanding each other in order to best co-produce the service.

Additionally, the adoption and usage of cloud infrastructure services can be considered as a process innovation customized to the idiosyncratic context and needs of the customer. Cloud infrastructure services are a true general purpose technology (Bresnahan and Trajtenberg 1995) that must be adapted to each firms use case for it to generate value. In short, delivering business services such as public cloud computing infrastructure services creates unique challenges for service providers that do not exist in the consumer service delivery environment.

In our research setting, the provider's customers use its hardware resources and also choose (and switch between) two levels of technology support, *basic* and *full*. Full support differs from basic support in that when offering it the provider educates buyers and helps them in their service co-production processes, whereas basic support only deals with simple quality of service issues; details about the service offering are described later in section 2.3.2. We first develop a parsimonious service co-production model that examines a buyer's tradeoffs when choosing between basic/no support and full support, and the corresponding optimal use. The insights from the analytical model are used to motivate our hypotheses.

To test our hypotheses we collect unique data from a major global public cloud infrastructure services provider. Our rich data consist of 22,179 firms that used the provider's public cloud infrastructure service at some point between March 2009 and

August 2012. Our econometric approach uses fixed effects panel data models and a difference-in-difference identification strategy to compare buyers' use of the service before adopting full support, during their continued access to full support, and after switching from full to basic support. We find that buyers who adopt and continue having access to the provider's full support use, on average, 188% more of the service relative to customers who have only had access to basic support, indicating that the business value of technology support is very significant. Furthermore, we find evidence that full support customers continue to use more of the service even if they downgrade to a lower level of support. Former full support customers continue using, on average, 77% more of the service compared to those who have never accessed full support.

Our difference-in-difference identification strategy assumes that unobserved factors are changing in the same way for buyers who adopt full support (treated) as for those that never do (controls), and if this assumption is violated then our estimates of the causal effects of full support adoption on service use become inconsistent. A particular worry is reverse causality, i.e., the support choice decision may follow IT use. To address this concern, we conduct several falsification tests, include additional controls, and perform the following additional analyses. First, we run our models employing matched subsamples that are constructed using a coarsened exact matching (CEM) procedure (Blackwell et al. 2010) based on buyers' usage of the service before they upgrade from basic to full support. This supports the premise that buyers do not exhibit differential behavior before the treatment. Second, we leverage detailed data on buyers' support interactions. For example, we use online live chat sessions and support tickets as the basis for instruments for buyer decisions to upgrade to full support. The rationale for this instrument is that the occurrence of failures, and more importantly the support interactions that take place between buyers and the provider when working to overcome the problems they cause, can serve as a signal to buyers for the value of full support and so will increase the likelihood of upgrading. Third, we also use lagged values of our variables as instruments (Arellano and Bover 1995, Blundell and Bond 1998) in dynamic panel models using a generalized method of moments (GMM) estimation approach. We augment this latter approach with our support-based instruments. The estimates across these various subsamples and models are qualitatively consistent with our main findings.

We also investigate whether certain firm characteristics are complementary with the adoption of full support and IT use. Specifically, we focus on the role of firm size, a measure that has been shown to be correlated with technical sophistication (e.g., Rogers 1995). By interacting our measure of full support with buyer employment, we show that larger firms exhibit a greater marginal increase in their use of the service from adopting and having access to full support. They also continue to use more of the service than smaller firms after they opt to switch to basic support.

Last, we provide further evidence on the impact of technology support on IT use by examining alternative measures of infrastructure use. Specifically, we provide additional evidence that technology support helps buyers make better and more efficient use of the service by quantifying the effects that full support has on buyers' likelihood of deploying horizontally distributed and scalable architectures.

Given the massive number of firms in our data, our study provides a framework than can be employed by SST providers in a broad range of industries to analyze the big data they collect on buyer behavior, enabling them to measure how a buyer's use of the service is influenced by its access to technology support. We have worked closely with the provider's business analytics team and used our models to offer guidance and rigorously quantify the influence of their premium technology support on buyer use of the service and provider revenue. Moreover, we have also demonstrated to the provider how, by automatically parsing the content of buyers' support interactions (i.e., live chat sessions and support tickets), it can make inferences of otherwise unobserved time-varying factors that influence buyer behavior. Finally, we also developed and automated the computation of a cloud-specific metric useful to assess buyers' capabilities in exploiting the service's features, further allowing the provider to understand the impact of technology support on buyers' service co-production efforts.

2.2 Theory Development and Hypotheses

Research in the B2C service co-production setting has suggested that educating customers is an appropriate strategy for providers when the complexity of the service is high (Burton 2002). Moving to the B2B context, similar propositions have been made yet not empirically tested in the knowledge-intensive business services industry (e.g., IT consulting and software design), where clients' co-participation in service delivery is indispensable and their training and education is an important element needed to ensure successful outcomes (Bettencourt et al. 2002). In our setting, the provider attempts to educate its customers on how to best co-produce the service by offering them full support. In what follows, we present a parsimonious model that demonstrates how additional support may influence service consumption by improving a buyer's productivity.

2.2.1 Motivating Model

We assume that there is a continuum of heterogeneous buyers with type $\theta \sim U(0,1)$. One can think of this type as the size or technical sophistication of the buyer firm. Each buyer seeks to source a service from a provider on a per-period basis as an input to produce its own products or services. The provider offers two levels of services, $s \in \{b, f\}$, one without support, b, at the price of p_b, and the other with support, f, but at a premium price $p_f > p_b$ plus a fixed fee $F \geq 0$ per period. We assume p_b is the spot market price for the commodity service that is competitively determined by the market, and we further assume the provider sets p_f such that it reflects its marginal cost of providing support. In each period, the buyer decides the support level (s) and service consumption volume $q \geq 0$. The production function of each buyer firm takes a simple

Cobb-Douglas form, i.e., $\theta^{1-z}q^{z}$, being determined by its type (θ), consumption volume (q), and co-production output elasticity, $0 < z < 1$, which is jointly determined by the provider and the buyer. Here, we adapt the standard Cobb-Douglas form employed in the service co-production literature (e.g., Xue et al. 2007) to our B2B setting.

A fundamental assumption in our model is that buyers who currently opt for full support enjoy a higher co-production output elasticity than those who never opted for full support, i.e., $z_f > z_b$, where z_f and z_b represent the corresponding co-production elasticities for full and basic support, respectively. We argue that this is a highly plausible scenario because the provider-buyer interactions that take place when full support is received enable more efficient learning to occur. Similar learning through interactions between consultants and clients have been documented in other B2B services contexts (e.g., Bettencourt et al. 2002, Ko et al. 2005). Furthermore, if customers retain over time the same support level as the one chosen upon initial adoption, their co-production output elasticity will correspond to the chosen support level, i.e., $z = z_s$. If they upgrade from basic to full support, they also upgrade to $z = z_f$. However, if they had full support in the past but they downgraded to basic support (denoted as $f2b$), past learning from the provider allows buyers to operate under co-production elasticity $z = z_{f2b} \in (z_b, z_f]$, in spite of their current support level being b. We restrict our analytical exercise to an illustrative setting where $\max\left\{1, \frac{z_b}{p_b}\right\} < \frac{z_{f2b}}{p_b} < \frac{z_f}{p_f}$.

All buyers try to maximize their instantaneous utility. More precisely, a buyer of type θ solves the following constrained optimization problem:

$$\max_{q\geq 0, s\in\{b,f\}} u(\theta|q,s,z) = \theta^{1-z}q^{z} - p_s q - 1_{\{s=f\}}F,$$

where $1_{\{s=f\}}$ is the indicator function that captures the fact that the two-part tariff scheme occurs only under full support and z is either z_s or z_{f2b} depending on currently chosen support level and past usage of support, as previously discussed.

If at a given time the customer decides to adopt (or continue to use) the service (i.e., can get positive utility from it) and chooses support level s and enjoys co-production elasticity level z, her optimal use (i.e., consumption volume) is given by

$$q^*(\theta|s,z) = \theta\left(\frac{z}{p_s}\right)^{\frac{1}{1-z}}.$$

It follows that customers initially prefer full support if their type is sufficiently high, i.e., $\theta \geq \hat{\theta} = \frac{F}{(1-z_f)\left(\frac{z_f}{p_f}\right)^{\frac{z_f}{1-z_f}} - (1-z_b)\left(\frac{z_b}{p_b}\right)^{\frac{z_b}{1-z_b}}}$. We consider scenarios where $0 < \hat{\theta} < 1$ and focus exclusively on customers who choose full support in the beginning, i.e., $\theta \geq \hat{\theta}$. The following hypotheses are directly motivated from our model; they all hold true under our aforementioned assumptions (detailed proofs are included in the appendix).

2.2.2 Technology Support and IT Use

Das (2003) mentions that "for high-technology vendors, technical support is not only a competitive necessity, but also a potential source of revenue in markets where profits from product sales are increasingly restricted by price competition", suggesting support can be used as a mechanism to influence demand for commoditized or weakly differentiated services such as cloud infrastructure services. Our model-based hypothesis is that, given a buyer, her optimal use (i.e., service consumption volume) will be greater if she opts for support rather than for no support. Formally, our model predicts that,

HYPOTHESIS 1: $q^*(\theta|s=f, z=z_f) - q^*(\theta|s=b, z=z_b) \geq 0, \forall\, \theta \geq \hat{\theta}$.

This hypothesis implies that, all else equal, buyers who adopt and have continued access to full support use more service compared to similar buyers who only have access to basic support.

2.2.3 Organizational Learning

In addition to deciding to adopt the provider's full support, buyers can also choose, in the future, to drop it and switch from full to basic support. A potential reason why customers might switch to basic support is their learning from the provider which, in turn, enables them to achieve similar productivity as those enjoyed by full support customers, but on their own, without the need of interacting intensely with the provider through support and without paying the corresponding price premiums and fixed fees. This is consistent with prior research that has shown that once firms internalize knowledge available from external sources, their valuation of that external knowledge decreases relative to their valuation of their own internal knowledge (e.g., Menon and Pfeffer 2003).

A key assumption underlying this process is that buyers will not forget what they have learned, or at least not so quickly. We argue that the implementation of projects that have a direct impact on buyers' internal business processes, such as those in the professional services industry (e.g., consultancy) or the adoption and usage of IaaS, can be characterized as a process innovation customized to the idiosyncratic context and needs of the customer (e.g., Hitt et al. 2002). In such innovations, not forgetting is vital for continued success, and extant research has found that organizational forgetting rates in this context are near zero (Boone et al. 2008).

We conjecture that if former full support buyers (who switched to basic support) have learned from the provider and do not quickly forget that knowledge, then, given the lower prices of basic support, they will use more service than other basic support customers who have not had the opportunity of learning from the provider. In other words, when buyers can achieve productivity equal or at least similar to those of full support ($z = z_{f2b} \in \left(z_{b,} z_f\right]$) when consuming at basic support prices (p_b), they will end up using a higher volume of the input service than buyers who have only accessed basic support in the past. Formally, our model predicts that,

HYPOTHESIS 2: $q^*\big(\theta|s=b, z=z_{f2b}\big) - q^*(\theta|s=b, z=z_b) \geq 0, \forall\, \theta \geq \hat{\theta}$.

2.2.4 Firm Size and Organizational Learning

A recurring result in the literature is that firm size is correlated with the speed of new technology adoption and assimilation (Rogers 1995). Among other reasons, it is understood that larger firms have more slack resources, greater economies of scale, higher levels of professionalism, and easier access to external resources (e.g., Attewell 1992, Fichman 2000, Fichman 2001, Forman 2005), all of which enable them to adopt new technologies faster. Given their higher level of IT sophistication, one might expect that large firms' marginal benefits from having access to the provider's technology support would be low, as they have little to gain. Nonetheless, such a view would overlook larger firms' greater ability to co-produce the service jointly with the provider.

Larger firms often have greater levels of technical sophistication and related knowledge than smaller firms. Such knowledge will facilitate the absorption of new (but related) knowledge needed to innovate successfully (Fichman 2001). In other words, larger firms have a greater absorptive capacity – defined as "the ability of a firm to recognize the value of new, external information, assimilate it, and apply it to commercial ends" (Cohen and Levinthal 1990), that will enable them to better communicate and co-produce the service with the provider through technology support. We thus propose that the benefits of adopting and having access to full support should be stronger for larger rather than smaller firms. Specifically, as predicted by our model, the greater a buyer's size, the greater its service use increase associated with the adoption of and continued access to full support:

HYPOTHESIS 3: $\frac{\partial\left(q^*\left(\theta|s=f, z=z_f\right) - q^*\left(\theta|s=b, z=z_b\right)\right)}{\partial\theta} > 0, \forall\, \theta \geq \hat{\theta}$.

Additionally, if Hypothesis 2 holds, whereby buyers do not forget quickly what they have learned from the provider, and also per Hypothesis 3 larger buyers are able to keep their service co-production costs lower than smaller customers buyers after they switch to basic support, we expect that larger basic support buyers who have accessed full support will use more services than smaller ones who have also accessed full support in the past. Formally, our model predicts that,

HYPOTHESIS 4: $\frac{\partial\left(q^*(\theta|s=b,z=z_{f2b})-q^*(\theta|s=b,z=z_b)\right)}{\partial\theta} > 0, \forall\ \theta \geq \hat{\theta}.$

2.3 Research Setting

2.3.1 Cloud Computing Public Infrastructure Services

Cloud computing has been defined by the US National Institute of Standards and Technology as a "model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction" (Mell and Grance 2011)[2]. The pay-as-you-go nature of the service along with its rapid elasticity provides firms the opportunity to reduce idle computing capacity waste and eliminate the necessity of an up-front capital commitment in overprovisioning resources (Armbrust et al. 2010, Harms and Yamartino 2010). It has been envisioned by some scholars as a *general purpose technology* (GPT) (Bresnahan and Trajtenberg 1995) that will serve as a catalyst for innovation and an engine for economic growth (e.g., Brynjolfsson et al. 2010, Varian 2010, Varian 2011). Nonetheless, the current slow adoption rates of cloud infrastructure services do not reflect such expectations. Surveys have suggested that only 29% of small and medium-sized businesses (SMBs) were paying for one or more cloud services in 2010 (Microsoft and Edge Strategies 2011) and that in 2011 only 4% of IT professionals had implemented cloud infrastructure services for production applications (SearchDataCenter.com 2011).

---

[2] This constitutes the 16th and final version of "The NIST Definition of Cloud Computing."

More recently, InformationWeek found that the number of firms receiving services from a cloud provider only grew from 16% in 2008 to 33% in February 2012 (Wittmann 2012).

A potential reason for this slow adoption is that these services are not offered as fully outsourced, turn-key and ready-to-use solutions for firms. Rather, the self-service nature of the cloud requires firms to co-participate (Bitner et al. 1997) in the service delivery process. In the particular case of Infrastructure-as-a-Service (IaaS) offerings, the setting that we study, "the consumer does not manage or control the underlying cloud infrastructure but has control over operating systems, storage, and deployed applications; and possibly limited control of select networking components" (Mell and Grance 2011). In other words, cloud infrastructure services are a high contact service (Bitner et al. 1997, Chase 1978) in which the value derived from the service depends to a great extent on buyers' own capabilities and their own service co-production efforts.

Additionally, a 2011 survey found that only 25% of IT staff in global organizations had cloud experience with public infrastructure or platform-as-a-service, and 50% of the organizations claimed that their staff was "less than somewhat prepared to handle" these services (Symantec 2011). This suggests that most buyers are not well prepared to co-produce cloud services and that helping them overcome their co-production costs may be vital for the success of the cloud model. Together, the need for customers' co-participation in the service delivery process and the presence of significant adaptation costs make cloud infrastructure services an ideal context to test if technology support influences use of IT services.

2.3.2 Description of Service Offering

In our particular setting, the provider has recognized that the novelty of the service plus the complexities involved in deploying distributed architectures that best leverage the cloud's scalability may pose significant knowledge barriers to buyers attempting to use the service. In response to this, the cloud provider offers them the option to contract and access full support. We discuss first the pricing and terms of the infrastructure service offering, and then elaborate on what characterizes full support.

The provider's offering adheres tightly to NIST's definition of cloud computing (Mell and Grance 2011). As per NIST, one of the essential characteristics of cloud services is that they are on-demand. Buyers only pay for what they use, and nothing else: there are no sign-up fees, no minimum spending requirements, no periodical subscription fees and – since buyers can choose to not to use their service as well – there are no contract termination penalties either. Moreover, in the particular case of our provider, the computing resources are offered to buyers at fixed hourly rates that increase in server size or capacity, generally in a linear fashion. Servers' capacity is defined in terms of memory (GB of RAM), processing power (number of virtual CPUs), and local storage (GB space of local hard disk). The 3 parameters tend to vary together as a bundle, meaning that more of one is generally associated with more of the other two, yet prices are set and buyers usually make infrastructure sizing decisions in terms of memory.[3]

Another feature of the cloud is its rapid elasticity, whereby for buyers, "the capabilities available for provisioning often appear to be unlimited and can be appropriated in any quantity at any time" (Mell and Grance 2011). Buyers in our context can launch as many servers and of any size they want, when they want. There are no usage caps, with the only exceptions to this being that the provider may have limited hardware installed at its data centers or may take security measures to prevent misuse of its service (e.g., spamming). In other words, for legit buyers, there is no pre-defined cap

[3] For example, a buyer may pay $0.10 per hour to run a 2GB RAM server, and $0.20 per hour to run a 4GB RAM server, in both cases paying $0.05 per GB of RAM; these rates are fictitious, but are very similar to actual prices in the market. Servers of larger sizes (e.g., 30GB RAM) may have marginally lower per GB RAM rates (e.g., $0.04 per GB of RAM, so $1.2 per hour for 30GB RAM server), but not enough to be considered a volume discount, which are not available in any form. Also, due to licensing fees, servers running Windows or RedHat are slightly more expensive than those running Linux (e.g., $0.07 per GB of RAM). Finally, rates per GB of RAM of servers were fixed and did not change during our sample period.

or limit to how much they can choose to use the service.

The provider complements its infrastructure offering with full support, which is offered for a fixed price premium per server-hour used plus an additional fixed monthly fee.[4] The monthly fee is paid as a monthly subscription, which is a fee high enough to deter buyers with very low willingness to pay (i.e., bloggers that use a single very small server). There are no sign-up or termination fees for the full support service. The only explicit switching cost from one support level to another is technical rather than monetary: When downgrading from full support to basic support, because of technical limitations in the service offering, buyers must redeploy their servers on their own under the new support regime. The redeployment will involve launching new servers with virgin operating systems (i.e., "out of the box"), and then installing and configuring their business applications on them.

A prime goal of full support is to educate buyers on how to best use the cloud infrastructure service and adapt it to their idiosyncratic business needs. When receiving full support, buyers receive personalized guidance and training, and thus have the opportunity to learn directly from the provider's prior experience in deploying applications in the cloud. Buyers not willing to pay the price premiums will only receive a basic level of support which has limited scope in the sense that it is intended to aid buyers with issues concerning account management or overall performance of the infrastructure service. For example, while a full support buyer may be personally guided step by step on how to deploy a web server through phone conversations, live chat sessions or support tickets, basic support buyers will be referred to a knowledge base. Similarly, if a server failed, which happens much more frequently than in traditional datacenter settings given the commodity hardware employed and the multi-tenant

[4] For instance, using the same examples as in footnote above, instead of paying $0.10 per hour for a 2GB RAM server under basic support, a full support buyer would pay $0.12 more, i.e., $0.22 per hour. For the 4GB RAM server the full support buyer would pay $0.32 instead of $0.20 per hour.

architecture (i.e., multiple organizations' virtual servers are hosted in the same and shared physical server), the provider would work together with full support buyers in solving the issues, while basic support users would only be notified about the failure, if anything. Thus, basic support customers do not have fluid access to external knowledge from the provider and have to rely mostly on their internal capabilities to co-produce the service.

2.4 Empirical Model

2.4.1 Difference-in-Difference Fixed Effects

We employ linear fixed effects panel data models along with a difference-in-difference identification strategy to tease out the effects of the adoption of and continued access to full support on cloud use. The pay-per-use model provides cloud infrastructure buyers the freedom to pay only for the computing resources they consume. Since, as we just mentioned, the servers are priced based on the amount of memory they have, and memory is the basis for buyers' infrastructure sizing decisions, the amount of memory consumed by buyers over time is a direct measure of their use of cloud services. We compute the average GB of RAM used by a buyer per month and employ it as our dependent variable, which we call $Memory_{it}$. Given that the distribution of memory (servers) usage has a strong positive skew and that at times buyers may not consume any memory, we use the log of memory plus 1 ($lnMemory_{it} = ln(Memory_{it} + 1)$) as our dependent variable.

Our first model tests if the adoption or the prior access to full support is associated with greater memory use:

$$lnMemory_{it} = \alpha + \beta\ FullSupport_{it} + \gamma\ SwitchToBasic_{it} + \mu_i + \tau_t + l_{it} + \varepsilon_{it}. \quad (1)$$

Subscripts i and t index individual buyers (firms) and time periods (months) respectively. Parameter μ_i is the buyer fixed effect and τ_t is a vector of month fixed effects. We also include a vector of dummy variables, l_{it}, indicating in what month of its lifetime a buyer is when month t starts. This allows us to control for the possibility that buyers' use of the service may increase in a nonlinear fashion as they grow older and learn more about it. Parameter ε_{it} is our error term which we assume is correlated only within individual firms, but not across them.

$FullSupport_{it}$ is a binary variable that indicates if full support was adopted by customer i by time t. Thus, β identifies the effects on cloud use of adopting and having access to full support, and we expect it to be positive and significant per Hypothesis 1. $SwitchToBasic_{it}$ is a binary variable that is equal to one if the buyer does not have access to full support by the end of the focal month but was using full support at the start of the focal month or in some prior month(s). The γ coefficient identifies the durability of the effects of full support. If they are durable, then γ will be insignificant (suggesting behavior does not change) or negative and significant but with a low value relative to β (suggesting the effects of full support do not dissipate entirely). $\beta + \gamma$ will measure differences in use behavior between basic support buyers who accessed full support in the past and those who exclusively accessed basic support. If Hypothesis 2 holds, and buyers' prior access to full support sets them apart from those who only used basic support, then $\beta + \gamma$ should be positive and significant.

Our fixed effects model allows us to difference out unobserved time-invariant buyer-level heterogeneity that may influence both the choice of support type and IT use. Like any difference-in-difference model, our estimates rely on the identifying assumption that unobserved factors influencing use change similarly for full support adopters (treated) and non-adopters (controls) over time. We explore the validity of this assumption by running our models using matched subsamples constructed using a coarsened exact matching (CEM) procedure (Blackwell et al. 2010). Employing matching procedures reduces the dependence of our estimates on our model specification and also reduces endogeneity concerns when making causal inferences (Ho et al. 2007). As described in further detail below, we match firms based on their pre-upgrade memory consumption levels, their pre-upgrade frequency of infrastructure resizing (i.e., number of changes in their total memory use), and their intended use cases for the cloud service, industry, and size.

Further, we use exogenous failure events experienced by buyers as an instrument for their support choice decision. When this type of problem occurs, the support interactions that take place between buyers and the provider when working on overcoming them can serve as signal to buyers for the value of full support. Basic support buyers who, because of the failure, obtain experience in co-producing the service with a greater involvement from the provider, are more likely to upgrade to full support than buyers who do not have such experiences with the provider. However, such interactions on their own are unlikely to increase use of the provider's service.

Following the approach suggested by Wooldridge (2007) and Angrist and Pischke (2009), we employ a probit model that has the exogenous failures as regressors to generate predicted values for $FullSupport_{it}$, which we denote $FullSupport_{it}^{f}$. We then use the fitted value, $FullSupport_{it}^{f}$, as our instrument in a standard two-stage least squares (2SLS) estimation. Given that it is very hard to instrument for the downgrading support choice decision, captured in the $SwitchToBasic_{it}$ parameter in Model (1), for our second stage we use a simplified version of this model that excludes such parameter:

$$lnMemory_{it} = \alpha_1 + \beta_1\ FullSupport_{it} + \mu_i + \tau_t + l_{it} + \varepsilon_{1it}. \tag{2}$$

The parameter β_1 in Model (2) will have a slightly different interpretation than parameter β in Model (1), since here β_1 will identify the average memory usage at any point after upgrading from basic to full support, regardless if buyers downgrade afterwards or not.

Two additional concerns remain. First, buyers' likelihood of suffering from an exogenous failure increases with the number of servers they employ, which undermines the random assignment of the instrument. Second, there may be persistence in memory use levels, such that buyers' use in prior periods may strongly influence their use in the focal period. We first address both of these concerns by including lagged values of the dependent variable as a regressor and using standard fixed effects model. However, such an approach suffers from dynamic panel bias: it fails the strict exogeneity assumption necessary for consistent estimates in fixed effects models (Nickell 1981, Roodman 2009a). Although this bias decreases in the number of periods (Nickell 1981), and we have a long panel with $T = 42$, the bias remains a concern.

A solution to this issue involves using the System GMM and Difference GMM approaches that have evolved from the work of Anderson and Hsiao (1981), Arellano and Bond (1991), Arellano and Bover (1995) and Blundell and Bond (1998), and have seen increasing use in applied work in the management literature (e.g., Archak et al. 2011, Ghose 2009). This approach has the important added benefit that it allows us to treat $FullSupport_{it}$ as endogenous, and control for at least some unobserved time-varying factors by using $lnMemory_{it}$ and $FullSupport_{it}$'s lagged values and differences as instruments. We employ System GMM (Arellano and Bover 1995, Blundell and Bond 1998) in conjunction with the finite-sample correction proposed by Windmeijer (2005). Moreover, we also augment the instruments matrix with additional parameters based on the exogenous failures. We elaborate on our selection of the number of lags of the dependent variable included as regressors and the number of instruments used in our results section.

2.4.2 Role of Firm Size

In order to examine the role of buyer size, we model it using the total number of employees at the firm. We use 3 different variables for this: (1) a binary indicator that is turned on if the buyer is above the median employment $(EmploymentTop50_i)$, (2) a binary variable indicating if the buyer is in the top 25th percentile of the employment distribution ($EmploymentTop25_i$), and (3) the log of the number of employees

$(lnEmployment_i = ln(Employees_i))$. We only present the first variable in model below and use the remaining two for robustness checks:

$$
\begin{aligned}
lnMemory_{it} = \alpha + \beta_1\, FullSupport_{it} \quad &+ \beta_2\, FullSupport_{it} \times EmploymentTop50_i \\
&+ \gamma_1\, SwitchToBasic_{it} + \gamma_2\, SwitchToBasic_{it} \times EmploymentTop50_i \\
&+ \mu_i + \tau_t + l_{it} + \varepsilon_{it}
\end{aligned} \tag{3}
$$

To test our third hypothesis, which argues that the benefits of full support will be stronger for larger firms, we interact our employment measures with dummies for the adoption of and switch from full support. If this hypothesis holds then the coefficient β_2 should be positive and significant. Similarly, if as per Hypothesis 4 larger buyers are able to keep their service co-production costs lower than smaller buyers after the switch to basic support, then $\beta_2 + \gamma_2$ should be positive and significant as well.

2.5 Data and Sample Construction

2.5.1 Data

We have collected a unique data set on cloud infrastructure services use from a major public cloud provider. Our entire data set includes 79,619 customers/buyers that adopted the provider's services at some point between March 2009 and August 2012. Buyers can freely choose if they rely only on the provider's basic support or if they pay additional fees to receive full support. They can also switch from one type of support to another, and we observe when such switching occurs.

We exclude buyers who use the service very little or who do not change their cloud architecture configuration (i.e., do not resize their infrastructure).[5] Our identification assumption is that changes in use behavior over time are very similar

[5] Specifically, we exclude buyers who (1) only accessed basic support and (2) averaged 512 MB RAM/hour or less during their first 6 months (excluding 1st month) or (3) made no adjustments to size of their infrastructure during their first 6 months (excluding 1st month). An infrastructure resizing occurs in any launching, halting, or resizing of a server in the buyers' cloud infrastructure. We do not consider their behavior during their 1st month in our threshold because most buyers are setting up their infrastructure during this time.

between basic support buyers and future full support buyers, before the latter upgrade from basic to full support. The excluded set of buyers has very different time-varying profiles and, although we exclude them ex ante, they likely would also be excluded later by our CEM procedures. This intuition was captured in our motivating model, where the lower-type customers $(\theta < \hat{\theta})$ would never opt for full support. A total of 57,440 customers are dropped from the sample as a result of this procedure, though our results are robust to their inclusion (see appendix for descriptive statistics of the initial full sample as well as results of regressions using all buyers in the data).

Among the remaining 22,179 buyers in our baseline sample, 16,157 relied exclusively on basic support, 1,611 upgraded from basic to full support, of which 203 downgraded back, and 4,411 started off employing full support, of which 215 eventually downgraded too. The sample includes 368,606 buyer-month observations. Table 2.1 provides descriptive statistics of the cloud use time-varying parameters in our baseline sample; we will describe our second dependent variable $FractionParallel_{it}$ later in section 2.6.4, but include it here for completeness. Table 2.1 also shows that statistics contingent on buyers' support choice $(FullSupport_{it})$; difference in means t-tests for all parameters are significant at the 1% level.

Table 2.1: Descriptive Statistics of Time-Varying Variables (Baseline sample, N=368,606)

Support Type Used Observations	Full or Basic 368,606				$FullSupport_{it} = 0$ 309,544				$FullSupport_{it} = 1$ 59,062			
Variable	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max
$Memory_{it}$	7.88	31.37	0	2,284.54	7.26	30.92	0	2,284.54	11.11	33.41	0	1,917.40
$lnMemory_{it}$	1.348	1.040	0	7.734	1.296	1.008	0	7.734	1.621	1.152	0	7.559
$FullSupport_{it}$	0.160	0.367	0	1	0	0	0	0	1	0	1	1
$SwitchToBasic_{it}$	0.008	0.089	0	1	0	0	0	0	0.050	0.217	0	1
$FractionParallel_{it}$	0.121	0.266	0	1	0.120	0.264	0	1	0.130	0.276	0	1

In addition to the buyers' cloud use data, we have also collected data on the timing and content of all support interactions, i.e., online live chat sessions and support tickets, between the buyers and the provider, starting from October 2009. We offer their detailed description later on.

Finally, we have collected data from a survey administered to buyers upon signup of a new account. The survey is optional and administered as part of the online signup web form; the response rate is 43.4%, and we have not found systematic differences between respondents and non-respondents. The survey was first administered in June 2010, and we have all buyers' responses until February 2012. Although there can only be one survey response per account, since buyers can have multiple accounts, we may also have multiple responses per buyer. In our data we have 6,152 survey responses from 5,565 different buyers in the baseline sample, 431 of which changed their response to at least one item across their surveys. However, for 42.3% of the buyers with varying responses the time gap between the survey responses is too short (i.e., less than 3 months) as to suggest that the variance is due to changes in firms' sizes or goals. Given this, we do not rely on variance across responses for our analysis and rather only consider the 5,134 buyers that either have a single survey response or that have consistent responses across all their submissions. Further, we have not considered firm attributes in the survey as controls in our models since they do not vary over time and thus would be absorbed by the firm fixed effect. We use 3 of the items in the survey: the firms' total employment, their intended use case for the cloud infrastructure service, and their industry.

We use the measure of employment for Model (3); the survey asks buyers to indicate their range of employment and we convert the survey's ranges to numerical values by taking the mean value of each range (e.g., we convert "From 51 to 100" to 75). Descriptive statistics of employment and some of the categories used for our subsample matching procedures are shown in Table 2.2. In what follows we describe how we use these and other cloud use parameters in our subsample matching procedures.

2.5.2 Coarsened Exact Matching

As mentioned in our econometrics approach, we run our models on subsamples defined using a coarsened exact matching (CEM) procedure (Blackwell et al. 2010, Iacus et al. 2012). We consider buyers who adopted full support at any point in their lifetimes as treated, and those that relied exclusively on basic support as controls. As the extensive literature in matching points out, one goal of matching treated and control firms is to reduce endogeneity concerns (Ho et al. 2007). CEM has been used extensively in recent work to improve the identification of appropriate control groups in difference-in-difference estimation (e.g., Azoulay et al. 2011, Azoulay et al. 2010, Furman et al. 2012).

The main idea behind CEM is to temporarily coarsen each matching parameter into meaningful groups (e.g., ranges of memory usage), generate an exact match on the coarsened data, and then retain the original (un-coarsened) values of the matched data (Blackwell et al. 2010). CEM is particularly convenient for our setting because it is a nonparametric procedure that does not require the estimation of propensity scores. This is useful because, aside from the exogenous failures, we have limited data that would allow us to directly predict the likelihood of full support. Each unique vector formed by combinations of the coarsened covariates describes a stratum, such that each firm is assigned to a unique stratum, and only observations in strata where there are at least one treated and one control firm are retained and used in posterior analysis. Since the number of treated and control observations in each strata may be different, observations are weighted according to the size of their strata (Iacus et al. 2012). The differences in means between the treated and the controls across the various matching parameters shown in Table 2.2 are almost all statistically significant. However, the samples are perfectly balanced and any mean differences are eliminated once we apply the CEM weights, as shown in Table C.4 of our appendix. When exact matching is possible, such that for every treated observation there is a control observation identical to the first one across all possible covariates except for the treatment, a simple difference in means of the

Table 2.2: Descriptive Statistics of Parameters used for CEM before matching (5,134 buyers)

Buyer Role	All buyers				Controls				Treated			
Number of Buyers	5,134				3,875				1,259			
Variable	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max
$Employees_i$	195.7	1,102.4	2	10,000	164.7	1019.9	2	10,000	291.0	1320.3	2	10,000
$lnEmployment_i$	2.402	1.706	1.099	9.21	2.26	1.608	1.099	9.21	2.838	1.914	1.099	9.21
$EmpCat1_i$	0.656	0.475	0	1	0.692	0.462	0	1	0.546	0.498	0	1
$EmpCat2_i$	0.198	0.398	0	1	0.187	0.390	0	1	0.230	0.421	0	1
$EmpCat3_i$	0.050	0.218	0	1	0.044	0.204	0	1	0.071	0.256	0	1
$EmpCat4_i$	0.037	0.188	0	1	0.030	0.171	0	1	0.056	0.231	0	1
$EmpCat5_i$	0.060	0.237	0	1	0.047	0.213	0	1	0.097	0.296	0	1
UC_HU_i	0.463	0.499	0	1	0.469	0.499	0	1	0.447	0.497	0	1
UC_LU_i	0.591	0.492	0	1	0.573	0.495	0	1	0.647	0.478	0	1
UC_BO_i	0.189	0.391	0	1	0.195	0.396	0	1	0.169	0.375	0	1
UC_HO_i	0.092	0.289	0	1	0.093	0.290	0	1	0.088	0.284	0	1
UC_TD_i	0.293	0.455	0	1	0.323	0.468	0	1	0.203	0.402	0	1

dependent variables would provide an estimate of the causal effect of interest. Nonetheless, since it is nearly impossible to use exact matching in observational data and thus there is always a concern about the influence of omitted variables, we continue using our fixed effects panel data model to control for them.

