LEVERAGING MONOPOLY POWER BY DEGRADING INTEROPERABILITY:

Theory and evidence from computer markets*

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

This Draft: December 2010

Abstract. When will a monopolist have incentives to foreclose a complementary market by degrading compatibility/interoperability? We develop a framework where leveraging extracts more rents from the monopoly market by "restoring" second degree price discrimination. In a random coefficient model with complements we derive a policy test for when incentives to reduce rival quality will hold. Our application is to 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 which were particularly strong at the turn of the 21st Century.

JEL Classification: O3, D43, L1, L4.

Keywords: Foreclosure, Anti-trust, Demand estimation, interoperability

*We would like to thank Tim Besley, Cristina Caffarra, Francesco Caselli, Stefonis Clerides, Peter Davis, Georg von Graevenitz, Ryan Kellogg, Mark McCabe, Aviv Nevo, Pierre Regibeau, Bob Stillman, Mike Whinston and participants at seminars at CEPR, LSE, Mannheim, 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 compatibility issues (called "interoperability" in software markets). 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 server operating systems work poorly with Windows. 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*". Microsoft eventually lost the case leading to the largest fine in EU anti-trust history.

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 product but competes with other firms in a market for a second product 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. Indeed, whenever there was product differentiation in the second market, the monopolist could only benefit from the presence of other firms in this second market.[1] Following the Chicago tradition, there has been much work on trying to derive efficiency explanations for exclusionary practices that

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

were previously seen as anti-competitive.[2]

More recently, studies of exclusive dealing (see Bernheim and Whinston, 1998) and tying[3] have shown that rational foreclosure in market for complements is possible in well specified models.[4] 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.

This paper makes several contributions. We propose a new theory of foreclosure through interoperability degradation and applies it to the market for PCs and servers. The theory suggests a relatively straightforward policy-relevant test for foreclosure that can be used in many contexts. To implement the test we develop a structural econometric approach using detailed market level data (quarterly data from the US PC and server markets between 1996 and 2001), which requires extending a random coefficient model to complementary products. We find strong and robust incentives for Microsoft to have degraded interoperability as the competition authorities alleged.[5]

In our theory, the reduction of competition in the secondary (server) market allows the PC monopolist to more effectively price discriminate between customers with heterogeneous demand. 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 his ability

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

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

[4] 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.

[5] 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.

to price discriminate in this way. By reducing interoperability, the monopolist can reduce competition on the server market, re-establishing the ability to price discriminate.

Although the incentive can exist in theory, whether it binds in practice depends on the interplay between two effects. The PC monopolist benefits from reducing interoperability because he gains share in the server market. But because interoperability lowers the quality of rival servers, some customers will purchase fewer PCs, and this reduces his profits from the PC monopoly.

For the argument we are making, modelling the heterogeneity between buyers is essential 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). 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). We show theoretically and empirically how different assumptions over complementarity will affects foreclose incentives. For example, we show how overly strong restrictions on the assumed form of complementarity (e.g. not allowing a PC only purchase) can cause the econometrician to underestimate the scope for foreclosure.

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, section 6 the results and section 7 concludes. In the Appendices we give more details of derivations, data and estimation techniques.

2. The Basic Argument

Our basic approach is to measure the incentives at any point in time for a monopolist in a primary market to reduce the quality of a rival in an adjacent complementary market through changes in the features of its monopoly product. In our application we examine whether Microsoft had an incentive to degrade interoperability in its PC operating system to foreclose competition in the server operating system market. In this section we give an overview of how we identify this incentive.[6]

2.1. The Test for Foreclosure Incentives. To outline our approach, we first introduce some notation that will be maintained for the rest of the paper. There are J different types of PCs offered in the market. A buyer of PC j has to pay the price $\hat{p}_j$ for the hardware and ω for the operating system of the monopolist. We observe the vector of PC prices $\mathbf{p}_J = \hat{\mathbf{p}}_J + \omega \cdot \mathbf{1}$ with element $p_j = \hat{p}_j + \omega$. For servers we observe the corresponding vector of hardware/software total system prices $\mathbf{p}_K = \hat{\mathbf{p}}_K + \boldsymbol{\omega}_K$ with element $p_k = \hat{p}_k + \omega_k$, where $\hat{p}_k$ is the hardware price of server k and ω_k is the price for the operating system running on that server.[7] We use the notation $\omega_k = \omega_M$ when the server product k uses the PC monopolist's server operating system.[8] We parameterize the degree of interoperability of the operating system of server k with the monopolist's PC operating system as $a_k \in [0, 1]$. We set $a_k = 1$ for all servers that run the server operating system of the monopolist and $a_k = a \leq 1$ for servers with competing operating systems.

[6] In this paper we do not identify unequivocally whether forecosure actually took place in the market for work-group servers by such interoperability degradation, because we cannot separately identify whether the introduction of a new Microsoft operating system only enhanced the quality of Microsoft servers relative to others or whether the decreases in interoperability also decreased the effective quality of rival server operating systems. In the anti-trust case, the European Commission (2004) claimed that changes in Windows technology did seriously reduce interoperability and in this paper we examine whether such actions would have been consistent with Microsoft's incentives or not.

[7] Note that we treat two servers with different operating system as different server products even if the hardware is identical.

[8] We adopt the notational convention of omitting subscipt "M" for PC operating system prices and quantities as this is monopolized. We have to be explicit for servers because there are several players, so "M" denotes the server operating system of the PC monopolist.Over our sample period Apple, Linux and others had less than 5% of the market, so Microsoft could be considered the monopoly supplier.

Given the price vectors we can define demand functions for total demand for PCs, $q(\mathbf{p}_j, \mathbf{p}_k, a)$, and the total demand for the monopolist's server operating system as $q_M(\mathbf{p}_j, \mathbf{p}_k, a)$. Total profits of the monopolist are therefore:

$$\Pi(\mathbf{p}_j, \mathbf{p}_k, a) = (\omega - c)q(\mathbf{p}_j, \mathbf{p}_k, a) + (\omega_M - c_M)q_M(\mathbf{p}_j, \mathbf{p}_k, a), \tag{1}$$

where c and c_M are the corresponding marginal costs for the monopolist's PC and server operating system respectively.[9]

We are interested in the incentive of the monopolist to decrease the interoperability parameter a. By the envelope theorem, there will be such an incentive if:

$$(\omega - c)\left.\frac{dq(\mathbf{p}_j, \mathbf{p}_k, a)}{da}\right|_{\omega,\omega_M} + (\omega_M - c_M)\left.\frac{dq_M(\mathbf{p}_j, \mathbf{p}_k, a)}{da}\right|_{\omega,\omega_M} < 0 \tag{2}$$

The demand derivatives with respect to the interoperability parameter are total derivatives of the respective output measures holding the monopolist's operating system prices constant. This derivative contains the direct effect of interoperability on demand as well as the impact of the price responses to a change in interoperability by all rival software and hardware producers. Total demand for PC will increase with greater interoperability because of complementarity between PCs and servers. Greater interoperability means that some customers start purchasing more PCs as the monopolist's rival servers have become more attractive. At the same time we expect the demand for the monopolist's server operating system, q_M, to decrease when interoperability increases because some customers will switch to a server with the alternative operating system. The relative impact on server and PC operating system demand from interoperability degradation will therefore be critical to the incentives to foreclose. Rearranging terms we obtain that there is an incentive to decrease interoperability

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

at the margin if:

$$\frac{\omega_M - c_M}{\omega - c} > -\frac{\left.\frac{dq(\mathbf{p}_j,\mathbf{p}_k,a)}{da}\right|_{\omega,\omega_M}}{\left.\frac{dq_M(\mathbf{p}_j,\mathbf{p}_k,a)}{da}\right|_{\omega,\omega_M}} \tag{3}$$

On the left hand side of equation (3) we have 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 ($\omega - c$). We call the expression on the right hand side of (3) the "relative output effect" as it measures the relative impact of a change in interoperability on demand for the PC operating system (increases with interoperability) and the monopolist's server operating system (decreases with interoperability) respectively.

Our estimation approach is designed to verify whether the strict inequality (3) holds in the data. The question is why this is a good test for foreclosure incentives when one might expect an optimal choice of interoperability by the monopolist to lead to a strict equality. There are several reasons why we would expect a strict inequality in the data when there is a incentive to foreclose. First, it is costly to change operating systems to reduce the degree of interoperability and there are time lags between the design of the less interoperable software and its diffusion on the market. Second, other server operating system vendors such as Novell and Linux sought to overcome the reduction in interoperability through a variety of measures such as developing "bridge" products, redesigning their own software, reverse engineering, etc. Third, 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.[10] Protocols for these server markets may provide some

[10]Enterprise servers are high-end corporate servers that manage vast amounts of mission critical data in

interoperability that is necessary in those markets, which allows for some interoperabilty also to be established in the workgroup server market. This means the monopolist would want to reduce quality of the server rivals further if he could. Finally, since the late 1990s, anti-trust action in the US and EU may have slowed down Microsoft's desire to reduce interoperability. All these reasons suggest that in the presence of foreclosure incentives we should find a strict incentive to foreclose at the margin, which is why we focus our analysis on estimating the relative margin and output effects.

2.2. Measuring the Relative Margin Effect. The margins on PC and server 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 of operating systems that allow us to infer an order of magnitude, we have to estimate the operating system margins from the data. For this estimation we therefore have to impose a specific model of price setting. Given the complementarity between software and hardware as well as between PCs and server, the move order in price setting is important for determining the pricing incentives for the monopolist. We assume that the hardware and software companies set their prices simultaneously so that the price the software company charges is then 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.[11] Maximizing equation (1) with respect to the PC operating system price ω and the monopolist's server operating system price ω_M yields the first order conditions:

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.

[11] 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 relative margin effect. If the software company would move first setting prices and the hardware company second, the software company would have to take into account the price reactions of the hardware company.

$$q + (\omega - c)\frac{\partial q}{\partial \omega} + (\omega_M - c_M)\frac{\partial q_M}{\partial \omega} = 0 \tag{4}$$

$$q_M + (\omega - c)\frac{\partial q}{\partial \omega_M} + (\omega_M - c_M)\frac{\partial q_M}{\partial \omega_M} = 0 \tag{5}$$

Denoting $\frac{\partial q}{\partial \omega}\frac{1}{q} = \varepsilon_\omega$ as the semi-elasticity of the impact of a change in price (ω) on quantity demanded (q), we can solve equations (4) and (5) for the profit margins:

PC operating system margin:

$$(\omega - c) = -\frac{1}{\varepsilon_\omega}\left(\frac{1 - \frac{q_M}{q}\frac{\varepsilon_\omega^M}{\varepsilon_{\omega_M}^M}}{1 - \frac{\varepsilon_{\omega_M}}{\varepsilon_\omega}\frac{\varepsilon_\omega^M}{\varepsilon_{\omega_M}^M}}\right) \tag{6}$$

Monopolist's server operating system margin:

$$(\omega_M - c_M) = -\frac{1}{\varepsilon_{\omega_M}^M}\left(\frac{1 - \frac{q}{q_M}\frac{\varepsilon_{\omega_M}}{\varepsilon_\omega}}{1 - \frac{\varepsilon_{\omega_M}}{\varepsilon_\omega}\frac{\varepsilon_\omega^M}{\varepsilon_{\omega_M}^M}}\right) \tag{7}$$

There are four relevant semi-elasticities: the own-price elasticity of the operating systems of PCs (ε_ω) and the monopolist's servers ($\varepsilon_{\omega_M}^M$); and the cross price elasticities of the monopolist's server with respect to PC prices (ε_ω^M) and PCs with respect to the monopolist's server prices (ε_{ω_M}). The semi-elasticities that determine the right hand side of these two equations can be estimated from PC and server sales and price data. The operating systems margins and the relative margin effect can therefore be inferred from estimating the parameters of an appropriate demand system.

A first remark on (6) and (7) is that the price cost margins differ from the standard single product monopoly margins due to the ratios of cross- to own-price elasticities of PC and server operating system demands,$\frac{\varepsilon_{\omega_M}}{\varepsilon_\omega}$ and $\frac{\varepsilon_\omega^M}{\varepsilon_{\omega_M}^M}$.These ratios are strictly positive when there is complementarity between PCs and servers implying that mark-ups will be lower than under independent demand. In general, mark-ups will be affected both by the degree of competition in the market and by the degree of complementarity. As a benchmark case, suppose that PCs and servers are perfect complements which means that customers buy servers and PCs in fixed proportions (i.e. exactly w PCs for every server purchased). With competition between different server operating systems we should generally expect $\left|\varepsilon_{\omega}^{M}\right| < \left|w\varepsilon_{\omega_M}^{M}\right|$: the demand response of the monopolist's server operating system should be greater for an increase in the server price than the PC operating system price, because the latter leads to a price increase for all servers and therefore does not lead to substitution between servers due to relative price changes. In the limit, as the server operating system market becomes perfectly competitive, i.e. $\varepsilon_{\omega_M}^{M} \rightarrow -\infty$ and $\frac{\varepsilon_{\omega}^{M}}{\varepsilon_{\omega_M}^{M}} \rightarrow 0$, the PC operating system margin of the monopolist goes to the single product monopoly margin, i.e. $(\omega - c) \rightarrow -\frac{1}{\varepsilon_{\omega}}$. At the same time the server operating system margin goes to zero, i.e. $(\omega_M - c_M) \rightarrow 0$. Hence generally, we would expect $\frac{\varepsilon_{\omega}^{M}}{\varepsilon_{\omega_M}^{M}}$ to decrease as competition in the market for server operating systems increases. One other implication of this is that a naive estimation of PC operating system margins that ignored the complementarity between PCs and servers as the current literature does, will systematically generate incorrect results for estimated margins. Generally, we would expect margins to be over-stated by the failure to recognize the complementarity between PC and servers operating systems (see Werden, 2001, Schmalensee, 2000 and Reddy et al, 2001, for a discussion).

2.3. Measuring the Relative Output Effect. While the direct impact of a uniform quality reduction of all rivals on demand can be deduced directly from the demand estimates, the total output effect needs to take into account the pricing reactions of rival server operating system and hardware producers. To measure this indirect effect of a quality change on relative output we have to impose the assumption of profit maximizing behavior also for all software and hardware companies other than the monopolist. A server with lower quality will command lower prices in equilibrium. Furthermore, if PC demand is reduced as a result of lower server qualities, PC hardware sellers will also partly accommodate by reducing their prices in order to increase demand. These equilibrium price adjustments are crucial to measure the size of the relative output effect. We therefore compute the equilibrium pricing response of each hardware and software producer to a common change of quality in non-Microsoft servers given the estimated demand function assuming a Nash equilibrium in prices. These price responses can then be used to compute the relevant demand derivatives to determine the relative output effect.

To check the robustness of our results we also estimate reduced form equations for PC server operating systems that depends only on quality indices and the estimated price cost margins of the monopolist. The derivatives of this reduced form demand with respect to the quality indices can then be used directly to calculate the relative output effect. This approach avoids the strong structural assumptions we have to make in the first approach, but has more ambiguities of interpretation. We show that the qualitative conclusions of the two approaches are essentially the same (see Appendix F for details).

3. The Model of Demand and Theories of Foreclosure

In this section we develop the theory of foreclosure and show how different types of unobserved heterogeneity (i.e. unobserved by both the researcher and the firms) map into foreclosure incentives. The theoretical mechanisms that generate foreclosure incentives are all based on theories in which competition in the server market interferes with an (privately) optimal price discrimination strategy by the PC monopolist. By (partially) foreclosing the server market, the monopolist can increase rent extraction by using the price of the PC and server operating systems to target different types of customers. In this section we first develop the demand structure that we use in the estimation and then discuss the theoretical results on foreclosure. The latter allow us to interpret the differences in results we generate from the different estimated demand models.

3.1. The Model of Demand. We model individual demand as a discrete choice of "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.[12] As before, there are J producers of PCs and K producers of servers indexed by j and k respectively[13]. 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:

$$u_{ijk}(w_i) = w_i \left[x_j\beta_i + a_k y_k \gamma_i - \lambda_i [p_j + \frac{1}{w_i} p_k] + \xi_j + \xi_k + \xi_{jk} + \epsilon_{ijk} \right] \tag{8}$$

The total price for the workgroup is given by $w_i p_j + p_k$[14] and the income sensitivity of utility of buyer i is measured by λ_i. The characteristics of PC j are captured by the vector x_j and the characteristics of server and server software k are represented by the vector y_k. The vectors β_i and γ_i represent the marginal value of these characteristics to buyer i. We normalize quality by assuming that the interoperability parameter $a = 1$ whenever server producer k has the Windows operating system installed. We assume that $a_k = a < 1$ is the

[12]Assuming that the purchase decisions are only about the setup of a whole "workgroup" implies some important abstractions from reality. If server systems are used for serving one workgroup we effectively assume that the whole system is 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 ϵ_{ijk} in equation (8). 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 ϵ_{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 as little dependent on functional form as possible.

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

[14]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" 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.

same for all non-Microsoft servers. In the case of $j = 0$, (x_j, p_j) is the null vector , while in the case of $k = 0$, (y_k, p_k) is the null vector. These represent "workgroup" purchase without a server or without PCs respectively. The models we estimate differ in whether these choices are allowed, which captures different assumptions about the degree of complementarity. The terms ξ_j and ξ_k represent unobserved quality characteristics of the PC and server respectively, while ξ_{jk} represents an interaction effect between a specific PC and server type.

The term ϵ_{ijk} represents a buyer specific shock to utility for the particular workgroup selected. Assumptions on the distribution of this term among customers will model the degree of horizontal product differentiation between different workgroup offerings. In the empirical section we assume a double exponential distribution for this variable, which leads to a typical mixed logit demand formulation. Given that we make ϵ_{ijk} workgroup specific, the variables ξ_{jk} and ϵ_{ijk} capture all of the potential complementarities between the PCs and the servers in a workgroup. In the empirical section we generally assume that $\xi_{jk} = 0$ except for one model version in which ξ_{jk} is a common shift variable for utility whenever a buyer consumes PCs and servers together.

