Lock-In and Unobserved Preferences in Server Operating System Adoption: A Case of Linux vs. Windows*

Seung-Hyun Hong
Department of Economics
University of Illinois
hyunhong@ad.uiuc.edu

Leonardo Rezende
Department of Economics
PUC-Rio
lrezende@econ.puc-rio.br

August 30, 2008

Abstract

This paper investigates two leading factors – lock-in and unobserved preferences – in the decision to adopt server operating systems. Both factors imply a strong positive correlation between the current choice and the previous choice. To distinguish between these two factors, we construct our model based on plausible assumptions about unobserved preferences. The model generates conditional moment restrictions and also allows us to use a linear GMM estimator. Accordingly, we implement a GMM version of semiparametric discrete choice panel data methods (Arellano and Carrasco 2003). Using detailed establishment-level data, we find that once unobserved preferences are taken into account, the coefficient estimates for the previous choice are considerably small and statistically insignificant, suggesting that unobserved preferences, rather than lock-in, are more important in firms' decisions to adopt server operating systems. Further robustness checks are consistent with our main findings.

JEL classification: C23, L15, L17, L86

Keywords: Discrete choice; Panel data; GMM; Lock-in; Unobserved preferences

*We appreciate the NET Institute for financial support. We thank George Deltas, Bin Gu, Mark Jacobsen, Kyoo il Kim, Roger Koenker, and Zhongjun Qu for helpful discussion and comments. We also thank participants at the NET Institute conference and the Symposium on Econometric Theory and Application for helpful comments. All remaining errors are our responsibility. Corresponding author: Seung-Hyun Hong, 428 David Kinley Hall, 1407 West Gregory Drive, Urbana, IL 61801.

1 Introduction

The persistent dominance of Microsoft's Windows is often claimed to be due to high switching costs and lock-in, suggesting potential inefficiency. By contrast, it is also argued that consumers continue to use Windows because of its superior quality, or at least perceived better quality, implying potential efficiency in the operating system market. Despite its importance in the debate over the antitrust case against Microsoft,[1] however, few empirical studies have attempted to separate out lock-in from perceived better quality, partly because of lack of detailed individual data, but also because of the difficulty in identification – both lock-in and unobserved preferences imply a strong positive correlation between the current choice and the previous choice in computer operating systems.[2]

To fill this gap, we use establishment-level data in this paper and propose a new identification strategy to distinguish between lock-in (or state dependence) and unobserved preferences for operating systems. In particular, we focus on Linux and Windows, and examine establishments' decision[3] to adopt either operating system for computer servers. To this end, we use unbalanced panel data from 2000-2003 *Computer Intelligence Technology Database* (CITDB) collected by the Harte-Hanks Market Intelligence. The CITDB surveys over 100,000 establishments in the United States every year. It contains detailed information on establishment characteristics and ownership of information technologies such as operating systems for various computers. This detailed information allows us to examine establishment-level decisions to adopt server operating systems.

As expected, the descriptive statistics from the CITDB show significant positive correlations between the current choice of Windows and the previous choice of Windows, and similarly

[1]See, e.g., Bresnahan (2001) and Liebowitz and Margolis (1999).

[2]This problem reflects the well-known difficulty in identification between state dependence and unobserved heterogeneity. See, e.g., Heckman (1981a, 1981c) and Hsiao (2003).

[3]This paper considers the operating system adoption by establishments, but not necessarily by firms, since a firm may own multiple establishments, that is, branches in different locations. Nevertheless, firms and establishments are identical for about a half of our samples from the CITDB, as they are single facility firms. For this reason and for convenience, we use both terms – establishments and firms – interchangeably in this paper.

for Linux. We then estimate probit models for the adoption of server operating systems, allowing for various other factors to determine the adoption decision. Without accounting for unobserved heterogeneity, we still find seemingly strong state dependence in both Windows and Linux. This positive correlation is robust even after we control for various observed heterogeneity. Nonetheless, we cannot interpret this result as evidence for lock-in in server operating systems, as it is also consistent with the importance of unobserved preferences for either operating system.

To distinguish between lock-in and unobserved preferences, we therefore construct our model based on plausible assumptions about unobserved preferences. Specifically, we posit that each firm has the true preference for operating systems under full information. In realistic situations, however, firms are unlikely to have full information on technical features and the quality of operating systems, because of complexity in operating systems as well as frequent updates and releases of new versions. Therefore, it is plausible to assume that the true preference under full information is not fully observed even to firms themselves. Hence, firms need to take conditional expectation of the true preferences for operating systems based on their past experiences. This assumption leads to conditional moment restrictions based on the law of iterated expectation.[4] Inverting the equation for the adoption probability then allows us to exploit these moment conditions and use a linear GMM estimator. Note that our model is closely related to the semiparametric random effects models in Arellano and Carrasco (2003). As a result, we implement a GMM version of their discrete choice panel data method to estimate our model.[5]

We apply the method to 11 subsamples in the CITDB which we group based on the NAICS code. For most subsamples, we find that the coefficient estimates for the previous choice of

[4]The conditional moment conditions imply that conditional on the same information available at present, the expectation of the true preference taken at the next period should on average be equal to the expectation of the true preference taken at the current period.

[5]We also consider other related approaches to account for unobserved heterogeneity in dynamic discrete choice models. Most approaches including standard random effects models and conditional logit fixed effect models, however, impose very strong assumptions that are unlikely to hold in our application. See Section 3.3.

server operating systems are considerably small and statistically insignificant, suggesting that unobserved preferences, rather than lock-in, are more important in firms' decision to adopt server operating systems. Though the magnitudes of these coefficient estimates are small, we also find that the estimated values vary across different subsamples, which is likely to reflect heterogeneity in unobserved preferences. Further robustness checks are consistent with our main findings.

Our findings can be interpreted as evidence that perceived better quality, or unobserved preference, may explain the persistent dominance of Windows in the operating system market. Similarly, unobserved preferences can also explain the increasing popularity of Linux in server operating systems. Nevertheless, the CITDB is not intended to be representative of all firms in the United States, though its sample size is large. For this reason, we do not attempt to generalize our findings beyond the samples examined in this paper.

The paper is organized as follows. Section 2 describes our data and provides descriptive statistics. Section 3 constructs our model and examines identifying assumptions. This section also presents our estimation approach and discusses related issues. Section 4 reports the results and robustness check. Section 5 concludes the paper.

2 Data and Descriptive Statistics

2.1 Data Description

We use data from the *Computer Intelligence Technology Database* (CITDB) collected by Harte-Hanks Market Intelligence. The CITDB is a yearly survey of over 100,000 establishments in the United States. It contains detailed establishment-level data on the use of a variety of information and communication technologies. This dataset has been used in several papers (e.g. Bresnahan and Greenstein 1996; Bresnahan, Brynjolfsson, and Hitt 2002). In this paper, we focus on the period from 2000 to 2003, during which two major events in operating system markets occurred – Microsoft released Windows 2000 in February 17, 2000,[6] and Windows XP in October 25, 2001.[7] These releases are likely to have led most firms to decide on their operating systems and then upgrade or switch their operating systems during this period. For this reason, we use the 2000-2003 CITDB data.[8] Though our data cover four years, they should contain enough information to study firm's decisions to adopt operating systems.

The CITDB is useful for our purpose because it contains detailed information on establishment characteristics and ownership of computer hardware and software including operating systems. The unit of observation is an establishment in a year. The CITDB has attempted to survey the same establishment each year, so that the dataset contains panel information of many establishments. Because the survey is voluntary, however, some establishments did not respond to survey requests, and the CITDB has added new establishments each year. As a result, the total number of observations remains similar each year, but many establishments were not surveyed in every year.

We study the adoption of operating systems at the segment level. The CITDB groups computers into four segments: Internet servers; network servers; personal computers, not used for either Internet servers or network servers; and non-PCs not used for servers. In this paper, we consider three mutually exclusive segments: server, including both Internet servers and network servers[9]; PC, including personal computers that are used for standalone desktops or client computers connected to servers[10]; and non-PC, including mainframes, midrange, and workstations that are not used for servers. Note that we can only investigate the adoption of operating systems up to the segment level, since the information on operating system choices

[6] Refer to http://www.microsoft.com/presspass/press/2000/feb00/w2kavailablepr.mspx.

[7] Refer to http://www.microsoft.com/presspass/press/2001/oct01/10-25xpoverallpr.mspx.

[8] Harte-Hanks releases a new dataset every January, containing information collected in the previous year. Our reference year is the collection year, not the year of release; e.g. the 2000 dataset was released January 2001. We also obtained the 2004 data, but we do not use them in this paper because some of the key variable definitions are inconsistent with those in the 2000-2003 data.

[9] We combine Internet servers and network servers for two reasons: to simplify our analysis and to increase the size of samples with any kind of server.

[10] Some PCs can be used as servers, but such PCs are included in the server segment in our data.

at the individual computer level is not available in the CITDB. In other words, we observe which kinds of operating systems are used for computers in each segment, but we do not know exactly which operating system is running on each individual computer. The segment-level information is valuable, nonetheless, because most establishments in the CITDB tend to use only one kind of operating system for each segment and many of them use only a small number of computers for each segment, except for the PC segment.

Table 1 presents summary statistics of variables used in our analysis. We use Windows to denote Windows-family operating systems such as Windows 95, 98, ME, NT, 2000, 2003, and XP. Linux indicates not only various versions of Linux (e.g. Debian, Red-Hat, Mandrake, SuSE, etc.) but also Berkeley Software Distribution (BSD).[11] We use other to denote other operating systems including Mac OS X as well as a variety of proprietary Unix (e.g. Solaris, HP-UX, AIX). Because we consider three segments, we use the following notations to denote the choice of operating systems on each segment: server.linux for Linux on the server segment; pc.linux for Linux on the PC segment; non-pc.linux for Linux on the non-PC segment.[12]; and similarly for server.windows, pc.windows, and non-pc.windows.

At least three observations emerge from Table 1. First, Windows is dominant in the PC segment as well as in most server segments, except for the non-PC segment where other operating system is the most popular, probably because most non-PCs are IBM computers running IBM operating systems. Note also that the adoption of Windows has increased in most server segments. This may suggest that potential network effects between the PC segment and server segments could have led Windows to gain popularity even in server segments. However, it is also plausible that the quality of Windows in server operating systems has improved significantly, so that more firms have adopted Windows for their server operating systems.

[11] BSD is the Unix derivative developed by the University of California, Berkeley. BSD is not Linux and follows its own licensing agreement different from the GNU Public License. Nevertheless, we include BSD in the Linux category, because BSD is similar to Linux in that it is a Unix-like operating system and is available for free. The percentage of establishments using BSD, however, is negligible in our data.

[12] We also use internet.sv.linux for Linux on the Internet server segment, network.sv.linux for Linux on the network server segment.

Second, the adoption of Linux has increased in both server segments and the PC segment, though its share seems to be still moderate. Notice, however, that many establishments do not own server computers, implying that 2-3% of Linux adoption in Internet servers, for example, is translated to over 10% of the share for Internet server operating systems.[13] Third, the adoption of other operating systems has declined over time. One possibility for these trends is that firms may have switched to Linux, not from Windows, but from proprietary Unix. However, this does not imply that the competition occurs only between Linux and Unix, and Windows is irrelevant to Linux adoption. Note that quite a few establishments have switched from Windows to Linux, and many establishments have switched from Unix to Windows, suggesting that the competition between Linux and Windows might be indeed intense. Section 2.3 examines these possibility in more detail.

2.2 Sample Restriction

For our empirical analysis in the following sections, we restrict our sample in order to meet three considerations. First, we restrict our sample to the establishments that report which server operating systems they are using.[14] Establishments may not report information on server operating systems for two reasons: either because they do not have any server computer, or because they consider server operating systems unimportant. By excluding the former case, we implicitly assume that our analysis is conditional on establishments' ownership of computer servers. Though it would be interesting to know which operating system an establishment without any server would choose if it started to use a server, our analysis does not allow us to construct such a counterfactual. The latter case is a common problem in many survey data – respondents do not answer every question in the survey, either because they do not remember, or because they do not consider it important. The CITDB is not an exception in this regard.

[13] If some establishments own a large number of servers running only Linux, the actual market share of Linux in servers would be much higher. Hence, we may underestimate the market share for Linux.

[14] Among 487,512 observations in the CITDB, about 53.5% of them report information on operating systems for either Internet server or network server.

This problem can result in a potential underestimation of the number of establishments using each operating system. For lack of further information, we cannot account for this problem. To the extent that this potential measurement error occurs randomly, however, it may not affect the estimated market share of each operating system.

