Technical Support and Service Demand: Evidence from the Cloud

Job Market Paper

German Retana
Scheller College of Business
Georgia Institute of Technology
german.retana@scheller.gatech.edu

Abstract

We study how a provider's technical support influences service demand in a business-to-business setting. The provider offers two levels of support, basic and full, reflecting its involvement with the buyer in co-producing the service. Using a unique nano-data set on public cloud infrastructure services consumption by 20,298 firms from March 2009 to April 2012, fixed effects panel data models, and a difference-in-difference identification strategy, we find that buyers who opt for full support consume 137% more IT capacity than those who do not. Moreover, even if they eventually switch from full to basic support, they continue to consume 105% more IT capacity, providing evidence of knowledge transfer. Further, purchase of support is complementary with buyer size: larger firms who adopt full support increase their demand more and retain knowledge better than smaller ones. Finally, we examine the underlying knowledge transfer mechanisms and estimate their economic value.

Keywords: Analytics, service, cloud computing, co-production, knowledge transfer.

Technical Support and Service Demand: Evidence from the Cloud

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 demand for the service. For example, 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 usage of the services (e.g., 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. To our knowledge, no prior work has demonstrated whether a provider's explicit attempts to improve KSAs through customer education and support can influence customer demand for a service.

We aim to take a first step in narrowing this gap in understanding by examining whether provider technical support influences demand for a particular commodity SST, namely public cloud infrastructure services.[1] Our central research questions are: *Does a provider's technical support influence demand for its service?* If so, *how?* We address these questions in the context of firm decisions to consume public cloud computing infrastructure services. Our focus on the business context allows us to 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). Further, we investigate how our results vary with firm size and technical sophistication. This allows us to advance existing literature on whether external sources of knowledge—such as service providers—are complements or substitutes to internal knowledge sources such as firm IT capabilities (see Argote and Miron-Spektor (2011) for a review). Finally, the procedures we have

[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).

developed have practical applications. They can be used by SST providers in a broad range of industries to analyze big data they collect on buyer behavior, enabling them to measure how a buyer's demand is influenced by its use of technical support.

In our research setting, the provider's customers consume IT server capacity and also choose (and switch between) two levels of technical 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 buyer's tradeoffs when choosing between basic/no support and full support, and the corresponding demand function. 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 20,298 firms that used the provider's public cloud infrastructure service at some point between March 2009 and April 2012. Our econometric approach uses fixed effects panel data models and a difference-in-difference identification strategy to compare buyers' demand for 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 consume an average of 137% more IT capacity relative to customers who have only had access to basic support, indicating that the economic value of the knowledge transferred to buyers through support is very significant. Furthermore, we find evidence that the effects of such knowledge are durable: former full support customers continue to consume, on average, 105% more IT capacity compared to those who have never used full support. To address concerns about how omitted variable bias in our models might influence our results (e.g., how changes in unobserved business needs may drive both the choice of using full support and changes in our dependent variable), we run our main model employing matched subsamples that are constructed using a coarsened exact matching (CEM) procedure (Blackwell et al. 2010). As an additional robustness check, 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. 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 adoption of full support. 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 demand for the service from adopting and having access to full support. They also retain more knowledge than smaller firms if they opt to switch to basic support.

Last, we provide further evidence on the mechanisms through which support use influences demand for cloud services by examining alternative measures of infrastructure use. Specifically, we provide additional evidence that technical support helps buyers make better and more efficient use of the service by quantifying the effects that full support has on buyers' likelihood of making use of advanced features such as the cloud's scalability.

In addition to contributing to the literature on service co-production, our work also contributes to extant work investigating whether internal and external knowledge sources are complements or substitutes to each other. In the context of R&D efforts, for long it has been suggested that there is a complementarity relationship between them (e.g., Arora and Gambardella 1994; Cassiman and Veugelers 2006; Cohen and Levinthal 1989; Cohen and Levinthal 1990). Nonetheless, much less is known regarding their relationship in the context of process innovations. While some recent work related to process innovations has found evidence of a substitution effect between internal and external knowledge sources (e.g., Forman et al. 2008; Vega-Jurado et al. 2009), Chang and Gurbaxani (2012) show that IT outsourcing customers' level of IT intensity positively affects the benefit they derive from the provider's knowledge, suggesting a complementarity relationship. We also find a complementarity relationship between the provider-sourced knowledge and a buyer's internal knowledge stocks.

Given the massive number of firms in our data, our study provides a framework than can be employed by SST providers to use their detailed and extensive records of customer actions to examine the impact of their managerial decisions on customer behavior. We have worked closely with the provider's business analytics team and used our model to offer guidance and rigorously quantify the influence of their premium technical support on customer demand and revenue. Moreover, we have also developed and automated the computation of a cloud-specific metric useful to assess customers' capabilities in exploiting the service's features, further allowing the provider to understand the impact of the knowledge they have transferred to their customers on their service co-production efforts.

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.

2.1 Motivating Model

We assume that there is a continuum of heterogeneous user firms (or buyers) with types $\theta \sim U(0,1)$. One can think of this type as the size or technical sophistication of the 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 purchase quantity ($q \geq 0$). The production function of each buyer firm takes the Cobb-Douglas form, being determined by its type (θ), quantity purchased (q), and co-production output elasticity, $0 < z_s < 1$, which is jointly determined by the provider and the buyer. While the use of the Cobb-Douglas form is standard in the service co-production literature (e.g., Xue et al. 2007), we adapt it to our B2B setting.

The fundamental assumption in our model is that buyers who opt for support enjoy a higher co-production output elasticity than those who opt for no support: formally, $z_f > z_b$. We argue this is the case since the provider-customer interactions that take place when full support is received enable knowledge transfers to occur. Similar knowledge transfers through interactions are present between consultants and clients in other B2B services contexts (Bettencourt et al. 2002; Ko et al. 2005). We follow Darr and Kurtzberg (2000) in arguing that knowledge "transfer has occurred when a contributor shares knowledge that is *used* by an adopter", and define knowledge transfer as "the communication of knowledge from a source so that it is learned and applied by a recipient" (Ko et al. 2005). Thus, the transferred knowledge directly impacts buyers' service co-production processes, which directly affects their output from input q.

For support to be an attractive option at least for some buyers, we further assume its price does not exceed the benefit it generates for the buyer such that $p_f \leq z_f$, and its price increase over no support does not exceed the benefit increase for the buyer such that $\frac{p_f - p_b}{p_b} \leq \frac{z_f - z_b}{z_b}$ (or $p_f \leq \frac{z_f}{z_b} p_b$).

Each buyer solves the following constrained optimization problem:

$$\max_{s,q} u_s(q, z_s|\theta) = \theta^{1-z_s} q^{z_s} - p_s q - 1_{\{s=f\}} F,$$

where $1_{\{s=f\}}$ is the indicator function that captures the fact that the two-part tariff scheme occurs only under support. Once the customer decides to adopt the service and chooses support level s, her optimal consumption level is given by

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

It follows that customers prefer support if their type is sufficiently large, 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 on customers who choose support, i.e., $\theta \geq \hat{\theta}$. The following hypotheses are directly motivated from our model (they all hold true under the set of conditions as identified in our assumptions above).