Following BLP, we allow random coefficients on the parameter vector $\theta_i = (\beta_i, \gamma_i, \lambda_i)$ as well as heterogeneity in the size of work groups w_i (captured by a random coefficient on the server price, $\lambda_i^S \equiv \lambda_i / w_i$). We derive demand from the above utility function in the standard way (see Appendix A), the key assumptions being that ϵ_{ijk} comes from a double exponential distribution and that (θ_i, w_i) are multivariate normal distributions. We can then calculate market shares (for individual i) for PCs as:

$$s_{ij} = e^{\delta_j + \mu_{ij}} \sum_{k=0}^{K} \frac{e^{\delta_k + \mu_{ik}}}{1 + \sum_{j=1}^{J} \sum_{k=0}^{K} e^{\delta_j + \mu_{ij} + \delta_k + \mu_{ik}}} \tag{9}$$

and for servers as:

$$s_{ik} = e^{\delta_k+\mu_{ik}} \sum_{j=1}^{J} \frac{e^{\delta_j+\mu_{ij}}}{1+\sum_{j=1}^{J}\sum_{k=0}^{K} e^{\delta_j+\mu_{ij}+\delta_k+\mu_{ik}}} \tag{10}$$

where the mean utilities are:

$$\begin{aligned} \delta_j &= x_j\beta - \lambda^{PC} p_j + \xi_j^{PC}, \\ \delta_k &= a_k y_k \gamma - \lambda^{S} p_k + \xi_k^{S} \end{aligned} \tag{11}$$

and the "individual effects" are:

$$\begin{aligned} \mu_{ij} &= \sigma^{PC} x_j \nu_i^{PC} + \sigma_p^{PC} p_j \nu_{ip}^{PC} \\ \mu_{ik} &= \sigma^{S} y_k \nu_i^{S} + \sigma_p^{S} p_k \nu_{ip}^{S} \end{aligned} \tag{12}$$

The $(\nu_i^{PC}, \nu_{ip}^{PC}, \nu_i^{S}, \nu_{ip}^{S})$ is a vector of the normalized individual effects on the parameters and $(\sigma^{PC}, \sigma_p^{PC}, \sigma^{S}, \sigma_p^{S})$ is the vector of standard deviations of these effects in the population. As noted above we assume that the vector of normalized individual effects is drawn from a multivariate normal distribution with zero mean and an identity covariance matrix.[15] Notice that μ_{ij} depends on the interaction of customer specific preferences and product characteristics.

3.2. Implications of customer heterogeneity for incentives to degrade interoperabilty. With this additional structure on the demand we can gain more insight into what generates strictly positive server operating system margins. The sign of the relative margin effect is determined by the sign of the server margin in equation (7), which in turn depends only on:

[15] The choice of this distribution is ad hoc. Although the multivariate normal is the most popular choice (e.g. BLP, Nevo, 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.

$$1 - \frac{q}{q_M}\frac{\varepsilon_{\omega M}}{\varepsilon_\omega} = -\frac{1}{\varepsilon_\omega}\int\left[\frac{q(\theta,w)}{q} - \frac{q_M(\theta,w)}{q_M}\right][\bar{\varepsilon}_\omega - \varepsilon_\omega(\theta,w)]\,dP(\theta)d\Upsilon(w) \tag{13}$$

where

$$\bar{\varepsilon}_\omega = \int \varepsilon_\omega(\theta,w)dP(\theta)d\Upsilon(w)$$

and $P(\theta)$ and $P(w)$ is the population distribution function of θ and w.

Hence, the price cost margin on servers will be positive if the own price semi-elasticity of the PC operating system, $-\varepsilon_\omega(\theta)$, is positively correlated with $\left(\frac{q(\theta,w)}{q} - \frac{q_M(\theta,w)}{q_M}\right)$. This means that on average buyers with more elastic demand for PCs (a more negative $\varepsilon_\omega(\theta)$) than the aggregate elasticity of demand ($\bar{\varepsilon}_\omega$) have higher market share in PC purchases than in server purchases from the monopolist. Several things follow. First, the server margin will be zero if there is no heterogeneity. Then a monopolist does best by setting the price of the server at marginal cost and extracting all surplus through the PC operating system price. In this case there is no incentive to foreclose rivals servers. The monopolist maximizes the value of the market by having buyers use the highest value server for them and thus sets price of its own server to marginal cost. This is possible because all rent can be extracted through the PC price when there is no heterogeneity.

Proposition 1. *If there is no demand heterogeneity in the parameter vector (θ, w), then the "one monopoly profit" theory holds. The PC operating system monopolist sells the server operating system at marginal cost and extracts all rents with the PC price. The monopolist has no incentive to degrade interoperability.*

In order to generate foreclosure incentives it must therefore be the case that the optimal extraction of surplus for the monopolist involves making a margin on the server product. In that case, competition among server operating systems reduce the margin that can be earned on servers and thus restricts the ability of the PC monopolist to extract the maximal rent that can be generated from that monopoly. By limiting interoperability with rival servers, the monopolist can reduce the quality of the server rival and "restore" the margin on its own server. In the next sub-sections we explain what kind of heterogeneities generate such incentives and what kind of heterogeneities work against this effect.

3.3. Imperfect Complementarity. Second degree price discrimination incentives that lead to foreclosure require sorting by customers with low PC demand elasticities into buying the PC monopolist's server operating system and by customers with high PC demand elasticities into not buying the server. Our central model generates this feature by making the assumption that buyers do not need to buy a server in order to gain value from a workgroup of PCs. However, servers are complements to PCs in the sense that they only have value when they are consumed with PCs. We call this the *imperfect complementarity* case. To see how this setting naturally generates foreclosure incentives let us simplify the argument by assuming a very limited type of heterogeneity: buyers have different marginal valuations of server quality γ_i, which can be either $\bar{\gamma}$ or 0. We also assume that there is no other heterogeneity across consumers with respect to the server product. In particular, there is no horizontal product differentiation with respect to using a server (i.e. $\epsilon_{ijk} = \epsilon_{ij}$ for all k). For simplicity we analyze the model in a version in which there is only one brand of PC and one brand of server with the monopolist's server operating system. The latter assumption has no impact on the results.

The basic logic of proposition 2 below can then be easily understood from equation (13). Suppose the server operating system was priced at marginal cost $c_M < \bar{\gamma} y_M$ as in the no-heterogeneity case. Then the market share in server sales of customers from the pool with $\gamma_i = \bar{\gamma}$ is 1 and that of customers with $\gamma_i = 0$ is zero. At the same time the share of the $\gamma_i = 0$ group in PC sales is strictly positive. It follows that $\left(\frac{q(\gamma_i)}{q} - \frac{q_M(\gamma_i)}{q_M} \right)$ is strictly decreasing in γ_i. In other words, low γ customers have a higher share in PC sales than in server sales. At the same time, at any given prices for PCs and servers, the high γ customers have lower elasticity of demand for PCs because they gain more from buying the workgroup. This means that there is a positive correlation between the elasticity of demand of the type γ and the relative importance of that type in PC sales. By equation (13) this implies that the server price will be strictly greater than zero.

Now consider that there is competition on the server operating system market. By standard Bertrand arguments competition between the two server products will compete down the price of the lower quality product to no more than marginal cost and the higher quality firm can extract (at most) the additional value provided by its higher quality. If the rival's server product does not have too much higher quality it will extract all of the quality improvement over the monopolist's product in the server price. This means that a monopolist with the lower quality server will generate the same profit as setting the server price at marginal cost without competition and setting the conditionally optimal PC price. Fully foreclosing the competitor is therefore optimal even if the competitor has arbitrarily better quality than the monopolist. Similarly, for a firm with lower quality the monopolist cannot extract the full value of its own server quality but only the improvement over the quality of the rival. Hence, reducing the quality of the rival slightly will increase the ability to extract surplus. From this we obtain proposition 2:

Proposition 2. *Suppose that all heterogeneity between buyers is captured by $\epsilon_{ijk} = \epsilon_{ij}$ and $\gamma_i \in \{0, \bar{\gamma}\}$. Then a pure monopolist will sets the server operating system price strictly above marginal cost. Then there exists $\bar{y}_k > y_M$, such that for all $y_k \in (y_M, \bar{y}_k)$ it is optimal for the monopolist to foreclose a competitor by fully degrading the quality of a competing server OS.*

Proposition 2 holds because competition limits the ability of the PC operating system monopolist to optimally extract surplus. If customers with high elasticity of PC demand sort away from servers, then server sales can be used as a second degree price discrimination device, allowing the monopolist to extract more surplus from high server value/ low elasticity customers. This effect is necessary to generate a positive server margin for the PC monopolist.

Note that even where the foreclosure effect exists, it is not always the case that there are *marginal incentives* to foreclose. For example in the above model there are no marginal incentives to foreclose when the rival has higher quality. A small reduction in the quality has no effect on the profits of the monopolist in that case. Only a reduction below the quality of the incumbent will increase profits. In more general models there can even be a *negative marginal incentive to foreclose* when there are *global incentives to foreclose.* This arises from a vertical product differentiation effect. Locally a small increase in the quality of a higher quality rival can lead to higher profits for the monopolist by relaxing price competition as in a Shaked and Sutton (1992) style product differentiation model. Nevertheless, there may be incentives to dramatically reduce quality of the rival in order to increase profits even further. Our focus in the empirical analysis of the marginal incentives to foreclose may therefore lead to an underestimation of the true foreclosure incentives.

Note also that if the quality of the rival is high enough. The marginal value of the server will not be fully extracted by the competing server in order to increase sales. This then benefits the monopolist, so that the monopolist will want to achieve full interoperability with much better server products. There is therefore the possibility that there are no foreclosure incentives at all when the rival products are much better than the monopolist's server product.

3.4. The cases of Strong Complementarity and Free Complementarity. To test the robustness of the results from our central model we also look at a "strong" complementarity model in which it is assumed that PCs can only be used with a server. In this case, the mechanism of the previous subsection cannot generate foreclosure incentives because buyers with low valuation of sever services cannot substitute into pure PC workgroups. Nevertheless one can generate foreclosure incentives from a closely related mechanism. Suppose that initially the PC operating system monopolist has a competitor in the server operating system market that has lower quality. Although customers need a server with their PCs, these customers with low marginal valuation of server quality then have an incentive to substitute away to the lower quality server and combine it with PCs. The monopolist can again use the PC and server prices to discriminate between customers with different marginal valuations for servers. There will be a cost in such price discrimination because the monopolist looses the margin from selling his higher quality product, but if the gains from price discrimination are large enough, the presence of the rival server producer will strictly increase profits by enabling price discrimination. Suppose now that a competing high quality server supplier is added to the market (and assume as before that the server product is homogenous to that of the monopolist up to the quality differential). By the same argument as in the last subsection this will reduce the profit extraction possibility of the monopolist. The monopolist will have an incentive to reduce the quality of the rival server products below its own quality to reestablish the ability to extract rent through second degree price discrimination. We obtain:

Proposition 3. *Suppose that all heterogeneity between buyers is captured by $\epsilon_{ijk} = \epsilon_{ij}$ and $\gamma_i \in \{\underline{\gamma}, \bar{\gamma}\}$, $\bar{\gamma} > \underline{\gamma} > 0$, let ϕ be the proportion of the buyers with $\gamma_i = \underline{\gamma}$, and assume that the PC operating system monopolist faces a lower quality server product that is competitively supplied. Then there exists $\hat{\phi} > 0$, such that for all $\phi < \hat{\phi}$, the monopolist will set his server operating system price strictly above marginal cost. Furthermore, there exists $\bar{y}_{\hat{k}} > y_M$, such that for all $y_{\hat{k}} \in (y_M, \bar{y}_{\hat{k}})$ the monopolist will degrade the quality of server competitor $\hat{k}$ as long as interoperability can be specifically degraded for $\hat{k}$ only.*

It follows from this proposition that, relative to a model in which there is the option not to buy the server operating system, there will be much less scope for foreclosure in this strong complementarity case. The reason is that it is harder to generate a positive margin on the server because price discrimination is costly. It involves supplying the $\underline{\gamma}$ buyers with a lower quality product in order to extract more from the $\bar{\gamma}$ buyers. The relative margin effect will therefore be smaller. We should therefore expect uniformly lower foreclosure incentives in a model with strong complementarity.

We also estimate a polar case in which customers may both choose to only buy PCs or only a server in addition to the combination of servers and PCs ("free complementarity"). The idea is that there may be firms that keep their workgroups in place but purchase new servers. In this case second degree price discrimination will be *less powerful* because customers who buy a server can now come from both a group with low PC demand elasticity and no demand for PCs at all. We would therefore expect a smaller server margin and a smaller overall relative margin effect in such a model (which is what we find empirically).

3.5. Some Types of Heterogeneity Reduce Foreclosure Incentives. We have already established that foreclosure incentives are not an inevitable outcome. Our model is flexible enough to either find foreclosure incentives or not. However, it is worthwhile to note that the model does not even impose positive server margins. Some types of heterogeneity that we allow for in our model will induce the monopolist to implicitly subsidize server sales through a negative server margin. In that case, foreclosure of a more efficient server rival will not be profitable.

To see how this can arise we consider the strong complementarity model but allow for heterogeneity in the size of the workgroup w. In particular we assume that there are two possible workgroup sizes w and $\hat{w}$. We assume that there is no other heterogeneity. For any (ω, ω_M) define the two per PC system prices $\Omega(w) = \omega + \frac{1}{w}\omega_M$ and $\Omega(\hat{w}) = \omega + \frac{1}{\hat{w}}\omega_M$, where $\Omega(\hat{w}) = \Omega(w) + \omega_M \left[\frac{1}{\hat{w}} - \frac{1}{w}\right]$. The smaller work group systematically pays a higher per unit price because there are increasing returns with respect to the server purchase. Given that the elasticity of demand in a logit model is increasing in the price, this means that the smaller workgroup buyer will have a higher elasticity of demand at these prices as long as ω_M is strictly positive. This means that the relative market share in the PC market is smaller for the buyer with the higher elasticity of demand, implying a negative correlation between elasticity of demand and relative market share in the PC market. By (13) this implies that the server price will be set below the marginal cost of the server.

It is therefore perfectly possible for our model to generate positive or negative margins on the server operating system. Whether we find foreclosure incentives in the server operating system market is therefore entirely an empirical issue. Since the different effects we have discussed will depend, as we have shown, on the degree of complementarity we assume, we have estimated different models of complementarity to explore the robustness of the results.

4. Econometric Modeling Strategy

4.1. Baseline Model. The baseline model of demand follows from theory in the last section and can be empirically implemented in the standard fashion of BLP demand models. In the baseline model of imperfect complementarity we allow customers to select either w PCs or a "workgroup" of w PCs and one server. The specification of the demand system is completed with the introduction of an "outside good". Customers 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 $u_{i00} = \xi_0^{PC} + \xi_0^S + \sigma_0^{PC}\nu_{i0}^{PC} + \sigma_0^S\nu_{i0}^S + \epsilon_{i00}$, 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 terms in ν_{i0} accounts for the outside alternatives' unobserved variance.

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:

$$\delta_k = y_k\gamma_1 + \gamma_2 M + \gamma_3(My_k) - \lambda^S p_k + \xi_k \tag{14}$$

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 equation (8) 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.[16] 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 in equation (14), 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, $\xi = (\xi_j^{PC}, \xi_k^S)$, is obtained through the inversion of the market share equations after aggregating appropriately the individual customer'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 (for a detailed description of the algorithm followed see Appendix C). The weighting matrix in the GMM function was computed using a two-step procedure. To minimize the GMM function we used both

[16] 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.

the Nelder-Mead nonderivative search method and the faster Quasi-Newton gradient method based on an analytic gradient.[17] 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.[18] In our benchmark specification we draw a sample of 150 customers, 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 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

[17] 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.

[18] 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.

and rival models.[19]

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 unobservable tastes for the goods are correlated - that some firms just have a greater taste for "computing power". 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 modeling complementarity. 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,

[19] 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.

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 baseline 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. However, in the robustness section of our results we also present the a freely estimated complementarity/substitutability parameter following Song and Chintagunta (2006) ("free complementarity").

In our baseline model customers 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.

Under "strong complementarity", we write our previous model of market shares as:

$$s_{ij} = e^{\delta_j + \mu_{ij}} \sum_{k=1}^{K} \frac{e^{\delta_k + \mu_{ik}}}{1 + \sum_{j=1}^{J} \sum_{k=1}^{K} e^{\delta_j + \mu_{ij} + \delta_k + \mu_{ik}}} \tag{15}$$

where the outside summation is from 1 to K instead of 0 to K. Similarly for server market share the formula is:

$$s_{ik} = e^{\delta_k + \mu_{ik}} \sum_{j=1}^{J} \frac{e^{\delta_j + \mu_{ij}}}{1 + \sum_{j=1}^{J} \sum_{k=1}^{K} e^{\delta_j + \mu_{ij} + \delta_k + \mu_{ik}}} \tag{16}$$

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 as discussed in the theory sub-section above.

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:

$$u_{ijk} = \delta_j + \mu_{ij} + \delta_k + \mu_{ik} + \Gamma(d^{PC}, d^S) + \epsilon_{ijk} \tag{17}$$

This is identical to the baseline except we have included a term, $\Gamma(d^{PC}, d^S)$ that 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 customers. This utility structure allows us to model complementarity and/or substitution at the level of the good, i.e. PC or server, via $\Gamma(d^{PC}, d^S)$. The key element[20] 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. We identify the $\Gamma^{PC,S}$ parameter in the classical way, by using aggregate time series variation in server prices (in the PC demand equation) and time series variation in PCs prices (in the server demand equation).[21] Further model and estimation

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

[21] Song and Chintagunta (2006) also build on Gentzkow to allow for a common complementarity/substitution

details are given in Appendix E.

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.[22] 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 in the US market to match each observation with more detailed product characteristics.[23]

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.[24]

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,

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.

[22] Various datasets from IDC have been used both in the literature (Foncel and Ivaldi, 2005; Van Reenen, 2006; Pakes, 2003; Genakos, 2004)

[23] 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.

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

OS390/400 and a residual category). For servers key characteristics are also RAM, the number of rack slots,[25] 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 A1 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 A2 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 optimized, 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,

[25] 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.

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).[26]

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.[27]

6. Results

6.1. Main Results. We first turn to the demand estimates from a simple logit model and the full baseline random coefficients 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 the price and to test the different sets of instrumental variables discussed in the previous section for each product category separately. Table 1 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

[26] 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).

[27] Even Linux's 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).

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 a key performance metric, RAM, has a negative and significant coefficient, although the other quality measure, “benchmark”, has the expected positive and significant coefficient. Furthermore, the final row shows that the vast majority of products (85.5%) are predicted to have inelastic demands, which is clearly unsatisfactory.

Column (2) of Table 1 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.[28] 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 A3) 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 overidentification restrictions does reject, however, a common problem in this literature. In the last two columns we restrict the number of 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.

[28] 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.

Table 2 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 (a statistically acceptable restriction).[29] 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 server price falls dramatically moving from OLS to IV (e.g. from -0.040 in column (3) to -0.179 in column (4)).

