LEVERAGING MONOPOLY POWER BY LIMITING INTEROPERABILITY:

Theory and evidence from the computer market*

Christos Genakos†, Kai-Uwe Kühn‡ and John Van Reenen§

First Draft: January 2005

This Draft: October 2009

Abstract. When will a monopolist have incentives to leverage into a complementary market by degrading compatibility/interoperability? We develop a framework where arbitrage limits price discrimination, so leveraging becomes a method to extract more rents from the monopoly market by restoring second degree price discrimination. In a random coefficient model with complements we derive explicit conditions for when these incentives will hold. We implement our framework in the context of Microsoft's alleged strategic incentives to leverage market power from personal computer to server operating systems. We estimate a structural random coefficients demand system which allows for complements (PCs and servers). Our estimates suggest that there were incentives to reduce interoperability at the end of our sample period in the early 2000s, but not the mid 1990s when our sample begins.

JEL Classification: O3, L1, L4.

Keywords: Anti-trust, interoperability, leverage, Microsoft, complements, demand estimation, random coefficients.

*We would like to thank Tim Besley, Cristina Caffarra, Francesco Caselli, Stefonis Clerides, Peter Davis, Georg von Graevenitz, Aviv Nevo, Pierre Regibeau, Bob Stillman, Mike Whinston and participants at seminars at CEPR, LSE, Michigan, MIT, Pennsylvania and the NBER for helpful comments. Financial support has come from the ESRC Centre for Economic Performance. In the past Kühn and Van Reenen have acted in a consultancy role for Sun Microsystems.The usual disclaimer applies.

†University of Cambridge and Centre for Economic Performance

‡University of Michigan and CEPR

§Corresponding author: j.vanreenen@lse.ac.uk; London School of Economics, Centre for Economic Performance, NBER and CEPR

1. Introduction

Many antitrust cases revolve around interoperability/compatibility issues. For example, the European Microsoft case focused on the question of whether Microsoft deliberately reduced interoperability between its personal computer (PC) operating system - Windows, a near monopoly product - and rival server operating systems (a complementary market) to drive rivals from the market. Microsoft's share of server operating systems rose substantially from 20% at the start of 1996 to near 60% in 2001 (see Figure 1) and the European Commission (2004) alleged that at least some of this increase was due to a strategy of making rival software work poorly with Windows.[1] The possibility of such leveraging of market power from the PC to the server market seemed to be suggested by Bill Gates in a 1997 internal e-mail: "*What we're trying to do is to use our server control to do new protocols and lock out Sun and Oracle specifically....the symmetry that we have between the client operating system and the server operating system is a huge advantage for us*".

Such quotes could just be cheap talk and the rationality of such strategies has been strongly challenged in the past by the "Chicago School" critique of leverage theory (e.g. Bork, 1978). For example, suppose one firm has a monopoly for one good but competes with other firms in a market for another good and both goods are used in fixed proportions by customers. The Chicago school observed that the monopolist in the first market did not have to monopolize the second market to extract monopoly rents from the market. Indeed, whenever there was product differentiation in the second market, the monopolist in the first could only benefit from the presence of other firms.[2] Following the lead of authors in the Chicago tradition, there has been much work on trying to derive efficiency explanations for exclusionary practices that were previously seen as anti-competitive. For example, price discrimination has been put forward as an explanation for tying strategies.[3]

More recently, studies of exclusive dealing (see Bernheim and Whinston, 1998) and tying[4]

[1] Microsoft's Appeal against the 2004 Decision was rejected by the European Court of First Instance in September 2007.

[2] For a formal statement of this point, see Whinston (1990), Proposition 3.

[3] Bowman (1957), Adams and Yellen (1976), Schmalensee (1984), McAfee, McMillan and Whinston (1989).

[4] See Whinston (1990), Farrell and Katz (2000), Carlton and Waldman (2002) among others.

have shown that rational foreclosure in market for complements is possible in well specified models.[5] Most of these models have the feature that exclusionary strategies are not necessarily profitable in the short run. However, exclusionary strategies through their impact on investment, learning by doing, etc., can make current competitors less effective in the future, making the exclusionary strategy profitable.

In this paper we suggest a different theory of foreclosure through interoperability degradation that relies on an immediate benefit for the monopolist and apply this to the PC and server markets. The reduction of competition allows the monopolist to price discriminate between customers with heterogeneous demand elasticities for PCs. If customers with high elasticity of demand for PCs also have low willingness to pay for servers, server purchases can be used for second degree price discrimination. A monopolist both of PC and server operating systems would lower the price for the PC operating system and extract surplus from customers with inelastic PC demand by charging higher server operating system prices. Competition on the server market will limit the ability to price discriminate in this way. By reducing interoperability, the PC operating system monopolist can reduce competition on the server market, re-establishing the ability to price discriminate. We show in a simple example (sub-section 3.2) that in the absence of horizontal product differentiation between servers, this effect will lead the monopolist to foreclose rivals in the server market even if their product is *arbitrarily* better than the product of the monopolist. More generally, with product differentiation between servers (and server operating systems) reducing interoperability will have a negative impact on the PC sales of the monopolist and this second effect limits his incentives to degrade his rival's quality. The overall incentive to degrade depends on the balance of these two forces and so is an empirical issue.

The goal of this paper is to characterize and estimate the foreclosure incentives of the monopolist. Such incentives are generally very difficult to measure, but the price discrimination theory allows us to do so primarily on the basis of estimating demand elasticities. For the argument we are making, modelling the heterogeneity between buyers is essential

[5]See Rey and Tirole (2007) for a comprehensive review of this literature and Whinston (2001) for an informal survey in relation to some aspects of the U.S. vs. Microsoft case.

for generating foreclosure incentives. But a model of customer heterogeneity is also a central feature of recent approaches for estimating demand systems in differentiated product markets. We therefore first develop the theory on the basis of a discrete choice model with random coefficients as used in demand estimations by Berry, Levinsohn, and Pakes (1995, henceforth BLP) or Nevo (2001). Customer heterogeneity is captured by the distribution of the coefficients in the utility function across the population. We extend this approach to allowing complementarity between two markets and compare our results to those from existing approaches such as Gentzkow (2007) and Song and Chintagunta (2006).

In the context of this model we show the role of customer heterogeneity in generating foreclosure incentives. We also show that the correlation between the demand elasticity for PCs and willingness to pay for servers is critical for generating a positive markup on the server operating system of the PC operating system monopolist.

We then bring this model to the data. We use quarterly data from the US PC and server markets between 1996 and 2001 to estimate a structural demand system in the two linked markets. We use this to infer Microsoft's margins on PC and server operating systems. According to our model results Microsoft did have incentives to reduce interoperability in the early 2000s. This is precisely the prediction of the theory.[6]

The paper is structured in the following way. Section 2 gives the basic idea and section 3 presents the core theoretical results relating foreclosure incentives to price discrimination. Section 4 details the econometrics, section 5 the data and section 6 the results. Section 7 concludes. In the Appendices we give more details of derivations, data and estimation techniques.

2. The basic idea

In this section we develop the basic conditions that have to hold to generate incentives to foreclose competitors from the market by degrading interoperability. We will develop the theory exclusively focusing on the maximization problem of the PC operating system

[6]Hence, our static motivation complements dynamic theories, for example those based on applications network effects, that have been shown to generate anti-competitive incentives to extend monopoly (e.g. Carlton and Waldman, 2002). These dynamic effects only make our static foreclosure incentives stronger.

monopolist ("the monopolist" denoted $_M$), leaving the optimal decisions of other players in the background. This is the only information necessary to make inferences about the incentives for foreclosure in the empirical part of the paper. Notationally, this means that we only need to keep track of the monopolist's price of the PC operating system, ω, his server operating system price, ω_M, and the rival's server operating system price, $\omega_{M'}$.

Let $\mathbf{p}_J = \hat{\mathbf{p}}_J + \boldsymbol{\omega}_J$ be the vector of total price for a PC, with element p_j, which is the sum of the vector of hardware prices $\hat{\mathbf{p}}_J$ and the vector PC operating system prices $\boldsymbol{\omega}_J$. Since there is a monopolist in the operating system market for PCs $\omega_J = \omega \cdot \mathbf{1}$. Similarly let $\mathbf{p}_k = \hat{\mathbf{p}}_k + \boldsymbol{\omega}_k$ be the vector of total price for a server of model k, which can again be broken down in hardware and software prices. We use $k = M$ when referring to the monopolist's server. Define $a_k \in [0, 1]$ as an interoperability parameter affecting some quality characteristic of a server k. We set $a_M = 1$ for the monopolist and $a_k = a \leq 1$ for competitors in the server market. We define $q(\mathbf{p}_j, \mathbf{p}_k, a)$ as the total demand for PCs and $q_M(\mathbf{p}_j, \mathbf{p}_k, a)$ the demand for servers that run the monopolist's server operating system. The idea is that a is some interoperability parameter that can be affected by the monopolist. Clearly, increased interoperability will increase total PC demand but will decrease the demand for server operating systems offered by the monopolist. Total profits of the monopolist are then given by:

$$\Pi(\mathbf{p}_j, \mathbf{p}_k, a) = (\omega - c)q(\mathbf{p}_j, \mathbf{p}_k, a) + (\omega_M - c_M)q_M(\mathbf{p}_j, \mathbf{p}_k, a),$$

where ω_M is the monopolist's price for the server operating system and c and c_M are the corresponding marginal costs.[7]

We are ultimately interested in the marginal incentive of the monopolist to decrease interoperability. There will be such an incentive if:

$$(\omega - c) \left. \frac{dq(\mathbf{p}_j, \mathbf{p}_k, a)}{da} \right|_{\omega,\omega_M} + (\omega_M - c_M) \left. \frac{dq_M(\mathbf{p}_j, \mathbf{p}_k, a)}{da} \right|_{\omega,\omega_M} < 0$$

where the derivatives are total derivatives of the respective output measures holding the mo-

[7] The marginal cost can be thought of as being very close to zero in software markets.

nopolist's operating system prices constant. Hence, this derivative contains the direct effect of interoperability as well as the impact of the price responses to a change in interoperability by all rival software producers and all hardware producers. Rearranging terms we obtain that there is an incentive to decrease interoperability at the margin if:

$$\frac{\omega_M - c_M}{\omega - c} > -\frac{\left.\frac{dq(\mathbf{p}_j,\mathbf{p}_k,a)}{da}\right|_{\omega,\omega_M}}{\left.\frac{dq_M(\mathbf{p}_j,\mathbf{p}_k,a)}{da}\right|_{\omega,\omega_M}} \tag{1}$$

Intuitively, degrading interoperability increases the PC monopolist's server sales but reduces PC sales. This implies that the right hand side of equation (1) (which we call the "relative output effect") is always strictly positive. Incentives for interoperability degradation will be larger the more the reduction in quality of the rivals will lead to substitution towards the monopolist's server operating system and the less buyers refrain from buying PCs as a result of lower server qualities. We will estimate these quantities directly in Section 6.

On the left hand side of equation (1), the relative value of PC and server operating system sales matters (we call this the "relative margin effect"). Interoperability degradation will only be profitable if the margin on the server operating system of the monopolist ($\omega_M - c_M$) sufficiently exceeds the margin on the PC operating system ($\omega - c$). As we will show this can never be the case if there is no heterogeneity between customers. In that case the "one monopoly profit theory" holds and the monopolist will price the server at marginal cost. We will explain in the next section that positive margins on the server are the result of second degree price discrimination between customers with differential sensitivity to price. Heterogeneity in demand elasticities between populations among which the monopolist cannot directly discriminate will therefore be crucial in generating positive server operating system margins.

The margins on operating systems are essentially unobservable. For our econometric estimations we only have prices of PC's and servers bought inclusive of an operating system. While there do exist some list prices that allow us to infer an order of magnitude, as usual we will have to estimate margins from the data. This estimation will therefore use the profit maximizing pricing behavior of the monopolist to infer such margins. However, there are some modelling choices that have to be made. Given the complementarity between software and hardware as well as between PCs and server, the move order in price setting will be important for determining the pricing incentives for the monopolist. We will assume that the hardware and software companies set their prices simultaneously. Then the price the software company charges is directly added to whatever price the hardware company charges for the computer. This assumption seems consistent with what we observe in the market as Microsoft effectively controls the price of the software paid by end users through licensing arrangements[8]. The first order conditions for the monopolist are then given by:

$$q + (\omega - c)\frac{\partial q}{\partial \omega} + (\omega_M - c_M)\frac{\partial q_M}{\partial \omega} = 0 \tag{2}$$

$$q_M + (\omega - c)\frac{\partial q}{\partial \omega_M} + (\omega_M - c_M)\frac{\partial q_M}{\partial \omega_M} = 0 \tag{3}$$

Denoting $\frac{\partial q}{\partial \omega}\frac{1}{q} = \varepsilon_\omega$, the semi-elasticity of the impact of a change in price (ω) on quantity demanded (q), we can solve equations (2) and (3) for the profit margins:

$$(\omega - c) = \frac{\frac{q_M}{q}\varepsilon_\omega^M - \varepsilon_{\omega_M}^M}{\varepsilon_\omega \varepsilon_{\omega_M}^M - \varepsilon_{\omega_M}\varepsilon_\omega^M} \tag{4}$$

$$(\omega_M - c_M) = \frac{\frac{q}{q_M}\varepsilon_{\omega_M} - \varepsilon_\omega}{\varepsilon_\omega \varepsilon_{\omega_M}^M - \varepsilon_{\omega_M}\varepsilon_\omega^M} \tag{5}$$

The semi-elasticities on the right hand side of these two equations can be estimated on the basis of PC and server data. Note that the estimates of the margins for server and PC operating systems are affected by the interaction of the two markets. As our later analysis will show, either of the price cost margins can be negative. To obtain some preliminary intuition, consider the benchmark in which the server market becomes perfectly competitive,

[8] Our assumption greatly simplifies the analysis of the monopolist's problem. While the optimal software price does depend on the expected prices for the hardware, we do not have to solve for the pricing policies of the hardware producers to analyze the pricing incentives of the software firm. An alternative sequential set up would be if the software company moves first. Its pricing incentives are not affected by whether the software producer charges the hardware firm or if it charges the consumer directly. However in this case the pricing incentives of the software company have to take into account the price reactions of the hardware company. This would add an additional layer of complexity to the model which we currently abstract from.

i.e. $\varepsilon^{M}_{\omega_M} \rightarrow -\infty$. Then the PC operating system margin of the monopolist goes to $w - c \rightarrow -\frac{1}{\varepsilon_\omega}$ and the server operating system margin $\omega_M - c_M \rightarrow 0$. Hence, a naive estimation of PC operating system margins that ignored server margins would systematically generate incorrect results. Indeed, when there are incentives to use the server operating system price for second degree price discrimination, it would look as if the monopolist would not be fully exploiting its monopoly power.

Before we develop a theory of second degree price discrimination as an explanation for foreclosure incentives, we first discuss why we believe equation (1) holds with an inequality, rather than a strict equality, as one might think in the case where interoperability would be chosen optimally by the monopolist. First, there are time lags between the design of the less interoperable software and its diffusion on the market. Second, other server OS vendors such as Novell, Sun and more recently Linux sought to overcome the fall in interoperability through a variety of measures such as developing "bridge" products, redesigning their own software, reverse engineering, etc. Third, note that interoperability refers to one of many quality characteristics of the operating system. For any given quality characteristic the optimal choice may be zero interoperability, i.e. a corner solution. This means the monopolist would want to reduce quality of the server rivals further if he could. At the same time there are many reasons why it will be impossible for a monopolist to reduce all interoperability to zero, i.e. making rival server operating systems non-functional. One reason is that there are different server market segments. For example, in European Commission (2004) it was claimed that Microsoft had an incentive to exclude rivals in workgroup server markets (the market which we focus on), but not in the markets for web servers or enterprise servers.[9] Finally, since the late 1990s anti-trust action in the US and EU may have somewhat slowed down Microsoft's desire to reduce interoperability. All these reasons strongly suggest that in the presence of foreclosure incentives we should find a strict incentive to foreclose at the margin.

[9] Enterprise servers are high end corporate servers that manage vast amounts of mission critical data in large corporations. They need very high levels of security and typically use custom written written software. Web servers host the web-sites of companies and are also used for e-commerce.

3. Second Degree Price Discrimination and Foreclosure Incentives: A theoretical Framework

3.1. The Model of Demand. We model the demand for "workgroup" purchases. A buyer i of type w has demand for a PC workgroup which consists of w PCs and one server. We assume that each buyer can connect his workgroup to one server or not. There are J producers of PCs[10] and K producers of servers indexed by j and k respectively. The index $j = 0$ refers to a purchase that does not include PCs while $k = 0$ refers to an option that does not include a server. A buyer i with workgroup size w who buys the PCs from producer j and the server from producer k has conditional indirect utility:

$$u_{ijk}(w) = w\left[x_j\beta_i + A_k y_k \gamma_i - \lambda_i[p_j + \frac{1}{w}p_k] + \xi_j + \xi_k + \eta_{ijk}\right] \tag{6}$$

The total price for the workgroup is given by $wp_j + p_k$[11] and the income sensitivity of utility is measured by λ_i. The term $x_j\beta_i$ captures the quality assessment of buyer i about the PC from producer j. The characteristics of the PC are captured by the vector x_j while the number of PCs that the buyer needs for his workgroup are captured by w. The quality of the server purchased is reflected by the expression $A_k y_k \gamma_i$. The vector y_k represents the attributes of the server and the server software. In the case of $j = 0$, x_j is the null vector, while in the case of $k = 0$, y_k is the null vector. The diagonal matrix A_k captures the degree to which each dimension on the server interoperates with the PC operating system (Windows). We normalize quality by assuming that $A = I$ whenever server producer k has the Windows operating system installed. We assume that A_k is the same for all non-Microsoft servers. In the simplest case of one server characteristic we can think of non-Microsoft server quality as a which indicates the degree to which a server running a non-Windows operating system interoperates with Windows on PCs. The terms ξ_j and ξ_k represent PC model j and server

[10] For notationally simplicity we are associating one producer with one PC hardware type. In the empirical work we of course allow for multi-product firms.

[11] We can allow for two part tariffs by having p_k take the form $p_k(w) = p_{k1} + wp_{k2}$. This can allow for typical pricing structures in which there is a fixed price for the server operating system and a pricing component based on the number of users (i.e. w Client Access Licences - CALs - have to be purchased). We can accommodate such pricing without any problems in our approach. All that is really important for the pricing structure is that there is some fixed component to the pricing of the monopolist's server operating system. For simplicity we will exposit all of the analysis below ignoring licenses based on client user numbers.

specific model k unobserved factors in utility.

The term η_{ijk} represents a buyer specific shock to utility for the particular workgroup solution selected. Given that we make this term workgroup specific, this shock captures all of the potential complementarity between the PCs and the servers in a workgroup. For an individual i there is "complementarity" between workgroup PCs j and the workgroup server k if $\eta_{ijk} > \eta_{i0k} + \eta_{ij0}$. As a benchmark consider the case in which $\eta_{ijk} = \eta_{i0k} + \eta_{ij0}$ and η_{i0k} and η_{ij0} are independent. In this case there is no interaction between the valuation of the PCs and of the server so that the demand of customer i for PCs will be independent of the demand of customer i for the server. If, in contrast η_{ijk} are independent across jk pairs, there will be some customers who experience complementarity between a particular type of PC and a particular server type, while the opposite may be the case for others. The distribution of η_{ijk} models both a horizontal characteristic between different brands and heterogeneity in the complementarity between PCs and servers.[12] This heterogeneity in the valuation of complementarities plays a crucial role in our discrimination explanation of foreclosure (see next section).

In the empirical section we will put some more structure on the distribution of the η_{ijk} that allows us, in particular, to shift the mean complementarity effect between PCs and servers. Note that it is not clear a priori in which direction such an effect would go.[13]

Following BLP we allow random coefficients on the parameter vector $\theta_i = (\beta_i, \gamma_i, \lambda_i)$. Heterogeneity in the parameters in the population will be critical for our theory as we will show below, because foreclosure incentives arise in this set up from the ability of the monopolist to use PC and server operating system pricing for second degree price discrimination. This set up includes some important abstractions from reality. In particular, we assume that the purchase decisions are only about the setup of a whole "workgroup".[14]

[12] The distribution of the error term η_{ijk} therefore does not just model horizontal product differention between bundles that are substitutes for each other but also imperfect complementarities at the aggregate demand levels.

[13] The heterogeneity in complementarity would reflect for example the varying outside options of not buying a server, for example, because keeping an existing server might be a relatively attractive option. In such a case the complementarity realization would be low. Similarly, for a firm in which most employees have an existing printer, the value from networking through a server might be low etc. All of these effects will be captured through the distribution of the η_{ijk}.

[14] If server systems are used for serving one workgroup we effectively assume that the whole system is

3.2. A simple example. We start with a simple example with minimal heterogeneity. It illustrates how the desire to maintain price discrimination between different buyer types can lead to strong foreclosure incentives. We abstract for now from the purchase of the hardware associated with the purchase of the operating systems and discuss the issues as if only the PC operating system and the server operating system were involved. To bias our results as much as possible in the direction of the one monopoly profit theory we assume that the PC and server operating systems are perfect complements in the market in the sense that they are consumed in fixed proportions one to one, i.e. $w = 1$.[15] There is a single PC operating system, $J = 1$, and there are two server operating systems: M of the monopolist of the PC operating system and M' of the rival ($K = 2$).

We will assume there are two groups of customers. Type L customers have inelastic demand for PC's, type H customers have a high elasticity of demand. For concreteness, we can think of the type L customers as being large businesses, whereas the type H customers being small businesses.[16] To make the model as simple as possible we assume $\beta = \xi_j = \xi_k = 0$ for both groups and our assumption gives i.e. $\lambda^H > \lambda^L$. We also assume away all heterogeneity in the evaluation of server characteristics setting $\gamma = 1$. The central assumption necessary for the price discrimination result is that the more price sensitive type H group are less likely to purchase a server than the type L group. Or put it differently, that large businesses are more likely to buy a server than small businesses. Formally $\eta^L_{i0k} = \eta^L_{ij0} = -\infty$ for all buyers in that group, but $\eta^L_{ijk} > -\infty$ for $j \neq 0$ and $k \neq 0$ has a full distribution in the population type L customers. We also assume that servers are perfect substitutes up to the quality difference, i.e. $\eta^L_{ijM} = \eta^L_{ijM'}$ for all i and j. Servers have no value for type H

scalable by the factor $1/w$. Effectively, we are capturing all potential effects of pre-existing stocks of servers and PCs (together with their operating systems) by the distribution of η_{ijk}. Since we are assuming that this distribution is invariant over time, we are implicitly assuming that (*modulo* some time trend) the distribution of stocks of computers is essentially invariant. Also note that scalability of workgroups implies that we are not allowing for any difference in firm size directly. All such differences will be incorporated into the distribution of the η_{ijk} and the parameters $(\beta_i, \gamma_i, \lambda_i)$ including a (heterogenous) constant. The idea is to make the relationship between size and purchases less dependent on functional form. However, we can distinguish between large and small businesses in the data. We can therefore control to some extent for firm size by relying on different distributional patterns across the sub-groups for which we are segmenting the market.