Second, we do not use the observations whose information on computing technology was outdated. The CITDB does not survey all the establishments every year. For some observations, the CITDB reuses information collected in the previous year. If an establishment continues to use the same operating system as before, information on operating systems can be current even though it was collected in the previous year. On the other hand, if the establishment actually switched to different operating systems, using outdated information would result in a spurious positive correlation between the current choice and the previous choice. To avoid this problem, we use only those with up-to-date information.[15] For the initial observation of each establishment in our sample, there is no issue regarding reusing the same information. For this reason, we include the initial observation of each establishment as long as information on computing technology was collected within the last one year.

Third, we use only complete panels for our main analysis in Sections 3 and 4. Obviously, we cannot use information from establishments that are observed only once in our data. We further restrict our sample to complete panels in order to use the econometric methodology presented in Section 3. Because this restriction reduces the sample size for each year considerably, we additionally consider shorter panels of three consecutive years: 2000-2002 and 2001-2003. By doing so, we increase the sample size for each year and also check robustness of our findings.

To examine potential implications of our sample restriction, Table 2 presents the market share of each operating system from the unrestricted sample and the restricted samples as discussed above. To compute the market share, a dummy for each operating system is assigned

[15]Because the CITDB records when the survey on each establishment was conducted, we can find whether its information is outdated. Among the 260,796 observations with any kind of server operating system, about 70% of them report up-to-date information on computing technology.

to each observation, and the mean values are reported in the table. In panel B, for example, the mean of server.windows is 0.914 in 2000, indicating that 91.4% of establishments reported to use Windows in their server computers. Because an establishment can have more than one kind of operating system, the sum of columns (1)-(3) can be larger than 1. At least in terms of the market shares, the table shows that our restrictions do not seem to create systematically different samples from the others.[16] Nevertheless, we do not attempt to generalize our findings beyond the samples examined in our analysis.

2.3 Switching Pattern

Before we present our empirical analysis in the following sections, we examine basic patterns in our data, focusing on switching in server operating systems. Table 3 shows the results. Panel A in the table reports the share of firms in each year that follow the switching pattern specified in each column. For example, column 1 reports the share of firms that used Linux in the given year but had not used Linux in the previous year. The first three columns show switching patterns of Linux adoption. The next three columns present those of Windows adoption, and the last three columns show switching patterns for the adoption of other operating systems. Columns 1-2 show that while about 4-5% of firms started to use Linux each year, more firms continued to use Linux over time. Nevertheless, column 3 reports that those who discontinued using Linux have increased, suggesting that Linux could also lose its current market share.

Columns 4-6 show the strong dominance of Windows in server operating system markets. More than 90% of firms in 2001 used Windows in both 2000 and 2001, and more firms have continued to use Windows in the following years. However, this does not mean that most firms were passive and did not make their decisions to adopt or switch operating system during this period. Recall that we combine Windows-family operating systems together. As a result, those who continued to use Windows include a substantial number of firms upgrading to Windows

[16] One exception is that the market shares in panels B through F are about twice larger than those in panel A, but this is expected because approximately half of observations in the unrestricted sample do not report to own any kind of server computer.

2000 and Windows XP during this period.

Similar shares for all years are reported in panel B. These are the average of the shares presented in panel A. In addition to the dominance of Windows, panel B also shows potentially significant switching costs in the adoption of operating systems. Note that the percentage of firms who used the same operating systems in two consecutive years is higher than the percentage of those who either started or discontinued the use of any of the operating systems. Panel C reports similar shares by different industries. It shows substantial heterogeneity in Linux adoption across industries. For example, 18% of firms in the information sector (the first two digits of NAICS equal to 51) continued to use Linux and 7.4% of them started to use Linux, whereas only 4.2% of those in the manufacturing sector (the first two digits of NAICS equal to 31-32) continued to use Linux and only 2.6% of them started to use Linux. Considerable heterogeneity across industries is also observed in the adoption of Windows and other operating systems.

One additional observation from Table 3 is that more firms discontinued to use other operating systems than started to use other operating systems (columns 7 and 9). Though it is possible that most of firms who stopped using other operating systems could switch to Linux, the table does not provide any evidence. For this reason, we decompose these firms into those who started to use Linux and those who switched to Windows. Table 4 presents the shares of such firms among total observations. Columns 6 and 8 in the table report that approximately 0.44% of firms switched to Linux, while about 1.39% of observations switched to Windows. Hence, more firms switched from other operating systems to Windows than to Linux. Table 4 also shows that those who switched to Linux are not necessarily those who stopped using other operating systems. Columns 1-4 report that those switching from Windows to Linux or from Linux to Windows are not negligible.

Overall, these switching patterns suggest strong state dependence in the adoption of server operating systems. Nevertheless, quite a few firms did in fact switch their operating systems.

Moreover, unobserved preferences for Windows, rather than lock-in due to high switching costs, may explain why many firms decided to continue to use Windows-family operating systems and have upgraded to Windows 2000 and XP. To distinguish between lock-in and unobserved preferences, the next section presents our econometric approach.

3 Models and Estimators

3.1 Economic Determinants of the Adoption Decision

To investigate the factors that determine firms' decision to adopt server operating systems, we consider the following reduced-form function for the net benefit of adopting server operating system j, where $j \in \{\textsf{Linux, Windows, other}\}$. For establishment i at period t, the net benefit is given by

$$\pi_{ijt} = \gamma_{jt} + \sum_k \beta_{jk} y_{ik(t-1)} + \alpha_j x_{it} + \delta_j Z_{it} + u_{ijt} \qquad (i = 1, \ldots, N; t = 1, \ldots, T), \tag{1}$$

where $y_{ik(t-1)}$ is a binary variable indicating whether firm i adopted server operating system k at the previous period, x_{it} is a vector of predetermined variables related to non-server segments, Z_{it} is a vector of observed characteristics of the firm, γ_{jt} captures a time effect, and u_{ijt} is an unobserved component of the net benefit.

The main focus of this paper is to distinguish between lock-in and unobserved individual preferences for operating systems. In this respect, the coefficient for $y_{ij(t-1)}$ reflects lock-in or state dependence, since high switching costs might lock firms into the previous investment in Windows or Linux. Because an information system in business computing environments is composed of various interrelated components including computer hardware and networks, database, and application software, it would be difficult to change one component such as the operating system without changing other components in the information system. Unless there are strong reasons to reorganize the existing architecture, firms may continue to use the same operating system, hence resulting in state dependence in the choice of operating systems.

The presence of unobserved preferences, however, generates a positive correlation between y_{ijt} and $y_{ij(t-1)}$ as well. Firms may have heterogeneous preferences over different operating systems, depending on firm characteristics or their assessment of the quality of operating systems. Therefore, firms may continue to use Windows, not necessarily because of high switching costs, but because of their preferences for Windows. Since these preferences are unlikely to be observed, however, we posit that they are included in the error term u_{ijt} in (1).

In addition to these two determinants, other factors might also be important in adoption decision. First, there might be direct network effects between different computers within a sever segment as well as across different segments. Especially in the adoption of operating systems for servers, firms may experience network effects between the operating systems in the PC segment and the operating systems in the server segment. For example, if a firm uses only Windows for all the PCs, then the value of using Windows on servers might be higher because of the compatibility between PCs and servers. Hence, the operating systems choice in PCs may affect the adoption of server operating systems.

Second, indirect network effects may also influence the adoption of operating systems. Some application software programs may be used together with a particular operating system. For example, Apache, one of the popular Web servers, was commonly run under the Linux operating system, suggesting that the adoption of Apache may be another factor behind the adoption of Linux (see Fink (2002)).

Third, operating systems have been frequently updated, and new versions of software programs have been constantly released. In particular, the releases of Windows 2000 and Windows XP are major events in operating system markets in 2000 and 2001, and they are likely to have affected firms' decisions to adopt or switch operating systems during the periods studied in this paper. Though we cannot fully account for all other factors in this paper, we attempt to consider some of them in our empirical analysis and include x_{it}, Z_{it}, and γ_{jt} in (1).

In this paper, we assume that establishment i decides to adopt operating system j at period t if the net benefit is non-negative. In other words, $y_{ijt} = \mathbb{1}\{\pi_{ijt} \geq 0\}$. Note that we do not use multinomial models often used in empirical studies, because an establishment in our data can own multiple servers and thus adopt more than one operating system at the segment level. However, we do not attempt to model the joint decision of adopting multiple operating systems.[17] Instead, we consider the adoption of each operating system separately, since this approach still allows us to examine the extent of state dependence or lock-in in the adoption of operating system j. In the equation for Linux adoption, for example, the coefficient for $\mathsf{server.linux}_{t-1}$ captures the degree of lock-in. One caveat is that the coefficients for $\mathsf{server.windows}_{t-1}$ and $\mathsf{server.other}_{t-1}$ in Linux adoption do not entirely reflect costs associated with switching from Windows or other operating systems to Linux.[18]

3.2 Unobservables and Identification

In this paper, we assume that u_{ijt} in (1) is the composite error given by

$$u_{ijt} = E(\eta_{ij}|H_i^t) + \epsilon_{ijt}, \tag{2}$$

where $H_i^t = (H_{i1}, \ldots, H_{it})$, and $H_{it} = (\mathsf{server.linux}_{t-1}, \mathsf{server.windows}_{t-1}, \mathsf{server.other}_{t-1}, x_{it}, Z_{it})$. We presume that $E(\eta_{ij}|H_i^t)$ reflects firm i's unobserved preference for operating system j, and ϵ_{ijt} is the idiosyncratic error term capturing the rest of unobserved component of the net benefit function.

We posit that each firm has η_{ij}, denoting the true preference for operating system under full information in which all the quality and various attributes of operating system are fully known to the firm. In realistic situations, however, firms are likely to have partial information on technical features and the quality of operating systems, and will acquire more information

[17]We could use multivariate models in the sense that we extend multinomial models by treating a segment-level joint adoption of different operating systems as one choice. The key factor in this joint decision, however, is network effects within the same segment, which cannot be fully captured by the multivariate models. Investigating such network effects requires further modeling, which is not pursued in this paper.

[18]For example, some observation i may have $y_{ijt} = y_{ij(t-1)} = 1$, while $y_{ik(t-1)} = 1$ and $y_{ikt} = 0$ for $k \neq j$. In this case, establishment i did not exactly switch from k to j because $y_{ij(t-1)} = 1 y_{ik(t-1)} = 1$. Hence, the coefficient for $y_{ik(t-1)}$ does not necessarily reflect switching costs from k to j.

based on their experiences over time. Note that the quality of operating systems depends not only on operating systems themselves, but also on other components in the information systems connected to operating systems. Moreover, frequent updates and new releases of operating systems and various components in the information systems make it practically impossible for each firm to keep track of every new feature of operating systems in each period. Therefore, it is plausible to assume that η_{ij} is not fully observed even to firm i. As a result, firm i will revise its expectation (or forecast) of the true preference for operating system j based on its previous experiences H_i^t. Consequently, unobserved preference is captured by $E(\eta_{ij}|H_i^t)$ in (2).

Given this assumption, the net benefit function in (1) is rewritten as

$$\pi_{ijt} = \gamma_{jt} + \sum_k \beta_{jk} y_{ik(t-1)} + \alpha_j x_{it} + \delta_j Z_{it} + E(\eta_{ijt}|H_i^t) + \epsilon_{ijt}. \tag{3}$$

To identify the parameters, we further assume that the idiosyncratic errors ϵ_{ijt} follow an i.i.d. known distribution such as normal distribution, so that $\epsilon_{ijt}|H_i^t \sim \mathcal{N}(0, \sigma_t)$. We can then write the probability of the adoption conditional on the history as follows:

$$\Pr(y_{ijt} = 1|H_i^t) = \Phi\left(\frac{\gamma_{jt} + \sum_k \beta_{jk} y_{ik(t-1)} + \alpha_j x_{it} + \delta_j Z_{it} + E(\eta_{ijt}|H_i^t)}{\sigma_t}\right), \tag{4}$$

where Φ is the standard normal cdf. Let us denote $h_{jt}(H_i^t) \equiv \Pr(y_{ijt} = 1|H_i^t)$. The equation above can be inverted to obtain

$$E(\eta_{ij}|H_i^t) = \sigma_t \Phi^{-1}(h_{jt}(H_i^t)) - \gamma_{jt} - \sum_k \beta_{jk} y_{ik(t-1)} - \alpha_j x_{it} - \delta_j Z_{it}. \tag{5}$$

We can then identify the parameters in the model by using the following moment condition:

$$E(\nu_{ijt}|H_i^{t-1}) \equiv E[E(\eta_{ij}|H_i^t)|H_i^{t-1}] - E(\eta_{ij}|H_i^{t-1}) = 0, \tag{6}$$

where $\nu_{ijt} \equiv E(\eta_{ij}|H_i^t) - E(\eta_{ij}|H_i^{t-1})$. The conditional moment restriction in (6) should hold because of the law of iterated expectation. Our interpretation is that if we condition only on information available at $t-1$ (i.e. H_i^{t-1}), the expectation of η_{ij} taken at period t should on average be equal to the expectation of η_{ij} taken at period $t-1$. A firm may change its opinion about the quality of server operating system j after its experience at the present period, but conditional on its previous experience, the expected future opinion at the next period should be the same as the current opinion.

