OPENING NETWORKS TO COMPETITION:
The Regulation and Pricing of Access,
edited by David Gabel and David Weiman,
Kluwer Academic Publishers, 1998.

10

EXCLUSIONARY BEHAVIOR IN THE MARKET FOR OPERATING SYSTEM SOFTWARE: THE CASE OF MICROSOFT[1]

Glenn A. Woroch, Frederick R. Warren-Boulton
and Kenneth C. Baseman

1. INTRODUCTION AND SUMMARY

This chapter examines Microsoft's licensing practices for its MS-DOS and Microsoft Windows operating system software. Our main focus is on Microsoft's use of CPU (central processing unit, or per-processor) licenses under which an original equipment manufacturer (OEM) of personal computers pays a royalty for each machine it ships rather than for each unit of an operating system it installs. We also examine license provisions that require purchase of a minimum number of rights to install an operating system, Microsoft's tying of both technical support information and Windows to MS-DOS, and Microsoft's attempts to induce technical incompatibility between MS-DOS and its main competitor, DR-DOS.

We begin in section 2 with a brief description of the market for personal computer operating systems, and a history of Microsoft's licensing practices and technical design tactics. We also track the record of antitrust investigations of Microsoft, both here and abroad, that culminated in the signing of a consent decree with the Department of Justice.

In section 3 we examine the main efficiency argument for a CPU license, i.e., that it is a variant on the standard two-part tariff used to achieve first-degree price discrimination which is generally efficient and welfare-enhancing. Upon closer examination, however, we find that the CPU license is not equivalent to a two-part tariff. In this specific factual context, uniform linear prices may maximize profits for a secure monopoly, while a two-part tariff would be neither welfare enhancing nor—absent an exclusionary effect—profit maximizing. We conclude that Justice's attempt to eliminate CPU licenses was subverted by its own endorsement of volume discounts which can approximate lump-sum payments to any desired degree.

Section 4 turns to potential anticompetitive rationales for Microsoft's practices in the DOS market. We begin by observing that markets for many high technology products are characterized by a competitive process where a new product appears with a significantly superior technology or design and sweeps the field. By rapidly displacing the old product and its old technology, the new product achieves a large market share in a short time, earning high gross profit margins. This situation persists only until the dominant firm's product is itself displaced by another new product. This cycle of a new product with an innovative technology displacing an existing product with an old technology is a process of "creative destruction" in the race to be best. Firms achieve a dominant position, but that position is only transitory because, without artificial barriers to entry, today's dominant or monopoly firm and product can readily be dislodged by a new product developed by a competitor or a new entrant.

When the monopolist's position is protected by strategically erected barriers to entry, however, this displacement process can come to a halt. We examine the possibility that Microsoft has used a variety of exclusionary practices—notably nonlinear pricing and technical incompatibility—not to achieve its initial position but rather to retain that position against new competition. We conclude that, under the conditions present in the operating systems market, such practices can be, and in this instance have been, effective in limiting the growth and threatening the existence of entrants and rivals with small market shares. We conclude that Microsoft's anticompetitive behavior has reduced social welfare.

2. BACKGROUND

2.1 The Market for Personal Computer Operating Systems

2.1.1 Personal computer platforms

Our focus is on the market for packaged software that operates personal computers, and to a lesser extent, the software applications that run using those operating systems. To better understand the market for these products, we must delve into the economics and technology of the personal computer.

PCs can be decomposed into hardware and software components. Some components are essential: every computer system requires a microelectronic chip (usually called the central processing unit, or CPU) plus operating system (OS) software. The OS directs the stream of instructions requested by the applications software, while the CPU performs the numerical computations. Importantly, the CPU and the OS are almost always combined in fixed proportions: one of each is needed per system.

Once an OS is installed, a user can run many kinds of applications software.[2] The most popular packages do word processing, spreadsheet analysis, and database management. Increasingly popular is the use of a graphical user interface (GUI) that simplifies the management of the various applications. Both applications and GUIs are optional components of a personal computer system.

Personal computers are available in several "platforms" that differ in their hardware specifications. The so-called "IBM-compatible" PC is the predominant platform that evolved from the hardware and software specifications of the machine first introduced by IBM in 1981.

2.1.2 Industry structure

The supply of many components is highly concentrated. An overwhelming proportion of IBM-compatible PCs in use today are equipped with CPUs manufactured by the Intel

Corporation. The majority of existing PCs run on one version or another of the operating system sold by Microsoft Corporation. Sales of applications software and peripheral hardware components are far less concentrated.

Hundreds of OEMs assemble hardware components in various configurations (usually called "models"), distribute the machines through retail stores or mail order, and provide technical and repair service. In addition to a few large OEMs such as Compaq, Dell and AST Research in the U.S. and NEC, Toshiba and Hitachi in Asia, there is a host of small resellers. We can safely assume this segment of the market to be competitive.

In the early 1990s, the bulk of new PCs shipped in the U.S. (see table below) arrived loaded with some operating system, usually Microsoft's MS-DOS, and often with the Microsoft Windows interface as well. IBM ships its PCs with one of its own operating systems: PC-DOS or OS/2.[3] The only independent OS (i.e., compatible with, but not a clone or derivative of MS-DOS) were Digital Research Incorporated (DRI's) DR-DOS (which, with Novell's acquisition of DRI in 1991, became known as Novell DOS) and IBM's PC-DOS. Users could purchase OSs at retail stores or direct from the software publisher.

In 1992, it was estimated that the worldwide installed base of personal computers of all platforms totaled over 138 million (Bernstein Research, 1993). Of those, 72 percent were IBM-compatible. Less than a quarter of those machines were equipped with Microsoft Windows.

2.1.3 Supply conditions

Operating system software is very costly to develop and market. For instance, it has been estimated that IBM has spent over $2 billion developing OS/2. In comparison, reproducing and distributing operating system software is extremely cheap. As a result, fixed costs are enormous while marginal costs are negligible. The fixed costs are also largely sunk. The code itself is rarely of much value in other uses. Development teams accumulate expertise and reputation, only a portion of which can be redeployed into other projects.

Besides the irreversible investment in computer code, incumbents acquire sunk, or partially sunk, assets such as customer lists and brand name recognition. Furthermore, any new OS must be compatible with all the applications that were written to that "standard." User switching costs also limit the ability of new entrants to gain a toehold. Of course, these costs erect barriers only when the incumbent firm has a first-mover advantage. However, sunk costs ordinarily imply a first-mover advantage, at least for the current vintage of technology.[4]

2.1.4 History of PC operating systems

Dating back to 1976, Digital Research Incorporated sold a popular operating system, called "CP/M", for use on machines based on Intel's 8-bit 8080 chip. In 1980, in what has been called "the deal of the century," Microsoft paid a mere $100,000 for the rights to a CP/M derivative software package called "Disk Operating System," which, with minor modifications, became the initial MS-DOS. In 1981, when IBM launched its entry into the personal computer market, it selected Intel's new 16-bit 8088 chip as the CPU. It also chose to endorse Microsoft's MS-DOS as the operating system.

IBM's partnership with Microsoft later fell apart. But in the meantime, neither IBM nor DRI stopped developing their own operating systems.[5] Under the terms of the dissolution, IBM continued to develop MS-DOS, and eventually its own variant, PC-DOS, which it loaded on PCs bearing the IBM nameplate. In exchange, IBM agreed to pay Microsoft a royalty for a predetermined number of units.

Having been passed over by IBM, DRI went on to modify CP/M for the Intel 8086 chip, leading to its CP/M-86 product. Later it developed DOS PLUS and then DR-DOS. In April 1990, DRI introduced DR-DOS 5.0 to critical acclaim. Instantly, it began to make inroads into MS-DOS 4.0's market share. By year-end 1990, DR-DOS's share had increased to 10 percent of new OS shipments, leaving MS-DOS with 70 percent and IBM with 18 percent.[6]

Within a month of DR-DOS 5.0's inauguration, Microsoft announced development of MS-DOS 5.0. Curiously, it was to contain nearly all of the innovative features of the DRI product. Yet MS-DOS 5.0 was not commercially available until July 1991, more than a year after DR-DOS 5.0's release. Anticipation of the new Microsoft product, prolonged by continuous Microsoft statements indicating imminent availability, reined in growth of DR-DOS 5.0 sales (Sherer, 1990).

The emergence of the graphical interface played an important role in the events that followed. After repairing bugs in Microsoft Windows 3.0, Microsoft shipped Microsoft Windows 3.1 in April 1991. In that year, 18.5 percent of new PC shipments included Microsoft Windows along with MS-DOS. By 1992, that fraction jumped to 59.7 percent. Over that period, sales of MS-DOS (both with and without Microsoft Windows 3.1) rose 28.9 percent while sales of PC-DOS and DR-DOS fell 15.4 percent (see Table 1). By 1993, the market shares for operating systems on x86 PCs were 79 percent for MS-DOS, 13 percent for PC-DOS, 4 percent for OS/2, 3 percent for DR-DOS and 1 percent for UNIX.[7]

2. MICROSOFT'S PRACTICES

2.1 The CPU License

When first available, MS-DOS was sold to OEMs for a flat fee. Microsoft offered an unrestricted number of copies for $95,000 and, for a limited time, reduced that price by half.[8] Around 1983, Microsoft began to gear its license fees to the level of OEM sales. Each OEM contract was individually negotiated; an external price list never existed.

Over time, Microsoft phased in a new type of royalty contract. By 1992, the "CPU license" became the dominant sales arrangement, with 60 percent of Microsoft's operating system sales made under CPU licenses.[9] Under its terms, affiliated OEMs were required to pay a royalty for every CPU they shipped. Since each machine had a single CPU, the OEM paid for a copy regardless of whether the machine was preloaded with MS-DOS. Microsoft would sell DOS licenses to OEMs which refused the CPU license, but only at significantly higher prices.