Columns (4)-(6) of Table 2 experiment with different instrument sets (first stages are reported in full in Table A4). Empirically, the most powerful set of instruments were the number of models by the firm, the number of models produced by rivals firms and the sum of RAM by rivals (used in columns (2) and (6)). We use these instruments in all columns and also include the official series for quality-adjusted prices for semi-conductors and for hard-disks (two key inputs for servers) in columns (4) and (5). In addition, column (5) includes sums of rivals' characteristics (rack-optimized servers, numbers of racks and use of Unix). Although the parameter estimates are reasonably stable across the experiments, the F-test of excluded instruments indicates that the parsimonious IV set of column (6) is preferred, with a F-statistic of 12.9. In these preferred estimates 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 falls 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.[30]

Results from the baseline random coefficients model are reported in column (1) of Table 3.

[29] We also estimated models with higher order polynomials in time with qualitatively similar results.

[30] We also estimated models allowing other server characteristics to interact with the Microsoft dummy. These produced similar evidence that these characteristics were less highly valued when used with a non-Microsoft server. The other interactions were not significant, however, so we use the RAM interaction as our preferred specification.

The first two panels (A and B) report the mean coefficients for PCs and servers respectively. Almost all mean coefficients are significant and have the expected sign. The lower rows (C and D) report the results for the random coefficients. We allow random coefficients only on price and one other basic characteristic in our baseline specification - performance benchmark for PCs and RAM for servers.[31] Our results indicate that there is significant heterogeneity in the price coefficient, but not in the other characteristics (although the random 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. The implied hardware margins from the baseline model seem realistic for both PCs and servers. Assuming multi-product firms and Nash-Bertrand competition in prices for hardware firms, our derived median margin for the whole period is 16% for PCs and 34% for servers.[32]

Figure 2 plots the calculated relative output and margin effects based on these coefficients (together with the 90% confidence interval).[33] Server operating system margins are higher than PC operating system margins (as indicated by relative margins well in excess of unity) 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). The shaded area indicates where we estimate that Microsoft has significant incentives to degrade interoperability.

[31] We also estimated models allowing a random coefficient on the interaction of RAM with Microsoft. This was insignificant and the implied overall effects were similar so we keep to the simpler formulation here.

[32] Industry reports at the time put the gross profit margins of the top PC manufacturers in the range of 10%-20% and for server vendors in the range of 25%-54% (see International, Data Corporation, 1999a,b, 2000). This is also in line with other results in the literature. For example, Goeree (2008) using a different quarterly US data between 1996-1998 reports a median margin of 19% for PCs from her preferred model.

[33] Figure A1 in the Appendix plots the calculated relative output and margin effects together with the 95% confidence interval.

Three key findings stand out. First, looking at the period as a whole the relative margin effect exceeds the output effect from the end of 1996 onwards indicating incentives to degrade interoperability. Second, the two effects trends in opposite directions with relative output decreasing and the relative margin steadily increasing. By the end of our sample period in 2001, the difference between the two effects takes its largest value with relative margin clearly dominating the relative output effect. Third, the key point when the two lines diverge is around the end of 1999 and beginning of 2000. These dates coincide 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 interoperability limitations that were much more severe than any previous version on Windows (European Commission, 2004). As we will show later these three findings 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 relative margin effect 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. The relative interoperability effect is declining primarily because the aggregate number of servers sold was rising faster than the number of PCs which is related to the growth of Internet and the move away from mainframes to client-server computing (see Bresnahan and Greenstone, 1999). Thus, a marginal change in interoperabilty had a smaller effect on loss of PC quantity (relative to the gain in servers) in 2001 than in 1996.

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. Under strong complementarity customers can only buy the PC-server bundle or the outside good, i.e. they cannot purchase a standalone PC as in the baseline model of the previous sub-section. Column (2) of Table 3 presents the simplest version of strong complementarity where we assume a random coefficients on prices and quality 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. Columns (3) and (4) of Table 3 add progressively more random coefficients. The estimated mean coefficients retain their magnitude and significance and again there appears to be significant heterogeneity for the PC price and characteristics (column (4) also suggests some heterogeneity on the constant for servers).

Figure 3 plots the calculated relative margin and output effects and confidence interval based on the estimated results from column (4) of Table 3. Consistent with the theory (see Proposition 3) the relative margin effect is smaller than in the baseline case. After mid 1998 however, significant interoperability incentives still exist as the relative output effect remains low.

Free Complementarity. Our most general model is presented in the last column of Table 3 where we allow customers to purchase standalone servers (as well as standalone PCs, bundles of PCs and servers or the outside good) and let 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 PCs and servers 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 4 plots the relative output and margin effects and their confidence interval. Again, the relative margin value is somewhat lower than under the baseline model of column (1) which is consistent with proposition 3. Nevertheless, we still find incentives to degrade interoperability towards the end of the sample period. Given that this is a much more demanding specification, the consistency of results with our baseline case is reassuring.[34]

7. Robustness

Table 4 reports various robustness tests of the baseline model (reproduced in column (1) and in Figure 5A to ease comparisons) to gauge the sensitivity of the results to changes in our assumptions. We show that the basic qualitative result that there were incentives to degrade interoperability is robust. First, we vary the number of random draws following the Monte Carlo evidence from Berry, Linton and Pakes (2004) for the BLP model. In column (2) we increase the number of draws to 500 (from 150 in the baseline model). The estimated results are very similar to our baseline specification, the only exception being that the PC benchmark now has a significant random coefficient. Unsurprisingly the calculated relative output and margin effects in Figure 5B exhibit the same pattern as in Figure 5A.

In column (3) and (4) of Table 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 5C and 5D are similar.

In columns (5) and (6) of Table 4 we reduce the number of instruments used for both the PCs and servers. On the one hand, using the most powerful instruments increases the

[34] 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 (causing 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.

absolute value of the coefficients. For example, the mean coefficient on PC price increases from -3.301 in the 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 5E and 5F reveal a qualitative similar picture as before.

In the final two columns of Table 4 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 random coefficient on server price. As before, both the estimated coefficients and calculated effects in Figures 5G and 5H 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.

As a final robustness test we consider an alternative approach to estimating the relative output effect which considers only the "reduced form" residual demand equations for Microsoft servers and PCs. Since these will be a function of non-Microsoft quality (and other variables), we can use the coefficients on these to calculate the output effects of degradation directly. Appendix F gives the details and shows that the relative margins continue to lie far above the this simpler, less structural approach to estimating the relative output effect. So again, this suggests strong incentives for Microsoft 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 a structural econometric method based upon this theory.

In our model, 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 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 with random coefficients and allowing for complementarity. According to our results it seemed that Microsoft had incentives to decrease interoperability at the turn of the 21st century as alleged by competition authorities. 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 (see for example Lee, 2010). 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, 2010, 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] Bernheim, D. and Whinston, M. (1998) "Exclusive Dealing", *Journal of Political Economy*, 106(1), 64-103.

[6] 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.

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

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

[9] 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.

[10] Bloom, N., Draca, M. and Van Reenen, J. (2010) "Trade Induced Technical Change", LSE/Stanford *mimeo.*

[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] Bresnahan, T. and Greenstein, S. (1999) "Technological Competition and the Structure of the Computer Industry" *Journal of Industrial Economics*, pp. 1-40.

[15] 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.

[16] Davis, P. and Huse. C. (2009) "Estimating the Co-ordinated Effects of Mergers", *mimeo*, Stockholm School of Economics.

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

[18] Farrell, J. and Katz (2000) "Innovation, Rent Extraction and Integration in Systems" *Journal of Industrial Economics*, 48(4), 413-432

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

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

[21] Gentzkow, M. (2007) "Valuing New Goods in a Model with Complementarity: Online Newspapers", *American Economic Review*, 97(3), 713-744.

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

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

[24] 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.

[25] Hendel, Igal (1999) "Estimating Multiple Discrete Choice Models: an Application to Computerization Returns", *Review of Economic Studies*, 66(2), 423-446.

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

[27] International Data Corporation (1999a) *Server Operating Environment and Platform Report.*

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

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

[30] Lee, R. (2010) "Dynamic Demand Estimation in Platform and Two Sided Markets", New York University *mimeo.*

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

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

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

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

[35] Nevo, A. (2003) "New Products, Quality Changes and Welfare Measures Computed from Estimated Demand Systems" *Review of Economics and Statistics,* 85(2), 266-275.

[36] Oyer, P (1998) "Fiscal Year Ends and Nonlinear Incentive Contracts: The Effect on Business Seasonality", *Quarterly Journal of Economics*, 113(1), 149-185.

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

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

[39] 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.

[40] Rey, P. and Tirole, J. (2007) "A Primer on Foreclosure", *Handbook of Industrial Organization Volume III,* Armstrong, M. and Porter, R. (eds) 2145-2220.

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

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

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

[44] Shaked, A. and Sutton, J. (1992) "Relaxing price competition through product differentiation" *Review of Economic Studies*, 49 (1), 3-13.

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

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

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

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

[49] 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.

Not for Publication, Referee use only Appendices

A. Proofs

A.1. Deriving market shares. To derive the market shares for an individual i we follow BLP and allow random coefficients on the parameter vector $\theta_i = (\beta_i, \gamma_i, \lambda_i)$ as well as heterogeneity in the size of work groups w_i. The latter can be captured by a random coefficient on the server price $\lambda_i^S \equiv \lambda_i / w_i$. We derive demand from the above utility function in the standard way.

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 as:

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

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:

$$s_{jk}(w) = \int_{B_{jk}(x_j, y_k, p_j, p_k, a, w)} dP(\theta, \xi_j, \xi_k, \epsilon_{ijk} \mid w) \tag{18}$$

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 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 $s_k = \sum_{j=0}^{J} s_{jk}$.[35] 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 $q^S = L \int \sum_{k=1}^{K} s_k(w) d\Upsilon(w)$. We will assume in everything that follows that ϵ_{ijk} comes from a double exponential distribution, so that conditional on θ_i, $s_{jk}(\theta_i)$ has the familiar logit form.

$$s_{ij} = e^{\delta_j + \mu_{ij}} \sum_{k=0}^{K} \frac{e^{\delta_k + \mu_{ik}}}{1 + \sum_{j=1}^{J} \sum_{k=0}^{K} e^{\delta_j + \mu_{ij} + \delta_k + \mu_{ik}}} \tag{19}$$

[35] 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. This possibility is allowed in some of our estimations and not in others.

For servers this is

$$s_{ik} = e^{\delta_k + \mu_{ik}} \sum_{j=1}^{J} \frac{e^{\delta_j + \mu_{ij}}}{1 + \sum_{j=1}^{J} \sum_{k=0}^{K} e^{\delta_j + \mu_{ij} + \delta_k + \mu_{ik}}} \tag{20}$$

Own price elasticity for PC operating system

$$\varepsilon_{\omega} = -\int \frac{q(\theta_i)}{q} w \lambda s_{00}(\theta_i) dP(\theta) \tag{21}$$

Own price elasticity for monopolist's server operating system

$$\varepsilon_{\omega_M}^{M} = -\int \frac{q_M(\theta_i)}{q_M} \lambda [1 - s_M(\theta_i)] dP(\theta) \tag{22}$$

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

$$\varepsilon_{\omega_M} = -\frac{q_M}{q} \int \frac{q_M(\theta_i)}{q_M} w \lambda s_{00}(\theta_i) dP(\theta) \tag{23}$$

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

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

Derivation of individual specific elasticities

$$\begin{aligned} \varepsilon_{\omega}(\theta_i) &= \frac{1}{q(\theta_i)} w L(\theta_i) \frac{\partial \sum_{j=1}^{J} \sum_{k=0}^{K} s_{jk}(\theta_i)}{\partial \omega} \\ &= \frac{1}{q(\theta_i)} w L(\theta_i) \frac{\partial}{\partial \omega} \left[\frac{\sum_{j=1}^{J} \sum_{k=0}^{K} e^{\delta_j + \delta_k}}{1 + \sum_{j=1}^{J} \sum_{k=0}^{K} e^{\delta_j + \delta_k}} \right] \\ &= -\lambda s_{00}(\theta_i) \end{aligned} \tag{25}$$

and

$$\varepsilon_{\omega_M}(\theta_i) = \frac{1}{q(\theta_i)} w L(\theta_i) \frac{\partial \sum_{j=1}^{J} \sum_{k=0}^{K} s_{jk}(\theta_i)}{\partial \omega_M} \tag{26}$$

$$= -w \lambda s_{00}(\theta_i) \frac{q_M(\theta_i)}{q(\theta_i)} \tag{27}$$

$$\begin{aligned}
\varepsilon_{\omega}^{M}(\theta_i) &= \frac{1}{q_M(\theta_i)} M(\theta_i) \frac{\partial \sum_{j=1}^{J} \sum_{k \in M} s_{jk}(\theta_i)}{\partial \omega_M} \\
&= -w\lambda s_{00}(\theta_i) \qquad (28)
\end{aligned}$$

$$\begin{aligned}
\varepsilon_{\omega_M}^{M}(\theta_i) &= \frac{1}{q_M(\theta_i)} L(\theta_i) \frac{\partial \sum_{j=1}^{J} \sum_{k \in M} s_{jk}(\theta_i)}{\partial \omega_M} \\
&= -\lambda \sum_{k \notin M} s_k(\theta_i) \qquad (29)
\end{aligned}$$

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

$$\begin{aligned}
\varepsilon_{\omega} &= \int \frac{q(\theta_i)}{q} \varepsilon_{\omega}(\theta_i) dP(\theta) \\
&= -\int \frac{q(\theta_i)}{q} w\lambda s_{00}(\theta_i) dP(\theta) \qquad (30)
\end{aligned}$$

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

$$\begin{aligned}
\varepsilon_{\omega}^{M} &= \int \frac{q(\theta_i)}{q} \varepsilon_{\omega}^{M}(\theta_i) dP(\theta) \\
&= -\int \frac{q(\theta_i)}{q} w\lambda s_{00}(\theta_i) dP(\theta) \qquad (32)
\end{aligned}$$

$$\begin{aligned}
\varepsilon_{\omega_M}^{M} &= \int \frac{q_M(\theta_i)}{q_M} \varepsilon_{\omega_M}^{M}(\theta_i) dP(\theta) \\
&= -\int \frac{q_M(\theta_i)}{q_M} \lambda [1 - s_M(\theta_i)] dP(\theta) \qquad (33)
\end{aligned}$$

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

$$\begin{aligned}
\frac{q}{q_M}\varepsilon_{\omega_M} - \varepsilon_\omega &= \int \left[\frac{q(\theta_i)}{q} - \frac{q_M(\theta_i)}{q_M}\right][w\alpha s_{00}(\theta_i)]\,dP(\theta) \\
&= -\int \left[\frac{q(\theta_i)}{q} - \frac{q_M(\theta_i)}{q_M}\right][\varepsilon_\omega(\theta_i)]\,dP(\theta) \\
&= \int \left[\frac{q(\theta_i)}{q} - \frac{q_M(\theta_i)}{q_M}\right][\bar{\varepsilon}_\omega - \varepsilon_\omega(\theta_i)]\,dP(\theta)
\end{aligned} \tag{34}$$

where the last equality comes from subtracting

$-\int \left[\frac{q(\theta_i)}{q} - \frac{q_M(\theta_i)}{q_M}\right]\bar{\varepsilon}dP(\theta) = 0$ from the second line, where $\bar{\varepsilon} = \left(\int \varepsilon_\omega(\theta_i)dP(\theta)\right)$.

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

$$\begin{aligned}
\frac{q_M}{q}\varepsilon_\omega^M - \varepsilon_{\omega_M}^M &= \int \alpha w s_{00}(\theta_i)\left(\frac{M(\theta_i) - q_M(\theta_i)}{wM(\theta_i) - q(\theta_i)}\frac{q_M(\theta_i)}{q_M} + \frac{q_M(\theta_i)}{q(\theta_i)}\frac{q(\theta_i)}{q}\right)dP(\theta) \\
&\quad -\frac{q_M}{q}\int \alpha w s_{00}(\theta_i)\left[\frac{q(\theta_i)}{q} - \frac{q_M(\theta_i)}{q_M}\right]dP(\theta)
\end{aligned} \tag{35}$$

Proof of Proposition 1: Note first that $\varepsilon_\omega(\theta, w) = \bar{\varepsilon}_\omega$ when there is no heterogeneity. It follows that the expression in (13) is zero and the price cost margin on the server operating system must be zero for any set of PCs and servers offered in the market. But then by (2) the impact of an increase in the quality of a rival server operating system is $(\omega - c)\frac{dq(\mathbf{p}_j,\mathbf{p}_k,a)}{da}\mid_{\omega,\omega_M} > 0$. The inequality is strict since there will be some buyers who substitute from buying no workgroup at all to buying a workgroup with the server operating system of the rival when the server operating system quality of the rival is increased. Hence, a quality increase in a rival server operating system can only increase the profits of the PC operating system monopolist. QED.

For proposition 2 we specialize the model to having one brand of PC and one brand of server, each with an operating system provided by the monopolist. We also assume that (beyond differences in γ), there is no heterogeneity in the marginal valuation of a server product, i.e. $\epsilon_{ijk} = \epsilon_{ij0}$ for all k.