[15] All of these assumptions are made to make the benchmark model extremely simple. We relax all of them in the empirical model that we estimate below.

[16] Genakos (2004) shows in the context of the US PC market that indeed these customer segments have different aggregate demand elasticities, with large businesses being relatively more inelastic.

customers.[17] Given these assumptions, only type L customers have demand for servers, i.e. type L (on average) values a $\lambda_k y_k$ with $\lambda_{M'} = \lambda < \lambda_M = 1$. We can then write demand for the PC operating system for each group as $q^L(\lambda^L\omega - \max_k\{\lambda_k y_k - \lambda^L\omega_k\})$ and $q^H(\lambda^H\omega)$. We assume that the demand functions $q^H(\cdot)$ and $q^L(\cdot)$ are log-concave and the absolute magnitude of the PC price elasticity of demand for type H is greater than that of the type L's, i.e. $\frac{\partial q^H(\omega)}{\partial\omega}\frac{\omega}{q^H} < \frac{\partial q^L(\omega)}{\partial\omega}\frac{\omega}{q^L}$. This will be the essential driving force of our results. For simplicity, we assume that marginal costs for producing an operating system are zero.

The one monopoly profit theory. Let us first assume that the PC operating system monopolist can set different prices for the two types of customers. The price to the type H segment is then simply the monopoly price and we only have to consider the pricing behavior for type L customers. Equilibrium then has the following properties:

Proposition 1. *If $y_M > \lambda y_{M'}$ (i.e. the monopolist has higher quality than the rival) the monopolist sells the server operating system and makes the profits of a monopolist in both markets with server operating system quality level y_M. If $y_M < \lambda y_{M'}$ (i.e. the monopolist has lower quality than the rival) then the rival sells the server operating system in all Pareto undominated equilibria. There is a continuum of Pareto undominated equilibria (among the firms). The worst for the monopolist has the same profits as in the case $y_M > \lambda y_{M'}$, while the best has the monopoly profits of a firm that can offer the quality $\lambda y_{M'}$.*

Proof. See Appendix A ■

This proposition illustrates the one monopoly profit theory. The presence of a server rival can never reduce the profits of the monopolist in his monopoly market. Given the monopoly on the PC operating system, the monopolist can extract at least the monopoly profits achievable with his own system. Furthermore, complete exclusion of a superior technology is not possible.[18]

[17] We can achieve that by assuming at $\eta^H_{ijk} = -\infty$ for $k \neq 0$ and $\eta^H_{ij0} > -\infty$ with a non-degenerate distribution. This assumption is not made for purpose of realism. All that is needed in the general model in the econometric section is that the small business segment has a larger probability of choosing a no server option.

[18] There is a caveat to this statement: The profits of the alternative system are not fully extractable in all equilibria and this leads to the typical problem of oligopoly in complementary products. The overall price

The one monopoly theory result implies that there is no incentive to degrade the quality of a rival by, for example, decreasing interoperability with a rival's server operating system. To see this, note that quality degradation of the rival's product has no effect on profits if $y_M > \lambda y_{M'}$. Since the monopolist can always extract all benefits through the PC operating system price, the presence of a rival in the server operating system market does not limit market power of the PC operating system monopolist at all. For the case $y_M < \lambda y_{M'}$ a rigorous conclusion is more difficult, since we have a continuum of equilibria. However, a reduction of the quality of the rival can only harm the PC operating system monopolist in this case, since it eliminates the best available equilibria from the point of view of the monopolist and does not enhance monopoly profits in the worst case. Furthermore, all equilibria in which the monopolist can extract some of the rents from the rival's quality improvement use weakly dominated strategies for the monopolist. It is therefore standard in the Bertrand equilibrium literature to exclude such equilibria, leaving only the equilibrium with the lowest profit for the monopolist in the case of $y_M < \lambda y_{M'}$.[19] In this sense we then have a model in which a pure one monopoly profit theory holds.

Second degree price discrimination and foreclosure incentives. We now show that these conclusions dramatically change, when the PC operating system monopolist *cannot* price discriminate on its PC operating systems sales. First we consider the case when the PC monopolist also is the only player in the server market. Then we consider the case when the monopolist faces a rival in the server market

Monopoly in the server market. If the firm controlling the PC operating system also has a monopoly in the server operating system market, it could still achieve price discrimination through the use of the server price. To see this note that optimal price

of the equilibrium system is too high relative to the price a monopolist would charge given he owned the better software. This effect is there in all but the equilibrium in which the monopolist extracts all the profits. However, greater profit extraction by the monopolist would also generate problems of dynamic efficiency, because they would reduce the incentives to develop alternative products. This illustrates that, even in markets in which the one monopoly profit theory holds, concerns about the extraction of rents due to market power are not innocuous because of issues of dynamic investment incentives.

[19] More precisely, we exclude weakly dominated strategies that are not limit points of sequences of undominated strategies as, for example, Allen, Deneckere, Faith and Kovenock (2000).

setting by a monopolist would imply:

$$\omega^* = -\frac{1}{\varepsilon_\omega^H}$$

and

$$\omega^* + \omega_M^* = \frac{1}{\varepsilon_\omega^L}$$

where ω^* is the optimal PC price and ω_M^* is the monopolist's optimal server price.

If the high elasticity customers have a strictly more elastic demand function than the low elasticity customers, the price on the server operating system will exceed the server quality, i.e. $\omega_M^* > y_M$, in order to exploit the inelastic PC demand elasticity of the large business market segment. Hence, the monopolist is using the server purchases of the large business segment to achieve a form of second degree price discrimination.[20] Note, that second degree price discrimination requires raising the price of the server. What is critical here is that the absolute elasticity of demand is lower for the customer group with the higher willingness to pay for servers. In other words, price discrimination allows to extract more from customers with low λ and high value of adding a server to the workgroup (i.e. either strong complementarity or high γ).

Duopoly in the server market. In the case of duopoly in the server operating system market, competition eliminates the possibility of extracting rents through the server price and thus undermines the scope for second degree price discrimination. This generates an incentive to reduce the quality of the rival in order to restore the ability to price discriminate. To show this formally recall that firms simultaneously set operating system prices, ω, ω_M, and $\omega_{M'}$. Then customers decide which firm to purchase from.[21] As before, we consider two cases: (1) the monopolist has the higher quality server software ($y_M > \lambda y_{M'}$) and (2) the

[20] Since we have only two types of customers the outcome coincides with the outcome of third degree price discrimination in this case. If we had more groups with different demands we would still only have two prices and could not achieve the outcome of third degree price discrimination.

[21] This is the condition of a standard Bertrand game with differences in quality. Note that we do not assume that consumers split evenly between firms in the case of indifference, which would lead to an equilibrium existence problem. Instead we allow for the possibility that buyers all go to one firm despite indifference. The equilibria we obtain this way are the same as those in a model with a smallest money unit, when the size of the money unit goes to zero. These are therefore equilibria that are robust to the strategy space.

monopolist has the lower quality server software ($y_M < \lambda y_{M'}$).

First, consider the case in which the monopolist has the high quality software. The monopolist then cannot set a price exceeding its quality advantage, i.e. $\omega_M \leq y_M - \lambda y_{M'}$. Suppose otherwise, then $\omega_M - (y_M - \lambda y_{M'}) > 0$ and firm M' could induce the sale of its product at some strictly positive price $\omega_{M'}$ slightly below $\omega_M - (y_M - \lambda y_{M'})$ and make strictly positive profits. Since in any equilibrium in which firm M' makes sales $\omega_{M'} \geq 0$ it follows that the monopolist could move sales to itself by undercutting by some small amount, ι, i.e. setting $\omega_M = (y_M - \lambda y_{M'}) + \omega_{M'} - \iota > 0$, which would generate strictly positive profits on the server sales and slightly increase demand for PC operating systems to $q^L(\omega + \omega_M - \lambda y_{M'} - \iota)$. Hence, such a move is strictly profit improving, contradicting the assumption that firm M' makes the server sales. Therefore, in equilibrium $\omega_M \leq y_M - \lambda y_{M'}$. This shows that the monopolist is unable to fully extract monopoly rents from second degree price discrimination. Given the assumption of log-concavity of demand, demand elasticities are increasing in price and by standard arguments from third degree price discrimination it now follows that the price set by the monopolist in the PC operating system market is $\omega^e > \omega^*$. In the server market the price is $\omega_M^e = y_M - \lambda y_{M'}$, where $\omega_M^e < \omega^* - \omega^e + \omega_M$, where in all cases the superscript e indicates equilibrium values. The rival sets $\omega_{M'} = 0$.

Note that the rival limits the margin on the server operating system for the monopolist to $y_M - \lambda y_{M'}$. This in turn limits the ability of firm M to profits strictly below those under second degree price discrimination by a monopolist. It is now clear how interoperability degradation will help the monopolist. By reducing the quality of its rival through degraded interoperability with the PC operating system, the monopolist can increase its market power in the server operating system market, increasing the scope for price discrimination:

Proposition 2. *Suppose $y_M > \lambda y_{M'}$, (i.e. the monopolist has the higher quality server). Then the monopolist has an incentive to degrade the quality of its rival at least up to the point that $y_M - \lambda y_{M'} = \omega_M^*$.*

Proof. See Appendix A ■

Reducing a rival's quality increases the market power of the monopolist in the server market and allows it to get closer to the "full" monopolist's profit.[22] The only role of competition in the case $y_M > \lambda y_{M'}$ is to limit the rent extraction possibilities of the PC operating system monopolist on his server product. This simply prevents second degree price discrimination. While the degradation of interoperability is anti-competitive in the sense that it reduces the ability of firm M' to compete, all that competition achieves is to limit second degree price discrimination. Since the welfare effects of such discrimination are ambiguous one might doubt that there is any serious harm to welfare. However, in the case $y_M < \lambda y_{M'}$ the same incentives are at play and can potentially generate dramatic welfare effects:

Proposition 3. *Suppose $y_M < \lambda y_{M'}$ (i.e. rival has the higher quality server software). Then the PC operating system monopolist wants to reduce the quality of the rival's product to the level $\omega_M^* - y_M$. Interoperability is reduced whenever this allows the monopolist to reduce the rival's quality level below its own through degraded interoperability.*

As long as it is possible to make interoperability infeasible, it will be in the interest of the PC operating system monopolist to exclude an arbitrarily better product of the rival from the market in order to gain market power in the server market. The reason is that the monopolist cannot extract the benefits from the improved server operating system through the PC price in any form. His profits are the same if he uses his own product. But by reducing the quality of the competitors product through interoperability degradation it is again possible to price discriminate more effectively. The PC operating system monopolist is therefore willing to induce arbitrarily high social costs on buyers to achieve more market power. This effect of excluding superior technologies through interoperability degradation is the central social loss from such strategies.

The extreme result of this simple example arises because the only effect at work is the market share shifting effect of interoperability degradation. Since in equilibrium each firm

[22] We could rewrite the above model so that there are no differences in the qualities of the server operating system vendors but that their marginal cost of selling an operating system differed. All results would go through simply replacing the wording "high quality" by "low cost" and the word "low quality" by "high cost". Mathematically there is no difference between raising rival costs and reducing rival quality in this model.

extracts the full benefit of its quality improvement, demand for PC operating systems is essentially unaffected by reducing interoperability. This will be different if there is genuine product differentiation between server operating systems. However, the basic second degree price discrimination effect illustrated with this simplified model will be present in any general model in which there is some subset of customers with low price sensitivity λ and a high incremental value of adding a server.[23] It is that heterogeneity in the demands that drives the incentives for second degree price discrimination.

3.3. General Implications of the Model. The firm like the econometrician cannot estimate the demands of all customers separately. However, since sufficient heterogeneity in price sensitivity and PC/server complementarity creates the incentives for second degree price discrimination, we can still estimate foreclosure incentives from a model that allows unobserved heterogeneity over these preference parameters in the population. To derive demand, we first define the set of realizations of the unobserved variables that lead to the choice of a given system jk across all types of customers:

$$B_{jk}(x_j, y_k, p_j, p_k, a, w) = \{\theta_i, \xi_j, \xi_k, \eta_{ijk} | u_{ijk}(w) \geqslant u_{ilm}(w), for\ all\ l, m\}$$

Using the population distribution function $dP(\theta)$, we can aggregate demands to generate the probability that a buyer of workgroup size w will purchase system jk as:

$$s_{jk}(w) = \int_{B_{jk}(x_j, y_k, p_j, p_k, a, w)} dP(\theta_i, \xi_j, \xi_k, \eta_{ijk} | w) \tag{7}$$

where s_{jk} is the probability of buying a PC-server bundle jk. The total demand for PCs of type j from users of system jk is then given by $q_{jk} = L \int w s_{jk}(w) d\Upsilon(w)$, where $\Upsilon(w)$ is the population distribution of workgroup sizes and $L \int w d\Upsilon(w)$ is the maximum number of PCs that could possibly be sold to all buyers of all types. This means that L is the maximal number of potential workgroups (market size). To generate the demand for a PC of type

[23] Such individuals must always exhibit some complementarity between the products. But even if all buyers would have the same positive level of complementarity, heterogeneity in γ would allow for the type of price discrimination that drives our model.

j, we aggregate these demands across all server options to $q_j = L \int w s_j(w) d\Upsilon(w)$, where $s_j(w) = \sum_{k=0}^{K} s_{jk}(w)$. The demand for server k from users of system jk is analogously given by $q_k = L \int s_k(w) d\Upsilon(w)$ where $s_k = \sum_{j=0}^{J} s_{jk}$.[24] The demand for PC operating systems is then given by aggregating over all PC sales: $q = L \int w s(w) d\Upsilon(w)$, where $s = \sum_{j=1}^{J} s_j$. Let M be the set of server sellers k that run the server operating system sold by the same firm as the PC operating system. Then the demand for server operating systems for firm M is given by $q_M = L \int \sum_{k \in M} s_k(w) d\Upsilon(w)$ and the demand for all servers is given by $q^S = L \int \sum_{k=1}^{K} s_k(w) d\Upsilon(w)$.

The most fundamental prediction of the simplified model in Sub-section 3.2 was that price discrimination will lead to relatively higher server margins. In our empirical work we will not be able to observe operating system prices directly. However, we will use the structural model of demand for the monopolist's products infer the relative markup derived from equations (4) and (5) from the data:

$$\frac{\omega_M - c_M}{\omega - c} = \frac{\frac{q}{q_M}\varepsilon_{\omega_M} - \varepsilon_\omega}{\frac{q_M}{q}\varepsilon_\omega^M - \varepsilon_{\omega_M}^M} \tag{8}$$

This is one critical component for assessing the incentives for foreclosure through interoperability degradation. To see that the relative server margin is pushed up precisely by heterogeneity we can look more carefully at the exact form of this expression for our specific model of buyer heterogeneity (see Appendix A). The sign of ω_M is determined by noting that

$$\frac{q}{q_M}\varepsilon_{\omega_M} - \varepsilon_\omega = \int \left[\frac{q(\theta)}{q} - \frac{q_M(\theta)}{q_M} \right] \left[\bar{\varepsilon}_\omega - \varepsilon_\omega(\theta) \right] dP(\theta) \tag{9}$$

where

$$\bar{\varepsilon}_\omega = \int \varepsilon_\omega(\theta) dP(\theta)$$

Hence, the price cost margin on servers will be positive if the own price elasticity of the PC operating system, $\varepsilon_\omega(\theta)$, is positively correlated with $\frac{q(\theta)}{q} - \frac{q_M(\theta)}{q_M}$, i.e. if on average buyers

[24] Note that we are summing up from 0 to J here, because we allow for the possibility that a buyer has an existing PC work group and simply adds a server.

with more elastic demand (a more negative $\varepsilon_\omega(\theta)$) have higher market share in PC purchases than the monopolist's server purchases. This will happen if firms with more elastic demand (for example firms with higher λ) have a higher likelihood of purchasing PCs than servers. This is the general expression of the idea that firms that are more price sensitive have lower complementarity with servers and lower valuations for server quality.

Note that equation (9) will be zero if there is no heterogeneity. When there is no heterogeneity in demand, the monopolist does best by setting the price of the server at marginal cost and extracting all surplus through the PC operating system price. In that case there is no incentive to price discriminate and, hence, no incentive to foreclose competitors in the server market. The greater the price discrimination incentive, the greater the server markup will be.

For ω, the price of PC operating systems, we obtain that it is proportional to:

$$\begin{aligned} \frac{q_M}{q}\varepsilon_\omega^M - \varepsilon_{\omega_M}^M &= \int \alpha w s_{00}(\theta, w) \begin{pmatrix} \frac{M(\theta,w)-q_M(\theta,w)}{wM(\theta,w)-q(\theta,w)} \frac{q_M(\theta,w)}{q_M} \\ +\frac{q_M(\theta,w)}{q(\theta,w)} \frac{q(\theta,w)}{q} \end{pmatrix} dP(\theta) \\ &\quad -\frac{q_M}{q}\left[\frac{q}{q_M}\varepsilon_{\omega_M} - \varepsilon_\omega\right] \end{aligned} \tag{10}$$

The term in the second line implies that the server margin weighted by the relative quantity of servers to PCs reduces the PC operating system margin. Hence, the complementarity means that whenever there is a positive server margin, the PC operating margin will be smaller than what one would conclude from looking at the PC operating system market alone. Given estimates for the parameter vector θ and the parameters of the distribution $P(\theta)$ these expressions can be calculated directly. We can therefore infer the margins on the monopolist's operating systems from estimating demand alone.

4. Econometrics

4.1. Baseline Model. We formulate an empirical model of demand by aggregating a discrete choice model of individual consumer behavior that allows us to model complementarity using market level data. The baseline empirical model follows the theory by allowing customers to select either (i) the outside good, (ii) a PC or (iii) the bundle (a "workgroup" purchase of w PCs and one server).

For purposes of estimation it is useful to rewrite the conditional indirect utility, $u_{ijk}(\theta)$ in (6) as the sum of mean utility and individual specific effects. Denoting the unknown parameter vector as $\theta = (\theta_1, \theta_2, \theta_3, \theta_4)$ we have:

$$u_{ijk}(\theta) = \delta_j(\theta_1) + \mu_{ij}(\theta_2) + \delta_k(\theta_3) + \mu_{ik}(\theta_4) + \xi_{jk} + \epsilon_{ijk}. \tag{11}$$

The first term, δ_j, is mean utility derived from consuming PC j, which is common to all consumers. It is given by:

$$\delta_j = x_j\beta - \lambda p_j + \xi_j, \tag{12}$$

where x_j and β are vectors of the observed product characteristics and the associated taste parameters respectively, λ is the mean sensitivity of the brand to the price, p_j of PC j and ξ_j denotes utility derived from characteristics observed by the consumers and the firms, but not the econometrician.

Unobserved product characteristics include unquantifiable variables such as firm or brand reputation for reliability, prestige effects or after-sales service quality. Since these characteristics are observed by market participants, they will be correlated with the equilibrium prices causing the price coefficient to be biased towards zero. Instrumental variable techniques can not straighforwardly be applied, given that both p_j and ξ_j enter the market share equation in a nonlinear way. Berry (1994) develops a general method that allows the use of instrumental variables to a large class of discrete choice models.

The second term in equation (11), μ_{ij}, represents a deviation from the mean utility. This is individual specific and can be written as:

$$\mu_{ij} = \sum_h \sigma_h^{PC} x_{jh} \nu_{ih}^{PC} + \sigma_p^{PC} p_j \nu_{ip}^{PC} \tag{13}$$

where x_{jh} is the hth characteristic of product j, for $h = 1, ..., H$ and σ_h^{PC}, σ_p^{PC} are unknown coefficients. The vector $\nu_i^{PC} = \left(\nu_{i1}^{PC}, ..., \nu_{iH}^{PC}, \nu_{ip}^{PC}\right)$ represents each consumer's $H+1$ idiosyncratic tastes for the H observed characteristics and the associated price. It is drawn from a multivariate normal distribution with zero mean and an identity covariance matrix.[25] Notice that μ_{ij} depends on the interaction of consumer specific preferences and product characteristics. More precisely, each consumer i derives $(\beta_h + \sigma_h \nu_{ih}) x_h$ utility from every hth product characteristic. BLP show that allowing for substitution patterns to depend on consumer's heterogeneous tastes (i.e. $\eta_{ij} \neq 0$) is important for realistic demand elasticities.[26] For example, consumers who attach a higher utility to laptop computers would more likely substitute towards other laptops rather than desktops. The notation is symmetric for servers:

$$\delta_k = A_k y_k \gamma - \lambda p_k + \xi_k, \tag{14}$$

$$\mu_{ik} = \sum_h \sigma_h^S y_{kh} \nu_{ih}^S + \sigma_p^S p_k \nu_{ip}^S \tag{15}$$

The specification of the demand system is completed with the introduction of an "outside good". Consumers are allowed to not purchase any of the bundles offered by these firms. Otherwise, a uniform price increase would not change the quantities purchased. The indirect utility of the outside option is:

$$u_{i0} = \xi_0 + \sigma_0^{PC} \nu_{i0}^{PC} + \sigma_0^S \nu_{i0}^S + \epsilon_{i0}. \tag{16}$$

where the price of the outside good is normalized to zero. Since relative levels of utility cannot be identified, the mean utility of one good has to be normalized to zero. As is customary, we normalize ξ_0 to zero. The term ν_{i0} accounts for the outside alternatives' unobserved variance.

[25] The choice of this distribution is ad hoc. Although the multivariate normal is the most popular choice (e.g. BLP, Nevo, 2000, 2001), other possibilities have also been explored (e.g., Petrin, 2002). There is no evidence that the choice of this assumption affects the estimated coefficients in any fundamental way.

[26] When μ_{ij} is zero, the only source of heterogeneity among consumers is based on the i.i.d. ϵ_{ij}'s. In terms of elasticities, that implies that all the consumers have the same expected ranking over products. In other words, consumers would substitute more towards the most popular products independently of their characteristics and the characteristics of the products they bought previously.

Each consumer is assumed to purchase one good per period[27] from the available choice set, which provides her with the highest utility. Given the assumption on the distribution of ϵ_{ijk}, the probability that consumer i purchases PC j is given by the multinomial logit choice probability (McFadden, 1973)[28]:

$$s_{ij} = \sum_{k=0}^{K} \frac{e^{\delta_j+\mu_{ij}+\delta_k+\mu_{ik}}}{1+\sum_{l=1}^{J}\sum_{k=0}^{K} e^{\delta_j+\delta_k+\mu_{ij}+\mu_{ik}}} \tag{17}$$

$$= e^{\delta_j+\mu_{ij}} \sum_{k=0}^{K} \frac{e^{\delta_k+\mu_{ik}}}{1+\sum_{j=1}^{J}\sum_{k=0}^{K} e^{\delta_j+\mu_{ij}+\delta_k+\mu_{ik}}} \tag{18}$$

Market shares for each product, s_j (and s_k), are obtained by aggregating over customers and their vectors of unobservable tastes. Following the theory, we build in complementarity by allowing server purchases only in conjunction with PCs. We consider two alternative models in the extensions section: (i) retsricting the model by ruling out PC-only purchases ("strong complementarity"), and (ii) allowing server only purchases and estimating freely the degree of complementarity (or substitutability) between the two product categories ("free complementarity"). Our qualitative results are robust to both of these alternatives.