Under the CPU license, an OEM usually had to commit also to a minimum "requirement" R that approximates its annual shipments. The one-time charge for this requirement is computed using a negotiated per-unit price p multiplied by R.[10] If an OEM shipped a machine with a competing operating system, say PC-DOS or DR-DOS, it would receive no reduction in its payment to Microsoft. Consequently, an OEM which accepts a CPU license faces a zero marginal price for units of MS-DOS up to the minimum requirement.[11] In the event that an OEM exceeded its projected volume during the contract period, the per-unit fee p used to calculate the lump sum payment for the first R units would apply to each unit above R. Thus, once the contract is in place, the marginal price is 0 up to R units and p for additional units beyond R.

Table 1 New Shipments of Personal Computer Operating Systems (000s of units)

Company	Operating System	1990	1991	1992
Microsoft	MS-DOS	11,648	13,178	18,525
	w/ Windows	490	2,440	11,056
	w/o Windows	11,158	10,738	7,469
IBM	PC-DOS	3,031	3,003	2,315
DRI/Novell	DR-DOS	1,737	1,819	1,617
DOS Subtotal		28,064	31,178	22,847
Apple	Macintosh	1,411	2,204	2,570
UNIX	UNIX	357	582	797
IBM	OS/2	0	0	409
Other	NEC, etc.	5,079	4,628	4,458
Totals		23,450	25,702	31,080

Sources: Bernstein Research, International Data Corporation

Regardless of whether an OEM ends up shipping more or less than R PCs during the contract, the terms of the CPU license commit the OEM to pay for one unit of MS-DOS for each PC it ships. As a result, customers view themselves as paying double if they use other OSs. If the supplier of a competing OS offers to sell at a per-unit price m, the OEM will only buy the second OS if that OS has a quality advantage over MS-DOS valued at m or more.

Table 2 shows the marginal cost of a PC under various scenarios facing a PC manufacturer, which has signed a CPU license with Microsoft. Let z be the marginal cost of producing the machine excluding an operating system, let X be the number of PCs produced regardless of which operating system is installed, if any at all. When the OEM ships less than the requirement of R units, its marginal cost of using MS-DOS on the next PC is zero, compared with a marginal cost of m if it chooses another OS. After R PCs have been shipped, the marginal cost is p if the OEM uses MS-DOS, and it is $p + m$ if the OEM uses the other OS. In each case, the increment to marginal cost from using an alternative OS is m. If the machine is shipped "naked," then the total marginal cost is just the production marginal cost, z.

Table 2 Marginal Cost of Different PC Configurations

CPUs shipped by OEM	MC with MS-DOS		MC with Alternative OS	
	OS	PC	OS	PC
X < R	0.00	z	m	z + m
X > R	p	z + p	p + m	z + p + m

In 1992, the average license fee per copy of MS-DOS to a hardware OEM under these CPU licenses has been estimated at $15, far below the average retail price of an upgrade of $49 (Bernstein Research, 1993). All together in that year, Microsoft grossed $399 million on worldwide sales of 18,525,000 units of MS-DOS to OEMs and as upgrades.[12]

The typical CPU license ran for a period of 2 years. It was quite likely an OEM will finish the contract period with unused licenses, in which case the customer does not necessarily receive a credit for its unused units. Microsoft exercised its discretion as to when the OEM could carry forward its unused licenses from the prior year.[13]

In addition to the price incentives for exclusivity embodied in the CPU license, Microsoft was alleged to have responded with a variety of direct penalties should an OEM ship some of its machines with a competing operating system. First, the OEM could be prohibited from carrying forward unused MS-DOS licenses, or be required to renew the CPU license at equal or higher volumes to retain the carry-forward option. In this way, Microsoft's policy on carry forwards could establish a "tie" between each year's sales and the next year's sales.

Second, Microsoft's technical service and support could be withheld from an OEM which installed a competing OS. This practice can disadvantage an OEM which needs this information to match its hardware configuration (especially the choice of the microprocessor, the amount of RAM, and the graphics card) with the demands of the operating software.

Third, the price of Microsoft Windows was allegedly increased to OEMs which purchase OSs from someone other than Microsoft. As far back as the days when Microsoft Windows was called "Interface Manager," Microsoft established a connection between the terms of sale of MS-DOS and its graphical interfaces.[14] Microsoft cautioned OEMs against bundling competing multitasking interfaces (such as Quarterdeck's DESQview, VisiCorp's VisiOn and DRI's GEM) with PC hardware components such as hard disks. Discounts on Microsoft Windows were extended to OEMs which agreed to accept a CPU license for MS-DOS. Those who refused the CPU license or who did not use MS-DOS exclusively, could still purchase Windows, but at a much higher per-unit price.

2.2 Technical Incompatibilities

Coordination on technical standards is crucial between the OS developer and applications developers. Nowhere is this coordination more important than with the publication of the Applications Program Interfaces (APIs) which contain the technical specifications that permit applications programs to communicate with the operating system. Microsoft has left undocumented some of these interfaces. In principle, access to these APIs would allow Microsoft to write applications (such as for its MS Word word processor or its Excel spreadsheet) that work faster and with greater functionality. Furthermore, even if an applications developer were to discover and use these undocumented interfaces, Microsoft could, as long as they remain "unofficial," remove or alter them in later versions of the operating software, rendering parts of the applications useless.

Compatibility is also be crucial to the success of operating system software when it must work with programs that function as an intermediary between operating systems and applications programs such as Microsoft's Windows program. Competitors to MS-DOS need to be aware of the functionality of Microsoft Windows so that their products remain compatible with applications written for Windows. In several instances, Microsoft made it difficult for competitors, especially DRI/Novell's DR-DOS, to achieve compatibility with Microsoft Windows. Concerns over possible incompatibility between DR-DOS and Microsoft Windows resulted in significant declines in DR-DOS sales.

One way for applications programmers to insure compatibility with an operating system is to receive copies of the preliminary version of the software. Known as "beta testing," this gives applications developers an opportunity to fine tune the interaction between the two programs.

In a well-publicized episode, DRI was excluded from the beta testing of Microsoft Windows 3.1 and, subsequently, from the beta testing of Microsoft's Windows for Workgroups product. The importance of compatibility testing with the beta version became evident when applications developers using DR-DOS received error messages warning them of a potential incompatibility with Microsoft Windows. Upon installation, Microsoft Windows 3.1 checked whether the source of the underlying system and the extended memory manager were Microsoft products. If they were not, the user was informed that a problem was detected, and was asked to contact Microsoft's beta technical support for Microsoft Windows 3.1. This message appeared on the screen even though no actual compatibility problem was detected. Indeed, if users continued past the alleged error message, they would discover that Microsoft Windows 3.1 would run in conjunction with DR-DOS.[15] The error messages raised fears of incompatibility among developers and users who contemplated running Microsoft Windows with non-Microsoft OSs. Finally, Microsoft Windows disks included a "Readme" text file that cautioned users that "running Microsoft Windows 3.1 with an operating system other than MS-DOS could cause unexpected results or poor performance."[16]

2.3 Antitrust Action

Microsoft's practices first came to the attention of antitrust authorities in Korea. The Korean Fair Trading Commission launched an investigation that centered on use of the CPU license in Asia. In May 1992, the Korean FTC banned the use of CPU licenses in that country (Phang, 1992). That action was not very effective, however, because Microsoft then began offering customer-specific price schedules with steep "cliffs" (sharp average price reductions) at volumes close to the customer's requirements.

In June 1990, the U.S. Federal Trade Commission initiated a nonpublic (sic) investigation of Microsoft's practices. The investigation eventually focused on Microsoft's marketing practices for DOS and Windows.[17] Without ever acknowledging the investigation, the Commission twice voted on whether to seek a preliminary injunction requiring Microsoft to cease and desist from its marketing practices. Both times the outcome was a 2-2 tie, resulting in no FTC action.

But then, in an unprecedented move, the Antitrust Division of the Department of Justice (the Department) took up the case and, after extensive further investigation, negotiated a consent decree with Microsoft. On July 15, 1994 the Department filed a civil antitrust complaint along with a proposed Final Judgment to which Microsoft had consented (the Consent Decree).[18] Simultaneously, Microsoft consent to a settlement filed by the European Commission. Next, a Competitive Impact Statement (CIS) was filed as required under the Tunney Act.[19]

The case then took an even more startling twist when Judge Sporkin of the DC District Court refused to play the role of a "mushroom"[20] and rejected the decree as inadequate under the Tunney Act. The U.S. government and Microsoft jointly appealed Judge Sporkin's decision. The U.S. Court of Appeals for the District of Columbia has since upheld the consent decree.

3. THE CPU LICENSE, FIRST-DEGREE PRICE DISCRIMINATION AND QUANTITY DISCOUNTS

3.1 The CPU License as First-Degree Price Discrimination

At first glance it may appear that the CPU license is just a means to provide discounts to OEMs that purchased large volumes of MS-DOS. This is, however, not the case; because the OEM's average royalty payment for MS-DOS is based on the share of its machines shipped with MS-DOS, an OEM that purchases *more* MS-DOS could pay a *higher* per-unit price than one that purchases *fewer* units. This would happen if an OEM purchased more units of MS-DOS than some other OEM, but then proceeded to ship many machines that were loaded with an alternative OS.[21]

Nor can the CPU license be characterized as first-degree price discrimination in any meaningful sense. First-degree price discrimination occurs when a seller charges a two-part fee, consisting of a lump-sum payment for the right to purchase the product and a price for each unit equal to the marginal production cost. Where, as here, marginal cost is essentially zero, first-degree price discrimination requires a per-unit price of zero.[22] It is correct that an OEM which signs under the CPU license (or a take-or-pay license with X > output) has agreed to a lump-sum payment, with an (expected) zero marginal price for one year. However, since the size of the lump-sum payment is based on expected sales multiplied by a per-unit royalty, the OEM knows that if its sales increase, the (apparent) lump-sum payment next year will also increase proportionately (based on the per-unit royalty Microsoft will be charging in the next year). Thus, for any time horizon longer than one year, the CPU license is a tax on output; it is not first-degree price discrimination.[23]

This finding should not be surprising, however, since first-degree price discrimination would not be profitable to Microsoft (nor would it be welfare-enhancing) when compared to a per-unit royalty. Economists have long recognized the strong efficiency advantages of first-degree price discrimination when customers are final consumers so that their demands are independent. But these results do not carry over to intermediate inputs sold to competing downstream firms. In that case the demands of such customers (the firms in the downstream industry) are clearly not independent (i.e., if my rivals pay less for an input than I do, the price of the final product falls, reducing my demand for the input). Ordover and Panzar (1982) state the issue quite clearly.

