Lock-In and Unobserved Preferences in Server Operating Systems: 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

March 23, 2011

Abstract

This paper investigates to what extent the persistence of Microsoft's Windows in the market for server oprating systems is due to lock-in or unobserved preferences. While the hypothesis of lock-in plays an important role in the antitrust policy debate for the operating systems market, it has not been extensively documented empirically. To account for unobserved preferences, we use a panel data identification approach based on time-variant group fixed effects, and estimate the dynamic discrete choice panel data model developed by Arellano and Carrasco (2003). Using detailed establishment-level data, we find that once we account for unobserved preferences, the estimated magnitudes of lock-in are considerably smaller than those from the conventional approaches, suggesting that unobserved preferences play a major role in the persistence of Windows. Further robustness checks are consistent with our findings.

JEL classification: C23, L15, L17, L86

Keywords: Lock-in; Unobserved preference; Panel data; Discrete choice; Fixed effects; Random effects

*We appreciate the NET Institute for financial support. We thank George Deltas, Bin Gu, Mark Jacobsen, Kyoo il Kim, Roger Koenker, Thierry Magnac, Zhongjun Qu, and Joel Waldfogel for helpful discussion and comments. We are particularly grateful to the editors and two anonymous referees for their valuable comments and suggestions that significantly improved the manuscript. All remaining errors are our responsibility. Corresponding author: Seung-Hyun Hong, 112 David Kinley Hall, 1407 West Gregory Drive, Urbana, IL 61801.

1 Introduction

This paper investigates to what extent the persistence of Microsoft's Windows can be explained by lock-in or unobserved preferences. Despite the appearance of alternatives such as GNU/Linux, Windows has held a dominant position in the operating system market. This persistence may be interpreted as evidence of lock-in, potentially arising from a number of sources such as the costs of training personnel and upgrading hardware. The observed persistence, however, may also result from consumers' preference for Windows, in that Windows might be perceived as superior to other operating systems.

Distinguishing between lock-in and unobserved preferences is a crucial step toward any evaluation of Microsoft's role in the operating system market, as these two factors have opposite implications on consumer welfare. However, despite the vigorous debate over the antitrust case against Microsoft (e.g., Bresnahan (2001), Liebowitz and Margolis (1999)), few empirical studies have tried to distinguish lock-in from unobserved preferences for Windows. This gap is partly due to lack of detailed individual data, but also due to the difficulty in identification between state dependence and unobserved heterogeneity (Heckman, 1981a,c). In this paper, we address these issues by using detailed establishment-level[1] panel data on server operating systems,[2] and applying a panel data identification approach based on time-variant group fixed effects to distinguish between lock-in and unobserved preferences.

While few empirical studies have examined lock-in and unobserved preferences in the choice of operating systems, a large body of empirical literature in marketing and industrial organization has attempted to distinguish between state dependence and unobserved heterogeneity in product choices in various other markets.[3] Most studies in this literature have relied on the random effects approach, where unobserved heterogeneity is assumed to follow a parametric distribution. However, the random effects approach is subject to the initial conditions problem (see, e.g., Heckman (1981b), Hsiao (2003)). Because of the difficulty in addressing

[1] Throughout this paper, we use firms and establishments interchangeably to refer to business organizations.

[2] See Section 3.1 for a detailed description of our data. Note that our data also contain information on other segments such as personal computers and mainframes. However, we focus on the server segment, because we have strong evidence that firms are likely to have repeatedly made decisions on their server operating systems during our sample period, whereas similar evidence for other segments is weak. See Section 3.3 for more discussion.

[3] See, e.g., Chintagunta *et al.* (1991), Dube *et al.* (2009), Guadagni and Little (1983), Keane (1997), Osborne (2007), Seetharaman (2004), and Shum (2004).

this problem, many studies in marketing and industrial organization tend to assume the initial conditions to be truly exogenous, and arbitrarily initialize the beginning of the process (e.g., Keane (1997)). This strong assumption might not be needed if reasonably long panel data were available,[4] but this is not the case in a typical study using short panel data. In this respect, one could alternatively use the fixed effects approach, in which no distributional assumption is made for unobserved heterogeneity, hence bypassing the initial conditions problem. Nonetheless, as noted in Honoré and Tamer (2006), fixed effects methods are not available for many dynamic discrete choice panel data models, and even when they are available, the maintained assumptions tend to be strong. The sophisticated estimator proposed by Honoré and Kyriazidou (2000), for example, rules out time-specific effects.

The point of departure for our approach is that the conventional specification for fixed effects which are time-invariant and individual-specific is not necessarily the only specification for unobserved heterogeneity. We can instead consider a different specification where fixed effects are time-variant and group-specific. To the extent that firms' preferences for operating systems depend on related information and experiences, firms with the same observed histories would have similar preferences, since they are likely to have acquired similar information and experiences. Furthermore, unobserved preferences for a particular operating system may change over time as firms acquire more information and experiences each period. For this reason, we specify that unobserved preferences for operating systems are reflected by the time-variant group-specific fixed effects which are the same for firms with the same observed histories. Note that our specification is essentially equivalent to the framework of Arellano and Carrasco (2003). Hence, we "difference out" these fixed effects, using a semi-parametric approach developed by Arellano and Carrasco (2003), and further estimate a GMM version of their dynamic discrete choice panel data model.

Using the balanced panel data from the *Computer Intelligence Technology Database*, we find that most of the conventional approaches, including those based on random effects, yield estimates of strong positive dependence between current and previous choices with respect to the use of server operating systems. Once we allow for the time-variant group-specific fixed

[4]Goldfarb (2006), for example, exploited unusually long panel data and estimated household-specific regressions, hence avoiding the random effects approach. He also found that the random effects models commonly used in the literature overestimated the switching costs, relative to his household-specific regressions.

effects, however, the estimated magnitudes of lock-in are considerably smaller than those from the conventional approaches. Though our estimates do not necessarily reject the significance of lock-in in Windows usage, they do imply that unobserved preferences account for a considerable part of the observed persistence. These findings are robust to further checks to address potential issues in our model specifications and our data. Because our data are not intended to be representative of all firms in the United States, we do not attempt to generalize our findings beyond the sample analyzed in this paper. Nevertheless, our findings suggest that unobserved preferences may be indeed an important factor in explaining the persistent dominance of Windows.

The paper is organized as follows. Section 2 explains our model and estimation methods. In Section 3, we describe our data and report descriptive statistics. Section 4 presents our estimation results and further provides robustness checks. Section 5 concludes the paper.

2 Econometric Framework

2.1 The Model

To investigate the factors that determine firms' choices of operating systems in the server segment, we begin with the net payoff from using server operating system j, $j = 1, \ldots, J$. Specifically, we consider the following reduced-form function for N firms observed T consecutive time periods:

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

where π_{ijt} is the net payoff from using product j at period t, γ_{jt} captures a time effect, $y_{ik(t-1)}$ is a binary variable indicating whether firm i used server operating system k at the previous period, x_{it} is a vector of binary indicator variables for using operating systems in other segment at the previous period (e.g., an indicator variable for using Windows in the personal computer segment), Z_{it} is a vector of observed characteristics of the firm, such as the number of desk workers, and u_{ijt} is an unobserved component of the net payoff.