We match buyers based on 5 attributes: (1) level of IT use (i.e., memory use), (2) frequency of cloud infrastructure resizing (i.e., how often buyers launch a server, halt a server, or resize an existing one), (3) employment, (4) intended use case for the cloud infrastructure service, and (5) industry. For the matching process, we only consider treated buyers who started using the cloud service with basic support and upgraded to full support later on. This allows us to match the upgraders to controls based on their usage behavior before they adopted full support, had the controls adopted full support in the same month of their lifetime interactions with the provider. This approach, which is similar to the one implemented by Azoulay et al. (2010) and Singh and Agrawal (2011), ensures treated firms do not exhibit differential usage behavior before they adopt full support relative to controls.

IT Use and Frequency of Infrastructure Resizing. In regards to overall use (i.e., memory use) and frequency of infrastructure resizing, when creating our baseline sample we had already discarded basic support users with very small and/or rather static deployments. Nonetheless, even among the remaining buyers there is considerable variation in these two parameters. For average memory usage, we set our cutoff points at standard server sizes: 512MB, 1GB, 2GB, 4GB, 8GB, 16GB, 32GB and 64GB of RAM. For frequency of infrastructure resizing we base our cutoff points on percentiles of the distribution: the 25th percentile is a single change to the size of the deployment, the 50th percentile is 3 changes, the 75th percentile is 9 changes, and the 95th percentile is 43 changes. In total, we have 9 categories of memory usage and 5 categories of frequency of infrastructure resizing to match on.

Employment. The employment, intended use case, and industry data are all collected from the signup survey. For the employment cutoff points, we broadly rely on the ranges used in the survey. Among the customers with consistent survey responses across all their accounts, 66% indicated they have 10 or fewer employees ($EmpCat1_i$), so we use 10 as our first cutoff point. Another 20% indicated they have between 11 and 50 employees ($EmpCat2_i$), making this our next cutoff point. We subdivide the remaining 15% of customers in three bins each accounting for roughly 5% of our sample: from 51 to 100 ($EmpCat3_i$), from 101 to 250 ($EmpCat4_i$), and greater than 250 ($EmpCat5_i$). Detailed descriptive statistics of each category (e.g., $EmpCatN_i$) are shown in Table 2.2.

Intended Use Case. The intended use case is collected by a multiple choice question (i.e., "Mark all that apply") that asked customers to "Please indicate what solution(s) you intend to use [the cloud infrastructure service] for." The 20 options available to buyers are very specific, and finding matches across such specific use cases would be extremely hard. Instead, we group the specific use cases into 3 more general use cases based on two dimensions: if the use case is related to back office or front office applications, and, in the latter case, if it is likely that the volume of usage for the use case is predictable or not. Our first general use case, which we call “High Usage Uncertainty” (UC_HU_i), includes customer-facing websites that are prone to unpredictable variance in their volume of usage. Examples of such use cases are social media sites, online gaming sites, online publishing sites, rich media sites (e.g., audio or video), and other Software-as-a-Service (SaaS) offerings. Our second general use case, “Low Usage Uncertainty” (UC_LU_i), includes customer-facing websites used for regular operation of the firm that have steady or at least predictable use levels. Examples are corporate websites, collaboration platforms, online portals, and e-commerce sites. We chose to include e-commerce sites in this general use case since, although it may have a high variance, seasonality makes the peaks and valleys of the demand fairly predictable. Finally, our “Back Office Applications” general use case (UC_BO_i), includes applications or systems used internally for business operations. Examples are a company’s intranet and systems used for accounting, customer relationship management, human resources, supply chain management, or backup. We additionally consider web hosting services (UC_HO_i) and running test and development environments ($\mathrm{UC}\_TD_{\mathrm{i}}$) as additional general use cases. Altogether, we have 5 general use cases, and the proportion of firms that marked each of them is shown in Table 2.2.

Industry. Finally, we incorporate an additional question on buyers’ industries in the survey to make an even more stringent match of treated buyers to controls. Although the survey item does not follow any standard industry categorization (e.g., NAICS or SIC codes), it does provide information on buyers’ broad industries. The most popular industries are IT services (15.75%), web development or design (11.11%), software (10.67%), e-commerce (9.01%), consulting (5.60%), SaaS (5.32%), advertising (5.56%), and entertainment (3.75%). This field also allows respondents to enter free text, which highly increases the number of categories that can be used for matching; there are over 280 different industries in the data.

Among the 5,134 buyers for which we have all this data, 1,259 are treated and 3,875 are potential controls. Using the 5 criteria described above, we develop 3 different weighted matched subsamples. The first, which we call CEM1, uses the memory usage, the frequency infrastructure resizing, the employment, and the general use cases as matching criteria (i.e., all except industry). The process produces 294 strata with at least one treated and one control firm in them. We have an average of 1.1 treated and 8 control buyers per stratum. CEM1 sample has 2,685 buyers, of which 320 upgrade from basic to full support, and the rest exclusively use basic support. For our next matched sample, CEM2, we drop memory use as matching criteria and incorporate industry. We drop pre-upgrade memory use to mitigate any potential concerns on matching based on a parameter directly tied to our dependent variable—although this is what prior work does (Azoulay et al. 2011, Azoulay et al. 2010). We also integrate industry, which as mentioned before is highly granular and thus makes matching much more stringent. CEM2 has a total of 2,029 buyers, with an average of 1 treated buyer and 6.1 controls per stratum. Finally, we use all possible matching criteria in CEM3, which thus is our most stringent matching outcome. This subsample has only 687 buyers, and matches an average of 1 treated buyer to 3.4 controls per stratum. The full details of the subsamples construction and CEM procedures are offered in our online appendix.

2.6 Results

2.6.1 Effects of Technology Support on IT Use

We present the results for Model (1) using the baseline sample in Column (1) of Table 2.3. Consistent with Hypothesis 1, the results indicate that buyers who adopt and have access to full support use, on average, 187.7% (i.e., $e^{1.057} - 1$) more memory than buyers who have access to basic support. Also, the test that the sum of the coefficients for $FullSupport_{it}$ and $SwitchToBasic_{it}$ in Column (1) of Table 2.3 are different from zero

Table 2.3: Baseline Results for Tests of Hypotheses 1 and 2

Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Sample	Baseline	CEM1	CEM2	CEM3	CEM1		
Model	Basic Model				Falsification Tests		Additional Controls
$FullSupport_{it}$	1.057*** (0.029)	1.067*** (0.057)	1.075*** (0.058)	1.055*** (0.084)	1.090*** (0.061)	1.086*** (0.064)	0.981*** (0.056)
$SwitchToBasic_{it}$	-0.488*** (0.060)	-0.578*** (0.141)	-0.752*** (0.146)	-0.679*** (0.122)	-0.579*** (0.141)	-0.579*** (0.141)	-0.581*** (0.148)
$AdoptFullIn2_{it}$					0.081** (0.041)		
$AdoptFullIn4_{it}$						0.037 (0.044)	
$BusinessGrowth1_{it}$							0.372*** (0.049)
$BusinessGrowth2_{it}$							0.229*** (0.057)
Constant	0.230*** (0.024)	-0.402 (0.427)	-0.302 (0.473)	-0.688* (0.392)	-0.397 (0.425)	-0.398 (0.426)	-0.354 (0.380)
Observations	368,606	48,725	37,837	13,262	48,725	48,725	48,725
Buyers	22,179	2,685	2,029	687	2,685	2,685	2,685
R^2	0.251	0.321	0.336	0.397	0.321	0.321	0.337
Upgrade change $(e^{\hat{\beta}} - 1)$	187.7%	190.8%	192.9%	187.2%	197.6%	196.3%	166.8%
Downgrade change $(e^{\hat{\beta}+\hat{\gamma}} - 1)$	76.7%	63.1%	38.0%	45.6%	66.8%	66.1%	49.2%
$\hat{\beta} + \hat{\gamma} = 0$ test p-value	0.000	0.000	0.022	0.002	0.000	0.000	0.005

Dependent variable is $lnMemory_{it}$. All regressions include calendar (τ_t) and lifetime time dummies (l_{it}). Robust standard errors, clustered on customers, in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

is statistically significant at the 1% level, indicating that even after buyers have switched from full to basic support, they continue using, on average, 76.7% (i.e., $e^{1.057-0.488} - 1$) more memory than buyers who never accessed to full support. This result provides support for our second hypothesis that suggests that the positive effects of technology support on IT use are durable.

These findings are economically significant for the service provider, as can be seen by computing their implications for average (monthly) revenue per user (ARPU)[6] as follows. While a buyer consuming memory at the median of the distribution generates an ARPU of $64.60, buyers who opt for full support generate an ARPU of $185.85. Moreover, buyers who switch to basic support continue contributing an ARPU of $114.15. Considering the tens of thousands of firms using cloud infrastructure services, offering full support to buyers has significant revenue implications for the provider.The results with our various CEM-based subsamples, shown in columns (2) through (4) of Table 2.3, are consistent with those obtained with the baseline sample. The percentage changes in memory use associated with the upgrade from basic to full support range between 187.2% and 192.9%. Similarly, the results suggests that basic support users who had access to full support in the past continue using an average of 38.0% to 63.1% more memory than those who have exclusively accessed to basic support. In all what follows we continue basing our analysis on models ran using the CEM1 subsample. We chose this subsample over the baseline subsample because the matching procedure, along with its weights, allows us to better compare treated and control groups. Further, the CEM1 sample has more observations than CEM2 and CEM3, making it less prone to small

[6] During our sample period, Amazon Web Services' Elastic Compute Cloud (EC2), the public IaaS with the largest market share and thus with the dominant price-setting position, offered small 1.7 GB RAM servers at $0.08/hour and medium 3.75 GB RAM servers at $0.16/hour (source: aws.amazon.com). Based on these rates, we compute the mid-point price for 1 GB RAM server/hour, our measurement unit for $Memory_{it}$, and set the market price (p_b) of 1 GB RAM/hour at $0.045. The median $Memory_{it}$ in our data is 2 GB RAM/hour, and we multiply it by 720 to get a median monthly memory usage of 1,440 GB RAM hours. We do not use the mean $Memory_{it}$ because of the strong positive skew of its distribution. With this we estimate median ARPU at $64.60, and then multiply it by corresponding marginal effects.

sample issues.

2.6.2 Robustness Checks

Our use of the matching procedures increases confidence in our identifying assumption that there do not exist unobserved time-varying factors that differentially affect IT use of our treatment and control groups. However, in this section we further probe concerns of omitted variable bias and simultaneity through a series of robustness checks.

Falsification Tests

We first perform a falsification test to verify if there is any significant change in buyer behavior in the months immediately preceding the adoption of full support. We examine whether buyers' memory use before the adoption of full support is similar among buyers who will adopt full support and those that will continue using basic support. For this, we add 2 variables to Model (1). Parameters $AdoptFullIn2_{it}$ and $AdoptFullIn4_{it}$ are dummy variables equal to 1 in the 2 and 4 months (respectively) immediately before the adoption of full support. Thus, for example, if a customer adopts full support in $t = 10$, then $AdoptFullIn4_{it} = 1$ for $t = 6, \ldots, 9$, and is equal to 0 otherwise.

We present our results with these new parameters in columns (5) and (6) of Table 2.3. We find that customers tend to consume between 3.8% (i.e., $e^{0.037} - 1$) and 8.4% (i.e., $e^{0.0.81} - 1$) more memory in the months preceding the adoption of full support. These coefficients are positive and significant. However, their magnitude is much lower compared to the magnitude of the coefficient for $FullSupport_{it}$, which indicates the change in behavior once full support is adopted. Thus, it is unlikely that our results solely reflect changing unobservables that influence both IT use and full support, such as a previously planned increase in use.

Controlling for Business Growth

In addition to the cloud infrastructure usage and survey data described earlier, we have also gained access to the timing and content of the support interactions (i.e., online live chat sessions and support tickets) between the buyers and the provider. As an additional robustness test, we search buyers' support interactions for requests that are indicative of business growth, and use them as a control in our main model. While these controls may themselves be correlated with unobservables that influence cloud infrastructure use, they represent an additional observable proxy for factors influencing demand. We use them along the spirit of prior work such as Altonji et al. (2005): if adding these variables results in a significant decline in the measured effects of $FullSupport_{it}$, then that would provide evidence that time-varying unobservables significantly influence our results.

To operationalize these business growth-related controls, we search for support requests in which buyers ask for assistance in installing technical components of online web applications that are required when deploying a new system[7]. We also search for requests associated with increasing the provider-imposed limits on API calls to the infrastructure system, which are a clear signal of increasing activity in buyers' servers. Next, we create two dummy variables, $BusinesGrowth1_{it}$ and $BusinesGrowth2_{it}$, which are turned on whenever the buyer i has had at least 1 or 2 of these requests (respectively). We report the results of Model (1) with the inclusion of these parameters in column (7) of Table 2.3. We note that although the coefficient for $FullSupport_{it}$ is relatively smaller than that reported in prior specifications, the percentage change

[7] We search for the following requests: (i) to install Secure Sockets Layer (SSL) certificates, which are used to establish secure, encrypted connections between web servers and web browsers, and are essential elements of any web page that handles visitors' private information (e.g., credit card information); (ii) to add a new IP address to an existing server, which is needed to install the SSL certificate; (iii) to send a Certificate Signing Request (CSR), a core element of public key infrastructure (PKI); or (iv) registering a Sender Policy Framework (SPF) record, which is needed to send emails without being flagged as spammers. Section D.1 of Appendix D elaborates further on the coding process.

attributable to the upgrade from basic to full support is still high (i.e., 166.8%). This suggests that although unobserved growth in buyers' memory usage may be affecting our results, its role does not appear to be strong enough to overturn our findings. However, we are aware that more should be done in ruling out this alternative explanation, and thus use procedures with instrumental variables next.

Instrumental Variables Approach

We have also used the support interaction data to identify when buyers suffer from exogenous failures in using the cloud service. As expressed before, these exogenous shocks force the buyer to interact with the provider, which serves as a useful signal of the providers' service capabilities. In particular, buyers discover that by adopting full support and interacting more closely with the provider, they can reduce their total cost of solving their complications, resulting in a greater use of the cloud service.

We identify three different types of exogenous failures: (1) generalized outages across the cloud infrastructure service, such as those caused by a bug in the provider's cloud management platform, and that are generally reported on the providers "service status" webpage; (2) network-related failures such as when a specific node in the provider's infrastructure, generally belonging to some buyer, is suffering from a distributed denial of service attack (DDoS) or when a particular networking hardware device has failed; and (3) problems in which buyers suffer degraded performance due to a problem in the physical host in which their virtual machine runs. The last type of problems is generally associated with excessive read/write (or input/output) operations on the hard disks, either by the buyer (e.g., by some unexpected bug in their own applications such as a memory overflow that causes swapping) or by another buyer whose virtual machine lives in same physical server (e.g., a "noisy neighbor"). These problems could also be associated with a failure of the physical hardware (e.g., a hard disk failure).

Using a process identical to construct the business growth parameter above, we create 3 vectors of dummies indicating the number of failures of each type that a buyer i has experienced by t. Specifically, let $FailOutageN_{it}$, $FailNetworkN_{it}$, and $FailHostN_{it}$ be dummies that are turned on if buyers have experienced at least N failures of each corresponding type by time t. Given that these failures may have differential effects on the likelihood of upgrading for less experienced buyers, we also interact these dummies with an indicator of buyers still being in their first semester (i.e., first 6 months) since signup. Our indicator of this is $Semester1_{it}$, and it is equal to 1 if t corresponds to any of the first 6 months in buyer i's lifetime. In this section we comment on our results using 2 dummies of each type, yet our results are consistent using 1 or 3 of them (they are included in our online appendix along with all other descriptive statistics mentioned in this section).

Given the binary nature of our endogenous variable, we first follow the approach suggested by Wooldridge (2007) and Angrist and Pischke (2009), and as a first step in our estimation process, we use the vector of failure-related indicators and their interaction with $Semester1_{it}$ in a probit model using $FullSupport_{it}$ as dependent variable.

We use each failure type independently in columns (1) through (3) in Part C of Table 2.4, and all 3 types of failures in column (4).[8] The results suggest that, as proposed, all failure types are positively associated with buyers' likelihood of adopting full support. We use the probit model to generate the fitted values of $FullSupport_{it}$, which we denote as $FullSupport_{it}^{f}$.[9] Next, we employ $FullSupport_{it}^{f}$ as our instrument for $FullSupport_{it}$ in a 2SLS estimation procedure. The first stage results are reported in Part

[8] We ran the probit model with the excluded instruments and the monthly calendar and lifetime dummies. However, given the high singularity of the variance matrix caused by the strong presence of zeroes in the exogenous failure indicators, we use semester lifetime dummies rather than monthly ones.

[9] The descriptive statistics for $FullSupport_{it}^{f}$ are shown in Part B of Table 2.4. Its mean value is lower than the 0.160 reported in Table 2.2 for the baseline sample because of the exclusion from the sample of buyers who use full support immediately upon signup.

Table 2.4: Probit for $FullSupport_{it}$ and First Stage Results with fitted $FullSupport_{it}^{f}$

Column	(1)	(2)	(3)	(4)
Failure Types	Outage	Network	Host	All 3
Part A. First Stage Regression of Fitted $FullSupport_{it}^{f}$ on Real $FullSupport_{it}$				
$FullSupport_{it}^{f}$	0.644***	0.888***	0.621***	0.636***
	(0.073)	(0.221)	(0.100)	(0.068)
Observations	48,725	48,725	48,725	48,725
Buyers	2,685	2,685	2,685	2,685
R^2	0.140	0.114	0.125	0.143
First Stage F-Stat	77.310	16.107	38.931	86.746
Part B. Descriptive Statistics of $FullSupport_{it}^{f}$				
Mean	0.078	0.078	0.078	0.078
Std. Dev.	0.087	0.063	0.079	0.092
Min	0.000	0.000	0.000	0.000
Max	0.861	0.625	0.623	0.918
Part C. Coefficients of Probit with $FullSupport_{it}$ as dependent variable				
$FailOutage1_{it}$	0.990***			0.769***
	(0.042)			(0.045)
$FailOutage2_{it}$	0.679***			0.601***
	(0.079)			(0.083)
$FailOutage1_{it}$	-0.419**			-0.586***
$\times Semester1_{it}$	(0.196)			(0.216)
$FailOutage2_{it}$	-0.574			-0.766
$\times Semester1_{it}$	(0.581)			(0.591)
$FailNetwork1_{it}$		0.712***		0.264***
		(0.066)		(0.073)
$FailNetwork2_{it}$		0.035		-0.910***
		(0.234)		(0.266)
$FailNetwork1_{it}$		-0.213		0.070
$\times Semester1_{it}$		(0.357)		(0.341)
$FailHost1_{it}$			0.433***	0.352***
			(0.027)	(0.028)
$FailHost2_{it}$			0.535***	0.185***
			(0.042)	(0.046)
$FailHost1_{it}$			0.122	0.186*
$\times Semester1_{it}$			(0.096)	(0.098)
$FailHost2_{it}$			-0.102	0.209
$\times Semester1_{it}$			(0.218)	(0.226)
Constant	-0.724***	-0.528***	-0.787***	-0.883***
	(0.109)	(0.105)	(0.109)	(0.112)
Observations	48,425	48,425	48,425	48,425
Pseudo-R2	0.130	0.092	0.121	0.143

Linear regressions in Part A include monthly calendar and semester lifetime time dummies. Robust standard errors, clustered on customers, in parentheses. Descriptive statistics in Part B correspond to $FullSupport_{it}^{f}$ within CEM1 and after considering periods in which population was not yet at risk (see text for details). Part C shows coefficients of Probit regressions that include monthly calendar dummies and semester (6-month) lifetime dummies. Robust standard errors in parentheses. Coefficient for $FailNetwork1_{it} \times Semester1_{it}$ is dropped out of model since parameter is always equal zero.

$^{*} p < 0.10, ^{**} p < 0.05, ^{***} p < 0.01.$

A of Table 2.4. The values of the F-statistic for the excluded instruments ranges between 16.11 and 86.75, and in all cases are significant at the 1% level. It is evident that $FullSupport_{it}^{f}$ is positively associated with the real $FullSupport_{it}$. The second stage results using $lnMemory_{it}$ as our dependent variable are reported in columns (2) through (5) of Table 2.5. As baseline for comparison, this table shows in column (1) the result of Model (2), which excludes the $SwtichToBasic_{it}$ parameter, using standard fixed effects. The coefficients for $FullSupport_{it}$ are high relative to models without instruments, yet still consistent with our main hypothesis. Since our model is exactly identified, having $FullSupport_{it}^{f}$ as the only excluded instrument for $FullSupport_{it}$, we do not report the Hansen (1982) J statistic for these models.

While the failure events identified through the support interactions are completely unexpected to the buyer, their exogeneity can be questioned if one considers that buyers with a greater number of servers are more likely to suffer at least one failure in any of their servers. In other words, the past failures may be influenced by past usage. In the next section, we employ models that add lagged use as additional controls in our instrumental variable regressions.

Dynamic Panel Estimation and Endogenous Adoption Decisions

We continue exploring unobserved time-varying factors that may influence our findings. In particular we examine how persistence in our dependent variable affects our findings, given that memory usage is recent past periods may have a strong influence on memory usage during the focal period. As suggested in our presentation of our econometrics approach, we first ran a fixed effects model using varying number of lags for the dependent variable, and attained qualitatively similar results to our baseline models. For reasons that will be explained below, we only present the results using 4 lags of $lnMemory_{it}$ in column (6) of Table 2.5. We confirm our suspicion that the current value of the variable is strongly influenced by its past values; this is reflected in the large

Table 2.5: Results with Instrumental Variables and Dynamic Panels for $lnMemory_{it}$

Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Model	Basic Model					Dynamic Panel Model				
Estimation Procedure	FE	2SLS				FE	System GMM			
$FullSupport_{it}$	1.019***	2.482***	2.948***	3.793***	2.881***	0.310***	0.087***	0.052***	0.089***	0.047**
	(0.055)	(0.294)	(1.027)	(0.592)	(0.327)	(0.022)	(0.029)	(0.019)	(0.031)	(0.020)
$lnMemory_{it-1}$						0.964***	0.976***	1.046***	0.970***	1.047***
						(0.024)	(0.069)	(0.065)	(0.069)	(0.064)
$lnMemory_{it-2}$						-0.192***	-0.029	-0.085	-0.025	-0.084
						(0.038)	(0.084)	(0.082)	(0.083)	(0.081)
$lnMemory_{it-3}$						-0.004	-0.037	-0.019	-0.038	-0.008
						(0.034)	(0.035)	(0.109)	(0.036)	(0.108)
$lnMemory_{it-4}$						0.004	0.039**	0.041	0.043**	0.033
						(0.018)	(0.017)	(0.082)	(0.018)	(0.081)
Constant	-0.359					0.253***	-0.038	0.096	-0.053	0.085
	(0.440)					(0.056)	(0.073)	(0.075)	(0.109)	(0.075)
Observations	48,725	48,725	48,725	48,725	48,725	37,991	37,991	37,991	37,991	37,991
Buyers	2,685	2,685	2,685	2,685	2,685	2,657	2,657	2,657	2,657	2,657
Failure-based IVs	-	Outage	Network	Host	All 3	-	-	-	All 3	All 3
Lags of first differences used as IVs							All avail.	Least Possible	All avail.	Least Possible
Total Number of IVs							859	168	870	179
Hansen J Statistic p-value							0.995	0.623	0.984	0.357
Upgrade change ($e^{\widehat{\beta}} - 1$)	177%	1097%	1807%	4339%	1683%	36.3%	9.0%	5.3%	9.3%	4.9%

Dependent variable is $lnMemory_{it}$. All regressions include calendar (τ_t) and lifetime time dummies (l_{it}). Columns (1) through (6) show robust standard errors, clustered on customers, in parentheses. System GMM models in columns (7) through (10) have robust standard errors that use Windmeijer's (2005) finite sample correction.

$^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01.$

Hansen J statistic not reported for 2SLS estimations in columns (2) through (5) as model is exactly identified.

System GMM estimations in columns (7) through (10) consider $FullSupport_{it}$ as endogenous. Given AR(2) in the errors, they all use the 2nd lag of the first difference of $lnMemory_{it}$ and $FullSupport_{it}$ as their instruments for the levels equation. Columns (7) and (9) use all available lags of $lnMemory_{it}$ and $FullSupport_{it}$ as instruments for the first differences equation, from the 3rd lag until the end of the panel. Columns (8) and (10) only use the 3rd lag of $lnMemory_{it}$ and $FullSupport_{it}$ as instruments for the differences equation. Additionally, columns (9) and (10) augment the instruments matrix by considering the same vector of exogenous failure-related instruments shown in column (4) of Table 2.4 and described in section 2.6.2.

size and statistical significance for the lagged dependent variables. Nonetheless, even after controlling for this, we find that the memory usage still increases 36.3% (i.e., $e^{0.310} - 1$) with the adoption of and continued access to full support ($FullSupport_{it}$). While the magnitude of the coefficient is much lower than that in prior specifications, the signs and statistical significance of the parameter continues to hold and support Hypotheses 1. The drop in the coefficient's magnitude was expected since the inclusion of lagged dependent variables is known to suppress the explanatory power of other covariates, especially if they are trending as our support choice indicators are (Achen 2000).

In implementing our System GMM estimation procedures, we first select the appropriate number of lags of the dependent variable to be included as regressors. Following a process similar to that executed by Chen et al. (2013), we selected the number of lags by first choosing a number of lags that is consistent with our phenomena of interest and then test for serial correlation in the errors and the validity of the overidentifying restrictions. We chose to use 4 lags of $lnMemory_{it}$ based on the provider's belief that it takes customers about 4 months to stabilize their behavior. In our first run, using all available instruments (from the 1st lag of the first differences and from the 2nd lag of the values up to the end of the panel), the Arellano and Bond (1991) serial correlation test indicated that we do not only have the expected 1st order serial correlation but also have 2nd order serial correlation. As a result, we cannot use the 1st lag of the variables first differences nor the 2nd lag of the variables values as instruments. However, we can still rely on the variables' 2nd lag of their first difference as instruments for the levels equation and their 3rd and posterior lags as instruments for the levels equation (Cameron and Trivedi 2010).

We show the model with all available instruments in column (7) of Table 2.5, which passes the Hansen (1982) J test for the validity of our overidentifying restrictions with $\chi^2(793) = 694.03,\ p = 0.995$. We also verified we did not suffer from 3rd or higher orders or serial correlation in our errors. The coefficient for $FullSupport_{it}$ suggests an increase in memory usage of 9.0% (i.e., $e^{0.087} - 1$). Then to avoid the problem of over fitting the model with too many instruments (Roodman 2009b), we gradually reduced the number of lags used as instruments until we found the least number of instruments under which we still passed the instrument validity Hansen J test. We found that we can limit our model to the use of the 3rd lag of $lnMemory_{it}$ and $FullSupport_{it}$. Such model is reported in column (8) of Table 2.5. We once again pass all specification tests, and we continue finding a positive and significant effect for full support, this time representing an increase in memory usage of 5.3% (i.e., $e^{0.052} - 1$). Moreover we also reduced the total number of instruments from 859 to just 168. Next, we augment our instrument matrix for these same model specifications with the exogenous failure-based instruments used in column (4) of Part C of Table 2.4. The results with this augmented instrument matrix are shown in columns (9) and (10) of Table 2.5, with all and the least number of instruments respectively. Results do not vary much relative to those already discussed in columns (7) and (8).

2.6.3 How do Firm Size and Technology Support Interact to Shape IT Use?

Our results that test Hypotheses 3 and 4 are presented in Table 2.6. Column (1) of Table 2.6 presents the same result already shown in column (2) of Table 2.3, and serve as reference. The interpretation of the coefficients of the interactions with the dummy variables $EmploymentTop50_i$ and $EmploymentTop25_i$ in columns (2) and (3) of Table 2.6 is straightforward, but in order to better understand the effects of the interactions with $lnEmployment_i$, we evaluate the percentage changes in memory ($lnMemory_{it}$) due to turning the support choice dummies on at one standard deviation below or above mean $lnEmployment_i$ (see footer of Table 2.6 for details).

The results provide strong support for both Hypotheses 3 and 4. The tests that the sums of all corresponding coefficients in the analysis here are different from zero are

Table 2.6: Are the Results of $FullSupport_{it}$ Stronger for Large Firms?

Column	(1)	(2)	(3)	(4)
Dependent Variable		$lnMemory_{it}$		
$FullSupport_{it}$	1.067***	0.835***	1.011***	0.848***
	(0.057)	(0.082)	(0.063)	(0.093)
$SwitchToBasic_{it}$	-0.578***	-0.473***	-0.728***	-0.822***
	(0.141)	(0.164)	(0.150)	(0.182)
$FullSupport_{it}$ $\times EmploymentTop50_i$		0.382***		
		(0.111)		
$SwitchToBasic_{it}$ $\times EmploymentTop50_i$		-0.167		
		(0.265)		
$FullSupport_{it}$ $\times EmploymentTop25_i$			0.257*	
			(0.145)	
$SwitchToBasic_{it}$ $\times EmploymentTop25_i$			0.774***	
			(0.252)	
$FullSupport_{it}$ $\times lnEmployment_i$				0.087***
				(0.032)
$SwitchToBasic_{it}$ $\times lnEmployment_i$				0.108*
				(0.060)
Constant	-0.402	-0.395	-0.393	-0.414
	(0.427)	(0.411)	(0.405)	(0.416)
R^2	0.321	0.324	0.324	0.324
Upgrade change, small buyers $\left(e^{\hat{\beta}_1}-1\right)$	190.8%	130.5%	174.9%	145.7% [a†]
Downgrade change, small buyers $\left(e^{\hat{\beta}_1+\hat{\gamma}_1}-1\right)$	63.1%	43.7%	32.7%	15.1% [a†]
$\hat{\beta}_1+\hat{\gamma}_1=0$ test p-value	0.000	0.027	0.049	-
Upgrade change, large buyers $\left(e^{\hat{\beta}_1+\hat{\beta}_2}-1\right)$		237.6%	255.5%	214.1% [a‡]
Downgrade change, large buyers $\left(e^{\hat{\beta}_1+\hat{\beta}_2+\hat{\gamma}_1+\hat{\gamma}_2}-1\right)$		78.0%	272.1%	99.6% [a‡]
$\hat{\beta}_1+\hat{\beta}_2+\hat{\gamma}_1+\hat{\gamma}_2=0$ test p-value		0.004	0.000	-

Dependent variable is $lnMemory_{it}$. All regressions include calendar (τ_t) and lifetime time dummies (l_{it}). Robust standard errors, clustered on customers, in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$. Coefficients $\hat{\beta}_1$, $\hat{\beta}_2$, $\hat{\gamma}_1$ and $\hat{\gamma}_2$ correspond to the estimated parameters of Model (3). All regressions use CEM1 subsample, which has 48,725 buyer-month observations from 2,685 buyers.

[a] Since $Memory_{it}=e^{lnMemory_{it}}-1$, per Model (3), the percentage change in $Memory_{it}$ is given by $e^{\hat{\beta}_1 FullSupport_{it}+\hat{\gamma}_1 SwitchToBasic_{it}+lnEmployees_i(\hat{\beta}_2 FullSupport_{it}+\hat{\gamma}_2 SwitchToBasic_{it})}-1$.

Mean $lnEmployment_i$ is $\bar{x}=2.002$, and its standard deviation is $\sigma^2=1.416$.

† These are marginal effects computed for small buyers, using $lnEmployment_i=\bar{\mathrm{x}}-\sigma^2$.

‡ These are marginal effects computed for large buyers, using $lnEmployment_i=\bar{\mathrm{x}}+\sigma^2$.

statistically significant at the 1% level. We base our first analysis on columns (2) and (3) of Table 2.6. We find that while large firms increase their memory use between 237.6% (i.e., $e^{0.835+0.382} - 1$) and 255.5% (i.e., $e^{1.011+0.257} - 1$) upon adoption of full support, smaller ones only increase their use between 130.5% (i.e., $e^{0.835} - 1$) and 174.9% (i.e., $e^{0.1.011} - 1$). Using our continuous measure of employment, column (4) suggests small buyers ($lnEmployment_i = \bar{x} - \sigma^2$) grow their usage by 145.7%, large ones $lnEmployment_i = \bar{x} + \sigma^2$) do so by 214.1%. Therefore, the coefficient estimates in Table 2.6 show strong support for Hypothesis 3. Similarly, while large former full support buyers continue using between 78.0% and 272.1% after they have switched to basic support relative to pure basic support buyers, small firms only use between 15.1% and 43.7%. These results provide strong support for Hypothesis 4.

Using the results in Column (2) of Table 2.6, we estimate how the increase in contributed ARPU due to the adoption and usage of full support varies with firm size, relative to a buyer at median memory consumption level (see footnote 6). We estimate that while large firms increase their ARPU from $64.60 to $218.09, small firms only increase their ARPU to $148.90. Similarly, regarding their ARPU after they switch to basic support, we estimate that large firms have an ARPU of $114.99, which is still higher than that of small firms who switch, $92.83, and much higher than that of the median buyer, $64.60. In sum, we find strong evidence of complementarity effects between firm size and the technology support.

2.6.4 Effects of Technology Support on Efficiency of IT Use

As mentioned in the presentation of our motivating analytical model, our fundamental assumption is that buyers who opt for full support enjoy a higher co-production output elasticity than those who opt for basic support, $z_f > z_b$. Moreover, adopters of full support continue enjoying benefits after switching to basic support because of what they have learned through their interactions with the provider. In this section we test if, in accordance with this assumption, buyers make better and more efficient use of the advanced cloud specific infrastructure as a result of having access to full support. An advantage of cloud infrastructure services is that we can partially observe z_s via certain attributes of buyers' deployments. One such proxy is the complexity of a buyer's deployments that serves to assess how proficient a buyer is in making use of the infrastructure. In general, if full support helps buyers co-produce more efficiently, then we should expect that full support buyers employ architectures with greater levels of complexity. We explain this assertion and offer a test of it in the discussion below.

Although the on-demand nature of the service along with its rapid elasticity provides firms the opportunity to reduce idle computing capacity waste and eliminate the necessity of an up-front capital commitment in overprovisioning resources (Armbrust et al. 2010, Harms and Yamartino 2010), doing so requires firms to scale their infrastructure in a cost-efficient manner. There are essentially two ways of growing an IT infrastructure: vertically, or up, and horizontally, or out (Garcia et al. 2008, Michael et al. 2007, Reese 2009, p. 176). Scaling vertically implies increasing the capacity of a server or spreading out the IT stack across several servers, in either case having at most one server per function. While this approach is easy to implement, growth in vertical scaling is capped by the maximum server capacity available. In contrast, under horizontal scaling several servers perform functions in parallel and this scaling method offers virtually unlimited growth potential. Buyers may prefer to scale horizontally for other reasons. Given the relatively high likelihood of a commodity cloud server failing, an IT infrastructure architecture designed for cloud environments will optimally have its workloads distributed across several nodes, rather than all concentrated in a single node (Reese 2009). However, despite its advantages, horizontal scaling also presents challenges associated with load balancing and session management across servers, among others (Casalicchio and Colajanni 2000, Cherkasova 2000, and interviews with cloud experts at IBM Thomas J. Watson Research Center, Yorktown Heights, New York, and a major technological research univeristy). Therefore, having servers that work in parallel increases the complexity of the architecture and, at the same time this signals a better and more efficient use of advanced cloud features. As a result of these increased efficiencies and complexity, we use the fraction of servers running in parallel as a measure that proxies for a buyer's skill at using cloud computing. Note that this measure is independent of memory use, our first dependent variable: a buyer can consume a large volume with none of its servers functioning in parallel, case in which the fraction is zero, or a small volume with all of its servers functioning in parallel, which makes the faction equal to 1. In the context of our motivating model, evidence of an increased fraction of servers running in parallel under full support provides additional evidence of our assumption that $z_f > z_b$.