> . . . we recast the welfare analysis of the simple two-part tariff using the classical model of perfect competition in which all firms are identical and free entry and exit ensures that the equilibrium output price is equal to minimum average cost. In this context we discover that two-part tariffs are not generally desirable from a welfare standpoint . . . . This is due to the fact that the entry fee, instead of acting as a "lump sum levy," affects both the equilibrium number of firms and their output level. This new distortion must be balanced against the losses due to a unit price in excess of marginal cost.

Where, as here, the input (the OS) is used in fixed proportions with the output (the PC) and the downstream industry is a classic competitive industry with U-shaped average cost curves, Ordover and Panzar find that a strong theoretical result obtains: a monopoly seller of the input would find any two-part tariff, including an all-or-nothing offer where marginal cost to the buyer is zero, ***less profitable than a uniform per-unit fee***. In addition, the uniform per-unit fee results in ***higher economic welfare than any two part-tariff***. As Ordover and Panzar (1982) put it,

> Most surprisingly, for the empirically relevant class of production processes in which the purchased input is required in fixed proportion to output, we discover that a two-part tariff is never optimal from either a profit or welfare maximizing standpoint. (p. 660.)
>
> The intuition behind this result is rather straightforward. It is well known that under fixed proportions an upstream uniform pricing monopolist can extract all the profits which an integrated uniform pricing monopolist could reap. Since competition downstream ensures that a uniform price prevails in the final product market, there can be nothing to gain from introducing a two-part tariff; optimal choice of [the per-unit price] allows the monopolist to earn the maximum possible under such circumstances. There is something to lose, however, since an entry fee $e > 0$ causes the downstream firms to operate at an inefficiently large scale. Total (upstream plus downstream) costs are not minimized and a portion of this dead-weight burden falls on the monopolist. Viewed another way, this result reveals the futility of attempting to impose a seemingly nondistortionary

lump-sum levy c on a perfectly competitive industry with free entry and exit. (pp. 666-67)

In short, even if Microsoft's CPU license (or equivalent volume discounts) did impose a true lump-sum payment, there would be no efficiency or welfare gain that could provide a defense for such a pricing system. Nor would such a licensing system be profitable for Microsoft to impose on OEMs even if those OEMs somehow did not recognize the link between their sales and the lump-sum royalties they paid. Both theory and the available evidence would indicate, therefore, that Microsoft's CPU license (or its equivalent in the form of a volume discount) is not a form of first-degree price discrimination.

While the CPU license does not produce a positive output effect (i.e., encourage efficient utilization of a zero-marginal-cost input), it does have a significant substitution effect. The CPU license induces substitution of MS-DOS for an alternative OS. While this may be privately profitable, the social gain is zero even if it did not induce the exit of rival operating systems such as DR-DOS, with its attendant expected effects on raising the MS-DOS license fees. MS-DOS and any other OS have a near-zero social marginal cost in use. Thus, to the extent that the CPU license induces substitution of MS-DOS for an alternative OS, no cost saving results. Indeed, if, as appears to be the case, other OSs offer greater value than that of MS-DOS, the substitution of MS-DOS for alternative OSs actually reduces efficiency and total welfare even in the short run. The adverse effects on social welfare are even greater in the long run, since the exclusionary nature of the CPU license will deter investments in competing OSs.

Finally, we should note that two other efficiency defenses for CPU licenses were raised, at least during the course of the FTC investigation. CPU licenses, it was argued, might be an effective way both to deter unscrupulous OEMs from engaging in under-reporting the number of units of MS-DOS installed and to reduce software piracy by OEMs, retailers and/or computer users. By reducing the number of "naked" machines shipped by OEMs, a CPU license could eliminate the incentive to engage in piracy and fraud. An examination of the historical record leads us to conclude, however, that the prevention of piracy and fraud is not a plausible explanation for why CPU licensing was introduced. Even more telling, however, is that the CPU license is no more effective at deterring piracy or fraud than are other available but unused non exclusionary alternatives such as a "credited-CPU" licenses.[24]

3.2 The DOJ Consent Decree and Quantity Discounts

Microsoft's practices did not allow an OEM to reduce its total payments to Microsoft if it installed a competing operating system on some of its machines. The Department of Justice's complaint and the CIS clearly state that such contracts are illegal and explains their exclusionary and anticompetitive nature. The consent decree does define and ban three types of contracts—per-processor licenses, lump-sum pricing, and minimum commitments—under which there is no reduction whatsoever in an OEM's total payments to Microsoft when the OEM installs a competing operating system on some of its machines. Nevertheless, the consent decree explicitly permits schemes that amount to *near* per-processor pricing, i.e., extreme quantity discounts that can have the same effect, or an effect sufficient to exclude a competitor.[25]

Thus, if Microsoft sets a royalty of $2.5 million to an OEM with a projected output of 100,000 machines, this would be lump-sum pricing. But if Microsoft sets a royalty of $2.499 million plus $0.01 for each unit of MS-DOS installed, this is not lump-sum pricing and would not be banned by the decree. To eliminate any possible confusion on this issue, the consent decree explicitly allows for license arrangements that embody volume discounts.[26] Thus, our hypothetical sales contract ($2.499 million for the first unit of MS-DOS, one cent for each additional unit) is explicitly legal.

The Department was aware of the potential for anticompetitive uses of quantity discounts.[27] Lacking evidence that Microsoft used volume discounts to foreclose competitors, the Department permitted this practice.[28] However, as long as CPU licenses are available to Microsoft, using quantity discounts to achieve exclusion would be redundant and unnecessary so that one should hardly expect to see them used. Only when CPU licenses are prohibited would we expect to see Microsoft turn to sales practices with an equally exclusionary. This is just what had occurred in Korea in 1992, after the Korean FTC investigated and banned the use of CPU licensing by Microsoft.[29] And, even if the Department did not believe when it entered into the consent decree that Microsoft would turn to exclusionary volume discounts, they should question their belief after the first report of Microsoft turning to such discounts.[30]

4. MARKET-POWER RATIONALES FOR CPU LICENSES

4.1 Workable Competition in Technology Markets with Rapid Technological Change

Economic theory would predict highly volatile market shares under a set of conditions that, to varying degrees, have often characterized PC software markets. Consider a market where numerous potential entrants face no *ex ante* barriers to entry into the development of a new technology: entrepreneurs, usually scientists or engineers themselves, put together teams of scientists and engineers, financed internally from their past successes or from venture capital, with access to a common pool of basic technology and to learning acquired at their previous firms. These new ventures incur significant sunk costs to develop a higher-quality technology that (we shall assume) is protected by laws that cover intellectual property to the optimal extent.[31] The new technology may be simply licensed to users (as to OEMs in the case of software) or embodied in a new product using manufacturing facilities available from competitive firms in a number of markets (e.g., software duplicators, or packagers for shrink-wrapped sales of software at retail). The products embodying these alternative technologies are mutually exclusive in the sense that a customer will almost always use only one operating system on any PC.

When two other conditions also hold, we would expect to observe "competitive" or "socially optimal" performance. First, firms in this market take their competitors' prices as given and unaffected by their own actions, and try to undercut their rivals' (quality adjusted) prices as long as that price exceeds their own marginal cost. Second, customers can costlessly switch among the products of rival suppliers.

Given these two conditions, we would expect to observe that (1) a new technology or product will be developed if (and only if) the expected value of the cost of development is less than the expected value of the increase in the value to consumers of this technology over the prior technology; (2) the price of the old technology (e.g., the license or royalty fee) will fall to zero upon introduction of the new technology; (3) the price of the new technology will equal the difference in value between the old and the new technologies; and (4) market share will rapidly shift from 100 percent for the old technology to 100 percent for the new technology.

While such competition may seem tough on the players, it can still generate very large rewards to the winners needed to cover the risks and costs of development, and results in even greater benefits to consumers since as each new generation appears, the value added by the prior generation is passed on directly to consumers.[32] It is efficient in terms of production and distribution: a technology is developed if and only if it adds more value than it costs to develop, and that technology is priced, like all products in a competitive market, just below the marginal cost of its next best substitute (the prior technology) plus the value of the quality differential. The case of "perfect" competition thus provides a benchmark for evaluating performance in any particular case.

When the two above conditions do not hold, performance may suffer. For instance, if it is costly for consumers to switch to the new technology, and heterogeneous consumers face different costs and benefits from switching, the old technology will retain market share at a positive price. The new technology will sell at a markup higher than its quality differential over the old technology. Similarly, if both technologies are owned by the same firm, the implicit price of the old technology will not fall all the way to zero. Again, the new technology will sell at a higher price than the quality differential, although it may still be profitable for the firm to set relative prices so as to encourage migration to the new technology.