In this paper, we focus on two main factors: lock-in and unobserved preferences in firms' decisions to use Windows (or Linux), where all versions of Windows are classified as the same operating system, and likewise for Linux. The degree of lock-in in server operating

system j is captured by β_{jj} in (1). Our specification of unobserved preferences is provided in Section 2.2. We consider the binary choice model in which the previous decisions also determine the current decision on whether to use operating system j. Specifically, our model assumes that firm i decides to use operating system j at period t if the net payoff is non-negative, i.e., $y_{ijt} = \mathbb{1}\{\pi_{ijt} \geq 0\}$. We acknowledge that it is important to study the joint decision of adopting multiple operating systems in the server segment, but we do not model such a decision because the key factor in this joint decision is network effects within the same segment, which requires a different modeling framework.[5] That said, we use the binary choice model framework, since it still allows us to examine the two main factors.

Three caveats are in order. First, our measure of lock-in reflects any state dependence consistent with the definition of y_{ijt} above.[6] For example, a significant positive value of β can result from direct or indirect costs of installing new software and retraining personnel, costs of acquiring new software licenses, or technical requirements that entail incompatibility between different operating systems. However, we do not attempt to identify a specific cause of lock-in.

Second, the durable good nature of operating systems implies that firms might continue to use the same operating systems without making active decisions, in which case we might include observations with spurious positive correlations between y_{ijt} and $y_{ij(t-1)}$ due to firms' inaction, thus suggesting an upward bias in our estimate of β.[7] However, the direction of the bias suggests that this concern does not seem to be serious in our case, because once we account for unobserved preferences, we do not find evidence of significant lock-in in Section 4.1. Nevertheless, we acknowledge that the durable good nature of operating systems is difficult to address, given that actual decisions are not directly observed. As a partial solution, we additionally select observations that are highly likely to have made active adoption or replacement decisions, and check robustness of our findings in Section 4.2.[8]

Third, we use $y_{ij(t-1)}$ as the summary of the past decisions in our main specification. This

[5]It is worth pointing out that multinomial discrete choice models are not applicable to this market, as choices are not mutually exclusive; firms can elect to use several different operating systems simultaneously.

[6]However, β does not capture costs associated with upgrading within the same operating system, since we do not distinguish between different versions of the same operating systems.

[7]If we also take the durable goods aspect seriously, β is likely to depend on a lapse of time since the initial adoption, in which case our estimate of the time invariant β should be interpreted as the mean or an approximation of the distribution of the coefficients over time.

[8]Given that β might reflect other factors not related to switching costs, we consider another partial solution by including proxies for prices of server operating systems. We appreciate a referee for suggesting this solution.

approach is often used in empirical work, though other specifications of the past decisions have been used as well. Because of a concern that the coefficient on $y_{ij(t-1)}$ may not fully capture lock-in, however, we also consider alternative specifications with further lagged dependent variables in our robustness check. A related concern is that firms may adopt a new operating system for one computer and test it before they adopt it for the entire server segment, in which case it might be difficult to interpret the coefficient on $y_{ij(t-1)}$ as capturing lock-in. By including the decisions before period $t-1$, we can partially address this concern. To fully address this concern, we further drop the firms that are likely to have tested server operating systems, and check the robustness of our results.

2.2 Unobserved Preferences

The fundamental difficulty in estimating the extent of lock-in is the presence of unobserved preference. 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 not observed, they are included in the error term u_{ijt} in (1). Hence, without imposing any assumption on u_{ijt}, we cannot distinguish state dependence from unobserved heterogeneity. In this regard, the key assumption in this paper is that u_{ijt} is a composite error given by

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

where $H_i^t = (H_{i1}, \ldots, H_{it})$, $H_{it} = (y_{i1(t-1)}, \ldots, y_{iJ(t-1)}, x_{it}, Z_{it})$, and conditional on H_i^t, ϵ_{ijt} follows a known distribution. Specifically, we assume that $\epsilon_{ijt}|H_i^t \sim$ i.i.d. $\mathcal{N}(0, \sigma_t)$. 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 payoff function.

We suppose that each firm has η_{ij}, denoting the true quality of operating system j perceived by firm i (or true preference of firm i) under full information. However, when firms make decisions to use operating systems, they are unlikely to have full information on technical features and qualities of operating systems. Based on their experiences at each period, firms will rather acquire more information on the true quality of operating systems over time, and their actual preferences would be revised accordingly. Therefore, we assume that η_{ij} is not

fully observed even to firm i. Firm i then acts on its expectation of the true preference for operating system j based on its previous history H_i^t. Hence, unobserved preferences are captured by $E(\eta_{ij}|H_i^t)$ in our formulation.

Note that the model in (1) and our assumptions on u_{ijt} are essentially the same as those in Arellano and Carrasco (2003), where their specification of unobserved heterogeneity is interpreted as a semi-parametric random effects specification. Though we use the same specification as in Arellano and Carrasco (2003), we consider another interpretation of this specification, given that $E(\eta_{ij}|H_i^t)$ is an unknown variable fixed for a given value of H_i^t. That is, $E(\eta_{ij}|H_i^t)$ can be interprted as a fixed effect that is the same for firms with the same observed history. Therefore, if we consider groups that are defined in terms of their histories H_i^t, then we can think of $E(\eta_{ij}|H_i^t)$ as the group-specific fixed effects that may vary over time. Strictly speaking, $E(\eta_{ij}|H_i^t)$ is not exactly a fixed effect because it depends on H_i^t which is random. Nevertheless, we call it a time-variant group-specific fixed effect, not only because we do not impose any distributional assumption on $E(\eta_{ij}|H_i^t)$, but also because this interpretation provides simpler intuition behind the identification of our model.

We acknowledge that this group-specific fixed effect is less general than standard fixed effects, in that it is not individual-specific.[9] In other respect, however, it is more flexible than time-invariant individual fixed effects, since it varies over time, depending on different histories. To the extent that firms' preferences for operating systems depend on related information and experiences, firms with the same history would have similar preferences, because they are likely to have acquired similar information and experiences. Moreover, unobserved preferences for a particular operating system may change over time as firms acquire more information and experiences at each period. Therefore, it is plausible to assume that unobserved preferences for an operating system are captured by time-variant group-specific fixed effects for firms with the same history.

[9]As a referee pointed out, our identifying assumption based on time-variant group-specific fixed effects is particularly strong for observations in the early period of our samples. This issue cannot be fully addressed, given our data. However, as a partial solution, we additionally consider shorter panels with different initial years, and check the robustness of our findings. These additional results are reported in the Web appendix.