Proof of Proposition 2: We first show that $\frac{q(\gamma)}{q} - \frac{q_M(\gamma)}{q^M}$ is strictly decreasing in γ and that $-\varepsilon_\omega(\gamma)$ is also strictly decreasing. This establishes that in the absence of a rival server product $\omega_M > c_M$. We then show that if there is a firm with a higher (but not too high) server quality in the market, the monopolist will always want to foreclose it. First note that:

$$\begin{aligned}\frac{1}{y_M}\left[\frac{\partial q(\gamma)/\partial\gamma}{q}-\frac{\partial q_M(\gamma)/\partial\gamma}{q_M}\right] &= \frac{wq_M(\gamma)-s_M(\gamma)q(\gamma)}{q}-\frac{q_M(\gamma)(1-s_M(\gamma))}{q_M} \\ &< \frac{q_M(\gamma)-s_M(\gamma)\frac{q(\gamma)}{w}}{q_M}-\frac{q_M(\gamma)(1-s_M(\gamma))}{q_M} \\ &= \frac{q_M(\gamma)(1-s_M(\gamma))-s_M(\gamma)(\frac{q(\gamma)}{w}-q_M(\gamma))-q_M(\gamma)(1-s_M(\gamma))}{q^M} \\ &= \frac{-s_M(\gamma)(\frac{q(\gamma)}{w}-q_M(\gamma))}{q_M}<0\end{aligned} \tag{36}$$

where the first inequality follows from the fact that $q/w > q_M$, since all buyers of servers buy w PCs, but there are some buyers of PCs who do not buy servers. The next equality follows by expanding $q_M(\gamma) - s_M(\gamma)\frac{q(\gamma)}{w}$ and the last inequality follows because $(\frac{q(\gamma)}{w} - q_M(\gamma)) > 0$ by the same argument we used for $q/w > q_M$. Now note that

$$-\varepsilon_\omega(\gamma) = \lambda s_{00} = \lambda\frac{1}{1+e^{\beta x_j-\lambda p_j}+e^{\beta x_j-\lambda p_j+\gamma y_M-\frac{\lambda}{w}p_M}},$$

which is decreasing in γ. It follows that $\frac{q(\gamma)}{q} - \frac{q_M(\gamma)}{q_M}$ and $-\varepsilon_\omega(\gamma)$ move in the same direction in γ, which implies by (13) that $\omega_M > c_M$. Now consider a server rival k entering the market. We only consider equilibria in which firms set prices no lower than marginal cost. First consider the case $ay_k < y_M$. Suppose for contradiction that $0 < \gamma ay_k - \frac{\lambda}{w}p_k < \gamma y_M - \frac{\lambda}{w}p_M$. Then the buyer only purchases the server M and locally the offering of rival k does not matter for the price incentives of the monopolist, so that M is either at a monopoly optimum. In that case k would have an incentive to set p_k below $-\frac{w}{\lambda}\gamma[y_M - ay_k] + p_M$ as long as this is strictly positive. Otherwise the quality is too low to affect the monopolist. If (p_j, p_M) are not optimal in the monopoly case, then there is an incentive to slightly change p_M and p_j, so that this would not be an equilibrium either. Now suppose $\gamma ay_k - \frac{\lambda}{w}p_k > \gamma y_M - \frac{\lambda}{w}p_M \geq 0$. Then there exists ε such that the monopolist can undercut to $p'_M = \frac{w}{\lambda}[\gamma y_M - \gamma ay_k] + p_k - \varepsilon > 0$. At this price the surplus gained from the server purchase is virtually unchanged for any buyer purchasing the server, so that PC demand is unchanged. But now the monopolist gains a margin of $\gamma y_M - \gamma ay_k + p_k - \varepsilon$ on the server sales as well. Hence, there is a profitable deviation as long as $p_k > -w(\gamma y_M - \gamma ay_k)$. Hence, in equilibrium, $p_M = w[\gamma y_M - \gamma ay_k]$.It follows that decreasing a strictly increases $p_M = \gamma w[y_M - ay_k]$ as long as the quality of the new producer is high enough to impose a competitive constraint. Now consider the case $ay_k > y_M$. If $p_k > w\gamma[ay_k - y_M] + p_M$ then k makes no sale but can always make a sale by charging $w\gamma[ay_k - y_M] + p_M - \varepsilon$. This will be profitable if $p_M > -w\gamma(ay_k - y_M)$. Hence, in equilibrium $p_k \leq w\gamma[ay_k - y_M]$ and $p_M = 0$.

Now let ϕ be the proportion of buyers with $\gamma = 0$ and let p_j^* be the non-discriminatory price that maximizes $(p_j - c)(\phi q(0) + (1-\phi)q(\gamma))$. Suppose that the monopolist sets $p_j =$

$p_j^*$and $p_M = 0$. Then the best response of k is to set k as to solve:

$$\max_{p_k \in [0, \frac{w}{\lambda}\gamma[y_M - ay_k]]} p_k q(\gamma, p_k, p_j^*)$$

Since $p_k \leq \frac{w}{\lambda}\gamma[ay_k - y_M]$, there exists $\bar{y}_k > y_M$ such that for all $y_k < \bar{y}_k$ and a, $q + p_k \frac{\partial q}{\partial p_k} > 0$. Hence for quality differences that are not too large firm k will set the maximal feasible price. This implies that the net benefit of a buyer from the server is given by $\frac{w}{\lambda}\gamma y_M$. Hence, setting p_j^* is a best response and the profits for the monopolist are the same as if setting a single price. Since the best response functions are contractions on the relevant domain, this equilibrium is unique. We have earlier shown that the monopolist makes higher profits by price discriminating (namely setting a positive margin on the server), it follows that profits are reduced in this equilibrium. Note that slightly reducing quality of k does not lead to higher profits. only when $ay_k < y_M$ do profits increase. It follows that $a = 0$ is the optimal choice of interoperability if $y_k < \bar{y}_k$.

Proof of Proposition 3: We assume that the competitive server product will be offered at a price of zero. The buyer will therefore get a rent from buying the competitive server product equal to $\gamma_i y_k$. Since $y_M > y_k$, the PC operating system monopolist can, at given price p_j, charge any server price below $\frac{w}{\lambda}\underline{\gamma}[y_M - y_k]$ and sell to all buyers who purchase a server. Note that whenever the server price is below this benchmark, the buyers only care about the total system price $p_j + \frac{p_M}{w}$. Any candidate equilibrium in which the monopolist makes all the server sales can therefore be induced setting $p_M = \frac{w}{\lambda}\underline{\gamma}[y_M - y_k]$. Let p_j^* be the price of the PCs in such a candidate equilibrium. It solves the first order condition

$$\begin{array}{c}(p_j + \frac{w}{\lambda}\underline{\gamma}[y_M - y_k] - c - c_M)\phi\frac{\partial D_{\underline{\gamma}}}{\partial p_j} + \phi D_{\underline{\gamma}} \\ +(p_j + \frac{w}{\lambda}\underline{\gamma}[y_M - y_k] - c - c_M)(1-\phi)\frac{\partial D_{\bar{\gamma}}}{\partial p_j} + (1-\phi)D_{\bar{\gamma}} = 0\end{array}$$

Since a $\bar{\gamma}$ customers have a rent that exceeds that of $\underline{\gamma}$ customers by $(\bar{\gamma} - \underline{\gamma})y_k$, it follows that $-\varepsilon_\omega(\underline{\gamma}) > -\varepsilon_\omega(\bar{\gamma})$. And hence,

$$(p_j + \frac{w}{\lambda}\underline{\gamma}[y_M - y_k] - c - c_M)(1-\phi)\frac{\partial D_{\bar{\gamma}}}{\partial p_j} + (1-\phi)D_{\bar{\gamma}} > 0.$$

The monopolist could therefore gain profits on $\bar{\gamma}$ customers by increasing the price above $p_M = \frac{w}{\lambda}\underline{\gamma}[y_M - ay_k]$. Let $p_M^{**}(p_j) \in (\frac{w}{\lambda}\underline{\gamma}[y_M - y_k], \frac{w}{\lambda}\bar{\gamma}[y_M - y_k]]$ be the price that maximizes $\pi_{\bar{\gamma}}(p_j, p_M) = (p_j + p_M - c - c_M)D_{\bar{\gamma}}(p_j, p_M)$ for given p_j. Then by setting the price to $p_M^{**}(p_j^*)$ the monopolist gains from $\bar{\gamma}$ types and looses margin on $\underline{\gamma}$ types for a net effect of:

$$(1-\phi)\int_{\frac{w}{\lambda}\underline{\gamma}[y_M - ay_k]}^{p_M^{**}(p_j^*)} \frac{\partial \pi_{\bar{\gamma}}(p_j, \xi)}{\partial p_M} d\xi - (\frac{w}{\lambda}\underline{\gamma}[y_M - y_k] - c_M)\phi D_{\underline{\gamma}} \tag{37}$$

Since $p_M^{**}(p_j^*) > \frac{w}{\lambda}\underline{\gamma}[y_M - y_k]$ for all ϕ, it follows that (37) strictly exceeds zero for ϕ small enough. Since this also holds for lower p_j, the monopolist will in equilibrium leave $\underline{\gamma}$ types to buy lower quality servers in order to better price discriminate between $\underline{\gamma}$ and $\bar{\gamma}$ types.

Now introduce a server competitor with $y_{\hat{k}} > y_M$. Note that the monopolist would not be able to make a server sale in equilibrium at any price p_M above $\frac{w}{\lambda}\underline{\gamma}[y_M - y_k]$ because $\underline{\gamma}$ consumers would not buy and firm $\hat{k}$ would sufficiently reduce the price to make all the sales since it has higher quality. But then the rent from the server purchase of a $\bar{\gamma}$ producer is strictly larger than ay_k. This implies that an increase in p_j would increase profits for the monopolist from $\bar{\gamma}$ consumers. Hence, the monopolist can reduce p_M by dp_M and increase p_j by dp_j, leave the profits on the $\underline{\gamma}$ consumers unchanged, and increase profits on the $\bar{\gamma}$ consumers. Hence, $p_M = 0$ in equilibrium and $p_{\hat{k}} \leq \frac{w}{\lambda}\bar{\gamma}[y_{\hat{k}} - y_M]$, leaving a rent of at least $\frac{w}{\lambda}\bar{\gamma}y_M$. The optimal p_j that maximizes profits when the M server is offered at 0 prices and a uniform price is set. Call this price $\hat{p}_j$. This must be equal to $p_j^* + \frac{w}{\lambda}\underline{\gamma}[y_M - y_k]$. As in the proof of proposition 2 it will now be true that there exists $\bar{y}_{\hat{k}}$ such that $\hat{p}_k = \frac{w}{\lambda}\bar{\gamma}[y_{\hat{k}} - y_M]$ for all $y_{\hat{k}} \in (y_M, y_{\hat{k}})$ if $\hat{p}_j$ is charged. Furthermore, the net rent left from the server at the price $\hat{p}_k$ is the same as the rent if the product $\hat{k}$ were not around. It follows by the same arguments as those of proposition 2 that the monopolist will want to fully exclude all firms $\hat{k}$ with quality levels $y_{\hat{k}} \in (y_M, \bar{y}_{\hat{k}})$.

B. Data Appendix

As noted in the Data section, quarterly data on quantities and prices[36] 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,[37] 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.[38] 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

[36]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.

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

[38]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 similar. 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. These server operating systems are divided into six non-Windows categories: Netware, Unix, Linux, VMS, OS390/400 and a residual. 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). 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 very rarely purchased by customers 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. Estimation Algorithm Details

In this section we describe in detail the algorithm followed for the estimation of the baseline model.

Define $\widetilde{\theta} \equiv (\sigma^{PC}, \sigma^{S}, \sigma_p^{PC}, \sigma_p^{S})$, 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 $\delta = (\delta_j, \delta_k)$, $\xi = (\xi_j^{PC}, \xi_k^{S})$, $\nu_i = \left(\nu_i^{PC}, \nu_i^{S}\right)$ 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 $\widetilde{\theta}$.

[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.

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

Step 2. Given (δ, μ_i), calculate individual customer 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 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.

D. 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.

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

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

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

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. Finally, remember that $p_j = \hat{p}_j + \omega$ and for servers $p_k = \hat{p}_k + \omega_k$.

In order to calculate the relative output effect, note that this derivative contains the direct effect of interoperability on demand as well as the impact of the price responses to a change in interoperability by all rival software producers and all hardware producers. In other words, the nominator of the relative output effect is given by:

$$\begin{aligned} \frac{dq(\mathbf{p}_j, \mathbf{p}_k, a)}{da} &= \sum_{j=1}^{J} \frac{ds_j}{da} L_{PC} = \\ &= \sum_{j=1}^{J} \left(\frac{\partial s_j}{\partial a} + \frac{\partial s_j}{\partial \hat{p}_j}\frac{\partial \hat{p}_j}{\partial a} + \frac{\partial s_j}{\partial \hat{p}_k}\frac{\partial \hat{p}_k}{\partial a} + \frac{\partial s_j}{\partial \omega_k}\frac{\partial \omega_k}{\partial a} \right) L_{PC} \end{aligned} \tag{38}$$

where the first term inside the parenthesis is the direct effect, the second term is the indirect PC hardware effect, the third term is the indirect server hardware effect and the last one is the indirect server non-Microsoft software effect.

Similarly the denominator of the relative output effect is given by:

$$\begin{aligned} \frac{dq_M(\mathbf{p}_j, \mathbf{p}_k, a)}{da} &= \sum_{k=1}^{K} \frac{ds_k}{da} L_S = \\ &= \sum_{k=1}^{K} \left(\frac{\partial s_k}{\partial a} + \frac{\partial s_k}{\partial \hat{p}_j}\frac{\partial \hat{p}_j}{\partial a} + \frac{\partial s_k}{\partial \hat{p}_k}\frac{\partial \hat{p}_k}{\partial a} + \frac{\partial s_k}{\partial \omega_k}\frac{\partial \omega_k}{\partial a} \right) L_S \end{aligned} \tag{39}$$

Each term inside the parenthesis in (38) is given below:

$$\frac{\partial s_j}{\partial a} = \frac{1}{ns}\sum_{i=1}^{ns} \frac{\partial s_{ij}}{\partial a} = \frac{1}{ns}\sum_{i=1}^{ns} s_{ij} \frac{\sum_{k=1}^{K} \widetilde{\gamma_k} e^{V_{ik}}}{(1+W_{ik})\left[1+W_{ij}(1+W_{ik})\right]}$$

where $\widetilde{\gamma_k} = \gamma\left(MEM_k * (1-M_k)\right)$.

$$\begin{aligned}
\frac{\partial s_j}{\partial \hat{p}_j} &= \frac{1}{ns}\sum\nolimits_{i=1}^{ns}\frac{\partial s_{ij}}{\partial \hat{p}_j} = \\
&= \frac{1}{ns}\sum\nolimits_{i=1}^{ns} s_{ij}(1-s_{ij})(\lambda^{PC}+\sigma_p^{PC}\nu_{ip}^{PC}) \\
\frac{\partial s_j}{\partial \hat{p}_k} &= \frac{1}{ns}\sum\nolimits_{i=1}^{ns}\frac{\partial s_{ij}}{\partial \hat{p}_k} = \\
&= \frac{1}{ns}\sum\nolimits_{i=1}^{ns} s_{ij}(\lambda^{PC}+\sigma_p^{PC}\nu_{ip}^{PC})\frac{e^{V_{ik}}}{(1+W_{ik})\,(1+W_{ij}+W_{ij}W_{ik})} \\
\frac{\partial s_j}{\partial \omega_k} &= \frac{1}{ns}\sum\nolimits_{i=1}^{ns}\frac{\partial s_{ij}}{\partial \omega_k} = \\
&= \frac{1}{ns}\sum\nolimits_{i=1}^{ns} s_{ij}(\lambda^{PC}+\sigma_p^{PC}\nu_{ip}^{PC})\frac{e^{V_{ik}}}{(1+W_{ik})\,(1+W_{ij}+W_{ij}W_{ik})}
\end{aligned}$$

Finally, we calculate the derivatives of prices w.r.t. a numerically using the pricing function of the supply side for hardware (PC and server) and non-Microsoft software producers assuming a Nash-Bertrand equilibrium.

Specifically, assume that each of the F multiproduct PC hardware firms has a portfolio, Γ_f, of the $j = 1, ..., J$ different products in the PC market. Then the profit function of firm f can be expressed as

$$\Pi_f = \sum_{j\in\Gamma_f} (p_j - mc_j)\, M s_j(p),$$

where $s_j(p)$ is the predicted market share of brand j, which depends on the prices of all other brands, M is the market size and mc_j is the constant marginal cost of production. Assuming that there exists a pure-strategy Bertrand-Nash equilibrium in prices and that all prices that support it are strictly positive, then the price p_j of any product produced by firm f must satisfy the first-order condition

$$s_j(p) + \sum_{r\in\Gamma_f} (p_r - mc_r)\frac{\partial s_r(p)}{\partial p_j} = 0$$

This system of J equations can be inverted to solve for the marginal costs. Define,

$$\Delta_{jr} = \begin{cases} -\partial s_j(p)/\partial p_r, & \text{if } j \text{ and } r \text{ are produced by the same firm } (j, r = 1, ..., J), \\ 0, & \text{otherwise,} \end{cases}$$

then we can write the above FOC in vector notation as:

$$\begin{aligned} s(p) - \Delta(p)(p - mc) &= 0 \\ p &= mc + \Delta(p)^{-1} s(p). \end{aligned}$$

Given our demand estimates, we calculate the estimated markup. We then compute numerically the derivatives $\partial \hat{p}_j / \partial a$ using "Richardson's extrapolation". We follow the same methodology to calculate the derivatives $\partial \hat{p}_k / \partial a$ and $\partial \omega_k / \partial a$.