The resulting deviation from the perfect competition model is not necessarily inefficient to the extent it reflects real costs of learning and equipment. But if owners of the current technology are allowed to erect artificial barriers to the entry of a new technology, those suppliers will earn too much, opportunities for technical change will suffer, and consumers will be harmed.

One might expect something close to the result of the competitive model in operating systems because the industry appears characterized by *ex ante* barriers to entry that are low enough for these industries to be workably competitive (absent exclusionary practices). Given the combination of high fixed development costs and low marginal production and distribution costs, the competition resulting from entry can have a dramatic effect on the profits of the first mover. Not surprisingly, therefore, the incumbent has a strong incentive to make life difficult for subsequent entrants, either by directly increasing their costs or by reducing the attractiveness of their product to consumers, and to do so as soon as possible.

Under certain conditions, it may be possible for a first mover to maintain or even extend its dominant position through certain price and nonprice strategies that seek to exclude or handicap its smaller rivals in dealing with its immediate customers. The goal of such a strategy, rather than to assist in achieving the original large market share which requires having, at least for a while, the first-best technology, would be to artificially preserve that status. The four conditions described below appear to hold in the market for operating systems, where Microsoft successfully maintained an overwhelming market share against competition from a product regarded by many software experts as technically superior. The conditions are:

(1) Buyers at the next level downstream (i.e., OEMs), can be presented with an all-or-nothing choice by the dominant firm that compels them to deal either exclusively or not at all with the dominant firm;[33]

(2) While buyers would be interested in purchasing rivals' products for some of their requirements, they are unwilling to rely exclusively on those rivals' products: at least some of the dominant firm's product is important or even essential to many or even all the downstream firms;

(3) The substitute product requires significant fixed sunk costs to develop, maintain or expand, so that some significant minimum market share is essential for entry or expansion and the market is not contestable (substantial sunk costs are lost in a failed entry attempt); and

(4) The costs to the dominant firm of forcing exclusivity on the downstream firms are relatively low.

4.2 Microsoft's Pricing and Marketing Strategies

Let us now turn to each of the four conditions for exclusivity to be an effective strategy against smaller rivals. Our first condition was that immediate buyers can be presented with an all-or-nothing choice by the dominant firm that compels them either to deal exclusively with the dominant firm or not at all. Here, Microsoft can induce OEMs that wish to incorporate MS-DOS in any of their PCs to use MS-DOS exclusively through either of two policies. First, Microsoft can set per-unit MS-DOS prices that are so high relative to CPU rates as to make selecting the per-unit "option" economically infeasible: the OEM that wishes to use any MS-DOS will in effect be required to sign a CPU contract.[34] The CPU license (or a policy of inducing large carry forwards) then provides a strong economic incentive (a zero cost to the OEM for using MS-DOS at the margin) for the OEMs to use MS-DOS exclusively. Second, Microsoft can refuse to sell Windows to an OEM that purchases any alternatives to MS-DOS, and can cut off the OEM from technical information and other services provided to "favored" OEMs. This imposes a direct penalty on the OEM for using an alternative DOS in addition to the pricing incentive created by the CPU contract.

Our second condition was that, while buyers would be interested in purchasing rivals' products for some of their requirements, they are unwilling to rely exclusively on rivals' products: at least some of the dominant firm's product is very important or even essential to many or even all the downstream firms. In this case, OEMs are very reluctant to purchase OSs exclusively from sources other than Microsoft, at least in the short run, for several reasons. First, requiring a sudden and complete switch from one OS to another imposes real costs that could be avoided under a more gradual transition. Second, actual or threatened technical incompatibility between other Microsoft products, such as Microsoft Windows, and competing versions of DOS results in at least some of the OEM's customers insisting on MS-DOS. Third, withdrawal of Microsoft support services to any OEM that does not enter into a CPU contract (or that purchases DOS from a source other than Microsoft) would impose what is in effect a lump-sum penalty for switching. Finally, Microsoft can refuse to sell Microsoft Windows to an OEM unless that OEM also purchases MS-DOS through a CPU contract.

Our third condition was that the substitute product requires significant fixed sunk costs to develop, maintain or expand, so that some significant minimum market share is essential for entry or expansion. In this case, given the large non-sunk fixed costs of remaining in the DOS market, any alternative to MS-DOS must either achieve a critical minimum market share, exit the market, or be subsidized indefinitely through other operations of the rival firm.

Our final condition was that the costs to the dominant firm of forcing exclusivity on the downstream firms are relatively low. Here, the cost to Microsoft of excluding rivals from the DOS market is small as long as the share of those rivals remains small. By requiring a CPU contract, Microsoft runs the risk of losing an OEM's entire purchases to a competitor. As long as MS-DOS remains essential, however, no OEM would refuse the CPU contract, and the cost to Microsoft is minimal.

The cost to Microsoft of tying Microsoft Windows to MS-DOS is also very small. Microsoft might have to sacrifice some sales of Microsoft Windows to customers for whom the value of Microsoft Windows is very low, but who would buy it to use with a rival's OS but not with MS-DOS. But until a rival achieves a significant share of the OS market, tying of Windows to other Microsoft products or services (or simply making Microsoft Windows technically incompatible with any rival OS) would, again, impose minimal costs on Microsoft.

Our analysis thus concludes that, as compared with other strategies for maintaining market share—such as cutting prices or merging with entrants—implementing exclusionary practices can be a relatively cost-effective strategy against an entrant which has a superior technology but whose market share is very small. It can thus be characterized as a "fight them on the beaches" or (less kindly) as "economic infanticide." The higher the market share of the entrant, the greater the costs and the less the benefits of this strategy to the established firm. Once, or if, the entrant reaches a critical market share, however, the incumbent can be expected to switch to the alternative defensive strategies or, if the entrant's technology is strictly superior and user switching costs are not significant, to simply abandon the field.

REFERENCES

Baseman, Kenneth C., Frederick R. Warren-Boulton, and Glenn A. Woroch (1995). "Microsoft Plays Hardball: Use of Nonlinear Pricing and Technical Incompatibility to Exclude Rivals in Operating Software." *Antitrust Bulletin* 40, 265-315.

Bernstein Research (1993). "FTC Investigation of Microsoft." New York: Sanford C. Bernstein & Co.

Clark, Don and Laurie Hays (1994). "Microsoft's New Marketing Tactics Draw Complaints." *Wall Street Journal* (12 December), B5.

Goldman-Rohm, Wendy (1993). "Will the FTC Come to Its Senses About Microsoft's Mischief?" *Upside* 5, 10-27.

Manes, Stephen and Paul Andrews (1993a). *Gates: How Microsoft's Mogul Reinvented an Industry and Made Himself the Richest Man in America*. New York: Simon and Schuster.

Manes, Stephen and Paul Andrews (1993b). "Microsoft Monopoly." *Upside* 5, 10-18.

Ordover, Janusz and John Panzar (1982). "On the Nonlinear Pricing of Inputs." *International Economic Review* 6, 659-675.

Phang (1992). "Microsoft Deals Not Fair: Korea." *Asia Computer Weekly* (May), 11-17.

Sherer (1990). "Microsoft Outlines DOS 5.0 to Ward Off DR-DOS." *PC Week* 7, 4.

U.S. Department of Justice (1994). "Proposed Final Judgment and Competitive Impact Statement." 59 Fed. Reg. 42845 (proposed 19 August).

U.S. v. Microsoft Corp. (1995). No. 94-1564 (D.D.C. filed 15 July), unpublished.

U.S. v. Microsoft Corp. (1995). 56 F. 3rd 1448 (C.A.C.D.).

Warren-Boulton, Frederick R. (1977). "Vertical Control by Labor Unions." *American Economic Review* 67, 309-322.

Warren-Boulton, Frederick R., Kenneth C. Baseman, and Glenn A. Woroch (1995a) "Current Copyright Protection Can Make Economic Sense." *The Computer Lawyer* 12, 18-28.

Warren-Boulton, Frederick R., Kenneth C. Baseman, and Glenn A. Woroch (1995b). "The Economics of Intellectual Property Protection for Software: The Proper Role for Copyright." *StandardView* 3, 68-78.

NOTES

1. This paper has been adapted from a previously published article, see note 28 *infra*. We would like to express our appreciation for helpful comments and other assistance to Paul Dennis, David Gabel, Linnet Harlan and Sturge Sobin. We are also grateful to the participants of the Columbia Institute for Tele-Information's "Seminar on Sustaining Competition in Network Industries through Regulating and Pricing Access," especially Janusz Ordover and Bobby Willig, as well as to participants of a July 1995 session of the Western Economic Association conference, especially Ben Klein and Bob Levinson.

2. A PC also requires a layer of software that stands between the CPU and the OS. Called the BIOS, or basic instruction operating system, this code is burned into the machine's ROM (read only memory) chip.

3. OS/2 combines OS software and a GUI in one program.

4. If there are no cost complementarities across vintages of technology, then the requirement to sink substantial investment in software development will not convey an advantage to the firm that succeeded in the first generation of a technology when competes to develop subsequent generations.

5. There were two other significant MS-DOS derivatives. For a while, Compaq Computer had shipped its machines with its own Compaq DOS, and NEC developed NEC-DOS, a proprietary operating system that, until recently, dominated the Japanese market.

6. See Bernstein Research (1993), Exhibit 2, p. 10.

7. See Bernstein Research (1993), Exhibit 2.

8. Manes and Andrews (1993a). Citations are to the edited and condensed version in Manes and Andrews (1993b).

9. The percentage of Microsoft's operating system sales made under CPU agreements rose from 20 percent in FY 1989 to 22 percent in 1990, 27 percent in FY 1991 and 50 percent in FY 1992. By FY 1993, 60 percent of MS-DOS sales to OEMs and 43 percent of Windows sales to OEMs were covered under CPU agreements. See note 22 *infra*.