To further understand the intuition behind identification, let us rewrite (5) as

$$Y_{ijt} = \sum_k \beta_{jk} y_{ik(t-1)} + \xi_{ijt}, \tag{7}$$

where $Y_{ijt} \equiv \sigma_t \Phi^{-1}(h_{jt}(H_i^t))$, $\xi_{ijt} \equiv E(\eta_{ij}|H_i^t)$, and we drop γ_{jt}, $\alpha_j x_{it}$, and $\delta_j Z_{it}$ in (5) to simplify the equation. Treating Y_{ijt} as observed variables, we could run ordinary least squares regressions to estimate β_j. However, the estimates would be inconsistent because ξ_{ijt} and $y_{ij(t-1)}$ are correlated. The idea behind identification is then to remove this correlation by first-differencing together with instrumental variables. To see this, let us consider the first-difference of (5) and rearrange it to obtain

$$\Delta Y_{ijt} = \sum_k \beta_{jk} \Delta y_{ik(t-1)} + \nu_{ijt}, \tag{8}$$

where $\Delta Y_{ijt} \equiv Y_{ijt} - Y_{ij(t-1)}$, and $\Delta y_{ij(t-1)} = y_{ij(t-1)} - y_{ij(t-2)}$. Given (8), we can think of H_i^{t-1} as instruments for $\Delta y_{ij(t-1)}$, because H_i^{t-1} determines $y_{ij(t-1)}$ and $y_{ij(t-2)}$, while (6) implies that H_i^{t-1} is not correlated with ν_{ijt}. For this reason, the conditional moment restriction in (6) addresses spurious state dependence due to unobserved preferences, thereby achieving identification.

3.3 Other Related Approaches

The model and the assumptions above are closely related to those in Arellano and Carrasco (2003). Therefore, we estimate our model parameters using the methodology proposed by Arellano and Carrasco (2003) – henceforth, the AC method. Before we present our estimation approach, we briefly consider two common approaches that might be used to distinguish between state dependence and unobserved preferences in discrete choice panel data models.[19] As

[19]See, e.g., Arellano and Honoré (2001), and Hsiao (2003) for more detailed literature review.

discussed below, these standard approaches require very strong assumptions that are unlikely to be plausible in our context, which motivates our application of the AC method.

The first approach to account for unobserved heterogeneity under discrete choice models is to treat η_{ij} as fixed effects (i.e. not imposing distributional assumptions for η_{ij}), while assuming a logit model for the idiosyncratic error terms. This method relies on conditional maximum likelihood methods and exploits the functional form of a conditional logit in order to difference out the fixed effects. Honoré and Kyriazidou (2000) extend this method to the case with predetermined variables. The identification of their method, however, requires conditioning the analysis on observations where the dependent variable follows specific patterns, namely $(y_{ij1}, y_{ij2}, y_{ij3}, y_{ij4}) = (0, 0, 1, 1)$ or $(1, 1, 0, 0)$, and $(y_{ij1}, y_{ij2}, y_{ij3}, y_{ij4}) = (0, 1, 0, 1)$ or $(1, 0, 1, 0)$. The problem of applying this method to our data is that we rarely observe the latter case. Few firms experiment with the same operating system by not using it at the first period, using it at the second period, and not using it again at the third period, and finally using it again at the fourth period. As a result, we cannot apply this approach to our data.

The second, so called the random effect approach, is to impose known distributional assumptions for the unobserved heterogeneity η_{ij} and integrate η_{ij} out in the likelihood function. For a model with predetermined variables such as lagged variables $y_{ij(t-1)}$, however, this approach yields inconsistent estimates. Even if η_{ij} actually follows an i.i.d. normal distribution, it is correlated with $y_{ij(t-1)}$, since η_{ij} also determined $y_{ij(t-1)}$ at period $t-1$. Consequently, the random effect η_{ij} cannot be simply integrated out. Alternatively, one could consider the following likelihood for an establishment with $T+1$ observations as

$$L_i = \int \prod_{t=1}^{T} \Pr(y_{ijt} | H_i^{t-1}, \eta_{ij}) f(y_{ij0} | \eta_{ij}) dG(\eta_{ij}), \tag{9}$$

where η_{ij} follows distribution $G(\eta_{ij})$, $f(y_{ij0}|\eta_{ij})$ denotes the marginal probability of the initial choice in the data given η_{ij}, and we assume $H_i^t = (y_{ij0}, y_{ij1}, \ldots, y_{ijt})$ for simplicity. The key difficulty of using (9) is how to specify the distribution of the initial condition given η_{ij}.

One could assume that y_{ij0} is independent of η_{ij}, but this assumption is very strong because we do not observed the very beginning of the process and y_{ij0} should be determined by η_{ij} and the history before y_{ij0}. One could also assume that $f(y_{ij0}|\eta_{ij})$ represents a steady-state distribution, but such a stationary distribution has been found only in limited special cases.[20]

In contrast to the random effect approach, the AC method for our model implicitly specifies the conditional distribution of η_{ij} given the initial condition y_{ij0}, as opposed to $f(y_{ij0}|\eta_{ij})$ in (9). The corresponding likelihood function under our assumptions in the previous section is then given by

$$L_i = \Pr(y_{ij0}) \int \prod_{t=1}^{T} \Pr(y_{ijt}|H_i^{t-1}, \eta_{ij}) dG(\eta_{ij}|y_{ij0}).$$

Therefore, the AC method allows for dependence between η_{ij} and y_{ij0}, while leaving the initial conditions of the process unrestricted.[21] Another difference between the random effect approach and the AC method is that the distributional assumption for η_{ij} is the key aspect of the random effect approach, whereas the AC method does not impose any parametric assumption on the distribution of η_{ij} and instead treats η_{ij} only through a nonparametric conditional expectation of η_{ij} given H_i^t. Because this semi-parametric approach is based on plausible assumptions as discussed in 3.2, while other related approaches rely on assumptions that are unrealistic and too strong in our context, we use the AC method in our application.

3.4 Estimation

Our estimation approach is based on the conditional moment restriction in (6), which naturally leads to a GMM estimation for our model. For this reason, we implement a GMM version of the AC method in our application. To explain the method, we first simplify the notation in this section. Though our actual estimation uses y_{ijt}, where $j \in \{\text{Linux, Windows, other}\}$, we suppress the subscript j in this section and consider the adoption decision denoted by

[20] See Heckman (1981b) and Wooldridge (2005) for more discussion on the initial conditions problem in dynamic discrete choice panel data models.

[21] See Wooldridge (2005) for related approaches and further discussion on the advantages of modeling the distribution of the unobserved heterogeneity conditional on the initial conditions.

y_{it}. Note also that x_{it} is a vector of indicator variables for adoption in non-server segments at the previous period, and we drop Z_{it} in our estimation because we estimate the model separately for relatively homogeneous subsamples. We then have a discrete random vector $H_{it} = (y_{it}, x_{it})$. Let H_{it} have a finite support of L points. The vector $H_i^t = (H_{i1}, \ldots, H_{it})$ then takes on L^t different values ϕ_l^t $(l = 1, \ldots, L^t)$. Hence, we can write the probability of the adoption conditional on the history ϕ_l^t as $h_t(\phi_l^t) = \Pr(y_{it} = 1 | H_i^t = \phi_l^t) \equiv p_l^t$. Let us also define $d_{il}^t = 1\!\!1\{H_i^t = \phi_l^t\}$. We thus have $h_t(H_i^t) = \sum_{l=1}^{L^t} d_{il}^t p_l^t$.

Given these notations, (6) implies the following moments

$$E(d_{il}^{t-1} \nu_{it}) = 0 \qquad (l = 1, \ldots, L^{t-1}), \tag{10}$$

which can be written as

$$E\left\{d_{il}^{t-1}\left[\sigma_t \Phi^{-1}(h_t(H_i^t)) - \sigma_{t-1}\Phi^{-1}(h_{t-1}(H_i^{t-1})) - \Delta\gamma_t - \beta\Delta y_{i(t-1)} - \alpha\Delta x_{it}\right]\right\} = 0, \tag{11}$$

where $\Delta\gamma_t = \gamma_t - \gamma_{t-1}$, $\Delta y_{i(t-1)} = y_{i(t-1)} - y_{i(t-2)}$,[22] and $\Delta x_{it} = x_{it} - x_{i(t-1)}$. To use the moments (11) in our estimation, let us further define

$$\psi_{il}^{t-1}(p, \theta) = d_{il}^{t-1}\left(\sigma_t \Phi^{-1}(h_t(H_i^t)) - \sigma_{t-1}\Phi^{-1}(h_{t-1}(H_i^{t-1})) - \Delta\gamma_t - \beta\Delta y_{i(t-1)} - \alpha\Delta x_{it}\right), \tag{12}$$

where θ is a vector of parameters to be estimated, and p is a vector of p_l^t's, $\forall t, l$. Because the moment condition (11) should hold for each t and l, we consider a vector of ψ_{il}^t's given by

$$\psi_i(p, \theta) = \begin{pmatrix} \psi_i^1(p, \theta) \\ \vdots \\ \psi_i^{T-1}(p, \theta) \end{pmatrix}, \quad \text{where} \quad \psi_i^t(p, \theta) = \begin{pmatrix} \psi_{i1}^t(p, \theta) \\ \vdots \\ \psi_{iL^t}^t(p, \theta) \end{pmatrix}.$$

We then write the sample orthogonality conditions as

$$\frac{1}{N}\sum_{i=1}^{N} \psi_i(p, \theta). \tag{13}$$

[22] The presence of $\Delta y_{i(t-1)}$ in the moment condition implies that identification of the parameters would require at least three observations available for each establishment.

Note that the orthogonality conditions above contain p_l^t that is unknown but can be non-parametrically estimated from the data. For this reason, we estimate the model parameters using a two-step approach, in which the first step estimates p_l^t by using the orthogonality conditions given by $E[d_{il}^t(y_{it} - p_l^t)] = 0$ $(l = 1, \ldots, L^t)$. This leads to cell-specific sample frequency estimators $\hat{p}_l^t = \frac{1}{\sum_{i=1}^N d_{il}^t} \sum_{i=1}^N y_{it} d_{il}^t$. In the second step, we then estimate θ using a GMM estimator given by

$$\hat{\theta} = \arg\min_{\theta} \left[\frac{1}{N} \sum_{i=1}^{N} \psi_i(\hat{p}, \theta) \right]' A_N \left[\frac{1}{N} \sum_{i=1}^{N} \psi_i(\hat{p}, \theta) \right], \tag{14}$$

where we replace p with $\hat{p}$, and A_N is a weighting matrix.

Before we present the results below, we need to discuss several practical issues in applying the AC method. First, $\psi_i(p, \theta)$ is supposed to be the $(\sum_{t=1}^{T-1} L^t) \times 1$ vector, but many cells of the history ϕ_l^t may be empty. We thus include only the sample moments for the histories actually observed in the data. The actual dimension of $\psi_i(p, \theta)$, denoted by M, will then be far less than $\sum_{t=1}^{T-1} L^t$. Second, some cells may contain very few observations, in which case there may be small sample biases in the estimate $\hat{p}_l^t$ for those cells. Arellano and Carrasco (2003) suggest to drop cells containing very few observations. We follow their suggestion but also check robustness of the results by experimenting with different cutoffs for dropping cells containing few observations.