Finally, to connect the empirical framework with the theoretical model, we model the interoperability parameter (a) as a multiplicative effect that customers derive from having a Microsoft (M) server:

$$\delta_k = y_k\gamma_1 + \gamma_2 M + \gamma_3(My_k) - \lambda p_k + \xi_k$$

where M is a dummy variable equal to one if the server runs a Microsoft operating system and zero otherwise. In that way, the interoperability parameter is captured by a combination of the estimated coefficients and therefore we can calculate the "relative output effect" in one step. Given this parameterization, the relationship between the utility foundation of

[27] Although this assumption seems reasonable for home or small business users, it might not be applicable to the large business segment. Hendel (1999), for example, observes PC purchases of large firms and models explicitly the choice of multiple products. However, without more dissagregated information his techniques cannot be applied to the current data. Hence, if this phenomenon is widespread this model can be seen as a first approximation to the true choice model.

[28] In principle we could allow a ξ_{jk} to reflect unobserved bundle-specific utility. However, the dataset does not identify the proportions of server of brand k being used by customers who purchased brand j of PCs. Thus we consider estimation of the simpler model where we abstract from the unobserved "cross effects".

equation (6) and the estimates is that $\gamma_3 = \gamma(1-a)$ and $\gamma_1 = a\gamma$, where $0 \leq a \leq 1$ is the interoperability parameter.[29] If there were no interoperability limitations between between Microsoft and non-Microsoft operating systems ($a = 1$), then γ_3, the coefficient on the interaction variable, would be insignificantly different from zero.

4.2. Estimation. Our estimation strategy closely follows the spirit of the BLP estimation algorithm, but modifies it so that multiple product categories can be accommodated. In essence, the algorithm minimizes a nonlinear GMM function that is the product of instrumental variables and a structural error term. This error term, defined as the unobserved product characteristics, $\xi = (\xi_j,\ \xi_k)$, is obtained through the inversion of the market share equations after aggregating appropriately the individual consumer's preferences. However, the presence of multiple product categories means that we need to compute the unobserved term, ξ, via a category-by-category contraction mapping procedure.

Define $\widetilde{\theta} \equiv (\sigma_h^{PC}, \sigma_h^{S}, \sigma_p^{PC}, \sigma_p^{S})$, the vector of non-linear parameters, i.e., the random coefficients on characteristics and price for PCs and servers. Let r be the set of variables that we are allowing non-linear parameters (e.g. x_j, y_k, p_j, p_k). Let $\delta = (\delta_j, \delta_k)$, $\xi = (\xi_j, \xi_k)$, $\nu_i = \left(\nu_i^{PC}, \nu_i^{S}\right)$ and $\mu_i = (\mu_{ij}, \mu_{ik})$.

Our iterative procedure is as follows:

Step 0: Draw the idiosyncratic taste terms ν_i (these draws remain constant throughout the estimation procedure) and starting values for $\widetilde{\theta}$.

Step 1. Given $(r, \widetilde{\theta})$, calculate μ_i.

Step 2. Given (δ, μ_i), calculate individual consumer product market shares for PCs and servers and aggregate to get market shares for each brand. We use a smooth simulator by integrating the logit errors analytically.

Step 3. Given $\widetilde{\theta}$, we need to numerically compute the mean valuations, δ, that equate the observed to the predicted brand market shares. Due to complementarity between the PCs and servers, we compute each product category's mean valuation conditional on the other

[29] We allow γ_2 to be freely estimated as it could reflect the higher (or lower) quality of Windows compared to other operating systems. Alternatively, γ_2 could also reflect interoperability limitations. We examine this possibility in a robustness exercise.

category's mean valuation. Specifically, it consists of the following sequentially iterative substeps:

Substep 3.0 Make an initial guess on δ and set $\delta_{old} = \delta$.

Substep 3.1 Compute δ_j given δ_k using BLP's contraction mapping. Update δ.

Substep 3.2 Compute δ_k given δ_j using BLP's contraction mapping and update δ.

Substep 3.3 Check if $\delta_{old} =$ updated δ. If yes, go to step 4. Otherwise, set $\delta_{old} = \delta$ and go to substep 3.1.

Step 4. Given δ, calculate ξ and form the GMM.

Step 5. Minimize a quadratic form of the residuals and update.

We also estimated two other variants of this algorithm. The first one reiterates one additional time substeps 3.1 and 3.2 to make sure that there is no feedback from PCs to server mean valuations. This variant takes slightly more computational time. The second variant instead of updating the mean valuations for each product category in substeps 3.1 and 3.2, always uses the initial estimates (taken from the simple logit IV regression). This variant takes more computational time, but it is more robust to starting values.[30] The weighting matrix in the GMM function was computed using a two-step procedure. To minimize the GMM function we used both the Nelder-Mead nonderivative search method and the faster Quasi-Newton gradient method based on an analytic gradient. We combine all these methods to verify that we reached a global instead of a local minimum.

Standard errors corrected for heteroskedasticity are calculated taking into consideration the additional variance introduced by the simulation.[31] In our benchmark specification we draw a sample of 150 consumers, but we also experiment with more draws in our robustness section. Confidence intervals for nonlinear functions of the parameters (e.g., relative output and relative margin effects) were computed by using a parametric bootstrap. We drew repeatedly from the estimated joint distribution of parameters. For each draw we computed

[30] In all contraction mappings, we defined a strict tolerance level: for the first hundred iterations the tolerance level is set to 10E-8, while after every 50 iterations the tolerance level increases by an order of ten.

[31] We do not correct for correlation in the distrurbances of a given model across time as this is unlikely to be material. First, because firm fixed effects are included in the estimation. Second, because there is a high turnover of products, with each brand model observation having a very short lifecycle compared to other durables like autos.

the desired quantity, thus generating a bootstrap distribution.

4.3. Identification and instrumental variables. Identification of the population moment condition is based on an assumption and a vector of instrumental variables. Following BLP we assume that the unobserved product level errors are uncorrelated with the observed product characteristics. We can therefore use functions of observed computer and server characteristics (in particular sums of characteristics for the firm across all its products and sums of the characteristics of competing firms). Given the previous exogeneity assumption, characteristics of other products will be correlated with price, since the markup for each model will depend on the distance from its nearest competitors. To be precise, for both PCs and servers we use the number of products produced by the firm and the number produced by its rivals as well as the sum of various characteristics (PCs: speed, RAM, hard drive; servers: RAM, rack optimized, number of racks, number of models running unix) of own and rival models.[32]

We also examine the robustness of our results by varying the type of instruments used. First, we experimented using alternative combinations of computer characteristics. As we show in the robustness section, our results qualitatively remain unaffected. Second, we use hedonic price series of computer inputs, such as semi-conductor chips, which are classic cost shifters. The results are robust to these two alternative sets of instruments, but they were less powerful in the first stage. Finally, we followed Hausman (1996) and Hausman et al (1994) and used model-level prices in other countries (such as Canada, Europe or Japan) as alternative instruments. These instruments were powerful in the first stage, but there was evidence from the diagnostic tests that these instruments were not valid (see Genakos, 2004 and Van Reenen, 2004, for more discussion).

Finally, one important limitation of using aggregate data is that we cannot separate true complementarity (or substitutability) of goods from correlation in customers' preferences (see Gentzkow, 2007). Observing that firms that buy PCs also buy servers might be evidence that the two product categories in question are complementary. It might also reflect the fact that

[32] All PC instruments were calculated separately for desktops and laptops following the spirit of the Bresnahan, Stern and Trajtenberg (1997) study of the PC market.

unobservable tastes for the goods are correlated - that some firms just have a greater taste for "computing power".[33] However, notice that for our purposes such a distinction does not make a major difference to the theoretical results - so long as there is a correlation between customers' heterogeneous preferences for PCs and their probability of buying servers, the incentive to leverage can exist.

4.4. Alternative approaches to modelling complementarity. Recent work by Gentzkow (2007) and Song and Chintagunta (2006), also provide empirical oligopolistic models that allow complementarity across product categories. Gentzkow (2007) was the first to introduce a complementarity parameter in a discrete setting. By observing individual purchase level data, he is able to model the correlation in demand between on-line and off-line versions of the *Washington Post* in a flexible way that allows for rich substitution patterns. Song and Chintagunta (2006), extend Gentzkow by allowing for a common complementarity/substitution parameter across product categories and apply it on aggregate data. Our basic model is more restrictive in that complementarity between PCs and servers is built in rather than estimated. This choice was driven both by our understanding of how the market for "workgroup" purchases operates (firms buy servers not to use them on a stand alone basis but to coordinate and organize PCs), but also from data considerations.[34] However, in the robustness section of our results we also experiment by estimating a variant of the Song and Chintagunta model that allows for a freely estimated complementarity/substitutability parameter.

In our baseline model consumers are assumed to buy either a PC, a bundle of a server and PC or the outside good. We also analyze two alternative empirical models: (i) one that assumes "strong" complementarity between the two product categories: i.e. firms buy either a bundle or nothing, and (ii) a more general model that allows the data to determine the degree of complementarity or substitutability between the two products.

[33] Combining our aggregate market share data with detailed firm level choice data should allow us in principle to separate these two effects, which is something we are plannning for future work using the micro Harte-Hanks dataset.

[34] The parameter that governs compelmentarity in Song and Chintagunta (2006) is identified from the time dimension. The limited time span (21 quarters) means that empirically it is really hard to identify precisely such a parameter in our data.

Under “strong complementarity”, we write our previous model as:

$$s_{ij} = e^{\delta_j + \mu_{ij}} \sum_{k=1}^{K} \frac{e^{\delta_k + \mu_{ik}}}{1 + \sum_{j=1}^{J} \sum_{k=1}^{K} e^{\delta_j + \mu_{ij} + \delta_k + \mu_{ik}}} \tag{19}$$

where we are summing up now from 1 to J or K for both PC and servers. The rest of the assumptions and estimation details remain the same as before. Note that this assumption restricts the data more in favor of rejecting any degradation incentives.

Under the "free complementarity" model a bundle includes one and only one alternative model from each product category, (j, k). Denote d^{PC} an indicator variable that takes the value of one if any PC is purchased and zero otherwise; similarly we define d^S to be the indicator for servers. Each customer i, maximizes utility by choosing at each point in time, t, the bundle of products, (j, k), with the highest utility, where utility is given by:

$$u_{ijkt} = V_{ijkt} + \Gamma_t(d_{PC}, d_S) + \epsilon_{ijkt} \tag{20}$$

The first term, V_{ijkt}, is customer and model specific as it is allowed to vary given the choice of particular brands and consumer's characteristics. The second term, $\Gamma(d_{PC}, d_S)$, is specific to the goods bought (PC or server) in the sense that it is not affected by a choice of particular brand once (d_{PC}, d_S) is given and does not vary across consumers. Finally, note that the error term (ϵ_{ijkt}) is bundle-specific. This utility structure allows us to model complementarity and/or substitution at the level of the good, i.e. PC or server, via $\Gamma(d_{PC}, d_S)$.

More specifically, we assume that the utility each consumer derives from the bundle is equal to the sum of utilities of each model involved (for the rest of this sub-section we drop the t subscript to simplify exposition) and is given by:

$$V_{ijk} = \delta_j + \mu_{ij} + \delta_k + \mu_{ik} \tag{21}$$

where $\delta_j, \delta_k, \mu_{ij}, \mu_{ik}$ are the same as before. We also assume that the utility derived from purchasing the empty bundle (the outside good), $(0, 0)$, is zero.

The key element[35] in the $\Gamma(d_{PC}, d_S)$ function is the parameter on $d^{PC}d^S$ (i.e. the indicator of whether a customer buys both a PC and a server), which we label the complementarity parameter, $\Gamma_{PC,S}$. This last parameter is symmetric, i.e. $\Gamma_{PC,S} = \Gamma_{S,PC}$ and captures the extra utility that a customer obtains from consuming these two products together over and above the utility derived from each product independently. We define $\Gamma_{PC,S}$ to be positive for a pair of complements and negative for a pair of substitutes. This model borrows directly from the work of Gentzkow (2007), who was the first to introduce a similar parameter in a discrete setting. Our utility model is more general in that we allow for random coefficients on the model characteristics and prices (Gentzkow does not have price variation in his data). More importantly, our model is designed to be estimated with aggregate market level data. The paper closest to ours is Song and Chintagunta (2006), who also extend Gentzkow, to allow for a common complementarity/substitution parameter and apply it on store level data for detergents and softeners. We differ from Song and Chintagunta in three ways: (i) we specify a different brand and consumer part of the utility that is closer to the original BLP specification, (ii) we use a different set of instruments to address the issue of price endogeneity and (iii) we implement a more robust estimation method. Further model and estimation details are given in Appendix D.

5. Data

Quarterly data on quantities and prices between 1996Q1 and 2001Q1 was taken from the PC Quarterly Tracker and the Server Quarterly Tracker, two industry censuses conducted by International Data Corporation (IDC). The Trackers gather information from the major hardware vendors, component manufacturers and various channel distributors and contains information on model-level revenues and transaction prices.[36] Unfortunately, the information on computer characteristics is somewhat limited in IDC so we matched in more detailed PC and server characteristics from several industry datasources and trade magazines. We concentrate on the top fourteen computer hardware producers with sales in large businesses

[35] There are also linear terms in d^{PC} and d^S.

[36] Various datasets from IDC have been used both in the literature (Foncel and Ivaldi, 2001; Van Reenen, 2004; Pakes, 2003; Genakos, 2004)

in the US market to match each observation with more detailed product characteristics.[37]

For PCs the unit of observation is distinguished into form factor (desktop vs. laptop), vendor (e.g. Dell), model (e.g. Optiplex), processor type (e.g. Pentium II) and processor speed (e.g. 266 MHZ) specific. In terms of characteristics we also know RAM (memory), monitor size and whether there was a CD-ROM or ethernet card included. A key PC characteristic is the performance "benchmark" which is a score assigned to each processor-speed combination based on technical and performance characteristics.[38]

Similarly, for servers a unit of observation is defined as a manufacturer and family/model-type. We also distinguish by operating system, since (unlike PCs) many servers run non-Windows operating systems (we distinguish six other categories: Netware, Unix, Linux, VMS, OS390/400 and a residual category). For servers key characteristics are also RAM, the number of rack slots,[39] whether the server was rack optimized (racks were an innovation that enhanced server flexibility), motherboard type (e.g. Symmetric Parallel Processing - SMP), and chip type (CISC, RISC or IA32). Appendix B contains more details on the construction of our datasets.

Potential market size is tied down by assuming that firms will not buy more than one new PC for every worker per year. The total number of employees in large businesses is taken from the US Bureau of Labour Statistics. Results based on different assumptions about the potential market size are also reported.

Table 1 provides sales weighted means of the basic variables for PCs respectively that are used in the specifications below. These variables include quantity (in actual units), price (in $1,000), benchmark (in units of 1,000), memory (in units of 100MB)as well as identifiers for desktop, CD-ROM and ethernet card. Similarly, Table 2 provides sales weighted means of the basic variables that are used for servers. These variables include quantity (in actual units), price (in $1,000), memory (in units of 100MB), as well as identifiers for rack opti-

[37]These manufacturers (in alphabetical order) are: Acer, Compaq, Dell, Digital, Fujitsu, Gateway, Hewlett-Packard, IBM, NEC, Packard Bell, Sony, Sun, Tandem and Toshiba. Apple was excluded due to the fact that we were unable to match more detail characteristics in the way its processors were recorded by IDC.

[38]Benchmarks were obtained from the CPU Scorecard (www.cpuscorecard.com). Bajari and Benkard (2005) were the first to use this variable.

[39]Rack mounted servers were designed to fit into 19 inch racks. They allow multiple machines to be clustered or managed in a single location and enhance scalability.

mized, motherboard type, each operating system used and number of racks. The choice of variables was guided by technological innovation taking place during the late 1990s, but also developments and trends in related markets (e.g. Ethernet for internet use or CD-ROM for multimedia).

There was a remarkable pace of quality improvement over this time period. Core computer characteristics have improved dramatically exhibiting average quarterly growth of 12% for "benchmark" and RAM. New components such as the Ethernet cards that were installed in only 19% of new PCs at the start of the period were standard in 52% of PCs by 2001. CD-ROM were installed in 80% of new PCs in 1996 but were ubiquitous in 2001. Furthermore, technological progress is accompanied by rapidly falling prices. The sales-weighted average price of PCs fell by 40% over our sample period (from \$2,550 to under \$1,500).[40]

Similar trends hold for the server market. Core characteristics, such as RAM, exhibits an average quarterly growth of 12% over the sample period, the proportion of servers using rack-optimization rose from practically zero at the start of the period to 40% by the end. The average price of servers fell by half during the same period (from \$13,523 to \$6,471). More importantly, for our purposes, is the dramatic rise of Windows on the server from 20% at the start of the sample to 57% by the end. As also seen in Figure 1, this increase in Windows' market share comes mainly from the decline of Novell's Netware (down from 38% at the start of the sample to 14% by the end) and, to a lesser extent of the various flavors of Unix (down from 24% to 18%). The only other operating system to have grown is open source Linux, although at the end of the period it had under 10% of the market.[41]

6. Results

6.1. Main Results. We turn now on the demand estimates from a simple logit model and the full model, before discussing their implications in terms of the theoretical model. The simple logit model (i.e. $\mu_{ij} = \mu_{ik} = 0$) is used to examine the importance of instrumenting

[40] There is an extensive empirical literature using hedonic regressions that documents the dramatic declines in the quality adjusted price of personal computers. See, for example, Berndt and Rappaport (2001) and Pakes (2003).

[41] Even Linux' limited success, despite being offered at a zero price, is mainly confined to server functions at the "edge" of the workgroup such as web-serving rather than the core workgroup taskd of file and print and directory services (see European Commission, 2004, for more discussion).

the price and to test the different sets of instrumental variables discussed in the previous section. Table 3 reports the results for PCs obtained from regressing $\ln(s_j) - \ln(s_0)$ on prices, characteristics and firm dummies. The first two columns include a full set of time dummies, whereas the last four columns include only a time trend (a restriction that is not statistically rejected). Column (1) reports OLS results: the coefficient on price is negative and significant as expected, but rather small in magnitude. Many coefficients have their expected signs - more recent generations of chips are highly valued as is an Ethernet card or CD-ROM drive. But the key performance metric, RAM, has a negative and significant coefficient, although the other quality measure, "benchmark", has a positive and significant coefficient. Furthermore, the vast majority of products (85.5%) are predicted to have inelastic demands, which is clearly unsatisfactory.

Column (2) uses the sum of the number of products and their observed characteristics offered by each firm and its rivals as instrumental variables. Treating price as endogenous greatly improves the model - the coefficient on price becomes much more negative and most other coefficients have now the expected signs.[42] Most importantly, under 1% of models now have inelastic demands.

Columns (3) and (4) report the same comparison between the OLS and IV results when we include a time trend instead of a full set of time dummies. Again, as we move from OLS to IV results, the coefficient on price becomes much more negative leaving no products with inelastic demands and all the other coefficients on PC characteristics have the expected sign. For example, both benchmark and RAM have now positive and significant coefficients and virtually all products have now elastic demands. In terms of diagnostics, the first stage results (reported in full in Table A1) indicate that the instruments are quite powerful: the F-statistic of the joint significance of the excluded instruments is 8.8 in column (2) and 27.2 in column (4). The Hansen-Sargan test of over-identification restrictions does reject, however, a common problem in this literature. In the last two columns we restrict the number of

[42] The only exception is monitor size which we would expect to have a positive coefficient whereas it has a small negative coefficient. This is likely to arise from the introduction of more advanced and thinner monitors of the same size introduced in 1999-2001. These are not recorded separately in the data.

instruments dropping hard disks in column (3) and also speed in column (4). Focusing on a sub-set of the more powerful instruments further improves our results. In the last column, for example, the first stage F-test is 40.62, moving the price coefficient further away from zero, leaving the no PC with inelastic demand.

Table 4 reports similar results from the simple logit model for the server data. In columns (1) and (2) the OLS and IV results are again reported based on regressions that include a full set of time dummies, whereas the latter four columns include instead a time trend. The price terms are significant, but with a much lower point estimate than PCs, indicating less customer sensitivity to price. Consistent with the PC results the coefficient on price falls dramatically moving from OLS to IV (e.g. from -0.040 in column (3) to -0.179 in column (4)). Looking at the preferred estimates of column (6) we find that RAM, the number of racks (an indicator of scalability) and type of chip appear to be significantly highly valued by customers. Most importantly, the estimated proportion of inelastic model demands has fallen dramatically from over 80% in column (3) to 22% in column (6). Notice also that the coefficient on the interaction of Windows and RAM is always positive and significant which is consistent with the idea of some interoperability constraints. As with PCs, the instruments (reported in full in Table A2) have power in the reduced form: the F-Statistic is 19 in column (2) and 13 in column (6).[43]

Results from the full (baseline) model are reported in Table 5. The first two panels report the mean coefficients for PCs and servers respectively. Almost all coefficients are significant and have the expected sign with benchmark the only notable exception. The last two panels report the results for the random coefficients. We allow random coefficients only on price and one other basic characteristic in our baseline specification (benchmark for PCs and RAM for servers). Our results indicate that there is significant heterogeneity in prices across the population of large businesses, but not in the other two characteristics (although the random

[43] The reason why estimated elasticities are somewhat lower for PCs than in the rest of the literature (Goeree, 2008; Foncel and Ivaldi, 2005), is because we are focusing on PC sales to large businesses whereas these papers focus on household purchases. Estimates using the same model for the home and small business sector exhibit higher elasticities (see Genakos, 2004). Similarly, the fact that server elasticities are lower compare to the rest of the literature (Ivaldi and Lörincz, 2008; Davis and Huse, 2009) is because we concentrate attention to the US (as opposed to the world market) and the segment of the market for workgroup servers (which is the more populus, less expensive segment).

coefficient for PC benchmark has a large value and is, in several robustness tests, significantly different from zero - see below). This indicates that for servers at least, characteristics are primarily vertically product differentiated at least for the larger firms who are the customer type we focus on here.

Figure 2 plots the calculated relative output and margin effects based on these coefficients (Table A3 and Figure A1 in the Appendix reports the numbers together with the 90% confidence interval). Unsurprisingly, server margins are higher than PC margins which reflects the finding that customers are less sensitive to server price than to PC prices. The positive value of the relative output effect indicates that reducing interoperability has a cost to Microsoft which is the loss of PC demand (due to complementarity). Three other key findings stand out. First, at the beginning of our sample period in 1996, the relative output effect is much higher than the relative margin which, according to our model indicates that Microsoft had no incentives to reduce interoperability. This is consistent with industry reports that interoperability was high during this period. Second, looking at the 1996-2001 period as a whole the two effects follow opposite directions with relative output steadily decreasing and the relative margin steadily increasing throughout our sample. By the end of our sample period in 2000 and 2001 the relative margin effect clearly dominates the relative output effect. Third, the key point when the two lines diverge is around the beginning of 2000, coinciding with the release of the new Microsoft operating system (Windows 2000). The European anti-trust case hinged precisely on industry reports that Windows 2000 contained severe interoperability limitations that were much more severe than any previous version on Windows. As we will show later these three facts are robust to alternative empirical models of complementarity and a battery of robustness tests.