10. It is paid to Microsoft over the course of the year with an initial installment at the beginning of the year.

11. From every indication, the implicit per-unit charges and requirement levels vary across the contracts signed by different OEMs.

12. See *Ibid.*, Exhibit 3. In that year, Microsoft's sales of Windows through OEM and upgrades totaled $599 million.

13. Whether the unit is marginal or inframarginal, its value to the OEM equals the reduction in license fees when that unit is applied to next year's purchases. Of course, to assess its current value, one must discount for time and for the likelihood that the additional unit will actually be needed.

14. See *supra* note 10.

15. "The only error was that the customer was running Microsoft Windows on a competitor's version of DOS;" Manes and Andrews (1993b).

16. Microsoft refused to address compatibility problems with DRI. Microsoft boldly defended its action, claiming it had no responsibility to assist an operating systems competitor. Microsoft's actions went beyond refusal to assist a competitor, however, as it had engaged in commercial sabotage. See Rohm (1993).

17. At one time, the FTC staff was also investigating whether the relationship between Microsoft's operating systems and applications divisions created remediable competitive problems in markets for applications software.

18. *United States v. Microsoft Corp.*, No. 94-1564 (D.D.C. filed 15 July 1995). Amended versions of the Proposed Final Judgment and the Competitive Impact Statement were filed with the court on July 27, 1994.

19. Proposed Final Judgment and Competitive Impact Statement, 59 Fed. Reg. 42845 (1994) (proposed August 19, 1994).

20. The court noted that "Tunney Act courts are not mushrooms to be placed in a dark corner and sprinkled with fertilizer." Microsoft, 1995 U.S. Dist. LEXIS 1654, p. 42.

21. The actual price paid per unit could be thus higher even if the royalty fee itself incorporated volume discounts. Moreover, if units beyond the required minimum R are sold at a per-unit charge, the marginal price jumps from zero to a positive level once PC production exceeds R, so that purchases beyond the requirements level incur a quantity *premium* (see Table 2, above). Average price is the more typical yardstick for measuring nonlinearity of prices, and in the case of a CPU license, average price falls through the range up to the minimum requirements. Thereafter it may rise or fall depending on whether the average price at the requirements level is lower or higher, respectively, than the per-unit charge for additional sales.

22. It also requires that the lump-sum payment from each OEM be tailored so as to be less than the incremental profit that OEM earns from substituting MS-DOS in place of the next-best alternative.

23. For a discussion of the incentives of an input monopolist to substitute an output tax for above-marginal-cost pricing of the input when inputs can be used in variable proportions, see Warren-Boulton (1977).

24. For an extensive analysis of these two efficiency defenses for Microsoft's CPU license, see Baseman, Warren-Boulton, and Woroch (1995).

25. The core provisions are found in sections IV(H) and II(F) of the Proposed Final Judgment.

26. See Sections IV(F) and IV(H) of the Proposed Final Judgment.

27. See the section of the CIS on "Alternatives to the Proposed Final Judgment."

28. "The Department . . . does not have evidence that Microsoft has, to date, in fact structured its volume discounts to achieve anticompetitive ends." The Competitive Impact Statement.

29. The resulting pricing schedule left the Korean OEMs with essentially no option but to deal exclusively with Microsoft.

30. On December 12, 1994 *The Wall Street Journal* reported that in August, just after the consent decree was signed, Microsoft proposed a contract to Vobis, the German PC maker, that estimated its annual shipments of 88 models at about 475,000 and quoted a Windows price of $28 a copy based on that total. When the chairman of Vobis tried to negotiate a discount based on lower estimated sales, in order to accommodate customers that might ask for OS/2, Microsoft's response was that Vobis would have to pay $83 for each machine under a per-copy license.

31. The optimal degree of protection for intellectual property—in particular, the optimal scope for patent or copyright protection in the computer hardware and software industries—is a matter of considerable debate that we can only touch on here. This article is focused exclusively on the horizontal effects of Microsoft's practices, and so we do not express an opinion here as to the merits of the vertical aspects of the antitrust case against Microsoft. We have dealt with similar issues (i.e., network externalities, sunk investments by users, *de facto* standards and interface specifications) in an analysis of the proper role for copyright in software. See Warren-Boulton, Baseman, and Woroch (1995a) and (1995b).

32. In effect, firms earn a normal (i.e., competitive) risk-adjusted return on their investment, while the value of the underlying opportunity is passed on to consumers.

33. For this condition to hold, arbitrage among OEMs must be uneconomic.

34. Microsoft can also structure its Windows pricing to an OEM in such a fashion as to make it very difficult for OEMs to avoid a Windows CPU contract.

Rand Journal of Economics
Vol. 16, No. 1, Spring 1985

Standardization, compatibility, and innovation

Joseph Farrell*

and

Garth Saloner**

There are often benefits to consumers and to firms from standardization of a product. We examine whether these standardization benefits can "trap" an industry in an obsolete or inferior standard when there is a better alternative available. With complete information and identical preferences among firms the answer is no; but when information is incomplete this "excess inertia" can occur. We also discuss the extent to which the problem can be overcome by communication.

1. Introduction

■ Many goods are "compatible" or "standardized" in the sense that different manufacturers provide more interchangeability than is logically necessary. For instance, CBS and NBC television can be received on the same set; GTE Telephone subscribers can talk to AT&T subscribers; some—though far from all—computer programs written for one computer can be run on another; different manufacturers' nuts and bolts can be used together; and there are fewer types of sparkplug than there are models of automobile.[1]

It is clear that, other things being equal, there are important benefits of such standardization. That is presumably why government smiles on the development of such standards, for instance through the National Bureau of Standards, the British Standards Institute, etc.[2] Consumers benefit in a number of ways. There may be a direct "network externality" in the sense that one consumer's value for a good increases when another consumer has a compatible good, as in the case of telephones or personal computer software. There may be a market-mediated effect, as when a complementary good (spare

* GTE Labs and Massachusetts Institute of Technology.
** Massachusetts Institute of Technology.
Helpful discussions with Therese Flaherty and Julio Rotemberg and financial support from the National Science Foundation (grant SES-8308782) are gratefully acknowledged.

[1] Other examples of industrial standardization include plugs and sockets (not internationally standardized; and in the United States the "polarized" plug is making headway), typewriter keyboards, the ASCI character sets for computers, 35 mm. film, light bulbs, records and record players, etc. Some examples of commodities that might usefully be standardized, but are not, include: video cassette recorders, many auto parts, etc. A source of some interesting history is Hemenway (1975).

[2] The bulk of standardization, however, seems to be done through voluntary industry committees (Kindleberger, 1983). This encourages us in our interpretation of standardization as owing mainly to network externalities as felt by producers. It has also attracted at least some scrutiny by antitrust authorities (U.S. Federal Trade Commission, 1983).

parts, servicing, software . . .) becomes cheaper and more readily available the greater the extent of the (compatible) market. There may be a benefit to having a thicker second-hand (used) market. Finally, compatibility may enhance price competition among sellers.

All these except the last will feed back into producers' incentive to make their products compatible. In addition, some kinds of standardization will allow producers to get inputs more cheaply by exploiting economies of scale in the production of those inputs. In fact, most standardization is voluntary, rather than government-imposed, and comes about because of these "network externalities" among producers: other things being equal, a producer will often prefer to make his product compatible with his rivals'. This incentive does not, however, necessarily correspond exactly to social benefits.

Katz and Shapiro (1983) develop an oligopoly model in which consumers value a product more highly when it is "compatible" with other consumers' products. They call this effect "network externalities." In this framework they analyze the social and private incentives for firms to produce compatible products or to switch from incompatible to compatible products. They find, for example, that a dominant firm may choose to remain incompatible with a rival because it will suffer a substantial decline in market share if it becomes compatible, since that would increase the value to consumers of its rival's product.

Although standardization has important social benefits, as outlined above, it may have important social costs as well. Apart from the reduction in variety, which is unfortunate if different buyers would prefer different types of product, there is another possible cost, less well accounted for in the market, which is the subject of this article. Intuitively, it is plausible that the industry, once firmly bound together by the benefits of compatibility or standardization, will be inclined to move extremely reluctantly to a new and better standard because of the coordination problems involved. For example, Hemenway (1975) reports that the National Bureau of Standards declined to write interface standards for the computer industry because it feared that such standards would retard innovation. And many investigators believe that the standard "QWERTY" typewriter keyboard is inferior to alternatives such as the Dvorak, even when retraining costs are considered: the reason for its persistence is (supposedly) the overwhelming benefit from compatibility.[3] In this article we study the possibility that this "excess inertia" impedes the collective switch from a common standard or technology to a possibly superior new standard or technology.[4]

In Section 2 we study a simple model where it is common knowledge that the firms are identical, and where they decide sequentially whether to change to the new technology. A somewhat surprising result emerges: if all firms would benefit from the change, then all *will* change! In other words, there is no excess inertia impeding the change. Both unanimity and complete information are necessary for this result, however. We discuss the complete-information model with different preferences, but the focus of the article is on the incomplete-information model.

In Section 3 we allow for incomplete information about the "eagerness" of each firm to switch to the new technology. The equilibria that arise resemble bandwagons. Firms that strongly favor the change switch early, while those that only moderately favor wait

---

[3] David (1984) cites a U.S. Navy study which found that the payback period for retraining typists with the Dvorak keyboard was only ten days. This implies present values of time savings very much in excess of plausible costs for converting the physical stock of typewriters, especially since golfball typewriters would only require a new golfball and some stickers for the keys, while word-processing computers can also be cheaply converted.

[4] Arthur (1983) has modelled the evolution of a standard in an industry with network externalities and shows how, in a simple model, the realization of early random events can affect the standard chosen. Our work is concerned with the behavior of an industry that has already adopted a standard and is considering switching to a new one.

to see whether others will switch and then get on the bandwagon if it in fact gets rolling. If that happens, some who oppose the change will ultimately adopt it. Among those who first get on the bandwagon are some types of firms that will regret switching if in fact they are not followed. They sufficiently favor the change, however, to be willing to take that risk; the compensating benefit is the hope that they will precipitate the bandwagon effect.