To calculate the relative margin

$$\frac{\omega_M - c_M}{\omega - c} = \frac{\frac{q}{q_M}\varepsilon_{\omega_M} - \varepsilon_\omega}{\frac{q_M}{q}\varepsilon_\omega^M - \varepsilon_{\omega_M}^M} = \frac{\frac{q}{q_M}(\sum_{j=1}^J \sum_{k=1}^K \frac{\partial s_j}{\partial p_k}\frac{1}{s_j}) - (\sum_{j=1}^J \sum_{j=1}^J \frac{\partial s_j}{\partial p_j}\frac{1}{s_j})}{\frac{q_M}{q}(\sum_{k=1}^K \sum_{j=1}^J \frac{\partial s_k}{\partial p_j}\frac{1}{s_k}) - (\sum_{k=1}^K \sum_{k=1}^K \frac{\partial s_k}{\partial p_k}\frac{1}{s_k})}$$

where the derivatives for the PCs are:

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

Similarly, the derivatives for the servers are:

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

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 $\widetilde{\theta} \equiv (\sigma, \sigma^S, \sigma_p, \sigma_p^S)$:

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

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 as follows. The derivatives for the PCs w.r.t the mean valuations are:

$$\begin{aligned} \frac{\partial s_j}{\partial \delta_j} &= \frac{1}{ns} \sum_{i=1}^{ns} \frac{\partial s_{ij}}{\partial \delta_j} = \frac{1}{ns} \sum_{i=1}^{ns} s_{ij}(1 - s_{ij}) \\ \frac{\partial s_j}{\partial \delta_d} &= \frac{1}{ns} \sum_{i=1}^{ns} \frac{\partial s_{ij}}{\partial \delta_d} = -\frac{1}{ns} \sum_{i=1}^{ns} s_{ij} s_{id} \\ \frac{\partial s_j}{\partial \delta_k} &= \frac{1}{ns} \sum_{i=1}^{ns} \frac{\partial s_{ij}}{\partial \delta_k} = \frac{1}{ns} \sum_{i=1}^{ns} s_{ij} \frac{e^{V_{ik}}}{(1 + W_{ik})(1 + W_{ij} + W_{ij} W_{ik})} \end{aligned}$$

The derivatives for the servers w.r.t the mean valuations are:

$$\begin{aligned} \frac{\partial s_k}{\partial \delta_k} &= \frac{1}{ns} \sum_{i=1}^{ns} \frac{\partial s_{ik}}{\partial \delta_k} = \frac{1}{ns} \sum_{i=1}^{ns} s_{ik}(1 - s_{ik}) \\ \frac{\partial s_k}{\partial \delta_m} &= \frac{1}{ns} \sum_{i=1}^{ns} \frac{\partial s_{ik}}{\partial \delta_m} = -\frac{1}{ns} \sum_{i=1}^{ns} s_{ik} s_{im} \\ \frac{\partial s_k}{\partial \delta_j} &= \frac{1}{ns} \sum_{i=1}^{ns} \frac{\partial s_{ik}}{\partial \delta_j} = -\frac{1}{ns} \sum_{i=1}^{ns} s_{ik} \frac{e^{V_{ij}}}{W_{ij}(1 + W_{ij} + W_{ij} W_{ik})} \end{aligned}$$

The derivatives for the PCs w.r.t the non-linear parameters are:

$$\begin{aligned}
\frac{\partial s_j}{\partial \sigma} &= \frac{1}{ns}\sum\nolimits_{i=1}^{ns}\frac{\partial s_{ij}}{\partial \sigma} = \frac{1}{ns}\sum\nolimits_{i=1}^{ns} s_{ij}\left(x_j v_i - \sum\nolimits_{d=1}^{J} s_{id}x_d v_i\right) \\
\frac{\partial s_j}{\partial \sigma_p} &= \frac{1}{ns}\sum\nolimits_{i=1}^{ns}\frac{\partial s_{ij}}{\partial \sigma_p} = \frac{1}{ns}\sum\nolimits_{i=1}^{ns} s_{ij}\left(p_j v_{ip} - \sum\nolimits_{d=1}^{J} s_{id}p_d v_{ip}\right) \\
\frac{\partial s_j}{\partial \sigma^S} &= \frac{1}{ns}\sum\nolimits_{i=1}^{ns}\frac{\partial s_{ij}}{\partial \sigma^S} = \frac{1}{ns}\sum\nolimits_{i=1}^{ns} s_{ij}\frac{\sum_{m=1}^{K} e^{V_{ik}} y_m v_i^S}{(1+W_{ik})(1+W_{ij}+W_{ij}W_{ik})} \\
\frac{\partial s_j}{\partial \sigma_p^S} &= \frac{1}{ns}\sum\nolimits_{i=1}^{ns}\frac{\partial s_{ij}}{\partial \sigma_p^S} = \frac{1}{ns}\sum\nolimits_{i=1}^{ns} s_{ij}\frac{\sum_{m=1}^{K} e^{V_{ik}} p_m v_{ip}^S}{(1+W_{ik})(1+W_{ij}+W_{ij}W_{ik})}
\end{aligned}$$

The derivatives for the servers w.r.t the non-linear parameters are:

$$\begin{aligned}
\frac{\partial s_k}{\partial \sigma^S} &= \frac{1}{ns}\sum\nolimits_{i=1}^{ns}\frac{\partial s_{ik}}{\partial \sigma^S} = \frac{1}{ns}\sum\nolimits_{i=1}^{ns} s_{ik}\left(y_k v_{ih}^S - \sum\nolimits_{m=1}^{K} s_{im}y_d v_i^S\right) \\
\frac{\partial s_k}{\partial \sigma_p^S} &= \frac{1}{ns}\sum\nolimits_{i=1}^{ns}\frac{\partial s_{ik}}{\partial \sigma_p^S} = \frac{1}{ns}\sum\nolimits_{i=1}^{ns} s_{ik}\left(p_k v_{ip}^S - \sum\nolimits_{m=1}^{K} s_{im}p_d v_{ip}^S\right) \\
\frac{\partial s_k}{\partial \sigma} &= \frac{1}{ns}\sum\nolimits_{i=1}^{ns}\frac{\partial s_{ik}}{\partial \sigma} = \frac{1}{ns}\sum\nolimits_{i=1}^{ns} s_{ik}\frac{\sum_{d=1}^{J} e^{V_{ij}} x_d v_i}{W_{ij}(1+W_{ij}+W_{ij}W_{ik})} \\
\frac{\partial s_k}{\partial \sigma_p} &= \frac{1}{ns}\sum\nolimits_{i=1}^{ns}\frac{\partial s_{ik}}{\partial \sigma_p} = \frac{1}{ns}\sum\nolimits_{i=1}^{ns} s_{ik}\frac{\sum_{d=1}^{J} e^{V_{ij}} p_d v_{ip}}{W_{ij}(1+W_{ij}+W_{ij}W_{ik})}
\end{aligned}$$

We also calculated the standard errors based on this Jacobian.

E. Estimation details for "free complementarity" model

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, customers and time. Given this assumption on the error term, for each customer 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:

$$\Pr(d_{PC} = 1 |\ x_j, y_k, i) = \frac{W_{ij}W_{ik}e^{\Gamma_{PC,S}} + W_{ij}}{W_{ij}W_{ik}e^{\Gamma_{PC,S}} + W_{ij} + W_{ik} + 1} \tag{40}$$

This follows because

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

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

The conditional brand choice probability for PC j is given by:

$$\Pr(j|\ d_{PC}=1,x_j,y_k,i)=\frac{\exp(V_{ij})}{W_{ij}}, \tag{41}$$

The unconditional brand choice probability is obtained by multiplication:

$$\Pr(j=1|\ x_j,y_k,i)=\Pr(d_{PC}=1|\ x_j,y_k,i)*\Pr(j|\ d_{PC}=1,x_j,y_k,i). \tag{42}$$

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 $\overline{\theta_2}\equiv(\sigma^{PC},\sigma^S,\sigma_p^{PC},\sigma_p^S)$ 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 $\delta=(\delta_j,\delta_k)$, $\xi=(\xi_j^{PC},\xi_k^S)$, $\nu_i=\left(\nu_i^{PC},\nu_i^S\right)$ 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 (41) for PCs and servers.

Step 3. Given $(\delta,\mu_i,\Gamma_{PC,S})$ calculate the marginal probabilities of equation (40) for PCs and servers.

Step 4. Calculate the unconditional brand probabilities of equation (42) 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 δ_{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.

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.

F. Estimating the relative output effect through a residual demand approach

As an alternative way to estimate the relative output effect, $-\frac{\left.\frac{dq(\mathbf{p}_j,\mathbf{p}_k,a)}{da}\right|_{\omega,\omega_M}}{\left.\frac{dq_M(\mathbf{p}_j,\mathbf{p}_k,a)}{da}\right|_{\omega,\omega_M}}$, we resort to a method that makes as little assumptions as possible about the maximization behavior of rivals to Microsoft in the server market. In essence, we estimate the residual demand functions for Microsoft's PC operating system demand q and server operating system demand q_M. This means that we are looking at the demands when all other players in the market are setting their equilibrium prices. This residual demand function will depend on the characteristics of PCs that are sold, as well as the PC operating system, the characteristics of Microsoft and non-Microsoft servers. We consider a "reduced form" estimation of PC and server quantities, as well as on the operating system prices of Microsoft ω and ω_M. Note that the derivatives of residual demand with respect to interoperability a corresponds precisely to the derivatives we need to calculate the relative output effect.

One worry is that changes in interoperability are fairly infrequent and hard to observe. However, given the assumption that server characteristics enter the indirect utility function linearly, the ratio of the derivatives is the same for any common marginal change in a given quality characteristics of rival servers. We can therefore exploit the quality variation in rival servers to identify the relative output effect. A further complication is that the number of observations to identify the relevant parameters is much lower than for our demand estimation, because we cannot exploit any cross-sectional variation in our data. For that reason we construct quality indices (following Nevo, 2003) for rival servers, Microsoft servers, and PCs in order to reduce the number of parameters to be estimated. We thus obtain estimating equations:

$$q_t = \lambda_1^{PC} I(\overline{y}_{kt,k\in M}) + \lambda_2^{PC} I(\overline{y}_{kt,k\notin M}) + \lambda_3^{PC}\omega_M + \lambda_4^{PC}\omega_t + \lambda_5^{PC} I(\overline{x}_t) + \varsigma_t^{PC} \tag{43}$$

$$q_M = \lambda_1^{M} I(\overline{y}_{kt,k\in M}) + \lambda_2^{M} I(\overline{y}_{kt,k\notin M}) + \lambda_3^{M}\omega_M + \lambda_4^{M}\omega_t + \lambda_5^{M} I(\overline{x}_t) + \varsigma_t^{M} \tag{44}$$

where $I(\overline{y}_{kt,k\in M})$ is an index of quality of servers running Microsoft server OS, $I(\overline{y}_{kt,k\notin M})$ is an index of quality of servers running non-Microsoft servers OS, and $I(\overline{x})$ is an index of PC quality. Since ω_t and ω_M are essentially unobservable, we replace them with the implied values from equations (6) and (7) evaluated at our estimated demand parameters.[40]

Given that variation in any quality characteristic will generate the same ratio of quantity changes, this will be true for variation in a quality index as well. We can therefore identify the relative output effect from the coefficients on the rival server quality index, λ_2^M and λ_2^{PC}. Hence, we estimate the relative output effect of interoperability as:

$$-\frac{\left.\widehat{\frac{dq(\mathbf{p}_j,\mathbf{p}_k,a)}{da}}\right|_{\omega,\omega_M}}{\left.\frac{dq_M(\mathbf{p}_j,\mathbf{p}_k,a)}{da}\right|_{\omega,\omega_M}} = -\frac{\widehat{\lambda_2^{PC}}}{\widehat{\lambda_2^{M}}} \tag{45}$$

The results for our residual demand estimations are presented in Table A5. The first column reports a regression of the quantity of Microsoft server operating systems sold against non-Microsoft server quality (as proxied by server memory[41]), a time trend and a seasonal dummy for the winter quarter.[42]

Microsoft server quality, PC quality[43], Microsoft server operating system prices and the PC Windows operating system price (see equation (43)). The signs of the coefficient is in line with expectations: non-Microsoft server quality is negatively and significantly correlated with Microsoft sales. Column (2) repeats the exercise for PCs and shows that higher non-Microsoft server quality is positively associated with PC demand, although the standard error is large. The implied interoperability effect is shown in the bottom row as 3.454, which is far below the relative margins in the Figures (at least at the end of the period). Since the quality variable is potentially endogenous we instrument it with its own lagged values in columns (3) and (4)[44]. This strengthens the results, with the coefficient on rival quality rising (in absolute value) for servers and falling closer to zero for PCs. This suggests that a degradation strategy would have low cost for Microsoft in terms of lost PC sales. The relative output effect is only 0.28 for the IV specification.

[40] For PC OS we also experimented with using the quality adjusted price index of Abel et al (2004).

[41] This was found to be the most important characteristic in the analysis of server demand in Van Reenen (2004). We experimented with other measures of server quality such as speed, but these did not give any significant extra explanatory power in the regressions.

[42] Demand is unusually high in this quarter because it coincides with the end of the fiscal year. Performance bonuses are usually based on end of fiscal year sales, so this generates a bump in sales (see Oyer, 1998, for systematic evidence on this effect).

[43] We build a quality index of PCs based on our estimates of Table 3 following Nevo (2003).

[44] The instrument has power in the first stage with an F-statistic of over 10 as shown at the base of the columns.

We include a host of other controls in the last two columns - Microsoft's own server quality, operating system quality, PC operating system prices and Microsoft's server operating system price. The coefficients remain correctly signed and the implied relative output effect remains below 4. Unsurprisingly, given the low degrees of freedom in the time series, the standard errors are large.

Overall, the "reduced form" results in Table A5 are consistent with the more structural approach taken in the paper. On average the relative output effect is small and certainly much smaller than the increase in margins from reducing interoperability.

TABLE 1 - RESULTS FOR SIMPLE LOGIT FOR PC MARKET SHARE

	(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)
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]
1^{st} 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%

Source: Authors' calculations based on the IDC Quarterly PC Tracker data corresponding to sales and prices of PC models for large business customers matched to more detailed PC characteristics from several industry datasources and trade magazines, US market (1996Q1-2001Q1).
Notes: Demand estimates from a simple logit model based on 3,305 observations. "Benchmark" are numbers assigned to each processor-speed combination based on technical and performance characteristics. "Generation" dummies indicate common technological characteristics shared among central processing units. All regressions include a full set of nine hardware vendor dummies. Columns (2) and (4) use BLP-type instruments: the number of the same form factor own-firm products, the number of the same form factor products produced by rival firms, the sum of the values of the same characteristics (speed, RAM and hard disk) of other products of the same form factor offered by the same firm and the sum of values of the same characteristics of all same factor products offered by rival firms. In the last two columns, we restrict the number of instruments dropping hard disks in column (3) and also speed in column (4). Full first stage results can be found in Table A3 of the Appendix. "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 2 - RESULTS FOR SIMPLE LOGIT FOR SERVER MARKET SHARE

	(1)	(2)	(3)	(4)	(5)	(6)
Estimation method	OLS	IV	OLS	IV	IV	IV
Dependent variable	$\ln(S_{kt})-\ln(S_{0t})$	$\ln(S_{kt})-\ln(S_{0t})$	$\ln(S_{kt})-\ln(S_{0t})$	$\ln(S_{kt})-\ln(S_{0t})$	$\ln(S_{kt})-\ln(S_{0t})$	$\ln(S_{kt})-\ln(S_{0t})$
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)
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%

Source: Authors' calculations based on the IDC Quarterly Server Tracker data corresponding to sales and prices of server models matched to more detailed server characteristics from several industry datasources and trade magazines, US market (1996Q1-2001Q1).

Notes: Demand estimates from a simple logit model based on 2,967 observations. All regressions include a full set of nine hardware vendor dummies. Columns (2) and (6) use BLP-type instruments: the number of own-firm products, the number of products produced by rival firms, the sum of RAM of products offered by rival firms. In columns (4) and (5) we also experiment with additional instruments based on server characteristics (sum of Rack and Rack Optimized of products offered by rival firms and sum of Unix own-firm models) and input prices (quality adjusted indices for semi-conductors and hard disks). Full first stage results can be found in Table A4 of the Appendix. "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 3 - RESULTS FROM ALTERNATIVE MODELS

	(1)	(2)	(3)	(4)	(5)
Estimation method	GMM	GMM	GMM	GMM	GMM
Empirical model	baseline model	strong complementarity	strong complementarity	strong complementarity	"free" complementarity
PANEL A: PC - Means					
Price	-3.301***	-3.102***	-3.057***	-2.844***	-3.314***
	(0.629)	(0.256)	(0.258)	(0.326)	(0.592)
Benchmark	0.021	0.145	0.059	-0.401	-0.153
	(1.243)	(0.429)	(0.477)	(0.284)	(1.176)
RAM	0.760**	0.965***	0.973***	1.016***	0.801***
	(0.316)	(0.232)	(0.245)	(0.296)	(0.303)
CD-ROM	0.275**	0.268**	0.271**	0.281**	0.278**
	(0.130)	(0.132)	(0.133)	(0.134)	(0.131)
Ethernet	0.423***	0.426***	0.438***	0.445***	0.444***
	(0.134)	(0.114)	(0.115)	(0.125)	(0.131)
5th Generation	2.783***	3.056***	3.055***	3.080***	2.821***
	(0.395)	(0.438)	(0.406)	(0.453)	(0.399)
6th Generation	4.053***	4.496***	4.517***	4.579***	4.132***
	(0.574)	(0.536)	(0.515)	(0.622)	(0.567)
7th Generation	2.709***	3.258***	3.285***	3.301***	2.733***
	(0.606)	(0.623)	(0.637)	(0.697)	(0.629)
Constant	-3.426***	0.832	0.798	0.224	-3.368***
	(0.708)	(0.799)	(0.644)	(0.817)	(0.704)
PANEL B: Server - Means					
Price	-0.282***	-0.231***	-0.233***	-0.256***	-0.674***
	(0.089)	(0.039)	(0.040)	(0.046)	(0.155)
RAM	0.173***	0.160***	0.163***	0.181***	0.208***
	(0.057)	(0.040)	(0.041)	(0.046)	(0.066)
Windows	0.794*	0.742**	0.755**	0.939**	1.543***
	(0.451)	(0.346)	(0.360)	(0.401)	(0.483)
Windows × RAM	0.077**	0.075**	0.076***	0.083***	0.133***
	(0.034)	(0.028)	(0.028)	(0.031)	(0.039)
Symmetric Parallel Processor	1.787***	1.765***	1.777***	1.924***	2.620***
	(0.390)	(0.278)	(0.286)	(0.322)	(0.408)
Rack Optimized	-0.185	-0.230	-0.240	-0.318	-0.373
	(0.234)	(0.217)	(0.220)	(0.240)	(0.273)
Number of Racks	0.060	0.064**	0.064**	0.077***	0.140***
	(0.039)	(0.029)	(0.028)	(0.032)	(0.034)
Constant	-5.814***	-10.649***	-10.596***	-10.389***	-8.096***
	(0.228)	(0.297)	(0.212)	(0.389)	(0.669)

	(1)	(2)	(3)	(4)	(5)
PANEL C: PC - Standard Deviations					
Price	0.916**	0.728***	0.702***	0.593***	0.902***
	(0.363)	(0.064)	(0.094)	(0.014)	(0.338)
Benchmark	1.321	1.176***	1.218***	1.484***	1.450*
	(0.822)	(0.170)	(0.191)	(0.026)	(0.752)
Constant				0.021	
				(0.023)	
PANEL D: Server - Standard Deviations					
Price	0.048**	0.001	0.002	0.001	0.162***
	(0.024)	(0.011)	(0.013)	(0.009)	(0.042)
RAM	0.014		0.005	0.007	0.027
	(0.104)		(0.013)	(0.070)	(0.091)
Constant				0.806***	
				(0.240)	
$\Gamma^{PC,S}$ parameter					2.647**
					(1.271)
GMM Objective (df)	75.613 (10)	75.111 (12)	70.344 (10)	74.293 (8)	57.493 (9)

Source: Authors' calculations based on the IDC Quarterly PC and Server Tracker data matched to more detailed PC and Server characteristics from several industry datasources and trade magazines, US market (1996Q1-2001Q1).