Our analysis of the complexity of buyer infrastructure deployments is based on an automated analysis of the names given by buyers to their servers. We develop an algorithm that compares the names of the servers being run by each buyer at the end of every day during our sample and check if we find servers with names very similar to each other.[10] Our assumption is that if we find two or more servers with very similar names, they will very likely be performing the same function in parallel (e.g., `web1.domain.com` and `web2.domain.com`). If we find different sets of servers with similar names, we count them all together as functioning in parallel (e.g., `web1`, `web2`, and `web3`, and `database1` and `database2`, are 5 servers working in parallel). At the end of each day in buyer i's lifetime, we count the number of servers with similar names and divide the count by the total number of servers being run, and then average the metric over month t. The resulting average fraction of servers running in parallel is captured in our new dependent variable, $FractionParallel_{it}$ (see Table 2.1 for descriptive statistics).

[10] Specifically, we consider two server names to be similar to each other if they have a Levenshtein (1966) Distance that is less or equal to two, meaning that one server's name can be made equal to the other by editing (inserting, deleting or substituting) 2 characters (letter or numbers) or less.

We estimate the exact same models described in our empirical approach (section 2.4) but substitute $FractionParallel_{it}$ for $lnMemory_{it}$ as the dependent variable. Overall, our results are consistent with respect to what we found before when using the IT service use dependent variable, providing additional evidence that full support enables customers to use the cloud more effectively.

Results shown in Table 2.7 show that customers who have adopted and continue having access to full support have a fraction of servers working in parallel that is between 9.6 and 10.8 percentage points higher than that of basic support users. Further, the results also indicate that former full support customers have a fraction of servers working in parallel that is between 6.5 and 9.0 percentage points higher after they switch to basic support. Also, the test that the sum of the coefficients for $FullSupport_{it}$ and $SwitchToBasic_{it}$ is different from zero is also statistically significant at the 1% level for all columns except column (7), where the test of the sum being different from zero is statistically significant but at the 5% level. As described in section 2.3.2, a very important nuance of how the service is offered makes this result very meaningful: if full support buyers who downgrade desire to continue running the same set of applications under the new basic support regime, they must redeploy their entire infrastructure on their own. Therefore, if they continue using a high number of servers in parallel, in turn suggesting usage of a horizontally scalable deployment, it must be the case that they set it up entirely on their own. Together, these results are consistent with our model assumptions that consumers learn from the provider through full support.

As before, we implemented a 2SLS model with exogenous failures as instruments for $FullSupport_{it}$. Column (6) of Table 2.8 shows the result of Model (2) using standard fixed effects, and columns (7) through (10) of the same table show the second stage results of 2SLS using the first stage results reported in Table 2.4; this is the same first stage used when we had $lnCapacity_{it}$ as dependent variable. The fixed effect estimate in in column (6) shows that the exclusion of the $SwitchToBasic_{it}$ parameter does not alter

Table 2.7: Results for Tests of Effects of Full Support on Efficiency of IT Use

Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Sample	Baseline	CEM1	CEM2	CEM3	CEM1		
Model	Basic Model				Falsification Tests		Additional Controls
$FullSupport_{it}$	0.096***	0.107***	0.106***	0.108***	0.108***	0.106***	0.098***
	(0.006)	(0.014)	(0.014)	(0.019)	(0.014)	(0.015)	(0.014)
$SwitchToBasic_{it}$	-0.032***	-0.030	-0.025	-0.018	-0.030	-0.030	-0.030
	(0.012)	(0.028)	(0.031)	(0.028)	(0.028)	(0.028)	(0.028)
$AdoptFullIn2_{it}$					0.002		
					(0.010)		
$AdoptFullIn4_{it}$						-0.003	
						(0.010)	
$BusinessGrowth1_{it}$							0.040***
							(0.013)
$BusinessGrowth2_{it}$							0.021
							(0.017)
Constant	0.023***	-0.210	-0.114	-0.422***	-0.209	-0.210	-0.204
	(0.006)	(0.131)	(0.120)	(0.120)	(0.131)	(0.130)	(0.129)
Observations	368,606	48,725	37,837	13,262	48,725	48,725	48,725
Buyers	22,179	2,685	2,029	687	2,685	2,685	2,685
R^2	0.026	0.037	0.042	0.060	0.037	0.037	0.040
Downward change $(\hat{\beta} + \hat{\gamma})$	0.065	0.077	0.081	0.090	0.078	0.076	0.068
$\hat{\beta} + \hat{\gamma} = 0$ test p-value	0.000	0.006	0.010	0.004	0.006	0.009	0.014

Dependent variable is $FractionParallel_{it}$. All regressions include calendar (τ_t) and lifetime time dummies (l_{it}). Robust standard errors, clustered on customers, in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.8: Results with Instrumental Variables and Dynamic Panels for $FractionParallel_{it}$

Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Model	Basic Model					Dynamic Panel Model				
Estimation Procedure	FE	2SLS				FE	System GMM			
$FullSupport_{it}$	0.105***	0.388***	0.242*	0.423***	0.415***	0.039***	0.013**	0.006	0.014**	0.007
	(0.013)	(0.081)	(0.132)	(0.108)	(0.078)	(0.005)	(0.006)	(0.005)	(0.007)	(0.005)
$FractionParallel_{it-1}$						0.888***	0.807***	0.792***	0.806***	0.794***
						(0.014)	(0.035)	(0.041)	(0.034)	(0.040)
$FractionParallel_{it-2}$						-0.181***	0.025	0.050	0.024	0.047
						(0.011)	(0.034)	(0.041)	(0.033)	(0.040)
Constant	-0.207					0.023	-0.008	-0.101	-0.007	-0.112
	(0.130)					(0.031)	(0.025)	(0.109)	(0.026)	(0.108)
Observations	48,725	48,725	48,725	48,725	48,725	43,355	43,355	43,355	43,355	43,355
Buyers	2,685	2,685	2,685	2,685	2,685	2,684	2,684	2,684	2,684	2,684
Failure-based IVs	-	Outage	Network	Host	All 3	-	-	-	All 3	All 3
Lags of first differences used as IVs							All avail.	Least Possible	All avail.	Least Possible
Total Number of IVs							798	542	809	553
Hansen J Statistic p-value							0.980	0.290	0.969	0.271

Dependent variable is $FractionParallel_{it}$. All regressions include calendar (τ_t) and lifetime time dummies (l_{it}). Columns (1) through (6) show robust standard errors, clustered on customers, in parentheses. System GMM models in columns (7) through (10) have robust standard errors that use Windmeijer's (2005) finite sample correction.
$^{*} p < 0.10$, $^{**} p < 0.05$, $^{***} p < 0.01$.

Hansen J statistic not reported for 2SLS estimations in columns (2) through (5) as model is exactly identified.

System GMM estimations in columns (7) through (10) consider $FullSupport_{it}$ as endogenous. Given AR(2) in the errors, they all use the 2nd lag of the first difference of $FractionParallel_{it}$ and $FullSupport_{it}$ as their instruments for the levels equation. Columns (7) and (9) use all available lags of $FractionParallel_{it}$ and $FullSupport_{it}$ as instruments for the first differences equation, from the 3rd lag until the end of the panel. Columns (8) and (10) only use the 3rd to 12th lags of $FractionParallel_{it}$ and the 3rd to 13th lags of $FullSupport_{it}$ as instruments for the differences equation.. Additionally, columns (9) and (10) augment the instruments matrix by considering the same vector of exogenous failure-related instruments shown in column (4) of Table 2.4 and described in section 2.6.2.

our findings concerning the effect of $FullSupport_{it}$. When using the instrumented $FullSupport_{it}$, we find that the fraction of servers running parallel grows between 24.2 and 41.5 percentage points after buyers upgrade from basic to full support.

Continuing with the same models used for our firs dependent variable, we also implement a dynamic panel data model using GMM estimation. Using 4 lags of the dependent variable and all available instruments, the Arellano and Bond (1991) test again found 2nd order serial correlation, so we must again rely on the 2nd lag of the variables' first difference 3rd or later lags of the their values as instruments. After adjusting for this, this time we found that we could reduce the number of lags of the dependent variable included as regressors from 4 to just 2. The results of the fixed effects dynamic panel model with 2 lags of $FractionParallel_{it}$ included as regressors is shown in column (6) of Table 2.8, and suggest that the fraction of servers running in parallel grows by 3.9 percentage points once buyers upgrade from basic to full support. We then report the System GMM estimation with all available instruments in column (7) of Table 2.8, which suggests the fraction increases by 1.3 percentage points. Then, we reduce the number of lags of the variables used as instruments until we find the valid model that uses the smallest number of lags of $FractionParallel_{it}$ and $FullSupport_{it}$. The specification used in column (8) of Table 2.8 uses the 2nd lag of the parameters' first difference as instruments for the levels equation, and also uses the 3rd through 12th lags of $FractionParallel_{it}$ and the 3rd through 13th lags of $FullSupport_{it}$ as instruments for the first differences equation. Although the estimation with all available instruments in column (7) yields statistically significant coefficients for $FullSupport_{it}$, column (8) with the reduced number of instruments does not. For completeness, estimation results in columns (9) and (10) of Table 2.8 employ an augmented instruments matrix that incorporates the exogenous service failure events as additional instruments.

We now turn to the role of firm size in moderating how full support influences customers' architecture complexity. We report the regression results in columns (1) through (4) of Table 2.9. While larger firms that adopt full support increase their fraction of servers running in parallel by 11.2 to 14.2 percentage points, smaller firms only do so by 5.4 to 10.6 percentage points. That is, the increase due to the adoption of full support is about twice as much for larger firms than for smaller ones. Finally, and very interestingly, we find that the effects of full support on buyers architecture complexity after they downgrade from full to basic support is positive for all firms, but especially for the very large ones (i.e., top 25th percentile). Using the $EmploymentTop50_i$ indicator in column (6), which divides the sample at the median employment, there is not a clear differential effect post-downgrade between firms below and firms above the median. The t-test for $\hat{\beta}_1 + \hat{\beta}_2 + \hat{\gamma}_1 + \hat{\gamma}_2 = 0$ is only significant at the 10% level (p-value = 0.057). Similar outcome is found when comparing linear combinations of the coefficients in column (8) using values of $lnEmployment_i$ one standard deviation above and below the mean $lnEmployment_i$; the test does not find evidence of any difference in the fraction of servers running in parallel post-down grade between small and large firms. However, if we use the $EmploymentTop25_i$ indicator in column (6), which cutoffs the sample at the top 25th percentile of employment (and that is more than 1 standard deviation above the mean), there is stronger effect for these very large firms than for the rest. Very large former full support buyers have a proportion of servers working in parallel 15.1 percentage points above the baseline of those who never used full support, while smaller former full support buyers are only 5.9 percentage points above this baseline. This suggests the effects of full support on buyers architecture complexity are durable for all firms, but especially so to the largest ones.

Table 2.9: Are the Results for $FractionParallel_{it}$ Stronger for Large Firms?

Column	(1)	(2)	(3)	(4)
Dependent Variable		$FractionParallel_{it}$		
$FullSupport_{it}$	0.107***	0.054***	0.106***	0.066***
	(0.014)	(0.017)	(0.016)	(0.020)
$SwitchToBasic_{it}$	-0.030	0.030	-0.046	-0.016
	(0.028)	(0.036)	(0.031)	(0.049)
$FullSupport_{it}$		0.088***		
$\times EmploymentTop50_i$		(0.025)		
$SwitchToBasic_{it}$		-0.100*		
$\times EmploymentTop50_i$		(0.053)		
$FullSupport_{it}$			0.007	
$\times EmploymentTop25_i$			(0.030)	
$SwitchToBasic_{it}$			0.085	
$\times EmploymentTop25_i$			(0.068)	
$FullSupport_{it}$				0.016**
$\times lnEmployment_i$				(0.006)
$SwitchToBasic_{it}$				-0.005
$\times lnEmployment_i$				(0.020)
Constant	-0.210	-0.204	-0.211	-0.209
	(0.131)	(0.133)	(0.131)	(0.132)
R^2	0.037	0.040	0.038	0.038
Upgrade change, small buyers $(\hat{\beta}_1)$	0.107	0.054	0.106	0.076 †
Downgrade change, small buyers $(\hat{\beta}_1 + \hat{\gamma}_1)$	0.077	0.084	0.059	0.056 †
$\hat{\beta}_1 + \hat{\gamma}_1 = 0$ test p-value	0.006	0.035	0.058	-
Upgrade change, large buyers $(\hat{\beta}_1 + \hat{\beta}_2)$		0.142	0.112	0.122 ‡
Downgrade change, large buyers $(\hat{\beta}_1 + \hat{\beta}_2 + \hat{\gamma}_1 + \hat{\gamma}_2)$		0.071	0.151	0.088 ‡
$\hat{\beta}_1 + \hat{\beta}_2 + \hat{\gamma}_1 + \hat{\gamma}_2 = 0$ test p-value		0.057	0.012	-

All regressions include calendar (τ_t) and lifetime time dummies (l_{it}). Coefficients $\hat{\beta}_1$, $\hat{\beta}_2$, $\hat{\gamma}_1$ and $\hat{\gamma}_2$ correspond to the estimated parameters of Model (3). All regressions use CEM1 subsample, which has 48,725 buyer-month observations from 2,685 buyers.

Robust standard errors, clustered on customers, in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Mean $lnEmployment_i$ is $\bar{x} = 2.002$, and its standard deviation is $\sigma^2 = 1.416$.

† These are marginal effects computed for small buyers, using $lnEmployment_i = \bar{x} - \sigma^2$.

‡ These are marginal effects computed for large buyers, using $lnEmployment_i = \bar{x} + \sigma^2$.

2.7 Conclusion

Using a unique and rich data set on public cloud infrastructure services consumption by 20,298 firms over the period from March 2009 to April 2012, our study is the first to examine how a provider's technology support influences buyer *post-adoption* IT use. We examine how a buyer's access to support influences its usage and use effectiveness of on-demand IT infrastructure. Our estimates of the positive impact of offering technology support are economically significant. Buyers who adopt and access full support use, on average, 188% more of the service than those who only access basic support. We also find evidence that customers who switch from full to basic support continue using an average of 77% more of the service than buyers who only had access to basic support throughout our sample period. These findings directly impact the provider's topline performance. While a median buyer generates an ARPU of $64.60, buyers who opt for full support generate an ARPU of $185.85 and former full support buyers who switch to basic support continue contributing an ARPU of $114.15.

Our research has broader implications for business analytics and operations for service-oriented technology industries such as cloud computing. Cloud computing democratizes IT infrastructures and allows small firms to have access to computing infrastructures previously available only to larger ones (Varian 2011). A byproduct of this is that the cloud context offers a unique opportunity to examine the IT investment of small startups who, given their low levels of IT spending, have not adequately been captured by traditional data sources and so have not been frequently studied. In our research, we observe data on the actual usage of an IT service by tens of thousands of very small firms (i.e., less than $1M in revenue and less than 100 employees).

On the other hand, the democratization of IT also implies that cloud services providers face the complex challenge of offering a service amenable to a wide variety of use cases and business needs for a very heterogeneous customer base (Venters and Whitley 2012). Moreover, the self-service nature of the service induces uncertainty in service outcomes, which strongly depend on customers' traditionally unknown capabilities in co-producing the service (Chase 1978). Our work with a provider's business analytics team has aimed to help them deal with these challenges by showing them how they can exploit their detailed records of customers' behavior to rigorously and cost effectively examine their managerial decisions' impact on customers' behavior and, in turn, on their own internal operations. We believe our approach can be applied by providers with similar data sets in a wide array of B2B self-service technologies that pose knowledge barriers to adopters. In particular, we have shown that technology support has the potential to increase usage for IaaS and to help buyers make use of it more effectively, enabling them to overcome the knowledge barriers that have engendered the slow rates of cloud adoption we see today.

Although, to our knowledge, this is the first study that empirically examines drivers of usage for cloud infrastructure services, doing so does not come without some inherent limitations. These limitations, nonetheless, may be overcome by future research through additional data collection. One limitation of this study is our inability to directly observe the value that buyers derive from their utilization of cloud infrastructure services. For example, since we do not observe buyers' financial and operational performance, we cannot follow prior literature on IT value (e.g., Aral et al. 2006, Brynjolfsson and Hitt 1995, Brynjolfsson and Hitt 1996, Hitt et al. 2002) in capturing the impact of the adoption and usage of cloud services on firm performance. Similarly, while one of the most commonly mentioned benefits of cloud computing is its ability to reduce idle IT capacity waste (Armbrust et al. 2010, Harms and Yamartino 2010), we cannot capture these IT capacity savings since we do not observe buyers' IT investments outside of the provider's cloud. In our research, we observe buyers' revealed preferences to employ cloud infrastructure services over some other IT infrastructure alternative, such as a self-run data center. Under the assumption that buyers are economically rational entities, we correspondingly also assume that buyers' demand for IT use in the cloud is a proxy for the value they derive from it. This presents an exciting future research opportunity that combines cloud usage data with metrics on firm performance and in-house IT infrastructure. Finally, our work has used a variety of approaches to identify the effects of technology support on service use, including matching methods, falsification exercises, instrumental variables, and dynamic panel models. Future work could further probe our identification assumptions by conducting field experiments with buyers in which the provider varies specific elements of the technology support it offers.

Regarding future research avenues, it is particularly interesting to note that recent work in a B2C context found that SSTs are not producing significantly greater customer satisfaction levels compared to physical channels, which suggests that if individuals continue using the SST it is because of high switching costs (Buell et al. 2010). In our B2B setting, and in particular given the standardized and commoditized nature of our studied service, technology support has the effect of greatly reducing customers' switching costs. Therefore, it is likely that technology support does not only increase the IT use as we have shown, but is also affecting customers' satisfaction, as that is the only reason for them staying. Future work may survey customers or conduct field experiments to assess the extent to which technology support is tied to customers' satisfaction. Another fruitful research area will be to explore the role of other parties (e.g., third-party service providers) in the cloud computing ecosystem. Finally, from the provider's perspective, there are abundant opportunities to measure how changes in service contract terms impact buyer behavior and provider revenue.

CHAPTER 3

EARLY PROACTIVE EDUCATION, CUSTOMER RETENTION, AND DEMAND FOR TECHNOLOGY SUPPORT

3.1 Introduction

Academics and practitioners have long recognized service customers' role as both recipients and producers, or co-producers, of the service delivered, particularly in the context of high contact services where customers are deeply involved in the creation of the service (Chase 1978). Examples of such services are online self-service technologies (SSTs), for which research has consistently shown that customers' knowledge, skills and abilities in co-producing the service are a key determinant of their adoption and continued usage (e.g., Xue and Harker 2002, Xue et al. 2011, Xue et al. 2007). In other contexts, such as financial services that require high involvement from customers, surveys have also shown that customers' expertise is positively associated with their loyalty (Bell and Eisingerich 2007).

Given the positive relationship between customer's capabilities and their adoption and use of SSTs, it is not surprising that providers make themselves available to answer questions from their customers and assist them in their service co-production efforts. A common channel used for this is *reactive* technical support—customer-initiated interactions with the provider in which the latter assists customers in deriving a greater utility from the service. Prior work has suggested that the accessibility to external knowledge sources and support are important in determining users' decisions to adopt a new IT product (Li et al. 2005, Morgan and Finnegan 2007). Ongoing research in the context of cloud infrastructure services, an SST with a relatively high level of technical sophistication, has also shown the positive link between a provider's (reactive) assistance to customers in their co-production efforts and their consumption of the service (Retana et al. 2013). However, much less is known about the potential benefits for providers of offering *proactive* (i.e., provider-initiated) assistance.

In this research, we take a first step in exploring one form of such proactive engagements, namely *early proactive education* (EPE), and how it influences customer behavior in the context of public cloud infrastructure services (a very high-contact SST).[11] We define EPE as any provider-initiated effort to increase customers' service co-production-related knowledge and skills immediately after service adoption. This may enable customers to derive a greater utility from the service. We distinguish it from reactive education, which could be offered through standard reactive technical support channels (e.g., a contact center), and which has been the focus of prior work (e.g., Field et al. 2012, Retana et al. 2013). We also distinguish it from proactive sales or cross-selling engagements (e.g., Aksin and Harker 1999, Gurvich et al. 2009), because in our context the education is offered by technical staff and not by sales representatives. Moreover, we specifically examine proactive education (Challagalla et al. 2009) that is offered as soon as customers adopt the service and are taking their initial steps in co-producing the service (i.e., start adapting the service to their idiosyncratic needs). In our empirical study, we attempt to offer insights into the following research question: *What are the effects of EPE on customers' retention and demand for technology support during the early stage of their co-production processes?*

Our focus on the early stages of customers' co-production processes is motivated by both prior literature and practice. To our knowledge, there is little research to date that has examined the role of service providers in assisting their customers in their co-production processes (Field et al. 2012, Retana et al. 2013), and none has focused on such phenomena immediately following the adoption of a service. We seek to narrow this gap

[11] Public cloud infrastructure services, or public Infrastructure-as-a-Service (IaaS), are a B2B SST in which on-demand computing and storage resources (i.e., servers) are offered on a pay-as-you-go basis (Mell and Grance 2011).

in understanding. Additionally, in our setting, more customers abandon the service during their first week than in any other week in their lifetimes, which makes retention during this period critical. Finally, customers' demand for technology support is frontloaded in the sense that they ask most of their questions in the periods immediately following adoption. This signals that customers face important service co-production costs during the initial setup or ramp-up stages relative to the rest of their lifetimes. Moreover, for the provider, this also implies that customers are most costly to serve soon after adoption, and, thus, reducing their demand for technology support during this stage may yield significant operational cost savings.

The potential influence of EPE on customer retention is not a trivial matter, especially in contexts such as ours where there are no contracts that lock customers in any way. On one hand, EPE can have a positive effect on retention. Prior work suggests education increases perceived service quality and satisfaction, both drivers of retention (Eisingerich and Bell 2008, Sharma and Patterson 1999). EPE may also lead to retention by setting realistic expectations for customers about the service features and performance, which in turn may lead to increased satisfaction of IS users and continued usage of IS (Bhattacherjee 2001, McKinney et al. 2002). Finally, education can also make customers more efficient in using the service, a factor shown to influence retention for other SSTs (Xue et al. 2007). On the other hand, there are those who suggest that educating customers may make them more capable and willing to consider alternate options in the market, which in turn may have a negative impact on retention (Fodness et al. 1993, Nayyar 1990). In sum, the early and proactive engagement may foster retention by increasing customers' perceived service quality, setting appropriate expectations, and aiding customers surpass the initial ramp-up stage, but it can also make them quickly realize the limitations of the service and consider defecting. Moreover, since soon after adoption the customers have not yet made any significant investment in co-producing the service (e.g., deployed a production application in the cloud) nor are there any contracts tying them to use the service for any period of time, they can switch away from the provider with ease. In other words, switching costs are very low and we can initially rule them out as the reason why customers would continue using the SST (Buell et al. 2010, Jones and Sasser 1995).

A similar set of opposing forces exists in regards to customer education and its effect on customer demand for (reactive) technology support. While education can make customers more efficient (e.g., they need less input to produce the service output) and in turn reduce the costs of serving them (Xue and Harker 2002), proactive education can also lead to escalated expectations, whereby customers continue expecting and seeking constant assistance from the provider to the extent that they become overly dependent on the provider (Challagalla et al. 2009). In other words, it is uncertain if EPE will reduce customers' demand for (reactive) technology support by making them self-sufficient, or if it increases it by making them provider-dependent. Our research setting, in which there is no cost for the customer to demand as much technology support as it wants (i.e., ask the provider as many questions as it wants), makes accessing the provider's support a very compelling and economically rational choice for customers, thus increasing the likelihood of the latter effect occurring.

To address our research question we collected unique data from a field experiment ran by a major public cloud computing infrastructure services provider during October and November 2011. Upon signup, 366 customers selected at random out of 2,673 customers that opened an account during this period received the field experiment's treatment: EPE. The treatment consisted in a short phone call followed up by a support ticket through which the provider offered initial guidance on how to use the basic features of the service. After the proactive engagement, the treated customers could continue interacting with the provider through reactive support, which was the only channel for technology support available for the non-treated control customers since signup. Our empirical strategy leverages the random assignment of the treatment and employs survival analysis and count data models to examine the differences in retention and demand for reactive technology support, respectively, between the two customer groups early on in their lifetimes. Our robustness checks thoroughly examine and validate the random assignment assumption, critical to our identification strategy.

We find that treated customers' hazard rate (i.e., number of customers who leave the service per unit of time) is about 49.60% lower than that of controls during the first week after adoption, and they are 3.1 percentage points more likely than the controls to survive through their first week. We argue that this is the case because customers' exposure to EPE increases customer satisfaction and enables them to derive a greater value from the service. Also, even just becoming familiar with how to use the basic functionalities of the service already constitutes a co-production skill that would be lost if they switched away. Moreover, our finding has a strong managerial implication for the provider. On average, 34.3% of new adopters abandon the service before 8 months of use. However, 18.8% of them (or 6.4% of all adopters) abandon during the first week, which is much more than in any other week. In other words, more customers abandon the service during the first week than during any other week in their lifetimes. Therefore, by improving customer retention during this critical stage in customers' lifetimes, EPE has a relevant positive impact on the overall size of the customer base.

We also test the effect of EPE on customers' early demand for technology support, as measured by the number of questions they ask to the provider through online live chat sessions and support tickets in the weeks following adoption. We use automated text parsing algorithms to distinguish between support interactions that correspond to questions on how to use the service (e.g., how to configure a server) and troubleshoot issues associated with the quality of the service offered by the provider (e.g., an unexpected hardware failure on the provider's end). EPE reduces the average number of questions asked during customers' first week after adoption by 19.55%. We argue that this occurs because in the early stages of the co-production process the provider can preempt customers' most frequently asked questions (FAQs). This is, again, an important economic benefit for the provider. Customers' demand for support is strongest when they are just starting to use the service and the drop in the number of questions implies a reduction in one of its major operational costs: the human labor-intensive offering of reactive technology support.

In addition to contributing to the services literature by advancing the ongoing debates concerning the potential dual effects of EPE on retention and demand for technology support, our work contributes by being, to our knowledge, the first to empirically study the effects of proactive education on customer behavior. In the marketing field, only recently was the concept of proactive post-sales service introduced, "which can be contrasted with the more prevalent approach of providing post-sales service in response to customer-initiated contacts, or reactive post-sales service" (Challagalla et al. 2009). Education can be considered one form of such service. In the operations field, and in particular within the context of contact centers, the offering of outbound (proactive) education is an unexplored alternative to combine or blend inbound (customer-initiated) and outbound (provider-initiated) calls. It has been suggested that the outbound calls can be used to call back customers (e.g., Armony and Maglaras 2004a, Armony and Maglaras 2004b), attend low priority work that can be postponed (e.g., Bhulai and Koole 2003, Gans and Zhou 2003), or cross-selling (e.g., Aksin and Harker 1999, Gurvich et al. 2009). Yet, to our knowledge, there is little research studying whether outbound calls can be an effective mean to educate users on how to co-produce a service. In a broader context, to our knowledge, we are also the first to measure education's effects on retention based on actual usage of a service and not just on customers' forward looking intentions to continue using a service captured through surveys (e.g., Bell and Eisingerich 2007, Huang and Zhou 2012).

Our study has very important managerial implications beyond our studied cloud infrastructure services context. The cloud is not the only setting in which initial technical difficulties affect customer retention and customers' technical abilities should not be taken for granted. For example, in the context of online learning programs, it has been suggested that, in addition to the actual content of the courses, first-time e-learners face the challenge of dealing with the technology and the interface of the e-learning sites (Muilenburg and Berge 2005). This challenge may lead to early attrition, especially "when technical support is not immediately available or easily accessed" (Tyler-Smith 2006). In other words, there exist other service contexts where SST providers can benefit significantly from proactively engaging customers, and in particular offering EPE. Moreover, the provider's EPE effort in our study is noteworthy as it challenges the general premise of cloud services being fully self-serviced, on-demand offerings with minimal interaction between customers and service providers (Mell and Grance 2011). Our research suggests that the cloud industry may actually benefit from not being exclusively "self-service."

3.2 Theory Development and Hypotheses

3.2.1 Early Proactive Education (EPE)

A key consideration that service providers must take into account when designing their offerings is their customers' co-participation in the service delivery process (Bitner et al. 1997, Mills and Morris 1986). Providers understand that the greater the level of contact the customer has with the service system, the more the quality and consistency of the service delivered depend on the customers' own skills and abilities (Chase 1978, Frei 2006). Self-service options (e.g., SSTs) that require no special skills have been suggested as an alternative to mitigate customers' influence on the service delivery output (Frei 2006). However, even in contexts with relatively simple service co-production processes, like online banking self-service portals, customers' capabilities are still very important. Research on this SST has shown that customers' capabilities are key determinants of their adoption and continued usage of the service portals (e.g., Xue and Harker 2002, Xue et al. 2011, Xue et al. 2007).

Moreover, there are emerging SSTs where the offering is far from a ready-to-use, turn-key solution, but more akin to a general purpose technology (Bresnahan and Trajtenberg 1995) that demands considerable adaptation or co-production efforts from customers for it to suffice their requirements. An excellent example of these are cloud computing infrastructure services, which some scholars (e.g., Brynjolfsson et al. 2010) have compared to a general purpose technology in part because of the wide variety of use cases they can be applied to (Venters and Whitley 2012). In such scenarios, where the co-production of the service requires relatively high levels of expertise, an alternate strategy to deal with customers' potential lack of skills is to educate them.

We define *customer education* broadly as any effort from the provider to increase its customers' service co-production related knowledge and skills, which, in turn, enable customers to derive a greater utility from the service.[12] A potential channel to educate customers, which is particularly suitable for technology services, is through the offering of technical support. We distinguish education through technical support from other more formal forms of education such as user training sessions or certification programs. Education through technical support occurs on a one-to-one basis. Recent research on the banking industry has explored the value of one-to-one, face-to-face interactions with customers in increasing their effectiveness in using the service (Field et al. 2012).

Moreover, technology support is generally offered through contact centers and in an inbound or reactive manner, in the sense that the one-on-one interactions occur when customers initiate them through a phone call, an online live chat session, email, or by submitting a support ticket. Instead, we focus on provider-initiated or proactive technology support, such that it is a means to offer *proactive customer education*

[12] See Burton (2002) and Aubert (2007) for thorough discussions of customer education and its definition.

(Challagalla et al. 2009). We emphasize that the proactive engagement has an educational purpose, and not a sales-oriented one, as is common in contact centers (e.g., Aksin and Harker 1999, Gurvich et al. 2009). More specifically, we focus on *early proactive education*, or EPE, defined as any provider-initiated effort to increase its customers' service co-production related knowledge and skills immediately after service adoption, which may enable them to derive a greater utility from the service. We discuss the potential tradeoffs for the provider in offering EPE in the following sections, as we develop our hypotheses.

3.2.2 EPE and Customer Retention

A recurring result in the services literature is the positive effect of customer education on perceived service quality (e.g., Bell and Eisingerich 2007, Burton 2002, Sharma and Patterson 1999). Extant research has also consistently shown that satisfaction, in turn, leads to customer loyalty (e.g., Bell and Eisingerich 2007, Sharma and Patterson 1999, Zeithaml et al. 1996).

In the particular context of IS, EPE can increase satisfaction and loyalty by helping customers match their expectations regarding the features of the SST and their early experiences with the service. Research based on the expectation-confirmation theory (Oliver 1980) that examines user satisfaction with IS has found that users are more satisfied if their expectations are confirmed by their experiences (Bhattacherjee 2001, McKinney et al. 2002), which in turn motivates them to continue using a service (Bhattacherjee 2001). Users also perceive a greater benefit from using IS when their experiences match their expectations (Staples et al. 2002). Moreover, in the context of internet-based services there is an added complication in managing customers' expectations given how rapid technologies evolve and thus how fast experiences may differ from expectations (Liu and Khalifa 2003). Such a challenge makes engaging customers early in their lifetimes through EPE especially valuable for the provider.

EPE can also incentivize customers to use a service by making them more efficient through a reduction in their early service co-production costs. Rather than requiring customers to invest in experimenting and learning how to use the basic functionalities of the service on their own, via EPE a provider can take that burden off customers, or at least make their initial ramp-up process less cumbersome. Prior work in the online banking context has found that more efficient customers are less likely to abandon an SST (Xue et al. 2007).

However, there exist circumstances in which education could lead to attrition rather than retention (Bell and Eisingerich 2007, Huang and Zhou 2012), particularly if it is offered soon after adoption and when there are near zero switching costs. When customers learn from the provider, the information asymmetry between them gets reduced and the former may be motivated to evaluate other alternatives in the market (Fodness et al. 1993, Nayyar 1990), thus increasing their likelihood of leaving. For example, EPE can make customers aware of some limitations of the service they did not know before the treatment. The increase in the likelihood of leaving will be especially important if EPE is not sufficiently effective in driving satisfaction. Prior work has suggested that customers who are not necessarily satisfied with an SST continue using it because of switching costs (Buell et al. 2010, Jones and Sasser 1995). However, in our setting, not only are there no contracts that lock customers in for a certain period of time (e.g., a subscription), but also early in their lifetimes customers have not yet incurred any large co-production effort (e.g., invested in deploying an application in the cloud service) that may represent some form of switching barrier. Therefore, in the absence of any switching costs, unless EPE has a strong influence on satisfaction, the risk of attrition is particularly high.

Despite the potential negative effects of EPE on customer retention, we suggest that EPE will have a positive effect on customer retention during the early stages of their co-production process because (i) customers will derive more value and will be more satisfied from using a service they understand better due to the treatment and, additionally, (ii) even just becoming familiar with the service and learning how to use its basic functionalities already constitutes a co-production skill learned that would be lost if they left. In other words, the treatment generates a small yet important switching cost that motivates customers to continue using the service. Formally:

HYPOTHESIS 1: EPE is positively associated with customer retention in the period immediately after adoption of the service.

3.2.3 EPE and Demand for Technology Support

Educating and improving the efficiency of customers in using the service can lead to a reduction in costs for the provider since it will employ less labor and other resources in the service delivery process (Xue and Harker 2002). In the particular context of technology support contact centers, the provider's initial investment in EPE could potentially lead to a reduction in later reactive support costs by reducing the number of questions asked by customers through reactive support channel (e.g., customers call in less frequently). For example, by guiding customers on how to navigate through the service control panel, the provider can preempt questions regarding its functionality as customers initiate their service co-production processes.