In our model with incomplete information, we show that there is always excess inertia. Two types of excess inertia occur. In the first, and the most striking, which we call symmetric inertia, the firms are unanimous in their preference for the new technology and yet they do not make the change. This arises when all the firms only moderately favor the change, and hence are themselves insufficiently motivated to start the bandwagon rolling, but would get on it if it did start to roll. As a result, they maintain the *status quo*. In the second type of inertia ("asymmetric inertia") the firms differ in their preferences over technologies, but the total benefits from the switch would exceed the total costs. As before, this inertia arises because those in favor are not sufficiently in favor to start the bandwagon rolling.

Symmetric inertia is purely a problem of coordination. Hence, one might expect that, as in Farrell (1982), nonbinding communication of preferences and intentions may eliminate the inertia. We show in Section 4, however, that while this indeed eliminates the symmetric excess inertia, it exacerbates the problem of asymmetric inertia.

In Section 5 we present our conclusion and suggest avenues for future research.

2. A model with sequential decisions and complete information

■ One of the clearest features of noncooperative[5] standard setting is its bandwagon quality. When compatibility is an important consideration for a firm setting its product specifications, early movers can influence later movers' decisions: if firm 1 switches to a new standard, then firm 2 will find switching more attractive than if firm 1 had not switched. In this section we present a simple model of that effect in which firms' decisions are taken sequentially, and payoffs are common knowledge. We show that if, allowing for transition costs, all firms would prefer the industry to switch, then the only perfect equilibrium is that they all do so.

While the sequential timing may seem artificial, we can show that every equilibrium we derive is also an equilibrium in the simultaneous-move game. Moreover, the equilibrium if timing is endogenous (see below) is one of those we now consider.

Let $N = \{1, 2, \dots, n\}$ denote the set of firms in the industry. For any $j \in N$ and any $S \subseteq N$ containing j, we define $B_j(S, Y)$ as the net benefit to firm j from switching, together with the other firms in S, from the old standard (X) to the new one (Y), relative to its benefit if all firms stick with X. In other words, we normalize so that each firm gets zero benefit in the *status quo*. Then $B_j(S, Y)$ is the value to j of switching and having the other members of S switch. This is a present value, and net of any transition costs. Thus, firm j would favor a change by the entire industry if and only if $B_j(N, Y) > 0$.

We also define $B_j(S, X)$ for subsets S containing j, as j's payoff if j and the other members of S stay with X, while the members of $N \setminus S$ switch to Y. Thus, in particular, $B_j(N, X) = 0$ by normalization.

The basic *assumption of positive network externalities* can now be phrased by Assumption 1.

[5] To be clear, what we have in mind is that those producers who adhere to the standard do so purely because others do so. There is neither a standard-enforcing authority nor a system of binding though voluntary contracts to adhere to standards, though both of these possible institutions would be interesting to analyze.

Assumption 1. If $j \in S \subseteq S'$ and $k = X$ or Y, then $B_j(S, k) \leq B_j(S', k)$.

This says that, whatever j's choice, he prefers to have others make the same choice. This introduces the coordination considerations that are the focus of this article.

□ **Symmetric case.** In some of the work below, we assume that $B_j(S, X)$ depends only on the number of firms in S, and likewise for $B_j(S, Y)$. Thus, we can write the benefit functions as $B_j(m, k)$, where m is the total *number* of firms in S, i.e., the number making the choice that j makes. Moreover, we shall sometimes assume that the function $B_j(\cdot, \cdot)$ is the same for all j, so we can simply write $B(m, X)$ or $B(m, Y)$.

□ **The model.** The set N of firms is given, as are the alternative standards X and Y. All firms are initially at standard X. There are n periods to the game, which has perfect and complete information. (Since one firm has a decision each period, the number of periods is equal to the number of firms.) In period j, firm j decides whether to switch to Y. If S denotes the set of firms that do switch, then the payoffs are

$$B_j(S, Y) \quad \text{for} \quad j \in S$$

$$B_j(N \setminus S, X) \quad \text{for} \quad j \notin S.$$

Proposition 1. Suppose that, for each j,

$$B_j(N, Y) > B_j(\{j, j+1, \ldots, n\}, X). \tag{1}$$

Then the unique perfect equilibrium involves all firms' switching.

Proof. The condition (1) ensures that, for each j, if $1, \ldots, j-1$ have already switched, then j prefers to switch (if he believes all the rest would follow) rather than to stay (*whatever* his beliefs about how many others would then switch). Since j knows this is true for $j+1, \ldots, n$, he knows they will switch if he does; and so he will switch.

Notice that Proposition 1 does not use Assumption 1. Using that assumption yields the following result.

Corollary. If

$$B_j(N, Y) > B_j(N, X) \text{ for all } j, \tag{2}$$

then the unique perfect equilibrium involves all firms' switching. Therefore, *in this model, there can be no excess inertia in the symmetric sense that each firm prefers an overall industry switch but it fails to happen.*

Condition (1) is weaker than unanimity (2), however. So Proposition 1 tells us that players j, late in the game, sometimes switch, even though $B_j(N, Y) < B_j(N, X)$. Moreover, it is clear that there is no necessary relationship between $\sum_j [B_j(N, Y) - B_j(N, X)]$ and the outcome of the game: we can find excess inertia or its opposite if we make judgments based on adding benefits.

Being late in the game is a strategic disadvantage because of our assumption that each agent has only one chance to choose his standard; thus, early movers are able to commit. In a game of complete information, there is no countervailing value to waiting to see how things evolve. This is expressed by the following result.

Proposition 2. Given the preferences of all agents, each agent is better off (not necessarily strictly) moving earlier than moving later.[6]

Proposition 2 is proved in the Appendix. It uses only the presence of network

[6] From the timing of political primaries, this might be called the New Hampshire theorem.

externalities—Assumption 1. The essence of the proof is that having an earlier position gives power over later movers, and hence even earlier movers are obliged to treat one's preferences with more respect.

Intuitively, there is a benefit of commitment from moving early. In a general game, there can be a countervailing factor of "regret": once a von Stackelberg follower has moved, the leader would like to change his move, if he could.[7] In this game that does not happen: every sequential equilibrium would also be an equilibrium if firms decided simultaneously on their choices. The other factor which sometimes makes it desirable to move later in some other games, i.e., the fact that information may flow in, is also absent from this model, but is addressed in Section 3.

A simple example in which Proposition 2 holds strictly is provided by the following two-firm case:

Firm A

	$B_A(m, X)$	$B_A(m, Y)$
$m = 1$	-2	-1
$m = 2$	0	1

Firm B

	$B_B(m, X)$	$B_B(m, Y)$
$m = 1$	-2	-3
$m = 2$	0	-1

If A moves first, then he will switch and B will follow. If B moves first, however, he will not switch, and A will then not switch. It is easy to check the claim of Proposition 2 that each firm prefers the outcome that results from its moving first.

□ **Endogenous timing and a bias for switching.** Hitherto, we have had no essential strategic difference between X and Y, once switching costs were netted out from the benefits of Y. Each firm in turn could commit itself to X or to Y. We now discuss what will happen if a choice of Y is irreversible, while a choice "remain at X" is not. One reason this might be true is that remaining at X means a continuing and gradual replacement of plant, worker skills, etc., while a switch to Y, or a reversion to X, would involve a much greater cost. If this switching cost is substantial, a switch to Y will be seen as at least somewhat of a commitment, while remaining at X enables a firm to keep its options open. With this assumption we can remove the artificial assumption that firms make their decisions in a prespecified order. Instead, those who wish to choose Y go first, in effect. In view of Proposition 2, this will bias the outcome towards Y, in the sense that among the specified-order equilibria it is the one most inclined to Y that will occur.

To make this precise, we introduce the following notation. Let e be any perfect equilibrium with a prespecified order of moves. Write $S(e)$ for the set of firms that switch to Y in that equilibrium. Now define S^* to be the union of all the sets $S(e)$, where e ranges over all possible orders of moves.

Proposition 3. When timing is endogenous as above, then all firms in S^* switch to Y.

The proof of Proposition 3 is in the Appendix. Notice that Proposition 3 implies that with this form of endogenous timing, if all firms favor a switch ($B_j(N, Y) > 0$ for all j), then they will all switch. If no firm favors a switch, none will switch. But in intermediate cases there is a bias for switching.

[7] For example, in "matching pennies," moving first would be a disadvantage.

3. A model with incomplete information

■ The analysis of the previous section relies heavily on the assumption of complete information. This assumption seems somewhat unrealistic, however, especially in view of its strong implications. In reality, a firm will generally be uncertain whether it would be followed if it switched. In this section we study a somewhat different model in which we represent that uncertainty as incomplete information about the other firm's preferences. We also allow for endogenous timing of moves, as above, but find that in conjunction with the incomplete information this yields a richer set of possibilities than Proposition 3 would suggest.

Since we are explicit about incomplete information and differences among firms, we can write the benefit function as $B^i(\cdot, \cdot)$, where i denotes a firm's type, and where there is now no need to subscript $B(\cdot, \cdot)$, since any differences are captured in different values of i. Higher values of i will be taken to indicate stronger preferences for the change to technology Y. We take the set of types to be the unit interval, and we assume that all types are *a priori* equally probable, i.e., types are distributed uniformly on [0, 1]. (These assumptions are not restrictive and considerably simplify the exposition.) We also restrict attention to the two-period, two-firm case, although we shall see that having more than two periods would not change the results.[8]

There are thus two periods, 1 and 2, and each firm can switch at time 1 or time 2 or not at all. As in Section 2 we rule out reswitching. As we show in footnote 9, however, the equilibrium which we develop below with this assumption also has the property that no firm that switches in period 1 would want to revert.