2.3 Estimation

To estimate our model, we follow the method proposed by Arellano and Carrasco (2003) – henceforth, the AC method. To explain the application of the method, we begin with the notation used in this section. We suppress the subscript j, and consider the decision denoted by y_{it}. We drop Z_{it} and include only x_{it} which is a vector of indicator variables for using different operating systems in other segments at the previous period. Accordingly, we consider a discrete random vector $H_{it} = (y_{i(t-1)}, x_{it})$, where H_{it} has a finite support of L points. The vector $H_i^t = (H_{i1}, \ldots, H_{it})$ thus takes on L^t different values ϕ_l^t $(l = 1, \ldots, L^t)$. Let us define the group-specific dummy variable for observations with the same history ϕ_l^t by $d_{il}^t = \mathbb{1}\{H_i^t = \phi_l^t\}$. For each specific history ϕ_l^t, we denote the conditional choice probability by $p_l^t = \Pr(y_{it} = 1 | H_i^t = \phi_l^t)$. To denote the conditional choice probability in general, we use $h_t(H_i^t)$, so that $h_t(H_i^t) = \sum_{l=1}^{L^t} d_{il}^t p_l^t$.

The assumption on u_{it} in the previous section then implies that the probability of $y_{it} = 1$ conditional on the history H_i^t is given by

$$\Pr(y_{it} = 1 | H_i^t) = \Phi\left(\frac{\gamma_t + \beta y_{i(t-1)} + x_{it}\delta + E(\eta_i | H_i^t)}{\sigma_t}\right), \tag{3}$$

where Φ is the standard normal cdf. To estimate the probit model in (3), one might consider estimating $E(\eta_i | H_i^t)$ directly by including d_{il}^t for all possible histories ϕ_l^t. However, this approach is practically infeasible, given that L^t can be very large. As a result, we need to "difference out" these time-variant group-specific fixed effects. To this end, we invert (3) to obtain

$$E(\eta_i | H_i^t) = \sigma_t \Phi^{-1}(h_t(H_i^t)) - \gamma_t - \beta y_{i(t-1)} - x_{it}\delta. \tag{4}$$

We then define $\nu_{it} \equiv E(\eta_i | H_i^t) - E(\eta_i | H_i^{t-1})$ and note that from the law of iterated expectations, we have $E(\nu_{it} | H_i^{t-1}) = E[E(\eta_i | H_i^t) | H_i^{t-1}] - E(\eta_i | H_i^{t-1}) = 0$, which implies the following unconditional moments

$$E(d_{il}^{t-1} \nu_{it}) = 0 \qquad (l = 1, \ldots, L^{t-1}). \tag{5}$$

Plugging (4) into (5) then yields

$$E\left\{ d_{il}^{t-1} \left[\sigma_t \Phi^{-1}(h_t(H_i^t)) - \sigma_{t-1} \Phi^{-1}(h_{t-1}(H_i^{t-1})) - \Delta\gamma_t - \beta \Delta y_{i(t-1)} - \Delta x_{it}\delta \right] \right\} = 0, \tag{6}$$

where $\Delta\gamma_t = \gamma_t - \gamma_{t-1}$, $\Delta y_{i(t-1)} = y_{i(t-1)} - y_{i(t-2)}$, and $\Delta x_{it} = x_{it} - x_{i(t-1)}$.[10] Note that the moment condition in (6) does not depend on $E(\eta_i|H_i^t)$. Therefore, we can use (6) to estimate the main parameters, while accounting for time-variant group-specific fixed effects.

To use the moments (6) for our estimation, let us further define

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

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 in (6) should hold for each t and l, we consider

$$\frac{1}{N}\sum_{i=1}^{N}\psi_i(p,\theta), \tag{7}$$

where N is the number of firms in our data, and $\psi_i(p,\theta)$ is given by

$$\psi_i(p,\theta) = \left[(\psi_{i1}^1(p,\theta),\ldots,\psi_{iL}^1(p,\theta)),\ldots,(\psi_{i1}^{T-1}(p,\theta),\ldots,\psi_{iL^{T-1}}^{T-1}(p,\theta))\right]'.$$

Note that the dimension of $\psi_i(p,\theta)$ is supposed to be $(\sum_{t=2}^{T} L^{t-1}) \times 1$, 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 $\psi_i(p,\theta)$ will then be far less than $\sum_{t=2}^{T} L^{t-1}$, the number of potential different histories. Let M denote the number of the moment conditions actually used in the estimation, $M < \sum_{t=2}^{T} L^{t-1}$.

The sample orthogonality conditions in (7) contain p_l^t which is unknown but can be estimated non-parametrically 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 the cell-specific sample frequency estimator $\hat{p}_l^t = \frac{1}{\sum_{i=1}^{N} d_{il}^t}\sum_{i=1}^{N} y_{it}d_{il}^t$. In the second step, we replace p with $\hat{p}$, and estimate θ using a GMM estimator given by

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

[10] The moment condition in (6) suggests that we need at least three periods of data, because not only y_{it} but also $y_{i(t-1)}$ and $y_{i(t-2)}$ enter (6). The moment condition also suggests that while we can identify σ_t, β, and δ, we cannot identify γ_t, but only $\Delta\gamma_t$. Note also that we consider the moment condition for each history ϕ_i^t. Thus, the total number of moments is $\sum_{t=2}^{T} L^{t-1}$, where T is the final period observed in the data. For example, if we have data for four periods $(t = 0, 1, 2, 3)$, and $L = 8$, then period 0 provides information on y_0, and we obtain the difference between period 2 and period 1 (conditional on the history up to period 1), as well as that between period 3 and period 2 (conditional on the history up to period 2). Hence, the total number of moments from (6) is $8 + 8^2 = 72$.

where A_M is a $M \times M$ weighting matrix. In the empirical application, some cells may contain very few observations, in which case there may be small sample biases in $\hat{p}_l^t$ 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. In addition, we estimate the standard errors of the parameter estimates by using the formula given in Arellano and Carrasco (2003), which takes into account the first stage estimation errors in $\hat{p}$.

In our actual estimation, we estimate our binary choice model for Windows and separately for Linux, and our regressors include indicator variables for using different server operating systems as well as discrete variables for using different operating systems in other segments at period $t-1$. Though we include Z_{it} in the conventional approaches used for comparison in Section 4, it is not included in our estimation using the AC method, because most variables in Z_{it} such as the number of desk workers and total personal computers range from zero to a large number, hence increasing the number of possible histories considerably. In the next section, we provide more precise definition of our key variables.

3 Data

3.1 Data Description

We use the 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 *et al.* (2002)). In this paper, we focus on the period from 2000 to 2004, during which three major events in the operating system markets occurred – Microsoft released Windows 2000 in February 17, 2000, Windows XP in October 25, 2001, and Windows Server 2003 in April 24, 2003.[11] 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 2000-2004 CITDB data.[12] Though

[11] Refer to the Microsoft News Center, available at http://www.microsoft.com/presspass.

[12] Harte-Hanks releases a new dataset every January, containing information collected in the previous year. Our reference year is the collection year, not the release year; e.g. the 2000 dataset was released January 2001.

our data cover five years, we believe that they contain sufficient information to study firms' decisions on the use of server operating systems.

Nevertheless, one might be worried that the period of 2000-2004 coincided with the burst of the dot-com bubble. The concern is that firms might not have invested in software during this period, so that it would be difficult to study firms' decisions on software such as operating systems. To check this concern, we examine the annual changes in investment in software and other equipment provided by the U.S. Bureau of Economic Analysis. Table 1 presents these changes from 1997 to 2005. The first column shows that software investment had increased substantially until 2000. This increase slowed down slightly after 2000, but the level of total investment in software still remained high during the period of 2000-2004, suggesting that firms are unlikely to have reduced the level of investment in software despite the burst of the dot-com bubble. This pattern in software investment is even more evident when we compare that with those of other types of investment shown in the rest of the columns. Therefore, it is likely that firms have recurrently made decisions on the use of operating systems during this period.