Notes: Column (1) presents demand estimates from the baseline model as described in subsection 4.1 in the main text based on 6,272 observations. Parameters were estimated via a two-step GMM algorithm described in the estimation subsection 4.2. Columns (2)-(4) report demand estimates from different specifications of the "strong complementarity" model as described in subsection 4.4 in the main text and estimated via a two-step GMM algorithm similar to the baseline model. Column (2) presents the simplest version, where we assume a random coefficient on price and quality benchmark for PCs and only price for servers. Columns (3) and (4) add progressively more random coefficients. Column (5) reports results from the "free complementarity" model estimated via a two-step GMM algorithm as described in Appendix E. The freely estimated parameter $\Gamma^{PC,S}$ allows us to model the extra utility that a customer obtains from consuming these two products together over and above the utility derived from each product independently. This paremeter would be positive for a pair of complements and negative for a pair of substitutes. All specifications include all the characteristics in Tables 1 and 2, 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. For all specifications we used BLP-type instruments corresponding to the number of the own-firm and rival products, as well as the sum of the values of the same characteristics (PCs: speed, RAM and hard disk; servers: RAM, number of racks, racks optimized, Unix) of other products offered by the same or rival firms. 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 parenthesis below coefficients: *significant at 10%; **significant at 5%; ***significant at 1%.

TABLE 4 - ROBUSTNESS

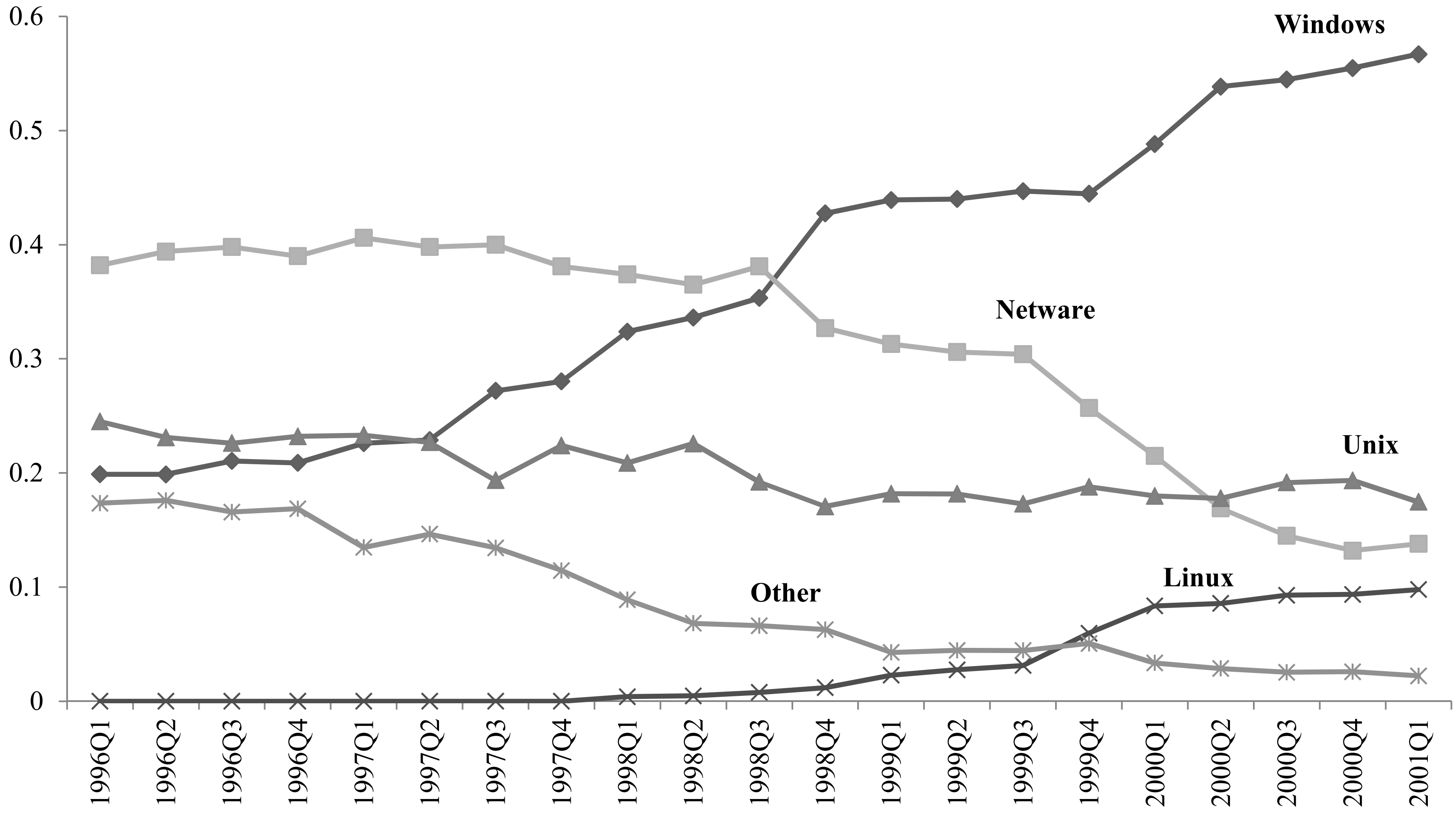
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Estimation method	GMM baseline model	GMM sample of 500 consumers	GMM potential mkt size reduced in half	GMM asymmetric potential mkt size	GMM reduced number of instruments	GMM reduced number of instruments	GMM include random coef. on constant	GMM reduce random coef. on servers
PANEL A: PC - Means								
Price	-3.301***	-2.795***	-3.002***	-3.353***	-3.622***	-5.598	-2.768***	-3.350***
	(0.629)	(0.501)	(0.411)	(0.604)	(0.676)	(3.882)	(0.555)	(0.635)
Benchmark	0.021	-2.503	0.070	0.088	-0.786	-1.388	-1.971*	0.020
	(1.243)	(1.572)	(0.493)	(1.114)	(2.355)	(5.770)	(1.152)	(1.229)
RAM	0.760**	1.088***	0.837***	0.747**	0.639*	0.244	0.753**	0.765**
	(0.316)	(0.284)	(0.278)	(0.303)	(0.348)	(0.568)	(0.320)	(0.312)
CD-ROM	0.275**	0.321***	0.261**	0.267**	0.304**	0.315	0.316**	0.275**
	(0.130)	(0.140)	(0.132)	(0.129)	(0.135)	(0.193)	(0.133)	(0.131)
Ethernet	0.423***	0.490***	0.486***	0.410***	0.403***	0.305	0.443***	0.424***
	(0.134)	(0.134)	(0.127)	(0.130)	(0.149)	(0.228)	(0.136)	(0.132)
5th Generation	2.783***	2.955***	2.811***	2.766***	2.869***	3.128***	3.153***	2.795***
	(0.395)	(0.461)	(0.401)	(0.391)	(0.400)	(0.547)	(0.491)	(0.394)
6th Generation	4.053***	4.517***	4.103***	4.007***	4.154***	4.296***	4.619***	4.066***
	(0.574)	(0.607)	(0.517)	(0.558)	(0.616)	(0.547)	(0.718)	(0.569)
7th Generation	2.709***	3.034***	2.757***	2.663***	2.529***	1.858	3.005***	2.702***
	(0.606)	(0.738)	(0.652)	(0.597)	(0.700)	(1.335)	(0.759)	(0.605)
Constant	-3.426***	-3.528***	-2.708***	-3.451***	-3.269***	-2.402	-6.319**	-3.379***
	(0.708)	(0.936)	(0.709)	(0.704)	(0.653)	(2.671)	(2.818)	(0.708)
PANEL B: Server - Means								
Price	-0.282***	-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	0.173***	0.203***	0.177***	0.162***	0.161***	0.180***	0.220**	0.174***
	(0.057)	(0.058)	(0.052)	(0.051)	(0.045)	(0.060)	(0.096)	(0.049)
Windows	0.794*	1.069*	0.828*	0.688	0.683**	0.888	1.342**	0.781*
	(0.451)	(0.556)	(0.460)	(0.431)	(0.342)	(0.737)	(0.590)	(0.436)
Windows × RAM	0.077**	0.092***	0.078**	0.072**	0.074**	0.085*	0.102**	0.076**
	(0.034)	(0.041)	(0.034)	(0.032)	(0.033)	(0.047)	(0.041)	(0.032)

Symmetric Parallel Processor	1.787***	2.015***	1.810***	1.698***	1.690***	1.858***	2.234***	1.773***
	(0.390)	(0.498)	(0.399)	(0.369)	(0.307)	(0.665)	(0.486)	(0.358)
Rack Optimized	-0.185	-0.266	-0.199	-0.145	-0.154	-0.208	-0.441	-0.176
	(0.234)	(0.267)	(0.237)	(0.223)	(0.203)	(0.329)	(0.296)	(0.227)
Number of Racks	0.060	0.084	0.063	0.056	0.055	0.074	0.094*	0.060
	(0.039)	(0.050)	(0.040)	(0.036)	(0.035)	(0.058)	(0.052)	(0.037)
Constant	-5.814***	-5.807***	-5.809***	-5.128***	-5.896***	-5.748***	-6.197***	-5.816***
	(0.228)	(0.260)	(0.228)	(0.223)	(0.294)	(0.269)	(0.566)	(0.228)
PANEL C: PC - Standard Deviations								
Price	0.916**	0.520*	0.758***	0.955***	1.140***	2.292	0.795***	0.938***
	(0.363)	(0.283)	(0.273)	(0.346)	(0.413)	(1.777)	(0.270)	(0.362)
Benchmark	1.321	2.794**	1.610***	1.282*	1.938	2.690	2.658***	1.332
	(0.822)	(1.102)	(0.444)	(0.771)	(1.532)	(3.199)	(0.910)	(0.812)
Constant							2.569	
							(1.839)	
PANEL D: Server - Standard Deviations								
Price	0.048**	0.062**	0.048*	0.042*	0.035	0.054*	0.049*	0.049**
	(0.024)	(0.031)	(0.025)	(0.025)	(0.037)	(0.030)	(0.028)	(0.024)
RAM	0.014	0.011	0.011	0.010	0.000	0.007	0.017	
	(0.104)	(0.090)	(0.106)	(0.103)	(0.159)	(0.312)	(0.145)	
Constant							0.930*	
							(0.528)	
GMM Objective (df)	75.613 (10)	68.583 (10)	71.783 (10)	88.356 (10)	54.146 (5)	46.723 (3)	56.899 (8)	79.292 (12)

Source: Authors' calculations based on the IDC Quarterly PC and Server Tracker data matched to more detailed PC and Server characteristics from several industry datasources and trade magazines, US market (1996Q1-2001Q1).

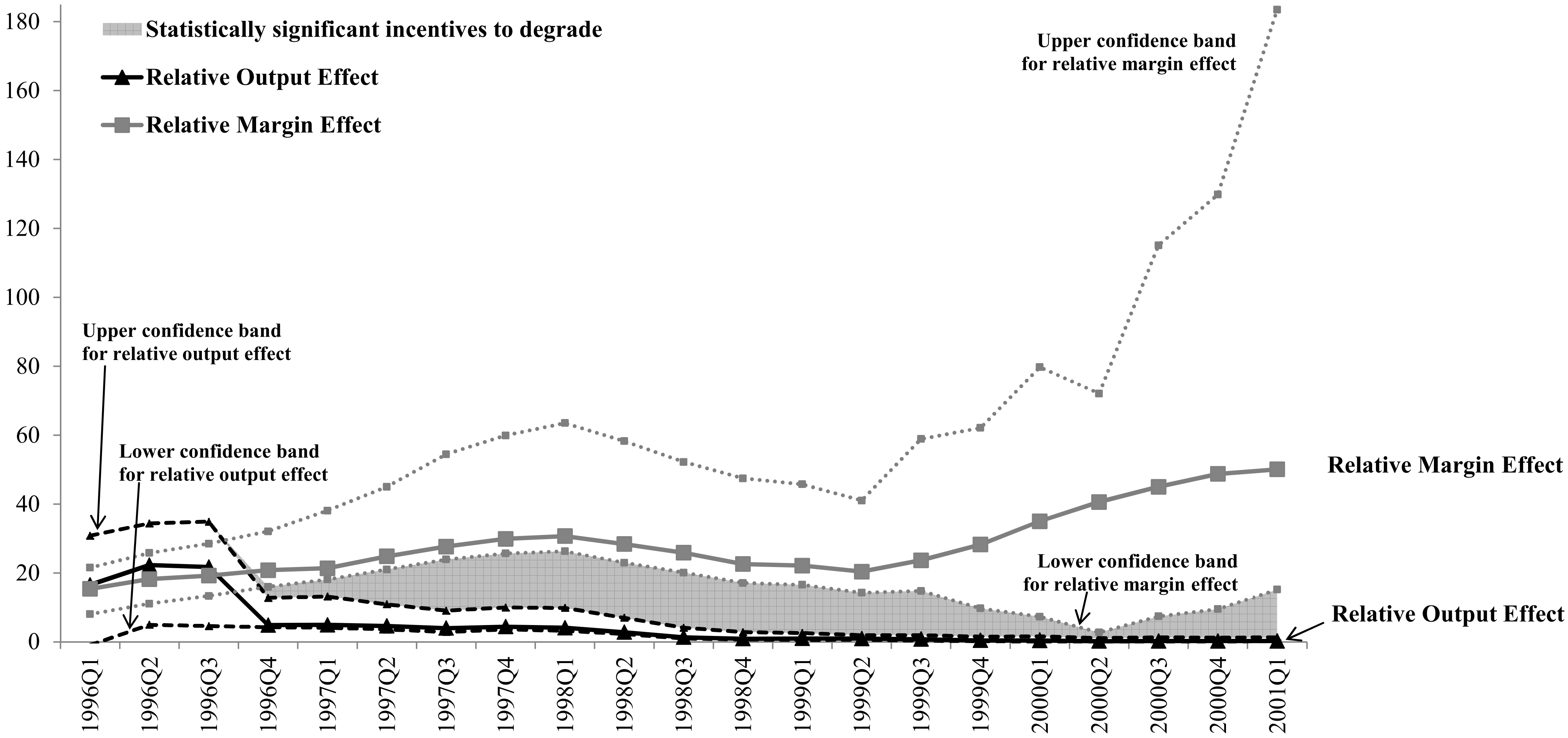
Notes: Demand estimates from the baseline model as described in section 4.1 in the text based on 6,272 observations. Parameters were estimated via a two-step GMM algorithm described in the estimation subsection 4.2. We include all the characteristics in Tables 1 and 2, i.e. for PCs: desktop, monitor size, CD-ROM, firm dummies and time trend; for servers: full set of operating system and firm dummies and time trend. In all columns (expect columns (5) and (6)) the instruments used are the same as in the baseline model in column (1). In column (2) we increase the number of draws relative to the baseline model to 500. In column (3) we assume that firms make a purchase decision every two years. In column (4), we assume that firms purchase a PC every year, whereas a server bundle every two years. In column (5) we reduce the number of instruments used to the ones corresponding to column (5) in Tables 1 and 2 for PCs and servers respectively. In column (6) we further reduce the instruments used for PCs to the ones corresponding to column (6) in Table 1. In column (7) we include a random coefficient on the constant for both PCs and servers. In the last column, we only allow for a random coefficient on price on servers. The standard errors take into account the variance introduced through the simulation by bootstrapping fifty times the relevant component of the variance in the moment conditions. Robust standard errors are reported in parenthesis below 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 market (1996Q1-2001Q1).
Notes: This plots the evolution of shares for different operating systems. 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.

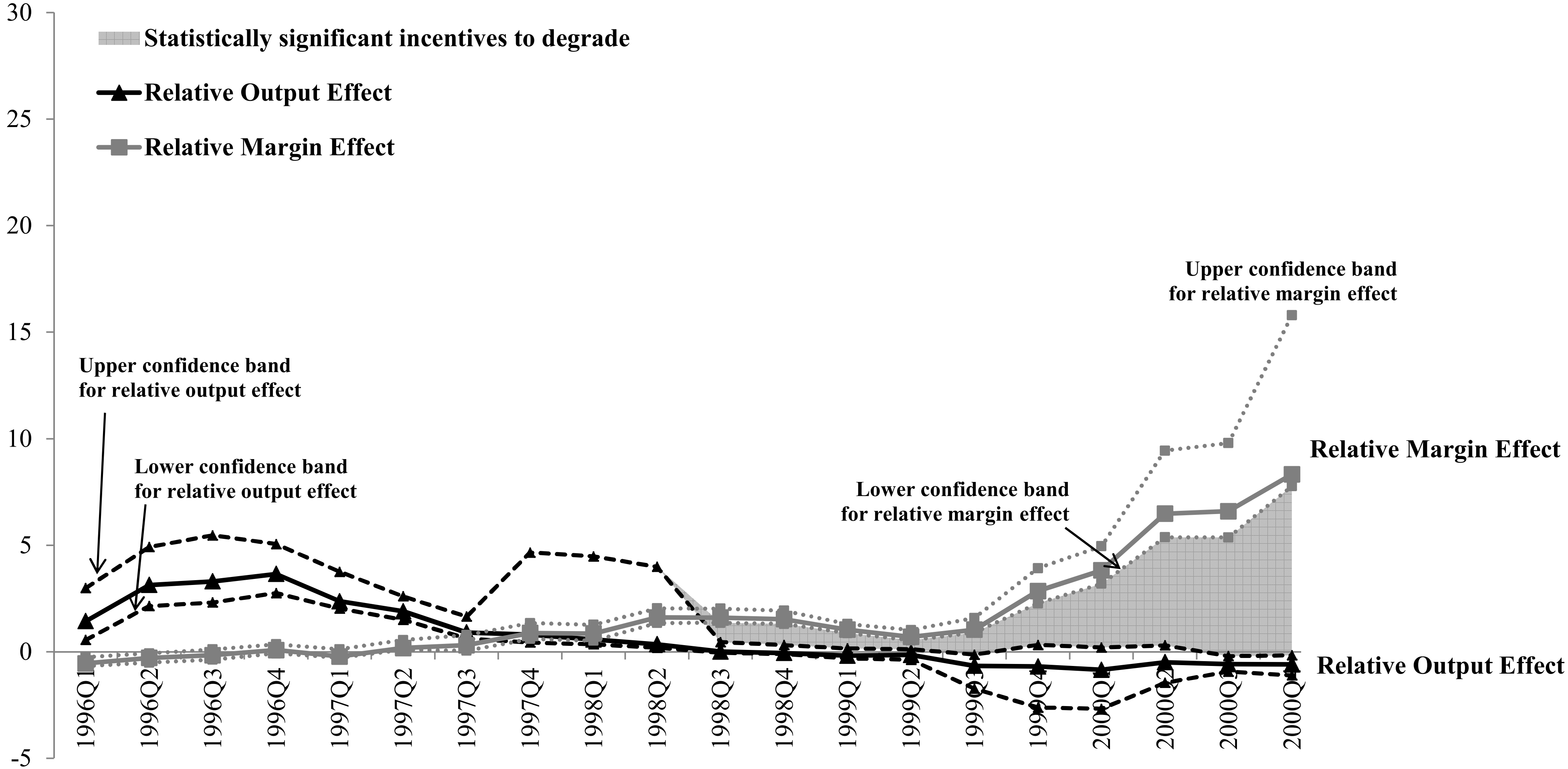
FIGURE 2: RELATIVE MARGIN AND OUTPUT EFFECTS (BASELINE MODEL)



Source: Authors' calculations based on the IDC Quarterly PC and Server Tracker data matched to more detailed PC and Server characteristics and the estimated coefficients in column (1) of Table 3.