Third, if the positive correlation between y_{it} and $y_{i(t-1)}$ is strong, and there are many cells with few observations, then an estimate of the asymptotic covariance matrix of the orthogonality condition (13) is likely to be singular or nearly singular. Note that a consistent estimate of the asymptotic covariance matrix used in Arellano and Carrasco (2003) is given by

$$\hat{W} = [I_M, -\hat{Q}] \left[\frac{1}{N} \sum_{i=1}^{N} \zeta_i(\hat{p}, \hat{\theta}) \zeta_i(\hat{p}, \hat{\theta})' \right] [I_M, -\hat{Q}]', \tag{15}$$

where I_M is the $M \times M$ identity matrix, $\hat{p}$ and $\hat{\theta}$ are consistent estimates of p and θ respectively, $\zeta_i(p, \theta) = [\psi_i(p, \theta)', h_i(p)']'$, and $Q = (\sum_i \partial \psi_i(p, \theta)/\partial p')(\sum_i \partial h_i(p)/\partial p')^{-1}$, and $h_i(p)$ is the

$M \times 1$ vector of $h_{il}^t \equiv d_{il}^t(y_{it} - p_l^t)$ for the histories actually occurred in the data. We can expand (15) as

$$\hat{W} = \frac{1}{N}\left(\sum_{i=1}^{N} \hat{\psi}_i \hat{\psi}_i' - \hat{Q}\sum_{i=1}^{N} \hat{h}_i \hat{\psi}_i' - \sum_{i=1}^{N} \hat{\psi}_i \hat{h}_i' \hat{Q}' + \hat{Q}\sum_{i=1}^{N} \hat{h}_i \hat{h}_i' \hat{Q}'\right). \tag{16}$$

To see how $\hat{W}$ can become singular, let us consider the m-th row of $\hat{W}$. Suppose that the m-th element of ψ_i is ψ_{il}^{t-1}, and that the cell for the history ϕ_l^{t-1} contains only one observation which is assumed to have $y_{it} = 1$ for all t. To simplify the expression, we further drop σ_t, γ_t, and $\alpha \Delta X_{it}$, which does not affect our argument below. Notice first that $\hat{p}_l^{t-1} = 1$ and $\hat{p}_k^t = 1$, where ϕ_k^t is assumed to be the history containing ϕ_l^{t-1}. Hence, $\hat{\psi}_{il}^{t-1} = d_{il}^{t-1}(\Phi^{-1}(\hat{p}_k^t) - \Phi^{-1}(\hat{p}_l^{t-1}) - \beta \Delta y_{i(t-1)}) = 0$, so that $\sum_i \hat{\psi}_{il}^{t-1} \hat{\psi}_i'$, the m-th row of the first term in (16), includes only zero as its element. Likewise, the m-th row of the third term in (16) contains only zero. Next, some tedious algebra shows that the m-th row of $\hat{Q}\hat{h}_i$ is given by $d_{il}^{t-1}\left(\frac{\partial \Phi^{-1}(\hat{p}_k^t)}{\partial p_k^t} \frac{d_{ik}^t}{\sum_i d_{ik}^t} \hat{h}_{ik}^t - \frac{\partial \Phi^{-1}(\hat{p}_l^{t-1})}{\partial p_l^{t-1}} \frac{d_{il}^{t-1}}{\sum_i d_{il}^{t-1}} \hat{h}_{il}^{t-1}\right) = 0$, since $\hat{h}_{ik}^t = \hat{h}_{il}^{t-1} = 0$. Accordingly, the m-th rows of the second and the fourth terms in (16) are zero. Therefore, $\hat{W}$ can be singular in this case. This possibility becomes more problematic if we wish to use the optimal weighting matrix which should be $\hat{W}^{-1}$. This further motivates dropping cells containing few observations. However, dropping cells entails additional issue discussed below.

The fourth practical issue is related to a consistent estimator of the asymptotic variance of $\sqrt{N}(\hat{\theta} - \theta)$ which is given by

$$\hat{V}_\theta = (\hat{D}_\theta' \hat{D}_\theta)^{-1} \hat{D}_\theta' \hat{W} \hat{D}_\theta (\hat{D}_\theta' \hat{D}_\theta)^{-1}, \tag{17}$$

where $\hat{D}_\theta = \frac{1}{N}\sum_{i=1}^{N} \partial \psi_i(\hat{p}, \hat{\theta})/\partial \theta'$, and we suppose that the GMM estimate $\hat{\theta}$ is estimated using $A_N = I_M$. To see the problem, consider the simplified case as above, in which $\theta = \beta$, there are many observations with $y_{it} = y_{i(t-1)}$, and many cells contain few observations. In this case, we have $\hat{D}_\theta' \hat{D}_\theta = \frac{1}{N^2}\sum_{t=2}^{T}\sum_{l=1}^{L^{t-1}}\left(\sum_{i=1}^{N} d_{il}^{t-1} \Delta y_{i(t-1)}\right)\left(\sum_{i=1}^{N} d_{il}^{t-1} \Delta y_{i(t-1)}\right)$. Because of the strong positive correlation between y_{it} and $y_{i(t-1)}$, many observations are likely to have $\Delta y_{i(t-1)} = 0$. Nevertheless, we will still have more cells with $\Delta y_{i(t-1)} \neq 0$ when using all the non-empty cells, than when dropping many cells containing few observations. However, if the remaining cells tend to have $\Delta y_{i(t-1)} = 0$ as we drop more cells with few observations, $\hat{D}_\theta'\hat{D}_\theta$ is likely to be close to zero. As a result, $\hat{D}_\theta'\hat{D}_\theta$ becomes singular or nearly singular, which suggests that we may not be able to estimate $\hat{V}_\theta$, or the estimated standard errors may be very large.[23]

The final issue is that it is not practically feasible to estimate $\Pr(y_{it} = 1|H_i^t)$ nonparametrically as a function of the history of many regressors in the net benefit function. Our approach to this issue is to estimate the model using relatively homogeneous subsamples, so that we can reduce the dimensionality in Z_{it}, a vector of variables capturing observed heterogeneity. Specifically, we consider 11 subsamples based on the NAICS code and estimate the model separately for each subsample.[24]

4 Results

4.1 Probit Results

Tables 5 and 6 report the results obtained by using conventional probit estimations (columns 1 and 2) as well as the AC method (column 3). These tables present the estimation results for the 2000-2003 complete panel. Note that we do not report the probit results for the 2000-2002 and 2001-2003 complete panels simply because the coefficient estimates for $\mathsf{server.windows}_{t-1}$ and $\mathsf{server.linux}_{t-1}$ from these samples are quite similar to those from the 2000-2003 sample. The results for the AC method using the 2000-2002 and 2001-2003 samples are discussed in Section 4.3 where we check robustness of our results.

Table 5 reports parameter estimates for Windows adoption, whereas Table 6 presents pa-

[23]To obtain more precise estimates, we can use the efficient GMM using the optimal weighting matrix $\hat{W}^{-1}$. As the third practical issue suggests, however, the presence of many cells with few observations causes difficulties in inverting $\hat{W}$. As a result, we use an identity matrix for A_N in our GMM estimation.

[24]Alternatively, one could also consider a semi-parametric approach to estimate $\Pr(y_{it}|H_i^t)$. For example, one might use a single-index model such that $\Pr(y_{it} = 1|H_i^t) = g(H_i^{t\prime}\lambda)$, where λ is a vector of coefficients corresponding to each element in H_i^t, and $g(\cdot)$ is an unknown function. Note, however, that the right hand-side of (4) is a function of H_i^t. As a result, using the single-index $H_i^{t\prime}\lambda$ implicitly imposes a structure which is unlikely to be internally consistent with the original model in (4). For this reason, we do not consider semi-parametric approach to estimate $\Pr(y_{it} = 1|H_i^t)$.

rameter estimates for Linux adoption. We group relatively homogenous samples based on kinds of business, and estimate the model separately for different subsamples. The main parameters of interest are the coefficients for $\mathsf{server.windows}_{t-1}$ in Windows adoption and $\mathsf{server.linux}_{t-1}$ in Linux adoption which potentially capture state dependence or lock-in. We also include predetermined variables for operating systems adoption in non-server segments, in a way to reflect potential network effects between servers and non-servers. Note, however, that we exclude some variables such as $\mathsf{pc.windows}_{t-1}$ and $\mathsf{non\text{-}pc.linux}_{t-1}$ for some samples, since they are almost all one (or zero), resulting in multicollinearity.

According to the results from conventional probit estimations (columns 1 and 2 in Table 5), the previous choice of Windows appears to be the most important factor to determine the current choice of Windows in server operating systems. The coefficient estimates for $\mathsf{server.windows}_{t-1}$ in Windows adoption are positive and statistically significant, and the estimated values are over 2. These results are consistent across different industries, and remain the same even after we account for observed heterogeneity by including various establishment-specific characteristics. Columns 1 and 2 in Table 6 show the same results in the case of Linux adoption. The coefficient estimates for $\mathsf{server.linux}_{t-1}$ are over 2 and statistically significant, which seems to be robust across different industries.

These results, however, only confirm strong positive correlations between the previous adoption choice and the current choice in server operating systems, because conventional probit does not take into account of unobserved preferences which can also explain why firms continued to use the same operating systems. To distinguish between lock-in and unobserved preferences, we therefore use the AC method, and the next section reports the results.

4.2 AC Method Results

The results obtained by using the AC method are reported in column 3 in Tables 5 and 6. In Windows adoption, the magnitudes of the coefficient estimates for $\mathsf{server.windows}_{t-1}$

vary across different subsamples. Nevertheless, the estimates are considerably smaller than those reported in columns 1 and 2, and they are statistically insignificant. All the coefficient estimates for $\mathsf{server.windows}_{t-1}$ therefore point to one conclusion: the previous use of Windows does not seem to be the most important factor that explains the current use of Windows across various industries. These results suggest that firms continued to use Windows more likely due to unobserved preferences for Windows, and that the adoption of Windows at the previous period did not necessarily lock firms into Windows.

The estimates for Linux adoption in Table 6 show similar results. The magnitudes of the coefficient estimates for $\mathsf{server.linux}_{t-1}$ differ across various industries. The estimates, nonetheless, are substantially smaller than those in columns 1 and 2 which are estimated using conventional probit, and they are statistically insignificant, suggesting that unobserved preferences, rather than lock-in, may explain strong positive correlations between the previous choice and the current choice in server operating systems.

Another observation from the estimates in Tables 5 and 6 is that when we use conventional probit, the coefficient estimates for $\mathsf{server.windows}_{t-1}$ in Windows adoption and $\mathsf{server.linux}_{t-1}$ in Linux adoption are quite similar across different industries. In contrast, when we account for unobserved preferences by using the AC method, these coefficients vary across industries, potentially reflecting heterogeneity in unobserved preferences. Because different experiences in using operating systems are likely to result in different opinion about server operating systems, and firms in different industries are unlikely to have similar experiences, the extent of unobserved preferences is likely to vary across industries, which is consistent with our result.

One more observation from Tables 5 and 6 is that the estimates of the asymptotic standard errors of the coefficient estimates tend to be large. As discussed in Section 3.4, however, potentially large standard errors are expected, given that the positive correlations between y_{ijt} and $y_{ij(t-1)}$ are strong in our data. To obtain smaller standard errors, one could attempt to use efficient GMM. However, as we discuss in Section 3.4, the optimal weighting matrix does not seem to be feasible to compute in our case because of the singularity in the covariance matrix.

4.3 Robustness Check

The main results obtained by using the AC method are that the coefficient estimates for $\mathsf{server.windows}_{t-1}$ in Windows adoption and $\mathsf{server.linux}_{t-1}$ in Linux adoption are statistically insignificant and considerably smaller than those obtained from conventional probit, and that their estimated values vary across different industries, potentially reflecting heterogeneity in unobserved preferences for server operating systems. In this section, we check robustness of these results with respect to sample restrictions and estimation procedures.

Recall that the restrictions in Section 2.2 reduce the sample size in the 2000-2003 complete panel. For this reason, we consider shorter panels with more observations and estimate the model using the same AC method. Table 7 presents the results for the 2000-2002 and 2001-2003 data, in addition to the 2000-2003 data. Panel A reports the coefficient estimates for $\mathsf{server.windows}_{t-1}$ in Windows adoption, and panel B reports the estimates for $\mathsf{server.linux}_{t-1}$ in Linux adoption. The table shows that the estimated values are not the same across different samples. The main results remain the same, nevertheless, in that most coefficient estimates are statistically insignificant and substantially smaller than those from conventional probit, and they also vary across industries.

In terms of estimation procedures, there are several practical issues in applying the AC method as discussed in Section 3.4. The key problem is related to dropping cells with very few observations. In this regard, we consider various cutoff numbers for dropping cells. Tables 8 and 9 report the coefficient estimates using different cutoffs. The coefficient estimates are not the same, as we change the cutoffs. Nonetheless, most coefficient estimates are statistically insignificant and much smaller than those from conventional probit. Therefore, the main results do not change regardless of key changes in our estimation procedures.

Still, one may worry that the results in our robustness check appear to suggest that our estimates are sensitive to sample selection and estimation procedures. Note, however, that firms in our model form their unobserved preferences based on their past experiences. Different samples or different moment conditions conditional on different histories imply that the degree of unobserved preferences is likely to vary considerably as we change our samples or drop more cells. To the extent that unobserved preferences are important in the adoption decision, we should expect that the coefficient estimates for $y_{ij(t-1)}$ vary across different samples and may change as we drop more cells. Therefore, the results from robustness check are in fact consistent with our main findings that unobserved preferences are important.

Additionally, we note that as we increase the cutoff numbers, the estimates of the asymptotic standard errors tend to become larger, and for some samples, they cannot be computed because $\hat{D}_\theta' \hat{D}_\theta$ in (17) is singular. This is consistent with our discussion in Section 3.4.

5 Conclusion

In this paper, we use detailed establishment-level data on the adoption of server operating systems, and examine two leading factors – lock-in and unobserved preferences – in the decision to adopt Windows or Linux as operating systems for server computers. To distinguish between these two factors, we construct our model based on plausible assumptions about unobserved preferences. The model generates conditional moment restrictions and also allows us to use a linear GMM estimator. Accordingly, we implement a GMM version of the semiparametric discrete choice panel data methods developed by Arellano and Carrasco (2003).