The CITDB is useful for our purpose because it contains detailed information on establishment characteristics and the ownership of computer hardware and software such as 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. Thus, the number of observations remains similar each year, but many establishments were not surveyed in every year. In our application, we focus on the observations with panel information.

We study the use 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[13]; PC, including personal computers that are used for standalone desktops or

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

client computers connected to servers[14]; and non-PC, including mainframes, midrange, and workstations that are not used for servers. Note that we can only investigate the usage of operating systems up to the segment level, since the information on operating system choices 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.[15]

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).[16] 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; and similarly for server.windows, pc.windows, and non-pc.windows.

3.2 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 firms that report the information on the use of server operating systems.[17] Firms may not report the information on server operating systems for two reasons: either because they do not have any server computer, or because they do not regard server operating systems as important. By excluding the former case, we implicitly assume that our analysis is conditional on firms' ownership of either an Internet server or a network server. The latter case is a common problem in many survey data

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

[15]For related statistics, see Table 2 in Section 3.2.

[16]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.

[17]Among 607,781 observations in the CITDB, about 54% of them report information on operating systems for either Internet server or network server.

– 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. This problem can result in a potential underestimation of the number of firms 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 firms every year. For some observations, the CITDB reuses information collected in the previous year. If a firm continues to use the same operating system as before, the information on operating systems can be current even though it was collected in the previous year. On the other hand, if the firm 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 observations with up-to-date information.[18] For the initial observation of each firm in our sample, there is no issue regarding reusing the same information. For this reason, we include the initial observation of each firm as long as the information on computing technology was collected either on the same year or the year before.

Third, we use only balanced panel data for our main analysis. Obviously, we cannot use information from firms that are observed only once in our data. We further restrict our sample to balanced panels of all five years in order to use the econometric methodology described in Section 2.[19] Hence, our main analysis uses the balanced panel data for 2000-2004. Though we do not attempt to generalize our findings beyond the sample examined in our analysis, it is still useful to check whether the main data used in our analysis are significantly different from the original sample. In the Web appendix, we compare summary statistics of the main data with those of the original unbalanced panel data, and do not find that our main data are systematically different from the overall sample.

[18] Because the CITDB records when the survey on each firm was conducted, we can find whether its information is outdated. Among the 328,109 observations with any kind of server operating system, about 68% of them report up-to-date information on computing technology.

[19] Since this restriction reduces the sample size for each year considerably, we additionally consider shorter panels of four consecutive years: 2000-2003 and 2001-2004. The results using these samples are reported in the Web appendix.

3.3 Descriptive Statistics

Before we present our estimation results in the next section, we examine basic descriptive statistics and several issues with respect to our data. We begin by considering Table 2. The table clearly shows that Windows is dominant in the server segment. In this table, the shares of each operating system are the mean values of the dummy variable for whether a firm uses the given operating system for the given segment. Because a firm can use more than one kind of operating system, the sum of shares for server.windows, server.linux, and server.other can be larger than one. Firms may use multiple operating systems either because of the complementarity between different operating systems, or because of potential testing – for example, a firm may use Windows for all servers, except one server for which it installs Linux to test whether Linux would meet its need. Since this kind of testing raises some concerns as discussed in Section 2.1, we attempt to identify the observations that might test an operating system, and exclude them from our analysis. That said, Panel B of Table 2 shows that most firms in our data use only one kind of operating systems for the server segment.

Table 3 presents the changes in the use of operating systems and the number of computers in each segment over time. Three observations emerge from Table 3. First, the dominance of Windows is persistent in both the server segment and the PC segment, except for the non-PC segment in which other operating systems are the most popular, presumably because most non-PCs are IBM computers running IBM operating systems. The persistent dominance of Windows can be explained by either lock-in or unobserved preferences for Windows operating systems, which we investigate further in the next section.

Second, the total number of server computers has increased over time. If a firm purchased a new server computer, it is likely to have made a decision on its server operating system. The increase in total.server throughout the sample period thus suggests that firms in our data are likely to have repeatedly made decisions on their server operating systems, which is one reason why we focus on the server segment. Another reason for focusing on the server segment is that a substantial fraction of firms have adopted either an Internet server computer or a network server computer for the first time during our sample period. This is shown in Panel A of Table 4 which reports that about 32.3% of firms have adopted server computers for the first time. For example, if a firm did not use an Internet server until 2002, then there is

no previous decision on whether to use a particular operating system for an Internet server before 2002. Hence, the adoption decision of this firm in 2002 is less likely to depend on the previous decisions.[20] In contrast, the proportion of firms that adopted PCs for the first time is insignificant in Table 4, although total.pc is increasing over time in Table 3. Notice also that total.non-pc is decreasing in Table 3, although the fractions of firms that adopted non-PC for the first time are not negligible in Table 4. Therefore, it is unclear whether firms have made decisions on their operating systems for PCs or non-PCs frequently during our sample period, which is the other reason why we focus only on the server segment.

Third, the use of Linux has increased in both the server segment and the PC segment in Table 3, while the use 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 a proprietary Unix operating system. However, it is also possible that firms have switched from Windows to Linux while others have simultaneously switched from Unix to Windows.

To examine these possibilities, we compute the fraction of firms that switched from an operating system to a different operating system, where switching means that a firm used an operating system before, and then stopped using it, while starting to use a different operating system at the same period. Table 4 presents the results. Panel B shows that more firms switched from Windows to Linux than from other operating systems to Linux, and that a nontrivial number of firms switched from other operating systems to Windows, thus suggesting that the presence of Windows has also affected the usage of Linux. Panel B also shows that a significant fraction of firms did switch from one operating system to another operating system in the server segment.

Firms' decisions on server operating systems are not limited to switching their operating systems. They also include updating one version to another version of the same operating system. Panel C of Table 4 reports the fractions of firms that updated their operating systems, where updating means that a firm stopped using a version of an operating system (say, Windows 2000), and started to use a different version of the same family of the operating system (say, Windows 2003).[21] The table shows that about 85.9% of firms updated either Windows or

[20]Note that the marketing literature often relies on brand switching induced by price discounts as exogenous events for the identification of state dependence (see, e.g., Dube, Hitsch, and Rossi (2008)). Although we do not emphasize this, the first-time adoption of a server could be considered similar events that may shift $y_{i(t-1)}$.

[21]In our data, there are 9 different versions of Windows in each segment. In the Internet server segment, there

Linux during our sample period. Therefore, both Panels B and C suggest that most firms in our data indeed made decisions on either switching or updating at least once during our sample period. In our robustness checks, we also restrict our sample to the firms that made the usage decision more frequently.

Panel D of Table 4 presents the proportion of firms that might have tested an operating system in the server segment, where testing an operating system means that a firm has used it for a single year while also continuing to use a different operating system for the entire sample period. The table shows that only a small fraction of the firms tested an operating system during our sample period, and thus, the possibility of testing is unlikely to be critical in our data.