Nevertheless, education, and in particular EPE, could also have the opposite effect. EPE can lead customers to realize early on that the provider constitutes a reliable, fast and easy-to-access knowledge source, especially if, as in our context, there are no additional fees associated with contacting the provider. Thus, customers who have revived EPE may become more aware of the provider's support capabilities and realize that it is much more efficient or convenient for them to constantly contact the provider for assistance, instead of attempting to search knowledge bases or experiment to solve their issues on their own. This, in turn, would result in customers becoming overly-dependent on the provider (Challagalla et al. 2009) and increase customers' demand for technology support. A similar result was found in the context of insurance services where presenting customers with more information increased, rather than decreased, the number of calls they made to a call center (Kumar and Telang 2012).

Although the risk for this last effect exists, we argue that customers' demand for technology support (i.e., the number of questions they ask through reactive support channels) soon after adoption will be reduced by EPE because of the provider's ability to preempt the questions customers generally have during this stage of their lifetimes (e.g., the frequently asked questions, or FAQs). Moreover, at this stage customers will not have developed any dependency habits that will lead them to increase their demand for technology support. We thus hypothesize:

HYPOTHESIS 2: EPE is negatively associated with customer demand for technology support in the period immediately after adoption of the service.

3.3 Research Setting and Field Experiment

Our research examines the effects of EPE on customer churn and demand for technology support by analyzing the outcome of a field experiment executed by a major cloud provider during October and November 2011. We will describe our data in detail later in section 3.5, but elaborate on the characteristics of our context and the field experiment here.

Cloud computing has been defined by the US National Institute of Standards and Technology as a "model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction" (Mell and Grance 2011).[13] We place emphasis on the last component of the definition, which is also one of these services' essential characteristics: cloud services are on-demand and self-serviced, whereby

[13] This constitutes the 16th and final version of "The NIST Definition of Cloud Computing."

customers are expected to use the service on their own and not have to interact with the provider to do so. Moreover, in cloud infrastructure services, or Infrastructure-as-a-Service offerings, such as that of our provider, "the consumer does not manage or control the underlying cloud infrastructure but has control over operating systems, storage, and deployed applications; and possibly limited control of select networking components" (Mell and Grance 2011). In other words, there is a significant technical burden on customers for the appropriate service co-production process that starts with aspects as simple as understanding the new technical jargon (e.g., Amazon Web Services 2013, Shinder 2010). Given these characteristics, it is reasonable to suspect that not all customers will find it easy to start using these services as they adopt them nor will they have clear expectations of what they will experience as they use the service. These potential gaps in customers' skills and expectations are the reasons why engaging customers through EPE may have a positive effect on their behavior.

In our setting, the provider's new customers sign up for the service and open an account through an online form without any cost; they only pay when they start consuming hardware resources (e.g., launch a server). Then, a few minutes after signup, they receive a call by an agent of a verification team that, to the best of his/her abilities, ensures the new account was opened by a legitimate customer (e.g., a customer that will not use the service to spam). Agents of the provider's verification team call prospective customers following a simple first-come first-serve (FCFS) queue. If they pass the verification process, customers can start using the on-demand servers, although without any explicit starting guidance from the provider aside from online documentation and manuals.

For the field experiment, a few designated agents of the verification team performed additional tasks beyond verifying the legitimacy of the new adopters: they applied the EPE treatment. Because the incidence of treatment is determined without any a priori information from the customers and by the FCFS queue, we can almost consider it as applied by random assignment. There were, however, small variations in the proportion of agents applying the treatment at different times during the field experiment. The number of designated agents applying the treatment relative to the total size of the verification team varied across work shifts, across days of the week, and also over weeks of the year. This, in turn, slightly altered the likelihood of receiving the treatment contingent on the time of adoption. Since the random assignment of the treatment lies at the core of our econometric approach, we will implement several controls to account for this later in section 3.4.

The designated agents who applied the treatment prolonged the verification call and followed it up with a support ticket in order to explain to the new customers what the basic features of the cloud service are and how to employ them. The provider's overall goal in offering EPE was to proactively facilitate customers' access to the knowledge they needed to get started in the use of the service. Specifically, the treatment had three components in addition to the fraud verification process: confirming product fit, setting expectations, and educating customers.[14] The first two clearly address potential negative expectations disconfirmation. As suggested before in our theory development, this should be positively associated with customer retention. In regards to customer education, during the call and through the support ticket the agent sought to teach the customer aspects such as how to access and use the online control panel, how to setup and access her first server, and how to make a backup of that server, among other basic topics. Although these constitute only basic functionalities of the service, the EPE prevented customers from having to investigate and learn them on their own, thus lowering their co-production costs and increasing their efficiency, as well as providing them with a skill level that they would relinquish, at least partially, if they opted to switch to some other provider.

When asked about the rationale for offering EPE, an executive from the provider

[14] For full details of the topics covered as well as the template used for the support ticket, please refer to Appendix F.

noted that their reactive technology support agents were approached too often by customers with very basic questions regarding the service's features, and not necessarily by young customers. The executive manifested two concerns about this that EPE would solve. First, if customers were asking these basic questions several weeks after adoption, then it was very likely that they were not playing well their part in the service co-production process. That is, they were self-servicing themselves a degraded service experience that could be having negative effects on their satisfaction and increasing their risk of churning. Second, any question that can be pre-empted and addressed in a proactive manner represents a reduction in the demand for reactive technical support, which given its uncertainty is more costly to offer. Proactively answering customers' questions when idle agents are available to call them is less costly than staffing sufficient technicians to cope with the peaks in the uncertain demand for reactive support.[15] After the provider concluded the field experiment, it decided to continue applying the EPE treatment to all newly adopting customers.

3.4 Empirical Models

3.4.1 Survival Analysis

In order to test the effects of EPE on customer retention we employ both non-parametric and semi-parametric survival analysis methods. However, we start with simpler linear probability and probit models.

Our first approach consists of examining the effect of the treatment on the likelihood of a customer surviving up to a certain age. In particular, let $SurvivedWeek1_i$ and $SurvivedMonth1_i$ be binary indicators that are turned on if the customer uses the service (survives) for at least 1 week (7 days) or 1 month (30 days), respectively. We use

[15] This is an aspect widely discussed in the literature associated with contact centers that blend (combine) inbound (i.e., customer-initiated) and outbound (i.e., provider-initiated) engagements with customers. See Aksin et al. (2007) for a review.

them as our dependent variables in both a linear probability and a probit model as follows (we use $SurvivedWeek1_i$ below, yet the model is the same with $SurvivedMonth1_i$):

$$SurvivedWeek1_i = \alpha + \beta\ EPE_i + \delta\ AdoptControls_i + \varepsilon_i, \quad \text{and} \qquad (1a)$$

$$\Pr\ (SurvivedWeek1_i = 1) = \Phi(\alpha + \beta\ EPE_i + \delta\ AdoptControls_i + \varepsilon_i). \qquad (1b)$$

Parameter EPE_i is our main regressor of interest. It is a binary variable equal to 1 if customer i received the EPE treatment, and is 0 otherwise. Thus, under the assumption of a randomly assigned treatment, the coefficient β identifies EPE's effect and is expected to be positive per Hypothesis 1.

We additionally have a set of control dummies, $AdoptControls_i$, that account for two aspects associated with customers' time of adoption (Appendix H contains figures used in this discussion). First, they control for potential unobserved systematic differences across customers contingent on their time of adoption. Second, as described before the number of agents applying the treatment changed slightly over time (see Figure H.4). Consequently, and given the FCFS queue, the proportion of newly adopting customers that was treated and the likelihood of any new customer receiving the treatment also varied slightly over time. Controlling for this latter aspect is critical for our random assignment assumption.

Our first controls are in the vector $AdoptHour_i$, which consists of 23 dummies, one for each hour of the day at the provider's time zone (we leave the 24th hour as the base level). On the customers end, we learned through interviews with the provider that customers adopting during regular business hours (e.g., from 8:00 a.m. to 6:00 p.m.) tend to be systematically different from those adopting after hours.[16] While the former will likely be working at a firm, the latter may be individuals working on personal projects. On the provider's end, our examination of the data revealed that between 5% and 14% of

[16] Despite our data is from a global provider, the concept of "office hours" remains valid as both the provider and the vast majority of its customer base are in the United States. Moreover, in order to not fix our dummies to any particular time zone within the United States, we use hourly dummies instead of a single "office hours" dummy.

adopting customers were treated from 12:00a.m. to 8:00 a.m., between 10% and 23% of adopting customers were treated between 8:00 a.m. and 6:00 p.m., and between 9% and 14% of adopting customers were treated from 6:00 p.m. to 12:00 a.m. (see Figure H.5). Therefore, the chances of being called by an agent that is applying the treatment vary slightly across the work shifts. Our vector of hourly dummies intends to control for both of these considerations.

Our second element in $AdoptControls_i$ is a single binary variable, $AdoptWeekday_i$, that is equal to 1 if adoption occurred from Monday through Friday and is zero for weekend signups. The rationale about systematic differences in customers depending on whether they adopt during a weekday or on the weekend is the same as the aforementioned rationale for controlling for adoption during business hours and after business hours. We also found a higher proportion of treated signups during the weekends (see Figure H.6). Our weekday control accounts for both of these potential issues.

Finally, we have a vector of weekly dummies, denoted as $AdoptWeek_i$. This vector has the goal of controlling for universal time shocks such as how close the time of adoption (which occurs between October and November 2011) is to the 2011 Holiday season. The vector also controls of a regime change on the provider's end whereby starting on November 13th (week 47 of the year) a greater proportion of agents in the verification team were applying the treatment than before (see Figure H.4 and Figure H.6).

In sum, we have:

$$\delta\ AdoptControls_i = \delta_1\ AdoptHour_i + \delta_2\ AdoptWeekday_i + \delta_3\ AdoptWeek_i.$$

Next, we employ non-parametric survival analysis to determine the overall effect of the treatment on customer retention. For this, we represent the rate at which customers fail (churn) at time t through the hazard function

$$h(t) = \lim_{\Delta t \to 0} \frac{\Pr(t < T < t + \Delta t | T \geq t)}{\Delta t},$$

where T is the nonnegative random variable denoting the time to a failure (churn) event. We use the log-rank (Mantel and Haenszel 1959) and Wilcoxon (Breslow 1970, Gehan 1965) tests for the equality of hazard functions between the treated and control customer groups. The log-rank test is particularly powerful when the hazards are not equal but instead are proportional to one another, as we shall assume and test later, while the Wilcoxon test places more weight to earlier failure times (Cleves et al. 2010). This upfront weighting is important since, as we mentioned earlier in our introduction and will discuss in great detail below, in our context retention is weakest (i.e., likelihood of failure is highest) during the early stages of customers' lifetimes.

Nevertheless, neither of these non-parametric approaches tests for the equality of the survivor functions at some point in time; they test for the equality across the entire timespan of the data. In order to distinguish the time-varying effects of the treatment (e.g., its decay) we must make some parametric assumptions. In the Cox (1972) proportional hazard model, the hazard for the i^{th} customer at time t is

$$h(t|X_i) = h_0(t) \exp(X_i \beta_X).$$

In this model, we assume that all individuals are subject to the same underlying baseline hazard, $h_0(t)$, yet we make no assumptions regarding its functional form. Instead, we simply assume that the treatment and other covariates in the vector X_i influence the baseline hazard in a multiplicative (proportional) way. We first employ a model to test if the treatment has a negative effect on the overall hazard, still without any time-varying considerations:

$$h(t|X_i) = h_0(t) \exp(\, \beta \, EPE_i + \delta \, AdoptControls_i \,). \tag{2a}$$

With this model, $e^{\hat{\beta}} - 1$ will be the estimated percentage change in the hazard (failure) rate caused by the treatment. Hypothesis 1 suggests $\hat{\beta} < 0$ since a decrease in the hazard rate implies an increase in customer retention. Model (2a) allows quantifying the overall effect of the treatment on the customers' hazard rates, conditional on the treatment and potential unobserved systematic differences which we believe may be associated with the customers' time of adoption. However, it is reasonable to expect the treatment's effect to decay over time. In other words, the marginal effect of β in Model (2a) should be smaller as customers grow older.

We can model how the effect of the treatment decays over time by employing a vector of weekly age dummies and interacting it with the treatment parameter (Cleves et al. 2010). By using a vector of dummies (i.e., one per week in customers' lifetime) rather than some other functional form (i.e., a linear or logged tenure parameter), we let the data determine how exactly the treatment's effect decays. Let t index days in a customer's lifetime. Then, let

$EPE\_Weekn_{it} = EPE_i \times 1[(n-1) \times 7 < t \leq n \times 7]$, $n = 1, 2, 3, \ldots, p$, and

$EPE\_OtherWeeks_{it} = EPE_i \times 1[t > p \times 7]$.

be those dummies. Thus, for all treated customers (i.e., $EPE_i = 1$), $EPE\_Week1_{it} = 1$ for week 1 and is zero otherwise, $EPE\_Week2_{it} = 1$ for week 2 and is zero otherwise, and so forth until week p. Then, $EPE\_OtherWeeks_{it} = 1$ only for any remaining weeks beyond the p^{th} week (i.e., $t > p \times 7$). Note that the number of weeks in which $EPE\_OtherWeeks_{it}$ is turned on depends on the number of individual weekly dummies used (p). All these dummies are turned off for the controls. Finally, let our model with time-varying parameters be:

$$h(t|X_i) = h_0(t) \times \exp\left(\sum_{n=1}^{p} \beta_n\, EPE\_Weekn_{it} + \beta_{t>p} EPE\_OtherWeeks_{it} + \delta\, AdoptControls_i\right). \quad \text{(2b)}$$

In Model (3), each β_n will identify the treatment's effect during week n, and $\beta_{t>p}$ the effect in weeks after the p^{th} week. Our expectation as per Hypothesis 1 is that only the estimates of the coefficients for the early weeks (i.e., $\hat{\beta}_n$ for low n) will be negative and significant. We use a sufficiently large p (i.e., $p = 8$, little under 2 months), such that it is reasonable to expect the treatment having no to little effect past p.

3.4.2 Demand for Technology Support

To estimate the effect of the treatment on the demand for technical support, we employ count data models that have the number of questions asked by customers during the initial stages of their lifetimes as dependent variable. We describe what exactly constitutes a question and offer descriptive statistics later in section 3.5.2. For now we note that these questions constitute customers' demand for reactive technology support and, given that it is a human-intensive operation, represent a very important cost driver for the provider. Moreover, we focus on customers' early questions (i.e., one or two weeks after adoption) since this is when customers ask most of their questions and also the time frame in which EPE is expected to have a negative effect on their frequency, as per Hypothesis 2.

Count data models, such the Poisson and negative binomial models we employ, account for the number of questions asked being a nonnegative integer value. However, since most customers do not ask any questions at all, our distribution has a large number of zeroes and hence more than likely suffers from overdispersion (i.e., the conditional variance is larger, and not equal, to the conditional mean). To account for this, we relax the equivariance assumption of the Poisson model and employ the quasi-maximum likelihood approach that uses a robust variance-covariance matrix for the Poisson maximum likelihood estimator. We also use the negative binomial model (with quadratic variance), which despite making more assumptions on the functional form of the distribution than the Poisson model, may fit our data better as it explicitly models overdispersion as well as a longer right tail in the probability distribution (Cameron and Trivedi 2010).

Let $QuestionsWeek1_i$ and $QuestionsWeek2_i$ represent the number of questions asked by customer i during its first week or first two weeks after signup, respectively. Since both the Poisson and the negative binomial models have the same conditional means, we present the same model for both estimation methods (we show the model with $QuestionsWeek1_i$, yet the model is the same with $QuestionsWeek2_i$):

$$E(QuestionsWeek1_i|X_i) = \exp(\alpha + \beta\ EPE_i + \delta\ AdoptControls_i + \varepsilon_i). \qquad (3)$$

Parameter β identifies the treatment's effect on the demand for technology support, and we expect $\hat{\beta} < 0$ per Hypothesis 2. However, we acknowledge that our specification may suffer from attrition bias since we are comparing volumes of questions over certain periods of time and some customers may leave the service before the completion of such periods. Moreover, since EPE (negatively) influences attrition, parameter β in model (3) captures the treatment's effect on both the number of questions and on attrition. Hence, we further acknowledge that our specification cannot completely separate these two effects.

3.5 Data

3.5.1 Data Source and Sample Construction

We have collected detailed data on a field experiment executed by a major public cloud infrastructure services provider from October 11 to November 28 of 2011. For every account opened during this period, we observe if it was treated, its use of the on-demand infrastructure services, and the timing and content of all support interactions (i.e., online live chat sessions and support tickets) between the customer and the provider. We observe the latter two aspects, server usage and support interactions, up to August 15, 2012 (i.e., between 8 and 9 months of history depending on day of adoption); this is relevant to our identification of failure (churn) as will be discussed shortly.

During the duration of the field experiment the provider opened 4,739 new accounts for its cloud infrastructure service, of which 744 received the EPE treatment. However, not all new accounts are suitable for our analyses. In what follows we describe the cropping criteria used to construct our final sample with 2,673 customers, 366 of which were treated.

First, the accounts can be opened by customers under two different support service levels: full and basic.[17] Full support involves frequent interactions between the customers and their assigned account managers, whereas basic support is limited to addressing quality of service issues. Thus, it is unlikely that the EPE treatment will have a substantial effect on full support customers as the treatment would be considered just the first of many rich and frequent interactions. Therefore, we focus on customers that exclusively used basic support through their lifetimes and drop those who started with or upgraded to full support (86% of upgraders adopt full support within their first month since adoption). After the exclusion, we are left with 4,194 basic support accounts (596 treated and 3,598 controls).[18]

Second, not all accounts are opened with the intention of using the service. Some of the accounts have very short lifetimes (i.e., less than 1 day) or never launch a server. Through interviews with the provider, we learned this is a common occurrence. It is often the case that a customer (e.g., an intern at a firm) opens an account simply to check if the provider's platform supports some particular feature. The customer opens the account, checks for the availability of the feature, and then never launches a server. Other customers do intend to use the cloud service, but not for legit purposes (e.g., for spamming). When detected by the provider, they do not pass the verification process. Finally, some accounts are opened by the provider's own internal staff to test the functionality of their system or for support agent training purposes. Indicating their corresponding count in parentheses, we exclude from our sample the accounts that did not live for more than 1 day (312), never launched a server (596), were flagged as illegal

[17] Chapter 2 of this dissertation discusses these two support levels and their impact on customer behavior.
[18] We note that, although the exclusion of the full support customers aids our identification of the treatment effect by reducing heterogeneity in the sample, our results remain consistent if they are kept in the sample, as we show in Appendix I.

accounts (138), or were opened by a provider's employee (71). We point out that the accounts' attributes are not mutually exclusive. The process removes a total of 923 accounts, leaving 3,395 accounts (457 treated and 2,938 controls).

Finally, some of these new accounts may correspond to a new account of a pre-existing customer. Since experienced customers that are now opening a new account presumably already achieved a basic level of proficiency with the service, EPP should have negligible effect on them. We use the contact information of the accounts to determine if any of the new accounts may belong to a pre-existing customer, and drop them from the sample. We are left with 2,673 customers (366 treated and 2,307 controls) who constitute our sample.

3.5.2 Measurement of Variables and Descriptive Statistics

We present in Table 3.1 the descriptive statistics of the variables used in our analysis and discuss their construction next.

Our main covariate of interest is the treatment, captured in EPE_i. Around 13.7% of the sample (366 customers) received the early education treatment upon signup.

Table 3.1: Descriptive Statistics

Customer Group Number of Customers	All Customers 2,673				Controls 2,307				Treated 366				t-test of mean difference	
Variable	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max	Diff.	p-value [a]
EPE_i	0.137	0.344	0	1	0	0	0	0	1	0	1	1		
$SurvivedWeek1_i$	0.936	0.245	0	1	0.931	0.253	0	1	0.964	0.185	0	1	-0.033	0.016
$SurvivedMonth1_i$	0.888	0.316	0	1	0.881	0.324	0	1	0.932	0.253	0	1	-0.051	0.004
$Failed_{it}$	0.374	0.484	0	1	0.383	0.486	0	1	0.314	0.465	0	1	0.069	0.011
$QuestionsWeek1_i$	0.591	1.445	0	18	0.609	1.497	0	18	0.481	1.054	0	8	0.128	0.115
$QuestionsWeek2_i$	0.759	1.851	0	25	0.777	1.896	0	25	0.647	1.536	0	15	0.129	0.215

[a] Reported p-value corresponds to two-tail t-test. Difference for $QuestionsWeek1_i$ with a one-tail t-test is significant with p-value 0.058.

Next we have the parameters associated with customer retention, which all depend on the accurate identification of a failure (churn) event. This can be potentially difficult in our on-demand cloud services context as customers may cease use of the service without necessarily closing their account. Given this, we identify the moment in which a customer stops using the service (the failure) by the customer's last observed usage of a cloud server or last observed support interaction with the provider, whichever comes last. We acknowledge that some of the customers marked as failed may return after our observed period. However, we have no reason to believe that such noise can be systematically associated with the treatment, particularly since our observation period ends between 8 and 9 months after adoption. Parameters $SurvivedWeek1_i$ and $SurvivedWeek2_i$ are binary variables that indicate if customer i uses the service for at least 1 week (i.e., $t > 7$) or 1 month (i.e., $t > 30$), respectively. Parameter $Failed_{it}$ is another binary variable that is set to 1 if the customer i stopped using the service at time t, and is 0 otherwise. During our observed period, 37.4% (999) of the customers in the sample failed, which implies that 62.6% (1,674) of the customers are retained and survive until the end of the data. The mean of our survival indicators suggest 93.6% of customers survive past the first week and 88.8% of do so past the first month.

However, in order to better examine when customers are churning we use the Kaplan and Meier (1958) survivor function estimate, which we plot in Figure 3.1. The vertical axis represents that proportion of customers that are still using the service, and the horizontal axis represents the days in customers' lifetimes (t). The steeper slope during the early periods indicates that a customer is more likely to fail (i.e., abandon the service) during early stages of its lifetime than later. Given a set of customers that adopt the service at a certain point in time, the estimated survivor function suggests that the provider is losing 5% of them by day 3 in their lifetimes, and 10% by day 20. From day 20 onwards, the provider loses about 3.2% (81.11 customers) of those adopters, in a trend that declines from 4.8% early on down to 2.3% after 9 months of age.

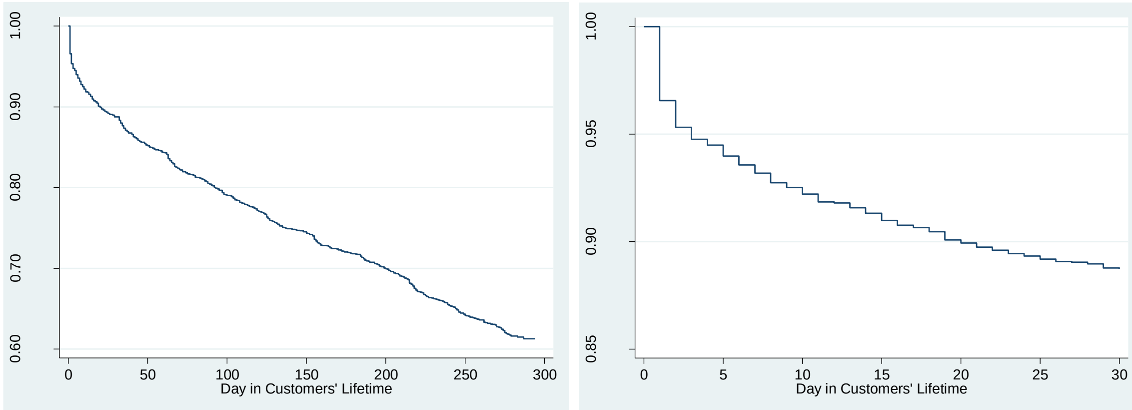

Figure 3.1: Kaplan-Meier Survival Estimates for all data (left) and first 30 days (right)

We operationalize customers' demand for technology support through variables $QuestionsWeek1_i$ and $QuestionsWeek2_i$, which represent the count of the number of questions asked by customers during their first week and first 2 weeks (i.e., 7 and 14 days) since adopting the service, respectively. In order to consider a support interaction as a question it must satisfy two requirements. First, it must have been initiated by the customer. While all chats can only be initiated by a customer, some support tickets are announcements or alerts sent out by the provider through the support ticketing system. To identify such announcements, we scanned the tickets' content and flagged as announcement those that were either identical (i.e., exact same subject and content) or that followed a certain template (e.g., an automated message where the customers' name in the first line is based on the account information but the rest of the content is identical). The list of subjects used to identify these provider-initiated tickets is presented in Appendix G. Second, the support interaction must not represent a response to an exogenous and unexpected failure in the service offering (e.g., a physical component of the provider's hardware fails). Such an interaction would be associated with troubleshooting the incident and not with a customer's inquiry on how to use or configure a feature of the cloud service. We took special care in identifying problems caused by the provider and not problems that arise when customers are trying to perform some action (e.g., change networking configuration of a server). For details on how we identified these failures, please refer to the support interactions flagged as $FailOutage$, $FailNetwork$, and $FailHost$ in section D.1 of Appendix D.

Before discussing the descriptive statistics of the number of questions we comment on two issues concerning their adequacy in capturing customers' demand for technology support. First, the provider interacts with its customers through three support channels of which we only observe two: we observe all online live chat sessions and support tickets in customers' lifetimes, but we do not observe their phone calls. More precisely, we observe the aggregate volume of phone calls but are unable to link each individual phone call to a specific customer. However, analysis of the aggregate data indicates that roughly 60% of support interactions occur through chats, 20% through tickets, and 20% through phone calls. Thus, although we do not observe phone calls, we are only missing a small proportion of the total number of support interactions. Moreover, we know that some phone calls are followed up by a support ticket, such as when the support agent wants to transmit some information to the customer (e.g., some step-by-step guide on how to configure some component of the infrastructure), which in turn means we do capture the interaction through the support ticket. Second, the count of support interactions does not offer insight into the complexity or topic of the questions asked, attributes that may affect the provider's cost of offering the reactive support. Although we have made an effort to cleanly identify support interactions that constitute questions, our counts consider all questions to be equally costly to answer. This can be remedied by further data analysis, as we comment in our conclusion.

As mentioned before, the distribution of the questions asked is frontloaded relative to customers' lifetimes. The mean number of questions during the first and second weeks of customers' lifetimes are 0.591 and 0.179, respectively, while the metric drops below 0.103 for all other weeks. Table 3.2 and Figure 3.2 show descriptive statistics for the number of questions asked by customers through their first 8 weeks.

Table 3.2: Number of Questions Asked by Week in Customers' Lifetimes

Week in Lifetime	Customers [a]	Mean	Obs. (%) with zero value	Median	75th Percentile	90th Percentile	95th Percentile
1 [b]	2,673	0.591	1,930 (72.2%)	0	1	2	3
2	2,501	0.179	2,286 (91.4%)	0	0	0	1
3	2,448	0.103	2,304 (94.1%)	0	0	0	1
4	2,404	0.093	2,285 (95.0%)	0	0	0	0
5	2,380	0.066	2,284 (96.0%)	0	0	0	0
6	2,343	0.065	2,259 (96.4%)	0	0	0	0
7	2,307	0.042	2,246 (94.4%)	0	0	0	0
8	2,284	0.046	2,216 (97.0%)	0	0	0	0
1 and 2 [c]	2,673	0.759	1,824 (68.2%)	0	1	2	4

[a] Count of customers using service at least during the first day of each week.

[b] Same as $QuestionsWeek1_i$. [c] Same as $QuestionsWeek2_i$.

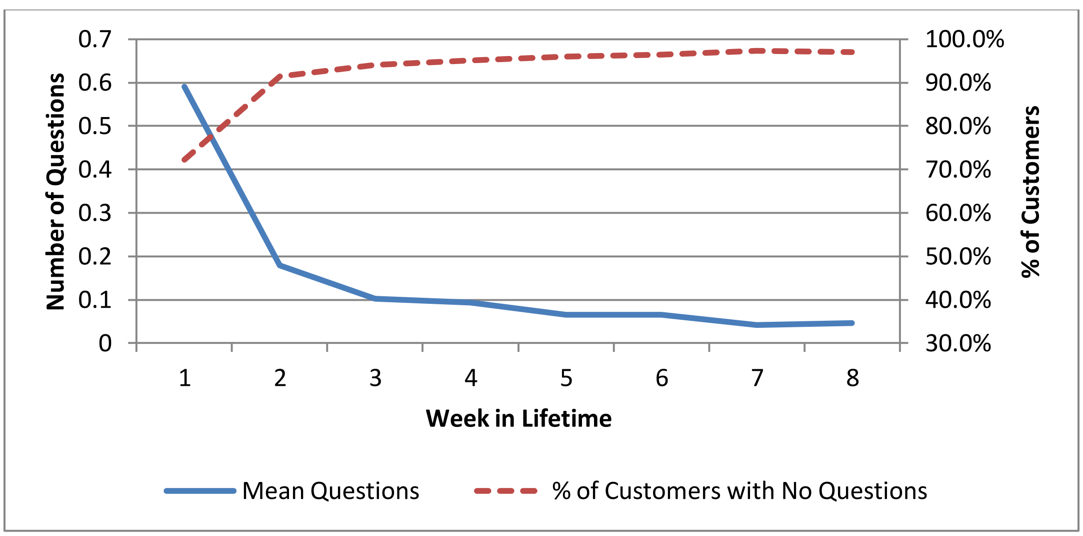

Figure 3.2. Number of Questions Asked by Week in Customers' Lifetimes

3.6 Results

3.6.1 Survival

The results attained using our various models all offer support for our first hypothesis: The EPE treatment has a positive effect on customer retention. Simple tests of the difference in means of our three retention-related dependent variables ($SurvivedWeek1_i$, $SurvivedMonth1_i$, and $Failed_{it}$) show early evidence of this (see two right-most columns of Table 3.1).

We present our initial results with the linear probability model (1a) and the probit model (1b) in columns (1) through (4) of Table 3.3. The marginal effect of the treatment for each model type is reported in the lower section of the table. Since EPE_i is a binary indicator, we compute the average marginal effect using the finite-difference method to appropriately capture the discrete change in the probability when the covariate changes from 0 to 1. Columns (1) and (2) suggest the treatment increases the likelihood of a customer surviving at least its first week of using the service between 3.1 and 3.2 percentage points. These results show that the treatment is effective in increasing customer retention during the early days after adoption. To put this estimate in perspective, it is useful to recall that mean retention for the sample after the first week (i.e., mean $SurvivedWeek1_i$) is 93.6% (see Table 3.1), so EPE brings survival rate much closer to 100% during the period in which it is most likely for customers to churn.

If we extend our analysis to survival through at least the first month we get very similar results. Columns (3) and (4) indicate treated customers are between 5.2 and 5.3 percentage points more likely to survive past the end of their first month. Although the effect is greater in magnitude than that for the first week, the probability of survival does not grow too much relative to that of the first week, which is initial evidence of decay over time of EPE's effectiveness.

Table 3.3: Survival Results

Column	(1)	(2)	(3)	(4)	(5)
Dependent Variable	$SurvivedWeek1_i$		$SurvivedMonth1_i$		$Failed_{it}$
Model	LPM	Probit	LPM	Probit	Cox Prop. Hazard
EPE_i	0.032***	0.308**	0.053***	0.323***	-0.254**
	(0.011)	(0.133)	(0.015)	(0.109)	(0.102)
Observed Failures					999
Marginal Effect of EPE_i	0.032	0.031	0.053	0.052	
% Change in Hazard $\left(e^{\hat{\beta}} - 1\right)$					-22.46%

All regressions use the 2,673 customers in the sample and include hourly, weekday, and weekly dummies. Robust standard errors in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Our result for the overall effect of EPE on the hazard rate employing the Cox proportional hazard model is presented in column (5) of Table 3.3. Under the assumption that EPE has a constant marginal effect throughout customers' lifetimes, the treatment slows the hazard rate by 22.46% (i.e., $e^{-.254} - 1$). This proportion reflects how much more treated customers were retained relative to the controls from adoption until the end of our data.

Before we investigate in greater detail the decay of the treatment's effect, we explore the robustness of our econometric specification. Our identification strategy hinges on the random assignment of the treatment. In particular, we assume that the provider did not consider any customer attributes to choose which customers received the treatment. To validate this assumption, we use data from an optional survey administered to customers at the moment of signup. In this survey, customers may indicate some attributes of themselves such as their size (i.e., employment), industry, and their intended use case for the cloud infrastructure services (e.g., e-commerce site, social media site, or back office application). The answers to this survey, along with the contact information of the account holder, are the only pieces of information the provider has about its new customer before the account verification process. The survey had a 22.6% response rate and, since the survey is administered as part of the online signup process before the verification agent calls (and potentially applies the treatment), the likelihood of responding is not associated with the treatment. We did not find any systematic differences in the customers' attributes (i.e., in the answers to the survey's items) between the control and the treated customers (see section H.1 of Appendix H for details).

As an additional robustness check for our random assignment assumption, in what follows we run the Cox proportional hazard model in column (5) of Table 3.3 using alternative sets of time-of-adoption controls. As described before, although the provider did not use any information to choose to which customers give EPE, the proportion of treated customers did vary through the field experiment's timespan. Our goal in using alternate controls is to further explore the drivers of the likelihood of receiving the treatment as well as is to de-saturate the model which at this time has 34 dummies as controls (i.e., 23 hours, 1 weekday, and 7 weeks). The use of too many dummies, particularly when they are collinear with each other, makes it more difficult to identify the effect of the covariate of interest (Hall et al. 2007).

We perform three changes on the $AdoptContorls_i$ vector (please refer to prior section 3.4.1 for details on its initial construction). In this paragraph we again refer to figures in Appendix H. First, we substitute the vector of hourly dummies $AdoptHour_i$ for three 8-hour shift dummies, $AdoptShift_i$, since a daily work shift pattern is clearly present in the data (see Figure H.4 and Figure H.5). Second, we substitute the individual weekly dummies, $AdoptWeek_i$, for a binary indicator, $AdoptRegime_i$, that signals the regime change observed in the data starting on week 47 when more agents started applying the treatment (see Figure H.4 and Figure H.6). Moreover, since some of the new agents worked on different days of the week than the ones already applying the treatment, the regime change does not only affect the weekly ratio of treated signups but also has implications regarding the proportion of treated customers per day of the week. For example, before week 47 no customers were treated on Sundays (Figure H.8), while many customers were treated on Sundays starting on week 47 (Figure H.9). Therefore, our third change implies interacting the regime indicator ($AdoptRegime_i$) with our weekday indicator ($AdoptWeekday_i$). We show the results with all possible permutations of the aforementioned controls in Table 3.4. Column (1) shows the same result in column (5) of Table 3.3 and serves as baseline for comparison; we also include a version without any adoption controls in column (9). The results across the adoption control permutations are consistent with each other, which strengthens our assumption that the treatment is as good as random.

We additionally followed the recommendations by Cleves et al. (2010) and performed various tests for our proportional-hazards assumption. Specifically, we performed a link test and confirmed that the coefficient β_2 in the model $X_i\beta_X = \beta_1(X_i\hat{\beta}_X) + \beta_1(X_i\hat{\beta}_X)^2$ is insignificant. We also interacted our treatment parameter EPE_i with time t and confirmed insignificance of the interaction. Finally, we confirmed that our predicted hazards have a non-zero slope in a generalized linear regression of the scaled Schoenfeld (1982) residuals.