2.2 Technical Support and IT Services Demand

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. The hypothesis is that, given a buyer, her optimal service quantity will be greater if she opts for support rather than for no support. Formally,

HYPOTHESIS 1: $q^*(p_f, z_f|\theta) - q^*(p_b, z_b|\theta) \geq 0, \forall\, \theta \geq \hat{\theta}$.

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

2.3 Organizational Learning, Forgetting, and Durability of Knowledge Transferred

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 switch to basic support is if they have learned from the provider and are now able to achieve similar productivity levels 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. This is consistent with prior research that has shown that once firms internalize knowledge transferred 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. 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 have learned from the provider and do not quickly forget what they have learned, then, given the lower prices of basic support, they will exhibit greater service demand than other basic support customers who have not had the opportunity of learning from the provider. In other words, when buyers can achieve productivity levels equal or at least similar to those of full support (z_f) when consuming at basic support prices (p_b), they will demand more services than customers who have only accessed basic support. Formally,

HYPOTHESIS 2: $q^*\left(p_b, z_f|\theta\right) - q^*(p_b, z_b|\theta) \geq 0, \forall\, \theta \geq \hat{\theta}$.

2.4 Firm Size and Organizational Learning and Forgetting

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. If internal capabilities are substitutes for external ones, one might expect that large firms' marginal benefits from having access to the provider's technical support would be low, as they have little to gain. Nonetheless, such a view would overlook larger firms' greater ability to derive value from external knowledge sources.

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 obtain greater benefits from the knowledge transferred to them via technical support. Researchers studying knowledge transfers between providers and customers in the IT industry have shown that absorptive capacity facilitates knowledge transfer. For example, Ko et al. (2005) find that customers' absorptive capacity is positively associated with knowledge transfer between consultants and clients in the context of enterprise information systems implementation. Similarly, recent work by Chang and Gurbaxani (2012) finds that firms with stronger IT investments benefit more from their provider's knowledge spillovers in the context of IT outsourcing. We thus propose that the benefits of adopting and having access to full support should be stronger for larger rather than smaller firms. Specifically, the greater a buyer's size, the greater its service demand increase associated with the adoption of and continued access to full support:

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

Additionally, if Hypothesis 2 holds, whereby buyers do not forget what they have learned from the provider, and also per Hypothesis 3 larger firms are more efficient than smaller ones at transforming knowledge into output, then one might consequently expect that larger firms are able to retain more knowledge than smaller ones after they switch from full to basic support. In other words, again because of larger firms' greater levels of related knowledge and absorptive capacity, we expect that larger basic support buyers who have accessed full support will demand more services than smaller ones who have also accessed full support in the past. Formally,

HYPOTHESIS 4: $\frac{\partial\left(q^*(p_b,z_f|\theta)-q^*(p_b,z_b|\theta)\right)}{\partial\theta} > 0, \forall\ \theta \geq \hat{\theta}.$

3 Research Setting: 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). 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. 2009; 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). More recently, InformationWeek Reports 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, there is industry insight suggesting that knowledge barriers are very likely a key service co-production cost inhibiting adoption. 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 knowledge barriers make cloud infrastructure services an ideal context to test if and how technical support influences demand for services.

In our particular setting, the provider has recognized that the novelty of the service plus the complexities involved in scaling IT capacity may pose significant knowledge barriers to buyers. In response to this, the provider offers them the option to contract and access full support for a price premium on the IT capacity (i.e., servers) hourly rates plus a fixed monthly fee. A prime goal of full support is to educate buyers on how to best use the IaaS and adapt it to their idiosyncratic business needs. Thus, when receiving full support, customers have the opportunity to learn from the provider's prior experience in deploying applications in cloud architectures. 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 customers with issues concerning account management or the overall performance of the infrastructure service. Thus, basic support customers do not have access to external knowledge from the provider and have to rely only on their internal knowledge stocks to co-produce the service.

4 Empirical Model

We employ linear fixed effects panel data models along with a difference-in-difference identification strategy to tease out the effects of adoption of full support on cloud demand. The pay-per-use model provides cloud infrastructure buyers the freedom to purchase (consume) and pay only for the computing resources they need. Thus, we can use customers' *IT capacity demand* ($Capacity_{it}$) as a direct measure of their demand for the service, and use it as our dependent variable. We calculate it as average server computing capacity (i.e., CPUs and GB of RAM) per month. Given that the distribution of IT capacity consumption has a strong positive skew and that at times buyers may not consume any capacity, we use the log of capacity plus 1 ($lnCapacity_{it} = ln(Capacity_{it} + 1)$) as our dependent variable.

Our first model tests if the usage or the prior usage of full support is associated with greater IT capacity demand:

$$lnCapacity_{it} = \alpha + \beta\ FullSupport_{it} + \gamma\ SwitchToBasic_{it} + \mu_i + \tau_t + l_{it} + \varepsilon_{it}. \tag{1}$$

Subscripts i and t index individual buyers (firms) and time periods (months) respectively. Parameter μ_i is the customer fixed effect and τ_t is a vector of month fixed effects. We also include a dummy variable, 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 customers' IT capacity may increase in a nonlinear fashion as they grow older and learn more about the service. Parameter ε_{it} is our error term which we assume is correlated only within individuals, 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 demand 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 customer 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 buyers retain the knowledge acquired during the time they used full support then γ will be insignificant (suggesting behavior does not change) or negative and significant but with a very low value relative to β (suggesting the effects of full support do not dissipate entirely). $\beta + \gamma$ will measure differences in behavior between basic support customers who accessed full support in the past and those who exclusively accessed basic support. If Hypothesis 2 holds, and customers' prior access to full support sets them apart from those who only used basic support, then $\beta + \gamma$ should be positive and significant.

We model buyer size 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 customer is above the median employment ($EmploymentTop50_i$), (2) a binary variable indicating if the customer 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:

$$\begin{aligned} lnCapacity_{it} = \alpha &+ \beta_1\ FullSupport_{it} + \beta_2\ FullSupport_{it} \times EmploymentTop50_i \\ &+ \gamma_1\ SwitchToBasic_{it} + \gamma_2\ SwitchToBasic_{it} \times EmploymentTop50_i \\ &+ \mu_i + \tau_t + l_{it} + \varepsilon_{it} \end{aligned} \tag{2}$$

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 customers are able to keep their service co-production costs lower than smaller customers after the switch to basic support, then $\beta_2 + \gamma_2$ should be positive and significant as well.

Our fixed effects model allows us to difference out unobserved time-invariant buyer-level heterogeneity that may influence both the demand for support and for IT capacity. Like any difference-in-difference model, our estimates rely on the identifying assumption that unobserved factors influencing IT capacity demand change similarly for treatment and control groups over time. While in our baseline model we use all firms who have not yet adopted full support as the control group, to explore the validity of this assumption we also run 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 intended use cases for the cloud service, industry, size, and service demand in the early periods of the interactions with the service provider.

We further address the omitted variable bias concerns by employing dynamic panel data models using GMM estimation procedures (Arellano and Bover 1995; Blundell and Bond 1998). This approach allows us to overcome the limited availability of data that could be used to construct instruments for buyer-specific time-varying factors by using lagged values of the dependent variable and the endogenous regressors (and the lags of their first differences) as instruments. Additionally, the dynamic panel framework, where lagged values of the dependent variable are used as regressors, allows us to control for the fact that buyers IT capacity demands in prior periods may strongly influence their demand in the focal period.