If we let S denote the action "switch" and let D denote "do not switch," a strategy for player j can be described by the pair

$$\sigma_1^j: [0, 1] \to \{S, D\} \qquad \text{and} \qquad \sigma_2^j: [0, 1] \times \{S, D\} \to \{S, D\},$$

i.e., the second-round move is conditioned on the player's own type and the opponent's first-period move. Here σ_t describes the strategy for period t and maps the set of player types and history of play to date into the possible actions the firm can take. (Strictly speaking σ_2^i should be conditioned also on whether player i switched at time 1. A player who did switch at time 1 has no further decisions to make, however, and hence σ_2^i can be simplified as above without ambiguity.)

We make the following assumptions, which are illustrated in Figure 1:

Assumption 1. $B^i(2, k) > B^i(1, k)$, $k = X$ and Y.

Networks are beneficial. (This is Assumption 1 of Section 2, rephrased for the current setting.)

Assumption 2. $B^i(2, Y)$ and $B^i(1, Y)$ are continuous and strictly increasing in i.

---

[8] With n firms, suppose that there were $m > n$ periods, and in period i there were both positive probability that some would switch and positive probability that none would switch. With symmetric strategies, if none switched in period i, then every firm would become uniformly more pessimistic about others' willingness to switch, and therefore (having decided against switching at period i) would never switch. If a firm were going to switch after receiving the bad news, this would mean it was going to switch anyway, but the strategy of waiting is dominated by switching immediately. This means that n periods suffice to analyze the n-firm case.

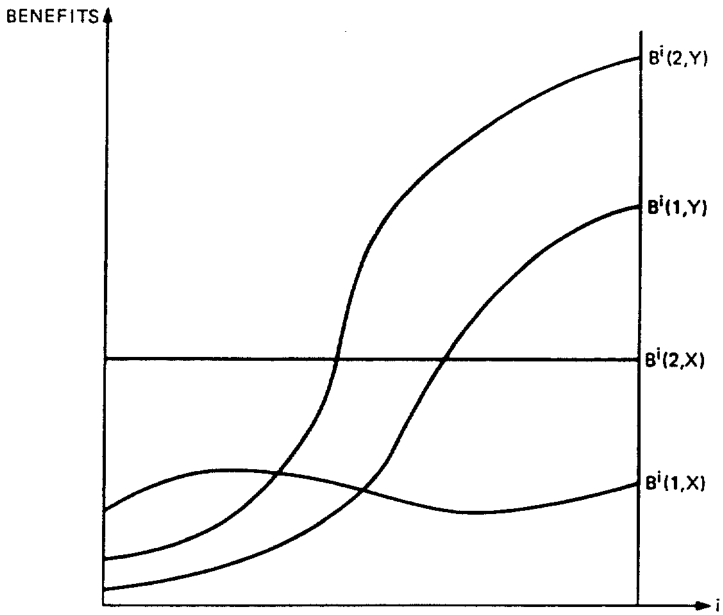
[9] We could assume that it is prohibitively costly to switch back to X in period 2 after switching to Y in period 1. See the brief discussion in Section 2. We could alternatively investigate the condition on the B function to ensure that such reswitching would never occur: anticipating the notation about to be developed, a sufficient condition is

$$\frac{i^* - \bar{i}}{i^*} B^{i^*}(2, Y) + \frac{\bar{i}}{i^*} B^{i^*}(1, Y) \geq \frac{i^* - \bar{i}}{i^*} B^{i^*}(1, X) + \frac{\bar{i}}{i^*} B^{i^*}(2, X).$$

Using the definition of i^*, and the fact that $B^{i^*}(2, X) = 0 \geq B^{i^*}(1, X)$, we can see that this condition is always satisfied. We do not fully understand this remarkable conclusion.

FIGURE 1

EXAMPLES OF BENEFIT FUNCTIONS THAT SATISFY THE ASSUMPTIONS OF THE MODEL



This assumption captures what is meant by a "type": higher types (indexed by higher values of i) are more eager to switch to Y, both unilaterally and if the other firm also switches.

Assumption 3. $B^1(1, Y) > 0$ and $B^0(2, Y) < B^0(1, X)$.

Unilateral switching is worthwhile for at least one possible type of firm, and (at the other end of the spectrum) there are some types who would rather remain alone with the old technology than join the other firm with the new technology. This assumption also implies that for intermediate values of i, a firm's decision will at least sometimes depend on its predecessor's decision: this is what makes the model interesting.

Assumption 4. $B^i(2, Y) - B^i(1, X)$ is monotone in i.

If a firm of type i' prefers a combined switch to Y to remaining alone with technology X, then so do all firms with $i > i'$. In other words, if i' would follow a lead, then so would $i > i'$.

A helpful analogy is a political "bandwagon" effect. Politicians considering what position to take on an issue are concerned not only with how strongly they feel about it, but perhaps also with how likely it is that their stand will become the majority view. Intuitively, we might expect vigorous opponents to oppose the issue regardless of their expectations. Staunch supporters might commit themselves without waiting to see whether it seems that theirs will become the popular view. A more "political" middle group may wait awhile to test the political waters, declaring themselves to be "for" the measure if the bandwagon begins to roll and "against" otherwise. Thus a "*bandwagon strategy*" for a firm can be defined by a pair $(i^*, \bar{i})$ with $i^* > \bar{i}$ such that: (i) if $i \geq i^*$, the firm switches at time 1; (ii) if $i^* > i \geq \bar{i}$, the firm does not switch at time 1, and then switches at time 2 if (and only if) the other firm switched at time 1; and (iii) if $i < \bar{i}$, the firm never switches.

A "*bandwagon equilibrium*" is defined to be a perfect Bayesian Nash equilibrium in which each firm plays a bandwagon strategy. In what follows we shall concentrate on symmetric bandwagon equilibria, i.e., those for which $(\bar{i}, i^*)$ is the same for each player. Asymmetric bandwagon equilibria only exist for some specifications of the benefit functions, and will come in mirror-image pairs if they occur. Accordingly, we expect them not to be focal. On the other hand, using only the fairly weak Assumptions 1–4, we show below that a unique symmetric bandwagon equilibrium exists and that there are no equilibria that are not bandwagon equilibria.

First, let $\bar{i}$ be defined by $B^{\bar{i}}(1, X) = B^{\bar{i}}(2, Y)$. Thus, any firm with type $i < \bar{i}$ would prefer remaining with the "old" technology to switching to the "new" technology, even if the other firm switched. Clearly, such a firm will never switch. On the other hand, a firm with $i > \bar{i}$ would switch in the second period if the other firm had already switched (and assuming that switching back is known to be precluded). This essentially describes behavior in the second period.[10] Using this, we can now analyze the first period.

Define $f(i) = iB^i(2, Y) - \bar{i}[B^i(2, Y) - B^i(1, Y)]$. Let $I = \{i: f(i) = 0\}$.

Lemma 1. (a) $f(i) < 0 \ \forall \ i \leqslant \bar{i}$; (b) $f(i)$ is strictly increasing in $i \ \forall \ i > \bar{i}$; (c) $f(1) > 0$; (d) I contains exactly one point (which we call i^*); and (e) $i^* \in (\bar{i}, 1)$.

Proof. (a) For $i \leqslant \bar{i}$, $\bar{i}B^i(2, Y) \geqslant iB^i(2, Y)$. Also $\bar{i}B^i(1, Y) < \bar{i}B^{\bar{i}}(1, Y) < \bar{i}B^{\bar{i}}(2, Y) = \bar{i}B^{\bar{i}}(1, X) < B^{\bar{i}}(2, X) = 0$. So $(iB^i(2, Y) - \bar{i}B^i(2, Y)) + \bar{i}B^i(1, Y) < 0 \ \forall \ i \leqslant \bar{i}$. (b) Immediate since $(i - \bar{i}) > 0$ for $i > \bar{i}$ and since $B^i(2, Y)$ and $B^i(1, Y)$ are strictly increasing. (c) $f(1) = B^1(2, Y)[1 - \bar{i}] + \bar{i}B^1(1, Y)$. But $\bar{i} < 1$ (since $B^1(2, Y) > 0$) and $B^1(2, Y) > B^1(1, Y) > 0$. (d)–(e) Since $f(i)$ is strictly increasing and continuous on $(\bar{i}, 1]$ with $f(\bar{i}) < 0$ and $f(1) > 0$, there exists exactly one $\bar{i} < i^* < 1$ for which $f(i^*) = 0$.

Lemma 2. $B^{i^*}(1, Y) < 0$ and $B^{i^*}(2, Y) > 0$.

Proof. $i^*B^{i^*}(2, Y) = \bar{i}B^{i^*}(2, Y) - \bar{i}B^{i^*}(1, Y)$ by the definition of i^*. Therefore $B^{i^*}(1, Y) = (\bar{i} - i^*)B^{i^*}(2, Y)/\bar{i}$. Now $\bar{i} > 0$ and $\bar{i} < i^*$ imply that $B^{i^*}(2, Y)$ and $B^{i^*}(1, Y)$ have opposite signs. But then $B^i(2, y) > B^i(1, Y)$ gives the result.

These lemmas are illustrated in Figure 2.

We can now prove the following.

Proposition 4. With $\bar{i}$ and i^* as defined above, a unique symmetric bandwagon equilibrium exists.

Proof. There are three actions to consider:

a_1: switch at time 1
a_2: switch at time 2 if and only if opponent switched at time 1
a_3: do not switch at time 2 even if opponent switched at time 1.