If we decompose the underlying causes of the time series changes in the effects, then the rise in relative margins appears to be driven by the increase in the absolute value of the PC elasticity, reducing PC margins relative to servers. This is likely to be caused by the increasing "commodification" of PCs over this time period linked to the increasing entry of large numbers of PC brands by low cost manufacturers (e.g. Dell and Acer) as the industry matured and cheaper production sites in Asia became available.

6.2. Alternative empirical models of complementarity. We now move to the two alternative models. The first (strong complementarity) restricts the form of complementarity in the baseline model and the second (free complementarity) relaxes it.

Strong Complementarity. In column (1) of Table 6 presents the simplest version of strong complementarity where we assume a random coefficients on price and benchmark for PCs and only price for servers. The mean coefficients are estimated more precisely than in the baseline model and there seems to be significant heterogeneity in both price and benchmark for PCs but not in servers. Figure 3A plots the calculated relative margin and output effects. A similar qualitative picture emerges as before: relative output dominates at the beginning of the period, but steadily decreases over time, whereas relative margin follows the exact opposite direction. The beginning of 2000 seems to play an even bigger role in helping relative margin dominate the relative output.

Columns (2) and (3) add progressively more random coefficients. The estimated mean coefficients retain their magnitude and significance. Again, there appears to be significant heterogeneity for the PC price and characteristics coefficient and column (3) suggests some heterogeneity on the constant for servers. The plots of relative output and margin effects in Figures 3B and 3C reveal a qualitative very similar picture as before. Table A4 and Figure A2 in the Appendix reports the numbers together with the 95% confidence interval for the two effects. As we can see (apart from the very last period), the more precisely estimated coefficients and more restrictive model translates into tighter bounds for the two effects showing a significant incentive to reduce interoperability by the end of the period.

Free Complementarity. Our most general model is presented in the last column of Table 7 where we allow customers to purchase standalone servers (as well as standalone PCs, bundles of PCs and servers or the outside good) and complementarity to be freely estimated through the parameter $\Gamma_{PC,S}$. The estimated $\Gamma_{PC,S}$ parameter is positive and significant, confirming our previous assumption and intuition that the two product categories are complementary. The mean and random coefficients all exhibit similar patterns to the baseline results with evidence of significant heterogeneity in price (for servers and PCs) and significant heterogeneity in customers' valuation of PC quality (benchmark) but not server quality (RAM). Figure 3D plots the relative output and margin effects. Results at the beginning of the period are more mixed now, but again towards the end of our sample the relative margin dominates the relative output following a very clear trend, just as was the case for the baseline model. Given that this is a much more demanding specification, the consistency of results with our baseline case is reassuring.[44]

7. Robustness

Table 7 reports various robustness tests of the baseline model in Table 5 to gauge sensitivity of the results to changes in assumptions. In the first two columns we vary the number of random draws following the Monte Carlo evidence from Berry, Linton and Pakes (2004) for the BLP model. In column (1) we increase the number of draws to 250 (from 150 in the baseline model) and to 500 in column (2). The estimated results are very similar to our baseline specification, the only exception being that the PC benchmark now has a significant random coefficient. Not surprisingly the calculated relative output and margin effects in Figures 4A and 4B exhibit the same pattern as in Figure 2.

In column (3) and (4) we make different assumptions about the potential market size. In column (3) we assume that firms will only make a purchase decision to give all employees a computer every two years, essentially reducing the potential market size by half. In column (4) we assume that the potential market size is asymmetric, whereby firms purchase a PC every year whereas they purchase a server bundle every two years. In both experiments the estimated coefficients are hardly changed in Figure 4C and 4D are similar.

In columns (5) and (6) we reduce the number of instruments used for both the PCs and servers. On the one hand, using the most powerful instruments increases the absolute value of the coefficients. For example, the mean coefficient on PC price increases from -3.301 in the

[44]The reason why we do not use this model as our baseline is because estimation of the free complementarity was significantly slower to converge and more sensitive to starting values (resulting in convergence problems). Since identification of both the random coefficients and the $\Gamma_{PC,S}$ parameter come solely from time variation, these problems are hardly surprising given the limited time span of our data.

baseline model to -3.622 and -5.598 in columns (5) and (6) respectively. On the other hand, using fewer instruments means that we are reducing the number of identifying restrictions and this is reflected in higher standard errors. As a result very few coefficients are significant in column (6). Despite these differences, Figures 4E and 4F reveal a qualitative similar picture as before.

In the final two columns of Table 7 we experiment using different random coefficients. In column (7), we add a random coefficient on the constant in both equations. The estimated coefficients indicate no significant heterogeneity for the outside good at the 5% level for either PCs or servers. In column (8) we reduce the number of estimated random coefficients by allowing only a coefficient on server price. Both the estimated coefficients and calculated effects in Figures 3G and 3H look similar to our baseline specification: at the beginning of our sample the relative output effect dominates the relative margin effect, but by the end of 2000 the ordering is clearly reversed indicating strong incentives from Microsoft's perspective to reduce interoperability.

8. Conclusions

In this paper we examine the incentives for a monopolist to degrade interoperability in order to monopolize a complementary market. These type of concerns are very common in foreclosure cases such as the European Commission's landmark 2004 Decision against Microsoft. Structural econometric approach to examining the plausibility of such foreclosure claims have generally been unavailable. This paper seeks to provide such a framework developing both a new theory and an econometric method based upon this theory.

The incentive to reduce rival quality in a secondary market comes from the desire to more effectively extract rents from the primary market that are limited *inter alia* by the inability to perfectly price discriminate. We have detailed a general model of heterogeneous demand (encompassing BLP) and derived empirically tractable conditions under which a monopolist would have incentives to degrade interoperability. We implemented our method in the PC and server market estimating demand parameters allowing for complementarity. It seemed that Microsoft had incentives to decrease interoperability in the 2000s, but not in the mid

1990s. In our view, the combination of theory with strong micro-foundations and detailed demand estimation is the correct way to confront complex issues of market abuse.

There are limitations over what we have done and many areas for improvement. First, our model is entirely static, whereas it is likely that dynamic incentives are also important in leveraging (e.g. Carlton and Waldman, 2002). An important challenge is how to effectively confront such dynamic theoretical models with econometric evidence (e.g. Lee, 2009). Second, we have used only market-level data but detailed micro-information on the demand for different types of PCs and servers could lead to improvements in efficiency (see Bloom, Draca and Van Reenen, 2009, for examples of such detailed IT data). Although we have gone some of the way in the direction of endogenising one characteristic choice (interoperability decisions) there is still a long way to go.

References

[1] Abel, J., Berndt, E. and White, C. (2007) "Price Indices for Microsoft's Personal Computer Software Products" in Berndt, E. and Hulten, C. (eds) *Hard to Measure Goods and Services*, Chicago: Chicago University Press

[2] Adams, W. and Yellen, J. (1976) "Commodity bundling and the burden of monopoly", *Quarterly Journal of Economics*, 90, 475 – 498.

[3] Allen, B,. Deneckere, R., Faith, T. and Kovenock, D. (2000) "Capacity Pre-Commitment as a Barrier to Entry: A Bertrand-Edgeworth Approach", *Economic Theory* 15(3), 501-530.

[4] Bajari, P. and L. Benkard (2005) "Demand Estimation with Heterogeneous customers and Unobserved Product Characteristics: A Hedonic Approach", *Journal of Political Economy*, 113 (6), 1239-1276.

[5] Bloom, N., Draca, M. and Van Reenen, J. (2009) "Trade Induced Technical Change", LSE/Stanford mimeo.

[6] Bernheim, D. and Whinston, M. (1998) "Exclusive Dealing", *Journal of Political Economy*, 106(1), 64-103

[7] Berndt, E., and Rappaport (2001) "Price and quality of desktop and mobile personal computers: a quarter century overview" *American Economic Review*, 91(2), 268-71

[8] Berry, S. (1994) "Estimating Discrete-Choice Models of Product differentiation", *RAND Journal of Economics*, 25, 242-262.

[9] Berry, S., Levinson, J., and Pakes, A. (1995) "Automobile Prices in Market Equilibrium", *Econometrica*, 63, 841-90.

[10] Berry, S., Linton, O., and Pakes, A. (2004) "Limit Theorems for Estimating the Parameters of Differentiated Product Demand Systems", *Review of Economic Studies*, 71(3), 613-654.

[11] Bork, R. (1978) *The Antitrust Paradox,* New York: Free Press

[12] Bowman, W. (1957) "Tying arrangements and the leverage problem" *Yale Law Journal* 67, 19-36

[13] Bresnahan, T., S. Stern and M. Trajtenberg (1997). "Market Segmentation and the Sources of Rents from Innovation: Personal Computers in the late 1980s", *RAND Journal of Economics*, 28, S17-S44.

[14] Carlton, D. and Waldman (2002), "The Strategic Use of Tying to Preserve and Create Market Power in Evolving Industries", *RAND Journal of Economics*, 33 (2), 194-220.

[15] Davis, P. and Huse. C. (2009) "Estimating the co-ordinated effects of mergers", mimeo, Stockholm School of Economics.

[16] European Commission (2004) "Commission Decision of 23.04.04 relating to the Case COMP/C-3/37.792 Microsoft"

[17] Farrell, J. and Katz (2000) "Innovation, Rent extraction and integration in systems" *Journal of Industrial Economics*, 48(4), 413-432

[18] Foncel, J. and M. Ivaldi (2005) "Operating System Prices in the Home PC Market", *Journal of Industrial Economics,* 53(2) 265-297.

[19] Genakos, C. (2004) "Differential Merger Effects: The Case of the Personal Computer Industry", *mimeo*, London Business School.

[20] Gentzkow, M. (2007) “Valuing new goods in a model with complementarity: Online newspapers”, *American Economic Review*, 97(3), 713-744

[21] Goeree, A. S. (2008) “Limited Information and Advertising in the US Personal Computer Industry”, *Econometrica,* 76(5) 1017-1074.

[22] Hausman J. A., G. Leonard and J. D. Zona (1994) “Competitive Analysis with Differentiated Products”, *Annales D’ Economie at de statistique*, 34, 159-180.

[23] Hausman J. A. (1996). “Valuation of New Goods Under Perfect and Imperfect Competition” in T. F. Bresnahan and R. Gordon, eds., *The Economics of New Goods.* Chicago: National Bureau of Economic Research.

[24] Hendel, Igal (1999) “Estimating Multiple discrete choice models: an application to computerization returns” *Review of Economic Studies*, 66(2), 423-446.

[25] International Data Corporation (1998) *IDC Server Market Methodology, Database Structure and Definitions* IDC Report No. 15909R

[26] International Data Corporation (1999) *Server Operating Environment and Platform Report*

[27] International Data Corporation (1999) *Client Operating Environments: 1999 Worldwide markets and trends* IDC Report No. 19477

[28] Ivaldi, M. and S. Lörincz (2008) “Implementation Relevant Market Tests in Antitrust Policy: Applications to Computer Servers”, *Review of Law and Economics*, forthcoming.

[29] Lee, R. (2009) “Vertical Integration and Exclusivity in Platform and Two-Sided markets”, New York University mimeo

[30] McFadden, D. (1973) "Conditional Logit Analysis of Qualitative Choice Behavior", in P. Zarembka, eds., *Frontiers of Econometrics*, New York, Academic Press.

[31] McAfee, J., McMillan, J. and Whinston, M. (1989) “Multiproduct Monopoly, Commodity Bundling, and Correlation of Values”, *Quarterly Journal of Economics*, 104 (2) 371-383

[32] Nelder, J. A. and R. Mead (1965) “A Simplex Method for Function Minimization”, *Computer Journal*, 7, 308-313.

[33] Nevo, A. (2001) “Measuring Market Power in the Ready-to-Eat Cereal Industry”, *Econometrica*, 69, 307-342.

[34] Pakes, A. (2003) “A Reconsideration of Hedonic Price Indices with an Application to PC’s”, *American Economic Review*, 93 (5), 1578-1596.

[35] Petrin A. (2002) “Quantifying the Benefits of New Products: The Case of the Minivan”, *Journal of Political Economy*, 110, 705-729.

[36] Reddy, B., Evans, D., Nichols, A. and Schmalensee, R. (2001) “A Monopolist would still charge more for Windows: A Comment on Werden”, *Review of Industrial Organization*, 18, 263-268.

[37] Rey, P. and Tirole, J. (2007) “A Primer on Foreclosure”, *Handbook of Industrial Organization Volume III,* Armstrong, Mark and Porter, Robert (eds) 2145-2220

[38] Schmalensee, R. (1984) “Gaussian demand and commodity bundling”, *Journal of Business*, 57(1) 211-30.

[39] Song, I. and Chintagunta, P. (2006) “Measuring cross-category price effects with aggregate store data” *Management Science,* 52 (10) 1594-1609

[40] Schmalensee, R. (2000) "Anti-trust issues in Schumpeterian Industries" *American Economic Review*, 90, 192-196

[41] Van Reenen, J. (2004) "Is there a market for workgroup servers?", Centre for Economic Performance Discussion Paper No.650.

[42] Van Reenen, J. (2006) "The Growth of Network Computing: Quality adjusted price changes for network servers" *Economic Journal*, 116, F29-F44

[43] Werden, G. (2001) "Microsoft's pricing of Windows and the economics of derived demand monopoly" *Review of Industrial Organization*, 18, 357-262.

[44] Whinston, M. (1990) "Tying, Foreclosure, and Exclusion", *American Economic Review,* 80(4) 837-59

[45] Whinston, M. (2001) "Exclusivity and Tying in U.S. v. Microsoft: What We Know and Don't Know" *Journal of Economic Perspectives* (15), Spring, 63-80.

Appendices: Not for Publication, Referee use only

A. Proofs of propositions

Proof. Proof of Proposition 1

Suppose $y_M > ay_{M'}$. Then the best outcome a monopolist owning both technologies could get is to set $\omega + \omega_M$ such that it maximizes $(\omega + \omega_M) D(\omega + \omega_M - v_M)$. Let this solution be denoted by W^*. Consider any equilibrium price offered by firm M', $\omega_{M'}$. The monopolist can always set $\omega_M < y_M - ay_{M'} + \omega_{M'}$ and $\omega = W^* - \omega_M$ to make all the sales and achieve the monopoly price. In particular, there is no trembling hand perfect equilibrium with $\omega_{M'} < 0$, so that there is typically a strictly positive ω_M for which this is possible. Hence, the unique equilibrium outcome has the PC operating system monopolist sell the bundle at the monopoly price.

Now suppose $y_M < ay_{M'}$. It is easy to construct equilibria in which firm M' is excluded and the monopolist makes the same profits as if the competitor were not in the market. However, any such equilibrium must be Pareto dominated by another equilibrium that allows the monopolist to make the same profits and allows firm M' to make a strictly positive profit. We therefore concentrate on equilibria in which the better technology is offered. Fix $\omega_{M'} \in [0, ay_{M'} - y_M]$. We show that any such price can be charged in a sub-game perfect equilibrium of the game. Consider the following strategy for the monopolist: Set $\omega_M = \omega_{M'} - [ay_{M'} - y_M]$ and set ω to maximize:

$$\omega D(\omega + \omega_{M'} - ay_{M'})$$

All customers buy the server operating system from firm M'. This clearly is an equilibrium. The monopolist cannot improve on the outcome by inducing its own server product to be purchased (since $\omega_{M'} - ay_{M'} < y_M$). Given that M sets $\omega_M = \omega_{M'} - [ay_{M'} - y_M]$, the rival M' has no incentive to deviate from $\omega_{M'}$.

■

Proof. Proof of Proposition 2 Profits of the monopolist in equilibrium are given by:

$$\Pi(\omega^e, \omega^e_M(a)) = \omega^e D^S(\omega^e) + (\omega^e + \omega^e_M(a))D(\omega^e + \omega^e_M(a) - y_M)$$

where $\omega^e_M(a) = y_M - ay_{M'}$. By the envelope theorem

$$\frac{d\Pi}{da} = -y_{M'}D(\omega^e + \omega^e_M(a) - y_M)\left[1 + \frac{(\omega^e + \omega_M(a))}{D(\omega^e + \omega_M(a) - y_M)}\frac{\partial D(\omega^e + \omega_M(a) - y_M)}{\partial(\omega + \omega_M)}\right]$$

which is negative since $\omega^e + \omega^e_M < \omega^* + \omega^*_M$ whenever the constraint of $\omega_{M'} = 0$ is binding, i.e. when $\omega^e_M = y_M - ay_{M'}$.

■

Derivation of aggregate elasticities

Own price elasticity for PC operating system

$$\varepsilon_\omega = -\int \frac{q(\theta_i)}{q} w\lambda s_{00}(\theta_i)dP(\theta) \tag{22}$$

Own price elasticity for monopolist's server operating system

$$\varepsilon^M_{\omega_M} = -\int \frac{q_M(\theta_i)}{q_M}\lambda[1 - s_M(\theta_i)]dP(\theta) \tag{23}$$

Cross price elasticity for PC operating system with respect to monopolist's server operating system price

$$\varepsilon_{\omega_M} = -\frac{q_M}{q}\int \frac{q_M(\theta_i)}{q_M} w\lambda s_{00}(\theta_i)dP(\theta) \tag{24}$$

Cross price elasticity for monopolist's server operating system with respect to PC operating system price

$$\varepsilon^M_\omega = -\int \frac{q(\theta_i)}{q} w\lambda s_{00}(\theta_i)dP(\theta) \tag{25}$$

Derivation of individual specific elasticities

$$\begin{aligned}
\varepsilon_\omega(\theta_i) &= \frac{1}{q(\theta_i)} wL(\theta_i)\frac{\partial \sum_{j=1}^J \sum_{k=0}^K s_{jk}(\theta_i)}{\partial \omega} \\
&= \frac{1}{q(\theta_i)} wL(\theta_i)\frac{\partial}{\partial \omega}\left[\frac{\sum_{j=1}^J \sum_{k=0}^K e^{\delta_j + \delta_k}}{1 + \sum_{j=1}^J \sum_{k=0}^K e^{\delta_j + \delta_k}}\right] \\
&= -w\lambda s_{oo}(\theta_i)
\end{aligned} \tag{26}$$

and

$$\varepsilon_{\omega_M}(\theta_i) = \frac{1}{q(\theta_i)} wL(\theta_i) \frac{\partial \sum_{j=1}^{J} \sum_{k=0}^{K} s_{jk}(\theta_i)}{\partial \omega_M} \tag{27}$$

$$= -w\lambda s_{00}(\theta_i) \frac{q_M(\theta_i)}{q(\theta_i)} \tag{28}$$

$$\begin{aligned} \varepsilon_{\omega}^{M}(\theta_i) &= \frac{1}{q_M(\theta_i)} M(\theta_i) \frac{\partial \sum_{j=1}^{J} \sum_{k\in M} s_{jk}(\theta_i)}{\partial \omega_M} \\ &= -w\lambda s_{00}(\theta_i) \end{aligned} \tag{29}$$

$$\begin{aligned} \varepsilon_{\omega_M}^{M}(\theta_i) &= \frac{1}{q_M(\theta_i)} L(\theta_i) \frac{\partial \sum_{j=1}^{J} \sum_{k\in M} s_{jk}(\theta_i)}{\partial \omega_M} \\ &= -\lambda \sum_{k\notin M} s_k(\theta_i) \end{aligned} \tag{30}$$

To generate the aggregate elasticities we simply need to add up the frequency weighted individual elasticities:

$$\begin{aligned} \varepsilon_{\omega} &= \int \frac{q(\theta_i)}{q} \varepsilon_{\omega}(\theta_i) dP(\theta) \\ &= -\int \frac{q(\theta_i)}{q} w\lambda s_{00}(\theta_i) dP(\theta) \end{aligned} \tag{31}$$

$$\begin{aligned} \varepsilon_{\omega_M} &= \int \frac{q(\theta_i)}{q} \varepsilon_{\omega_M}(\theta_i) dP(\theta) \\ &= -\frac{q_M}{q} \int \frac{q_M(\theta_i)}{q_M} w\lambda s_{00}(\theta_i) dP(\theta) \end{aligned} \tag{32}$$

$$\begin{aligned} \varepsilon_{\omega}^{M} &= \int \frac{q(\lambda,\beta,\gamma,w)}{q} \varepsilon_{\omega}^{M}(\theta_i) dP(\theta) \\ &= -\int \frac{q(\lambda,\beta,\gamma,w)}{q} w\lambda s_{00}(\theta_i) dP(\theta) \end{aligned} \tag{33}$$

$$\begin{aligned} \varepsilon_{\omega_M}^{M} &= \int \frac{q_M(\theta_i)}{q_M} \varepsilon_{\omega_M}^{M}(\theta_i) dP(\theta) \\ &= -\int \frac{q_M(\theta_i)}{q_M} \lambda[1 - s_M(\theta_i)] dP(\theta) \end{aligned} \tag{34}$$

We can then determine the sign of ω_M and ω_{OS} by noting that

$$\begin{aligned}\frac{q}{q_M}\varepsilon_{\omega_M} - \varepsilon_{\omega_{OS}} &= \int \left[\frac{q(\theta_i)}{q} - \frac{q_M(\theta_i)}{q_M}\right][w\alpha s_{oo}(\theta_i)]\, dP(\theta) \\ &= -\int \left[\frac{q(\theta_i)}{q} - \frac{q_M(\theta_i)}{q_M}\right][\varepsilon_\omega(\theta_i)]\, dP(\theta) \\ &= \int \left[\frac{q(\theta_i)}{q} - \frac{q_M(\theta_i)}{q_M}\right][\bar{\varepsilon}_\omega - \varepsilon_\omega(\theta_i)]\, dP(\theta) \qquad (35)\end{aligned}$$

where the last equality comes from subtracting $-\int \left[\frac{q(\theta_i)}{q} - \frac{q_M(\theta_i)}{q_M}\right]\bar{\varepsilon}_\omega dP(\theta) = 0$ from the second line where

$$\bar{\varepsilon}_\omega = \int \varepsilon_\omega(\theta_i) dP(\theta)$$

For ω, the price of PC operating systems we obtain that it is proportional to:

$$\begin{aligned}\frac{q_M}{q}\varepsilon^M_\omega - \varepsilon^M_{\omega_M} &= \int \alpha w s_{00}(\theta_i)\left(\frac{M(\theta_i) - q_M(\theta_i)}{wM(\theta_i) - q(\theta_i)}\frac{q_M(\theta_i)}{q_M} + \frac{q_M(\theta_i)}{q(\theta_i)}\frac{q(\theta_i)}{q}\right) dP(\theta) \\ &\quad -\frac{q_M}{q}\int \alpha w s_{00}(\theta_i)\left[\frac{q(\theta_i)}{q} - \frac{q_M(\theta_i)}{q_M}\right] dP(\theta) \qquad (36)\end{aligned}$$

B. Data Appendix

As noted in the Data section, quarterly data on quantities and prices[45] between 1995Q1 and 2001Q1 was taken from the PC and Server quarterly trackers conducted by International Data Corporation's (IDC). The PC tracker provided disaggregation by manufacturer, model name, form factor,[46] chip type (e.g. 5th Generation) and processor speed bandwidth (e.g. 200-300 MHz). Similarly the server tracker provides disaggregation by manufacturer, model name, chip type (Risc, Cisc, Intel) and operating system. Basic characteristics are also available on CPU numbers, CPU capacity, whether the server was rack optimized and the number of racks. In order to obtain more detailed product characteristics we matched each observation in the IDC dataset with information from trade sources such as the Datasources catalogue and various computer magazines.[47] In order to be consistent with the IDC definition of price, we assign the characteristics of the median model per IDC observation if more than two models were available. The justification for this choice is that we preferred to keep the transaction prices of IDC, rather than substitute them with the list prices published in the magazines. An alternative approach followed by Pakes (2003) would be to list all the available products by IDC observation with their prices taken from the magazines and their

[45] Prices are defined by IDC as "the average end-user (street) price paid for a typical system configured with chassis, motherboard, memory, storage, video display and any other components that are part of an "average" configuration for the specific model, vendor, channel or segment". Prices were deflated using the Consumer Price Index from the Bureau of Labor Statistics.