Table 3.4: Survival Results with Varying Sets of Time-of-Adoption Controls

Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
EPE_i	-0.254**	-0.259**	-0.245**	-0.252**	-0.255**	-0.261**	-0.245**	-0.251**	-.0239**
	(0.102)	(0.101)	(0.101)	(0.100)	(0.102)	(0.101)	(0.101)	(0.100)	(0.098)
% Change in Hazard $(e^{\hat{\beta}} - 1)$	-22.46%	-22.86%	-21.75%	-22.24%	-22.53%	-22.94%	-21.71%	-22.19%	-21.24%
Controls Used									
$AdoptHour_i$	✓		✓		✓		✓		
$AdoptShift_i$		✓		✓		✓		✓	
$AdoptWeekday_i$	✓	✓	✓	✓	✓	✓	✓	✓	
$AdoptWeek_i$	✓	✓			✓	✓			
$AdoptRegime_i$			✓	✓			✓	✓	
$AdoptWeekday_i \times AdoptWeek_i$					✓	✓			
$AdoptWeekday_i \times AdoptRegime_i$							✓	✓	

All regressions employ Cox Proportional Hazard model and use all 2,673 customers in sample. There are 999 observed failures.
Robust standard errors, clustered on customers in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

3.6.2 Decay of Treatment Effect on Survival

In this section we explore how long after adoption does the treatment still influences customer retention.

We start addressing this question using the log-rank (Mantel and Haenszel 1959) and Wilcoxon (Breslow 1970, Gehan 1965) tests for the equality of hazard functions on subsamples constrained by minimum tenure. Specifically, we gradually remove the early churners in order to determine after what age the two functions are not distinguishable from one another. The results of this process are shown in Table 3.5. The log-rank test fails to distinguish the hazard functions (i.e., p-value > 0.10) when customers have already lived 4 days (i.e., churn on day 5 or later, if they do). Meanwhile, the Wilcoxon test that places a stronger weight on the early periods, fails to distinguish the hazard functions if all customers have lived at least 3 days (i.e., churn on day 4 or later).

Table 3.5: Non-parametric Tests of Survival Constraining Sample by Minimum Tenure

Minimum Tenure	Customers at risk	Observed Failures		Log-rank Test			Wilcoxon Test		
		Control	Treated	E(Control)	E(Treated)	p-value	E(Control)	E(Treated)	p-value
1 day	2,673	884	115	857.39	141.61	0.0155	854.39	141.61	0.0143
2 days	2,581	799	108	777.99	129.01	0.0455	777.99	129.01	0.0503
3 days	2,548	767	107	749.58	124.42	0.0913	749.58	124.42	0.1128
4 days	2,533	752	107	736.69	122.31	0.1344	736.69	122.31	0.1751
7 days	2,501	725	102	709.20	117.80	0.1156	709.20	117.80	0.1465

The E(·) columns indicate the expected number of failures per customer group if both had the same hazard function.

This phenomenon is best appreciated graphically. Figure 3.3 shows the estimated Kaplan and Meier (1958) survivor functions for the treated and control groups of customers. The left panel shows the estimates for the full sample and there is a clear vertical separation between the two survivor functions. However, it is also evident that the separation between the two functions grows in the early periods and remains

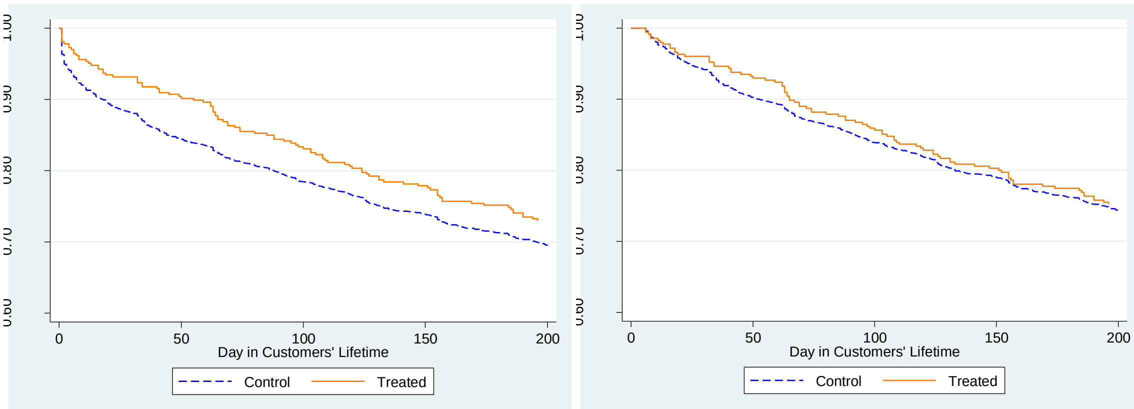

Figure 3.3: Kaplan-Meier Survival Estimates by Treatment for all customers (left) and for customers who lived at least 7 days (right)

Table 3.6: Decay of Treatment Effect on Survival Results

Column	(1)	(2)	(3)	(4)	(5)	(6)
EPE_Week1_{it}	-0.685** (0.287)	-0.685** (0.287)	-0.685** (0.287)	-0.685** (0.287)	-0.685** (0.287)	-0.685** (0.287)
EPE_Week2_{it}		-0.269 (0.434)	-0.269 (0.434)	-0.269 (0.434)	-0.269 (0.434)	-0.269 (0.434)
EPE_Week3_{it}			-0.273 (0.474)	-0.273 (0.474)	-0.273 (0.474)	-0.273 (0.474)
EPE_Week4_{it}				-1.358 (1.021)	-1.358 (1.021)	-1.358 (1.021)
EPE_Week5_{it}					-0.087 (0.479)	-0.087 (0.479)
EPE_Week6_{it}					-0.634 (0.602)	-0.634 (0.602)
EPE_Week7_{it}						-1.330 (1.023)
EPE_Week8_{it}						0.022 (0.627)
$EPE\_OtherWeeks_{it}$	-0.183* (0.109)	-0.177 (0.112)	-0.172 (0.116)	-0.146 (0.117)	-0.126 (0.123)	-0.102 (0.126)

All regressions employ Cox Proportional Hazard model, use the 2,673 customers in the sample, and include hourly, weekday, and weekly dummies. There are 999 observed failures.

Robust standard errors, clustered on customers in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

relatively constant thereafter. On the right panel we show the same estimates but after constraining our sample to customers who have lived at least one week (i.e., 7 days). Similar to the prior analysis, the two functions almost overlap.

Turning to a semi-parametric approach, we now employ Model (2b) and include time-varying parameters in our Cox proportional hazard model. The results with this model are shown in Table 3.6. Column (1) uses a single indicator for week one (i.e., $p = 1$ in the model's formulation). The coefficient for EPE_Week1_{it} represents a 49.60% (i.e., $e^{-0.685} - 1$) drop in the hazard rate. This implies that the EPE causes treated customers to fail about half as fast as the controls during the first week of their lifetimes.

Parameter $EPE\_OtherWeeks_{it}$ in column (1) is also negative and statistically significant, albeit it is only significant at the 10% level. The coefficient suggests that from week 2 onwards the treatment still reduces the hazard rate only by 16.71% (i.e., $e^{-0.183} - 1$), an effect much smaller than that found during the first week. Moreover, once we include an indicator for week 2 after the treatment, as in column (2) and the following ones, such effect vanishes. In other words, the treatment has no measurable effect during week 2 nor afterwards as the insignificance of EPE_Week2_{it}, the other weekly indicators, and $EPE\_OtherWeeks_{it}$ suggests.

In sum, both our non-parametric and our semi-parametric analyses indicate that the decay in EPE's effect is very fast and does not last more than a week.

3.6.3 Demand for Technical Support

We now test if the treated customers ask fewer questions during the initial stages of their service co-production processes. As with the retention-related parameters, we started by running simple t-tests of differences in the means of $QuestionsWeek1_i$ and $QuestionsWeek2_i$. Although the tests do not find a statistically significant difference (see right columns of Table 3.1), we must recall that these variables suffer from very strong overdispersion, in the sense that the (unconditional) variances are much greater than the means. The high variances (i.e., long tails) may explain the result, which is why we implement models that account for overdispersion.

The results of our Poisson and negative binomial model (3) for the number of questions asked by customers are presented in Table 3.7. In addition to reporting the coefficient for EPE_i, we also report the percentage change and the discrete change (computed using the finite difference method) in the number of questions asked caused by EPE. Column (1), with the Poisson specification, indicates the treatment reduces the number of questions asked by customers during their first week by 19.55%, or an average of 0.119 questions less. This is important for the provider since, as discussed before, this is when customers ask most of their questions.

Table 3.7: Results for Number of Questions Asked

Column	(1)	(2)	(3)	(4)
Dependent Variable	$QuestionsWeek1_i$		$QuestionsWeek2_i$	
Model	Poisson	Negative Binomial	Poisson	Negative Binomial
EPE_i	-0.218* (0.123)	-0.273** (0.122)	-0.164 (0.132)	-0.211* (0.127)
Percentage Change [a]	-19.55%	-23.89%	-15.12%	-19.04%
Discrete Change [b]	-0.119	-0.146	-0.117	-0.149
Squared correlation between actual and fitted number of questions	0.010	0.009	0.013	0.012

All regressions use the 2,673 customers in the sample and include hourly, weekday, and weekly dummies. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

[a] $e^{\hat{\beta}} - 1$.

[b] $E(QuestionsWeekn|EPE_i = 0) - E(QuestionsWeekn|EPE_i = 1)$ holding all other covariates' values at their means

Using these results and following recommendation by Cameron and Trivedi (2010), we test the null hypothesis of equidispersion, $H_0: \alpha = 0$, against the alternative of overdispersion, $H_a: \alpha > 0$, in the equation

$$Var(QuestionsWeek1_i|X_i) = E(QuestionsWeek1_i|X_i) + \alpha^2 E(QuestionsWeek1_i|X_i).$$

We reject the null and confirm a strong presence of overdispersion (i.e., $\hat{\alpha} = 3.98$, p-value less than 0.001). This supports our choice of using robust standard errors (i.e., the quasi-maximum likelihood approach) as well as the consideration of the negative binomial model.

Column (2), with the negative binomial specification, suggests a reduction of 23.89% in the number of questions, or 0.146 questions less, a slightly stronger yet qualitatively consistent estimate of the treatment's effect relative to that in column (1). Moreover, both specifications have a very similar squared correlation between the fitted number of questions and the actual number of questions (i.e., 0.010 and 0.009), suggesting both models provide a similar fit for the conditional mean (Cameron and Trivedi 2010). However, if we compare the fitted probability distributions of $QuestionsWeek1_i$ produced by both models, we find that the negative binomial model is much more accurate in predicting lower counts of numbers of questions (i.e., $QuestionsWeek1_i \leq 2$). In particular, the negative binomial is much more accurate in predicting when no questions or a single question occurs (i.e., $QuestionsWeek1_i = 0$ or $QuestionsWeek1_i = 1$). Overall, while the mean difference between the actual and fitted counts by the Poisson model is 0.405, the same difference is only 0.026 for the negative binomial model. All these results, which are described in detail in section J.1 of Appendix J, suggest that the negative binomial model is a better choice if we were interested in the entire distribution. However, given our focus on the conditional mean (i.e., the treatment's effect on the overall demand), we still prefer to base our conclusions on the more conservative estimate of the Poisson model. Both of these models are robust to the use of the alternate sets of time-of-adoption controls employed before in Table 3.4, indicating there is no bias induced by the time of adoption and that our identification strategy remains valid (see section J.2 of Appendix J).

In columns (3) and (4) of Table 3.7 we use the same two specifications but having the number of questions asked during the first two weeks as dependent variable, $QuestionsWeek2_i$. We don't find any measurable effect of EPE using the Poisson model in column (3). Additionally, using the negative binomial and relative to the prior result in column (2), in column (4) we find a weaker effect of the treatment (in terms of statistical significance and magnitude). We also checked if the results with these models are robust to alternate sets of time-of-adoption controls. The effect of the treatment estimated with the Poisson model in column (3) never became significant, as expected. However, the weak effect observed using the negative binomial model in column (4) was lost as soon as we performed any change in the control vector. In other words, the only specification in which the treatment exhibited a statistically significant effect (at least at the 10% level) is that shown in column (4) using the original set of controls. In sum, we do not find conclusive evidence that EPE is effective in reducing the number of questions asked by customers during their first two weeks. Rather, as suggested above, the difference between the two customer groups is only present during the first week. This result is consistent with our previous findings in that the effect of the treatment is only measureable during the first week.

3.7 Conclusion

Leveraging a field experiment executed by a public cloud infrastructure services provider, our study is the first to quantify the effects of customer education, and in particular early proactive education (EPE), on customer retention and demand for technology support early on in customers' service co-production processes. Our estimates of EPE's effect on customer behavior are economically significant. During the first week, which is when customers are most likely to abandon the service, customers who receive

EPE are retained about twice as much as customers who do not. Moreover, since customer retention is relatively strong and stable after the initial ramp up stage, by improving customer retention when customers are just starting to learn how to use the service EPE has a positive effect on the overall size of the customer base in the long-run. Additionally, on average, the treated customers ask 19.55% fewer questions during their first week since adoption relative to the controls. Since the offering of technology support is a very costly and human labor-intensive endeavor, the reduction in the number of support requests represents an important cost reduction. In sum, by offering EPE, the provider positively affects its bottom line by both increasing revenues (from having a larger customer base) and reducing technology support costs.

Considering how simple the EPE treatment is, it is worth discussing why it has a measureable effect on customer behavior. After all, a successful application of the EPE treatment is nothing more than a short phone conversation followed up by a support ticket, so its impact is somewhat surprising. After discussing the issue at large with executives, analysts, and also agents who applied the treatment at the provider’s premises, their explanation for the positive effects is that customers value that someone is taking the time to reach out to them to assist them, which is consistent with our theoretical premise that EPE increases customer satisfaction. EPE also reduces customers’ service co-production costs. Regardless of customers’ a priori technical capabilities, the adoption of any new service will necessarily imply having to learn some unique features of the new service. In the particular case of cloud infrastructure services, even seasoned system administrators will be unfamiliar with the provider’s web-based control panel until they see it for the first time. Therefore, simple guidance such as being walked through the control panel prevents customers from having to climb learning curves on their own, in turn also preventing potential frustrations that may arise from even simple issues such as not finding an option hidden in some menu. Customers value having someone telling them right away where in the control panel the options are as well as answering any other questions they may have about the service. The value of the assistance for customers is enhanced by their alternative, which would be having to invest time in finding manuals and knowledge forums on their own.

The findings regarding EPE’s positive impact on customer early behavior can be generalized beyond the cloud infrastructure services context and, in particular, to other service settings where customers can sign up for and abandon the service for free. An example discussed before, that suffers from early retention as the cloud does, is that of online learning programs (Muilenburg and Berge 2005, Tyler-Smith 2006). The costless exit from the service makes any effort needed to get started with the service significant.

Despite this chapter’s contributions, it is still subject to some limitations which, however, may be overcome by future research through additional data analysis and collection. For example, as we acknowledged when presenting our econometric approach, our model for the number of questions asked by customers potentially suffers from attrition bias. Additional data in regards to customer attributes that may lead to attrition (e.g., level of IT skills, pre-existing IT infrastructure, goals when opening their account, length of project) may serve to construct controls that mitigate the potential bias. Some of these attributes are already captured through the signup survey, but the survey is limited as it does not include sufficiently detailed items and has a low response rate.

Additionally, in this chapter we have considered all customer support requests as equal with respect to the provider’s cost in addressing them. Supplemental data concerning the time it takes support agents to address each request and the agent’s level of technical expertise may serve to produce a better estimate of the provider’s costs, in turn allowing better identification of the effects of EPE on them. Moreover, similar information with respect to the costs of offering EPE (e.g., duration of each EPE engagement) would allow balancing the cost of offering EPE with its (negative) effect on reactive support costs. The net payoff of EPE per customer on the provider’s bottom line is an important and interesting matter for future research.

Finally, while in this study we assume that one of the reasons why treated customers ask fewer questions is because the treatment improves their service co-production skills for the period immediately after adoption, further examination of the support interactions may serve to assess the validity of this claim. For example, by text mining the transcripts of the online live chat sessions and tickets, we could test if the type of questions asked, and in particular their level of technical complexity, changes between the treated and the controls. It may be the case that even though the number of questions asked during the first couple of weeks does not change radically, as we have found thus far, the questions asked by the treated customers are more sophisticated ones. In other words, the questions may be associated with customers' better service co-production knowledge and skills and refer to more advanced configuration issues than the basic topics already addressed through EPE. Since more efficient customers are less costly to serve and more likely to continue using SSTs (Xue et al. 2007), this would further demonstrate EPE's benefits.

Future research may address questions associated with the feasibility of an EPE-based business strategy, particularly considering how fast costs can grow. When proactively approaching customers, the provider is engaging a much broader customer base than the one it would if it only offered reactive support (Challagalla et al. 2009). In other words, a customer who was not going to approach the provider through reactive support becomes more costly to serve because of the proactive investment. In settings with viral adoption patterns this may not be sustainable. A potential way of addressing this issue is by examining varying levels of EPE and determining what is the minimum (i.e., less costly) treatment that can be offered that still produces the desired outcomes. Future field experiments, similar to the one used in this study, can serve this purpose.

APPENDIX A

FOR CHAPTER 2: PROOFS OF MOTIVATING HYPOTHESES

Solving the utility maximization via FOC, if buyers use the input service at a given moment in time, their instantaneous optimal consumption volume is given by

$$q^*(\theta|s,z) = \theta\left(\frac{z}{p_s}\right)^{\frac{1}{1-z}}.$$

Plugging optimal consumption volume into the utility function under basic and full support scenarios, we obtain the following values for instantaneous utilities at adoption time:

$$u_b^*(\theta) = u^*(\theta|q = q^*(\theta|s = b, z = z_b), s = b, z = z_b) = (1 - z_b)\left(\frac{z_b}{p_b}\right)^{\frac{z_b}{1-z_b}}\theta \text{ and}$$

$$u_f^*(\theta) = u^*\left(\theta\middle|q = q^*\left(\theta\middle|s = f, z = z_f\right), s = f, z = z_f\right) = \left(1 - z_f\right)\left(\frac{z_f}{p_f}\right)^{\frac{z_f}{1-z_f}}\theta - F.$$

The individual rationality constraint is automatically satisfied for u_b^* under its corresponding optimal consumption volume. In order for some users to prefer full over basic support, we need the incentive compatibility constraint to hold (i.e., $u_f^* > u_b^*$) which is equivalent to $F < \theta\left[\left(1 - z_f\right)\left(\frac{z_f}{p_f}\right)^{\frac{z_f}{1-z_f}} - (1 - z_b)\left(\frac{z_b}{p_b}\right)^{\frac{z_b}{1-z_b}}\right]$.

Denote $\hat{\theta} \equiv \frac{F}{(1-z_f)\left(\frac{z_f}{p_f}\right)^{\frac{z_f}{1-z_f}} - (1-z_b)\left(\frac{z_b}{p_b}\right)^{\frac{z_b}{1-z_b}}}$. As mentioned in the main text, we focus on scenarios where $0 < \hat{\theta} < 1$, or, equivalently $\left(1 - z_f\right)\left(\frac{z_f}{p_f}\right)^{\frac{z_f}{1-z_f}} - (1 - z_b)\left(\frac{z_b}{p_b}\right)^{\frac{z_b}{1-z_b}} > F > 0$, further assuming $\max\{1, \frac{z_b}{p_b}\} < \frac{z_{f2b}}{p_b} < \frac{z_f}{p_f}$.

PROOF OF HYPOTHESIS 1: We need to show $q^*\left(\theta\middle|s = f, z = z_f\right) > q^*(\theta|s = b, z = z_b) \Leftrightarrow \left(\frac{z_b}{p_b}\right)^{\frac{1}{1-z_b}} < \left(\frac{z_f}{p_f}\right)^{\frac{1}{1-z_f}}$. If $z_b < z_f$ then $1 - z_b > 1 - z_f$ and $\frac{1}{1-z_b} < \frac{1}{1-z_f}$. Given the assumed property $\max\left\{1, \frac{z_b}{p_b}\right\} < \frac{z_{f2b}}{p_b} < \frac{z_f}{p_f}$, we have two cases:

Case (1): $1 < \frac{z_b}{p_b} < \frac{z_f}{pf}$. Then we raise two numbers greater than 1 to powers that are also ordered—the ordering remains.

Case (2): $\frac{z_b}{p_b} < 1 < \frac{z_f}{pf}$. Then $\left(\frac{z_b}{p_b}\right)^{\frac{1}{1-z_b}} < 1 < \left(\frac{z_f}{p_f}\right)^{\frac{1}{1-z_f}}$.

PROOF OF HYPOTHESIS 2: We need to show $q^*\left(\theta|s=b, z=z_{f2b}\right) > q^*(\theta|s=b, z=z_b) \Leftrightarrow \left(\frac{z_b}{p_b}\right)^{\frac{1}{1-z_b}} < \left(\frac{z_{f2b}}{p_b}\right)^{\frac{1}{1-z_{f2b}}}$. The proof is similar to the proof of Hypothesis 1 (replacing z_f with z_{f2b}) and we omit it for brevity.

PROOF OF HYPOTHESIS 3: Note that $q^*\left(\theta|s=f, z=z_f\right) - q^*(\theta|s=b, z=z_b) = \theta\left[\left(\frac{z_f}{p_f}\right)^{\frac{1}{1-z_f}} - \left(\frac{z_b}{p_b}\right)^{\frac{1}{1-z_b}}\right]$. It follows that $\frac{\partial\left(q^*(\theta|s=f,z=z_f) - q^*(\theta|s=b,z=z_b)\right)}{\partial\theta} = \left(\frac{z_f}{p_f}\right)^{\frac{1}{1-z_f}} - \left(\frac{z_b}{p_b}\right)^{\frac{1}{1-z_b}} > 0$ per Hypothesis 1.

PROOF OF HYPOTHESIS 4: Note that $q^*\left(\theta|s=b, z=z_{f2b}\right) - q^*(\theta|s=b, z=z_b) = \theta\left[\left(\frac{z_{f2b}}{p_b}\right)^{\frac{1}{1-z_{f2b}}} - \left(\frac{z_b}{p_b}\right)^{\frac{1}{1-z_b}}\right]$. It follows that $\frac{\partial\left(q^*(\theta|s=b,z=z_{f2b}) - q^*(\theta|s=b,z=z_b)\right)}{\partial\theta} = \left(\frac{z_{f2b}}{p_b}\right)^{\frac{1}{1-z_{f2b}}} - \left(\frac{z_b}{p_b}\right)^{\frac{1}{1-z_b}} > 0$ per Hypothesis 2.

APPENDIX B

FOR CHAPTER 2: DESCRIPTION OF SUBSAMPLES AND CEM PROCEDURES

We work with several subsamples dependent on data availability and the application of the Coarsened Exact Matching (CEM) procedure. The "Full" sample constitutes our entire dataset, without any buyers excluded. For the "Baseline" subsample we have excluded buyers who (1) only accessed basic support and (2) averaged 512 MB RAM/hour or less during their first 6 months (excluding 1st month) or (3) made no adjustments to size of their infrastructure during their first 6 months (excluding 1st month).. An infrastructure resizing occurs on any launching, halting, or resizing of a server in the customers' cloud infrastructure. We do not consider their behavior during their 1st month in our threshold because most customers are setting up their infrastructure during this time. All other subsamples are subsets of the "Baseline" subsample. We only have visibility into the buyers' support interactions with the provider starting on October 2009, which constrains our "Support" subsample. The support data is needed to construct the instruments (see section 2.6.2 for details). The "CEMn" subsamples correspond to the usage of different matching criteria in the CEM procedures. Each of the 3 different CEM-based subsamples employs different matching criteria, described in see section 2.5.2.

Using the criteria above, several different matched subsamples were developed using different permutations of the criteria. It was deemed that all matches using 3 or less criteria, in addition to the time of upgrade, were too coarse (i.e., too many buyers per stratum). We then employ matched subsamples based on 4 or more criteria, as described in 2.5.2. We consider customers who exclusively use basic support as controls, and those who start with basic support and then upgrade to full support as treated.

Table B.1: General Description of Subsamples

Sample		Full	Baseline	Support	CEM1	CEM2	CEM3
Buyers in Sample		79,619	22,179	20,040	2,685	2,029	687
Panel Start		March'09	March'09	Oct'09	Oct'09	Oct'09	Oct'09
Panel End		Aug'12	Aug'12	Aug'12	Aug'12	Aug'12	Aug'12
Panel Length (months)		42	42	35	35	35	35
Buyer-month observations		1,073,998	368,606	298,539	48,725	37,837	13,262
Buyers' support choice behavior	Only use basic support	73,594	16,157	14,338	2,365	1,732	526
	Start with basic, upgrade to full support	1,409	1,408	1,132	275	258	136
	Start with basic, upgrade to full, and downgrade to basic	205	203	159	45	39	25
	Start with full, downgrade to basic	215	215	215	Excluded	Excluded	Excluded
	Only full support	4,196	4,196	4,196	Excluded	Excluded	Excluded
Data components available	Cloud infrastructure usage and support choice data	Yes	Yes	Yes	Yes	Yes	Yes
	Survey data used for CEM	Incomplete	Incomplete	Incomplete	Yes	Yes	Yes
	Support interaction data used to construct IVs	Incomplete	Incomplete	Yes	Yes	Yes	Yes
CEM procedure applied?		No	No	No	Yes	Yes	Yes

Table B.2. Description of Matching Criteria used in CEM Procedures

Abbreviation	Description	# of Categories	Categories
Emp	Employment	5	0-10, 11-50, 51-100, 101-250, >250
UC	General use cases (can have more than 1)	5	High variance, low variance, back office, hosting, test & dev
Mem	Memory usage in months before upgrade	9	<0.5, 0.5-1, 1-2, 2-4, 4-8, 8-16, 16-32, 32-64, >64
Adj	Frequency of infrastructure resizing in months before upgrade	5	0, 1-2, 3-9, 10-43, >43
Ind	Industries	258	Popular ones have 11% to 15% of observations
t-upg	Upgrade month for treated, and month in lifetime for controls	40	One per month; longest delay in upgrading is 40 months.

Table B.3: Coarsened Exact Matching (CEM) Procedure Outcomes

Subsample	Strata	Strata	Matched Customers / Total Customers			Average Customers per Stratum			Parameters used for matching					
	w/cust.	w/matches	Controls	Treated	Both	Controls	Treated	Both	Emp	UC	Mem	Sca	Adj	t-upg
CEM1	11,876	294	2365/3800	320/ 400	2685/4200	8.0	1.1	9.1	Yes	Yes	Yes	Yes	No	Yes
CEM2	22,268	284	1732/3800	297/ 400	2029/4200	6.1	1.0	7.1	Yes	Yes	No	Yes	Yes	Yes
CEM3	36,908	157	526/3800	161/ 400	687/4200	3.4	1.0	4.4	Yes	Yes	Yes	Yes	Yes	Yes

APPENDIX C

FOR CHAPTER 2: DESCRIPTIVE STATISTICS

C.1 Time-Varying Parameters

Table C.1: Descriptive Statistics of Time-Varying Variables

Sample	Full				Baseline				Support				CEM1				CEM2				CEM3			
Buyers	79,619				22,179				20,040				2,685				2,029				687			
Observations	1,073,998				368,606				298,539				48,725				37,837				13,262			
Variable	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max
$Memory_{it}$	3.4	19.2	0	2,284.5	7.9	31.4	0.0	2,284.5	7.3	27.4	0.0	2,284.5	5.2	17.2	0.0	675.9	6.7	22.1	0.0	675.9	5.1	12.1	0.0	329.0
$lnMemory_{it}$	0.746	0.871	0	7.734	1.348	1.040	0	7.734	1.343	1.014	0	7.734	1.218	0.894	0	6.518	1.302	0.986	0	6.518	1.221	0.920	0	5.799
$FullSupport_{it}$	0.055	0.228	0	1	0.160	0.367	0	1	0.183	0.387	0	1	0.078	0.268	0	1	0.093	0.291	0	1	0.155	0.362	0	1
$SwitchToBasic_{it}$	0.003	0.052	0	1	0.008	0.089	0	1	0.009	0.093	0	1	0.007	0.082	0	1	0.007	0.084	0	1	0.013	0.113	0	1
$FractionParallel_{it}$	0.058	0.198	0	1	0.121	0.266	0	1	0.120	0.267	0	1	0.106	0.251	0	1	0.116	0.258	0	1	0.104	0.246	0	1

C.2 Time-Invariant Parameters used for CEM Procedures

Table C.2: Descriptive Statistics of Parameters used in Survey Data for CEM before matching (5,134 buyers)

Buyer Role Number of Buyers		All Buyers 5,134				Controls 3,875				Treated 1,259				t-test of mean difference	
Variable	Description	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max	Diff.	p-value
$Employees_i$	Number of employees	195.7	1,102.4	2	10,000	164.7	1019.9	2	10,000	291.0	1320.3	2	10,000	126.3	0.000
$lnEmployment_i$	$ln(Employees_i)$	2.402	1.706	1.099	9.21	2.26	1.608	1.099	9.21	2.838	1.914	1.099	9.21	0.578	0.000
$EmpCat1_i$	$1[Employees_i \leq 10]$	0.656	0.475	0	1	0.692	0.462	0	1	0.546	0.498	0	1	-0.146	0.000
$EmpCat2_i$	$1[11 \leq Employees_i \leq 50]$	0.198	0.398	0	1	0.187	0.390	0	1	0.230	0.421	0	1	0.044	0.001
$EmpCat3_i$	$1[51 \leq Employees_i \leq 100]$	0.050	0.218	0	1	0.044	0.204	0	1	0.071	0.256	0	1	0.027	0.000
$EmpCat4_i$	$1[101 \leq Employees_i \leq 250]$	0.037	0.188	0	1	0.030	0.171	0	1	0.056	0.231	0	1	0.026	0.000
$EmpCat5_i$	$1[250 < Employees_i]$	0.060	0.237	0	1	0.047	0.213	0	1	0.097	0.296	0	1	0.049	0.000
UC_HU_i	1[High Uncertainty UC]	0.463	0.499	0	1	0.469	0.499	0	1	0.447	0.497	0	1	-0.021	0.185
UC_LU_i	1[Low Uncertainty UC]	0.591	0.492	0	1	0.573	0.495	0	1	0.647	0.478	0	1	0.073	0.000
UC_BO_i	1[Back Office UC]	0.189	0.391	0	1	0.195	0.396	0	1	0.169	0.375	0	1	-0.026	0.043
UC_HO_i	1[Hosting UC]	0.092	0.289	0	1	0.093	0.290	0	1	0.088	0.284	0	1	-0.005	0.613
UC_TD_i	1[Test & Development UC]	0.293	0.455	0	1	0.323	0.468	0	1	0.203	0.402	0	1	-0.120	0.000

Table C.3: Descriptive Statistics of Parameters in CEM1 Matched Sample <u>without</u> Weights (2,685 buyers)

Buyer Role		All Buyers				Controls				Treated				t-test of mean difference	
Number of Buyers		2,685				2,365				320					
Variable	Description	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max	Diff.	p-value
$Employees_i$	Number of employees	110.2	830.8	2	10,000.00	94.5	779.7	2	10,000.00	226.4	1134.3	2	10,000.00	131.8	0.008
$lnEmployment_i$	$ln(Employees_i)$	1.974	1.372	1.099	9.21	1.901	1.292	1.099	9.21	2.513	1.772	1.099	9.21	0.612	0.000
$EmpCat1_i$	$1[Employees_i \leq 10]$	0.780	0.415	0	1	0.798	0.401	0	1	0.641	0.481	0	1	-0.158	0.000
$EmpCat2_i$	$1[11 \leq Employees_i \leq 50]$	0.151	0.358	0	1	0.145	0.352	0	1	0.200	0.401	0	1	0.055	0.009
$EmpCat3_i$	$1[51 \leq Employees_i \leq 100]$	0.026	0.159	0	1	0.022	0.145	0	1	0.059	0.237	0	1	0.038	0.000
$EmpCat4_i$	$1[101 \leq Employees_i \leq 250]$	0.011	0.105	0	1	0.010	0.098	0	1	0.022	0.147	0	1	0.012	0.052
$EmpCat5_i$	$1[250 < Employees_i]$	0.032	0.176	0	1	0.026	0.159	0	1	0.078	0.269	0	1	0.052	0.000
UC_HU_i	1[High Uncertainty UC]	0.495	0.500	0	1	0.495	0.500	0	1	0.491	0.501	0	1	-0.005	0.880
UC_LU_i	1[Low Uncertainty UC]	0.598	0.490	0	1	0.595	0.491	0	1	0.622	0.486	0	1	0.027	0.349
UC_BO_i	1[Back Office UC]	0.093	0.291	0	1	0.089	0.285	0	1	0.125	0.331	0	1	0.036	0.036
UC_HO_i	1[Hosting UC]	0.034	0.181	0	1	0.029	0.167	0	1	0.072	0.259	0	1	0.043	0.000
UC_TD_i	1[Test & Development UC]	0.236	0.425	0	1	0.234	0.423	0	1	0.253	0.435	0	1	0.019	0.446

Table C.4: Descriptive Statistics of Parameters in CEM1 Matched Sample with Weights (2,685 buyers)

Buyer Role Number of Buyers		All Buyers 2,685				Controls 2,365				Treated 320				t-test of mean difference	
Variable	Description	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max	Diff.	p-value
$Employees_i$	Number of employees	226.8	1135.0	2	10,000.00	278.1	1368.9	2	10,000.00	226.4	1134.3	2	10,000.00	-51.7	0.827
$lnEmployment_i$	$ln(Employees_i)$	2.512	1.771	1.099	9.21	2.478	1.837	1.099	9.21	2.513	1.772	1.099	9.21	0.034	0.926
$EmpCat1_i$	$1[Employees_i \leq 10]$	0.641	0.480	0	1	0.641	0.480	0	1	0.641	0.481	0	1	0.000	1.000
$EmpCat2_i$	$1[11 \leq Employees_i \leq 50]$	0.200	0.400	0	1	0.200	0.400	0	1	0.200	0.401	0	1	0.000	1.000
$EmpCat3_i$	$1[51 \leq Employees_i \leq 100]$	0.059	0.236	0	1	0.059	0.236	0	1	0.059	0.237	0	1	0.000	1.000
$EmpCat4_i$	$1[101 \leq Employees_i \leq 250]$	0.022	0.146	0	1	0.022	0.146	0	1	0.022	0.147	0	1	0.000	1.000
$EmpCat5_i$	$1[250 < Employees_i]$	0.078	0.268	0	1	0.078	0.268	0	1	0.078	0.269	0	1	0.000	1.000
UC_HU_i	1[High Uncertainty UC]	0.491	0.500	0	1	0.491	0.500	0	1	0.491	0.501	0	1	0.000	1.000
UC_LU_i	1[Low Uncertainty UC]	0.622	0.485	0	1	0.622	0.485	0	1	0.622	0.486	0	1	0.000	1.000
UC_BO_i	1[Back Office UC]	0.125	0.331	0	1	0.125	0.331	0	1	0.125	0.331	0	1	0.000	1.000
UC_HO_i	1[Hosting UC]	0.072	0.258	0	1	0.072	0.258	0	1	0.072	0.259	0	1	0.000	1.000
UC_TD_i	1[Test & Development UC]	0.253	0.435	0	1	0.253	0.435	0	1	0.253	0.435	0	1	0.000	1.000

APPENDIX D

FOR CHAPTER 2: SUPPORT INTERACTIONS AND CONSTRCUTION OF INSTRUMENTS

D.1 Support Interactions Coding Process

The content of the support interactions between the provider and its customers was used to identify three types of exogenous failures experienced by buyers. The following are the keywords and phrases used to identify each of these types of interactions. All support interactions that matched some keyword or phrased were visually examined to rule out false positives.