Without accounting for unobserved preferences, we find a seemingly robust positive correlation between the current choice and the previous choice. Once unobserved preferences are taken into account, however, we find that the coefficient estimates for the previous choice become considerably smaller and statistically insignificant, suggesting that unobserved preferences, rather than lock-in, are more important in firms' decisions to adopt server operating systems. Though the magnitudes of these coefficient estimates are small, we also find that the estimated values vary across different subsamples, likely reflecting heterogeneity in unobserved preferences. These findings are robust to sample restrictions and estimation procedures.

Our data include many establishments across various regions and sectors. However, they are not intended to be representative of all establishments in the United States. Consequently, we do not attempt to generalize our findings beyond the samples examined in this paper. For this reason, it would be interesting to investigate the degrees of lock-in and unobserved preferences by using other sources of datasets. Lastly, our findings also motivate further structural modeling to shed light on what underlies unobserved preferences, and what other factors are important in the adoption of operating systems.

References

Arellano, M. and Honoré, B. "Panel Data Models: Some Recent Developments." *Handbook of Econometrics, Vol. 5*, edited by J. Heckman and E. Leamer. Amsterdam: North-Holland, pp. 3229-3296.

Arellano, M. and Carrasco, R. "Binary Choice Panel Data Model with Predetermined Variables." *Journal of Econometrics*, 2003, *115*, pp. 125-57.

Bresnahan, T. "Network Effects and Microsoft." Manuscript. Stanford University, 2001.

Bresnahan, T and Greenstein, S. "Technical Progress and Co-invention in Computing and in the Uses of Computers." *Brookings Papers on Economic Activity: Microeconomics*, 1996, pp. 1-77.

Bresnahan, T., Brynjolfsson, E., and Hitt, L. "Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence" *Quarterly Journal of Economics*, 2002, *117*, pp. 339-376.

Fink, M. *The Business and Economics of Linux and Open Source*, New Jersey: Prentice Hall PTR, 2002.

Heckman, J. "Statistical Models for Discrete Panel Data." *Structural Analysis of Discrete Panel Data with Econometric Applications*, edited by C. Manski and D. McFadden. Cambridge, MA: MIT Press. 1981a, pp. 114-78.

Heckman, J. "The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process." *Structural Analysis* *of Discrete Panel Data with Econometric Applications*, edited by C. Manski and D. McFadden. Cambridge, MA: MIT Press. 1981b, pp. 179-95.

Heckman, J. "Heterogeneity and State Dependence." *Studies in Labor Markets*, edited by S. Rosen. University of Chicago Press. 1981c, pp. 91-139.

Honoré, B. and Kyriazidou, E. "Panel Data Discrete Choice Models with Lagged Dependent Variables." *Econometrica*, 2000, *70*, pp. 839-74.

Hsiao, C. *Analysis of Panel Data.* Cambridge, United Kingdom: Cambridge University Press. 2003.

Liebowitz, S. and Margolis, S. *Winners, Losers, & Microsoft: Competition and Antitrust in HIgh Technology.* The Independent Institute. 1999.

Varian, H. and Shapiro, C. "Linux Adoption in the Public Sector: An Economic Analysis." Manuscript. University of California, Berkeley, 2003.

Wooldridge, J. "Simple Solutions to the Initial Conditions Problem in Dynamic, Nonlinear Panel Data Models with Unobserved Heterogeneity." *Journal of Applied Econometrics*, 2005, *20*, pp. 39-54.

Table 1: Summary Statistics of Variables for Each Year[a]

Year	2000	2001	2002	2003
	(1)	(2)	(3)	(4)
server.windows[b]	0.46	0.44	0.51	0.52
server.linux	0.04	0.05	0.06	0.07
server.other	0.13	0.13	0.13	0.12
internet.sv.windows	0.13	0.15	0.17	0.15
internet.sv.linux	0.02	0.02	0.03	0.03
internet.sv.other	0.03	0.03	0.03	0.03
network.sv.windows	0.40	0.37	0.45	0.46
network.sv.linux	0.02	0.03	0.04	0.05
network.sv.other	0.11	0.11	0.11	0.11
pc.windows	0.84	0.85	0.88	0.90
pc.linux	0.04	0.07	0.08	0.10
pc.other	0.09	0.06	0.06	0.05
non-pc.windows	0.02	0.02	0.04	0.03
non-pc.linux	0.00	0.00	0.00	0.00
non-pc.other	0.28	0.21	0.20	0.16
perl[c]	0.01	0.01	0.01	0.01
apache[d]	0.02	0.03	0.03	0.03
#pc[e]	159.17	158.54	168.22	176.32
#non-pc	2.71	2.27	2.29	1.97
#Internet.server	0.53	0.64	0.85	0.85
#network.server	4.79	4.74	4.96	5.14
#pc.server	4.25	4.41	4.96	5.17
#employees[f]	316.25	299.01	298.42	297.53
#white.collar.workers	175.51	167.19	172.24	174.04
#desk.workers	137.75	129.85	134.85	133.73
#programmers	3.10	3.31	3.23	3.10
#it.workers	n/a	4.61	6.18	6.81
#Internet.users	61.08	68.63	72.14	77.27
#Internet.developers	0.60	0.65	0.70	0.73
revenue[g] (in $million)	68.90	64.28	62.05	60.50
#observations	120,880	124,324	120,984	121,324

[a]The table reports the mean of each variable.
[b]Dummy equal to 1 if Windows is installed on either Internet server or network server.
[c]Dummy equal to 1 if Perl is installed on any computer in the establishment.
[d]Dummy equal to 1 if Apache is installed on any computer in the establishment.
[e]Total number of PCs that are not used for any server.
[f]Total number of employees in the establishment.
[g]The amount of revenue for each establishment estimated by Harte-Hanks.

Table 2: Yearly Market Share of Each Operating System for Different Samples[a]

Year	server. windows	server. linux	server. other	pc. windows	pc. linux	pc. other	non-pc. windows	non-pc. linux	non-pc. other	#obs.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
		A. Unrestricted sample								
2000	0.456	0.036	0.133	0.839	0.044	0.087	0.021	0.003	0.279	120,880
2001	0.444	0.050	0.133	0.854	0.066	0.062	0.025	0.004	0.211	124,324
2002	0.510	0.064	0.133	0.878	0.085	0.057	0.036	0.005	0.196	120,984
2003	0.517	0.074	0.124	0.896	0.099	0.050	0.034	0.005	0.162	121,324
		B. Sample with any server operating system								
2000	0.914	0.072	0.267	0.920	0.071	0.116	0.031	0.005	0.380	60,282
2001	0.881	0.099	0.265	0.897	0.103	0.080	0.036	0.006	0.288	62,641
2002	0.897	0.112	0.234	0.909	0.121	0.073	0.051	0.007	0.259	68,786
2003	0.908	0.130	0.218	0.918	0.136	0.063	0.047	0.007	0.216	69,087
		C. Sample with any server o/s and with up-to-date information								
2000	0.914	0.072	0.267	0.920	0.071	0.116	0.031	0.005	0.380	60,282
2001	0.878	0.103	0.258	0.978	0.116	0.081	0.040	0.006	0.274	37,065
2002	0.903	0.117	0.229	0.984	0.135	0.076	0.060	0.007	0.272	46,001
2003	0.911	0.140	0.219	0.982	0.155	0.062	0.047	0.008	0.205	40,119
		D. Complete panel: 2000-2003[b]								
2000	0.923	0.081	0.283	0.953	0.072	0.122	0.032	0.003	0.413	11,010
2001	0.926	0.111	0.262	0.983	0.125	0.089	0.043	0.004	0.302	11,010
2002	0.936	0.140	0.246	0.989	0.151	0.076	0.064	0.007	0.293	11,010
2003	0.938	0.159	0.234	0.984	0.175	0.059	0.049	0.008	0.203	11,010
		E. Complete panel: 2000-2002								
2000	0.920	0.078	0.280	0.949	0.071	0.117	0.031	0.003	0.406	18,061
2001	0.925	0.106	0.260	0.982	0.122	0.086	0.043	0.004	0.299	18,061
2002	0.933	0.133	0.245	0.988	0.148	0.073	0.064	0.007	0.293	18,061
		F. Complete panel: 2001-2003								
2001	0.900	0.108	0.262	0.980	0.116	0.093	0.040	0.005	0.301	18,435
2002	0.923	0.135	0.246	0.987	0.148	0.079	0.064	0.008	0.283	18,435
2003	0.929	0.155	0.234	0.984	0.172	0.062	0.050	0.008	0.204	18,435

[a] To compute the market share, a dummy for each operating system is assigned to each observation, and the table reports its mean. Because an establishment can have more than one kind of operating system, the sum of columns (1)-(3) can be larger than 1.

[b] Complete panels include only those with any server operating system and also with up-to-date information.

Table 3: Switching Patterns of the Adoption of Operating Systems in Servers[a]

Industry	Year	[Linux_{t-1}, Linux_t] [0,1] (1)	[1,1] (2)	[1,0] (3)	[Windows_{t-1}, Windows_t] [0,1] (4)	[1,1] (5)	[1,0] (6)	[Other_{t-1}, Other_t] [0,1] (7)	[1,1] (8)	[1,0] (9)	total obs.
A. All Industries and By Year											
	2001	0.041	0.064	0.014	0.023	0.899	0.019	0.030	0.226	0.049	27,578
	2002	0.046	0.080	0.019	0.041	0.873	0.020	0.038	0.207	0.057	37,244
	2003	0.046	0.095	0.022	0.034	0.887	0.019	0.034	0.190	0.050	35,968
B. All Industries and All Years											
		0.045	0.081	0.019	0.034	0.885	0.019	0.034	0.206	0.052	100,790
C. By Industry and All Years											
agri_utility.1-2[b]		0.030	0.046	0.016	0.026	0.910	0.017	0.031	0.153	0.046	4,453
manufacture.31-32		0.026	0.042	0.012	0.028	0.904	0.017	0.029	0.172	0.045	10,211
manufacture.33		0.038	0.060	0.016	0.027	0.906	0.019	0.030	0.194	0.050	14,624
retail.4		0.035	0.060	0.017	0.036	0.865	0.019	0.033	0.206	0.046	7,856
information.51		0.074	0.180	0.028	0.048	0.789	0.032	0.047	0.302	0.058	7,071
financial.52-53		0.031	0.043	0.012	0.032	0.903	0.013	0.027	0.153	0.049	7,063
professional.54-56		0.049	0.095	0.020	0.030	0.898	0.017	0.031	0.165	0.050	9,989
education.61		0.074	0.159	0.031	0.048	0.846	0.027	0.050	0.303	0.067	12,674
medical.62		0.039	0.055	0.016	0.029	0.913	0.013	0.030	0.177	0.050	8,589
arts_service.6-7		0.034	0.054	0.017	0.031	0.890	0.022	0.029	0.158	0.049	2,554
government		0.045	0.065	0.017	0.031	0.898	0.016	0.033	0.217	0.054	13,283

[a]The table reports the share of the establishments that follow the switching pattern specified in each column. The sample includes only those that reported information on server operating systems. The sample also excludes the observations whose information on computing technology was outdated.

[b]The number after the name of industry denotes the first digit (or first two digits) of NAICS.

Table 4: Switching between Different Operating Systems in Servers[a]

Industry	Year	$\text{Windows}_{t-1}=1$, $\text{Windows}_t=0$		$\text{Linux}_{t-1}=1$, $\text{Linux}_t=0$		$\text{Other}_{t-1}=1$, $\text{Other}_t=0$			
		$L_{t-1}=1$, $L_t=1$	$L_{t-1}=0$, $L_t=1$	$W_{t-1}=1$, $W_t=1$	$W_{t-1}=0$, $W_t=1$	$L_{t-1}=1$, $L_t=1$	$L_{t-1}=0$, $L_t=1$	$W_{t-1}=1$, $W_t=1$	$W_{t-1}=0$, $W_t=1$
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. All Industries and By Year									
	2001	0.0030	0.0037	0.0111	0.0019	0.0041	0.0048	0.0387	0.0073
	2002	0.0036	0.0032	0.0123	0.0050	0.0045	0.0049	0.0375	0.0167
	2003	0.0040	0.0023	0.0158	0.0041	0.0054	0.0035	0.0323	0.0159
B. All Industries and All Years									
		0.0036	0.0030	0.0132	0.0038	0.0047	0.0044	0.0360	0.0139
C. By Industry and All Years									
agri_utility.1-2		0.0027	0.0022	0.0115	0.0040	0.0038	0.0040	0.0305	0.0121
manufacture.31-32		0.0020	0.0027	0.0075	0.0037	0.0024	0.0018	0.0303	0.0131
manufacture.33		0.0025	0.0030	0.0118	0.0031	0.0030	0.0031	0.0360	0.0127
retail.4		0.0032	0.0028	0.0115	0.0045	0.0029	0.0037	0.0298	0.0151
information.51		0.0098	0.0048	0.0174	0.0054	0.0109	0.0081	0.0368	0.0139
financial.52-53		0.0010	0.0014	0.0078	0.0034	0.0024	0.0034	0.0317	0.0156
professional.54-56		0.0041	0.0035	0.0154	0.0029	0.0048	0.0049	0.0355	0.0125
education.61		0.0067	0.0043	0.0237	0.0043	0.0110	0.0084	0.0474	0.0155
medical.62		0.0015	0.0028	0.0095	0.0045	0.0022	0.0041	0.0353	0.0141
arts_service.6-7		0.0031	0.0043	0.0110	0.0043	0.0035	0.0035	0.0301	0.0172
government		0.0029	0.0017	0.0125	0.0032	0.0032	0.0035	0.0390	0.0137

[a]The table reports the share of the establishments that follow the switching pattern specified in each column. The sample includes only those that reported information on server operating systems. The sample also excludes the observations whose information on computing technology was outdated. Note that L_t (or W_t) denotes the use of Linux (or Windows) at period t.