4 Results

4.1 Main Estimation Results

To investigate whether the persistent dominance of Windows can be explained by lock-in or unobserved preferences, we estimate the model presented in Section 2. Table 5 reports the main estimation results using the 2000-2004 balanced panel data described in Section 3.2. We additionally consider shorter panels as well as separating the sample by industry subgroups. The findings from these additional samples are similar to those in Table 5 and are reported in the Web appendix.

In Table 5, the last column presents the results from the AC method, and all other columns report the results from the conventional approaches which include the probit model, the random effects probit model, the logit model, and the conditional logit model. To estimate the random effects probit model, we assume a normal distribution for unobserved heterogeneity. For the conditional logit model, we assume that all regressors including the lagged dependent variables are strictly exogenous, and apply the standard method (see, e.g., Chamberlain (1980)). The AC method reports the results from using the sample orthogonality conditions with cells containing at least 4 observations, which is the cutoff used by Arellano and Carrasco (2003). We also

are 10 versions of Linux and 43 versions of other operating system. In the network server segment, there are 10 versions of Linux and 54 versions of other operating systems. As for the PC segment, there are 10 versions of Linux and 30 versions of other operating systems. In the non-PC segment, there are 7 versions of Linux and 53 versions of other operating systems. To identify updating, we check the changes in the use of these versions.

consider different cutoffs, but our main findings from Table 5 remain valid. These results and other detailed results on the AC method are presented in the Web appendix as well.

The results for the Windows estimation are presented in Panel A of Table 5. The coefficients on $\mathsf{server.windows}_{t-1}$ are intended to capture the extent of lock-in. For all probit models, the coefficient estimates on $\mathsf{server.windows}_{t-1}$ range between 2.10 and 2.17, and all of them are statistically significant. To interpret these estimates in terms of the contributions to choice persistence, the table also presents the marginal effect of changing $y_{ij(t-1)}$ from 0 to 1 on the probability of $y_{ijt} = 1$. Specifically, we compute the difference in the predicted probabilities $\Pr(y_{ijt} = 1|y_{ij(t-1)} = 1) - \Pr(y_{ijt} = 1|y_{ij(t-1)} = 0)$ for each observation, and report their means and standard deviations.[22] The coefficient estimate of 2.17 in the first column then implies that using Windows in the previous period increases the likelihood of using Windows in the current period by 49% on average. The strong positive correlation between the current decision and the previous decision, nevertheless, can be also explained by unobserved preferences which generate a potential positive bias in the coefficient estimate on $\mathsf{server.windows}_{t-1}$. This bias is only slightly reduced by including various firm-specific characteristics or by allowing for random effects.

Because we consider the conditional logit model to allow for the individual-specific fixed effects, we also consider the logit model as a benchmark. The logit coefficient estimate is 3.85 and is also statistically significant. In contrast, the conditional logit estimate is 0.49. Because the standard conditional logit does not allow for the lagged dependent variables, we do not believe that this estimate is consistent. Nonetheless, this estimate suggests that the magnitude of fixed effects can be large, and thus, unobserved preferences may explain a considerable part of the positive correlation between the current decision and the previous decision.

The AC method allows for the time-variant group-specific fixed effects as discussed in Section 2.2. The results from the AC method show that the coefficient estimate on $\mathsf{server.windows}_{t-1}$ is 0.66 and statistically significant. Thus, lock-in seems to be a nontrivial factor. However, its magnitude is much smaller than those from the probit models, suggesting that unobserved preferences are likely to be more important. In terms of the marginal effect, the coefficient

[22] To compute the marginal effect for the AC method, we follow the procedure described in Arellano and Carrasco (2003), in which the marginal effect is computed for each observation. For this reason, the marginal effect for the conventional approaches is also computed for each observation.

estimate of 0.66 implies that using Windows in the previous period increases the likelihood of using Windows in the current period by 8% on average. To the extent that the estimated marginal effects from the probit models reflect both lock-in and unobserved preferences, and that the marginal effect from the AC method reflects only lock-in, our estimates suggest that lock-in accounts for 0.08/0.49, or about 16% of the persistence in Windows, whereas unobserved preferences account for about 84% of the Windows persistence.

Panel B of Table 5 reports the estimation results for Linux usage. Similar to the results for Windows usage, the coefficient estimates on $\mathsf{server.linux}_{t-1}$ are positive and statistically significant for all probit models and the logit model, while it is much smaller for the conditional logit model. In contrast, the coefficient estimate from the AC method is -0.30 and not statistically significant, and thus, the corresponding marginal effect is negligible. This result suggests that the degree of lock-in is not substantial in Linux usage, whereas unobserved preferences are still important. The table also shows that the AC method estimates for $\mathsf{server.windows}_{t-1}$ and $\mathsf{server.other}_{t-1}$ are respectively -0.16 and 0.39, suggesting that switching from Windows or other operating systems to Linux may not entail significant costs.

4.2 Robustness Checks

The previous section shows that the choice persistence in the server operating systems is largely explained by unobserved preferences, rather than lock-in. However, our main model assumes that firms make decisions on the use of their operating systems each year, which may raise two potential concerns in our findings, given the durable goods nature of the operating systems. First, actual decisions might not have been made frequently, so that we may treat a firm's inaction as its decision to choose the same operating system. Second, our measure of lock-in may reflect other factors not related to switching costs.

The first concern implies that our data include observations with spurious positive correlations in their choices over time, hence suggesting an upward bias in our estimate for lock-in. Note that the direction of this bias is in our favor, given that we do not find a substantial degree of lock-in. Nevertheless, we attempt to address this concern by selecting only observations that are highly likely to have made decision recurrently. To this end, we first check whether a firm started or discontinued a version of an operating system at period t, which indicates

that the firm indeed made a decision on the use of its operating system at period t. We then consider the following two selected samples. The first sample includes only firms that have made decisions on their operating systems in the server segment for at least two years. The second sample includes firms that have made decisions in the server segment each year. In both these samples, we also exclude those who are likely to have tested an operating system.[23]

Table 6 presents the results using these selected samples. Panel A reports the results for the Windows estimation using the first sample, and Panel B reports the results using the second sample. Similarly, the results for the Linux estimation using the first and second sample are presented in Panels C and D, respectively. In the table, as we restrict our samples, the coefficient estimates on $y_{ij(t-1)}$ become smaller, which seems reasonable because we are essentially including only those who switched or updated operating systems more frequently.[24] Nevertheless, our main findings do not change even when we use selected samples. That is, the positive correlations estimated from the conventional approaches become far less significant once we allow for the time-variant group-specific fixed effects, suggesting that the positive correlations between the current decisions and the previous decisions are largely explained by unobserved preferences.