5 Data

We have collected a unique data set on cloud infrastructure services consumption from a major public cloud provider. Our entire data set includes 68,107 customers that adopted the provider's services at some point between March 2009 and April 2012. Customers 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 customers who consume very low levels of IT capacity or who do not change their architecture configuration. These are customers who are using the cloud as a low cost (fixed capacity) web hosting service (e.g., to host a small personal blog), and for whom the adoption of full support would have no effect since they have no intentions of growing their IT capacity demand. A similar intuition was already derived from our motivating model, where the lower-type customers $(\theta < \hat{\theta})$ would not opt for full support.[2] 47,809 customers are dropped from the sample as a result of this procedure, though our results are robust to their inclusion (these results are available upon request).

Among the remaining 20,298 customers in our baseline sample, 15,225 relied exclusively on basic support, 3,816 relied exclusively on full support, and 1,257 accessed both types of support during their observed lifetimes. The sample includes 293,661 customer-month observations. Table 1 provides descriptive statistics of the time-varying parameters in our sample.

Table 1. Descriptive Statistics of Time-Varying Variables (Baseline sample, N=293,661)

Variable	Description	Mean	Std. Dev.	Min	Max
$Capacity_{it}$	Average GB RAM /hour	7.49	28.63	0.00	1,653.00
$lnCapacity_{it}$	$ln(Capacity_{it} + 1)$	1.36	1.00	0.00	7.41
$FullSupport_{it}$	Indicator	0.13	0.34	0	1
$SwitchToBasic_{it}$	Indicator	0.02	0.15	0	1

We capture IT capacity demand, $Capacity_{it}$, as a customer i's average hourly consumption during month t measured in GB RAM/hour.[3] Our measure reflects the standard way cloud infrastructure providers price their services: hourly rates that increase in server capacity.

In addition to time-varying data on each buyer's usage of the cloud service, we have also collected data from an optional survey administered to buyers upon signup. The survey was first administered in June 2010, and we have all buyers' responses until February 2012. We use 3 of the items in the survey: their firms' total employment, their intended use case for the cloud infrastructure service, and their industry. We describe the employment data here, and defer description of the latter two items for when we discuss the development of the matched subsamples. 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 customer fixed effect.

In terms of employment, since a customer could have responded to the survey more than once, we only consider buyers that either have a single survey response (that is not "I don't know") or that have consistent responses across all their submissions. This leaves us with unique responses for 5,277 customers in the baseline sample. The survey asks customers to indicate their range of employment; we

[2] Specifically, we exclude customers 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) had no scaling activity during their first 6 months (excluding 1st month). A scaling activity is 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.

[3] For example, $Capacity_{i3} = 2.25$ means that customer i used, on average, a set of servers with a combined capacity of 2.25 GB RAM during month 3. This could be two 1 GB RAM servers and a 256 MB RAM server, or 9 servers, each with 256 MB of RAM.

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 are shown in Table 2.

Table 2. Descriptive Statistics of Employee Count-related Variables (N=5,277)

Variable	Description	Mean	Std. Dev.	Min	Max
$Employees_i$	Number of employees	204.81	1,127.64	2	10,000
$EmploymentTop50_i$	Indicator = 1 if above median, 0 otherwise	0.56	0.50	0	1
$EmploymentTop25_i$	Indicator = 1 if in top 25th percentile, 0 otherwise.	0.22	0.42	0	1
$lnEmployment_i$	$ln(Employees_i)$	2.43	1.73	1.10	9.21

6 Results

6.1 Effects of Adoption of Full Support on Capacity

We present the results for Model (1) in Column (1) of Table 3. Consistent with Hypothesis 1, the results indicate that customers who adopt and use full support consume, on average, 137% (i.e., $e^{0.8622} - 1$) more IT capacity than customers who use basic support. Also, the test that the sum of the coefficients for $FullSupport_{it}$ and $SwitchToBasic_{\ it}$ in Column (1) of Table 3 are different from zero is statistically significant at the 1% level, indicating that even after buyers have switched from full to basic support, they continue consuming, on average, 105% (i.e., $e^{0.8622-0.1435} - 1$) more capacity than customers who never accessed full support. This result provides support for our second hypothesis that suggests that customers retain and use the knowledge acquired during the time they accessed full support, and the positive effects of technical support on IT capacity demand 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)[4] as follows. While a buyer consuming IT capacity at the median of the distribution generates an ARPU of $64.60, buyers who opt for full support generate an ARPU of $153.00. Moreover, buyers who switch to basic support continue contributing an ARPU of $132.55. Considering the tens of thousands of firms using cloud infrastructure services, offering full support to buyers has significant revenue implications for the provider.

[4] 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 $Capacity_{it}$, and set the market price (p_b) of 1 GB RAM/hour of IT capacity at $0.045. The median $Capacity_{it}$ in our data is 2 GB RAM/hour, and we multiply it by 720 to get a median monthly IT capacity usage of 1,440 GB RAM hours. We do not use the mean $Capacity_{it}$ 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.

Table 3. Baseline Results for Tests of Hypotheses 1 and 2

Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample	Baseline	CEM1	CEM2	CEM3	Baseline			
Model	Basic Model				Basic Model with Falsification Tests		Dynamic Panel	
							FE	GMM
$FullSupport_{it}$	0.8622***	0.9712***	0.8351***	0.8835***	0.9032***	0.9282***	0.2898***	0.0381***
	(0.0337)	(0.0642)	(0.0777)	(0.0726)	(0.0374)	(0.0405)	(0.0046)	(0.0099)
$SwitchToBasic_{it}$	-0.1435***	-0.1782***	-0.2051***	-0.2319***	-0.1442***	-0.1446***	-0.0536***	-0.0080
	(0.0149)	(0.0338)	(0.0507)	(0.0809)	(0.0149)	(0.0149)	(0.0045)	(0.0060)
$AdoptFullIn2_{it}$					0.1496***			
					(0.0255)			
$AdoptFullIn4_{it}$						0.1416***		
						(0.0281)		
$lnCapacity_{it-1}$							0.9432***	1.0415***
							(0.0021)	(0.0331)
$lnCapacity_{it-2}$							-0.1834***	-0.1172***
							(0.0027)	(0.0376)
$lnCapacity_{it-3}$							0.0318***	0.0616***
							(0.0019)	(0.0071)
Constant	-3.1271***	-2.7667***	-3.0337***	-1.8308***	-3.1369***	-3.1376***	0.2221***	-0.0226
	(0.0739)	(0.3295)	(0.4442)	(0.4412)	(0.0739)	(0.0740)	(0.0198)	(0.0246)
N	293,661	45,246	15,065	24,745	293,661	293,661	234,089	234,089
R^2	0.239	0.300	0.342	0.348	0.240	0.240	0.747	
Customers	20,298	3,434	1,188	1,563	20,298	20,298	18,488	18,488

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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The GMM estimation in column (8) considers $FullSupport_{it}$ and $SwitchToBasic_{it}$ as endogenous. Given AR(2) in the errors, it uses the 3rd lag of $lnCapacity_{it}$, $FullSupport_{it}$ and $SwitchToBasic_{it}$ as their instruments for the difference equation, and uses the 2nd lag of the 3 variables' first difference as instruments for the levels equation. Total number of instruments is 216. Hansen (1982) specification test passed with $\chi^2(142) = 153.64,\ p = 0.238$. Robust standard errors use Windmeijer's (2005) finite sample correction.