Table D.1: Keywords and Phrases Searched for Support Interactions Coding

Support Interaction Type (s)	Description of Event	List of keywords or phrases
$BusinessGrowth$	In order to deploy a new website, buyers will generally need to install an SSL certificate (e.g., for credit card transactions) or setup an SPF record (e.g., to send emails); the latter also requires adding an additional IP address to a server. Also, if number of visitor increases, they may need to request permission to surpass pre-established API call limits. Any of these types of actions is suggestive of a buyer's business growth.	Additional IP, new IP, another IP, request IP, IP request, extra IP, add IP, second IP API limit SPF record, SSL certificate, CSR request, Generate CSR
$FailOutage$	Provider may suffer from generalized outages in different components of its service (e.g., memory leak in provider's cloud management system). Such generalized problems are announced in the provider's status webpage and/or announced to buyers.	Providers' service status URL, cloud status, outage, scheduled maintenance, undergoing maintenance
$FailNetwork$	Some node in the provider's infrastructure, generally belonging to some customer, is suffering from a distributed denial of service attack (DDoS) or some networking hardware device has failed.	Server does not respond to ARP requests, faulty switch, network issue in our data center, lb in error state, load-balancer hardware nodes, DDoS

Table D.1 (continued)

$FailHost$	Buyer is suffering degraded performance due to a problem in the physical host in which the buyer's virtual machine runs. Problems are generally associated with excessive read/write (or input/output) operations on the hard disks, either by the buyer (e.g., by some unexpected bug in their applications such as a memory overflow that causes swapping) or by another customer whose virtual machine lives in same physical server (e.g., a "noisy neighbor"). Problems could also be associated with failure of the physical hardware (e.g., a hard disk failure).	Consuming a significant amount of Disk I/O, very high disk I/O usage, iowait, iostat, swapping, swappers, swap space, extreme slowness, slowdown problems, hardware failure, degraded hardware, drive failing, drives failing, server outage, host failure, server is down, server down, site down, host became unresponsive, server unresponsive, server not responding, server is unresponsive, is hosted on has become unresponsive, problem with our server, host server, physical host, physical hardware, physical machine, host machine, failing hardware, hardware failure, imminent hardware issues, migrate your cloud server to another host, queued for move, issue on the migrations, host server of your cloud servers

D.2 Construction of Support-Based Parameters

Let $S = BusinessGrowth, FailOutage, FailNetwork, FailHost$ represent a type of support interaction identified through coding process. Let $NumS_{it}$ be the number of support interactions of type s counted for buyer i during month t. Further, let $AccS_{it}$ be the accumulated number of support interactions of type S that buyer i has experienced up to month t. Formally, $AccS_{ik} = \sum_{t=1}^{t=k} NumS_{it}$. Finally, construct indicators that are turned on when the total number of interaction is greater than $N = 1{,}2$, as $SN_{it} = 1[AccS_{it} \geq N]$. Then, for example, parameter $BusinessGrowth2_{it}$ will be equal to 1 if buyer i has accumulated at least 2 support interactions that have been coded as type $BusinessGrowth$ by month t.

D.3 Descriptive Statistics of Support Interactions

Table D.2: Descriptive Statistics of Support Interactions-based Parameters (Support and CEM1 subsamples)

Subsample	Support									CEM1								
Buyers in Sample	20,040									2,685								
Observations	298,539									48,725								
Variable	Mean	S.D.	Min	Med	75th p.	90th p.	95th p.	99th p.	Max	Mean	S.D.	Min	Med	75th p.	90th p.	95th p.	99th p.	Max
$NumBusinessGrowth_{it}$	0.018	0.166	0	0	0	0	0	1	12	0.017	0.151	0	0	0	0	0	1	4
$AccBusinessGrowth_{it}$	0.240	0.868	0	0	0	1	1	4	23	0.226	0.771	0	0	0	1	1	4	13
$BusinessGrowth1_{it}$	0.134	0.340	0	0	0	1	1	1	1	0.126	0.332	0	0	0	1	1	1	1
$BusinessGrowth2_{it}$	0.049	0.215	0	0	0	0	0	1	1	0.048	0.215	0	0	0	0	0	1	1
$NumFailOutage_{it}$	0.009	0.123	0	0	0	0	0	0	9	0.006	0.090	0	0	0	0	0	0	4
$AccFailOutage_{it}$	0.067	0.409	0	0	0	0	0	2	18	0.045	0.300	0	0	0	0	0	1	8
$FailOutage1_{it}$	0.044	0.205	0	0	0	0	0	1	1	0.031	0.174	0	0	0	0	0	1	1
$FailOutage2_{it}$	0.012	0.109	0	0	0	0	0	1	1	0.008	0.089	0	0	0	0	0	0	1
$NumFailNetwork_{it}$	0.002	0.059	0	0	0	0	0	0	10	0.001	0.039	0	0	0	0	0	0	3
$AccFailNetwork_{it}$	0.017	0.172	0	0	0	0	0	1	12	0.01	0.108	0	0	0	0	0	0	3
$FailNetwork1_{it}$	0.013	0.114	0	0	0	0	0	1	1	0.009	0.095	0	0	0	0	0	0	1
$FailNetwork2_{it}$	0.002	0.046	0	0	0	0	0	0	1	0.001	0.025	0	0	0	0	0	0	1
$NumFailHost_{it}$	0.024	0.197	0	0	0	0	0	1	21	0.020	0.170	0	0	0	0	0	1	11
$AccNumFailHost_{it}$	0.239	0.96	0	0	0	1	1	3	114	0.188	0.775	0	0	0	1	1	3	30
$NumFailHost1_{it}$	0.146	0.354	0	0	0	1	1	1	1	0.118	0.323	0	0	0	1	1	1	1
$NumFailHost2_{it}$	0.043	0.204	0	0	0	0	0	1	1	0.031	0.174	0	0	0	0	0	1	1

Table D.3: Descriptive Statistics of Support Interactions-based Parameters (CEM2 and CEM3 subsamples)

Subsample	CEM2									CEM3								
Buyers in Sample	2,029									687								
Observations	37,837									13,262								
Variable	Mean	S.D.	Min	Med	75th p.	90th p.	95th p.	99th p.	Max	Mean	S.D.	Min	Med	75th p.	90th p.	95th p.	99th p.	Max
$NumBusinessGrowth_{it}$	0.018	0.159	0	0	0	0	0	1	4	0.019	0.165	0	0	0	0	0	1	4
$AccBusinessGrowth_{it}$	0.233	0.785	0	0	0	1	2	4	9	0.257	0.869	0	0	0	1	2	5	9
$BusinessGrowth1_{it}$	0.126	0.332	0	0	0	1	1	1	1	0.129	0.335	0	0	0	1	1	1	1
$BusinessGrowth2_{it}$	0.050	0.218	0	0	0	0	1	1	1	0.055	0.227	0	0	0	0	1	1	1
$NumFailOutage_{it}$	0.007	0.103	0	0	0	0	0	0	4	0.011	0.127	0	0	0	0	0	0	4
$AccFailOutage_{it}$	0.057	0.348	0	0	0	0	0	2	8	0.072	0.414	0	0	0	0	0	2	8
$FailOutage1_{it}$	0.038	0.191	0	0	0	0	0	1	1	0.045	0.207	0	0	0	0	0	1	1
$FailOutage2_{it}$	0.011	0.104	0	0	0	0	0	1	1	0.014	0.119	0	0	0	0	0	1	1
$NumFailNetwork_{it}$	0.001	0.040	0	0	0	0	0	0	2	0.001	0.042	0	0	0	0	0	0	2
$AccFailNetwork_{it}$	0.011	0.111	0	0	0	0	0	1	3	0.013	0.117	0	0	0	0	0	1	2
$FailNetwork1_{it}$	0.010	0.101	0	0	0	0	0	1	1	0.012	0.109	0	0	0	0	0	1	1
$FailNetwork2_{it}$	0.001	0.026	0	0	0	0	0	0	1	0.001	0.025	0	0	0	0	0	0	1
$NumFailHost_{it}$	0.022	0.182	0	0	0	0	0	1	11	0.026	0.216	0	0	0	0	0	1	11
$AccNumFailHost_{it}$	0.209	0.854	0	0	0	1	1	3	30	0.242	1.089	0	0	0	1	1	4	30
$NumFailHost1_{it}$	0.126	0.332	0	0	0	1	1	1	1	0.136	0.343	0	0	0	1	1	1	1
$NumFailHost2_{it}$	0.035	0.185	0	0	0	0	0	1	1	0.041	0.198	0	0	0	0	0	1	1

APPENDIX E

FOR CHAPTER 2: ADDITIONAL RESULTS AND ROBUSTNESS CHECKS

E.1 Main Effect Results

Table E.1: Results with Basic Model for $lnMemory_{it}$ using Different Subsamples

Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Sample	Full		Baseline		Support		CEM1		CEM2		CEM3	
$FullSupport_{it}$	1.141***	1.194***	1.057***	1.103***	0.965***	0.998***	1.067***	1.090***	1.075***	1.085***	1.055***	1.083***
	(0.029)	(0.031)	(0.029)	(0.031)	(0.029)	(0.032)	(0.057)	(0.061)	(0.058)	(0.062)	(0.084)	(0.093)
$SwitchToBasic_{it}$	-0.443***	-0.444***	-0.488***	-0.488***	-0.430***	-0.430***	-0.578***	-0.579***	-0.752***	-0.753***	-0.679***	-0.681***
	(0.060)	(0.060)	(0.060)	(0.060)	(0.060)	(0.060)	(0.141)	(0.141)	(0.146)	(0.146)	(0.122)	(0.122)
$AdoptFullIn4_{it}$		0.245***		0.200***		0.117***		0.081**		0.038		0.089
		(0.023)		(0.023)		(0.023)		(0.041)		(0.042)		(0.061)
Constant	0.194***	0.193***	0.230***	0.227***	0.124	0.121	-0.402	-0.397	-0.302	-0.300	-0.688*	-0.683*
	(0.010)	(0.010)	(0.024)	(0.024)	(0.092)	(0.092)	(0.427)	(0.425)	(0.473)	(0.472)	(0.392)	(0.399)
Observations	1,073,998	1,073,998	368,606	368,606	298,539	298,539	48,725	48,725	37,837	37,837	13,262	13,262
Buyers	79,619	79,619	22,179	22,179	20,040	20,040	2,685	2,685	2,029	2,029	687	687
R^2	0.183	0.183	0.251	0.252	0.237	0.237	0.321	0.321	0.336	0.336	0.397	0.398
Upgrade change ($e^{\widehat{\beta}} - 1$)	213.1%	230.1%	187.7%	201.2%	162.5%	171.3%	190.8%	197.6%	192.9%	196.0%	187.2%	195.3%
Downgrade change ($e^{\widehat{\beta}+\widehat{\gamma}} - 1$)	101.0%	111.7%	76.7%	84.8%	70.9%	76.6%	63.1%	66.8%	38.0%	39.4%	45.6%	49.4%
$\hat{\beta} + \hat{\gamma} = 0$ test p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.022	0.019	0.002	0.001

Dependent variable is $lnMemory_{it}$. All regressions include calendar (τ_t) and lifetime time dummies (l_{it}).Robust standard errors, clustered on customers, in parentheses.
$^{*} p < 0.10, ^{**} p < 0.05, ^{***} p < 0.01.$

Table E.2: Results with Basic Model for $FractionParallel_{it}$ using Different Subsamples

Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Sample	Full		Baseline		Support		CEM1		CEM2		CEM3	
$FullSupport_{it}$	0.096***	0.098***	0.096***	0.100***	0.089***	0.090***	0.107***	0.108***	0.106***	0.105***	0.108***	0.110***
	(0.006)	(0.007)	(0.006)	(0.007)	(0.007)	(0.008)	(0.014)	(0.014)	(0.014)	(0.014)	(0.019)	(0.021)
$SwitchToBasic_{it}$	-0.031***	-0.031***	-0.032***	-0.032***	-0.027**	-0.027**	-0.030	-0.030	-0.025	-0.025	-0.018	-0.018
	(0.012)	(0.012)	(0.012)	(0.012)	(0.013)	(0.013)	(0.028)	(0.028)	(0.031)	(0.031)	(0.028)	(0.028)
$AdoptFullIn4_{it}$		0.013***		0.014***		0.003		0.002		-0.004		0.008
		(0.005)		(0.005)		(0.005)		(0.010)		(0.010)		(0.016)
Constant	0.009***	0.009***	0.023***	0.023***	-0.002	-0.002	-0.210	-0.209	-0.114	-0.114	-0.422***	-0.422***
	(0.002)	(0.002)	(0.006)	(0.006)	(0.022)	(0.022)	(0.131)	(0.131)	(0.120)	(0.120)	(0.120)	(0.120)
Observations	1,073,998	1,073,998	368,606	368,606	298,539	298,539	48,725	48,725	37,837	37,837	13,262	13,262
Buyers	79,619	79,619	22,179	22,179	20,040	20,040	2,685	2,685	2,029	2,029	687	687
R^2	0.017	0.017	0.026	0.026	0.024	0.024	0.037	0.037	0.042	0.042	0.060	0.060
Downward change $\hat{\beta} + \hat{\gamma}$	0.065	0.067	0.065	0.068	0.062	0.063	0.077	0.078	0.081	0.080	0.090	0.092
$\hat{\beta} + \hat{\gamma} = 0$ test p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.006	0.006	0.010	0.012	0.004	0.004

Dependent variable is $FractionParallel_{it}$. All regressions include calendar (τ_t) and lifetime time dummies (l_{it}).Robust standard errors, clustered on customers, in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table E.3: Results with Basic Model and Controlling for Business Growth using Different Subsamples

Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable	$lnMemory_{it}$				$FractionParallel_{it}$			
Subsample	Support	CEM1	CEM2	CEM3	Support	CEM1	CEM2	CEM3
$FullSupport_{it}$	0.884***	0.981***	0.985***	0.970***	0.080***	0.098***	0.097***	0.104***
	(0.028)	(0.056)	(0.058)	(0.084)	(0.007)	(0.014)	(0.014)	(0.020)
$SwitchToBasic_{it}$	-0.419***	-0.581***	-0.745***	-0.687***	-0.026**	-0.030	-0.025	-0.018
	(0.060)	(0.148)	(0.156)	(0.137)	(0.013)	(0.028)	(0.031)	(0.028)
$BusinessGrowth1_{it}$	0.454***	0.372***	0.392***	0.316***	0.046***	0.040***	0.054***	0.018
	(0.021)	(0.049)	(0.057)	(0.073)	(0.006)	(0.013)	(0.015)	(0.022)
$BusinessGrowth2_{it}$	0.223***	0.229***	0.268***	0.305***	0.034***	0.021	0.002	0.007
	(0.027)	(0.057)	(0.067)	(0.105)	(0.008)	(0.017)	(0.019)	(0.029)
Constant	0.124	-0.354	-0.233	-0.602**	-0.002	-0.204	-0.102	-0.416***
	(0.087)	(0.380)	(0.411)	(0.264)	(0.022)	(0.129)	(0.115)	(0.112)
Observations	298,539	48,725	37,837	13,262	298,539	48,725	37,837	13,262
Buyers	20,040	2,685	2,029	687	20,040	2,685	2,029	687
R^2	0.258	0.337	0.355	0.414	0.028	0.040	0.045	0.060
$e^{\hat{\beta}} - 1$ (cols. 1-4) or $\hat{\beta}$ (cols. 5-8)	142.1%	166.9%	167.9%	163.8%	0.080	0.098	0.097	0.104
$e^{\hat{\beta}+\hat{\gamma}} - 1$ (cols. 1-4) or $\hat{\beta} + \hat{\gamma}$ (cols. 5-8)	59.2%	49.2%	27.2%	32.7%	0.054	0.068	0.073	0.086
$\hat{\beta} + \hat{\gamma} = 0$ test p-value	0.000	0.005	0.112	0.040	0.000	0.014	0.018	0.006

All regressions include calendar (τ_t) and lifetime time dummies (l_{it}).Robust standard errors, clustered on customers, in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The "Full" and "Baseline" subsamples are not considered as the support interaction data is not available for all observations.

Table E.4: Results with Dynamic Panel Fixed Effects Models for $lnMemory_{it}$ using Different Subsamples

Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Subsample	Full		Baseline		Support		CEM1		CEM2		CEM3	
$FullSupport_{it}$	0.249***	0.240***	0.297***	0.282***	0.316***	0.309***	0.335***	0.337***	0.339***	0.344***	0.319***	0.318***
	(0.008)	(0.008)	(0.008)	(0.009)	(0.010)	(0.011)	(0.021)	(0.023)	(0.021)	(0.024)	(0.033)	(0.038)
$SwitchToBasic_{it}$	-0.156***	-0.172***	-0.156***	-0.173***	-0.147***	-0.168***	-0.270***	-0.283***	-0.317***	-0.330***	-0.236***	-0.240***
	(0.018)	(0.019)	(0.020)	(0.020)	(0.023)	(0.025)	(0.059)	(0.063)	(0.060)	(0.064)	(0.042)	(0.044)
$lnMemory_{it-1}$	0.987***	1.010***	0.954***	0.978***	0.931***	0.949***	0.965***	0.960***	0.971***	0.969***	0.999***	1.009***
	(0.004)	(0.006)	(0.006)	(0.008)	(0.006)	(0.009)	(0.018)	(0.024)	(0.021)	(0.029)	(0.021)	(0.028)
$lnMemory_{it-2}$	-0.156***	-0.202***	-0.150***	-0.196***	-0.159***	-0.196***	-0.190***	-0.190***	-0.191***	-0.200***	-0.222***	-0.266***
	(0.004)	(0.007)	(0.005)	(0.009)	(0.005)	(0.011)	(0.016)	(0.038)	(0.019)	(0.047)	(0.017)	(0.035)
$lnMemory_{it-3}$		0.047***		0.043***		0.037***		-0.004		-0.003		0.040*
		(0.005)		(0.007)		(0.008)		(0.034)		(0.041)		(0.023)
$lnMemory_{it-4}$		-0.011***		-0.005		-0.006		0.005		0.014		0.001
		(0.003)		(0.004)		(0.005)		(0.018)		(0.021)		(0.018)
Constant	0.181***	0.164***	0.280***	0.251***	0.339***	0.297***	0.290***	0.268***	0.385***	0.295***	0.474***	0.350***
	(0.010)	(0.010)	(0.015)	(0.015)	(0.024)	(0.022)	(0.095)	(0.061)	(0.071)	(0.058)	(0.076)	(0.090)
Observations	925,429	802,234	324,406	281,927	258,617	220,392	43,355	37,991	33,779	29,727	11,888	10,514
Buyers	63,476	56,429	21,573	20,020	19,440	17,919	2,684	2,657	2,029	2,011	687	686
R^2	0.790	0.794	0.773	0.779	0.734	0.733	0.779	0.764	0.793	0.779	0.811	0.801
Upgrade change ($e^{\hat{\beta}} - 1$)	28.3%	27.1%	34.6%	32.6%	37.2%	36.2%	39.8%	40.0%	40.4%	41.0%	37.6%	37.5%
Downgrade change ($e^{\hat{\beta}+\hat{\gamma}} - 1$)	9.8%	7.0%	15.1%	11.5%	18.4%	15.1%	6.7%	5.5%	2.2%	1.4%	8.7%	8.2%
$\hat{\beta} + \hat{\gamma} = 0$ test p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.249	0.379	0.711	0.819	0.034	0.036

Dependent variable is $lnCapacity_{it}$. All regressions include calendar (τ_t) and lifetime time dummies (l_{it}).Robust standard errors, clustered on customers, in parentheses.
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table E.5: Results with Dynamic Panel Fixed Effects Models for $FractionParallel_{it}$ using Different Subsamples

Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Sample	Full		Baseline		Support		CEM1		CEM2		CEM3	
$FullSupport_{it}$	0.026***	0.025***	0.032***	0.030***	0.033***	0.031***	0.040***	0.036***	0.040***	0.038***	0.043***	0.040***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.005)	(0.006)	(0.006)	(0.006)	(0.008)	(0.009)
$SwitchToBasic_{it}$	-0.011***	-0.014***	-0.010**	-0.013***	-0.007	-0.010*	-0.014	-0.015	-0.016	-0.016	-0.009	-0.011
	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)	(0.014)	(0.015)	(0.015)	(0.016)	(0.013)	(0.015)
$FractionParallel_{it-1}$	0.923***	0.949***	0.898***	0.928***	0.881***	0.908***	0.888***	0.901***	0.884***	0.888***	0.905***	0.894***
	(0.004)	(0.005)	(0.005)	(0.006)	(0.006)	(0.007)	(0.014)	(0.015)	(0.016)	(0.018)	(0.030)	(0.033)
$FractionParallel_{it-2}$	-0.159***	-0.202***	-0.165***	-0.202***	-0.172***	-0.204***	-0.181***	-0.198***	-0.182***	-0.183***	-0.188***	-0.194***
	(0.003)	(0.005)	(0.004)	(0.007)	(0.004)	(0.007)	(0.011)	(0.016)	(0.012)	(0.018)	(0.023)	(0.035)
$FractionParallel_{it-3}$		0.053***		0.051***		0.051***		0.064***		0.046***		0.070***
		(0.004)		(0.005)		(0.005)		(0.011)		(0.012)		(0.022)
$FractionParallel_{it-4}$		-0.024***		-0.027***		-0.032***		-0.055***		-0.047***		-0.051***
		(0.003)		(0.003)		(0.003)		(0.008)		(0.009)		(0.015)
Constant	0.021***	0.020***	0.033***	0.030***	0.058***	0.055***	0.024	0.016	0.055***	0.021	0.104***	0.083***
	(0.003)	(0.003)	(0.005)	(0.005)	(0.017)	(0.017)	(0.031)	(0.029)	(0.020)	(0.034)	(0.012)	(0.016)
Observations	925,429	802,234	324,406	281,927	258,617	220,392	43,355	37,991	33,779	29,727	11,888	10,514
Buyers	63,476	56,429	21,573	20,020	19,440	17,919	2,684	2,657	2,029	2,011	687	686
R^2	0.663	0.675	0.638	0.659	0.613	0.631	0.615	0.626	0.612	0.620	0.628	0.627
Downward change $\hat{\beta}+\hat{\gamma}$	0.015	0.011	0.023	0.018	0.026	0.021	0.026	0.022	0.024	0.022	0.034	0.029
$\hat{\beta}+\hat{\gamma}=0$ test p-value	0.000	0.009	0.000	0.000	0.000	0.000	0.060	0.127	0.112	0.163	0.009	0.051

Dependent variable is $FractionParallel_{it}$. All regressions include calendar (τ_t) and lifetime time dummies (l_{it}).Robust standard errors, clustered on customers, in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

E.2 Results for Role of Firm Size

Table E.6: Results Considering Firm Size for $lnMemory_{it}$ using Different Subsamples

Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Subsample	Survey Data Available (Baseline)				CEM2				CEM3			
$FullSupport_{it}$	1.025***	0.707***	0.937***	0.759***	1.075***	0.840***	1.005***	0.834***	1.055***	0.903***	1.012***	0.715***
	(0.050)	(0.067)	(0.057)	(0.083)	(0.058)	(0.082)	(0.064)	(0.098)	(0.084)	(0.115)	(0.085)	(0.149)
$SwitchToBasic_{it}$	-0.430***	-0.386***	-0.444***	-0.414**	-0.752***	-0.584***	-0.796***	-0.701***	-0.679***	-0.642***	-0.731***	-0.828***
	(0.105)	(0.109)	(0.117)	(0.186)	(0.146)	(0.159)	(0.155)	(0.202)	(0.122)	(0.145)	(0.107)	(0.259)
$FullSupport_{it}$		0.474***				0.391***				0.295*		
$\times EmploymentTop50_i$		(0.093)				(0.114)				(0.168)		
$SwitchToBasic_{it}$		-0.049				-0.290				-0.058		
$\times EmploymentTop50_i$		(0.185)				(0.272)				(0.237)		
$FullSupport_{it}$			0.280**				0.338**				0.349	
$\times EmploymentTop25_i$			(0.123)				(0.154)				(0.329)	
$SwitchToBasic_{it}$			0.065				0.277				0.633	
$\times EmploymentTop25_i$			(0.241)				(0.375)				(0.429)	
$FullSupport_{it}$				0.088***				0.097***				0.165***
$\times lnEmployment_i$				(0.027)				(0.035)				(0.063)
$SwitchToBasic_{it}$				-0.002				-0.022				0.100
$\times lnEmployment_i$				(0.077)				(0.083)				(0.142)
Constant	-0.195*	-0.188*	-0.192*	-0.190*	-0.302	-0.286	-0.267	-0.296	-0.688*	-0.729**	-0.717*	-0.750**
	(0.112)	(0.111)	(0.112)	(0.110)	(0.473)	(0.454)	(0.440)	(0.459)	(0.392)	(0.367)	(0.373)	(0.351)
Observations	87,964	87,964	87,964	87,964	37,837	37,837	37,837	37,837	13,262	13,262	13,262	13,262
Buyers	5,134	5,134	5,134	5,134	2,029	2,029	2,029	2,029	687	687	687	687
R^2	0.301	0.304	0.302	0.303	0.336	0.340	0.338	0.339	0.397	0.400	0.400	0.405
$e^{\hat{\beta}_1} - 1$	178.6%	102.9%	155.2%		192.9%	131.6%	173.1%		187.2%	146.7%	175.0%	
$e^{\hat{\beta}_1+\hat{\gamma}_1} - 1$	81.2%	37.9%	63.6%		38.0%	29.1%	23.2%		45.6%	29.9%	32.3%	
$\hat{\beta}_1 + \hat{\gamma}_1 = 0$ test p-value	0.000	0.008	0.000	0.081	0.022	0.114	0.163	0.500	0.002	0.046	0.006	0.647
$e^{\hat{\beta}_1+\hat{\beta}_2} - 1$		225.8%	237.6%			242.4%	282.8%			231.4%	289.9%	
$e^{\hat{\beta}_1+\hat{\beta}_2+\hat{\gamma}_1+\hat{\gamma}_2} - 1$		110.9%	131.1%			42.9%	127.8%			64.6%	253.2%	
$\hat{\beta}_1 + \hat{\beta}_2 + \hat{\gamma}_1 + \hat{\gamma}_2 = 0$ test p-value		0.000	0.000	0.001		0.092	0.013	0.164		0.011	0.005	0.256

All regressions include calendar (τ_t) and lifetime time dummies (l_{it}). Robust standard errors, clustered on customers, in parentheses. Coefficients $\hat{\beta}_1$, $\hat{\beta}_2$, $\hat{\gamma}_1$ and $\hat{\gamma}_2$ correspond to the estimated parameters of Model (3). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. “Survey Data Available” sample corresponds to buyers in “Baseline” sample for which the survey data (e.g., firm size) is available. Results with CEM1 subsample are in main text of the dissertation.

Table E.7: Results Considering Firm Size for $FractionParallel_{it}$ using Different Subsamples

Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Subsample	Survey Data Available (Baseline)				CEM2				CEM3			
$FullSupport_{it}$	0.106***	0.069***	0.103***	0.068***	0.106***	0.059***	0.106***	0.072***	0.108***	0.061***	0.106***	0.048
	(0.011)	(0.016)	(0.013)	(0.018)	(0.014)	(0.018)	(0.016)	(0.022)	(0.019)	(0.023)	(0.021)	(0.031)
$SwitchToBasic_{it}$	-0.007	0.026	-0.009	0.012	-0.025	0.036	-0.041	-0.011	-0.018	0.034	-0.018	0.052
	(0.020)	(0.029)	(0.024)	(0.036)	(0.031)	(0.038)	(0.034)	(0.047)	(0.028)	(0.043)	(0.030)	(0.062)
$FullSupport_{it}$		0.055***				0.079***				0.091**		
$\times EmploymentTop50_i$		(0.021)				(0.025)				(0.036)		
$SwitchToBasic_{it}$		-0.051				-0.108*				-0.106*		
$\times EmploymentTop50_i$		(0.040)				(0.059)				(0.055)		
$FullSupport_{it}$			0.009				-0.001				0.009	
$\times EmploymentTop25_i$			(0.024)				(0.031)				(0.055)	
$SwitchToBasic_{it}$			0.008				0.091				0.003	
$\times EmploymentTop25_i$			(0.044)				(0.080)				(0.083)	
$FullSupport_{it}$				0.012**				0.014*				0.029**
$\times lnEmployment_i$				(0.005)				(0.008)				(0.013)
$SwitchToBasic_{it}$				-0.007				-0.006				-0.036
$\times lnEmployment_i$				(0.013)				(0.018)				(0.027)
Constant	-0.057**	-0.056**	-0.057**	-0.056**	-0.114	-0.108	-0.116	-0.113	-0.422***	-0.424***	-0.422***	-0.423***
	(0.025)	(0.025)	(0.025)	(0.025)	(0.120)	(0.122)	(0.120)	(0.121)	(0.120)	(0.121)	(0.120)	(0.121)
Observations	87,964	87,964	87,964	87,964	37,837	37,837	37,837	37,837	13,262	13,262	13,262	13,262
Buyers	5,134	5,134	5,134	5,134	2,029	2,029	2,029	2,029	687	687	687	687
R^2	0.037	0.037	0.037	0.037	0.042	0.044	0.042	0.043	0.060	0.064	0.060	0.063
$\hat{\beta}_1$	0.106	0.069	0.103		0.106	0.059	0.106		0.108	0.061	0.106	
$\hat{\beta}_1 + \hat{\gamma}_1$	0.098	0.095	0.094		0.081	0.095	0.065		0.090	0.095	0.088	
$\hat{\beta}_1 + \hat{\gamma}_1 = 0$ test p-value	0.000	0.003	0.000	0.032	0.010	0.022	0.058	0.205	0.004	0.043	0.009	0.127
$\hat{\beta}_1 + \hat{\beta}_2$		0.124	0.112			0.138	0.105			0.151	0.115	
$\hat{\beta}_1 + \hat{\beta}_2 + \hat{\gamma}_1 + \hat{\gamma}_2$		0.098	0.110			0.066	0.155			0.079	0.100	
$\hat{\beta}_1 + \hat{\beta}_2 + \hat{\gamma}_1 + \hat{\gamma}_2 = 0$ test p-value		0.001	0.003	0.002		0.126	0.025	0.064		0.032	0.162	0.034

All regressions include calendar (τ_t) and lifetime time dummies (l_{it}). Robust standard errors, clustered on customers, in parentheses. Coefficients $\hat{\beta}_1$, $\hat{\beta}_2$, $\hat{\gamma}_1$ and $\hat{\gamma}_2$ correspond to the estimated parameters of Model (3). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. "Survey Data Available" sample corresponds to buyers in "Baseline" sample for which the survey data (e.g., firm size) is available. Results with CEM1 subsample are in main text of the dissertation.

E.3 Results with Instrumental Variables

Varying Number of Dummies per Support Interaction Type

In this section we run the process ran in Table 2.4 and Table 2.5 in the main text. However, instead of using 2 indicators for each of the 3 types of exogenous failure types identified, we use 1 or 3. We work within the same CEM1 subsample employed in the main text of the dissertation.

Table E.8: Probit for $FullSupport_{it}$ and First Stage Results with fitted $FullSupport_{it}^{f}$ (CEM1 Subsample, varying indicators)

Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of Indicators per Failure	1 Indicator				3 Indicators			
Failure Types	Outage	Network	Host	All 3	Outage	Network	Host	All 3
Part A. First Stage Regression of Fitted $FullSupport_{it}^{f}$ on Real $FullSupport_{it}$								
$FullSupport_{it}^{f}$	0.652***	0.885***	0.662***	0.651***	0.658***	0.893***	0.606***	0.645***
	(0.081)	(0.222)	(0.112)	(0.074)	(0.071)	(0.214)	(0.095)	(0.065)
Observations	48,725	48,725	48,725	48,725	48,715	48,718	48,725	48,708
Buyers	2,685	2,685	2,685	2,685	2,685	2,685	2,685	2,685
R^2	0.137	0.114	0.121	0.141	0.141	0.114	0.127	0.145
First Stage F Statistic	65.008	15.940	34.822	76.756	84.870	17.345	41.075	98.052
Part B. Descriptive Statistics of $FullSupport_{it}^{f}$								
Mean	0.078	0.078	0.078	0.078	0.078	0.078	0.078	0.078
Std. Dev.	0.084	0.063	0.074	0.089	0.087	0.064	0.081	0.092
Min	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Max	0.724	0.626	0.495	0.805	0.928	0.674	0.716	0.960

Table E.8 (continued)

Part C. Coefficients of Probit with $FullSupport_{it}$ as dependent variable								
$FailOutage1_{it}$	1.171***			0.954***	0.990***			0.769***
	(0.036)			(0.038)	(0.042)			(0.045)
$FailOutage2_{it}$					0.488***			0.383***
					(0.092)			(0.097)
$FailOutage3_{it}$					0.565***			0.576***
					(0.147)			(0.168)
$FailOutage1_{it}$	-0.589***			-0.688***	-0.419**			-0.598***
$\times Semester1_{it}$	(0.184)			(0.196)	(0.196)			(0.219)
$FailNetwork1_{it}$		0.714***		0.231***		0.712***		0.258***
		(0.064)		(0.072)		(0.066)		(0.075)
$FailNetwork2_{it}$						0.269		-1.001***
						(0.257)		(0.319)
$FailNetwork1_{it}$		-0.215		0.076		-0.213		0.075
$\times Semester1_{it}$		(0.357)		(0.341)		(0.357)		(0.342)
$FailHost1_{it}$			0.603***	0.413***			0.434***	0.354***
			(0.023)	(0.025)			(0.027)	(0.028)
$FailHost2_{it}$							0.275***	0.043
							(0.054)	(0.058)
$FailHost3_{it}$							0.542***	0.304***
							(0.069)	(0.074)
$FailHost1_{it}$			0.015	0.170*			0.121	0.184*
$\times Semester1_{it}$			(0.088)	(0.091)			(0.096)	(0.098)
$FailHost2_{it}$							0.021	0.298
$\times Semester1_{it}$							(0.257)	(0.262)
$FailHost3_{it}$							-0.027	-0.116
$\times Semester1_{it}$							(0.435)	(0.486)
Constant	-0.731***	-0.528***	-0.749***	-0.890***	-0.716***	-0.529***	-0.783***	-0.870***
	(0.110)	(0.105)	(0.108)	(0.112)	(0.109)	(0.105)	(0.108)	(0.112)
Observations	48,425	48,425	48,425	48,425	48,415	48,418	48,425	48,408
Pseudo-R2	0.127	0.092	0.114	0.140	0.131	0.092	0.123	0.144

Linear regressions in Part A include monthly calendar and semester lifetime time dummies. Robust standard errors, clustered on customers, in parentheses. Descriptive statistics in Part B correspond to $FullSupport_{it}^{f}$ within CEM1 and after considering periods in which population was not yet at risk (see text for details).). Part C shows coefficients of Probit regressions that include monthly calendar dummies and semester (6-month) lifetime dummies. Robust standard errors in parentheses. Some of the interaction terms (e.g., $FailOutage2_{it} \times Semester1_{it}$) are dropped out of model since parameter is always equal zero. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Using the first stage regressions in Part A of Table E.8, we run a second stage regression for $lnMemory_{it}$ in Table E.9 and for $FractionParallel_{it}$ in Table E.10. Each column in the latter two tables corresponds to the same column number in the first table.