As for the second concern, we acknowledge that our measure of lock-in does not solely reflect switching costs. In particular, note that our specification does not include prices of operating systems, and thus, our measure of lock-in may capture prices of new operating systems even for firms that do not consider updating or switching because their current operating systems are fairly new. To address this concern, ideally we would want to include interaction terms between $y_{ij(t-1)}$ and prices of operating systems.[25] However, the CITDB does not collect data on prices, and publicly available information on prices of server operating systems is limited. Though we obtained published prices previously posted on software vendors' web sites,[26] most

[23]Note that some of the observed switching activity may be due to small scale testing of operating systems, rather than actual switching to new operating systems. As in Section 3.3, we define testing in our data as using an operating system for a single year, while still using a different operating system for the entire sample period. In addition to the samples used in this section, we also consider a separate set of samples excluding only those who have tested an operating system at least once, and estimate the same models as in Table 5. The results are similar to those in Table 5, and are reported in the Web appendix.

[24]This observation suggests a potential upward bias in our main results, as discussed above.

[25]See, e.g., Shum (2004) for related specifications including prices.

[26]Though most software vendors delete their old web pages, this information can be still obtained from the Internet Archive, available at http://www.archive.org/web/web.php.

web sites note that actual prices may vary. In fact, many firms tend to use volume licensing, rather than paying published prices.[27] Moreover, published prices for standard or basic edition do not vary during the period studied in this paper.[28] As a result, we cannot use published prices.

Nevertheless, to the extent that firms tend to use volume licensing, actual prices might depend on the number of servers, in which case we may use the number of servers as proxies for prices of server operating systems. In our application, we use indicator variables for different ranges of the number of servers owned by a firm, in order to use cell sample frequencies for the AC method. Panels A-B of Table 7 report the results from using the number of servers as proxies for prices from volume licensing. Most coefficient estimates for $\mathsf{server.window}_{t-1}$ in Panel A and $\mathsf{server.linux}_{t-1}$ in Panel B are slightly smaller than those reported in Table 5, but our main findings remain the same in this table.

In addition to the concern related to the durable goods nature of the operating systems, one more concern is that lock-in may not be fully captured by $y_{ij(t-1)}$ only. To address this concern, we consider an alternative specification that includes $y_{ij(t-2)}$. We estimate similar models as in Table 5, and the results are presented in Panels C-D in Table 7. For all the probit models and the logit model, including $y_{ij(t-2)}$ slightly reduces the coefficient estimates on $y_{ij(t-1)}$, compared to those without $y_{ij(t-2)}$. In contrast, the results from the AC method show that the estimated degree of lock-in, captured either by $y_{ij(t-1)}$ or by $y_{ij(t-2)}$, is unlikely to be significant, hence implying that unobserved preferences are indeed important. We also include $y_{ij(t-3)}$, and additionally use shorter panels including $y_{ij(t-2)}$, but the main findings do not change. These additional results are reported in the Web appendix.

[27] See, e.g., BearingPoint (2004), for industry practices. This study was commissioned by Microsoft. BearingPoint (2004) considers typical purchase scenarios for medium and enterprise businesses over a five year period, in which organizations purchase licenses to server operating systems as well as (often annual) subscriptions to vendor support, and work directly with the vendor to receive reduced pricing. This study also shows that the prices of software licenses tend to account for a small portion of the overall licensing expenses for server operating systems. In typical scenarios, other recurrent expenses make up about 40%-72% of the budget for Windows, and 100% of the budget for Linux.

[28] For example, Windows 2000 server with 5 client access licenses (CALs) and standard edition of Windows 2003 server, plus 5 CALs, both cost $999 during the sample period. In addition, Linux can be downloaded for free, in which case its basic price can be zero.

5 Concluding Remarks

In this paper, we examine the persistence in the usage of server operating systems, and decompose its sources into two factors: lock-in and unobserved preferences. To account for unobserved preferences, we use a specification proposed by Arellano and Carrasco (2003), which can be interpreted as time-variant group fixed effects. We then difference out these fixed effects using a semi-parametric approach developed by Arellano and Carrasco (2003). Our results show that once we allow for unobserved preferences, the estimated degrees of lock-in are substantially smaller than those from the conventional approaches. Though these results suggest that lock-in may not be as significant as what is commonly believed in this industry, we do not wish to argue that this conclusion necessarily applies to other periods or other segments of the operating systems market – rather, our paper presents an instance where the observed persistence may not necessarily imply lock-in.

References

Arellano M, Carrasco R (2003). Binary choice panel data models with predetermined variables. *Journal of Econometrics* **115**:125–157.

BearingPoint (2004). Server operating system licensing and support cost comparison: Windows Server 2003, Red Hat Enterprise Linux 3 and Novell/SUSE Linux 8. Http://download.microsoft.com/download/5/2/9/529e82d1-3485-49a4-91cd-f7f8216731a2/BearingPoint.pdf.

Bresnahan T (2001). Panel data models: some recent developments. Discussion Paper 00-51, Stanford Institute for Economic Policy Research.

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

Bresnahan T, Greenstein S (1996). Technical progress and co-invention in computing and in the uses of computers. *Brookings Papers on Economic Activity. Microeconomics* **1996**:1–83.

Chamberlain G (1980). Analysis of covariance with qualitative data. *The Review of Economic Studies* **47**:225–238.

Chintagunta P, Jain D, Vilcassim N (1991). Investigating heterogeneity in brand preferences in logit models for panel data. *Journal of Marketing Research* **28**:417–428.

Dube J, Hitsch G, Rossi P (2009). State dependence and alternative explanations for consumer inertia. Working Paper w14912, NBER.

Goldfarb A (2006). State dependence at internet portals. *Journal of Economics & Management Strategy* **15**:317–352.

Guadagni P, Little J (1983). A logit model of brand choice calibrated on scanner data. *Marketing science* **2**:203–238.

Heckman J (1981a). Heterogeneity and state dependence. In Rosen S (ed.) *Studies in Labor Markets*, volume 31, 91–140, University of Chicago Press.

Heckman J (1981b). The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process. In Manski C, McFadden D (eds.) *Structural analysis of discrete data with econometric applications*, 179–195, MIT Press.

Heckman J (1981c). Statistical models for discrete panel data. In Manski C, McFadden D (eds.) *Structural analysis of discrete data with econometric applications*, 114–178, MIT Press.

Honoré B, Kyriazidou E (2000). Panel data discrete choice models with lagged dependent variables. *Econometrica* 839–874.

Honoré B, Tamer E (2006). Bounds on parameters in panel dynamic discrete choice models. *Econometrica* 611–629.

Hsiao C (2003). *Analysis of panel data*. Cambridge Univ Pr.

Keane MP (1997). Modeling heterogeneity and state dependence in consumer choice behavior. *Journal of Business & Economic Statistics* **15**:310–27.

Liebowitz S, Margolis S (1999). *Winners, Losers & Microsoft*. Independent Institute.

Osborne M (2007). Consumer learning, switching costs, and heterogeneity: A structural examination. E.A.G. Discussion Paper 07-10, U.S. Department of Justice.

Seetharaman P (2004). Modeling multiple sources of state dependence in random utility models: A distributed lag approach. *Marketing Science* **23**:263–271.

Shum M (2004). Does advertising overcome brand loyalty? Evidence from the breakfast-cereals market. *Journal of Economics & Management Strategy* **13**:241–272.