We also run Model (1) on matched subsamples defined using a coarsened exact matching (CEM) procedure (Blackwell et al. 2010; Iacus et al. 2012). We consider customers 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 IT capacity 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 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). 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 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 data panel model to control for them.

We match customers based on 4 attributes: (1) their level of IT capacity usage during their first 6 months (excluding the 1st month), (2) their level of scaling activity (i.e., how often they adjust their level of IT capacity usage) during their first 6 months (excluding the 1st month), (3) their intended use case for the cloud infrastructure service, and (4) their employment. The first 3 characterize the firms' business needs and service demand characteristics, while employment is associated with their IT capabilities. We discuss each in turn.

To create our baseline sample, we had already discarded basic support users with very small and rather static deployments. Nonetheless, even among the remaining customers there is considerable variation in IT capacity usage and scaling activity. In particular, the presence of strong outliers in terms of average IT capacity usage and scaling activity at the higher end of the spectrum prevents us from relying on automated coarsening algorithms that attempt to create equally-spaced cutoff points; doing so would group customers into few strata while leave many strata empty. Given this feature of our data, we employ our own manual cutoff points that make better use of the variation at the lower end of the distribution. For average IT capacity usage, we set our cutoff points at standard server sizes: 1GB, 2GB, 4GB, 8GB, 16GB, and 32GB. For scaling activity 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 8 changes, and the 95th percentile is 40 changes.

The intended use case and the employment data are both collected from the signup survey administered by the provider to its customers. 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 customers 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 IT capacity demand for the use case is predictable or not. Our first general use case, which we call "Back Office Applications", includes use cases associated with 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. Our second general use case, "Low Demand Uncertainty", includes customer-facing websites used for regular operation of the firm that have steady or at least predictable IT capacity demands. 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 "High Demand Uncertainty" general use case includes customer-facing websites that are prone to unpredictable variance in their IT capacity demand. 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 SaaS offerings. We additionally consider web hosting services and running test and development environments as additional general use cases. Altogether, we have 5 general use cases.

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, 65% indicated they have 10 or fewer employees, so we use 10 as our first cutoff point. Another 20% indicated they have between 11 and 50 employees, 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, from 101 to 250, and greater than 250.

The above exercise leaves us with a sample of only 5,040 customers, of which 1,213 are treated and 3,827 are potential controls. Through the CEM process we are able to match 1,016 treated buyers to 2,418 controls, forming a subsample of 3,434 matched customers. We denote this sample as CEM1 and present the results of Model (1) using this subsample in column (2) of Table 3. The results are not only consistent with our prior results using the baseline sample in column (1), but actually are slightly more economically significant. The coefficients suggest that the adoption and continued access to full support increases customers' IT capacity usage by 164% (i.e., $e^{0.9712} - 1$) and that, after switching to basic support, former full support customers still consume 121% (i.e., $e^{0.9712-0.1782} - 1$) more IT capacity compared to customers who never used full support. As before, the test that the sum of the coefficients for $FullSupport_{it}$ and $SwitchToBasic_{it}$ is different from zero is statistically significant at the 1% level.

As an additional check, we now incorporate an additional question on customer industries in the survey to make an even more stringent match of treated customers 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. Among the 5,040 customers in our CEM sample, 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%), Software-as-a-Service (SaaS) (5.32%), advertising (5.56%), and entertainment (3.75%). With this additional matching criterion, the CEM process matches 479 treated buyers to 709 controls, forming a subsample with 1,188 buyers that we denote as CEM2. The results with this new subsample are presented in column (5) of Table 3 and are consistent with prior results.

As a final matched sample analysis, we only consider treated customers 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 that implemented by Azoulay et al. (2010) and by Singh and Agrawal (2011), ensures treated firms do not exhibit differential demand for the service before they adopt full support relative to controls. We use as matching criteria the buyers' average IT capacity usage and their level of scaling activity from the 2nd month in their lifetimes (because of the abnormal setup-related behavior in their 1st months) until the month right before they adopted full support (again had the controls adopted full support at the same time as the treated firms). For simplicity and consistency, we employ the same coarsening cutoff points for IT capacity and scaling activity described above. We additionally consider the buyers' intended general use cases and employment as before. This final CEM procedure finds 1,331 appropriate controls for 205 upgrading (treated) customers, forming a subsample of 1,536 customers which we denote as CEM3. The results of Model (1) using the CEM3 subsample are presented in column (4) of Table 3, and are similar to those attained before.

6.2 Robustness Checks

6.2.1 Falsification Tests

Our use of the matching procedures described above increases confidence in our identifying assumption that there do not exist unobserved time-varying factors that differentially affect capacity demand for 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. We first perform a falsification test to verify if there is any significant change in customer behavior in the months immediately preceding the adoption of full support. We examine whether customers' behavior before the adoption of full support is similar among customers 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 3. We find that customers tend to consume between 15% (i.e., $e^{0.1416}-1$) and 16% (i.e., $e^{0.1496}-1$) more IT capacity 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 capacity demand and full support, such as a previously planned increase in capacity. However, to address these concerns further, in the next section we conduct a dynamic panel data analysis.

6.2.2 Dynamic Panel Estimation and Endogenous Adoption and Switching Decisions

Another concern is that there may be persistence in the demand for IT capacity. We first address this concern by including lagged values of the dependent variable as a regressor and estimate the model using standard fixed effects. The model has the following form:

$$lnCapacity_{it} = \alpha + \sum_{n=1}^{p} \lambda_n lnCapacity_{it-n} + \beta\ FullSupport_{it} + \gamma\ SwitchToBasic_{it} + \mu_i + \tau_t + l_{it} + \varepsilon_{it} \tag{3}$$

We ran this fixed effects model using varying number of lags for the dependent variable (i.e., $p = 1, \ldots, 5$), and attained qualitatively similar results. For reasons that will be explained below, we only present the results using 3 lags ($p = 3$) of $lnCapacity_{it}$ in column (7) of Table 3. We confirm our suspicion that the current value of the variable is strongly influenced by its past values; this is reflected in the large size and statistical significance for the lagged dependent variables. Nonetheless, even after controlling for this, we find that the IT capacity usage still increases 34% (i.e., $e^{0.2898}-1$) with the adoption of and continued access to full support ($FullSupport_{it}$). Furthermore, with respect to the effects of switching from full to basic support ($SwitchToBasic_{it}$), we also find that IT capacity usage only decreases marginally. Former full support customers continue consuming 27% (i.e., $e^{0.2898-0.0536}-1$) more capacity compared to basic support customers who never used full support; the test that the sum of the corresponding coefficients is different from zero is statistically significant at the 1% level. While the magnitudes of the coefficients in column (7) are much lower than those in our basic model in column (1) of Table 3, the signs and statistical significance of the parameters of interest continue to hold and support Hypotheses 1 and 2.