[46] Form factor means whether the PC is a desktop, notebook or ultra portable. The last two categories were merged into one.

[47] The magazines included PC Magazine, PC Week, PC World, Computer Retail Week, Byte.com, Computer User, NetworkWorld, Computer World, Computer Reseller News, InfoWorld, Edge: Work-Group Computing Report and Computer Shopper.

sales computed by splitting the IDC quantity equally among the observations. Although, clearly, both approaches adopt some ad hoc assumptions, qualitatively the results would probably be the same. Both list and transaction prices experienced a dramatic fall over this period and the increase in the number and variety of PCs offered would have been even more amplified with the latter approach. All nominal prices are deflated using the CPI.

For PCs, instead of using the seventeen processor type dummies and the speed of each chip as separate characteristic, we merge them using CPU "benchmarks" for each computer. CPU benchmarks were obtained from *The CPU Scorecard* (www.cpuscorecard.com). They are essentially numbers assigned to each processor-speed combination based on technical and performance characteristics. Our final unit of observation is defined as a manufacturer (e.g. Dell), model (e.g. Optiplex), form factor (e.g. desktop), processor type (e.g. Pentium II) and processor speed (e.g. 266 MHZ) combination with additional information on other characteristics such as the RAM, hard disk, modem/ethernet, CD-ROM and monitor size.

Similarly, for servers a unit of observation is defined as a manufacturer and family/model-type. We also distinguish by operating system, since (unlike PCs) many servers run non-Windows operating systems (we distinguish six other categories: Netware, Unix, Linux, VMS, OS390/400 and a residual category). For servers key characteristics are also RAM, the number of rack slots[48] whether the server was rack optimized (racks were an innovation that enhanced server flexibility), motherboard type (e.g. Symmetric Parallel Processing - SMP), and chip type (CISC, RISC or IA32). For more discussion of the datasets and characteristics see International Data Corporation (1998, 1999a,b) and Van Reenen (2004, 2006).

The PC data allows us to distinguish by end user. Since servers are not purchased by consumers and small firms, we condition on PCs purchased by firms with over 500 employees. Results were robust to changing this size threshold (see Genakos, 2004, for separate estimation by customer type).

Given the aggregate nature of our data, we assume that the total market size is given by the total number of employees in large businesses is taken from the Bureau of Labour Statistics. Results based on different assumptions about the potential market size are also reported in the robustness section.

C. Calculating the relative output effect, relative margin effect and standard errors

There is an incentive to decrease interoperability at the margin if:

$$\frac{\omega_M - c_M}{\omega - c} > -\frac{\left.\frac{dq(\mathbf{p}_j, \mathbf{p}_k, a)}{da}\right|_{\omega, \omega_M}}{\left.\frac{dq_M(\mathbf{p}_j, \mathbf{p}_k, a)}{da}\right|_{\omega, \omega_M}}$$

where the left hand side is the relative margin, whereas the right hand side is the relative output effect.

[48]Rack mounted servers were designed to fit into 19 inch racks. They allow multiple machines to be clustered or managed in a single location and enhance scalability.

In our baseline specification, individual PC and server market shares are given by:

$$\begin{aligned} s_{ij} &= e^{\delta_j+\eta_{ij}} \sum_{k=0}^{K} \frac{e^{\delta_k+\eta_{ik}}}{1+\sum_{j=1}^{J}\sum_{k=0}^{K} e^{\delta_j+\eta_{ij}+\delta_k+\eta_{ik}}} = \frac{e^{V_{ij}}}{\frac{1}{1+W_{ik}}+W_{ij}} = \frac{e^{V_{ij}}\left(1+W_{ik}\right)}{1+W_{ij}+W_{ij}W_{ik}} \\ s_{ik} &= \frac{e^{V_{ik}}}{\frac{1}{W_{ij}}+1+W_{ik}} = \frac{e^{V_{ik}}W_{ij}}{1+W_{ij}+W_{ij}W_{ik}} \end{aligned}$$

where $V_{ij}=\delta_j+\mu_{ij}$,$V_{ik}=\delta_k+\mu_{ik}$, $W_{ij}=\sum_{j=1}^{J}e^{V_{ij}}$, $W_{ik}=\sum_{k=1}^{K}e^{V_{ik}}$. To get the aggregate PC and server market shares $s_j=\frac{1}{ns}\sum_{i=1}^{ns}s_{ij}$ and $s_k=\frac{1}{ns}\sum_{i=1}^{ns}s_{ik}$, where ns is the number of drawn individuals.

Hence, to calculate the relative output:

$$\frac{dq\,/d\alpha}{dq_M\,/d\alpha} = \frac{\frac{ds_j}{d\alpha}L_{PC}}{\frac{ds_k}{d\alpha}L_S} = \frac{\frac{1}{ns}\sum_{i=1}^{ns} s_{ij}\frac{\sum_{k=1}^{K}\widetilde{\gamma_k}e^{V_{ik}}}{(1+W_{ik})[1+W_{ij}(1+W_{ik})]}}{\frac{1}{ns}\sum_{i=1}^{ns}\left(s_{ik}\widetilde{\gamma_k}-s_{ik}\frac{\sum_{k=1}^{K}\widetilde{\gamma_k}e^{V_{ik}}}{\frac{1}{W_{ij}}+1+W_{ik}}\right)}\frac{L_{PC}}{L_S}$$

where L_{PC}, L_S are the market sizes for PCs and servers respectively and $\widetilde{\gamma_k}=\gamma\left(MEM_k*(1-MSFT_k)\right)$.

To calculate the relative margin

$$\frac{\omega_M-c_M}{\omega-c} = \frac{\frac{q}{q_M}\varepsilon_{\omega_M}-\varepsilon_{\omega}}{\frac{q_M}{q}\varepsilon_{\omega}^{M}-\varepsilon_{\omega_M}^{M}} = \frac{\frac{q}{q_M}\left(\frac{\partial s_j}{\partial p_k}\frac{1}{s_j}\right)-\left(\frac{\partial s_j}{\partial p_j}\frac{1}{s_j}\right)}{\frac{q_M}{q}\left(\frac{\partial s_k}{\partial p_j}\frac{1}{s_k}\right)-\left(\frac{\partial s_k}{\partial p_k}\frac{1}{s_k}\right)}$$

where the derivatives for the PCs are:

$$\begin{aligned} \text{own price semi-elasticity} \quad &: \quad \frac{\partial s_j}{\partial p_j} = \frac{1}{ns}\sum_{i=1}^{ns}\frac{\partial s_{ij}}{\partial p_j} \\ &= \frac{1}{ns}\sum_{i=1}^{ns} s_{ij}(1-s_{ij})(\lambda^{PC}+\sigma_p^{PC}\nu_{ip}^{PC}) \\ \text{cross PC price semi-elasticity} \quad &: \quad \frac{\partial s_j}{\partial p_d} = \frac{1}{ns}\sum_{i=1}^{ns}\frac{\partial s_{ij}}{\partial p_d} \\ &\quad = -\frac{1}{ns}\sum_{i=1}^{ns} s_{ij}s_{id}(\lambda^{PC}+\sigma_p^{PC}\nu_{ip}^{PC}) \\ \text{cross PC-server semi-elasticity} \quad &: \quad \frac{\partial s_j}{\partial p_k} = \frac{1}{ns}\sum_{i=1}^{ns}\frac{\partial s_{ij}}{\partial p_k} \\ &= \frac{1}{ns}\sum_{i=1}^{ns} s_{ij}(\lambda^{PC}+\sigma_p^{PC}\nu_{ip}^{PC})\frac{e^{V_{ik}}}{\left(1+W_{ik}\right)\left(1+W_{ij}+W_{ij}W_{ik}\right)} \end{aligned}$$

Similarly, the derivatives for the servers are:

$$\begin{aligned}
\text{own price semi-elasticity} \quad &: \quad \frac{\partial s_k}{\partial p_k} = \frac{1}{ns}\sum\nolimits_{i=1}^{ns}\frac{\partial s_{ik}}{\partial p_k} = \frac{1}{ns}\sum\nolimits_{i=1}^{ns} s_{ik}(1-s_{ik})(\lambda^S + \sigma_p^S \nu_{ip}^S) \\
\text{cross PC price semi-elasticity} \quad &: \quad \frac{\partial s_k}{\partial p_m} = \frac{1}{ns}\sum\nolimits_{i=1}^{ns}\frac{\partial s_{ik}}{\partial p_m} = -\frac{1}{ns}\sum\nolimits_{i=1}^{ns} s_{ik}s_{im}(\lambda^S + \sigma_p^S \nu_{ip}^S) \\
\text{cross PC-server semi-elasticity} \quad &: \quad \frac{\partial s_k}{\partial p_j} = \frac{1}{ns}\sum\nolimits_{i=1}^{ns}\frac{\partial s_{ik}}{\partial p_j} = \frac{1}{ns}\sum\nolimits_{i=1}^{ns} s_{ik}(\lambda^S + \sigma_p^S \nu_{ip}^S)\frac{e^{V_{ij}}}{W_{ij}\left(1 + W_{ij} + W_{ij}W_{ik}\right.}
\end{aligned}$$

To compute the gradient of the objective function, we need the derivatives of the mean value $\delta = (\delta_j, \delta_k)$ with respect to the non-linear parameters $\widetilde{\theta} \equiv (\sigma_h^{PC}, \sigma_h^S, \sigma_p^{PC}, \sigma_p^S)$:

$$D\delta = \begin{pmatrix} \frac{\partial \delta_1}{\partial \widetilde{\theta}_1} \cdots & \frac{\partial \delta_1}{\partial \widetilde{\theta}_H} \\ \frac{\partial \delta_J}{\partial \widetilde{\theta}_1} \cdots & \frac{\partial \delta_J}{\partial \widetilde{\theta}_H} \\ \frac{\partial \delta_1}{\partial \widetilde{\theta}_1} \cdots & \frac{\partial \delta_1}{\partial \widetilde{\theta}_H} \\ \frac{\partial \delta_K}{\partial \widetilde{\theta}_1} \cdots & \frac{\partial \delta_K}{\partial \widetilde{\theta}_H} \end{pmatrix} = - \begin{pmatrix} \frac{\partial s_1}{\partial \delta_1} \cdots & \frac{\partial s_1}{\partial \delta_J} & \frac{\partial s_1}{\partial \delta_1} \cdots & \frac{\partial s_1}{\partial \delta_K} \\ \frac{\partial s_J}{\partial \delta_1} \cdots & \frac{\partial s_J}{\partial \delta_J} & \frac{\partial s_J}{\partial \delta_1} \cdots & \frac{\partial s_J}{\partial \delta_K} \\ \frac{\partial s_1}{\partial \delta_1} \cdots & \frac{\partial s_1}{\partial \delta_J} & \frac{\partial s_1}{\partial \delta_1} \cdots & \frac{\partial s_1}{\partial \delta_K} \\ \frac{\partial s_K}{\partial \delta_1} \cdots & \frac{\partial s_K}{\partial \delta_J} & \frac{\partial s_K}{\partial \delta_1} \cdots & \frac{\partial s_K}{\partial \delta_K} \end{pmatrix}^{-1} \begin{pmatrix} \frac{\partial s_1}{\partial \widetilde{\theta}_1} \cdots & \frac{\partial s_1}{\partial \widetilde{\theta}_H} \\ \frac{\partial s_J}{\partial \widetilde{\theta}_1} \cdots & \frac{\partial s_J}{\partial \widetilde{\theta}_H} \\ \frac{\partial s_1}{\partial \widetilde{\theta}_1} \cdots & \frac{\partial s_1}{\partial \widetilde{\theta}_H} \\ \frac{\partial s_K}{\partial \widetilde{\theta}_1} \cdots & \frac{\partial s_K}{\partial \widetilde{\theta}_H} \end{pmatrix}$$

where $\widetilde{\theta}_i$, $i = 1, ..., H$ denotes the i's element of the vector $\widetilde{\theta}$, which contains the non-linear parameters of the model. Given the smooth simulator used for the market shares, the above derivatives are given by:

$$
\begin{aligned}
\frac{\partial s_j}{\partial \delta_j} &= \frac{1}{ns}\sum_{i=1}^{ns}\frac{\partial s_{ij}}{\partial \delta_j} = \frac{1}{ns}\sum_{i=1}^{ns} s_{ij}(1-s_{ij}) \\
\frac{\partial s_j}{\partial \delta_d} &= \frac{1}{ns}\sum_{i=1}^{ns}\frac{\partial s_{ij}}{\partial \delta_d} = -\frac{1}{ns}\sum_{i=1}^{ns} s_{ij}s_{id} \\
\frac{\partial s_j}{\partial \delta_k} &= \frac{1}{ns}\sum_{i=1}^{ns}\frac{\partial s_{ij}}{\partial \delta_k} = \frac{1}{ns}\sum_{i=1}^{ns} s_{ij}\frac{e^{V_{ik}}}{(1+W_{ik})\,(1+W_{ij}+W_{ij}W_{ik})} \\
\frac{\partial s_k}{\partial \delta_k} &= \frac{1}{ns}\sum_{i=1}^{ns}\frac{\partial s_{ik}}{\partial \delta_k} = \frac{1}{ns}\sum_{i=1}^{ns} s_{ik}(1-s_{ik}) \\
\frac{\partial s_k}{\partial \delta_m} &= \frac{1}{ns}\sum_{i=1}^{ns}\frac{\partial s_{ik}}{\partial \delta_m} = -\frac{1}{ns}\sum_{i=1}^{ns} s_{ik}s_{im} \\
\frac{\partial s_k}{\partial \delta_j} &= \frac{1}{ns}\sum_{i=1}^{ns}\frac{\partial s_{ik}}{\partial \delta_j} = -\frac{1}{ns}\sum_{i=1}^{ns} s_{ik}\frac{e^{V_{ij}}}{W_{ij}\,(1+W_{ij}+W_{ij}W_{ik})} \\
\frac{\partial s_j}{\partial \sigma_h^{PC}} &= \frac{1}{ns}\sum_{i=1}^{ns}\frac{\partial s_{ij}}{\partial \sigma_h^{PC}} = \frac{1}{ns}\sum_{i=1}^{ns} s_{ij}\left(x_j v_{ih}^{PC} - \sum_{d=1}^{J} s_{id}x_d v_{ih}^{PC}\right) \\
\frac{\partial s_j}{\partial \sigma_p^{PC}} &= \frac{1}{ns}\sum_{i=1}^{ns}\frac{\partial s_{ij}}{\partial \sigma_p^{PC}} = \frac{1}{ns}\sum_{i=1}^{ns} s_{ij}\left(p_j v_{ip}^{PC} - \sum_{d=1}^{J} s_{id}p_d v_{ip}^{PC}\right) \\
\frac{\partial s_j}{\partial \sigma_h^{S}} &= \frac{1}{ns}\sum_{i=1}^{ns}\frac{\partial s_{ij}}{\partial \sigma_h^{S}} = \frac{1}{ns}\sum_{i=1}^{ns} s_{ij}\frac{\sum_{m=1}^{K} e^{V_{ik}}y_m v_{ih}^{S}}{(1+W_{ik})\,(1+W_{ij}+W_{ij}W_{ik})} \\
\frac{\partial s_j}{\partial \sigma_p^{S}} &= \frac{1}{ns}\sum_{i=1}^{ns}\frac{\partial s_{ij}}{\partial \sigma_p^{S}} = \frac{1}{ns}\sum_{i=1}^{ns} s_{ij}\frac{\sum_{m=1}^{K} e^{V_{ik}}p_m v_{ip}^{S}}{(1+W_{ik})\,(1+W_{ij}+W_{ij}W_{ik})} \\
\frac{\partial s_k}{\partial \sigma_h^{S}} &= \frac{1}{ns}\sum_{i=1}^{ns}\frac{\partial s_{ik}}{\partial \sigma_h^{S}} = \frac{1}{ns}\sum_{i=1}^{ns} s_{ik}\left(y_k v_{ih}^{S} - \sum_{m=1}^{K} s_{im}y_d v_{ih}^{S}\right) \\
\frac{\partial s_k}{\partial \sigma_p^{S}} &= \frac{1}{ns}\sum_{i=1}^{ns}\frac{\partial s_{ik}}{\partial \sigma_p^{S}} = \frac{1}{ns}\sum_{i=1}^{ns} s_{ik}\left(p_k v_{ip}^{S} - \sum_{m=1}^{K} s_{im}p_d v_{ip}^{S}\right) \\
\frac{\partial s_k}{\partial \sigma_h^{PC}} &= \frac{1}{ns}\sum_{i=1}^{ns}\frac{\partial s_{ik}}{\partial \sigma_h^{PC}} = \frac{1}{ns}\sum_{i=1}^{ns} s_{ik}\frac{\sum_{d=1}^{J} e^{V_{ij}}x_d v_{ih}^{PC}}{W_{ij}\,(1+W_{ij}+W_{ij}W_{ik})} \\
\frac{\partial s_k}{\partial \sigma_p^{PC}} &= \frac{1}{ns}\sum_{i=1}^{ns}\frac{\partial s_{ik}}{\partial \sigma_p^{PC}} = \frac{1}{ns}\sum_{i=1}^{ns} s_{ik}\frac{\sum_{d=1}^{J} e^{V_{ij}}p_d v_{ip}^{PC}}{W_{ij}\,(1+W_{ij}+W_{ij}W_{ik})}
\end{aligned}
$$

We also calculated the standard errors based on this Jacobian.

D. Estimation details of alternative models

Since utilities are defined over bundles of models across categories, the model cannot be directly taken to aggregate data. We need to derive marginal probabilities of purchase in each category and the conditional (on purchase) models choice probabilities. To derive these probabilities, we need to assume that the error term, ϵ_{ijkt}, is logit i.i.d. distributed across bundles, consumers and time. Given this assumption on the error term, for each consumer define $W_{ij} \equiv \sum_{j=1}^{J}\exp(\delta_j+\mu_{ij})$, the inclusive value for PCs, and $W_{iS} \equiv \sum_{k=1}^{K}\exp(\delta_k+\mu_{ik})$, the inclusive value for servers. Then, using the result derived in Song and Chintagunta (2006), the marginal probability for purchasing a PC is given by[49]:

[49] This follows because

$$\Pr(d_{PC} = 1|\ x_j, y_k, i) = \frac{W_{ij} W_{ik} e^{\Gamma_{PC,S}} + W_{ij}}{W_{ij} W_{ik} e^{\Gamma_{PC,S}} + W_{ij} + W_{ik} + 1} \tag{37}$$

Then the conditional brand choice probability for PC j is given by:

$$\Pr(j|\ d_{PC} = 1, x_j, y_k, i) = \frac{\exp(V_{ij})}{W_{ij}}, \tag{38}$$

The unconditional brand choice probability is obtained by multiplication:

$$\Pr(j = 1|\ x_j, y_k, i) = \Pr(d_{PC} = 1|\ x_j, y_k, i) * \Pr(j|\ d_{PC} = 1, x_j, y_k, i). \tag{39}$$

Market shares for each product, s_j (and s_k), are obtained by aggregating over customers and their vectors of unobservable tastes.

The estimation of this model follows a similar logic to the one estimated in the main text. The only major difference now is that we have an additional non-linear parameter apart from the random coefficients. Define $\overline{\theta_2} \equiv (\sigma_h^{PC}, \sigma_h^S, \sigma_p^{PC}, \sigma_p^S)$ then $\widetilde{\theta} \equiv (\overline{\theta_2}, \Gamma_{PC,S})$ is now the vector of non-linear parameters, i.e., i.e., the random coefficients on characteristics and price for PCs and servers and the complementarity parameter. Let r be the set of variables that we are allowing non-linear parameters (e.g. x_j, y_k, p_j, p_k). Let $\delta = (\delta_j, \delta_k)$, $\xi = (\xi_j, \xi_k)$, $\nu_i = \left(\nu_i^{PC}, \nu_i^S\right)$ and $\mu_i = (\mu_{ij}, \mu_{ik})$.

Our iterative procedure is as follows:

Step 0: Draw the idiosyncratic taste terms ν_i (these draws remain constant throughout the estimation procedure) and starting values for $\widetilde{\theta}$.

Step 1. Given $(r, \overline{\theta_2})$, calculate μ_i.

Step 2. Given (δ, μ_i), calculate the conditional probabilities of equation (38) for PCs and servers.

Step 3. Given $(\delta, \mu_i, \Gamma_{PC,S})$ calculate the marginal probabilities of equation (37) for PCs and servers.

Step 4. Calculate the unconditional brand probabilities of equation (39) and aggregate to get the market shares for each brand.

Step 5. Given $\widetilde{\theta}$, we need to numerically compute the mean valuations, δ, that equate the observed to the predicted brand market shares. Due to complementarity between the PCs and servers, we compute each product category's mean valuation conditional on the other category's mean valuation. Specifically, it consists of the following sequentially iterative substeps:

Substep 5.0 Make an initial guess on δ and set $\delta_{old} = \delta$.

Substep 5.1 Compute δ_j given δ_k using BLP's contraction mapping. Update δ.

Substep 5.2 Compute δ_k given δ_j and update δ.

Substep 5.3 Check if $\delta_{old} =$ updated δ. If yes, go to step 4. Otherwise, set $\delta_{old} = \delta$ and go to substep 5.1.

Step 6. Given δ, calculate ξ and form the GMM.

Step 7. Minimize a quadratic form of the residuals and update.

$$\Pr(d_{PC} = 1|\ x, i) = \frac{W_{PC}(e^{\Gamma(d_{PC}, d_S)} W_S + e^{\Gamma(d_{PC}, 0)})}{W_{PC}(e^{\Gamma(d_{PC}, d_S)} W_S + e^{\Gamma(d_{PC}, 0)}) + (e^{\Gamma(0, d_S)} W_S + e^{\Gamma(0,0)})}$$

and we normalize $g_{PC} = g_S = 0$.

We also estimated two other variants of this algorithm. The first one reiterates one additional time substeps 5.1 and 5.2 to make sure that there is no feedback from PCs to server mean valuations. This variant takes slightly more computational time. The second variant instead of updating the mean valuations for each product category in substeps 5.1 and 5.2, always uses the initial estimates (taken from the simple logit IV regression). This variant takes more computational time, but it is more robust to starting values. To minimize the GMM function we used both the Nelder-Mead nonderivative search method and the faster Quasi-Newton gradient method based on an analytic gradient. We combine all these methods to verify that we reached a global instead of a local minimum. Standard errors are based on the same analytic Jacobian and are corrected for heteroskedasticity taking also into consideration the additional variance introduced by the simulation.