Table 1: Investment in Software and Other Equipment[a]

Year	Software	Communication equipment	Office equipment	Nonmedical instruments
1997	101,659	53,355	5,629	18,065
1998	122,834	61,602	5,032	18,412
1999	151,497	74,765	3,650	18,386
2000	172,441	96,864	3,800	19,547
2001	173,681	90,648	4,781	20,326
2002	173,445	73,663	5,185	20,547
2003	185,576	75,357	8,156	20,202
2004	204,620	81,748	8,352	22,112
2005	218,004	83,181	8,571	23,555

[a]Source: National Income and Product Account Table 5.5.6U available at the U.S. Bureau of Economic Analysis website. The table reports the real private fixed investment in millions of chained (2005) dollars.

Table 2: Summary Statistics[a]

A. Shares of Operating Systems	
server.windows	0.93
server.linux	0.13
server.other	0.25
B. Kinds of OS in Servers	
windows only	0.76
linux only	0.04
other only	0.08
windows and linux	0.07
windows and other	0.14
linux and other	0.01

C. Firm Characteristics	
total.pc	251.7
total.non-pc	2.1
total.server	9.3
revenue (in $million)	59.6
employees	325.4
desk.workers	149.3
internet.users	113.4
internet.developers	0.8
programmers	3.4
#observations	36,690

[a]The table reports the mean of each variable in the 2000-2004 balanced panel data. The samples include only observations with any server operating system and with up-to-date information. The share is the mean of a dummy variable for whether an observation uses each operating system in the server segment.

Table 3: Changes in the Use of Operating Systems[a]

Variable	2000	2001	2002	2003	2004
server.windows	0.92	0.93	0.94	0.94	0.93
server.linux	0.08	0.12	0.15	0.17	0.15
server.other	0.29	0.27	0.25	0.24	0.22
pc.windows	0.96	0.98	0.99	0.98	0.98
pc.linux	0.07	0.13	0.15	0.18	0.17
pc.other	0.12	0.09	0.08	0.06	0.03
non-pc.windows	0.03	0.04	0.07	0.06	0.03
non-pc.linux	0.00	0.00	0.01	0.01	0.01
non-pc.other	0.42	0.31	0.31	0.22	0.12
total.pc	213.9	242.6	257.6	269.1	275.2
total.non-pc	2.34	2.25	2.71	1.98	1.04
total.server	7.10	8.14	9.65	10.32	11.07
#observations	7,338	7,338	7,338	7,338	7,338

[a]The table reports the mean of each variable.

Table 4: First-time Computer Adopters and Switching Patterns[a]

A. First-time Adoption		C. Updating	
adopted a server for the first time	0.323	updating in server.windows	0.646
adopted a PC for the first time	0.020	updating in server.linux	0.213
adopted a non-PC for the first time	0.239	updating in server.other	0.000
B. Switching in Servers		D. Testing	
switching from Windows to Linux	0.135	testing Linux in server	0.014
switching from other to Linux	0.066	testing Windows in server	0.009
switching from Linux to Windows	0.115	testing other in server	0.012
switching from other to Windows	0.182		
switching from Linux to other	0.043		
switching from Windows to other	0.120		

[a]The table first reports the fractions of the firms that did not have a computer in each segment and then adopted a computer in that segment for the first time during the sample period. The table next reports the fractions of the firms that updated, switched, or tested an operating system for the server segment during the sample period. Updating means that a firm stopped using a version of an operating system, and started to use a different version of the same family of the operating system. Switching means that a firm stopped using an operating system, and started to use a different operating system. Testing an operating system means that a firm did not use it before, and started to use it, and then stopped using it in the following year, while the firm also continued to use a different operating system for the entire sample period.

Table 5: Estimation Results[a]

	Probit				Random Effects Probit				Logit		Conditional Logit		AC Method	
	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.
A. Dependent Variable: Windows Use														
server.linux$_{t-1}$	-0.37	0.04	-0.39	0.04	-0.39	0.04	-0.40	0.04	-0.84	0.09	-0.30	0.15	0.18	0.17
server.windows$_{t-1}$	2.17	0.04	2.14	0.04	2.13	0.04	2.10	0.04	3.85	0.07	0.49	0.10	0.66	0.20
server.other$_{t-1}$	-0.56	0.03	-0.58	0.03	-0.59	0.03	-0.62	0.03	-1.29	0.07	-0.42	0.13	-0.38	0.08
pc.linux$_{t-1}$	-0.02	0.04	-0.04	0.05	-0.02	0.04	-0.04	0.04	-0.09	0.09	-0.04	0.14	0.74	0.10
pc.windows$_{t-1}$	0.40	0.08	0.42	0.08	0.21	0.04	0.21	0.05	0.76	0.15	0.49	0.25	0.77	0.09
pc.other$_{t-1}$	-0.09	0.05	-0.10	0.05	-0.14	0.04	-0.16	0.04	-0.21	0.10	-0.04	0.20	-0.83	0.08
non-pc.windows$_{t-1}$	0.13	0.08	0.12	0.08	0.14	0.08	0.12	0.08	0.23	0.16	-0.09	0.25	-0.05	0.08
non-pc.other$_{t-1}$	-0.04	0.03	-0.07	0.03	-0.04	0.03	-0.07	0.03	-0.15	0.07	0.02	0.12	-0.11	0.03
marginal effect of y_{t-1}	0.49	0.10	0.48	0.11	0.48	0.10	0.46	0.11	0.42	0.15	0.08	0.02	0.08	0.08
B. Dependent Variable: Linux Use														
server.linux$_{t-1}$	2.28	0.03	2.18	0.03	2.25	0.03	2.13	0.03	3.86	0.06	0.25	0.08	-0.30	0.17
server.windows$_{t-1}$	-0.02	0.04	-0.02	0.04	-0.43	0.04	-0.47	0.04	-0.05	0.09	0.10	0.15	-0.16	0.10
server.other$_{t-1}$	0.29	0.03	0.22	0.03	0.12	0.03	0.05	0.03	0.43	0.06	0.09	0.11	0.39	0.06
pc.linux$_{t-1}$	0.64	0.03	0.58	0.03	0.66	0.03	0.61	0.03	1.05	0.06	0.37	0.09	0.22	0.09
pc.windows$_{t-1}$	0.00	0.08	0.01	0.08	-1.25	0.04	-1.30	0.05	0.04	0.16	0.14	0.26	-0.15	0.07
pc.other$_{t-1}$	0.19	0.04	0.13	0.04	-0.10	0.04	-0.18	0.04	0.24	0.08	0.03	0.16	0.00	0.05
non-pc.windows$_{t-1}$	0.03	0.05	0.01	0.05	0.06	0.05	0.03	0.05	0.01	0.11	-0.02	0.16	-0.28	0.07
non-pc.other$_{t-1}$	0.10	0.03	0.08	0.03	0.08	0.03	0.06	0.03	0.14	0.05	-0.15	0.10	0.04	0.02
marginal effect of y_{t-1}	0.66	0.05	0.61	0.06	0.65	0.06	0.59	0.07	0.61	0.07	0.05	0.01	-0.02	0.03
additional control	no		yes		no		yes		yes		yes		no	

[a]All estimations use the 2000-2004 balanced panel data. The estimations with additional controls include the variables such as revenues, #IT employees, #programmers, #desk workers, Apache, total PCs, total non-PCs, total Internet servers, and total network servers. The coefficient estimates on these controls and time dummies are suppressed. The random effect probit models assume a normal distribution for the random effects, and calculate the likelihood functions using adaptive Gauss-Hermite quadrature. The conditional logit model assumes that all regressors are strictly exogenous. The AC method reports the estimation results using the sample orthogonality conditions with cells containing at least 4 observations. The estimates for $\Delta\gamma_t$ and σ_t are also suppressed in the AC method results. The marginal effect of y_{t-1} is the predicted probability under $y_{t-1} = 1$ minus that under $y_{t-1} = 0$, where y_{t-1} is $\mathsf{server.window}_{t-1}$ in Panel A and $\mathsf{server.linux}_{t-1}$ in Panel B. This is computed for each observation, and the table reports means and standard deviations.