The fixed effect Model (3) will suffer 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 $T = 38$, 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 added benefit that it allows us to treat $FullSupport_{it}$ and $SwitchToBasic_{it}$ as endogenous and use their 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).

We first select the appropriate number of lags of the dependent variable to be included as regressors (p). 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 $lnCapacity_{it}$ 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 2nd lag 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 assume our errors follow a MA(1) process in which $\varepsilon_{it} = \eta_{it} + \rho\eta_{it-1}$, where $|\rho| < 1$ and $E[\eta_{it}] = 0$. Given this, we cannot use the 2nd lag of the variables as instruments for the difference equation (since $E[lnCapacity_{it-2}\Delta\varepsilon_{it}] \neq 0$, where $\Delta\varepsilon_{it}$ is the first difference of ε_{it}) nor the 1st lag of their first difference as instruments for the levels equation (since $E[\Delta lnCapacity_{it-1}\varepsilon_{it}] \neq 0$, where $\Delta lnCapacity_{it}$ is the first difference of the $lnCapacity_{it}$). Nonetheless, we can still rely on the variables' 3rd and posterior lags and on the 2nd lag of their first difference as instruments (Cameron and Trivedi 2010). That is, we make the identification assumptions that $E[lnCapacity_{it-3}\Delta\varepsilon_{it}] = E[\Delta lnCapacity_{it-2}\varepsilon_{it}] = 0$. After this consideration, our model becomes

$$\begin{aligned} lnCapacity_{it} = \alpha + \sum_{n=1}^{p} \lambda_n lnCapacity_{it-n} + \beta\ FullSupport_{it} + \gamma\ SwitchToBasic_{it} \\ + \mu_i + \tau_t + l_{it} + \eta_{it} + \rho\eta_{it-1}. \end{aligned} \tag{4}$$

We ran model (4) using all available 1,135 instruments (from the 3rd lag to the end of the panel). We confirmed we still have 2nd order serial correlation but do not have 3rd or posterior order serial correlation. We then used the Hansen (1982) J test to test the validity of our overidentifying restrictions. We passed the Hansen J test with $\chi^2(1062) = 1033.26,\ p = 0.731$. We then checked if we could use fewer lags of $lnCapacity_{it}$ as regressors, and found that we could still pass the serial correlation test (with no 3rd or posterior order serial correlation) and the Hansen J test (with $\chi^2(1068) = 1061.84, p = 0.547$) if we used only 3 lags.

The second estimation decision to make is selecting an appropriate number of instruments 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 all 3 endogenous variables ($lnCapacity_{it}$, $FullSupport_{it}$ and $SwitchToBasic_{it}$) as instruments for the difference equation. We also use the 2nd lag of the 3 variables' first difference as instruments for the levels equation. When using such specification, we pass the Hansen J test with $\chi^2(142) = 153.64,\ p = 0.238$. Overall, we reduced the total number of instruments from 1,142 to 216.

The results for this System GMM specification are reported in column (8) of Table 3. As in column (7), we find that the values of $lnCapacity_{it}$ are strongly influenced by its prior values. In this model, while the magnitude of the coefficient from $FullSupport_{it}$ is reduced, it is still positive and significant (p-value = 0.000), which matches what was expected per Hypothesis 1. The coefficient of $SwitchToBasic_{it}$ is no longer significant (p-value = 0.182), which is different from our prior results. Nonetheless, rather than contradicting Hypothesis 2, this result provides stronger support for the hypothesis that basic support customers who accessed full support in the past demand more IT capacity than those who never accessed full support.

6.3 How do Firm Size and Full Support Interact to Shape Demand?

Our results that test Hypotheses 3 and 4 are presented in Table 4. Column (1) of Table 4 presents coefficients nearly identical to those in Column (1) of Table 3, suggesting the subsample of customers who responded to the survey is not distinct from our broader baseline sample. Next, we examine the interaction effects. The interpretation of the coefficients of the interactions with the dummy variables $EmploymentTop50_i$ and $EmploymentTop25_i$ is straightforward, but in order to better understand the effects of the interactions with $lnEmployment_{\mathrm{i}}$, we evaluate the percentage changes in IT capacity ($lnCapacity_{it}$) due to turning the support choice dummies on at different levels of $lnEmployment_i$, which we present in Table 5. While all models use the complete baseline sample used in Table 3, as a robustness check we have also re-estimated Model (2) after applying the CEM procedure on this subsample of survey respondents, with consistent findings.

The results provide strong support for both hypotheses. The tests that the sums of all corresponding coefficients in the analysis here are different from zero are statistically significant at the 1% level. We base our first analysis on columns (2) and (3) of Table 4. We find that while large firms increase their IT capacity demands between 174% (i.e., $e^{0.5541+0.4547} - 1$) and 197% (i.e., $e^{0.7798+0.3074} - 1$) upon adoption of full support, smaller ones only increase their demand between 74% (i.e., $e^{0.5541} - 1$) and 118% (i.e., $e^{0.7798} - 1$). Using our continuous measure of employment, columns (1) and (3) of Table 5 also show that larger firms increase their capacity by 161% while smaller ones only do so by 98%. In short, the coefficient estimates in Table 4 show strong support for Hypothesis 3.

Similarly, again based on Table 4, while large former full support customers continue consuming between 145% and 180% after they have switched to basic support relative to pure basic support users, small firms only consume between 37% and 79% more. Results in Table 5 show a similar difference between large and small firms using continuous measures of employment. Together, the results provide strong support for Hypothesis 4.

Using the results in Column (2) of Table 4, 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 IT capacity consumption level (see footnote 4). We estimate that while large firms increase their ARPU from \$64.60 to \$177.16, small firms only increase their ARPU to \$112.43. Similarly, regarding their ARPU after they switch to basic support, we estimate that large firms have an ARPU of \$158.15, which is very high relative to that of small firms who switch, \$88.65, and to that of the median buyer, \$64.60.

In sum, we find strong evidence of complementarity effects between firm size and the external knowledge made available to them by the provider.

Table 4. Are the results of full support stronger for large firms?

Column	(1)	(2)	(3)	(4)
$FullSupport_{it}$	0.8825***	0.5541***	0.7798***	0.6281***
	(0.0540)	(0.0767)	(0.0588)	(0.0878)
$SwitchToBasic_{it}$	-0.1558***	-0.2376***	-0.1968***	-0.2598***
	(0.0279)	(0.0462)	(0.0332)	(0.0454)
$FullSupport_{it} \times EmploymentTop50_i$		0.4547***		
		(0.1012)		
$SwitchToBasic_{it} \times EmploymentTop50_i$		0.1241**		
		(0.0519)		
$FullSupport_{it} \times EmploymentTop25_i$			0.3074**	
			(0.1296)	
$SwitchToBasic_{it} \times EmploymentTop25_i$			0.1408***	
			(0.0459)	
$FullSupport_{it} \times lnEmployment_i$				0.0796***
				(0.0270)
$SwitchToBasic_{it} \times lnEmployment_i$				0.0378***
				(0.0116)
Constant	-3.5448***	-3.5434***	-3.5501***	-3.5515***
	(0.3602)	(0.3565)	(0.3580)	(0.3567)
N	71,478	71,478	71478	7,1478
R^2	0.293	0.297	0.295	0.296
Customers	5,277	5,277	5,277	5,277