Table E.9: Second stage results for $lnMemory_{it}$ using $FullSupport^{f}_{it}$ from Table E.8 as IV (CEM1 Subsample, varying indicators)

Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of Indicators per Failure	1 Indicator				3 Indicators			
Failure Types Used	Outage	Network	Host	All 3	Outage	Network	Host	All 3
$FullSupport_{it}$	2.510***	2.955***	4.081***	2.921***	2.490***	2.854***	3.844***	2.936***
	(0.327)	(1.034)	(0.650)	(0.347)	(0.285)	(0.973)	(0.565)	(0.311)
Observations	48,725	48,725	48,725	48,725	48,715	48,718	48,725	48,708
Buyers	2,685	2,685	2,685	2,685	2,685	2,685	2,685	2,685

All regressions include monthly calendar and lifetime time dummies. Robust standard errors, clustered on customers, in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table E.10: Second stage results for $FractionParallel_{it}$ using $FullSupport^{f}_{it}$ from Table E.8 as IV (CEM1 Subsample, varying indicators)

Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of Indicators per Failure	1 Indicator				3 Indicators			
Failure Types Used	Outage	Network	Host	All 3	Outage	Network	Host	All 3
$FullSupport_{it}$	0.353***	0.242*	0.475***	0.394***	0.389***	0.243*	0.478***	0.430***
	(0.083)	(0.133)	(0.121)	(0.079)	(0.079)	(0.128)	(0.108)	(0.076)
Observations	48,725	48,725	48,725	48,725	48,715	48,718	48,725	48,708
Buyers	2,685	2,685	2,685	2,685	2,685	2,685	2,685	2,685

All regressions include monthly calendar and lifetime time dummies. Robust standard errors, clustered on customers, in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Varying Sample

We now use 2 dummies per support type to generate the fitted values of $FullSupport_{it}$, as in the main text of the dissertation, but we run it on different subsamples. We use only the subsample for which the support interactions data is available: Support, CEM2 and CEM3; results with CEM1 already appear in main text of the dissertation.

Table E.11: Probit for $\boldsymbol{FullSupport_{it}}$ and First Stage Results with fitted $\boldsymbol{FullSupport_{it}^{f}}$ (Varying subsamples, 2 indicators)

Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Subsample	Support				CEM2				CEM3			
Failure Types	Outage	Network	Host	All 3	Outage	Network	Host	All 3	Outage	Network	Host	All 3
Part A. First Stage Regression of Fitted $FullSupport_{it}^{f}$ on Real $FullSupport_{it}$												
$FullSupport_{it}^{f}$	0.344***	0.314***	0.307***	0.349***	0.652***	0.991***	0.571***	0.631***	0.627***	0.934***	0.584***	0.620***
	(0.025)	(0.060)	(0.029)	(0.024)	(0.072)	(0.217)	(0.101)	(0.068)	(0.087)	(0.258)	(0.129)	(0.087)
Observations	298,539	298,539	298,539	298,539	37,837	37,837	37,837	37,837	13,262	13,254	13,262	13,254
Buyers	20,040	20,040	20,040	20,040	2,029	2,029	2,029	2,029	687	687	687	687
R^2	0.077	0.063	0.067	0.078	0.165	0.140	0.147	0.167	0.220	0.186	0.200	0.221
First Stage F Statistic	188.813	27.644	110.455	210.269	81.360	20.900	32.039	85.348	52.199	13.089	20.496	50.321
Part B. Descriptive Statistics of $FullSupport_{it}^{f}$												
Mean	0.183	0.183	0.183	0.183	0.093	0.093	0.093	0.093	0.155	0.154	0.155	0.154
Std. Dev.	0.145	0.127	0.136	0.148	0.102	0.078	0.094	0.107	0.139	0.104	0.128	0.143
Min	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Max	0.951	0.903	0.817	0.997	0.895	0.658	0.670	0.936	0.904	0.703	0.705	0.936

Table E.11 (continued)

Part C. Coefficients of Probit with $FullSupport_{it}$ as dependent variable

$FailOutage1_{it}$	0.897***			0.764***	0.920***			0.710***	1.162***			0.913***
	(0.015)			(0.016)	(0.044)			(0.047)	(0.070)			(0.076)
$FailOutage2_{it}$	0.473***			0.386***	0.726***			0.648***	0.487***			0.379***
	(0.027)			(0.028)	(0.079)			(0.084)	(0.124)			(0.129)
$FailOutage1_{it}$ $\times Semester1_{it}$	0.183***			0.137***	-0.403*			-0.603***	-0.067			-0.084
	(0.037)			(0.039)	(0.210)			(0.233)	(0.320)			(0.338)
$FailOutage2_{it}$ $\times Semester1_{it}$	0.142			0.105	-0.584			-0.457				
	(0.093)			(0.096)	(0.590)			(0.526)				
$FailNetwork1_{it}$		0.433***		-0.040		0.710***		0.298***		0.704***		0.307**
		(0.026)		(0.030)		(0.069)		(0.072)		(0.107)		(0.123)
$FailNetwork2_{it}$		0.310***		0.022		0.120		-0.955***				
		(0.059)		(0.069)		(0.254)		(0.287)				
$FailNetwork1_{it}$ $\times Semester1_{it}$		0.118		0.228***		-0.200		0.059		-0.333		-0.037
		(0.072)		(0.081)		(0.366)		(0.346)		(0.399)		(0.390)
$FailNetwork2_{it}$ $\times Semester1_{it}$		0.424**		0.567**								
		(0.192)		(0.227)								
$FailHost1_{it}$			0.326***	0.226***			0.416***	0.342***			0.498***	0.401***
			(0.010)	(0.011)			(0.030)	(0.030)			(0.045)	(0.047)
$FailHost2_{it}$			0.405***	0.161***			0.554***	0.198***			0.553***	0.160**
			(0.015)	(0.017)			(0.045)	(0.049)			(0.069)	(0.077)
$FailHost1_{it}$ $\times Semester1_{it}$			0.176***	0.182***			0.133	0.192*			-0.141	-0.134
			(0.023)	(0.024)			(0.108)	(0.110)			(0.166)	(0.180)
$FailHost2_{it}$ $\times Semester1_{it}$			0.018	0.044			0.062	0.383			0.094	0.196
			(0.046)	(0.052)			(0.241)	(0.242)			(0.441)	(0.379)
Constant	-1.240***	-1.075***	-1.244***	-1.329***	-0.642***	-0.451***	-0.720***	-0.804***	-0.814***	-0.470***	-0.685***	-0.905***
	(0.031)	(0.030)	(0.031)	(0.032)	(0.117)	(0.112)	(0.118)	(0.120)	(0.185)	(0.173)	(0.175)	(0.193)
Observations	294,208	294,208	294,208	294,208	37,579	37,579	37,579	37,579	13,087	13,079	13,087	13,079
Pseudo-R2	0.144	0.118	0.132	0.149	0.143	0.104	0.133	0.156	0.144	0.097	0.129	0.154

Linear regressions in Part A include monthly calendar and semester lifetime time dummies. Robust standard errors, clustered on customers, in parentheses. Descriptive statistics in Part B correspond to $FullSupport_{it}^{f}$ within CEM1 and after considering periods in which population was not yet at risk (see text for details).). Part C shows coefficients of Probit regressions that include monthly calendar dummies and semester (6-month) lifetime dummies. Robust standard errors in parentheses. Some of the interaction terms (e.g., $FailOutage2_{it} \times Semester1_{it}$) are dropped out of model since parameter is always equal zero. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Using the first stage regressions in Part A of Table E.11, we run a second stage regression for $lnMemory_{it}$ in Table E.12 and for $FractionParallel_{it}$ in Table E.13. Each column in the latter two tables corresponds to the same column number in the first table.

Table E.12: Second stage results for $lnMemory_{it}$ using $FullSupport_{it}^{f}$ from Table E.11 as IV (Varying subsamples, 2 indicators)

Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Subsample	Support				CEM2				CEM3			
Failure Types Used	Outage	Network	Host	All 3	Outage	Network	Host	All 3	Outage	Network	Host	All 3
$FullSupport_{it}$	4.542***	6.825***	8.434***	5.373***	2.634***	2.334***	4.184***	3.051***	2.669***	2.449**	3.335***	2.827***
	(0.317)	(1.219)	(0.757)	(0.354)	(0.312)	(0.844)	(0.729)	(0.353)	(0.404)	(1.113)	(0.746)	(0.446)
Observations	298,381	298,381	298,381	298,381	37,837	37,837	37,837	37,837	13,262	13,254	13,262	13,254
Buyers	19,882	19,882	19,882	19,882	2,029	2,029	2,029	2,029	687	687	687	687
R^2	-0.745	-2.075	-3.398	-1.155	-0.170	-0.075	-0.970	-0.334	-0.282	-0.181	-0.696	-0.375

All regressions include monthly calendar and lifetime time dummies. Robust standard errors, clustered on customers, in parentheses.
$^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$

Table E.13: Second stage results for $FractionParallel_{it}$ using $FullSupport_{it}^{f}$ from Table E.11 as IV (Varying subsamples, 2 indicators)

Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Subsample	Support				CEM2				CEM3			
Failure Types Used	Outage	Network	Host	All 3	Outage	Network	Host	All 3	Outage	Network	Host	All 3
$FullSupport_{it}$	0.521***	0.609***	0.794***	0.589***	0.395***	0.305**	0.472***	0.426***	0.408***	0.363	0.324**	0.386***
	(0.058)	(0.165)	(0.096)	(0.058)	(0.083)	(0.136)	(0.128)	(0.084)	(0.111)	(0.235)	(0.146)	(0.110)
Observations	298,381	298,381	298,381	298,381	37,837	37,837	37,837	37,837	13,262	13,254	13,262	13,254
Buyers	19,882	19,882	19,882	19,882	2,029	2,029	2,029	2,029	687	687	687	687
R^2	-0.134	-0.196	-0.365	-0.181	-0.098	-0.039	-0.165	-0.124	-0.154	-0.108	-0.070	-0.131

All regressions include monthly calendar and lifetime time dummies. Robust standard errors, clustered on customers, in parentheses.
$^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$

APPENDIX F

FOR CHAPTER 3: GOALS AND CONTENT OF EPE TREATMENT

This appendix includes information handed by the cloud infrastructure services provider that describes their EPE treatment. This was the information used to train the agents applying the treatment.

F.1 General Information and Goals

F.1.1 Goals

The EPE treatment has the following goals:

1. Fraud verification process
2. Confirm product fit
3. Setting expectations
4. Customer education

F.1.2 Topics Covered in Customer Education Component

The following are the topics that customers should understand after having received education through the EPE treatment:

1. Basic Services: Usage of control panel, billing, account management.
2. Building in the Cloud: Provisioning servers, cloud storage service, cloud load balancer service
3. Security & Remote Access: Windows firewall and IP tables, SSH, and remote access
4. Managing Server Image: Setting automated backups, restoring from an image, cloning servers
5. DNS and Domain Management: Migrating domains, utilizing the DNS Control Panel, DNSaaS
6. Uploading Content: SFTP vs FTP, rsync, cloud storage service

F.2 Template of Support Ticket

After having spoken with the customer and applied the treatment, or having attempted to (in case the conversation could not be established), a support ticket following the template below is opened for the customer. The template has been redacted to comply with the provider's NDA.

Hello [First Name],

Thank you for your time today.

Per our discussion [Put comments about what was discussed. i.e., the customer wanted to install CPANEL although not supported and we told them it will not be supported in case you need to use it for documentation purposes.]

Here are a few additional things to remember:

Cloud Servers "without" [Full Support] – New Public and Private IP's will be assigned with a 5 IP Limit per Cloud Server.

With Cloud Servers, we manage the network, the hardware, and the virtualization layer. You get full control of your virtual instance—that means you call the shots when it comes to the OS, server applications and code. *** Please Note: Support will not be able to log into your Cloud Servers* **. If you need help, support can present you with a Knowledge base article or a forum link to help guide you to find a possible solution to your issue.

Backups?

Backups are extended snapshots. You are allowed up to 3 snapshots. You can create an image of any cloud server containing less than 80GB of data —and you can use this image to restore a server or clone a new one. You can create an unlimited number of images on-demand, or you can schedule an automatic daily or weekly image.

The snapshot feature does not specifically backup your database, the database will be included in your flat image file! If you would like to just backup the database you can run commands to export and zip contents of MS SQL and MySql databases, and use the API to upload to Cloud Files.

Keep in mind that we cannot restore individual files or directories from backup. We can only restore complete Cloud Servers. If you need something restored from a Cloud Server, you must fire up a new server, restore the data, pull what you need, and shut down the new server

Please be sure to check out our [online] Essentials guide for step-by-step directions to get you up and running, by visiting [URL].

I have sent you additional details with useful information via email.

Again thank you for choosing [Provider]!

Sincerely,

The [Provider] Cloud Team

APPENDIX G

FOR CHAPTER 3: TICKET SUBJECTS CONSIDERED PROVIDER-INITIATED SUPPORT INTERACTIONS

The provider frequently uses the support ticketing system to communicate with its customers. The following is the list of subjects of tickets that have been used to identify such provider-initiated support interactions. The list was built by identifying tickets that were identical to each other or that followed a template. The list of subjects presented here is not exhaustive, yet it does encompass all tickets that pertain the studied sample. Given our NDA with the provider, we use [Provider] and [Offering] to redact the name the provider and its cloud infrastructure service offering. We also use the percentage symbol (%) to represent wildcards that can substitute any other character(s).

- Welcome to [Offering] (various similar subjects)
- Welcome to [Provider] (various similar subjects)
- Getting Started with [Offering]
- %Excessive Swapping%
- %Excessive DNS Queries%
- %Excessive DNS Requests%
- Notice: End of Sale for certain Linux Distros
- [Feature of Offering] Incident
- Fedora 14 End Of Sale Notice
- Notice: Microsoft Security Bulletin
- Notice: [Offering] Server Migration Pending
- Announcing [Feature of Offering]

APPENDIX H

FOR CHAPTER 3: RANDOM ASSIGNMENT OF TREATMENT

This appendix provides supplemental information concerning the random assignment of the EPE treatment. We first verify that treated and control customers are similar to each other (e.g., have similar firm size), and then elaborate on the proportion of customers treated over time.

H.1 Attributes of Treated and Controls

Upon signup, customers are offered the opportunity to fill in an online survey in which they indicate some attributes about themselves such as their firm size, industry, and intended use case for the cloud service. A total of 605 of the 2,673 customers in the sample completed the survey (a 22.6% response rate). We show that treated and control customers do not vary significantly across any of these attributes, supporting the random treatment assignment assumption. The numbers in parentheses are the number of customers in each of the shown categories.

In the case of firm size, of the 605 completed the employment item, 15% were treated. The proportion of treated customers is very similar across all employment ranges, as shown in Figure H.1. The means difference test for employment between the groups was insignificant as well.

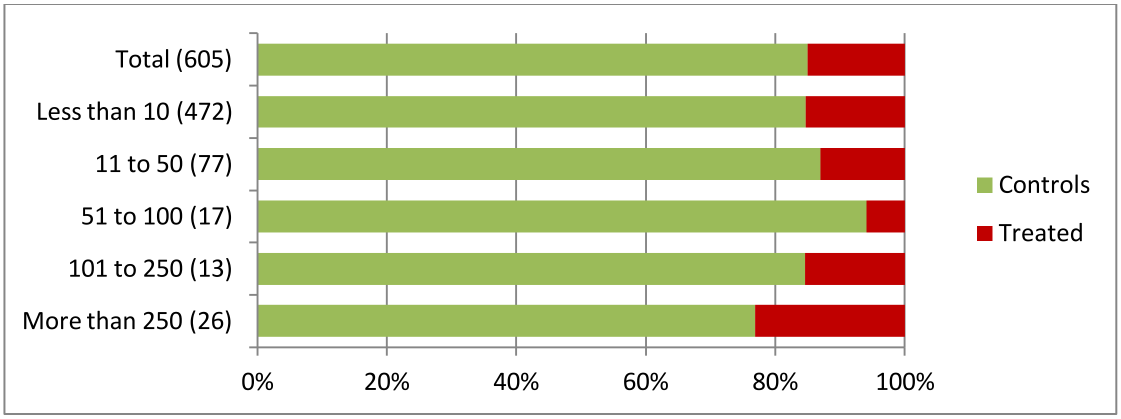

Figure H.1: Proportion of Controls and Treated per Customer Size (Employment)

Not all firms indicated the industry to which they belong, but we have data on 473 of them. Within these, 14% were treated. In Figure H.2 we show the proportion of treated customers is similar across the 10 most popular industries in the data.

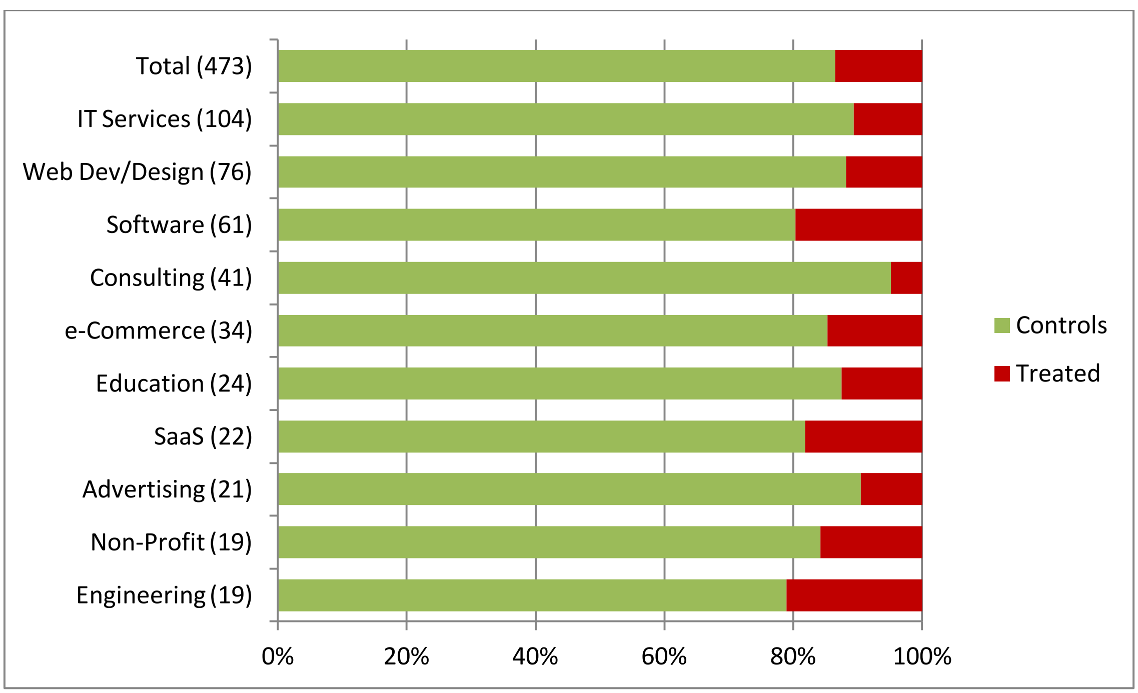

Figure H.2: Proportion of Controls and Treated per 10 Most Popular Industries

Finally, we have data on the customers' intended use case for the cloud service. We follow the procedures described in section 2.5.2 to categorize the use cases into 5 general categories. Since customers can choose more than a single use case, in this case we have more responses than customers, and have 968 responses. Once again, the proportion of treated customers is very similar across all use cases, as shown in Figure H.3 below.

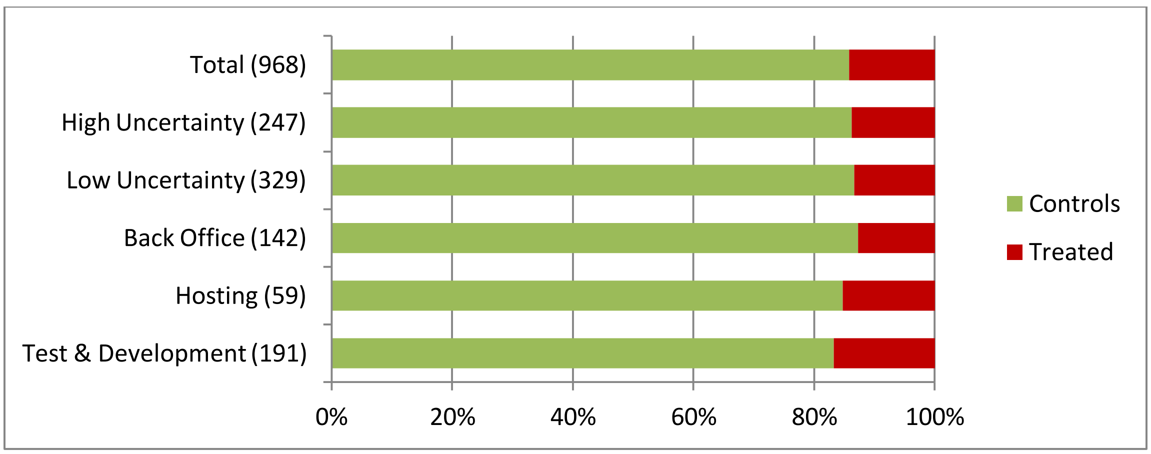

Figure H.3: Proportion of Controls and Treated per Intended Use Case for Service

H.2 Proportion of Treated Customers over Time

The first figure in this appendix shows the total number of agents in the verification team that were applying the treatment per unit of time.

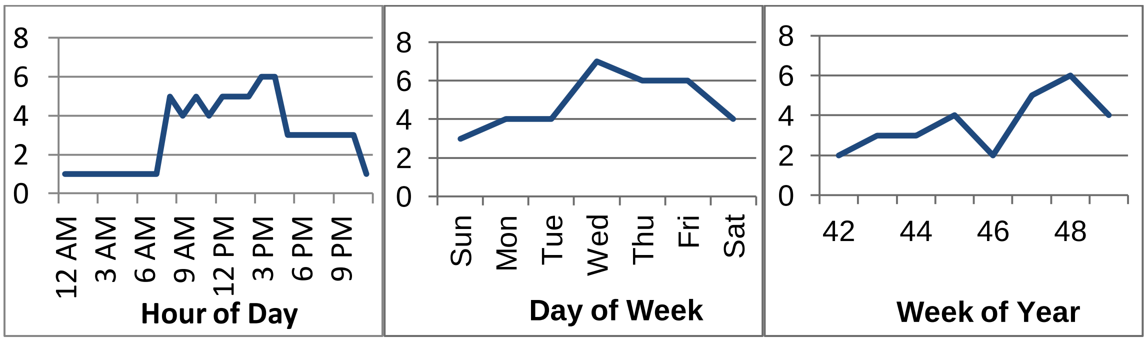

Figure H.4: Number of Agents Applying Treatment per Unit of Time

The remaining figures describe the number of customers adopting the service and the proportion of those treated over varying units of time. In all these figures, the shaded area is measured by the vertical axis on the left ("Number of Accounts") and represents the number of customers adopting by the unit of time in the horizontal axis. Within the shaded area, the red area represents the number of customers treated, while the green area represents the controls. We also plot the proportion of customers being treated during each unit of time, which is computed as the number of treated signups divided by total number of signups. This metric is represented by the blue line, for which the values are displayed on the right vertical axis.

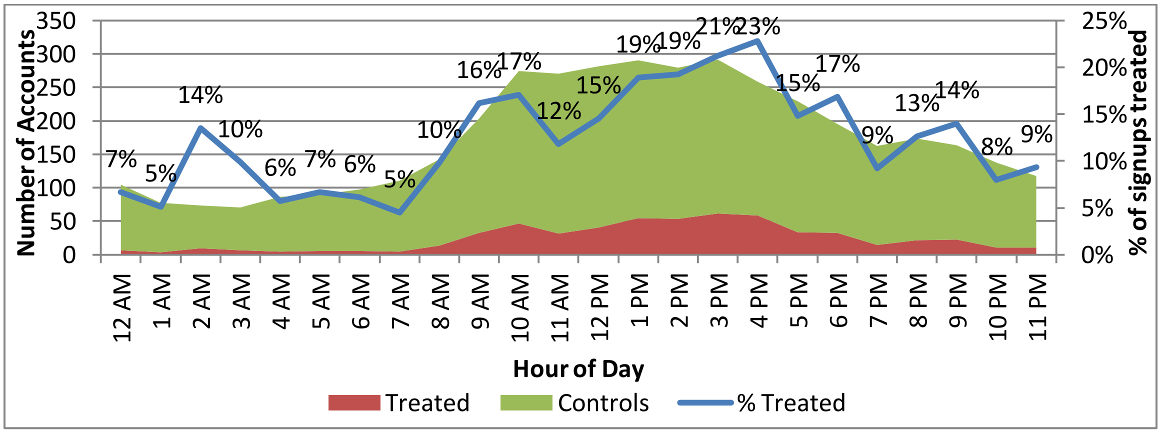

Figure H.5: Treatment by Hour of the Day

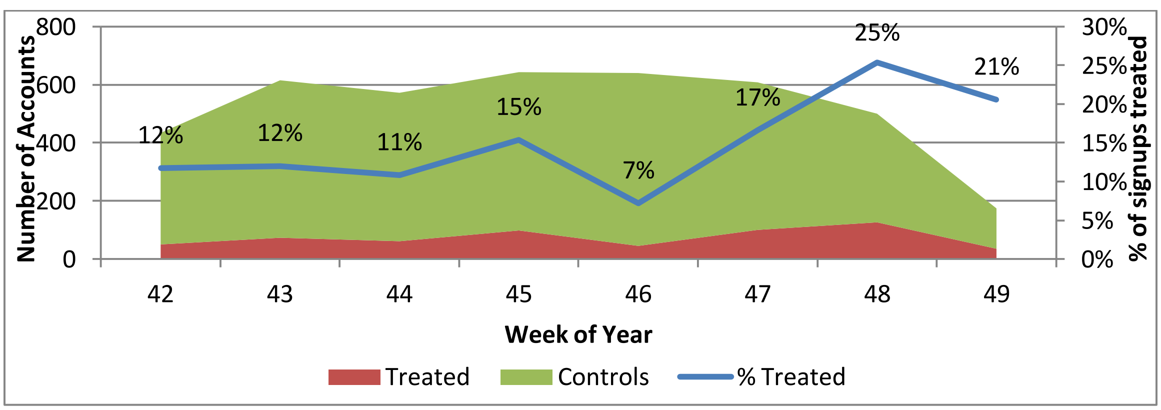

Figure H.6: Treatment by Week of the Year

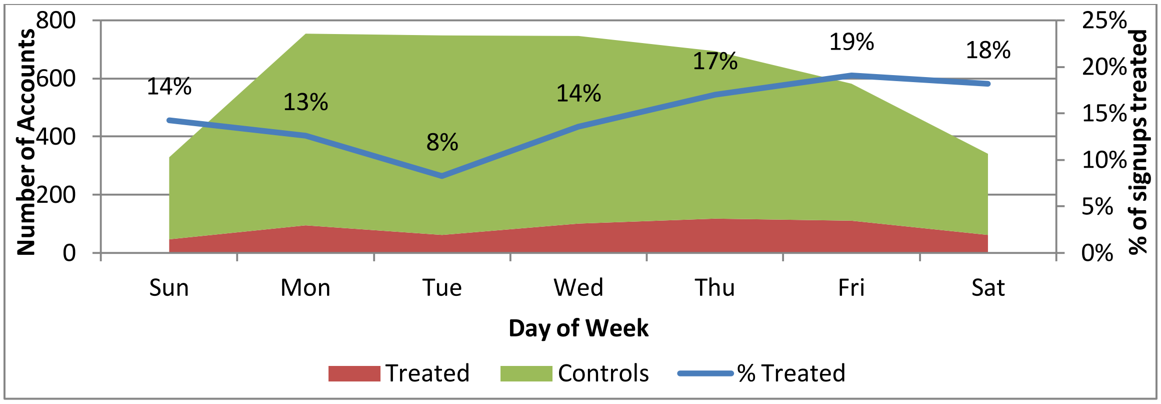

Figure H.7: Treatment by Day of the Week (considering entire experiment)

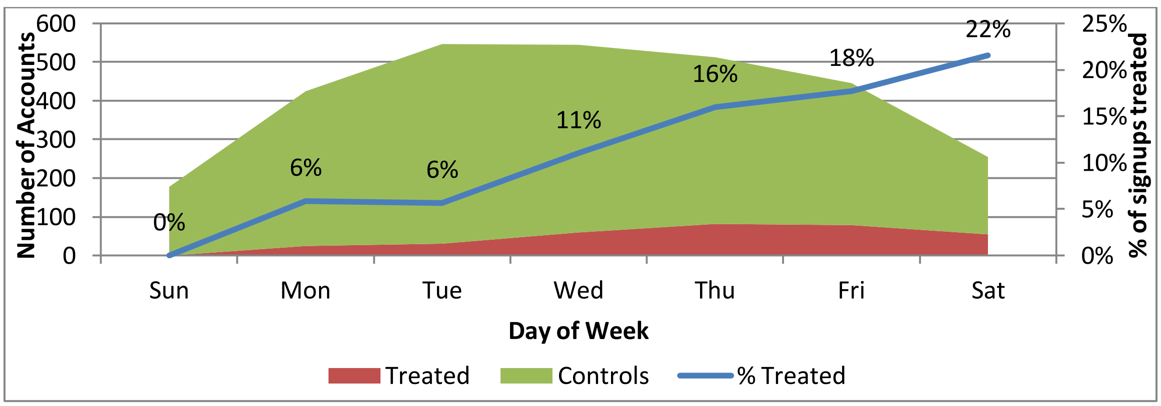

Figure H.8: Treatment by Day of the Week (Weeks 42 to 46)

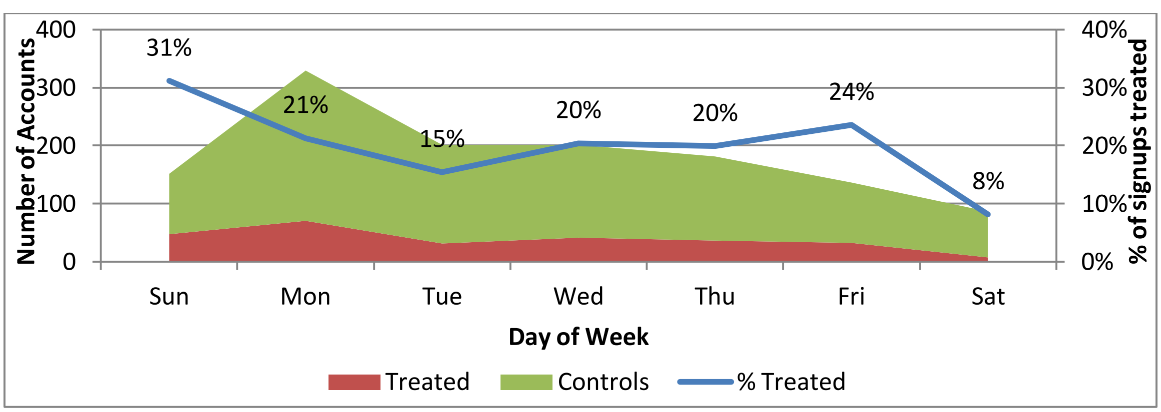

Figure H.9: Treatment by Day of the Week (Weeks 47 to 49)

APPENDIX I

FOR CHAPTER 3: EPE AND FULL SUPPORT CUSTOMERS

This appendix justifies the exclusion of the full support customers from the analyzed sample and shows that our results for the basic support customers are robust to their inclusion.

I.1 Descriptive Statistics

The excluded group includes 379 full support customers (that may own more than 1 account) of which 157 (41%) started off with full support since adoption, 130 (34%) upgraded from basic to full support within a week, and another 40 (11%) upgraded during the rest of their first month. In other words, 327 (86%) of them were using full support before the end of their first month. We show their descriptive statistics and compare them to the basic support customer in Table I.1.

Table I.1: Descriptive Statistics of Basic and Full Support Customers

Customer Group Number of Customers	All Customers 3,052				Basic Support 2,673				Full Support 379				t-test of mean difference	
Variable	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max	Diff.	p-value
EPE_i	0.153	0.360	0	1	0.137	0.344	0	1	0.266	0.443	0	1	-0.130	0.000
$SurvivedWeek1_i$	0.943	0.231	0	1	0.936	0.245	0	1	0.997	0.051	0	1	-0.062	0.000
$SurvivedMonth1_i$	0.901	0.299	0	1	0.888	0.316	0	1	0.992	0.089	0	1	-0.104	0.000
$Failed_{it}$	0.355	0.478	0	1	0.374	0.484	0	1	0.227	0.419	0	1	0.147	0.000
$QuestionsWeek1_i$	0.949	3.187	0	18	0.591	1.445	0	18	3.475	7.743	0	64	-2.883	0.000
$QuestionsWeek2_i$	1.274	4.245	0	25	0.759	1.851	0	25	4.910	10.300	0	90	-4.151	0.000

There were 101 (26.6%) treated customers in this group. However, only 3 out of the 379 abandoned the service during the first month (i.e., 99.2% survive past first 30 days). Therefore, it is unlikely EPE has any effect in their early retention. Furthermore, comparison of their mean number of questions asked during the first week (i.e., $QuestionsWeek1_i$) suggests the excluded group asks 5.27 times more questions, and 5.84 times more during the first two weeks (i.e., $QuestionsWeek2_i$), confirming that they interact much more frequently with the provider. Differences in means across all variables are statistically significant.

I.2 Customer Retention

If we run our models for customer retention (i.e., models (1a), (1b), and (2a)) using a sample that includes only the 379 full support customers, we find that the treatment has no significant effect on customer retention; this is expected as too few full support customers abandon the service. We can, however, include the full support customers into our sample to show that the results remain consistent. The results of our models for customer retention with the sample that includes both basic and full support customers (N=3,052) are shown in columns (1) through (5) of Table I.2. In columns (6) through (10) we augment the models by adding a new indicator, $FullSupport_i$, which is turned on if customer i belongs to the previously excluded group and used full support. All results for the basic support customers (i.e., the coefficient for EPE_i) are consistent with those shown in Table 3.3 in the main text. Moreover, it is evident that having access to full support (i.e, $FullSupport_i$=1) is strongly and positively associated with customer retention.