Table 5: Results for Windows Adoption from 2000-2003 Complete Panel

Variable	Probit				AC Method[a]	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
	(1)		(2)		(3)	
	A. NAICS 1-2: Agriculture, Utility, and Construction					
server.linux$_{t-1}$	-0.2400	0.2696	-0.5097	0.3078	0.2347	0.3608
server.windows$_{t-1}$	**2.1137**	**0.2095**	**1.9766**	**0.2308**	**0.3346**	**0.9682**
server.other$_{t-1}$	-0.7869	0.1743	-0.9168	0.2041	-0.1886	0.5102
pc.linux$_{t-1}$	-0.1149	0.2725	-0.3150	0.3110	-0.3940	0.6377
non-pc.other$_{t-1}$	0.0558	0.1697	-0.0624	0.1879	0.0700	0.5029
γ_{2001}	0.2487	0.2593	0.1869	0.3234	0.4986	0.2728
γ_{2002}	0.7109	0.2642	0.7263	0.3265	0.8674	1.1479
γ_{2003}	0.1411	0.2625	0.0414	0.3244	-0.6875	0.1497
	B. NAICS 31-32: Manufacturing					
server.linux$_{t-1}$	-0.5008	0.1726	-0.5263	0.1980	-0.0437	1.1261
server.windows$_{t-1}$	**2.1847**	**0.1239**	**2.0919**	**0.1332**	**1.2393**	**1.6234**
server.other$_{t-1}$	-0.4742	0.1110	-0.5019	0.1207	0.0878	1.1299
pc.linux$_{t-1}$	0.2031	0.1972	0.2373	0.2173	0.2750	1.4358
pc.windows$_{t-1}$	0.5063	0.2594	0.6117	0.2822	0.9504	0.7034
pc.other$_{t-1}$	0.1175	0.1691	0.0831	0.1775	-0.5589	0.5622
non-pc.windows$_{t-1}$	0.0394	0.2412	-0.2629	0.2577	-0.8566	1.2843
non-pc.other$_{t-1}$	-0.0244	0.0996	-0.0958	0.1078	-0.2896	0.4655
γ_{2001}	-0.4670	0.2908	-0.5999	0.3330	0.9501	0.3249
γ_{2002}	-0.5218	0.2960	-0.6655	0.3366	0.0316	0.2132
γ_{2003}	-0.2289	0.2970	-0.3337	0.3352	-0.0868	0.2355
	C. NAICS 33: Manufacturing					
server.linux$_{t-1}$	-0.5034	0.1150	-0.5635	0.1263	-0.3003	0.7235
server.windows$_{t-1}$	**2.1623**	**0.0990**	**2.1013**	**0.1027**	**0.3917**	**0.8205**
server.other$_{t-1}$	-0.5233	0.0870	-0.5675	0.0907	-0.5093	0.4769
pc.linux$_{t-1}$	0.1642	0.1330	0.2031	0.1423	0.4164	0.3801
pc.windows$_{t-1}$	0.1292	0.2641	0.2091	0.2628	0.0409	0.7717
pc.other$_{t-1}$	-0.1372	0.1537	-0.1786	0.1602	-0.3516	0.4151
non-pc.windows$_{t-1}$	-0.0104	0.1456	-0.0792	0.1509	-0.4431	1.6330
non-pc.other$_{t-1}$	-0.1034	0.0816	-0.1268	0.0856	0.0242	0.4098
γ_{2001}	0.0368	0.2787	-0.0588	0.2942	0.8127	0.2673
γ_{2002}	0.1156	0.2885	0.0279	0.3030	-0.0531	0.1397
γ_{2003}	0.0542	0.2924	-0.0330	0.3068	0.6459	0.5961
	D. NAICS 4: Wholesale, Retail, and Transportation					
server.linux$_{t-1}$	-0.2247	0.1923	-0.2996	0.2055	-0.6794	1.7110
server.windows$_{t-1}$	**2.3338**	**0.1458**	**2.2925**	**0.1515**	**0.6732**	**2.0650**
server.other$_{t-1}$	-0.6339	0.1298	-0.6647	0.1340	-0.5637	0.8612
pc.linux$_{t-1}$	0.1018	0.2070	0.0152	0.2140	0.0465	1.3663
pc.other$_{t-1}$	-0.1677	0.2448	-0.2530	0.2550	-1.1194	0.6952
non-pc.other$_{t-1}$	-0.0778	0.1222	-0.1020	0.1268	0.5359	0.8228
γ_{2001}	0.0778	0.1901	-0.0351	0.2569	0.7657	0.4858
γ_{2002}	0.0184	0.1920	-0.1014	0.2569	0.5516	0.8608
γ_{2003}	0.0044	0.1965	-0.1113	0.2613	0.0002	0.0738
additional control[b]	No		Yes		No	

[a]We drop the sample orthogonality conditions if the number of observations in a cell is less than one. Table 8 reports the results using different cutoffs for dropping a cell with few observations.

[b]Additional control includes revenue, #it.workers, #programmers, #desk.workers, apache, #pc, #non-pc, #Internet.server, #network.server, #pc.server, and dummies for population where establishments are located.

Table 5: Results for Windows Adoption (Continued)

Variable	Probit				AC Method	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
	(1)		(2)		(3)	
	E. NAICS 51: Information					
server.linux$_{t-1}$	-0.4975	0.1195	-0.4025	0.1325	-0.5576	2.4033
server.windows$_{t-1}$	**2.2486**	**0.1018**	**2.1982**	**0.1044**	**0.2241**	**1.5619**
server.other$_{t-1}$	-0.4048	0.1056	-0.4216	0.1089	-1.1356	1.1778
pc.linux$_{t-1}$	-0.0230	0.1228	-0.0352	0.1296	0.1494	1.9680
pc.other$_{t-1}$	-0.2789	0.1088	-0.3080	0.1136	-0.3684	1.1487
non-pc.other$_{t-1}$	0.1057	0.1048	-0.0039	0.1096	0.6916	0.6093
γ_{2001}	-0.2052	0.1442	-0.3748	0.1804	1.3414	2.0571
γ_{2002}	-0.1170	0.1439	-0.3213	0.1805	0.0002	0.2408
γ_{2003}	-0.0914	0.1468	-0.3025	0.1818	0.1822	0.2902
	F. NAICS 52-53: Finance and Real Estate					
server.linux$_{t-1}$	-0.4957	0.2618	-0.6325	0.2872	-0.8266	2.8209
server.windows$_{t-1}$	**2.1608**	**0.1913**	**2.1177**	**0.2054**	**0.1336**	**2.9960**
server.other$_{t-1}$	-0.5334	0.1696	-0.5960	0.1918	-0.1110	1.0768
pc.linux$_{t-1}$	-0.1447	0.2494	-0.1239	0.2925	-0.0736	0.9427
pc.windows$_{t-1}$	0.3811	0.3964	0.6651	0.4296	0.7105	0.5504
pc.other$_{t-1}$	-0.3540	0.3363	-0.2686	0.3802	-1.0831	0.5003
non-pc.other$_{t-1}$	0.2010	0.1533	0.1317	0.1659	-0.0850	0.5021
γ_{2001}	-0.1662	0.4276	-0.4036	0.5041	0.7028	0.3482
γ_{2002}	-0.1930	0.4391	-0.4370	0.5136	-0.2082	0.1219
γ_{2003}	-0.2039	0.4425	-0.3744	0.5169	0.3931	0.3215
	G. NAICS 54-56: Professional and Technical Services					
server.linux$_{t-1}$	-0.5688	0.1326	-0.7029	0.1508	0.1869	1.1036
server.windows$_{t-1}$	**2.1411**	**0.1283**	**2.0749**	**0.1351**	**0.4233**	**1.2557**
server.other$_{t-1}$	-0.6060	0.1212	-0.6916	0.1321	-0.6095	1.0794
pc.linux$_{t-1}$	-0.2390	0.1413	-0.2367	0.1495	-0.0154	0.9185
pc.windows$_{t-1}$	0.4084	0.2896	0.4214	0.3009	0.0009	1.7326
pc.other$_{t-1}$	-0.4057	0.1619	-0.4051	0.1682	-0.3417	0.7359
non-pc.windows$_{t-1}$	0.2611	0.2895	0.1900	0.3252	0.1944	0.6874
non-pc.other$_{t-1}$	0.1031	0.1232	0.0872	0.1333	-0.6277	0.8752
γ_{2001}	-0.1600	0.3107	-0.0448	0.3877	1.5533	1.2513
γ_{2002}	-0.0178	0.3232	0.1351	0.3955	-0.3126	0.1724
γ_{2003}	-0.0648	0.3252	0.0915	0.3988	-0.0768	0.2068
	H. NAICS 61: Educational Services					
server.linux$_{t-1}$	-0.3368	0.0792	-0.4037	0.0854	0.7710	1.1220
server.windows$_{t-1}$	**2.1196**	**0.0774**	**2.0707**	**0.0799**	**-0.0082**	**1.1219**
server.other$_{t-1}$	-0.3487	0.0759	-0.4059	0.0794	-0.8678	0.7999
pc.linux$_{t-1}$	0.1069	0.0908	0.0267	0.0939	-0.5691	1.3190
pc.windows$_{t-1}$	0.4299	0.1169	0.4319	0.1186	0.9639	1.1268
pc.other$_{t-1}$	0.0388	0.0787	0.0393	0.0810	-0.4217	0.6917
non-pc.other$_{t-1}$	0.1734	0.0801	0.0888	0.0850	-0.0591	0.3889
γ_{2001}	-0.5306	0.1436	-0.5162	0.1536	1.5476	1.3182
γ_{2002}	-0.3427	0.1496	-0.3128	0.1591	-0.0484	0.1374
γ_{2003}	-0.4502	0.1537	-0.4310	0.1632	-0.3870	0.1946
additional control	No		Yes		No	

Table 5: Results for Windows Adoption (Continued)

Variable	Probit				AC Method	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
	(1)		(2)		(3)	
	I. NAICS 62: Health Care					
server.linux$_{t-1}$	-0.2452	0.1756	-0.2213	0.1878	-0.4965	1.9200
server.windows$_{t-1}$	**2.3643**	**0.1547**	**2.3603**	**0.1638**	**0.5620**	**1.1913**
server.other$_{t-1}$	-0.4665	0.1308	-0.5420	0.1384	-0.3839	0.3157
pc.linux$_{t-1}$	-0.3100	0.1789	-0.3499	0.1966	-0.2737	1.0358
pc.windows$_{t-1}$	0.1951	0.3657	0.1727	0.4199	-0.0190	2.0328
pc.other$_{t-1}$	-0.1320	0.2459	-0.2082	0.2637	-0.5652	0.4995
non-pc.windows$_{t-1}$	0.3188	0.4237	0.3380	0.4553	-1.3198	0.3861
non-pc.other$_{t-1}$	-0.0990	0.1280	-0.1947	0.1427	0.2183	0.5326
γ_{2001}	-0.1606	0.3777	-0.1906	0.4406	0.9475	0.4020
γ_{2002}	0.0076	0.4017	-0.0564	0.4637	0.0004	0.2077
γ_{2003}	-0.0106	0.4084	-0.0822	0.4697	0.2981	0.5103
	J. NAICS 7-8: Arts, Entertainment, and Other Services					
server.linux$_{t-1}$	-0.3094	0.3350	-0.2624	0.3859	-0.1906	1.7818
server.windows$_{t-1}$	**2.4221**	**0.3181**	**2.7543**	**0.3923**	**-0.4006**	**2.0170**
server.other$_{t-1}$	-0.0447	0.2763	0.2093	0.3350	-0.3050	1.3032
pc.linux$_{t-1}$	-0.0206	0.3833	0.2345	0.4475	0.1024	0.7892
pc.other$_{t-1}$	0.6714	0.5965	0.6851	0.6714	-0.1593	0.6821
non-pc.other$_{t-1}$	0.3860	0.2781	0.3992	0.3206	0.0039	0.8322
γ_{2001}	-0.7061	0.3781	-0.5748	0.5365	0.1739	0.3552
γ_{2002}	-0.2756	0.3717	-0.0874	0.5440	0.3716	0.4918
γ_{2003}	-0.1466	0.3745	0.0805	0.5528	0.7662	1.7259
	K. NAICS 9: Public Administration					
server.linux$_{t-1}$	-0.3623	0.1085	-0.3667	0.1148	-0.1131	1.6549
server.windows$_{t-1}$	**2.2756**	**0.0896**	**2.2323**	**0.0925**	**1.2930**	**1.1326**
server.other$_{t-1}$	-0.6647	0.0844	-0.6886	0.0867	-0.7279	0.5209
pc.linux$_{t-1}$	-0.0570	0.1142	-0.0799	0.1194	-0.5443	1.7893
pc.windows$_{t-1}$	0.2253	0.2259	0.3430	0.2296	0.2014	0.5007
pc.other$_{t-1}$	0.0835	0.1565	0.0314	0.1631	0.0094	0.7401
non-pc.windows$_{t-1}$	0.3114	0.1822	0.1924	0.1910	-0.9059	1.2281
non-pc.other$_{t-1}$	-0.0148	0.0786	-0.0489	0.0818	0.1757	0.2997
γ_{2001}	-0.1203	0.2398	-0.2557	0.2568	0.7601	0.1748
γ_{2002}	0.0335	0.2455	-0.1033	0.2621	0.5906	0.5208
γ_{2003}	-0.1724	0.2471	-0.3187	0.2646	-0.2278	0.1518
additional control	No		Yes		No	