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 5. Changes in IT Capacity Demand from Adopting and Switching from Full Support at Different Levels of Firm Size

Column	(1)	(2)	(3)
Fixed level of $lnEmployment_i$ [b]	$\bar{x} - \sigma^2$	$\bar{x}$	$\bar{x} + \sigma^2$
$FullSupport_{it} = 1$	98.22	127.40	160.87
$SwitchToBasic_{it} = 1$	-20.86	-20.45	-20.03
$FullSupport_{it} = SwitchToBasic_{it} = 1$	56.87	80.90	108.61

[a] Since $Capacity_{it} = e^{lnCapacity_{it}} - 1$, per model (3), the percentage change in $Capacity_{it}$ is given by $e^{\beta_1 FullSupport_{it} + \gamma_1 SwitchToBasic_{it} + lnEmployees_i(\beta_2 FullSupport_{it} + \gamma_2 SwitchToBasic_{it})} - 1$.

[b] $\bar{x}$ and σ^2 denote the sample mean and standard deviation of $lnEmployment_i$, respectively. Results are based on column (4) of Table 4 and are expressed in percent change.

7 Exploration of Underlying Knowledge Transfer Mechanisms

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 the knowledge transferred to them. In this section we test if, in accordance with this assumption, customers make better use of the service 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 customers' deployments. One such proxy is the complexity of a buyer's deployments that serves to assess how proficient a customer is in making use of the service. In general, if full support helps customers co-produce better service outputs, then we should expect that full support customers 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. 2009; Harms and Yamartino 2010), doing so requires firms to scale their capacity 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. Users 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). Having more servers, and thus more moving parts, will increase the complexity of the architecture and at the same time signal a better use of the service. As a result of these increased efficiencies and complexity, we use horizontal scaling as a measure that proxies for a customer's skill at using cloud computing; in the context of our model, evidence of increased horizontal scaling under full support provides additional evidence of our assumption that $z_f > z_b$.

Our analysis of the complexity of customer deployments is based on an automated analysis of the names given by customers to their servers. We develop an algorithm that compares the names of the servers being run by each customer at the end of every day during our sample and check if we find servers with names very similar to each other.[5] Our assumption is that we can identify a horizontally scalable architecture (i.e., an architecture with high complexity) if we find two or more servers with very similar names, given that they will very likely be performing the same function in parallel (e.g., `web1.domain.com` and `web2.domain.com`). We measure the presence of any such pair of servers with similar names during month t in customer i's infrastructure with a new indicator variable, $Horizontal_{it}$, which is equal to 1if similar server names are found and is 0 otherwise. The indicator has a mean value of 0.23 and a standard deviation of 0.42.

We run the exact same regressions used for Models (1) and (2) but substitute $Horizontal_{it}$ for $lnCapacity_{it}$ as the dependent variable. Overall, our results are consistent with respect to what we found before when using the IT service demand dependent variable, providing additional evidence that full support enables customers to use the cloud more effectively.

Columns (1) through (6) of Table 6 show that customers who have adopted and continue having access to full support are 19.74 to 21.78 percentage points more likely to use complex, horizontally scalable architectures than basic support users. Further, the results also indicate that former full support customers are only 2.71 to 4.40 percentage points less likely to employ highly complex architectures 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. A very important nuance of how the service is offered makes this result more meaningful: if buyers desire to continue running the same set of applications under the new support regime, they must redeploy their entire

[5] Specifically, we consider two server names to be similar to each other if they have a Levenshtein Distance (Levenshtein 1966) 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 letters or less.

infrastructure on their own. Therefore, if they continue using a horizontally scalable deployment after having switched to basic support, it must be the case that they set it up entirely on their own, and in general we find that basic support customers who accessed full support in the past continue employing configurations with a greater complexity than customers who exclusively relied on basic support. Together, these results are consistent with our model assumptions that consumers learn from the provider through full support.

Table 6. Results for Tests of Effects of Full Support on Architecture Complexity

Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample	Baseline	CEM1	CEM2	CEM3	Baseline			
Model	Basic Model				Basic Model with Falsification Tests		Dynamic Panel	
							FE	GMM
$FullSupport_{it}$	0.1974***	0.2168***	0.2009***	0.2049***	0.2093***	0.2178***	0.0993***	0.0168***
	(0.0138)	(0.0281)	(0.0366)	(0.0310)	(0.0153)	(0.0164)	(0.0036)	(0.0052)
$SwitchToBasic_{it}$	-0.0271***	-0.0240*	-0.0440**	-0.0479	-0.0273***	-0.0275***	-0.0157***	-0.0066*
	(0.0062)	(0.0135)	(0.0213)	(0.0362)	(0.0062)	(0.0062)	(0.0036)	(0.0036)
$AdoptFullIn2_{it}$					0.0434***			
					(0.0104)			
$AdoptFullIn4_{it}$						0.0436***		
						(0.0110)		
$Horizontal_{it-1}$							0.6034***	0.7577***
							(0.0022)	(0.0562)
$Horizontal_{it-2}$							0.0367***	0.0731
							(0.0025)	(0.0454)
$Horizontal_{it-3}$							0.0121***	0.0139
							(0.0024)	(0.0357)
$Horizontal_{it-4}$							-0.0078***	0.0671**
							(0.0020)	(0.0309)
Constant	-3.1271***	-0.5233***	-0.2408	-0.7011***	-0.7316***	-0.7320***	0.0826***	-0.0133
	(0.0739)	(0.1270)	(0.2479)	(0.1488)	(0.0433)	(0.0433)	(0.0153)	(0.0251)
N	293,661	45,246	15,065	24,745	293,661	293,661	215,601	215,601
R^2	0.239	0.058	0.087	0.069	0.028	0.028	0.424	
Customers	20,298	3,434	1,188	1,563	20,298	20,298	17,561	17,561

Dependent variable is $Horizontal_{it}$. All regressions include calendar (τ_t) and lifetime time dummies (l_{it}).Robust standard errors, clustered on customers, in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The GMM estimation in column (8) considers $FullSupport_{it}$ and $SwitchToBasic_{it}$ as endogenous. Given AR(2) in the errors, it uses the 3rd lag of $Horizontal_{it}$, the 3rd through 7th lags of $FullSupport_{it}$, and the 3rd through 12th lags of $SwitchToBasic_{it}$ as instruments for the difference equation. It also uses the 2nd lag of the 3 variables' first difference as instruments for the levels equation. Total number of instruments is 387. Hansen (1982) specification test passed with $\chi^2(314) = 325.11,\ p = 0.321$. Robust standard errors use Windmeijer's (2005) finite sample correction.