Table I.2: Survival Results including Full Support Customers (No Interaction)

Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent Variable	$Survived\ Week1_i$		$Survuved\ Month1_i$		$Failed_{it}$	$Survived\ Week1_i$		$Survuved\ Month1_i$		$Failed_{it}$
Model	LPM	Probit	LPM	Probit	Cox Prop. Hazard	LPM	Probit	LPM	Probit	Cox Prop. Hazard
EPE_i	0.030*** (0.010)	0.322** (0.126)	0.052*** (0.013)	0.354*** (0.103)	-0.241*** (0.092)	0.024** (0.009)	0.264** (0.129)	0.043*** (0.013)	0.304*** (0.106)	-0.196** (0.093)
$FullSupport_i$						0.054*** (0.006)	1.244*** (0.318)	0.091*** (0.008)	1.142*** (0.213)	-0.580*** (0.110)
Observed Failures					1,085					1,085
Marginal Effect of EPE_i	0.030	0.029	0.052	0.051		0.024	0.024	0.043	0.044	
% Change in Hazard $(e^{\hat{\beta}} - 1)$					-21.40%					-17.79%

All regressions use the 3,053 (basic and full support) customers and include hourly, weekday, and weekly dummies. Robust standard errors in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

We can also interact the full support indicator with the treatment indicator (i.e., $EPE_i \times FullSupport_i$) and include the interaction in the model. This would test if the treatment affects full support customers differently than how it affects basic support customers. We show the results with it in Table I.3. Although the interaction term in columns (1) and (2) comes negative and significant, suggesting that the treatment and full support substitute for each other, we must note that the identification comes from the single full support customer who abandoned the service within the first week, who happens to have been treated. Similar situation arises when we use the first month survival indicator as dependent variable in the next two columns. The interaction term is significant with the linear probability model in column (3). However, it is no longer significant once if we use the more appropriate probit model in column (4). In column (5), as expected, the interaction terms has no effect in the overall hazard rate. Finally, we note that the treatment's effect for the basic support customers (i.e., the coefficient for EPE_i) is again consistent with prior results shown in Table 3.3 in the main text.

Table I.3: Survival Results including Full Support Customers (With Interaction)

Column	(1)	(2)	(3)	(4)	(5)
Dependent Variable	$SurvivedWeek1_i$		$SurvivedMonth1_i$		$Failed_{it}$
Model	LPM	Probit	LPM	Probit	Cox Prop. Hazard
EPE_i	0.032***	0.308**	0.053***	0.321***	-0.252**
	(0.011)	(0.133)	(0.015)	(0.109)	(0.102)
$FullSupport_i$	0.064***	4.345***	0.104***	1.237***	-0.675***
	(0.006)	(0.064)	(0.009)	(0.258)	(0.129)
$EPE_i \times FullSupport_i$	-0.041***	-3.801***	-0.055***	-0.408	0.403
	(0.015)	(0.389)	(0.019)	(0.456)	(0.252)
Observed Failures					1,085
Sum of Coefficients [a]	0.055	0.852	0.103	1.151	-0.524
Test of Sum $\neq 0$ p-value [a]	0.000	0.020	0.000	0.002	0.009

All regressions use the 3,053 (basic and full support) customers and include hourly, weekday, and weekly dummies. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

[a] Sum of 3 reported coefficients that represent effect of EPE on full support customers.

I.3 Demand for Technology Support

Next, we run our count-data models that have the number of questions asked as dependent variable. We again join the basic and full support customers in the sample and run an augmented version of the model that includes the $FullSupport_i$ variable and its interaction with the treatment variable.

Table I.4 shows the results for the number of questions asked during the first week. In columns (1) and (2), without controlling for the access to full support, the treatment has no statistically significant effect. Using the Poisson model, in columns (3) and (5), the inclusion of the $FullSupport_i$ variable does not change the results. However, the results with the negative binomial specifications in columns (4) and (6) do suggest a statistically significant effect for the treatment. In these cases, the results are consistent with those shown in Table 3.7 in the main text. It is also clear that the access to full support (i.e, $FullSupport_i$=1) increases the number of questions asked by customers.

Table I.4: Results for $QuestionsWeek1_i$ with Full Support Customers

Column	(1)	(2)	(3)	(4)	(5)	(6)
Model	Poisson	Negative Binomial	Poisson	Negative Binomial	Poisson	Negative Binomial
EPE_i	0.162	0.072	-0.141	-0.216**	-0.177	-0.259**
	(0.165)	(0.130)	(0.154)	(0.107)	(0.129)	(0.123)
$FullSupport_i$			1.773***	1.736***	1.764***	1.699***
			(0.119)	(0.110)	(0.137)	(0.126)
$EPE_i \times FullSupport_i$					0.060	0.157
					(0.261)	(0.241)
Squared correlation between actual and fitted number of questions	0.022	0.020	0.150	0.119	0.150	0.119

All regressions use the 3,052 (basic and full support) customers and include hourly, weekday, and weekly dummies. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Finally, in Table I.5 we show the results for the number of questions asked during the first 2 weeks. The treatment has no statistically significant effect, as was the case using the regular sample.

Table I.5: Results for $QuestionsWeek2_i$ with Full Support Customers

Column	(1)	(2)	(3)	(4)	(5)	(6)
Model	Poisson	Negative Binomial	Poisson	Negative Binomial	Poisson	Negative Binomial
EPE_i	0.203	0.132	-0.109	-0.171	-0.126	-0.197
	(0.158)	(0.138)	(0.148)	(0.110)	(0.137)	(0.128)
$FullSupport_i$			1.871***	1.857***	1.866***	1.833***
			(0.113)	(0.109)	(0.127)	(0.122)
$EPE_i \times FullSupport_i$					0.028	0.104
					(0.264)	(0.256)
Squared correlation between actual and fitted number of questions	0.019	0.017	0.157	0.122	0.157	0.122

All regressions use the 3,052 (basic and full support) customers and include hourly, weekday, and weekly dummies. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

APPENDIX J

FOR CHAPTER 3: MODEL COMPAISON AND ROBUSTNESS CHECKS FOR RESULTS FOR NUMBER OF QUESTIONS ASKED

J.1 Comparison of Fitted Probabilities of Poisson and Negative Binomial Models

We examine the performance of the Poisson and negative binomial models in fitting the probability distribution of $QuestionsWeek1_i$, as suggested by Cameron and Trivedi (2010) and using the software developed by Long and Freese (2006). First, we show in Table J.1 the distribution of $QuestionsWeek1_i$. We note that very few customers ask more than 6 questions during their first week, so we will limit our following analysis up to this point.

Table J.1: Distribution of $QuestionsWeek1_i$

$QuestionsWeek1_i$ (Count)	Frequency	Actual Probability	Cumulative Probability
0	1,930	72.20	72.20
1	412	15.41	87.62
2	146	5.46	93.08
3	83	3.11	96.18
4	40	1.50	97.68
5	19	0.71	98.39
6	16	0.60	98.99
7	4	0.15	99.14
8	5	0.19	99.33
9	5	0.19	99.51
10	4	0.15	99.66
11	2	0.07	99.74
12	2	0.07	99.81
14	2	0.07	99.89
15	1	0.04	99.93
17	1	0.04	99.96
18	1	0.04	100.00
Total	2,673	100.00	

Next, we compare the predicted probabilities by each of the two estimation procedures of each count occurring to the actual ones. Table J.2 shows the actual probability of each count (from 0 to 6 questions), the predicted probabilities by each model, their differences with respect to the actual values, and the Pearson Chi-Square statistic computed as $Count \times (Diff)^2/Predicted$, where $Count$ is the number of observations with a given number of questions, while $Diff$ and $Predicted$ are the values of the corresponding columns. We note that the Poisson model is much less accurate than the negative binomial model in predicting the probabilities of $QuestionsWeek1_i \leq 2$ occuring. For larger counts, both models perform similarly in terms of their accuracy. The models' performance can also be appreciated graphically. Figure J.1 plots the differences between the predicted and the actual probabilities, and it is evident that the Poisson model is farther from the actual values than the negative binomial model for the lower counts. Finally, the mean differences for the Poisson model is 0.405, while for the negative binomial model it is only 0.026 (these results are not reported in Table J.2).

Table J.2: Actual and Predicted Probabilities

$QuestionsWeek1_i$	Actual	Poisson Model			Negative Binomial Model		
(Count)	Probability	Predicted	Diff	Pearson	Predicted	Diff	Pearson
0	0.722	0.559	0.163	126.95	0.725	0.003	0.03
1	0.154	0.320	0.166	229.31	0.142	0.012	2.81
2	0.055	0.097	0.042	48.60	0.061	0.006	1.56
3	0.031	0.021	0.010	14.28	0.031	0.000	0.01
4	0.015	0.003	0.011	100.49	0.017	0.002	0.56
5	0.007	0.001	0.007	229.10	0.010	0.003	1.86
6	0.006	0.000	0.006	1420.73	0.006	0.000	0.02
Sum	0.990	1.000	0.405	2169.44	0.990	0.026	6.85

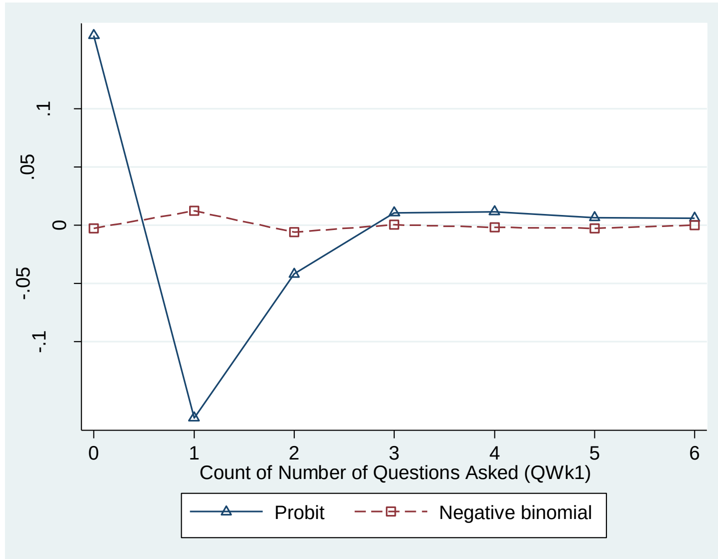

Figure J.1: Difference between Actual and Fitted Probabilities by Model

J.2 Results for Number of Questions Asked using Alternate Controls

The following results use the number of questions asked during the first week or first two weeks ($QuestionsWeek1_i$ or $QuestionsWeek2_i$) as dependent variable.

Table J.3: Results for $QuestionsWeek1_i$ Varying Sets of Time-of-Adoption Controls (Poisson Model)

Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
EPE_i	-0.218*	-0.213*	-0.208*	-0.206*	-0.227*	-0.222*	-0.208*	-0.206*	-0.236*
	(0.123)	(0.124)	(0.124)	(0.125)	(0.125)	(0.127)	(0.124)	(0.125)	(0.125)
Marginal Effect of EPE_i	-0.119	-0.117	-0.114	-0.113	-0.124	-0.121	-0.114	-0.113	-0.128
Controls Used									
$AdoptHour_i$	✓		✓		✓		✓		
$AdoptShift_i$		✓		✓		✓		✓	
$AdoptWeekday_i$	✓	✓	✓	✓	✓	✓	✓	✓	
$AdoptWeek_i$	✓	✓			✓	✓			
$AdoptRegime_i$			✓	✓			✓	✓	
$AdoptWeekday_i \times AdoptWeek_i$					✓	✓			
$AdoptWeekday_i \times AdoptRegime_i$							✓	✓	

Dependent variable is $QuestionsWeek1_i$. All regressions employ Poisson model and use all 2,673 customers in sample. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table J.4: Results for $QuestionsWeek1_i$ Varying Sets of Time-of-Adoption Controls (Negative binomial Model)

Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
EPE_i	-0.273**	-0.241*	-0.248**	-0.217*	-0.270**	-0.237*	-0.249**	-0.218*	-0.236*
	(0.122)	(0.125)	(0.123)	(0.126)	(0.122)	(0.125)	(0.123)	(0.126)	(0.125)
Marginal Effect of EPE_i	-0.146	-0.131	-0.134	-0.119	-0.145	-0.129	-0.135	-0.119	-0.128
Controls Used									
$AdoptHour_i$	✓		✓		✓		✓		
$AdoptShift_i$		✓		✓		✓		✓	
$AdoptWeekday_i$	✓	✓	✓	✓	✓	✓	✓	✓	
$AdoptWeek_i$	✓	✓			✓	✓			
$AdoptRegime_i$			✓	✓			✓	✓	
$AdoptWeekday_i \times AdoptWeek_i$					✓	✓			
$AdoptWeekday_i \times AdoptRegime_i$							✓	✓	

Dependent variable is $QuestionsWeek1_i$. All regressions employ negative binomial model and use all 2,673 customers in sample. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table J.5: Results for $QuestionsWeek2_i$ Varying Sets of Time-of-Adoption Controls (Poisson Model)

Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
EPE_i	-0.164	-0.163	-0.151	-0.153	-0.174	-0.173	-0.151	-0.153	-0.182
	(0.132)	(0.135)	(0.132)	(0.135)	(0.133)	(0.137)	(0.132)	(0.135)	(0.134)
Marginal Effect of EPE_i	-0.117	-0.117	-0.109	-0.110	-0.124	-0.123	-0.109	-0.110	-0.129
Controls Used									
$AdoptHour_i$	✓		✓		✓		✓		
$AdoptShift_i$		✓		✓		✓		✓	
$AdoptWeekday_i$	✓	✓	✓	✓	✓	✓	✓	✓	
$AdoptWeek_i$	✓	✓			✓	✓			
$AdoptRegime_i$			✓	✓			✓	✓	
$AdoptWeekday_i \times AdoptWeek_i$					✓	✓			
$AdoptWeekday_i \times AdoptRegime_i$							✓	✓	

Dependent variable is $QuestionsWeek2_i$. All regressions employ Poisson model and use all 2,673 customers in sample. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table J.6: Results for $QuestionsWeek2_i$ Varying Sets of Time-of-Adoption Controls (Negative binomial Model)

Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
EPE_i	-0.211*	-0.187	-0.190	-0.166	-0.211	-0.185	-0.192	-0.166	-0.182
	(0.127)	(0.133)	(0.128)	(0.134)	(0.128)	(0.134)	(0.128)	(0.134)	(0.134)
Marginal Effect of EPE_i	-0.149	-0.133	-0.135	-0.119	-0.148	-0.132	-0.136	-0.119	-0.129
Controls Used									
$AdoptHour_i$	✓		✓		✓		✓		
$AdoptShift_i$		✓		✓		✓		✓	
$AdoptWeekday_i$	✓	✓	✓	✓	✓	✓	✓	✓	
$AdoptWeek_i$	✓	✓			✓	✓			
$AdoptRegime_i$			✓	✓			✓	✓	
$AdoptWeekday_i \times AdoptWeek_i$					✓	✓			
$AdoptWeekday_i \times AdoptRegime_i$							✓	✓	

Dependent variable is $QuestionsWeek2_i$. All regressions employ negative binomial model and use all 2,673 customers in sample. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

REFERENCES

Achen, C. H. 2000. "Why Lagged Dependent Variables Can Supress the Explanatory Power of Other Independent Variables." Annual Meeting of the Political Methodology Section of the American Political Science Association.

Aksin, O. Z., P. T. Harker. 1999. To Sell or Not to Sell: Determining the Trade-Offs between Service and Sales in Retail Banking Phone Centers. *Journal of Service Research*. **2**(1) 19-33.

Aksin, Z., M. Armony, V. Mehrotra. 2007. The Modern Call Center: A Multi-Disciplinary Perspective on Operations Management Research. *Production and Operations Management*. **16**(6) 665-688.

Altonji, J. G., T. E. Elder, C. R. Taber. 2005. Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy*. **113**(1) 151-184.

Amazon Web Services. 2013. AWS Glossary. Retrieved Jul 19, 2013, http://docs.aws.amazon.com/general/latest/gr/glos-chap.html.

Anderson, T. W., C. Hsiao. 1981. Estimation of Dynamic Models with Error Components. *Journal of the American Statistical Association*. **76**(375) 598-606.

Angrist, J. D., J.-S. Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, Princeton, NJ.

Aral, S., E. Brynjolfsson, D. J. Wu. 2006. "Which Came First, IT or Productivity? Virtuous Cycle of Investment and Use in Enterprise Systems." Available at SSRN: http://ssrn.com/abstract=942291: pp. 1819-1839.

Archak, N., A. Ghose, P. G. Ipeirotis. 2011. Deriving the Pricing Power of Product Features by Mining Consumer Reviews. *Management Science*. **57**(8) 1485-1509.

Arellano, M., S. Bond. 1991. Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *The Review of Economic Studies*. **58**(2) 277-297.

Arellano, M., O. Bover. 1995. Another Look at the Instrumental Variable Estimation of Error-components Models. *Journal of Econometrics*. **68**(1) 29-51.

Armbrust, M., A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, M. Zaharia. 2010. A View of Cloud Computing. *Communications of the ACM*. **53-58**(4) 50-58.

Armony, M., C. Maglaras. 2004a. Contact Centers with a Call-Back Option and Real-Time Delay Information. *Operations Research*. **52**(4) 527-545.

Armony, M., C. Maglaras. 2004b. On Customer Contact Centers with a Call-Back Option: Customer Decisions, Routing Rules, and System Design. *Operations Research*. **52**(2) 271-292.

Attewell, P. 1992. Technology Diffusion and Organizational Learning: The Case of Business Computing. *Organization Science*. **3**(1) 1-19.

Aubert, B. 2007. "Customer Education: Definition, Measures and Effects on Customer Satisfaction." University of Newcastle.

Azoulay, P., J. S. Graff Zivin, B. N. Sampat. 2011. "The Difusion of Scientifc Knowledge Across Time and Space:Evidence from Professional Transitions for the Superstars of Medicine." National Bureau of Economic Research.

Azoulay, P., J. S. Graff Zivin, J. Wang. 2010. Superstar Extinction. *Quarterly Journal of Economics*. **25** 549-589.

Bell, S. J., A. B. Eisingerich. 2007. The Paradox of Customer Education: Customer Expertise and Loyalty in the Financial Services Industry. *European Journal of Marketing*. **41**(5/6) 466-486.

Bettencourt, L. A., A. L. Ostrom, S. W. Brown, R. I. Roundtree. 2002. Client Co-Production in Knowledge-Intensive Business Services. *California Management Review*. **44**(4) 100-128.

Bhattacherjee, A. 2001. Understanding Information Systems Continuance: An Expectation-Confirmation Model. *MIS Quarterly*. **25**(3) 351-370.

Bhulai, S., G. Koole. 2003. A Queueing Model for Call Blending in Call Centers. *Automatic Control, IEEE Transactions on*. **48**(8) 1434-1438.

Bitner, M. J., W. T. Faranda, A. R. Hubbert, V. A. Zeithaml. 1997. Customer Contributions and Roles in Service Delivery. *International Journal of Service Industry Management*. **8**(3) 193-205.

Blackwell, M., S. Iacus, G. King, G. Porro. 2010. cem: Coarsened Exact Matching in Stata. *Stata Journal*. **9**(4) 524-546.

Blundell, R., S. Bond. 1998. Initial Conditions and Moment Restrictions in Dynamic Panel Data Models. *Journal of Econometrics*. **87**(1) 115-143.

Boone, T., R. Ganeshan, R. L. Hicks. 2008. Learning and Knowledge Depreciation in Professional Services. *Management Science*. **54**(7) 1231-1236.

Breslow, N. 1970. A Generalized Kruskal-Wallis Test for Comparing K Samples Subject to Unequal Patterns of Censorship. *Biometrika*. **57**(3) 579-594.

Bresnahan, T., M. Trajtenberg. 1995. General Purpose Technologies: 'Engines of growth?'. *Journal of Econometrics*. **65**(83-108).

Brynjolfsson, E., L. M. Hitt. 1995. Information Technology as a Factor of Production: The Role of Differences Among Firms. *Economics of Innovation and New Technology*. **3**(4) 183-200.

Brynjolfsson, E., L. M. Hitt. 1996. Paradox lost? Firm-level Evidence on the Returns to Information Systems Spending. *Management Science*. **42**(4) 541-558.

Brynjolfsson, E., P. Hofmann, J. Jordan. 2010. Cloud Computing and Electricity: Beyond the Utility Model. *Communications of the ACM*. **53**(5) 32-34.

Buell, R. W., D. Campbell, F. X. Frei. 2010. Are Self-Service Customers Satisfied or Stuck? *Production and Operations Management*. **19**(6) 679-697.

Burton, D. 2002. Consumer Education and Service Quality: Conceptual Issues and Practical Implications. *Journal of Services Marketing*. **16**(2) 125-142.

Cameron, A. C., P. K. Trivedi. 2010. *Microeconometrics Using Stata, Revised Edition*. Stata Press, College Station, TX.

Casalicchio, E., M. Colajanni. 2000. Scalable Web Clusters with Static and Dynamic Contents, in: *IEEE International Conference on Cluster Computing, 2000.* pp. 170-177.

Challagalla, G., R. Venkatesh, A. K. Kohli. 2009. Proactive Postsales Service: When and Why Does It Pay Off? *Journal of Marketing*. **73**(2) 70-87.

Chase, R. B. 1978. Where Does the Customer Fit in a Service Operation? *Harvard Business Review*. **56**(6) 137.

Chen, H., P. De, J. Hu. 2013. IT-Enabled Broadcasting in Social Media: An Empirical Study of Artists' Activities and Music Sales. Available at SSRN: http://ssrn.com/abstract=2201430

Cherkasova, L. 2000. FLEX: Load Balancing and Management Strategy for Scalable Web Hosting Service, in: *Fifth IEEE Symposium on Computers and Communications (ISCC 2000)*. Antibes, France: pp. 8-8.

Cleves, M. A., W. W. Gould, R. G. Gutierrez. 2010. *An Introduction to Survival Analysis Using Stata*. Stata Press, College Station, TX.

Cohen, W. M., D. A. Levinthal. 1990. Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly*. **35**(1) 128-152.

Cox, D. R. 1972. Regression Models and Life-tables (with discussion). *Journal of the Royal Statistical Society, Series B*(30) 187-220.

Das, A. 2003. Knowledge and Productivity in Technical Support Work. *Management Science*. **49**(4) 416-431.

Eisingerich, A. B., S. J. Bell. 2008. Perceived Service Quality and Customer Trust: Does Enhancing Customers' Service Knowledge Matter? *Journal of Service Research*. **10**(3) 256-268.

Fichman, R. 2000. The Diffusion and Assimilation of Information Technology Innovations, in: *Framing the domains of IT management : projecting the future--through the past*, R.W. Zmud (ed.). Cincinnati, Ohio: Pinnaflex Education Resources, Inc.

Fichman, R. G. 2001. The Role of Aggregation in the Measurement of IT-Related Organizational Innovation. *MIS Quarterly*. **25**(4) 427-455.

Fichman, R. G., C. F. Kemerer. 1999. The Illusory Diffusion of Innovation: An Examination of Assimilation Gaps. *Information Systems Research*. **10**(3) 255-275.

Field, J. M., M. Xue, L. M. Hitt. 2012. Learning by Customers as Co-producers in Financial Services: An Empirical Study of the Effects of Learning Channels and Customer Characteristics. *Operations Management Research*. **5**(1-2) 43-56.

Fodness, D., B. E. Pitegoff, E. T. Sautter. 1993. From Customer to Competitor: Consumer Cooption in the Service Sector. *Journal of Services Marketing*. **7**(3) 18-25.

Forman, C. 2005. The Corporate Digital Divide: Determinants of Internet Adoption. *Management Science*. **51**(4) 641-654.

Frei, F. X. 2006. Breaking the Trade-off Between Efficiency and Service. *Harvard Business Review*. **84**(11) 92.

Frei, F. X. 2008. The Four Things a Service Business Must Get Right. *Harvard Business Review*. **86**(4) 70-80.

Furman, J. L., K. Jensen, F. Murray. 2012. Governing Knowledge in the Scientific Community: Exploring the Role of Retractions in Biomedicine. *Research Policy*. **41** 276-290.

Gans, N., Y.-P. Zhou. 2003. A Call-Routing Problem with Service-Level Constraints. *Operations Research*. **51**(2) 255-271.

Garcia, D. F., G. Rodrigo, J. Entrialgo, J. Garcia, M. Garcia. 2008. Experimental Evaluation of Horizontal and Vertical Scalability of Cluster-based Application Servers for Transactional Workloads, in: *8th International Conference on Applied*

Informatics and Communications (AIC'08). Rhodes, Greece: World Scientific and Engineering Academy and Society (WSEAS), pp. 29-34.

Gehan, E. A. 1965. A Generalized Two-sample Wilcoxon Test for Doubly Censored Data. *Biometrika*. **52**(3/4) 650-653.

Ghose, A. 2009. Internet Exchanges for Used Goods: An Empirical Analysis of Trade Patterns and Adverse Selection. *MIS Quarterly*. **33**(2) 263-291.

Gurvich, I., M. Armony, C. Maglaras. 2009. Cross-Selling in a Call Center with a Heterogeneous Customer Population. *Operations Research*. **57**(2) 299-313.

Hall, B. H., J. Mairesse, L. Turner. 2007. Identifying Age, Cohort, and Period Effects in Scientific Research Productivity: Discussion and Illustration Using Simulated and Actual Data on French Physicists. *Economics of Innovation and New Technology*. **16**(2) 159-177.

Hansen, L. P. 1982. Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica*. **50**(4) 1029-1054.

Harms, R., M. Yamartino. 2010. The Economics of The Cloud. Microsoft. http://www.microsoft.com/en-us/news/presskits/cloud/docs/the-economics-of-the-cloud.pdf

Hitt, L. M., D. J. Wu, X. Zhou. 2002. Investment in Enterprise Resource Planning: Business Impact and Productivity Measures. *Journal of Management Information Systems*. **19**(1) 71-98.

Ho, D., K. Imai, G. King, E. Stuart. 2007. Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference. *Political Analysis*. **15**(3) 199-236.

Huang, M., X. Zhou. 2012. Research on Customer Education Paradox and the Mechanism, in: *Computational Intelligence and Design (ISCID), 2012 Fifth International Symposium on*. IEEE, pp. 355-358.

Iacus, S. M., G. King, G. Porro. 2012. Causal Inference without Balance Checking: Coarsened Exact Matching. *Political Analysis*. **20**(1) 1-24.

Jones, T. O., W. E. Sasser. 1995. Why Satisfied Customers Defect. *Harvard Business Review*. **73**(6) 88-91.

Kaplan, E. L., P. Meier. 1958. Nonparametric Estimation from Incomplete Observations. *Journal of the American Statistical Association*. **53**(282) 457-481.

Ko, D.-G., L. J. Kirsch, W. R. King. 2005. Antecedents of Knowledge Transfer from Consultants to Clients in Enterprise System Implementations. *MIS Quarterly*. **29**(1) 59-85.

Kumar, A., R. Telang. 2012. Does the Web Reduce Customer Service Cost? Empirical Evidence from a Call Center. *Information Systems Research*. **23**(3) 721-737.

Levenshtein, V. I. 1966. Binary Codes Capable of Correcting Deletions, Insertions, and Reversals. *Soviet Physics Doklady*. **10**(8) 707-710.

Li, Y., C.-H. Tan, H.-H. Teo, A. Siow. 2005. A Human Capital Perspective of Organizational Intention to Adopt Open Source Software, in: *Proceedings of the 26th International Conferences on Information Systems (ICIS 2005)*. Las Vegas, NV: pp. 137-149.

Liu, V., M. Khalifa. 2003. Determinants of Satisfaction at Different Adoption Stages of Internet-based Services. *Journal of the Association for Information Systems*. **4**(1) 12.

Long, J. S., J. Freese. 2006. *Regression Models for Categorical Dependent Variables Using Stata*. Stata Press, College Station, TX.

Mantel, N., W. Haenszel. 1959. Statistical Aspects of the Analysis of Data from Retrospective Studies of Disease. *Journal of the National Cancer Institute*. **22**(4) 719.

McKinney, V., K. Yoon, F. M. Zahedi. 2002. The Measurement of Web-Customer Satisfaction: An Expectation and Disconfirmation Approach. *Information Systems Research*. **13**(3) 296-315.

Mell, P., T. Grance. 2011. "The NIST Definition of Cloud Computing," National Institute of Standards and Technology Information Technology Laboratory (ed.). Gaithersburg, MD.

Menon, T., J. Pfeffer. 2003. Valuing Internal vs. External Knowledge: Explaining the Preference for Outsiders. *Management Science*. **49**(4) 497-513.

Michael, M., J. E. Moreira, D. Shiloach, R. W. Wisniewski. 2007. Scale-up x Scale-out: A Case Study using Nutch/Lucene, in: *Parallel and Distributed Processing Symposium (IPDPS 2007). IEEE International*. pp. 1-8.

Microsoft, Edge Strategies. 2011. SMB Cloud Adoption Study Dec 2010 - Global Report. www.edgestrategies.com.

Mills, P. K., J. H. Morris. 1986. Clients as "Partial" Employees of Service Organizations: Role Development in Client Participation. *Academy of Management Review* 726-735.

Morgan, L., P. Finnegan. 2007. How Perceptions of Open Source Software Influecne Adoption: An Exploratory Study, in: *Proceedings of the 15th European Conference on Information Systems (ECIS 2007)*. St. Gallen, Switzerland: pp. 973-984.

Muilenburg, L. Y., Z. L. Berge. 2005. Student Barriers to Online Learning: A Factor Analytic Study. *Distance Education*. **26**(1) 29-48.

Nayyar, P. R. 1990. Information Asymmetries: A Source of Competitive Advantage for Diversified Service Firms. *Strategic Management Journal*. **11**(7) 513-519.

Nickell, S. 1981. Biases in Dynamic Models with Fixed Effects. *Econometrica*. **49**(6) 1417-1426.

Oliver, R. L. 1980. A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions. *Journal of Marketing Research*. **17**(4) 460-469.

Reese, G. 2009. *Cloud Application Architectures: Building Applications and Infrastructure in the Cloud*. O'Reilly Media.

Retana, G., C. Forman, S. Narasimhan, M. F. Niculescu, D. J. Wu. 2013. Technolgy Support and IT Use: Evidence from the Cloud. Available at SSRN: http://ssrn.com/abstract=2165649.

Rogers, E. M. 1995. *Diffusion of Innovations*. The Free Press, New York.

Roodman, D. 2009a. How to Do xtabond2: An Introduction to Difference and System GMM in Stata. *Stata Journal*. **9**(1) 86-136.

Roodman, D. 2009b. A Note on the Theme of Too Many Instruments. *Oxford Bulletin of Economics & Statistics*. **71**(1) 135-158.

Schoenfeld, D. 1982. Partial Residuals for the Proportional Hazards Regression Model. *Biometrika*. **69**(1) 239-241.

SearchDataCenter.com. 2011. Data Center Decisions 2011 Survey Special Report. TechTarget.

Sharma, N., P. G. Patterson. 1999. The Impact of Communication Effectiveness and Service Quality on Relationship Commitment in Consumer, Professional Services. *Journal of Services Marketing*. **13**(2) 151-170.

Shinder, D. 2010. Mini-glossary: Cloud Computing Terms You Should Know, TechRepublic, http://www.techrepublic.com/blog/the-enterprise-cloud/mini-glossary-cloud-computing-terms-you-should-know/.

Singh, J., A. Agrawal. 2011. Recruiting for Ideas: How Firms Exploit the Prior Inventions of New Hires. *Management Science*. **57**(1) 129-150.

Staples, D. S., I. Wong, P. B. Seddon. 2002. Having Expectations of Information Systems Benefits that Match Received Benefits: Does it Really Matter? *Information & Management*. **40**(2) 115-131.

Symantec. 2011. State of the Cloud Survey. Symantec.

Tyler-Smith, K. 2006. Early Attrition among First Time eLearners: A Review of Factors that Contribute to Drop-out, Withdrawal and Non-completion Rates of Adult Learners undertaking eLearning Programmes. *Journal of Online Learning and Teaching*. **2**(2) 73-85.

Varian, H. R. 2010. Computer Mediated Transactions. *American Economic Review*. **100**(2) 1-10.

Varian, H. R. 2011. Micromultinationals Will Run the World, in: *Foreign Policy*. http://www.foreignpolicy.com/articles/2011/08/15/micromultinationals_will_run_the_world.

Venters, W., E. Whitley. 2012. A Critical Review of Cloud Computing: Researching Desires and Realities. *Journal of Information Technology*. **27**(3) 179-197.

Windmeijer, F. 2005. A Finite Sample Correction for the Variance of Linear Efficient Two-step GMM Estimators. *Journal of Econometrics*. **126**(1) 25-51.

Wittmann, A. 2012. Cloud Computing Is Still In Its Adolescence. Retrieved Aug 24, 2012, http://www.informationweek.com/global-cio/interviews/cloud-computing-is-still-in-its-adolesce/232600928.

Wooldridge, J. 2007. Control Function and Related Methods, in: *NBER Summer Institute - What's New in Econometrics?* http://www.nber.org/minicourse3.html.

Xue, M., P. T. Harker. 2002. Customer Efficiency. *Journal of Service Research*. **4**(4) 253-267.

Xue, M., L. M. Hitt, P.-Y. Chen. 2011. Determinants and Outcomes of Internet Banking Adoption. *Management Science*. **57**(2) 291-307.

Xue, M., L. M. Hitt, P. T. Harker. 2007. Customer Efficiency, Channel Usage, and Firm Performance in Retail Banking. *Manufacturing & Service Operations Management*. **9**(4) 535-558.

Zeithaml, V. A., L. L. Berry, A. Parasuraman. 1996. The Behavioral Consequences of Service Quality. *Journal of Marketing*. **60**(2) 31-46.

VITA

GERMAN F. RETANA

German Retana enrolled in the IT Management PhD program in 2008. Since early in the program he started learning about the emerging cloud computing industry. For his research he works directly with the business analytics team of a major public Infrastructure-as-a-Service (IaaS) provider. He employs econometric methods to study how the provider's technology support influences the way tens of thousands of customers make use of the on-demand cloud servers. He has presented his work at academic conferences including the International Conference on Information Systems (ICIS), the Workshop on Information Systems and Economics (WISE), and the INFORMS Annual Meeting, among others, and has also presented his research to industry executives of the cloud provider he is working with and to the Atlanta cloud community. One of his research articles is under second round revision at *Management Science*.

At Georgia Tech, he designed and successfully taught 3 times the Data Business Communications course, with ratings of 4.8, 4.91, and 4.84 (out of 5.0) in each occasion, and was honored with the 2013 CETL / BP Outstanding Graduate Instructor award. At the Scheller College of Business, he was one of the 2 students (out of 45) who were awarded the 2012 Ashford Watson Stalnaker Memorial Prize for Student Excellence in the PhD program, which considered his performance in both research and teaching.

Before coming to Atlanta, German headed a 60+ users SAP implementation in Costa Rica. He also worked in a regional systems integration project that relied on remote computing technologies. As part of his MBA internship, he collaborated in an MIT Media Lab project that involved wireless connectivity in a rural town.

German holds an MBA (with distinction) from INCAE Business School - Costa Rica, a Master's degree in Project Management from the University for International Collaboration - Costa Rica, and a BS degree (honor graduate) in Computer Engineering from the Costa Rican Institute of Technology. He was also an MBA exchange student for a semester at the Leipzig Graduate School of Management (HHL) - Germany. German is a Costa Rican with a strong passion for his country and who enjoys playing soccer and dancing salsa very much.