Table 6: Results for Linux Adoption from 2000-2003 Complete Panel

Variable	Probit				AC Method[a]	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
	(1)		(2)		(3)	
	A. NAICS 1-2: Agriculture, Utility, and Construction					
server.linux$_{t-1}$	**2.4462**	**0.1699**	**2.3363**	**0.1790**	**0.4708**	**2.7302**
server.windows$_{t-1}$	0.7208	0.3325	0.5816	0.3289	-0.5417	3.5707
server.other$_{t-1}$	0.4711	0.1399	0.4369	0.1567	-0.0811	0.9249
pc.linux$_{t-1}$	0.8031	0.1724	0.8092	0.1792	0.2501	1.8975
non-pc.other$_{t-1}$	0.0978	0.1303	0.0441	0.1405	0.1332	0.4096
γ_{2001}	-2.9677	0.3589	-2.8164	0.3684	-0.6044	0.3128
γ_{2002}	-2.6777	0.3475	-2.5431	0.3551	-0.0310	0.2272
γ_{2003}	-2.8198	0.3642	-2.6649	0.3718	-0.6511	0.9045
	B. NAICS 31-32: Manufacturing					
server.linux$_{t-1}$	**2.3700**	**0.1306**	**2.2591**	**0.1380**	**0.6690**	**1.8449**
server.windows$_{t-1}$	-0.1815	0.1518	-0.2321	0.1638	0.3232	0.6683
server.other$_{t-1}$	0.3081	0.1015	0.2122	0.1127	0.0217	0.6340
pc.linux$_{t-1}$	0.8513	0.1285	0.8108	0.1373	-0.2063	0.5235
pc.windows$_{t-1}$	-0.1106	0.3103	0.1258	0.3607	0.6242	0.6470
pc.other$_{t-1}$	0.1360	0.1540	0.1225	0.1681	-0.1818	0.4789
non-pc.windows$_{t-1}$	0.4346	0.1679	0.4231	0.1836	0.4367	1.1497
non-pc.other$_{t-1}$	0.0942	0.0933	0.0888	0.1010	-0.0806	0.1606
γ_{2001}	-1.8947	0.3321	-2.2070	0.3872	-0.3834	0.1554
γ_{2002}	-1.8543	0.3369	-2.1662	0.3936	-0.5788	0.6278
γ_{2003}	-1.9292	0.3366	-2.1920	0.3902	0.5114	0.1834
	C. NAICS 33: Manufacturing					
server.linux$_{t-1}$	**2.3072**	**0.0819**	**2.1901**	**0.0863**	**0.4901**	**1.6051**
server.windows$_{t-1}$	0.1082	0.1246	0.0334	0.1330	-0.0678	0.8370
server.other$_{t-1}$	0.2470	0.0696	0.2048	0.0747	-0.0126	0.1597
pc.linux$_{t-1}$	0.7528	0.0819	0.7035	0.0860	-0.1949	0.6754
pc.windows$_{t-1}$	-0.2233	0.2453	-0.1543	0.2654	-0.4300	0.6909
pc.other$_{t-1}$	0.2053	0.1275	0.1467	0.1363	-0.4742	0.6055
non-pc.windows$_{t-1}$	0.1324	0.1094	0.0934	0.1157	-0.2713	1.1558
non-pc.other$_{t-1}$	0.0882	0.0639	0.0406	0.0692	0.1101	0.3544
γ_{2001}	-1.7877	0.2647	-1.9231	0.2946	-0.4295	0.2179
γ_{2002}	-1.7485	0.2714	-1.9012	0.3008	-0.0940	0.1367
γ_{2003}	-1.8741	0.2738	-1.9485	0.3030	-0.0217	0.1633
	D. NAICS 4: Wholesale, Retail, and Transportation					
server.linux$_{t-1}$	**2.1845**	**0.1263**	**2.0777**	**0.1318**	**0.8492**	**1.6556**
server.windows$_{t-1}$	0.0534	0.1645	-0.0212	0.1721	0.3706	1.0754
server.other$_{t-1}$	0.5155	0.1048	0.4638	0.1112	0.2723	0.2463
pc.linux$_{t-1}$	0.7675	0.1276	0.7075	0.1340	0.0821	2.0667
pc.other$_{t-1}$	-0.2636	0.2621	-0.3489	0.2856	-0.2731	0.7430
non-pc.other$_{t-1}$	0.1361	0.0975	0.1641	0.1033	0.1862	0.6439
γ_{2001}	-2.1205	0.1922	-2.2946	0.2524	-0.7405	0.5744
γ_{2002}	-2.1630	0.1939	-2.3271	0.2533	-0.1870	0.3861
γ_{2003}	-1.9746	0.1928	-2.1259	0.2532	-0.1967	0.7637
additional control[b]	No		Yes		No	

[a]We drop the sample orthogonality conditions if the number of observations in a cell is less than one. Table 9 reports the results using different cutoffs for dropping a cell with few observations.

[b]Additional control includes revenue, #it.workers, #programmers, #desk.workers, apache, #pc, #non-pc, #Internet.server, #network.server, #pc.server, and dummies for population where establishments are located.

Table 6: Results for Linux Adoption (Continued)

Variable	Probit				AC Method	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
	(1)		(2)		(3)	
	E. NAICS 51: Information					
server.linux$_{t-1}$	**2.4257**	**0.1029**	**2.2381**	**0.1088**	**0.3410**	**2.3016**
server.windows$_{t-1}$	-0.0945	0.1148	-0.0293	0.1214	-0.2352	1.9770
server.other$_{t-1}$	0.1891	0.0903	0.1456	0.0945	0.4268	0.9426
pc.linux$_{t-1}$	0.5573	0.0954	0.4402	0.0997	0.3949	1.3750
pc.other$_{t-1}$	0.0187	0.1053	0.0318	0.1111	-0.0487	1.0759
non-pc.other$_{t-1}$	0.0997	0.0884	0.1238	0.0937	-0.2579	1.1474
γ_{2001}	-1.6606	0.1456	-1.8918	0.1821	-0.9637	0.7164
γ_{2002}	-1.6480	0.1456	-1.8566	0.1825	0.2912	0.3641
γ_{2003}	-1.4767	0.1438	-1.6348	0.1771	-0.2066	0.3990
	F. NAICS 52-53: Finance and Real Estate					
server.linux$_{t-1}$	**2.5984**	**0.1780**	**2.5967**	**0.1889**	**0.1697**	**2.6402**
server.windows$_{t-1}$	0.0211	0.2386	-0.0077	0.2644	0.1740	0.9584
server.other$_{t-1}$	0.4080	0.1381	0.1754	0.1616	0.1253	0.8734
pc.linux$_{t-1}$	0.9052	0.1604	0.7169	0.1745	-0.0773	2.1277
pc.windows$_{t-1}$	-0.4127	0.3759	-0.3540	0.4123	0.0578	0.8361
pc.other$_{t-1}$	0.0830	0.3013	-0.0958	0.3377	0.3148	1.1802
non-pc.other$_{t-1}$	0.0472	0.1173	-0.0050	0.1305	-0.0199	0.2843
γ_{2001}	-1.7912	0.4053	-2.1278	0.4836	-0.0905	0.1875
γ_{2002}	-1.6484	0.4209	-1.9882	0.4973	0.1116	0.1115
γ_{2003}	-1.6554	0.4211	-2.0104	0.4988	-0.8208	1.1072
	G. NAICS 54-56: Professional and Technical Services					
server.linux$_{t-1}$	**2.3657**	**0.0968**	**2.2186**	**0.1013**	**0.1972**	**1.5065**
server.windows$_{t-1}$	-0.1857	0.1449	-0.2730	0.1547	-0.3436	1.2056
server.other$_{t-1}$	0.2133	0.0950	0.0323	0.1051	0.1449	1.1885
pc.linux$_{t-1}$	0.5903	0.0987	0.5547	0.1036	0.3032	2.1058
pc.windows$_{t-1}$	-0.0299	0.2827	0.1335	0.3058	-0.3010	1.5604
pc.other$_{t-1}$	0.0334	0.1412	-0.0287	0.1513	0.5460	0.6399
non-pc.windows$_{t-1}$	-0.1224	0.1655	-0.1588	0.1774	-0.3747	3.1179
non-pc.other$_{t-1}$	0.1808	0.0876	0.0671	0.0952	0.3613	0.4263
γ_{2001}	-1.5799	0.3060	-1.3959	0.3524	-0.5497	0.1341
γ_{2002}	-1.5707	0.3129	-1.4414	0.3594	0.0197	0.1263
γ_{2003}	-1.5869	0.3141	-1.4074	0.3604	0.1380	0.1153
	H. NAICS 61: Educational Services					
server.linux$_{t-1}$	**2.2183**	**0.0580**	**2.1225**	**0.0600**	**0.5555**	**0.7640**
server.windows$_{t-1}$	-0.0925	0.0828	-0.1003	0.0876	-0.1326	0.8790
server.other$_{t-1}$	0.2144	0.0547	0.1091	0.0581	-0.0614	0.4258
pc.linux$_{t-1}$	0.4491	0.0600	0.3696	0.0626	-0.0483	1.0855
pc.windows$_{t-1}$	0.2487	0.1216	0.2301	0.1283	-0.0266	1.0687
pc.other$_{t-1}$	0.0662	0.0584	0.0667	0.0606	-0.0974	0.2880
non-pc.other$_{t-1}$	0.2694	0.0545	0.1865	0.0579	0.4972	0.4272
γ_{2001}	-1.6688	0.1396	-1.6606	0.1517	-0.5995	0.1756
γ_{2002}	-1.6570	0.1422	-1.6591	0.1542	-0.3396	0.2545
γ_{2003}	-1.7626	0.1454	-1.7251	0.1568	-0.0118	0.2314
additional control	No		Yes		No	

Table 6: Results for Linux Adoption (Continued)