As before, we implement a dynamic panel data model using GMM estimation. We again begin by using 4 lags of the dependent variable as regressors. Using all available instruments, the Arellano and Bond (1991) test also found 2nd order serial correlation, so we must again rely on the 3rd or later lags of the variables' values and the 3rd lag of their first difference as instruments. The smallest number of lags of $Horizontal_{it}$ we can include in our regression is 4. In terms of instruments, the specification with the smallest numbers of instruments employs the 3rd lag of $Horizontal_{it}$, the 3rd through 7th lags of $FullSupport_{it}$, and the 3rd through 12th lags of $SwitchToBasic_{it}$ as instruments for the difference equation. We also use the 3rd lag of the 3 variables' first difference as instruments for the levels equation. With this specification we pass the Hansen (1982) instrument validity test with $\chi^2(314) = 325.11, p = 0.321$. Our total number of instruments is 387. We report our results for this model in column (8) of Table 6. Despite adding the lagged dependent variable, we still find that the adoption and continued access to full support increases the likelihood of using horizontal scaling by 1.68 percentage points. This likelihood only decreases marginally (0.66 percentage points, and only significant at the 10% level) when customers switch to basic support. Thus, we continue finding support for the hypothesized effects of full support.

Table 7. Are the results of full support on architecture complexity stronger for large firms?

Column	(1)	(2)	(3)	(4)
$FullSupport_{it}$	0.2104***	0.1151***	0.1803***	0.1187***
	(0.0232)	(0.0370)	(0.0269)	(0.0384)
$SwitchToBasic_{it}$	-0.0181	-0.0391**	-0.0366***	-0.0623***
	(0.0128)	(0.0184)	(0.0130)	(0.0187)
$FullSupport_{it} \times EmploymentTop50_i$		0.1320***		
		(0.0460)		
$SwitchToBasic_{it} \times EmploymentTop50_i$		0.0321		
		(0.0226)		
$FullSupport_{it} \times EmploymentTop25_i$			0.0897*	
			(0.0499)	
$SwitchToBasic_{it} \times EmploymentTop25_i$			0.0628**	
			(0.0262)	
$FullSupport_{it} \times lnEmployment_i$				0.0287***
				(0.0110)
$SwitchToBasic_{it} \times lnEmployment_i$				0.0160***
				(0.0060)
Constant	-0.9160***	-0.9156***	-0.9177***	-0.9187***
	(0.1214)	(0.1205)	(0.1202)	(0.1200)
N	71,478	71,478	71,478	7,1478
R^2	0.050	0.052	0.052	0.052
Customers	5,277	5,277	5,277	5,277

Dependent variable is $Horizontal_{it}$. 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 8. Changes in Likelihood of Employing Horizontally Scalable Architectures from Adopting and Switching from Managed Support at Different Levels of Firm Size

Column	(1)	(2)	(3)
Fixed level of $lnEmployees_i$ [a]	$\bar{x} - \sigma^2$	$\bar{x}$	$\bar{x} + \sigma^2$
$ManagedSupport_{it} = 1$	98.22	127.40	160.87
$SwitchToBasic_{it} = 1$	-20.86	-20.45	-20.03
$ManagedSupport_{it} = SwitchToBasic_{it} = 1$	56.87	80.90	108.61

[a] $\bar{x}$ and σ^2 denote the sample mean and standard deviation of $lnEmployees_i$, respectively. Results are based on column (4) of Table 7 and are expressed in changes in percentage points.

We now turn to the role of firm size in moderating how full support influences customers' architecture complexity. We report the regression results in Table 7 and compute the percentage changes in the likelihood of employing a horizontally scalable deployment at different levels of $lnEmployment_i$ in Table 8. While larger firms that adopt full support increase their likelihood of using horizontal scaling methods by 23.80 to 27.00 percentage points, smaller firms only do so by 11.51 to 18.03 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 while larger firms (on average) are able to continue using the same level of complexity in their architectures, smaller firms practically completely loose this capability. Post-switch to basic, large former full support customers continue being between 23.80 and 29.62 percentage point more likely to employ horizontal scaling than basic support customers. In contrast, small former full support customers are only between 7.6 and 14.37 percentage points more likely to do so. This suggests that larger firms are much more capable of retaining and not forgetting what they have learned from the provider.

8 Conclusion

Using a unique and rich nano-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 technical support influences buyer demand for its services. We examine how a buyer's access to support influences its usage of on-demand IT infrastructure. We also explore the underlying mechanisms behind the positive relationship between technical support and the service's demand, studying whether the value of the knowledge transferred from the provider to the customer is durable. Taking advantage of the near-commodity nature of the cloud, we provide evidence that technical support increases buyers' productivity via co-production, thereby increasing the demand for the provider's service.

Our estimates of the positive impact of offering technical support are economically significant. Buyers who adopt and access full support, consume, on average, 137% more IT capacity than those who only access basic support. We also find evidence that customers who switch from full to basic support continue consuming an average of 105% more IT capacity than customers who only had access to basic support throughout our entire sample period. These findings directly impact the provider's topline performance. While a buyer consuming IT capacity at the median of the capacity distribution generates an ARPU of $64.60, buyers who opt for full support generate an ARPU of $153.00 and buyers who switch to basic support continue contributing an ARPU of $132.55.

Our results also indicate that customers internalize and retain most of what they have learned from the provider: in other words, the effects of knowledge transfer on service demand are durable. We find strong evidence that larger firms are most responsive to the provider's technical support; their increases in IT capacity demand after adopting full support are more than twice that of smaller firms. These finding are particularly interesting since, although inter-organizational knowledge transfers have been studied in a variety of other contexts, much less is known regarding the durability of such knowledge transfers once the organizations limit or cease their interactions. Thus, our work not only shows that the effects of knowledge transfer are durable, but also that complementarities between external and internal knowledge sources are retained. Altogether our findings suggest that offering technical support is an effective differentiation strategy when offering commodity-like IT services.

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 nano-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 technical support has the potential to increase demand for IaaS and to help users make use of advanced cloud features, enabling them to overcome the knowledge barriers that have engendered the slow rates of cloud adoption we see today.

Our findings point to several opportunities for future research. We plan to explore in depth the underlying knowledge transfer mechanisms by taking advantage of nano-data from the provider to measure the value of its different Internet-based knowledge exchange channels—tickets and chats—and their interactions. 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.

References

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

Archak, N., Ghose, A., and Ipeirotis, P.G. 2011. "Deriving the Pricing Power of Product Features by Mining Consumer Reviews," *Management Science* (57:8), August 1, 2011, pp 1485-1509.

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

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

Argote, L., and Miron-Spektor, E. 2011. "Organizational Learning: From Experience to Knowledge," *Organization Science* (22:5), September/October 2011, pp 1123-1137.

Arora, A., and Gambardella, A. 1994. "Evaluating Technological Information and Utilizing It: Scientific Knowledge, Technological Capability, and External Linkages in Biotechnology," *Journal of Economic Behavior and Organization* (24:1), pp 91-114.

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

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

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

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

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

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

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

Boone, T., Ganeshan, R., and Hicks, R.L. 2008. "Learning and Knowledge Depreciation in Professional Services," *Management Science* (54:7), July 1, 2008, pp 1231-1236.