TABLE 1 - DESCRIPTIVE STATISTICS FOR PC DATA

Period	No. of models	Quantity	Price ($1000s)	Benchmark (1000s)	RAM (100MB)	CD-ROM	Ethernet	Monitor size (inches)	Desktop
1996Q1	104	6438.699	2.550	0.221	0.138	0.799	0.187	10.388	0.703
1996Q2	103	7823.198	2.437	0.240	0.151	0.863	0.254	11.089	0.706
1996Q3	99	8946.276	2.441	0.266	0.157	0.905	0.279	11.426	0.674
1996Q4	114	8034.009	2.437	0.294	0.178	0.889	0.236	11.845	0.628
1997Q1	129	7116.477	2.409	0.363	0.213	0.896	0.091	11.596	0.637
1997Q2	156	6806.709	2.255	0.424	0.248	0.919	0.127	11.209	0.692
1997Q3	181	6978.622	2.210	0.489	0.287	0.963	0.177	11.035	0.698
1997Q4	193	6485.918	2.123	0.531	0.321	0.931	0.217	10.626	0.709
1998Q1	204	5660.170	2.101	0.609	0.388	0.892	0.378	10.898	0.723
1998Q2	219	5452.665	2.019	0.695	0.430	0.936	0.335	11.705	0.708
1998Q3	215	6428.275	1.885	0.775	0.483	0.947	0.417	12.382	0.734
1998Q4	143	10258.830	1.896	0.914	0.595	0.884	0.453	13.447	0.749
1999Q1	131	10656.770	1.810	1.069	0.670	0.914	0.436	15.128	0.755
1999Q2	124	14062.890	1.705	1.124	0.701	0.926	0.454	16.137	0.763
1999Q3	113	15190.380	1.663	1.279	0.796	0.955	0.446	16.213	0.741
1999Q4	122	13123.920	1.619	1.487	0.938	0.973	0.401	15.757	0.727
2000Q1	152	9227.644	1.592	1.792	1.073	0.963	0.384	13.461	0.731
2000Q2	179	9047.285	1.585	2.001	1.091	0.972	0.418	13.481	0.719
2000Q3	194	9266.313	1.554	2.085	1.109	0.977	0.440	13.385	0.703
2000Q4	233	7365.650	1.555	2.206	1.110	0.986	0.513	13.453	0.707
2001Q1	197	8413.300	1.493	2.417	1.120	0.993	0.517	13.143	0.721
ALL	3305	8357.177	1.884	1.165	0.662	0.937	0.367	13.107	0.716

Source: International Data Corporation (IDC) Quarterly PC Tracker matched to more detailed PC characteristics from several industry datasources and trade magazines.

Notes: All the entries (except model numbers and quantity) are weighted by PC model sales. "Benchmark" is a score assigned to each processor-speed combination based on technical and performance characteristics (see CPU Scorecard: www. cpuscorecard.com).

TABLE 2 - DESCRIPTIVE STATISTICS FOR SERVER DATA

Period	No. of models	Quantity	Price ($1000s)	RAM (100MB)	Rack Optimize	Symmetrical Processor	Number of Racks	Windows	Netware	Unix	Linux
1996Q1	123	727.252	13.523	0.618	0.036	0.558	0.036	0.199	0.382	0.245	0.000
1996Q2	125	772.664	12.323	0.766	0.037	0.551	0.037	0.199	0.394	0.231	0.000
1996Q3	116	843.828	13.637	1.336	0.010	0.618	0.071	0.211	0.398	0.226	0.000
1996Q4	129	923.101	13.793	1.444	0.094	0.580	0.883	0.209	0.390	0.232	0.000
1997Q1	128	908.258	11.945	1.602	0.079	0.595	1.221	0.226	0.406	0.233	0.000
1997Q2	129	1112.605	11.671	1.671	0.103	0.684	1.808	0.229	0.398	0.227	0.000
1997Q3	134	1331.254	9.874	1.469	0.164	0.716	2.350	0.272	0.400	0.194	0.000
1997Q4	145	1322.752	10.830	1.793	0.119	0.753	2.582	0.280	0.381	0.224	0.000
1998Q1	153	1071.209	9.485	2.023	0.088	0.794	2.708	0.324	0.374	0.209	0.004
1998Q2	143	1154.790	9.113	2.222	0.057	0.779	3.115	0.336	0.365	0.226	0.005
1998Q3	145	1331.276	8.253	2.226	0.057	0.777	3.788	0.353	0.381	0.192	0.008
1998Q4	167	1523.964	7.434	2.666	0.108	0.818	3.855	0.427	0.327	0.171	0.012
1999Q1	151	1412.715	8.053	3.122	0.068	0.786	3.974	0.439	0.313	0.182	0.023
1999Q2	125	2105.560	7.942	3.267	0.079	0.871	4.135	0.440	0.306	0.182	0.028
1999Q3	131	2016.008	7.879	3.523	0.077	0.893	4.235	0.447	0.304	0.173	0.031
1999Q4	146	1840.541	7.166	3.938	0.122	0.878	4.013	0.445	0.257	0.188	0.060
2000Q1	150	1748.087	7.249	4.223	0.203	0.891	3.754	0.488	0.215	0.180	0.084
2000Q2	171	1881.368	7.115	4.478	0.329	0.886	3.527	0.539	0.169	0.178	0.086
2000Q3	162	2147.352	6.952	4.586	0.399	0.890	3.363	0.545	0.145	0.192	0.093
2000Q4	148	2270.491	6.748	4.807	0.417	0.877	3.495	0.555	0.132	0.193	0.094
2001Q1	146	1805.041	6.471	4.803	0.396	0.896	3.535	0.567	0.138	0.175	0.098
ALL	2967	1466.206	8.556	3.174	0.181	0.808	3.134	0.414	0.281	0.195	0.042

Source: International Data Corporation (IDC) Quarterly Server Tracker matched to more detailed Server characteristics from several industry data sources and trade magazines.
Notes: All the entries (except model numbers and quantity) are weighted by server model sales.

TABLE 3 - RESULTS FROM SIMPLE LOGIT FOR PCs

	(1)	(2)	(3)	(4)	(5)	(6)
Estimation method	OLS	IV	OLS	IV	IV	IV
Dependent variable	$\ln(S_{jt})-\ln(S_{0t})$	$\ln(S_{jt})-\ln(S_{0t})$	$\ln(S_{jt})-\ln(S_{0t})$	$\ln(S_{jt})-\ln(S_{0t})$	$\ln(S_{jt})-\ln(S_{0t})$	$\ln(S_{jt})-\ln(S_{0t})$
Price	-0.336***	-1.400***	-0.404***	-2.085***	-2.275***	-2.488***
	(0.037)	(0.281)	(0.037)	(0.204)	(0.239)	(0.258)
Benchmark	0.305***	0.953***	0.388***	1.153***	1.239***	1.336***
	(0.108)	(0.211)	(0.095)	(0.160)	(0.180)	(0.190)
RAM	-0.458***	0.339	-0.333***	0.920***	1.062***	1.221***
	(0.101)	(0.246)	(0.105)	(0.220)	(0.239)	(0.262)
CD-ROM	0.226**	0.257**	0.188*	0.278**	0.288**	0.299**
	(0.095)	(0.112)	(0.096)	(0.130)	(0.136)	(0.143)
Ethernet	0.140*	0.354***	0.105	0.463***	0.504***	0.549***
	(0.077)	(0.103)	(0.077)	(0.109)	(0.116)	(0.123)
Desktop	0.375***	-0.406*	0.273***	-0.908***	-1.042***	-1.192***
	(0.070)	(0.213)	(0.071)	(0.169)	(0.193)	(0.208)
5th Generation	1.068***	1.814***	0.894***	2.520***	2.704***	2.911***
	(0.244)	(0.364)	(0.229)	(0.379)	(0.410)	(0.432)
6th Generation	0.889***	2.314***	0.954***	3.652***	3.957***	4.299***
	(0.268)	(0.496)	(0.252)	(0.472)	(0.523)	(0.556)
7th Generation	1.112***	2.037***	1.084***	3.087***	3.313***	3.568***
	(0.395)	(0.526)	(0.385)	(0.561)	(0.595)	(0.626)
Monitor Size	-0.066***	-0.086***	-0.066***	-0.097***	-0.101***	-0.105***
	(0.008)	(0.009)	(0.008)	(0.010)	(0.010)	(0.011)
Trend			-0.051***	-0.368***	-0.404***	-0.444***
			(0.013)	(0.041)	(0.047)	(0.051)
Firm Dummies (9)	yes	yes	yes	yes	yes	yes
Time Dummies (21)	yes	yes	no	no	no	no
Identification		60.383		65.425	50.836	27.114
		[0.000]		[0.000]	[0.000]	[0.000]
1st Stage F-test		8.8		27.21	30.40	40.620
		[0.000]		[0.000]	[0.000]	[0.000]
Own Price Elasticities						
Mean	-0.73	-3.04	-0.88	-4.52	-4.94	-5.40
Standard deviation	0.31	1.28	0.37	1.90	2.07	2.27
Median	-0.68	-2.83	-0.82	-4.21	-4.60	-5.03
% inelastic demands	85.51%	0.70%	71.44%	0.03%	0.00%	0.00%

Notes: Based on 3,305 observations from the US PC market for large business customers. "Test of Over Identification" is the Hansen-Sargan test of over-identification for the IV regressions with the p-values in square parentheses. Robust standard errors are reported in parenthesis below coefficients: *significant at 10%; **significant at 5%; ***significant at 1%.

TABLE 4 - RESULTS FROM SIMPLE LOGIT FOR SERVERS

	(1)	(2)	(3)	(4)	(5)	(6)
Estimation method	OLS	IV	OLS	IV	IV	IV
Dependent variable	$\ln(S_{jt})-\ln(S_{0t})$	$\ln(S_{jt})-\ln(S_{0t})$	$\ln(S_{jt})-\ln(S_{0t})$	$\ln(S_{jt})-\ln(S_{0t})$	$\ln(S_{jt})-\ln(S_{0t})$	$\ln(S_{jt})-\ln(S_{0t})$
Price	-0.040***	-0.075***	-0.040***	-0.179***	-0.201***	-0.234***
	(0.003)	(0.020)	(0.003)	(0.031)	(0.035)	(0.041)
RAM	0.002	0.031*	0.002	0.116***	0.133***	0.161***
	(0.007)	(0.019)	(0.008)	(0.032)	(0.036)	(0.042)
Windows	-0.861***	-0.567***	-0.867***	0.305	0.484	0.766**
	(0.113)	(0.196)	(0.114)	(0.282)	(0.310)	(0.357)
Windows×RAM	0.013	0.025*	0.012	0.060***	0.068***	0.079***
	(0.011)	(0.014)	(0.011)	(0.023)	(0.025)	(0.029)
Symmetric Parallel Processor	0.474***	0.705***	0.474***	1.388***	1.528***	1.748***
	(0.081)	(0.156)	(0.081)	(0.224)	(0.246)	(0.284)
Rack Optimized	0.455***	0.337**	0.458***	-0.005	-0.076	-0.187
	(0.110)	(0.134)	(0.110)	(0.182)	(0.197)	(0.225)
Number of Racks	-0.009	0.006	-0.008	0.051**	0.060**	0.074***
	(0.010)	(0.013)	(0.010)	(0.022)	(0.024)	(0.028)
Linux	0.037	0.542	-0.033	1.995***	2.307***	2.795***
	(0.413)	(0.484)	(0.392)	(0.605)	(0.647)	(0.715)
Unix	-0.675***	0.351	-0.681***	3.393***	4.019***	5.000***
	(0.166)	(0.589)	(0.168)	(0.907)	(1.000)	(1.176)
OS390/400	-1.750***	-0.711	-1.717***	2.390**	3.020***	4.008***
	(0.204)	(0.611)	(0.204)	(0.936)	(1.037)	(1.218)
VMS	-1.961***	-1.620***	-2.009***	-0.610	-0.396	-0.059
	(0.255)	(0.330)	(0.257)	(0.574)	(0.639)	(0.734)
Other OS	-2.088***	-1.094*	-2.070***	1.874**	2.480**	3.429***
	(0.222)	(0.596)	(0.222)	(0.900)	(0.992)	(1.163)
Trend			-0.030***	-0.144***	-0.161***	-0.189***
			(0.007)	(0.025)	(0.028)	(0.032)
Firm Dummies (9)	yes	yes	yes	yes	yes	yes
Time Dummies (21)	yes	yes	no	no	no	no
Test of Over Identification		64.409		35.389	20.061	12.03
		[0.000]		[0.000]	[0.000]	[0.002]
1st Stage F-test		18.53		5.82	8.70	12.87
		[0.000]		[0.000]	[0.000]	[0.000]
Own Price Elasticities						
Mean	-0.63	-1.18	-0.63	-2.84	-3.18	-3.71
Standard deviation	0.62	1.15	0.62	2.79	3.12	3.64
Median	-0.44	-0.83	-0.44	-2.01	-2.25	-2.62
% inelastic demands	81.13%	57.40%	80.89%	32.02%	28.08%	22.11%

Notes: Based on 2,967 observations from the US server market. "Test of Over Identification" is the Hansen-Sargan test of over-identification for the IV regressions with the p-values in square parentheses. Robust standard errors are reported in parenthesis below coefficients: *significant at 10%; **significant at 5%; ***significant at 1%.

TABLE 5 - RESULTS FROM THE FULL MODEL

Estimation method	GMM
PC - Means	
Price	-3.301***
	(0.629)
Benchmark	0.021
	(1.243)
RAM	0.760**
	(0.316)
CD-ROM	0.275**
	(0.130)
Ethernet	0.423***
	(0.134)
5th Generation	2.783***
	(0.395)
6th Generation	4.053***
	(0.574)
7th Generation	2.709***
	(0.606)
Server - Means	
Price	-0.282***
	(0.089)
RAM	0.173***
	(0.057)
Windows	0.794*
	(0.451)
Windows×RAM	0.077**
	(0.034)
Symmetric Parallel Processor	1.787***
	(0.390)
Rack Optimized	-0.185
	(0.234)
Number of Racks	0.060
	(0.039)
PC - Standard Deviations	
Price	0.916**
	(0.363)
Benchmark	1.321
	(0.822)
Server - Standard Deviations	
Price	0.048**
	(0.024)
RAM	0.014
	(0.104)
GMM Objective (df)	75.613 (10)

Notes: Based on 6,272 observations from the PC and Server market. Parameters estimated via a two-step GMM algorithm described in the estimated section. We include all the characteristics in Tables 3 and 4, i.e. for PCs: desktop, monitor size, CD-ROM, firm dummies and time trend; for servers: full set of operating system and firm dummies and time trend.The standard errors take into account the variance introduced through the simulation by bootstrapping the relevant component of the variance in the moment conditions. Robust standard errors are reported in parentheses below estimated coefficients: *significant at 10%;   **significant at 5%; ***significant at 1%.

TABLE 6 - RESULTS FROM ALTERNATIVE MODELS

	(1)	(2)	(3)	(4)
Estimation method	GMM strong complementarity	GMM strong complementarity	GMM strong complementarity	GMM "free" complementarity
PC - Means				
Price	-3.102***	-3.057***	-2.844***	-3.314***
	(0.256)	(0.258)	(0.326)	(0.592)
Benchmark	0.145	0.059	-0.401	-0.153
	(0.429)	(0.477)	(0.284)	(1.176)
RAM	0.965***	0.973***	1.016***	0.801***
	(0.232)	(0.245)	(0.296)	(0.303)
CD-ROM	0.268**	0.271**	0.281**	0.278**
	(0.132)	(0.133)	(0.134)	(0.131)
Ethernet	0.426***	0.438***	0.445***	0.444***
	(0.114)	(0.115)	(0.125)	(0.131)
5th Generation	3.056***	3.055***	3.080***	2.821***
	(0.438)	(0.406)	(0.453)	(0.399)
6th Generation	4.496***	4.517***	4.579***	4.132***
	(0.536)	(0.515)	(0.622)	(0.567)
7th Generation	3.258***	3.285***	3.301***	2.733***
	(0.623)	(0.637)	(0.697)	(0.629)
Constant	0.832	0.798	0.224	-3.368***
	(0.799)	(0.644)	(0.817)	(0.704)
Server - Means				
Price	-0.231***	-0.233***	-0.256***	-0.674***
	(0.039)	(0.040)	(0.046)	(0.155)
RAM	0.160***	0.163***	0.181***	0.208***
	(0.040)	(0.041)	(0.046)	(0.066)
Windows	0.742**	0.755**	0.939**	1.543***
	(0.346)	(0.360)	(0.401)	(0.483)
Windows×RAM	0.075**	0.076***	0.083***	0.133***
	(0.028)	(0.028)	(0.031)	(0.039)
Symmetric Parallel Processor	1.765***	1.777***	1.924***	2.620***
	(0.278)	(0.286)	(0.322)	(0.408)
Rack Optimized	-0.230	-0.240	-0.318	-0.373
	(0.217)	(0.220)	(0.240)	(0.273)
Number of Racks	0.064**	0.064**	0.077***	0.140***
	(0.029)	(0.028)	(0.032)	(0.034)
Constant	-10.649***	-10.596***	-10.389***	-8.096***
	(0.297)	(0.212)	(0.389)	(0.669)

PC - Standard Deviations				
Price	0.728***	0.702***	0.593***	0.902***
	(0.064)	(0.094)	(0.014)	(0.338)
Benchmark	1.176***	1.218***	1.484***	1.450*
	(0.170)	(0.191)	(0.026)	(0.752)
Constant			0.021	
			(0.023)	
Server - Standard Deviations				
Price	0.001	0.002	0.001	0.162***
	(0.011)	(0.013)	(0.009)	(0.042)
RAM		0.005	0.007	0.027
		(0.013)	(0.070)	(0.091)
Constant			0.806***	
			(0.240)	
Γ parameter				2.647**
				(1.271)
GMM Objective (df)	75.111 (12)	70.344 (10)	74.293 (8)	57.493 (9)

Notes: Based on 6,272 observations from the PC and Server market. Parameters estimated via a two-step GMM algorithm described in the estimated section. We include all the characteristics in Tables 3 and 4, i.e. for PCs: desktop, monitor size, CD-ROM, firm dummies and time trend; for servers: full set of operating system and firm dummies and time trend.The standard errors take into account the variance introduced through the simulation by bootstrapping the relevant component of the variance in the moment conditions. Robust standard errors are reported in parentheses below estimated coefficients: *significant at 10%; **significant at 5%; ***significant at 1%.

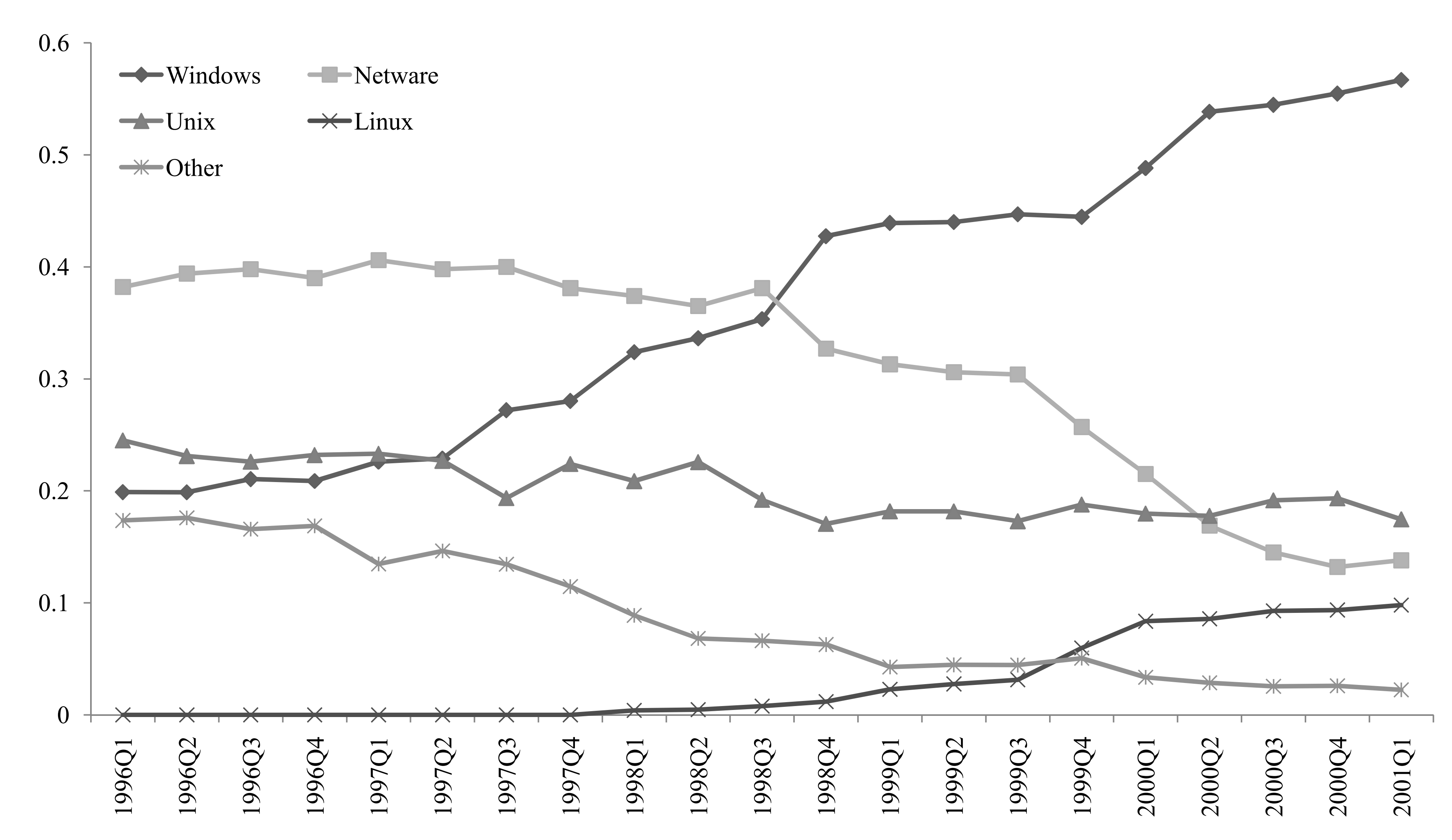
TABLE 7 - ROBUSTNESS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Estimation method	GMM sample of 250 consumers	GMM sample of 500 consumers	GMM bundle or PC purchase every two yrs	GMM PC every yr, Server every two yrs	GMM IV 6 PC, 5 Server	GMM IV 4 PC, 5 Server	GMM price, bench, con	GMM only server price
PC - Means								
Price	-3.853***	-2.795***	-3.002***	-3.353***	-3.622***	-5.598	-2.768***	-3.350***
	(0.673)	(0.501)	(0.411)	(0.604)	(0.676)	(3.882)	(0.555)	(0.635)
Benchmark	0.555	-2.503	0.070	0.088	-0.786	-1.388	-1.971*	0.020
	(0.610)	(1.572)	(0.493)	(1.114)	(2.355)	(5.770)	(1.152)	(1.229)
RAM	0.677***	1.088***	0.837***	0.747**	0.639*	0.244	0.753**	0.765**
	(0.249)	(0.284)	(0.278)	(0.303)	(0.348)	(0.568)	(0.320)	(0.312)
CD-ROM	0.257***	0.321***	0.261**	0.267**	0.304**	0.315	0.316**	0.275**
	(0.127)	(0.140)	(0.132)	(0.129)	(0.135)	(0.193)	(0.133)	(0.131)
Ethernet	0.351***	0.490***	0.486***	0.410***	0.403***	0.305	0.443***	0.424***
	(0.109)	(0.134)	(0.127)	(0.130)	(0.149)	(0.228)	(0.136)	(0.132)
5th Generation	2.805***	2.955***	2.811***	2.766***	2.869***	3.128***	3.153***	2.795***
	(0.405)	(0.461)	(0.401)	(0.391)	(0.400)	(0.547)	(0.491)	(0.394)
6th Generation	4.013***	4.517***	4.103***	4.007***	4.154***	4.296***	4.619***	4.066***
	(0.548)	(0.607)	(0.517)	(0.558)	(0.616)	(0.547)	(0.718)	(0.569)
7th Generation	2.724***	3.034***	2.757***	2.663***	2.529***	1.858	3.005***	2.702***
	(0.570)	(0.738)	(0.652)	(0.597)	(0.700)	(1.335)	(0.759)	(0.605)
Constant	-2.940***	-3.528***	-2.708***	-3.451***	-3.269***	-2.402	-6.319**	-3.379***
	(0.831)	(0.936)	(0.709)	(0.704)	(0.653)	(2.671)	(2.818)	(0.708)
Server - Means								
Price	-0.281***	-0.352***	-0.288***	-0.258***	-0.249***	-0.298**	-0.352***	-0.281***
	(0.089)	(0.133)	(0.094)	(0.085)	(0.081)	(0.131)	(0.113)	(0.086)
RAM	0.154***	0.203***	0.177***	0.162***	0.161***	0.180***	0.220**	0.174***
	(0.068)	(0.058)	(0.052)	(0.051)	(0.045)	(0.060)	(0.096)	(0.049)
Windows	0.724*	1.069*	0.828*	0.688	0.683**	0.888	1.342**	0.781*