Variable	Probit				AC Method	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
	(1)		(2)		(3)	
	I. NAICS 62: Health Care					
server.linux$_{t-1}$	**2.3016**	**0.1156**	**2.2475**	**0.1192**	**0.9844**	**0.6721**
server.windows$_{t-1}$	-0.3311	0.1633	-0.3419	0.1691	-0.1442	1.0240
server.other$_{t-1}$	0.1648	0.0968	0.1418	0.1003	0.2839	0.7332
pc.linux$_{t-1}$	0.8028	0.1184	0.7189	0.1222	0.2486	0.8643
pc.windows$_{t-1}$	-0.5905	0.2899	-0.6020	0.2933	-0.1069	2.0235
pc.other$_{t-1}$	0.3116	0.1637	0.3502	0.1693	-0.1742	0.5010
non-pc.windows$_{t-1}$	0.1769	0.2336	0.2486	0.2327	-0.8158	0.3564
non-pc.other$_{t-1}$	0.2265	0.0876	0.2058	0.0923	0.2738	0.1396
γ_{2001}	-1.1277	0.3191	-1.1803	0.3278	-0.4087	0.1611
γ_{2002}	-0.9138	0.3340	-0.9628	0.3436	0.1615	0.1192
γ_{2003}	-1.0088	0.3367	-1.0261	0.3457	-0.3028	0.2136
	J. NAICS 7-8: Arts, Entertainment, and Other Services					
server.linux$_{t-1}$	**2.4383**	**0.2671**	**2.4285**	**0.3052**	**0.5707**	**2.1140**
server.windows$_{t-1}$	0.0669	0.3103	0.0722	0.3453	0.0440	0.9308
server.other$_{t-1}$	0.4240	0.2182	0.4682	0.2526	0.2105	0.8526
pc.linux$_{t-1}$	0.6669	0.2944	0.6180	0.3251	0.4427	4.3052
pc.other$_{t-1}$	0.0561	0.3698	0.0653	0.3968	0.1469	1.6876
non-pc.other$_{t-1}$	-0.1430	0.2315	-0.2133	0.2730	-0.1002	0.2835
γ_{2001}	-1.7391	0.3520	-1.8171	0.4473	-0.3929	0.2827
γ_{2002}	-2.4236	0.3794	-2.4863	0.4732	-0.0885	0.3678
γ_{2003}	-2.4234	0.3866	-2.5864	0.4919	-0.8375	1.7445
	K. NAICS 9: Public Administration					
server.linux$_{t-1}$	**2.3387**	**0.0746**	**2.2644**	**0.0778**	**0.4771**	**1.6498**
server.windows$_{t-1}$	0.0656	0.1099	0.0153	0.1150	-0.3292	0.8094
server.other$_{t-1}$	0.2020	0.0641	0.1452	0.0672	-0.1631	0.2219
pc.linux$_{t-1}$	0.6681	0.0744	0.6039	0.0778	0.9199	0.8071
pc.windows$_{t-1}$	0.1550	0.2462	0.1080	0.2536	-0.2641	0.4442
pc.other$_{t-1}$	0.0583	0.1221	-0.1734	0.1368	-0.1584	0.9459
non-pc.windows$_{t-1}$	0.1287	0.1093	0.0218	0.1166	0.2032	0.8041
non-pc.other$_{t-1}$	0.1579	0.0587	0.1152	0.0618	-0.0632	0.1667
γ_{2001}	-2.0987	0.2614	-2.0302	0.2749	-0.2199	0.0795
γ_{2002}	-2.0717	0.2656	-2.0218	0.2799	-0.3989	0.1532
γ_{2003}	-2.1240	0.2669	-2.0421	0.2811	0.0088	0.1302
additional control	No		Yes		No	

Table 7: Robustness Check for Sample Selection[a]

Industry	2000-2003 Est.	S.E.	#obs.	2000-2002 Est.	S.E.	#obs.	2001-2003 Est.	S.E.	#obs.
	(1)			(2)			(3)		
	A. $\mathsf{server.windows}_{t-1}$ in Windows Adoption								
agri_utility.1-2	0.335	0.968	1,980	0.872	0.663	2,469	1.331	2.218	2,376
manufacture.31-32	1.239	1.623	4,352	1.389	1.083	5,607	-0.027	1.407	5,349
manufacture.33	0.392	0.821	6,304	0.431	0.958	7,974	1.242	1.075	7,767
retail.4	0.673	2.065	2,768	0.054	1.756	3,522	1.589	1.953	3,789
information.51	0.224	1.562	2,704	0.466	1.779	3,483	0.058	1.566	3,612
financial.52-53	0.134	2.996	2,624	0.189	1.972	3,531	0.548	1.731	3,573
professional.54-56	0.423	1.256	3,904	0.462	0.962	4,977	0.764	1.196	5,205
education.61	-0.008	1.122	6,152	0.607	1.190	7,152	0.947	0.898	7,767
medical.62	0.562	1.191	3,676	0.811	1.939	4,506	0.736	1.085	4,695
arts_service.6-7	-0.401	2.017	868	0.009	1.720	1,218	0.731	3.745	1,299
government	1.293	1.133	7,620	1.210	1.691	8,466	0.765	1.498	8,436
	B. $\mathsf{server.linux}_{t-1}$ in Linux Adoption								
agri_utility.1-2	0.471	2.730	1,980	1.069	1.972	2,469	1.030	1.634	2,376
manufacture.31-32	0.669	1.845	4,352	1.352	2.013	5,607	0.781	2.220	5,349
manufacture.33	0.490	1.605	6,304	0.427	1.071	7,974	0.398	1.132	7,767
retail.4	0.849	1.656	2,768	1.094	1.915	3,522	0.794	1.566	3,789
information.51	0.341	2.302	2,704	1.060	2.078	3,483	1.056	1.197	3,612
financial.52-53	0.170	2.640	2,624	0.162	2.089	3,531	-0.010	2.001	3,573
professional.54-56	0.197	1.507	3,904	0.811	2.335	4,977	0.868	1.716	5,205
education.61	0.556	0.764	6,152	0.708	0.730	7,152	0.947	0.401	7,767
medical.62	0.984	0.672	3,676	0.732	1.163	4,506	0.372	0.778	4,695
arts_service.6-7	0.571	2.114	868	0.654	0.903	1,218	1.566	2.932	1,299
government	0.477	1.650	7,620	-0.085	0.839	8,466	0.276	1.153	8,436

[a]The AC method is separately applied to the 2000-2003, 2000-2002, and 2001-2003 complete panels. The coefficient estimates for $\mathsf{server.windows}_{t-1}$ in Windows adoption and $\mathsf{server.linux}_{t-1}$ are reported separately for subsamples of different industries. We drop the sample orthogonality conditions if the number of observations in a cell is less than one.

Table 8: Robustness Check for Different Cutoff Numbers to Drop Cells[a]
($\mathsf{server.windows}_{t-1}$ in Windows Adoption)

	Drop the sample orthogonality conditions if											
	#obs. in a cell < 1			#obs. in a cell < 2			#obs. in a cell < 3			#obs. in a cell < 4		
Industry	Est.	S.E.	M^b	Est.	S.E.	M	Est.	S.E.	M	Est.	S.E.	M
agri_utility.1-2	0.335	0.968	482	-0.591	1.029	155	-1.017	0.564	98	0.462	0.534	75
manufacture.31-32	1.239	1.623	1,237	0.638	1.874	318	1.689	2.002	188	0.182	2.243	136
manufacture.33	0.392	0.821	1,935	-1.145	0.902	478	-0.710	0.923	268	1.087	0.966	191
retail.4	0.673	2.065	852	0.591	2.440	256	0.731	2.647	155	0.938	3.057	94
information.51	0.224	1.562	1,368	0.506	2.008	279	0.050	2.213	142	0.216	2.082	99
fin/acial.52-53	0.134	2.996	641	0.292	3.862	185	0.750	4.541	107	0.576	4.548	78
professional.54-56	0.423	1.256	1,307	-0.346	1.456	287	-0.189	1.532	157	0.094	1.771	109
education.61	-0.008	1.122	2,873	1.183	1.584	530	1.074	1.800	289	1.669	1.893	187
medical.62	0.562	1.191	1,047	1.451	1.416	265	0.291	1.367	148	0.138	1.048	112
arts_service.6-7	-0.401	2.017	297	0.651	1.170	70	-0.112	1.568	42	-0.348	n/a	33
government	1.293	1.133	2,098	0.588	1.224	546	0.975	1.249	331	1.805	1.257	237

	Drop the sample orthogonality conditions if											
	#obs. in a cell < 5			#obs. in a cell < 6			#obs. in a cell < 7			#obs. in a cell < 8		
Industry	Est.	S.E.	M	Est.	S.E.	M	Est.	S.E.	M	Est.	S.E.	M
agri_utility.1-2	0.506	0.617	60	-0.083	0.638	50	-0.286	0.647	40	-0.446	0.507	37
manufacture.31-32	0.469	2.310	109	0.640	2.322	89	1.329	1.880	77	1.775	n/a	69
manufacture.33	-0.032	0.824	153	-0.646	0.856	132	-0.681	0.859	110	-0.473	0.857	103
retail.4	0.419	3.324	75	0.075	3.976	64	0.178	1.068	55	0.373	n/a	46
information.51	1.134	2.009	73	0.436	2.437	56	0.131	2.440	47	1.622	1.053	36
fin/acial.52-53	0.255	5.098	64	1.361	5.108	55	-0.290	n/a	47	-0.106	n/a	38
professional.54-56	0.048	1.789	88	-0.124	1.613	73	-0.065	1.620	61	1.150	1.681	56
education.61	0.641	1.743	155	0.883	1.932	131	0.463	2.098	109	0.895	3.546	91
medical.62	0.965	0.426	87	0.552	0.447	77	0.424	n/a	68	0.573	n/a	61
arts_service.6-7	-0.211	n/a	30	-0.520	n/a	24	-0.557	n/a	17	1.420	n/a	15
government	0.878	1.307	187	0.907	1.313	155	0.892	1.320	131	1.668	1.322	119

[a] The AC method is applied to the 2000-2003 complete panel. Windows adoption is estimated separately using different cutoffs for dropping cells containing few observations. The coefficient estimates for $\mathsf{server.windows}_{t-1}$ are reported separately for subsamples of different industries. For some samples, the estimate of the asymptotic variance is not available because $\hat{D}_\theta' \hat{D}_\theta$ in (17) is singular.

[b] M denotes the number of the sample moment conditions actually used in the estimation.

Table 9: Robustness Check for Different Cutoff Numbers to Drop Cells[a]
(server.linux$_{t-1}$ in Linux Adoption)

	Drop the sample orthogonality conditions if											
	#obs. in a cell < 1			#obs. in a cell < 2			#obs. in a cell < 3			#obs. in a cell < 4		
Industry	Est.	S.E.	M^b	Est.	S.E.	M	Est.	S.E.	M	Est.	S.E.	M
agri_utility.1-2	0.471	2.730	482	0.280	3.524	155	0.791	4.142	331	0.644	4.907	75
manufacture.31-32	0.669	1.845	1,237	0.192	2.697	318	0.493	2.858	188	0.186	3.788	136
manufacture.33	0.490	1.605	1,935	0.921	1.999	478	0.543	2.163	268	0.640	2.254	191
retail.4	0.849	1.656	852	0.566	2.043	256	0.391	2.558	155	-0.338	3.179	94
information.51	0.341	2.302	1,368	0.617	3.237	279	-0.086	3.810	142	0.574	4.519	99
fin/acial.52-53	0.170	2.640	641	-0.384	3.661	185	-0.224	4.640	107	-0.351	5.198	78
professional.54-56	0.197	1.507	1,307	0.245	1.932	287	0.264	2.137	107	0.141	2.320	109
education.61	0.556	0.764	2,873	0.470	0.974	530	-0.055	1.063	289	0.086	1.150	187
medical.62	0.984	0.672	1,047	0.714	0.665	265	0.422	0.658	148	0.777	0.670	112
arts_service.6-7	0.571	2.114	297	0.501	0.812	70	1.236	n/a	42	-0.082	n/a	33
government	0.477	1.650	2,098	1.076	1.876	546	0.444	1.954	331	0.552	1.998	237

	Drop the sample orthogonality conditions if											
	#obs. in a cell < 5			#obs. in a cell < 6			#obs. in a cell < 7			#obs. in a cell < 8		
Industry	Est.	S.E.	M	Est.	S.E.	M	Est.	S.E.	M	Est.	S.E.	M
agri_utility.1-2	0.577	5.761	60	1.079	5.831	50	-0.810	8.352	40	0.130	8.347	37
manufacture.31-32	0.364	5.542	109	0.233	5.555	89	0.694	8.341	77	0.793	n/a	69
manufacture.33	0.795	2.486	153	0.792	2.655	132	1.017	3.243	110	0.286	3.244	103
retail.4	0.798	3.646	75	0.772	3.720	64	0.927	3.718	55	0.241	n/a	46
information.51	0.807	5.616	73	1.547	6.749	56	1.224	6.750	47	-0.137	7.874	36
fin/acial.52-53	-0.597	6.229	64	-1.432	8.833	55	-1.371	n/a	47	-0.107	n/a	38
professional.54-56	-0.014	2.574	88	-0.180	2.697	73	-0.036	3.357	61	0.364	3.333	56
education.61	-0.111	1.233	155	-0.058	1.370	131	-0.231	1.535	109	-0.560	1.557	91
medical.62	0.165	0.599	87	0.387	0.610	77	0.503	n/a	68	-0.047	n/a	61
arts_service.6-7	0.509	n/a	30	-0.981	n/a	24	0.281	n/a	17	0.353	n/a	15
government	0.252	2.157	187	1.251	2.201	155	0.882	2.247	131	0.941	2.408	119

[a]The AC method is applied to the 2000-2003 complete panel. Linux adoption is estimated separately using different cutoffs for dropping cells containing few observations. The coefficient estimates for server.linux$_{t-1}$ are reported separately for subsamples of different industries. For some samples, the estimate of the asymptotic variance is not available because $\hat{D}'_\theta \hat{D}_\theta$ in (17) is singular.

[b]M denotes the number of the sample moment conditions actually used in the estimation.