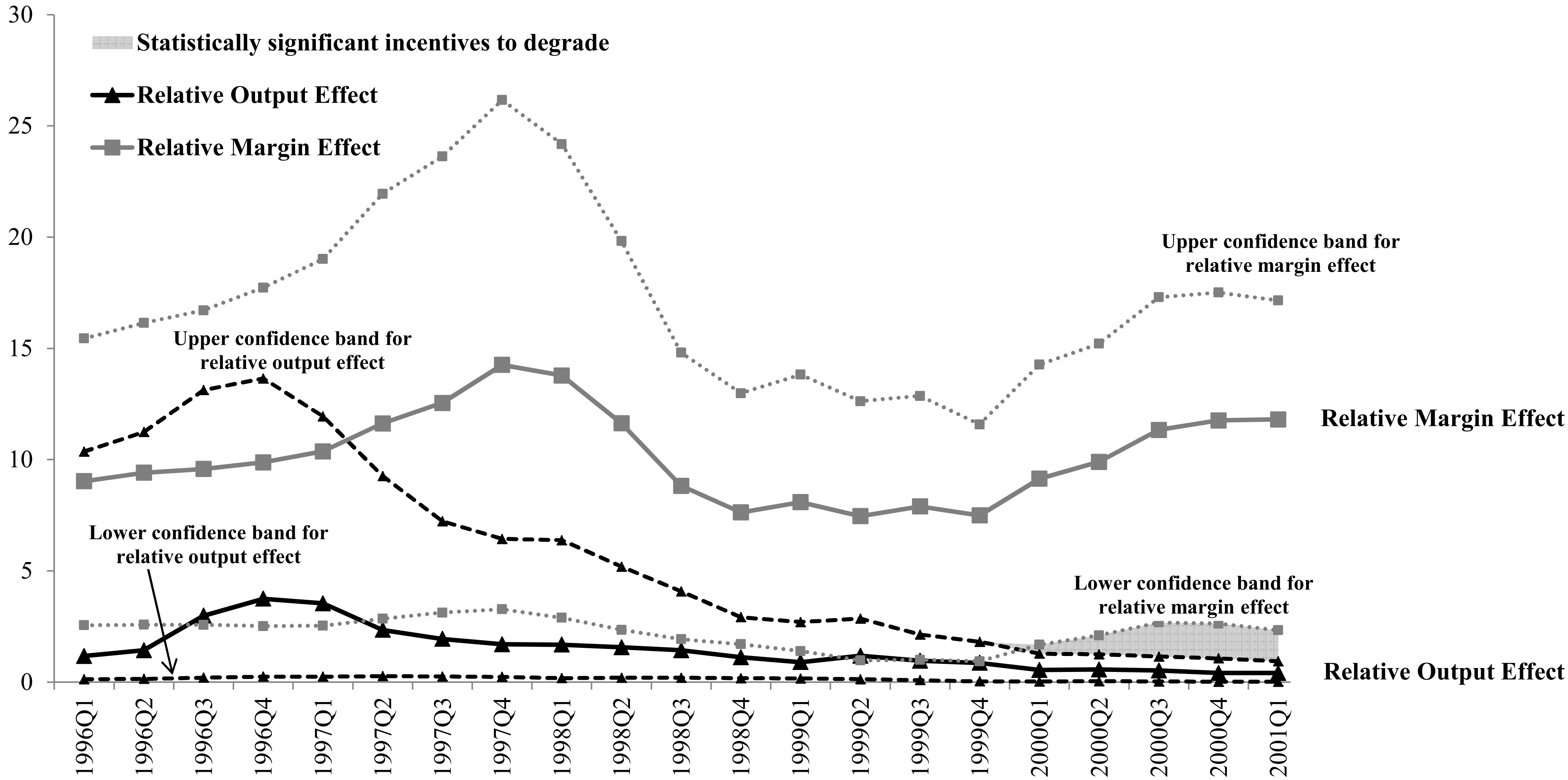
Notes: This plots the evolution of calculated relative output and margin effects based on the formulas provided in Appendix D and their 90% confidence interval. The shaded area highlights the period where the relative margin effect is statistically higher than the relative output effect, hence Microsoft had significant incentives to degrade according to our model.

FIGURE 3: RELATIVE MARGIN AND OUTPUT EFFECTS (STRONG COMPLEMENTARITY)



Source: Authors' calculations based on the IDC Quarterly PC and Server Tracker data matched to more detailed PC and Server characteristics and the estimated coefficients in column (4) of Table 3.
Notes: This plots the evolution of calculated relative output and margin effects based on the formulas provided in Appendix D and their 90% confidence interval. The shaded area highlights the period where the relative margin effect is statistically higher than the relative output effect, hence Microsoft had significant incentives to degrade according to our model.

FIGURE 4: RELATIVE MARGIN AND OUTPUT EFFECTS (FREE COMPLEMENTARITY)



Source: Authors' calculations based on the IDC Quarterly PC and Server Tracker data matched to more detailed PC and Server characteristics and the estimated coefficients in column (5) of Table 3.

Notes: This plots the evolution of calculated relative output and margin effects based on the formulas provided in Appendix D and their 90% confidence interval. The shaded area highlights the period where the relative margin effect is statistically higher than the relative output effect, hence Microsoft had significant incentives to degrade according to our model.

FIGURE 5: RELATIVE MARGIN AND OUTPUT EFFECTS (ROBUSTNESS)