	(0.393)	(0.556)	(0.460)	(0.431)	(0.342)	(0.737)	(0.590)	(0.436)
Windows×RAM	0.079***	0.092***	0.078**	0.072**	0.074**	0.085*	0.102**	0.076**
	(0.032)	(0.041)	(0.034)	(0.032)	(0.033)	(0.047)	(0.041)	(0.032)
Symmetric Parallel Processor	1.766***	2.015***	1.810***	1.698***	1.690***	1.858***	2.234***	1.773***
	(0.340)	(0.498)	(0.399)	(0.369)	(0.307)	(0.665)	(0.486)	(0.358)
Rack Optimized	-0.153	-0.266	-0.199	-0.145	-0.154	-0.208	-0.441	-0.176
	(0.206)	(0.267)	(0.237)	(0.223)	(0.203)	(0.329)	(0.296)	(0.227)
Number of Racks	0.066***	0.084	0.063	0.056	0.055	0.074	0.094*	0.060
	(0.031)	(0.050)	(0.040)	(0.036)	(0.035)	(0.058)	(0.052)	(0.037)
Constant	-5.807***	-5.807***	-5.809***	-5.128***	-5.896***	-5.748***	-6.197***	-5.816***
	(0.275)	(0.260)	(0.228)	(0.223)	(0.294)	(0.269)	(0.566)	(0.228)
PC - Standard Deviations								
Price	1.220**	0.520*	0.758***	0.955***	1.140***	2.292	0.795***	0.938***
	(0.326)	(0.283)	(0.273)	(0.346)	(0.413)	(1.777)	(0.270)	(0.362)
Benchmark	1.021**	2.794**	1.610***	1.282*	1.938	2.690	2.658***	1.332
	(0.512)	(1.102)	(0.444)	(0.771)	(1.532)	(3.199)	(0.910)	(0.812)
Constant							2.569	
							(1.839)	
Server - Standard Deviations								
Price	0.050*	0.062**	0.048*	0.042*	0.035	0.054*	0.049*	0.049**
	(0.028)	(0.031)	(0.025)	(0.025)	(0.037)	(0.030)	(0.028)	(0.024)
RAM	0.018	0.011	0.011	0.010	0.000	0.007	0.017	
	(0.084)	(0.090)	(0.106)	(0.103)	(0.159)	(0.312)	(0.145)	
Constant							0.930*	
							(0.528)	
GMM Objective (df)	80.934 (10)	68.583 (10)	71.783 (10)	88.356 (10)	54.146 (5)	46.723 (3)	56.899 (8)	79.292 (12)

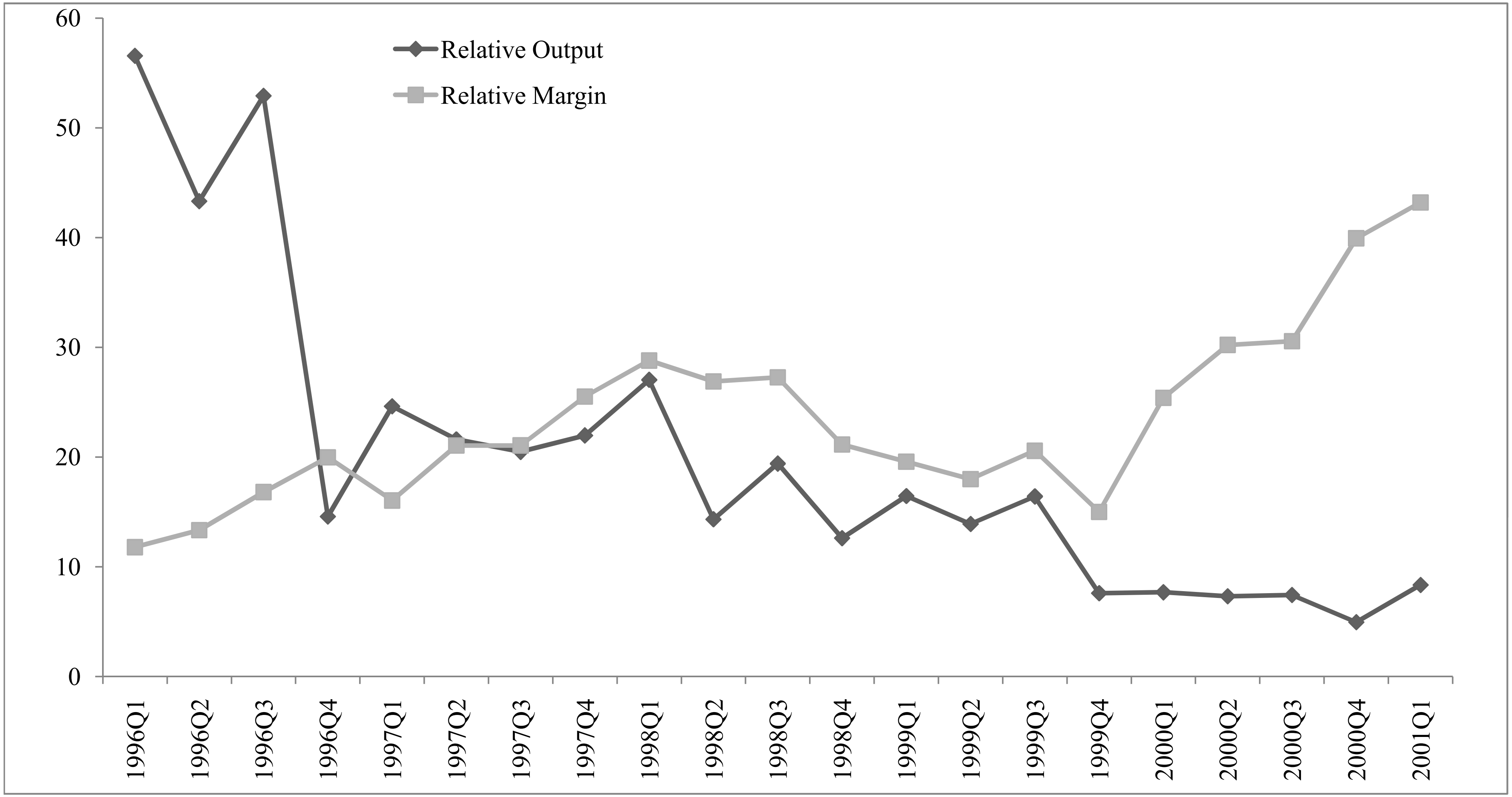
Notes: Based on 6,272 observations from the PC and Server market. Parameters estimated via a two-step GMM algorithm described in the estimated section. We include all the characteristics in Tables 3 and 4, i.e. for PCs: desktop, monitor size, CD-ROM, firm dummies and time trend; for servers: full set of operating system and firm dummies and time trend.The standard errors take into account the variance introduced through the simulation by bootstrapping the relevant component of the variance in the moment conditions. Robust standard errors are reported in parentheses below estimated coefficients: *significant at 10%; **significant at 5%; ***significant at 1%.

FIGURE 1: EVOLUTION OF MARKET SHARES FOR SOFTWARE VENDORS IN US (units)



Source: International Data Corporation (IDC) Quarterly Server Tracker survey, US only.
Notes: This plots the evolution of shares for different operating systems 1996-2001. Shares are measures in unit volumes. Other includes operating systems include IBM's OS390/400, Compaq's VMS and some other smaller non-Unix operating systems.

FIGURE 2: RELATIVE MARK-UP AND INTEROPERABILITY EFFECTS



Notes: This plots the evolution of relative output and markup effects based on the estimated coefficients from Table 5 and the formulas provided in Appendix C.

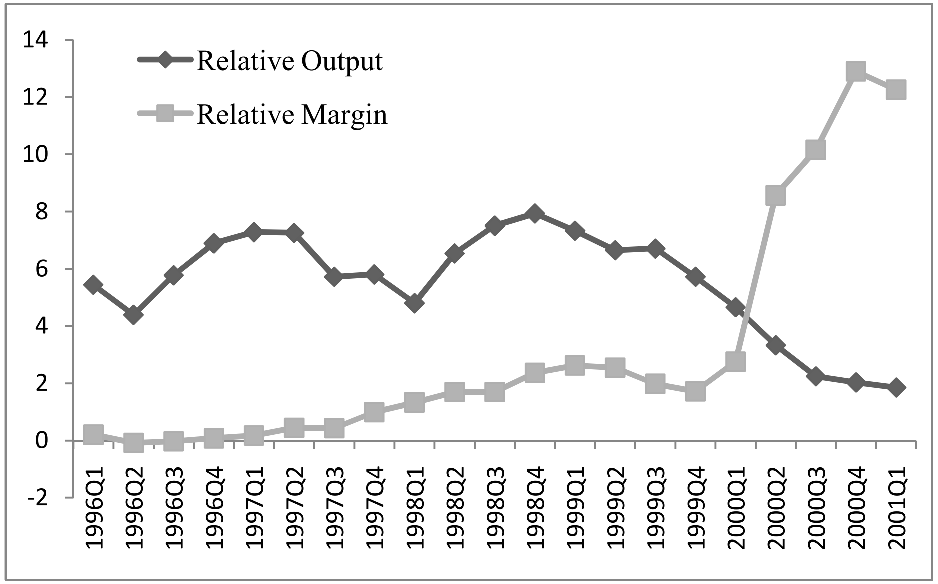


FIGURE 3A: COLUMN 1, TABLE 6

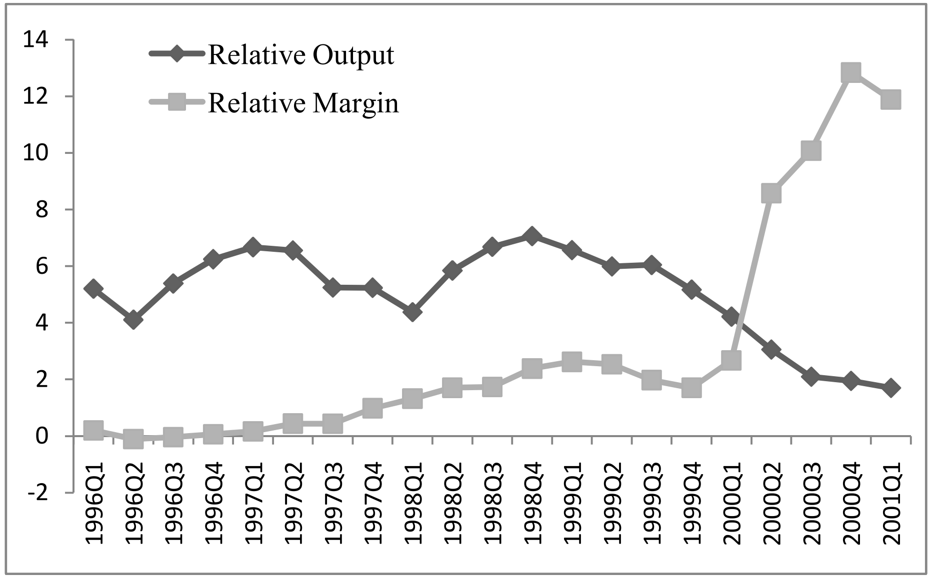


FIGURE 3C: COLUMN 3, TABLE 6

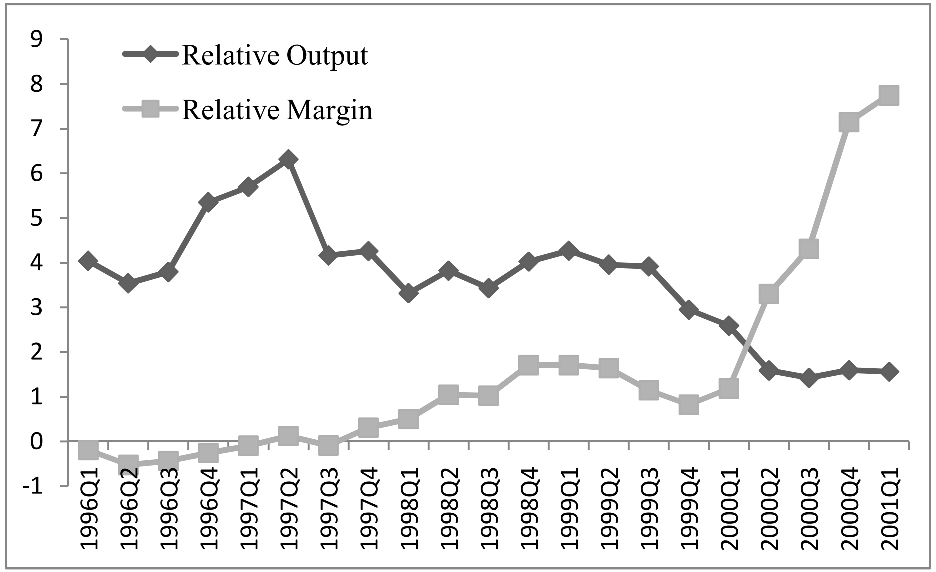


FIGURE 3B: COLUMN 2, TABLE 6

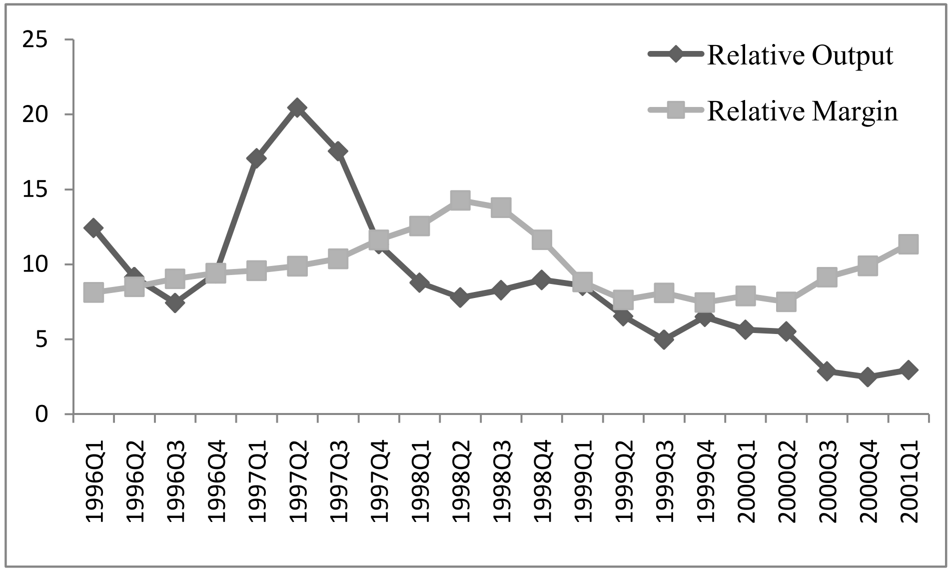


FIGURE 3D: COLUMN 4, TABLE 6

Notes: These plot the evolution of relative output and markup effects based on the estimated coefficients of each column from Table 6 and the formulas provided in Appendix C. Three quarter moving averages are shown here to smooth some outliers.