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

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

Burton, D. 2002. "Consumer Education and Service Quality: Conceptual Issues and Practical Implications," *Journal of Services Marketing* (16:2), pp 125-142.
Cameron, A.C., and Trivedi, P.K. 2010. *Microeconometrics Using Stata, Revised Edition*. Texas: Stata Press.
Casalicchio, E., and Colajanni, M. 2000. "Scalable Web Clusters with Static and Dynamic Contents," *IEEE International Conference on Cluster Computing, 2000.*, pp. 170-177.
Cassiman, B., and Veugelers, R. 2006. "In Search of Complementarity in Innovation Strategy: Internal R&D and External Knowledge Acquisition," *Management Science* (52:1), January 1, 2006, pp 68-82.
Chang, Y.B., and Gurbaxani, V. 2012. "Information Technology Outsourcing, Knowledge Transfer, and Firm Productivity: An Empirical Analysis," *MIS Quarterly* (forthcoming).
Chase, R.B. 1978. "Where Does the Customer Fit in a Service Operation?," *Harvard Business Review* (56:6), p 137.
Chen, H., De, P., and Hu, J. 2013. "IT-Enabled Broadcasting in Social Media: An Empirical Study of Artists' Activities and Music Sales." Working Paper available at SSRN: http://ssrn.com/abstract=2201430
Cherkasova, L. 2000. "Flex: Load Balancing and Management Strategy for Scalable Web Hosting Service," *Fifth IEEE Symposium on Computers and Communications (ISCC 2000)*, Antibes, France, pp. 8-8.
Cohen, W.M., and Levinthal, D.A. 1989. "Innovation and Learning: The Two Faces of R & D," *The Economic Journal* (99:397), pp 569-596.
Cohen, W.M., and Levinthal, D.A. 1990. "Absorptive Capacity: A New Perspective on Learning and Innovation," *Administrative Science Quarterly* (35:1), pp 128-152.
Darr, E.D., and Kurtzberg, T.R. 2000. "An Investigation of Partner Similarity Dimensions on Knowledge Transfer," *Organizational Behavior and Human Decision Processes* (82:1), pp 28-44.
Das, A. 2003. "Knowledge and Productivity in Technical Support Work," *Management Science* (49:4), pp 416-431.
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), pp 427-455.
Fichman, R.G., and Kemerer, C.F. 1999. "The Illusory Diffusion of Innovation: An Examination of Assimilation Gaps," *Information Systems Research* (10:3), pp 255-275.
Field, J.M., Xue, M., and Hitt, L.M. 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), pp 43-56.
Forman, C. 2005. "The Corporate Digital Divide: Determinants of Internet Adoption," *Management Science* (51:4), pp 641-654.
Forman, C., Goldfarb, A., and Greenstein, S. 2008. "Understanding the Inputs into Innovation: Do Cities Substitute for Internal Firm Resources?," *Journal of Economics & Management Strategy* (17:2), pp 295-316.
Furman, J.L., Jensen, K., and Murray, F. 2012. "Governing Knowledge in the Scientific Community: Exploring the Role of Retractions in Biomedicine," *Research Policy* (41), pp 276-290.
Garcia, D.F., Rodrigo, G., Entrialgo, J., Garcia, J., and Garcia, M. 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.
Ghose, A. 2009. "Internet Exchanges for Used Goods: An Empirical Analysis of Trade Patterns and Adverse Selection," *MIS Quarterly* (33:2), pp 263-291.

Hansen, L.P. 1982. "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica* (50:4), pp 1029-1054.
Ho, D., Imai, K., King, G., and Stuart, E. 2007. "Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference," *Political Analysis* (15:3), pp 199-236.
Iacus, S.M., King, G., and Porro, G. 2012. "Causal Inference without Balance Checking: Coarsened Exact Matching," *Political analysis* (20:1), pp 1-24.
Ko, D.-G., Kirsch, L.J., and King, W.R. 2005. "Antecedents of Knowledge Transfer from Consultants to Clients in Enterprise System Implementations," *MIS Quarterly* (29:1), pp 59-85.
Levenshtein, V.I. 1966. "Binary Codes Capable of Correcting Deletions, Insertions, and Reversals," *Soviet Physics Doklady* (10:8), pp 707-710.
Mell, P., and Grance, T. 2011. "The Nist Definition of Cloud Computing," National Institute of Standards and Technology Information Technology Laboratory (ed.). Gaithersburg, MD.
Menon, T., and Pfeffer, J. 2003. "Valuing Internal Vs. External Knowledge: Explaining the Preference for Outsiders," *Management Science* (49:4), April 1, 2003, pp 497-513.
Michael, M., Moreira, J.E., Shiloach, D., and Wisniewski, R.W. 2007. "Scale-up X Scale-Out: A Case Study Using Nutch/Lucene," *Parallel and Distributed Processing Symposium (IPDPS 2007). IEEE International*, pp. 1-8.
Microsoft, and Edge Strategies. 2011. "SMB Cloud Adoption Study Dec 2010 - Global Report," http://www.edgestrategies.com.
Nickell, S. 1981. "Biases in Dynamic Models with Fixed Effects," *Econometrica* (49:6), pp 1417-1426.
Reese, G. 2009. *Cloud Application Architectures: Building Applications and Infrastructure in the Cloud*. O'Reilly Media.
Rogers, E.M. 1995. *Diffusion of Innovations*. New York: The Free Press.
Roodman, D. 2009a. "How to Do Xtabond2: An Introduction to Difference and System Gmm in Stata," *Stata Journal* (9:1), pp 86-136.
Roodman, D. 2009b. "A Note on the Theme of Too Many Instruments," *Oxford Bulletin of Economics & Statistics* (71:1), pp 135-158.
SearchDataCenter.com. 2011. "Data Center Decisions 2011 Survey Special Report," TechTarget.
Singh, J., and Agrawal, A. 2011. "Recruiting for Ideas: How Firms Exploit the Prior Inventions of New Hires," *Management Science* (57:1), pp 129-150.
Symantec. 2011. "State of the Cloud Survey," Symantec.
Varian, H.R. 2010. "Computer Mediated Transactions," *American Economic Review* (100:2), pp 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.
Vega-Jurado, J., Gutiérrez-Gracia, A., and Fernández-de-Lucio, I. 2009. "Does External Knowledge Sourcing Matter for Innovation? Evidence from the Spanish Manufacturing Industry," *Industrial and Corporate Change* (18:4), August 1, 2009, pp 637-670.
Venters, W., and Whitley, E.A. 2012. "A Critical Review of Cloud Computing: Researching Desires and Realities," *Journal of Information Technology* (forthcoming).
Windmeijer, F. 2005. "A Finite Sample Correction for the Variance of Linear Efficient Two-Step Gmm Estimators," *Journal of Econometrics* (126:1), pp 25-51.
Wittmann, A. 2012. "Cloud Computing Is Still in Its Adolescence." *Information Week, Global CIO* Retrieved Aug 24, 2012, from http://www.informationweek.com/global-cio/interviews/cloud-computing-is-still-in-its-adolesce/232600928
Xue, M., and Harker, P.T. 2002. "Customer Efficiency," *Journal of Service Research* (4:4), pp 253-267.
Xue, M., Hitt, L.M., and Chen, P.-Y. 2011. "Determinants and Outcomes of Internet Banking Adoption," *Management Science* (57:2), pp 291-307.
Xue, M., Hitt, L.M., and Harker, P.T. 2007. "Customer Efficiency, Channel Usage, and Firm Performance in Retail Banking," *Manufacturing & Service Operations Management* (9:4), pp 535-558.