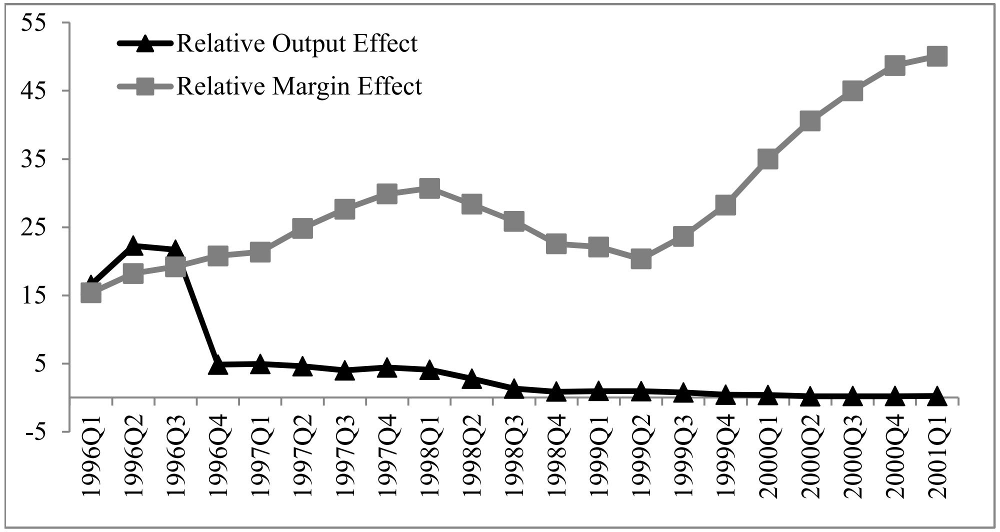


FIGURE 5A: COLUMN (1), TABLE 4

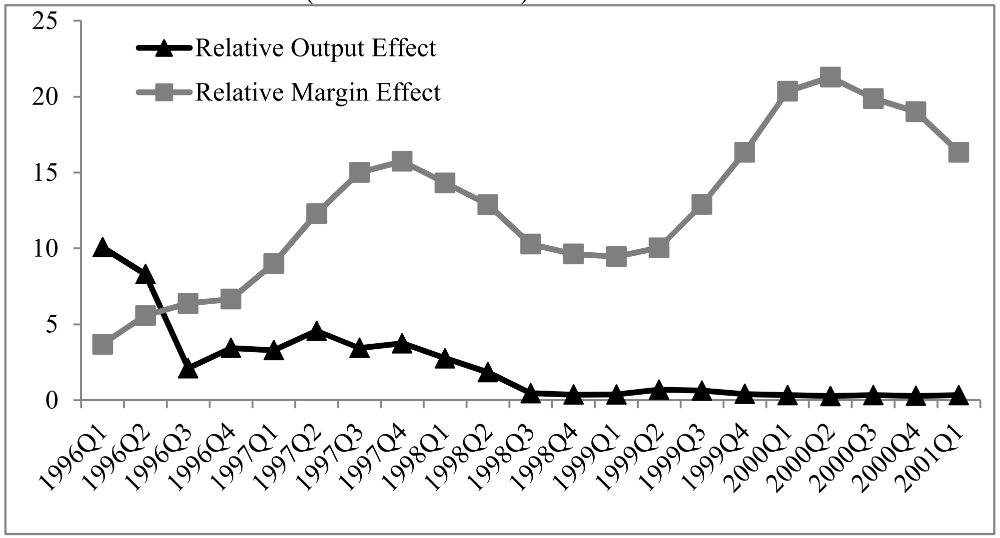


FIGURE 5B: COLUMN (2), TABLE 4

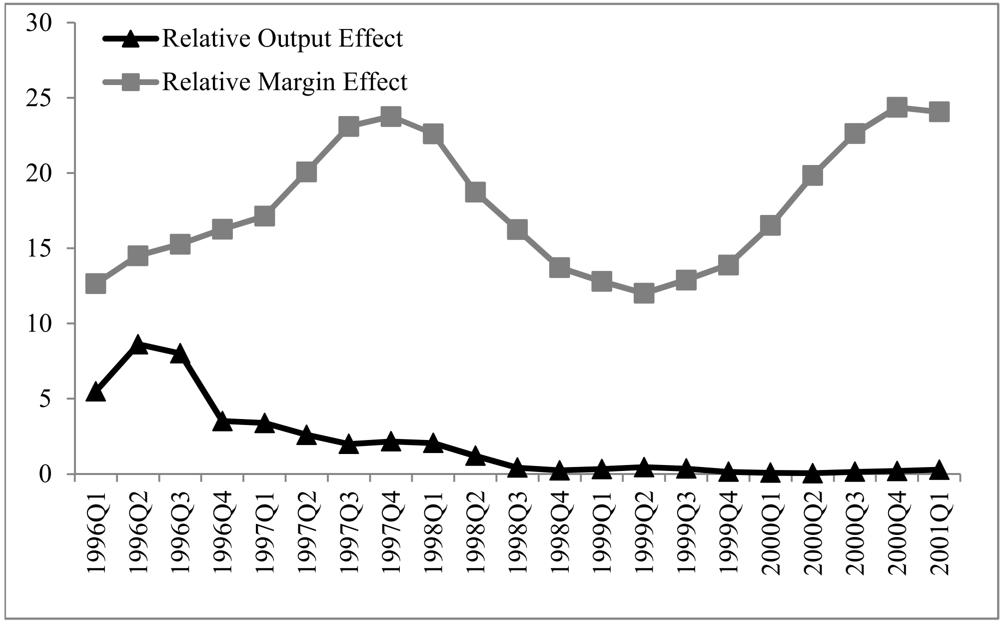


FIGURE 5C: COLUMN (3), TABLE 4

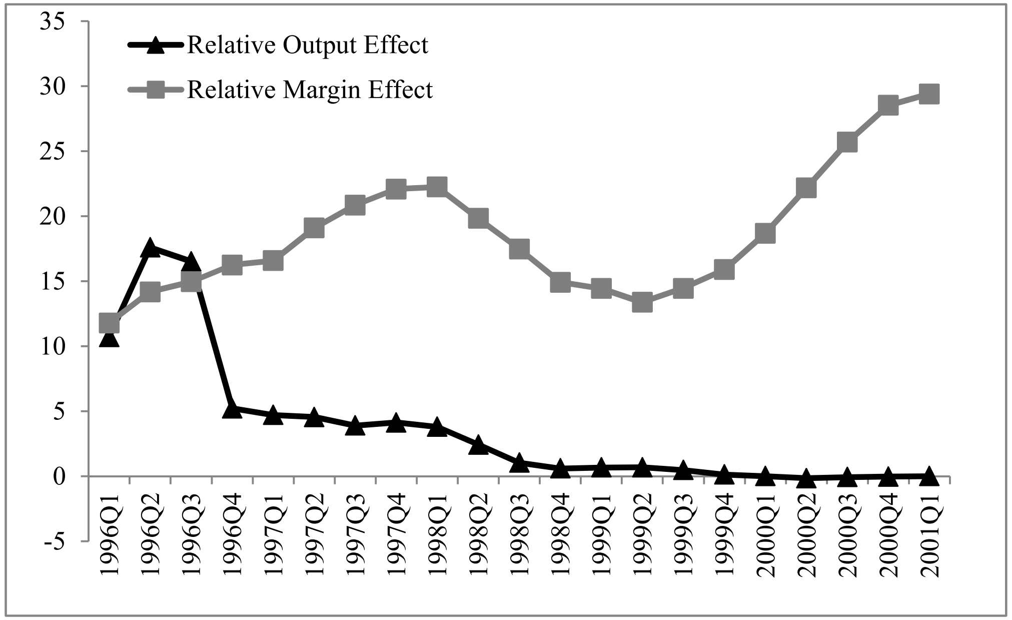


FIGURE 5D: COLUMN (4), TABLE 4

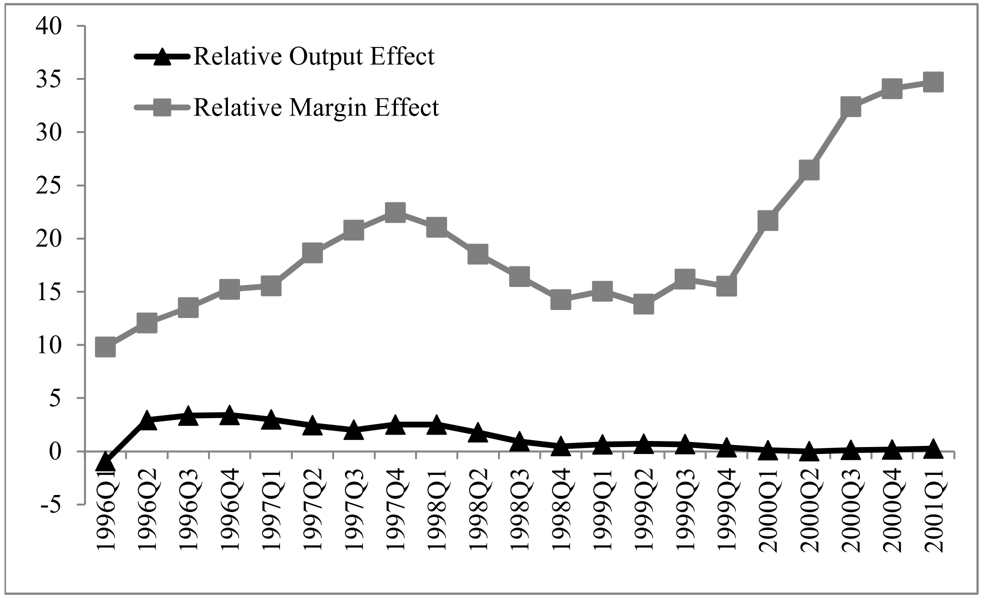


FIGURE 5E: COLUMN (5), TABLE 4

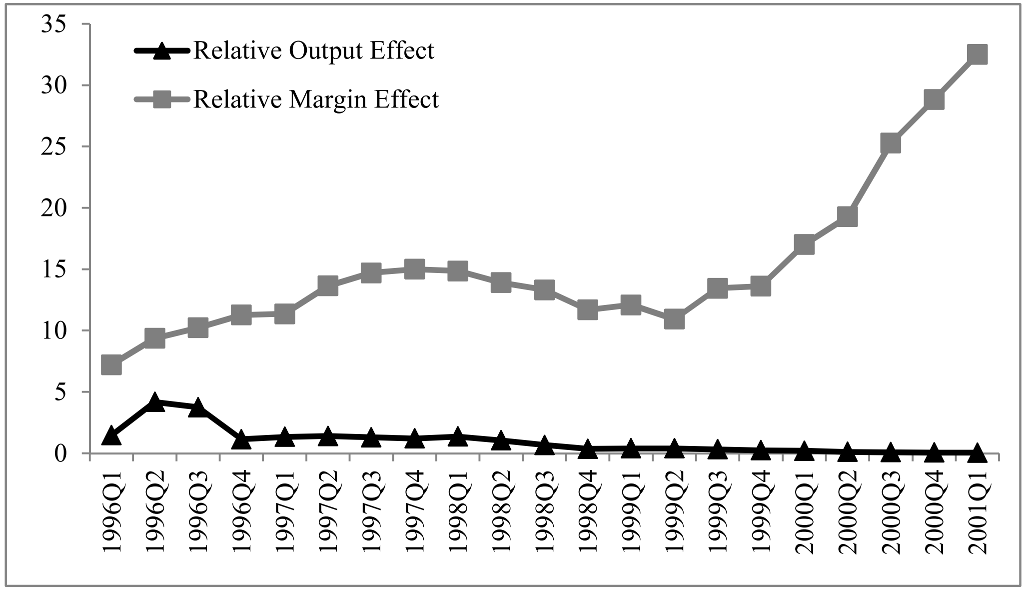


FIGURE 5F: COLUMN (6), TABLE 4

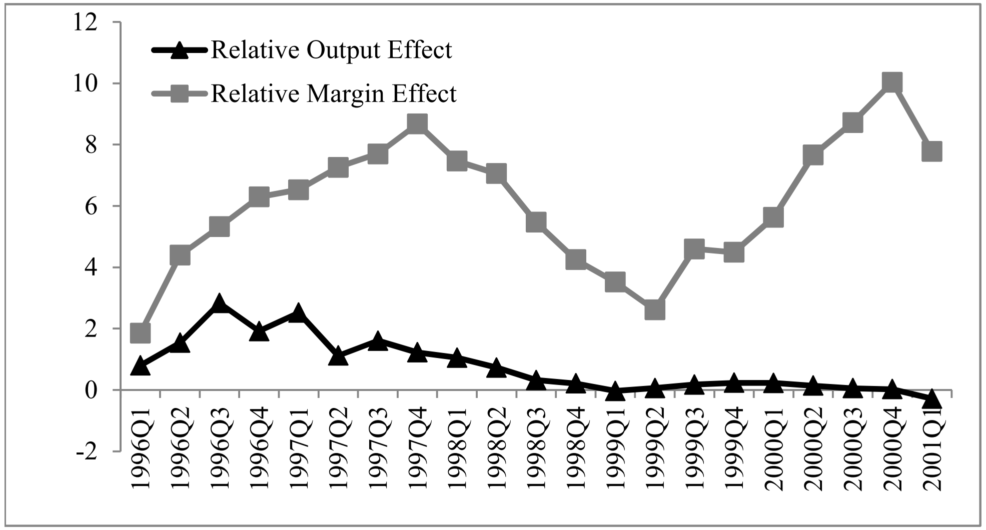


FIGURE 5G: COLUMN (7), TABLE 4

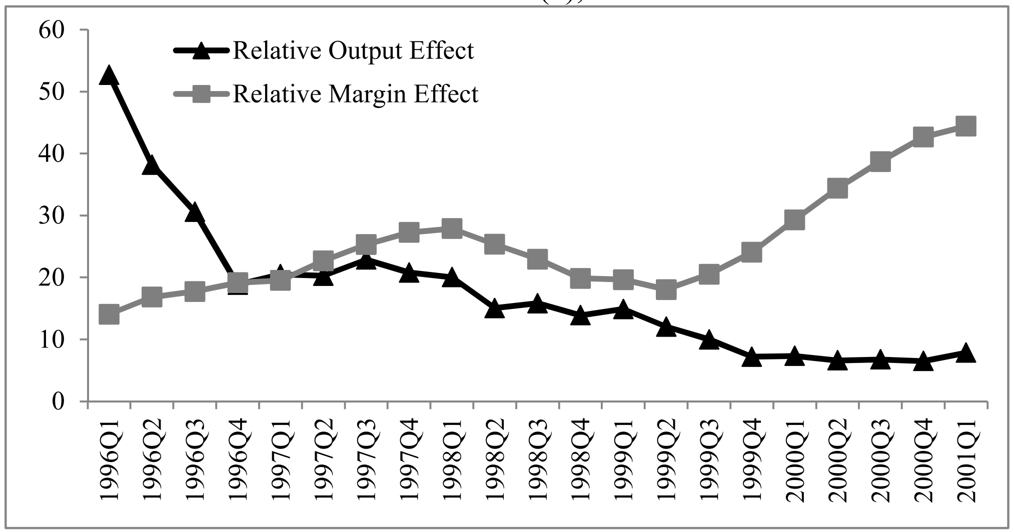


FIGURE 5H: COLUMN (8), TABLE 4

Source: Authors' calculations based on the IDC Quarterly PC and Server Tracker data matched to more detailed PC and Server characteristics and the estimated coefficients from Table 4.
Notes: These plot the evolution of calculated relative output and margin effects based on the formulas provided in Appendix D.

TABLE A1 - 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 for large business customers matched to more detailed PC characteristics from several industry datasources and trade magazines, US market (1996Q1-2001Q1).
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 A2 - DESCRIPTIVE STATISTICS FOR SERVER DATA

Period	No. of models	Quantity	Price ($1000s)	RAM (100MB)	Rack Optimized	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, US market (1996Q1-2001Q1).
Notes: All the entries (except model numbers and quantity) are weighted by server model sales.

TABLE A3 - LOGIT DEMAND FOR PCs - First Stage Results

	(1)	(2)	(3)	(4)
Instruments				
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 1. The regressions include all the exogenous variables in Table 1. 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 1. 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 A4 - 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)
Quality adjusted price indices for semi-conductors		-3568.634	-8592.337**	
		(5156.247)	(4135.553)	
Quality adjusted price indices for hard disks		-4989.916	-2513.099	
		(3507.393)	(3125.327)	
Sum of Rack Optimized in rival firm’s models		-203.671		
		(131.086)		
Sum of Racks in rival firm's models		-2.108		
		(5.148)		
Sum of Unix of firm’s models		64.556		
		(198.946)		

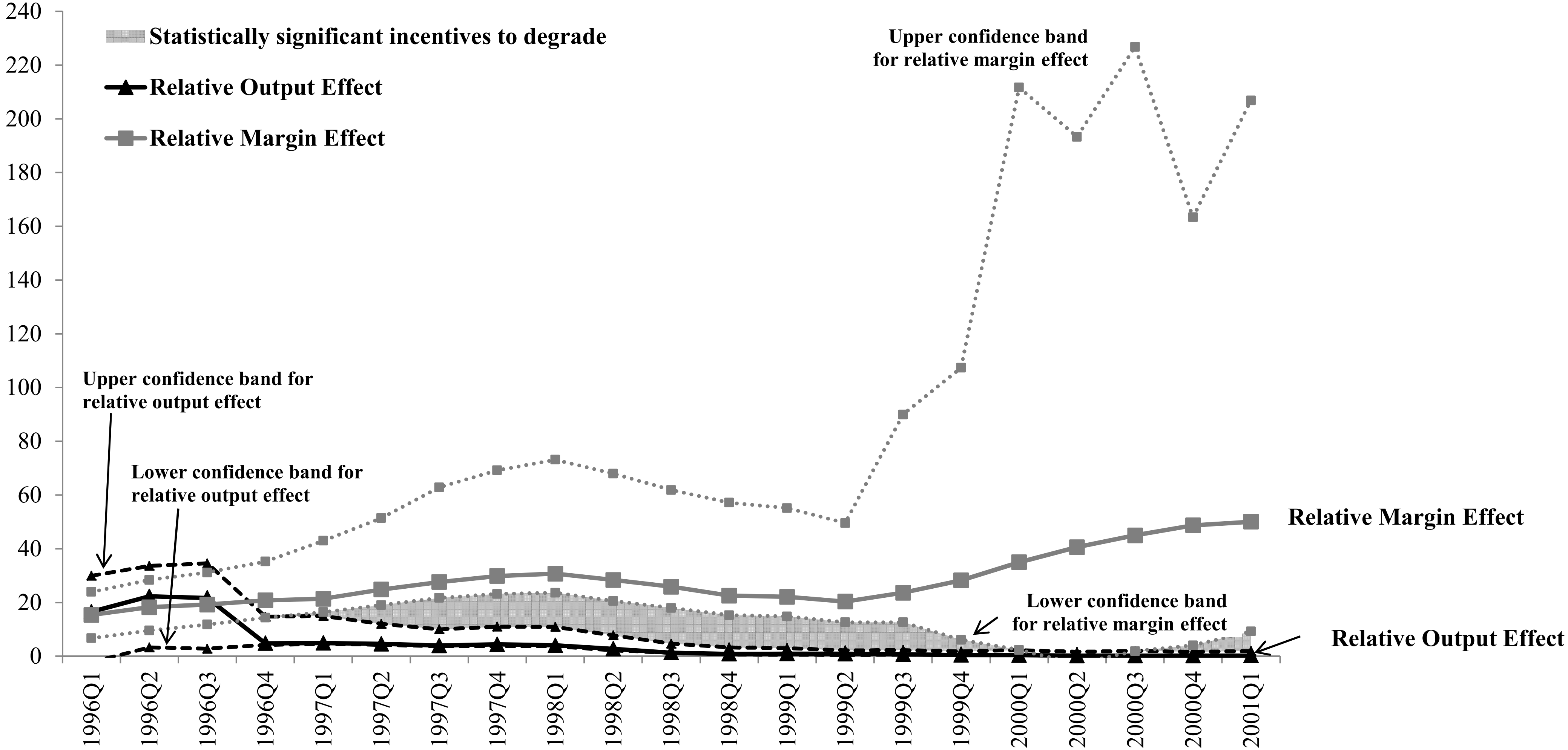
Notes: These are the first stage results from Table 2. The regressions include all the exogenous variables in Table 2. Column (1) corresponds to column (2), column (2) to column (4), column (3) to column (5) and column (4) to column (6) in Table 2. Coefficients and standard errors are multiplied by 1,000.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

TABLE A5 - REDUCED FORM ESTIMATES OF RESIDUAL INTEROPERABILITY EFFECT

	(1)	(2)	(3)	(4)	(5)	(6)
Estimation method	OLS	OLS	IV	IV	OLS	OLS
Dependent variable	Quantity sold of Windows servers	Quantity sold of PCs	Quantity sold of Windows servers	Quantity sold of PCs	Quantity sold of Windows servers	Quantity sold of PCs
Non-Microsoft Server quality	-4.110**	14.194	-6.142**	1.721	-4.504	17.960
	(1.804)	(17.479)	(2.754)	(19.950)	(2.985)	(25.760)
Other Controls	No	No	No	No	Microsoft server quality, PC quality, PC OS price, Microsoft server OS price	Microsoft server quality, PC quality, PC OS price, Microsoft server OS price
1st Stage F-test			10.2			
Implied relative interoperability effect	3.454		0.280		3.987	

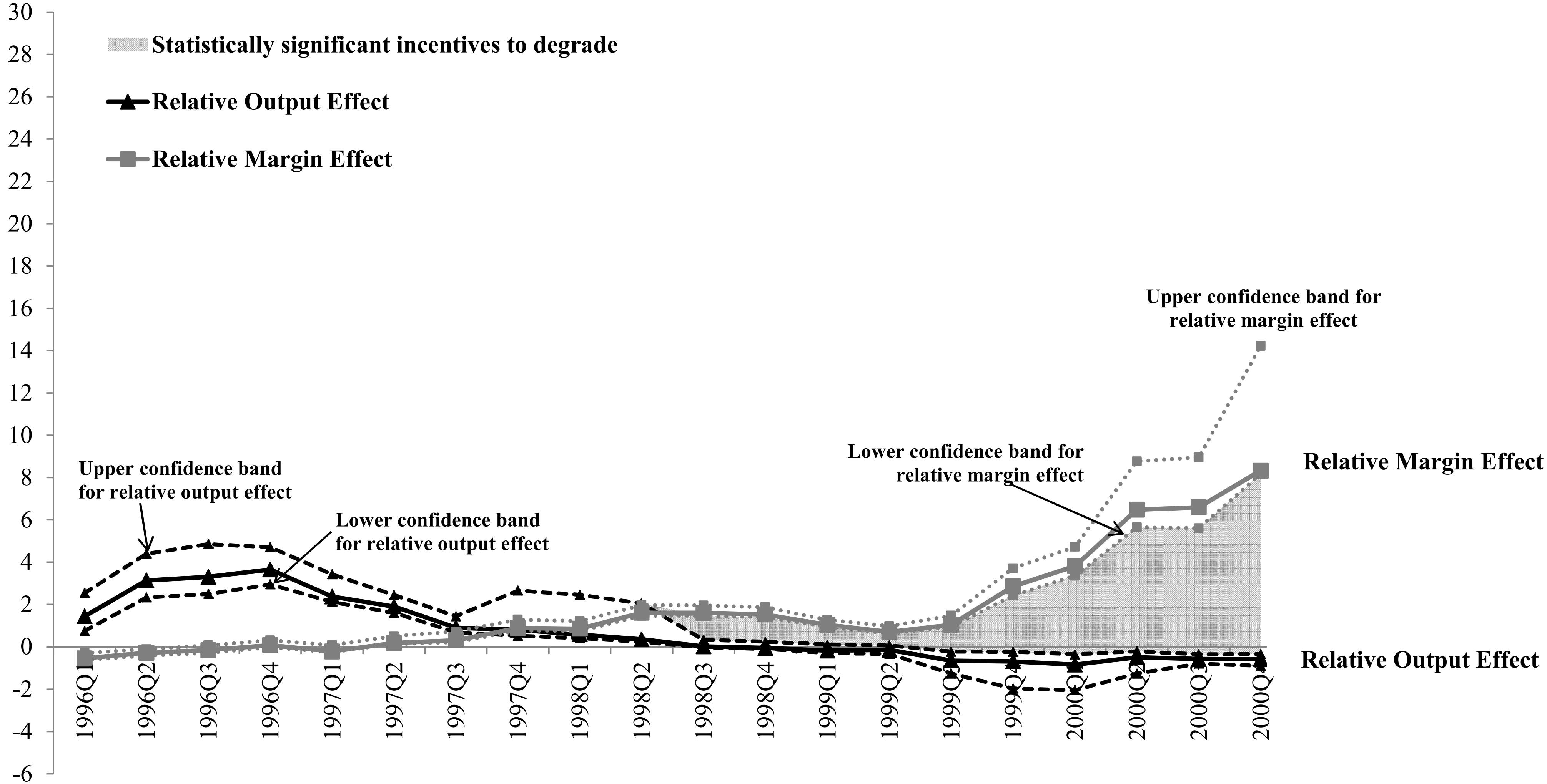
Source: Authors' calculations based on the IDC Quarterly PC and Server Tracker data matched to more detailed PC and Server characteristics from several industry datasources and trade magazines, US market (1996Q1-2001Q1).
Notes: All equations control for a winter quarter dummy and time trend. Columns (3) and (4) treat server quality as endogenous and instrument with lagged quality. The "implied interoperability effect" is calculated using the formula in Appendix F. Robust standard errors are reported in parenthesis below coefficients: *significant at 10%; **significant at 5%; ***significant at 1%.

FIGURE A1: RELATIVE MARGIN AND OUTPUT EFFECTS (BASELINE MODEL, 95% CONFIDENCE INTERVAL)



Source: Authors' calculations based on the IDC Quarterly PC and Server Tracker data matched to more detailed PC and Server characteristics and the estimated coefficients in column (1) of Table 3.
Notes: This plots the evolution of calculated relative output and margin effects based on the formulas provided in Appendix D and their 95% confidence interval. The shaded area highlights the period where the relative margin effect is statistically higher than the relative output effect, hence Microsoft had significant incentives to degrade according to our model.

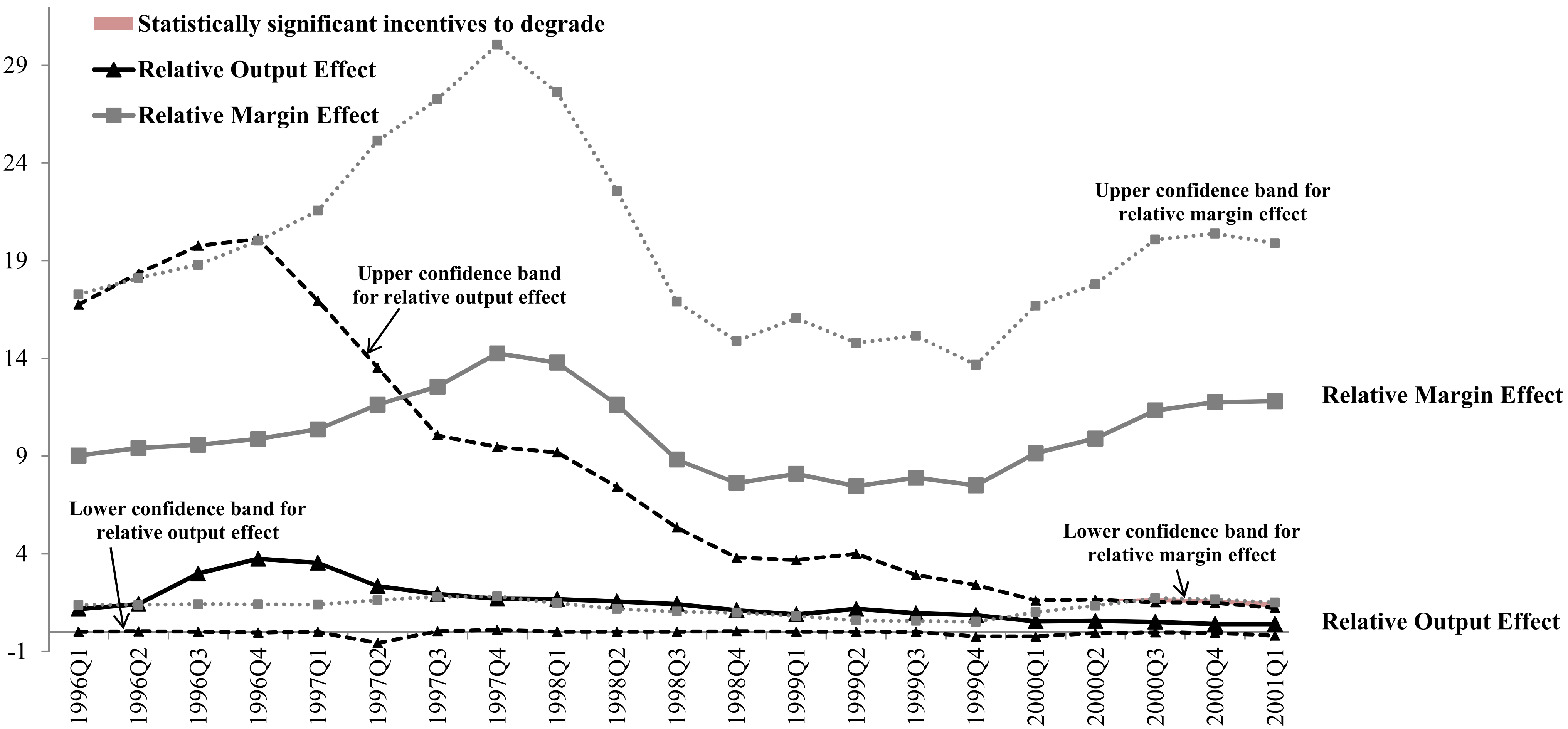
FIGURE A2: RELATIVE MARGIN AND OUTPUT EFFECTS (STRONG COMPLEMENTARITY, 95% CONFIDENCE INTERVAL)



Source: Authors' calculations based on the IDC Quarterly PC and Server Tracker data matched to more detailed PC and Server characteristics and the estimated coefficients in column (4) of Table 3.

Notes: This plots the evolution of calculated relative output and margin effects based on the formulas provided in Appendix D and their 95% confidence interval. The shaded area highlights the period where the relative margin effect is statistically higher than the relative output effect, hence Microsoft had significant incentives to degrade according to our model.

FIGURE A3: RELATIVE MARGIN AND OUTPUT EFFECTS (FREE COMPLEMENTARITY, 95% CONFIDENCE INTERVAL)



Source: Authors' calculations based on the IDC Quarterly PC and Server Tracker data matched to more detailed PC and Server characteristics and the estimated coefficients in column (5) of Table 3.

Notes: This plots the evolution of calculated relative output and margin effects based on the formulas provided in Appendix D and their 95% confidence interval. The shaded area highlights the period where the relative margin effect is statistically higher than the relative output effect, hence Microsoft had significant incentives to degrade according to our model.