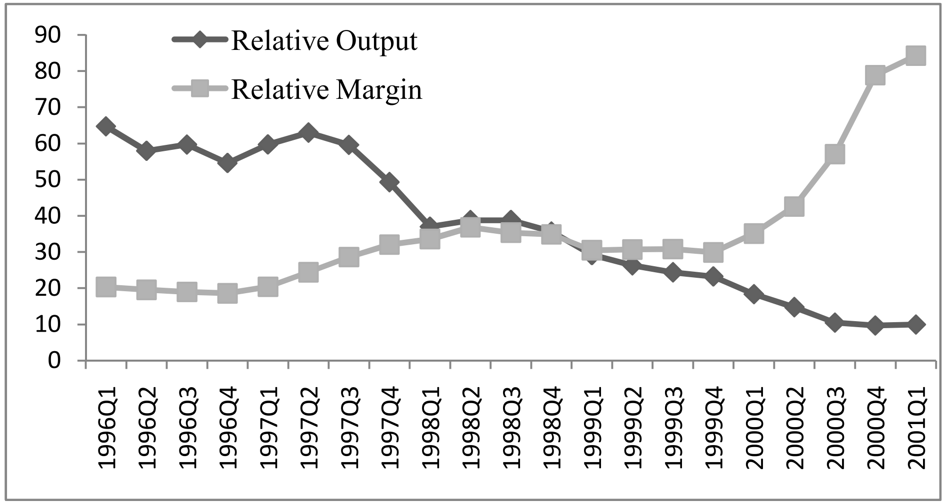


FIGURE 4A: COLUMN 1, TABLE 7

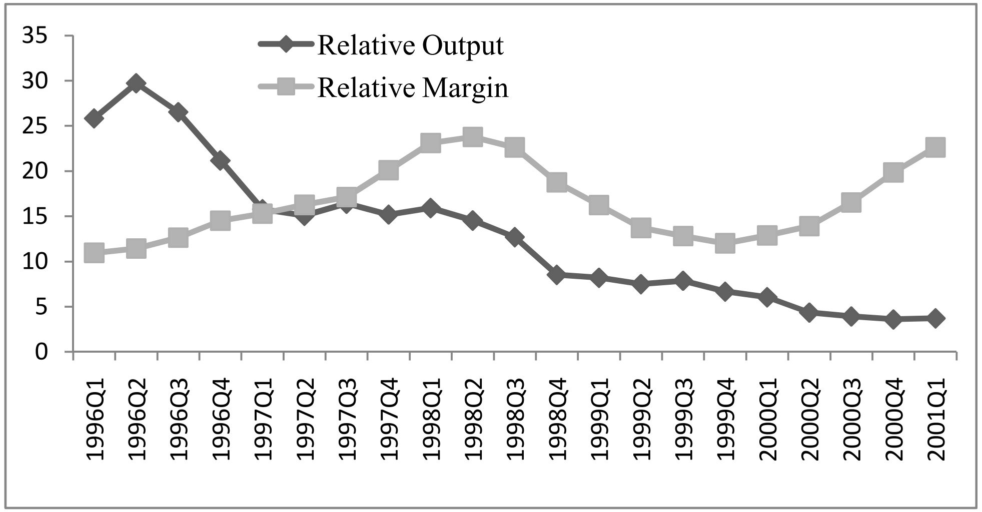


FIGURE 4C: COLUMN 3, TABLE 7

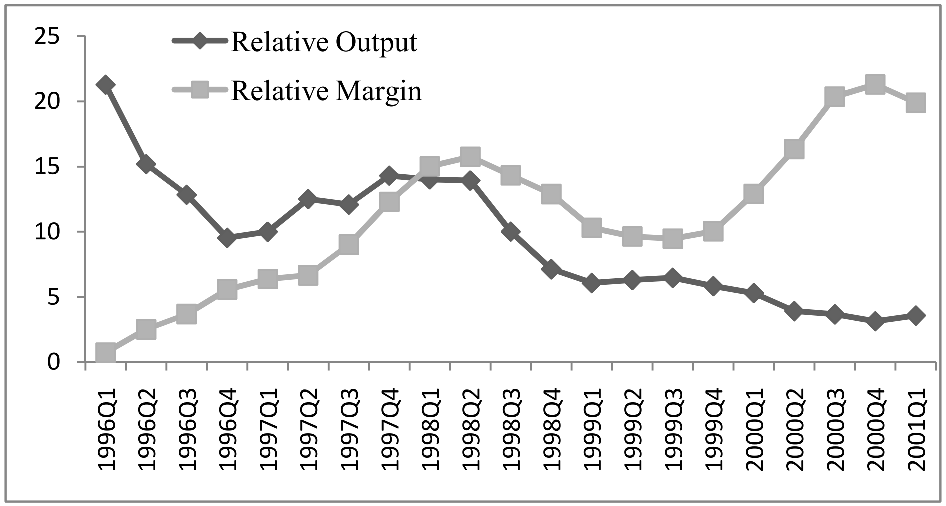


FIGURE 4B: COLUMN 2, TABLE 7

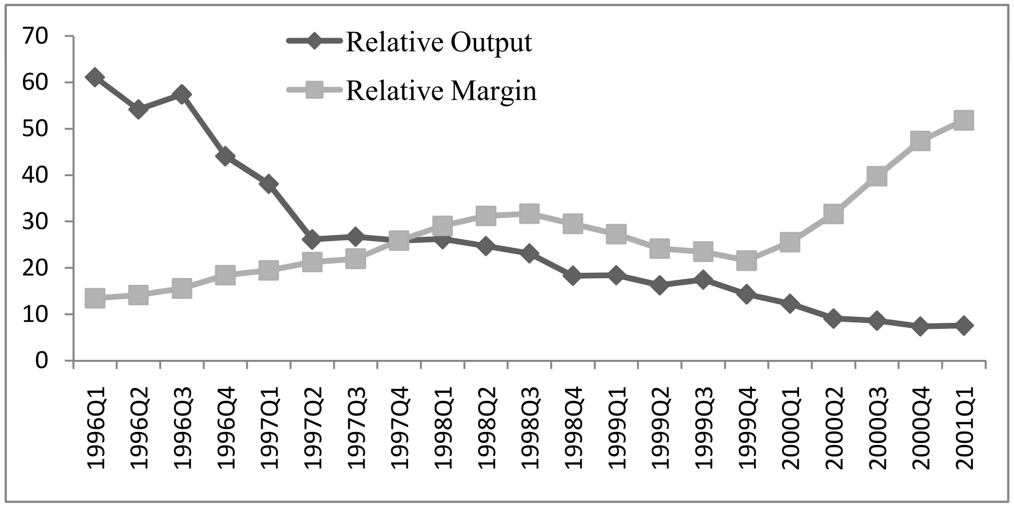


FIGURE 4D: COLUMN 4, TABLE 7

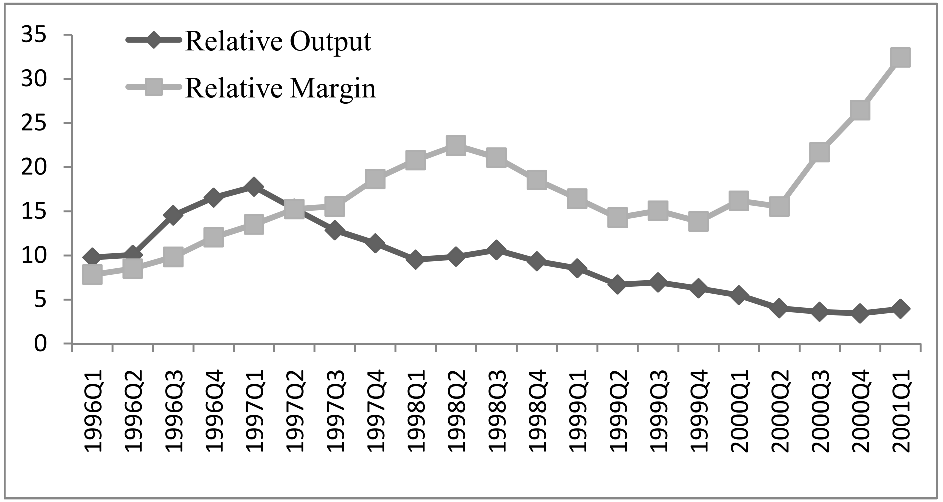


FIGURE 4E: COLUMN 5, TABLE 7

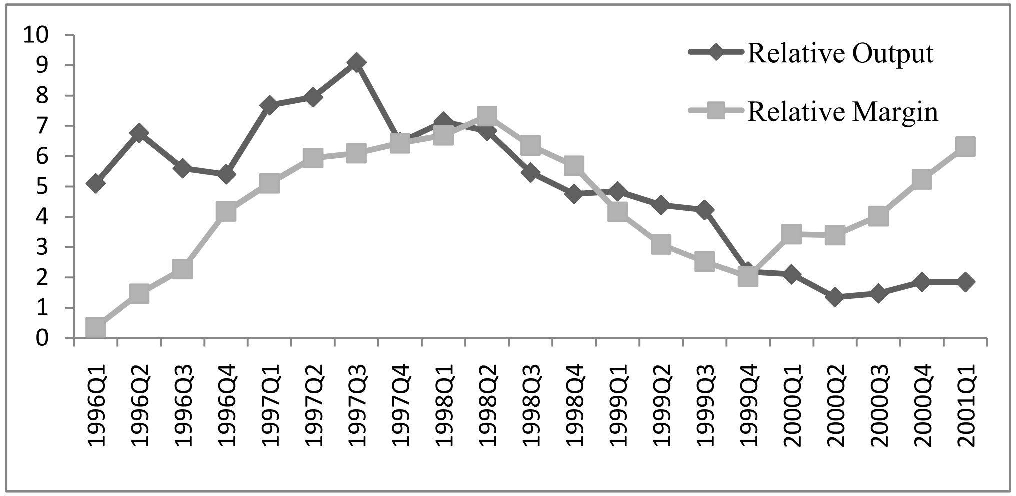


FIGURE 4G: COLUMN 7, TABLE 7

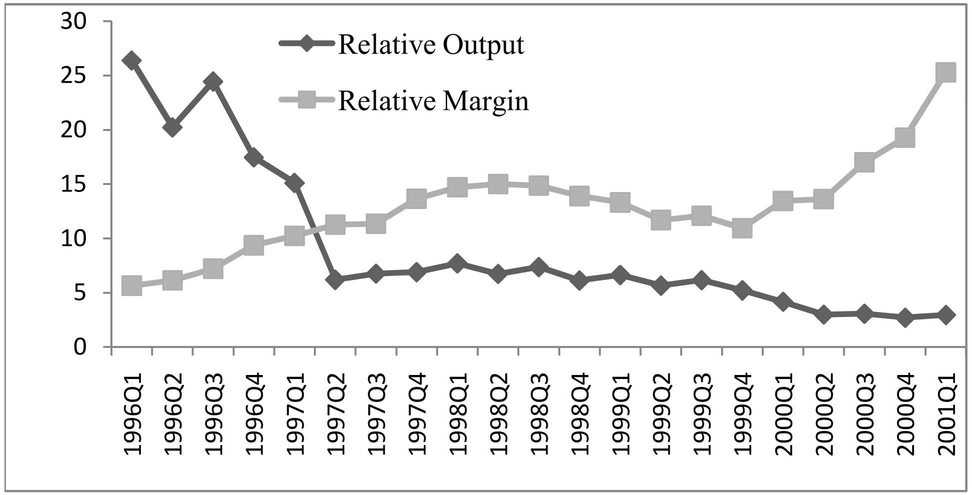


FIGURE 4F: COLUMN 6, TABLE 7

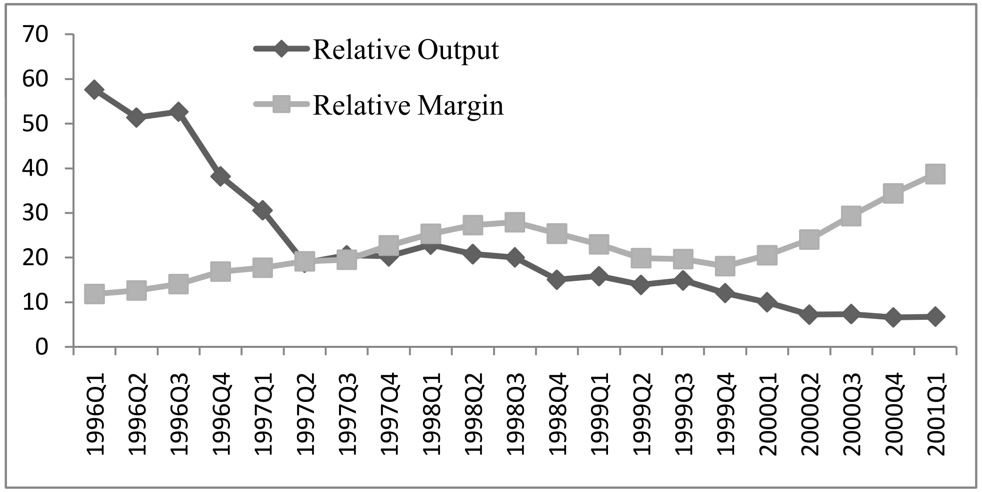


FIGURE 4H: COLUMN 8, TABLE 7

Notes: These plot the evolution of relative output and markup effects based on the estimated coefficients of each column from Table 7 and the formulas provided in Appendix C. Three quarter moving averages are shown here to smooth some outliers.

TABLE A1 - LOGIT DEMAND FOR PCs - First Stage Results

Instruments	(1)	(2)	(3)	(4)
Number of models produced by firm	38.422***	34.248***	30.549***	36.478***
	(6.513)	(5.145)	(5.212)	(4.713)
Number of models produced by other firms	11.180***	7.116***	4.676***	3.763***
	(3.836)	(1.167)	(1.125)	(1.115)
Sum of RAM of firm models	-0.100	-0.392***	-0.378***	-0.189***
	(0.168)	(0.130)	(0.130)	(0.079)
Sum RAM on rival firm's models	0.285**	-0.031	-0.031	-0.065***
	(0.128)	(0.064)	(0.064)	(0.018)
Sum of speed of firm's models	-0.088**	-0.014	0.051***	
	(0.044)	(0.029)	(0.023)	
Sum of other firms' model speed	-0.110***	-0.031**	-0.009	
	(0.035)	(0.011)	(0.010)	
Sum of hard disk space of own firm models	2.802***	2.623***		
	(0.925)	(0.872)		
Sum hard disk space of other firm's models	1.153***	0.860***		
	(0.338)	(0.210)		

Notes: These are the first stage results from Table 3. The regressions inculde all the exogenous variables in Table 3. Column 1 corresponds to column 2, column 2 corresponds to column 4, column 3 corresponds to column 5 and column 4 corresponds to column 6 in Table 3. Coefficients and standard errors are multiplied by 1,000. Robust standard errors are reported in parenthesis below coefficients: *significant at 10%; **significant at 5%; ***significant at 1%.

TABLE A2 - LOGIT DEMAND FOR SERVERS - First Stage Results

	(1)	(2)	(3)	(4)
Instruments				
Number of models produced by firm	-722.195	80.373**	91.446***	107.739***
	(931.887)	(39.395)	(26.849)	(25.253)
Number of models produced by other firms	-948.912	-89.467***	-81.724***	-58.440***
	(933.217)	(23.741)	(22.473)	(18.793)
Sum RAM on rival firm's models	0.031***	0.017***	0.010***	0.007***
	(0.005)	(0.006)	(0.002)	(0.002)
		-3568.634	-8592.337**	
		(5156.247)	(4135.553)	
		-4989.916	-2513.099	
		(3507.393)	(3125.327)	
Sum RAM on rival firm's models		-203.671		
		(131.086)		
Sum RAM on rival firm's models		-2.108		
		(5.148)		
Sum RAM on rival firm's models		64.556		
		(198.946)		

Notes: These are the first stage results from Table 4. The regressions inculde all the exogenous variables in Table 4. Column 1 corresponds to column 2 in Table 4, column 2 to column 4, column 3 to column 5 and column 4 to column 6 in Table 4. Coefficients and standard errors are multiplied by 1,000. Robust standard errors are reported in parenthesis below coefficients: *significant at 10%; **significant at 5%; ***significant at 1%.

TABLE A3 - RELATIVE OUTPUT AND MARGIN FROM THE FULL MODEL (90% CI)

Period	Relative Output	Relative Margin
1996Q1	56.6	11.8
	(30.5 - 230.3)	(6.4 - 19.8)
1996Q2	43.3	13.3
	(30.2 - 131.5)	(8.5 - 22.3)
1996Q3	52.9	16.8
	(25.5 - 141.1)	(12.6 - 29.5)
1996Q4	14.6	20.0
	(9.2 - 53.7)	(15.4 - 33.3)
1997Q1	24.6	16.0
	(16.5 - 114.9)	(12.3 - 28.1)
1997Q2	21.6	21.0
	(18.7 - 87.8)	(16.5 - 37.7)
1997Q3	20.5	21.1
	(13.8 - 72.2)	(17.9 - 49.8)
1997Q4	22.0	25.5
	(11.3 - 51.2)	(21.3 - 50.7)
1998Q1	27.0	28.8
	(21.9 - 85.6)	(24.6 - 65.8)
1998Q2	14.3	26.9
	(9.6 - 59)	(23.3 - 70.6)
1998Q3	19.4	27.3
	(10.8 - 58.9)	(21.6 - 60.9)
1998Q4	12.6	21.1
	(7.4 - 32.8)	(15.9 - 56.8)
1999Q1	16.4	19.6
	(8.4 - 42.2)	(15.4 - 50.6)
1999Q2	13.9	18.0
	(6.7 - 30)	(13.9 - 50.6)
1999Q3	16.4	20.6
	(7.6 - 29.7)	(16.6 - 51)
1999Q4	7.6	15.0
	(5.3 - 19)	(10.8 - 38.8)
2000Q1	7.7	25.4
	(6.5 - 18.6)	(18.1 - 139.2)
2000Q2	7.3	30.2
	(2.7 - 15.2)	(3.9 - 80.1)
2000Q3	7.4	30.5
	(3.4 - 13.6)	(9.3 - 136.7)
2000Q4	4.9	39.9
	(3.4 - 10.7)	(13.5 - 97)
2001Q1	8.3	43.2
	(2.6 - 12.1)	(22.7 - 335.2)

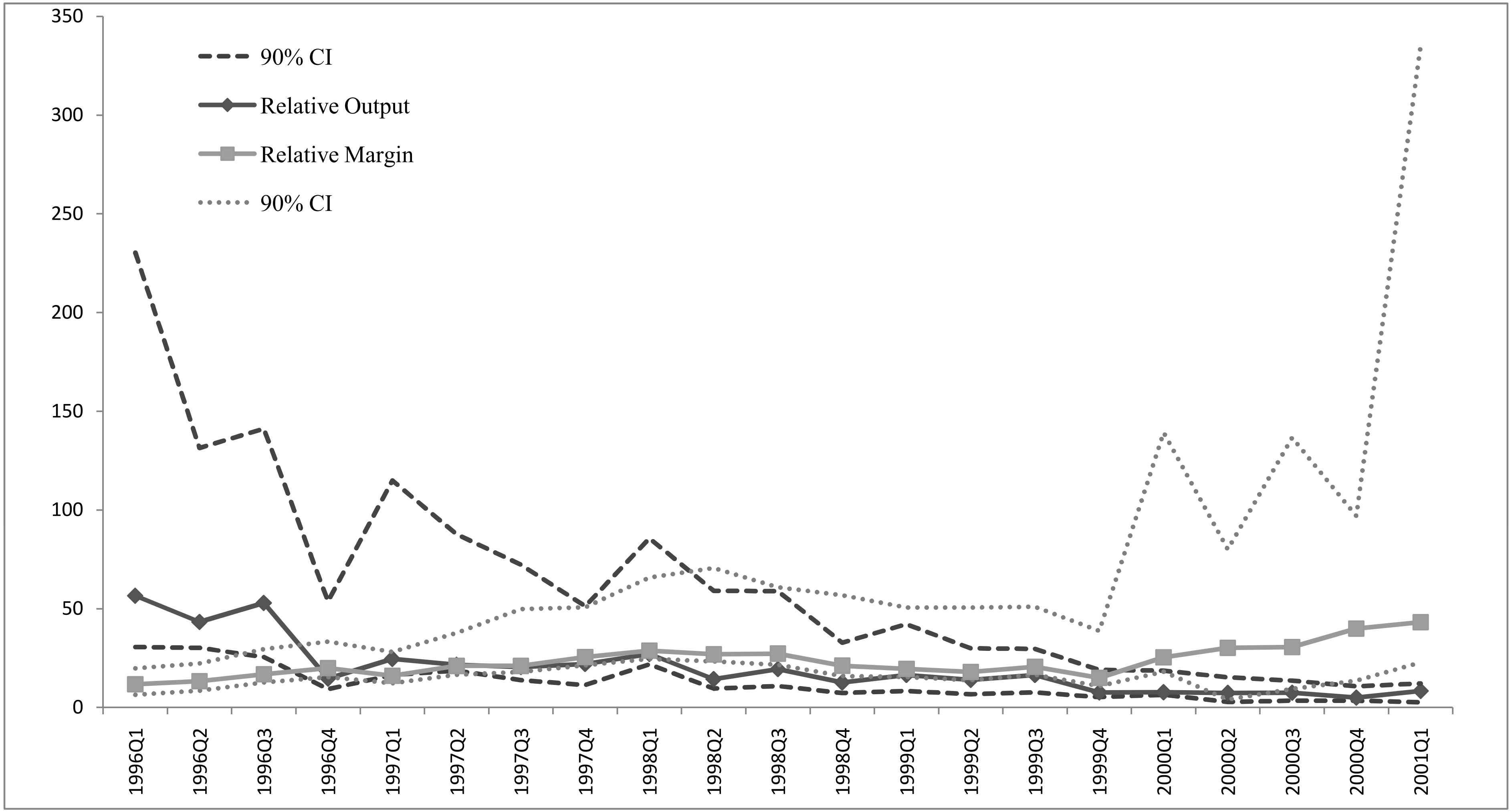
Notes: Calculated output and margin effects based on the estiamted coefficients in Table 5 and the formulas given in Appendix C. Confidence intervals were computed by using a parametric bootstrap, based on 2000 draws from the estimated joint distribution of parameters. For each draw we computed the desired quantity, thus generating a bootstrap distribution.

TABLE A4 - RELATIVE OUTPUT AND MARGIN FOR "STRONG COMPLEMENTARITY" MODEL (95% CI)

Period	Relative Output	Relative Margin
1996Q1	56.6	-0.2
	(30.5 - 230.3)	(-0.4 - 0.1)
1996Q2	43.3	-0.9
	(30.2 - 131.5)	(-1 - -0.7)
1996Q3	52.9	-0.3
	(25.5 - 141.1)	(-0.3 - -0.1)
1996Q4	14.6	0.4
	(9.2 - 53.7)	(0.1 - 0.6)
1997Q1	24.6	-0.4
	(16.5 - 114.9)	(-0.5 - -0.1)
1997Q2	21.6	0.4
	(18.7 - 87.8)	(0.3 - 0.6)
1997Q3	20.5	-0.3
	(13.8 - 72.2)	(-0.4 - -0.1)
1997Q4	22.0	0.8
	(11.3 - 51.2)	(0.5 - 1.2)
1998Q1	27.0	1.0
	(21.9 - 85.6)	(0.8 - 1.3)
1998Q2	14.3	1.4
	(9.6 - 59)	(1.2 - 1.6)
1998Q3	19.4	0.7
	(10.8 - 58.9)	(0.6 - 0.9)
1998Q4	12.6	3.0
	(7.4 - 32.8)	(2.6 - 3.6)
1999Q1	16.4	1.4
	(8.4 - 42.2)	(1.2 - 1.5)
1999Q2	13.9	0.5
	(6.7 - 30)	(0.5 - 0.6)
1999Q3	16.4	1.5
	(7.6 - 29.7)	(1.3 - 1.7)
1999Q4	7.6	0.4
	(5.3 - 19)	(0.2 - 0.7)
2000Q1	7.7	1.6
	(6.5 - 18.6)	(1.3 - 2.3)
2000Q2	7.3	7.9
	(2.7 - 15.2)	(5.4 - 8.7)
2000Q3	7.4	3.4
	(3.4 - 13.6)	(2.9 - 3.8)
2000Q4	4.9	10.1
	(3.4 - 10.7)	(7.7 - 15.1)
2001Q1	8.3	9.7
	(2.6 - 12.1)	(-44.2 - 194.1)

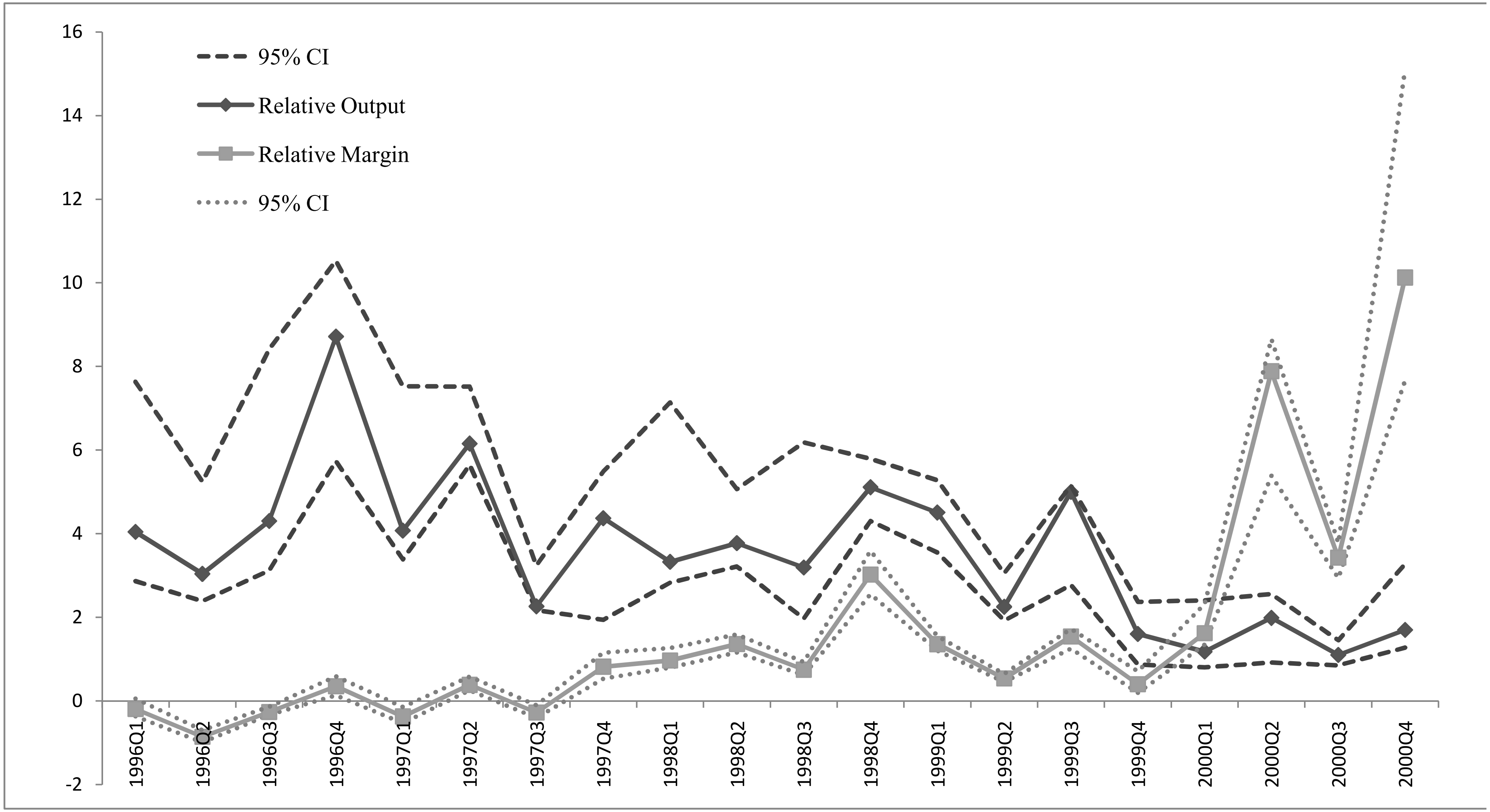
Notes: Calculated output and margin effects based on the estimated coefficients in column 3, Table 6 and the formulas given in Appendix D. Confidence intervals were computed by using a parametric bootstrap, based on 2000 draws from the estimated joint distribution of parameters. For each draw we computed the desired quantity, thus generating a bootstrap distribution.

FIGURE A1: RELATIVE MARK-UP AND INTEROPERABILITY EFFECTS WITH BOOTSTRAP S.E.



Notes: This plots the evolution of relative output and markup effects and their 90% CI based on Table A3.

FIGURE A2: RELATIVE MARK-UP AND INTEROPERABILITY EFFECTS



Notes: This plots the evolution of relative output and markup effects and their 95% CI based on Table A4.