Table 6: Robustness Check I: Use Selected Samples That Made More Frequent Decisions.[a]

	Probit				Random Effects Probit				Logit		Conditional Logit		AC Method	
	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.
	A. Windows Use Estimation: Include only firms that made decisions in server for at least two years													
server.linux$_{t-1}$	-0.40	0.04	-0.39	0.05	-0.41	0.04	-0.39	0.05	-0.81	0.09	-0.25	0.16	0.49	0.19
server.windows$_{t-1}$	1.79	0.04	1.77	0.04	1.78	0.04	1.75	0.04	3.19	0.08	0.25	0.11	0.69	0.21
server.other$_{t-1}$	-0.45	0.04	-0.47	0.04	-0.46	0.04	-0.48	0.04	-0.99	0.08	-0.42	0.14	-0.89	0.09
	B. Windows Use Estimation: Include firms that made decisions in server for all years													
server.linux$_{t-1}$	-0.23	0.08	-0.23	0.09	-0.23	0.08	-0.23	0.09	-0.44	0.17	-0.20	0.28	0.16	0.32
server.windows$_{t-1}$	1.16	0.10	1.11	0.10	1.15	0.11	1.08	0.11	2.02	0.18	-0.59	0.24	0.01	0.46
server.other$_{t-1}$	-0.36	0.08	-0.36	0.08	-0.34	0.08	-0.35	0.08	-0.75	0.16	-0.07	0.26	-0.61	0.21
	C. Linux Use Estimation: Include only firms that made decisions in server for at least two years													
server.linux$_{t-1}$	2.12	0.03	2.03	0.03	2.08	0.03	1.97	0.04	3.54	0.06	0.21	0.08	-0.20	0.22
server.windows$_{t-1}$	-0.06	0.05	-0.03	0.05	-0.39	0.04	-0.41	0.05	-0.05	0.10	0.17	0.16	-0.25	0.14
server.other$_{t-1}$	0.24	0.03	0.19	0.03	0.13	0.03	0.06	0.03	0.35	0.06	0.10	0.12	0.08	0.08
	D. Linux Use Estimation: Include firms that made decisions in server for all years													
server.linux$_{t-1}$	1.57	0.06	1.52	0.06	1.52	0.08	1.44	0.08	2.56	0.11	-0.14	0.14	-0.14	0.28
server.windows$_{t-1}$	-0.07	0.10	-0.02	0.11	-0.32	0.09	-0.32	0.10	-0.03	0.18	-0.21	0.28	-0.01	0.33
server.other$_{t-1}$	0.15	0.06	0.10	0.06	0.10	0.06	0.05	0.06	0.18	0.10	-0.13	0.19	-0.07	0.17
additional control	no		yes		no		yes		yes		yes		no	

[a]All estimations use the 2000-2004 balanced panel data, excluding the firms that tested operating systems in the server segment. Except for using more selected samples, all estimations use the same specifications as in Table 5, and other coefficient estimates are suppressed.

Table 7: Robustness Check II: Include Additional Regressors[a]

	Probit				Random Effects Probit				Logit		Conditional Logit		AC Method	
	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.
	A. Windows Use Estimation: Include proxies for volume licensing prices													
server.window$_{t-1}$	2.08	0.14	1.85	0.16	2.06	0.12	1.83	0.17	3.25	0.38	0.13	0.35	0.46	0.50
server.window$_{t-1}$×tot.sv.1	-0.06	0.10	0.19	0.15	-0.07	0.10	0.17	0.14	0.38	0.19	0.28	0.46	0.34	0.47
server.window$_{t-1}$×tot.sv.10	0.17	0.13	0.39	0.18	0.17	0.17	0.38	0.15	0.81	0.42	0.55	0.53	0.22	0.45
server.window$_{t-1}$×tot.sv.50	0.40	0.15	0.49	0.21	0.40	0.15	0.50	0.26	1.05	0.35	-0.13	0.26	-0.41	0.43
	B Linux Use Estimation: Include proxies for volume licensing prices													
server.linux$_{t-1}$	2.32	0.19	2.00	0.20	2.34	0.19	1.98	0.18	3.73	0.34	0.24	0.35	-0.25	0.35
server.linux$_{t-1}$×tot.sv.1	-0.25	0.17	0.03	0.14	-0.36	0.18	-0.05	0.13	-0.13	0.37	-0.41	0.47	0.06	0.42
server.linux$_{t-1}$×tot.sv.10	0.07	0.12	0.26	0.18	0.03	0.11	0.25	0.15	0.28	0.36	0.19	0.41	-0.31	0.41
server.linux$_{t-1}$×tot.sv.50	0.20	0.18	0.37	0.19	0.23	0.18	0.44	0.19	0.47	0.38	0.66	0.46	0.71	0.46
	C. Windows Use Estimation: Include more lagged dependent variables													
server.windows$_{t-1}$	1.82	0.05	1.81	0.06	1.78	0.06	1.77	0.06	3.27	0.10	-0.45	0.15	-0.09	0.21
server.windows$_{t-2}$	0.49	0.06	0.47	0.06	0.52	0.06	0.50	0.06	0.84	0.11	-2.33	0.27	0.09	0.11
	D. Linux Use Estimation: Include more lagged dependent variables													
server.linux$_{t-1}$	2.05	0.04	1.98	0.04	2.03	0.04	1.94	0.04	3.52	0.07	-0.83	0.12	-0.13	0.27
server.linux$_{t-2}$	0.38	0.04	0.33	0.05	0.37	0.05	0.36	0.05	0.58	0.08	-3.32	0.23	-0.07	0.10
additional control	no		yes		no		yes		yes		yes		no	

[a]All estimations use the 2000-2004 balanced panel data. Panels A-B report the estimation results from including proxies for server OS prices. Since the transaction prices are likely to be determined by volume licensing with OS companies, we use the indicators for the different number of servers as price proxies. In the table, tot.sv.1 (or tot.sv.10, or tot.sv.50) is the indicator for whether the number of servers is between 1 and 9 (or between 10 and 49, or above 50). The coefficient estimates for interactions between these dummies and y_{t-1} are reported. Panels C-D report the coefficient estimates from including additional lagged variables. Except for these additional regressors, all estimations use the same specifications as in Table 5, and other coefficient estimates are suppressed.