(There is a fourth possible action, a_4: switch at time 2 if opponent did not switch at time 1, but this is dominated by a_1.)[11]

Let $u^i(a_j)$ be the expected benefit to a firm of type i when it uses action a_j and when its opponent is using the bandwagon strategy $(\bar{i}, i^*)$. The proof proceeds in three steps:

(i) For $i > \bar{i}$, $u^i(a_1) - u^i(a_2)$ has the sign of $i - i^*$:

$$u^i(a_1) = B^i(2, Y)(1 - \bar{i}) + B^i(1, Y)\bar{i};$$

$$u^i(a_2) = B^i(s, X)i^* + (1 - i^*)B^i(2, Y) = (1 - i^*)B^i(2, Y).$$

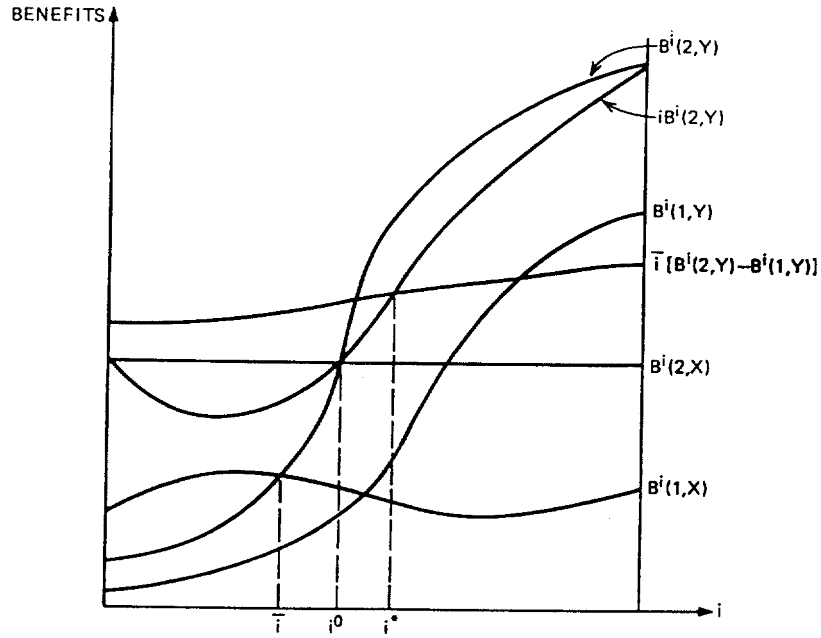
---

[10] The only thing left to specify is what happens in the second period if neither firm switched in the first. We show below (Proposition 5) that neither firm will switch. See also footnote 11.

[11] If a firm's opponent is of a type below $\bar{i}$, a_1 and a_4 yield the same payoff $B^i(1, Y)$. If the opponent is of a type above $\bar{i}$, a_1 yields $B^i(2, Y)$, whereas one can easily show that a_4 gives a positive probability of $B^i(1, Y)$, and complementary probability of $B^i(1, Y) < B^i(2, Y)$. This concludes the argument.

FIGURE 2

AN ILLUSTRATION OF THE DERIVATION OF THE CRITICAL LEVELS $\bar{i}$, i^0, AND i^*



Therefore, $u^i(a_1) - u^i(a_2) = f(i)$. The result follows from Lemma 1.

(ii) $u^i(a_2) - u^i(a_3)$ has the sign of $i - \bar{i}$:

$$u^i(a_2) = B^i(2, Y)(1 - i^*);$$

$$u^i(a_3) = B^i(1, X)(1 - i^*).$$

The result follows from Assumption 4 and the definition of $\bar{i}$.

(iii) If $i \leq \bar{i}$, a_3 is a dominant strategy. If $i > \bar{i}$, a_2 is preferred to a_3 (from (ii)) and if $i > i^*$, a_1 is preferred to a_2 (from (i)). Therefore, the bandwagon strategy $(\bar{i}, i^*)$ is the unique best response to the bandwagon strategy $(\bar{i}, i^*)$.

Finally, a symmetric equilibrium has $f(i^*) = 0$ by step (i). But then Lemma 1 implies that there is a unique symmetric bandwagon equilibrium. This proves Proposition 4.

Several features of the equilibrium can be observed directly from Figure 2. As Lemma 2 shows, there is a region below i^* where nonetheless $B^i(2, Y) > 0$. If both firms are of types that fall into this region, the switch will not be made, although it would have been made in a world of complete information and although both firms would then be better off. There is symmetric excess inertia! The intuition is clear. Both firms are fencesitters, happy to jump on the bandwagon if it gets rolling but insufficiently keen to set it rolling themselves.

In addition, there is also asymmetric excess inertia. One firm may be of the kind discussed above ($B^i(2, Y) > 0$ and $i < i^*$), but the other firm may have $B^i(2, Y) < 0$. There will *always exist* some cases where $B^i(2, Y) + B^{i'}(2, Y) > 0$ and $i, i' < i^*$. Here again the switch will not be made even though the sum of the benefits is positive. Finally, it is possible that the switch will be made even though the sum of the benefits is negative. This occurs when one of the firms favors the switch and, although the other opposes it strongly, the latter prefers switching to remaining alone with the old technology. Excess "momentum" of this kind will not always exist, but can occur for appropriately specified benefit functions.

Notice too that there are some types in the region $i^* > i > \bar{i}$ for which $B^i(2, Y) < 0$. These firms will switch if the other firm switches, but would have preferred that the new technology had not come along at all. If polled about their intentions *ex ante*, they would vehemently claim that they would not switch even if the other switched.[12] This motivates examining the question of communication, to which we turn in the next section.

There are also some types just above i^* for which $B^i(1, Y) < 0$. These types start the bandwagon rolling, but if it turns out that the other firm was of a type below $\bar{i}$ (so that their lead is not followed), they regret their decision *ex post*. Here, again, there is a straightforward intuition. Types in this range sufficiently favor technology Y that they risk starting the bandwagon even though they know with positive probability that they are up against an "intransigent" with type less than $\bar{i}$ and will end up worse off if this turns out to be so.

There are a number of interesting comparative static results. Consider increasing $B^i(2, Y)$ or decreasing $B^i(1, X)$ until $B^0(2, Y) > B^0(1, X)$ (removing Assumption 3), so that every type of firm would follow if the other firm switched. In that case $\bar{i} = 0$ and so $f(i) = iB^i(2, Y)$. Therefore, i^* is defined by $B^{i^*}(2, Y) = 0$. This means that in equilibrium if the switch is beneficial for both firms, they will both switch at time 1! Thus, in the absence of the intransigents with $i < \bar{i}$, symmetric excess inertia disappears. In addition, (trivially) the inertia that arises when only one firm favors the switch also disappears here. Excess momentum can, however, still arise. This bias in favor of switching arises from the assumption that switching back from Y to X cannot occur, just as in Proposition 3.

As one would expect, as $B^i(1, Y)$ increases towards $B^i(2, Y)$, i^* decreases until the point defined by $B^{i^*}(2, Y) = 0$. As $B^i(1, Y)$ decreases, i^* increases, and tends to 1 as $B^i(1, Y)$ becomes sufficiently low.

Finally we demonstrate that there are no equilibria that are not bandwagon equilibria.

Proposition 5. Any equilibrium strategy is a bandwagon strategy.

Proof. First, we have

$$\sigma_2(S, i) = \begin{cases} S & \text{if} \quad i \geq \bar{i} \\ D & \text{if} \quad i < \bar{i} \end{cases}$$

by perfectness. Further, $\sigma_2(D, i) = D$ for all i (see footnote 11). Consider firm 1's decision. Suppose it assesses probability $1 - q$ that firm 2 will switch at time 1. Then, if it waits until time 2, it earns $B^i(2, Y)(1 - q) + B^i(2, X)q = B^i(2, Y)(1 - q)$. If it switches at time 1, it earns $B^i(2, Y)(1 - \bar{i}) + \bar{i}B^i(1, Y)$. It pays to switch if

$$B^i(2, Y)q - \bar{i}[B^i(2, Y) - B^i(1, X)] \geq 0,$$

which is monotone in i. Therefore, if it is optimal for any type i' to switch at time 1, then it is also optimal for any higher type i'', $i'' > i'$. So any optimal strategy involves a cutoff at time 1. But then any equilibrium strategy is a bandwagon strategy.

4. The model with incomplete information and communication

■ The analysis of the previous section shows that incomplete information introduces excess inertia in which the new technology is not adopted even when adoption is favored

[12] The purpose of this lie would be to dissuade the other from switching, if the other had

$$B^i(2, Y) > 0 > B^i(1, Y).$$

by *both* firms. It seems plausible that allowing even a minimal amount of coordination between the firms would eliminate such "symmetric" or "Pareto" inertia. In particular, if we allow a single public statement by each firm as to whether it favors the switch before any actions are taken, this problem disappears. Any type i firm for which $B^i(2, Y) > 0$ would have no incentive to hide this fact and could be expected to announce truthfully. If both firms so announced, we would expect technology Y to be adopted. Similarly, any type of firm with $i < \bar{i}$ could be relied on to reveal its type truthfully. Only those types of firm for whom $B^i(2, Y) < B^i(2, X)$ and $B^i(1, X) < B^i(2, Y)$ should be expected to misreport. This is the group that would "jump on the bandwagon" once it got rolling but that would rather the bandwagon had not started rolling at all.

Formally, we model this by adding a period to the beginning of the two-period model of the previous section. At time 0 each firm (simultaneously) announces F or A ("for" or "against") the switch.[13] Time 1 and time 2 are then as before.

A strategy now stipulates for each type of firm what announcement to make and whether to switch at times 1 or 2 (as a function of all available information). We shall demonstrate below that the following strategies constitute a perfect Bayesian Nash equilibrium to this game with communication:

(1) Announce F if and only if $i \geq i^0$, where i^0 is defined by $B^{i^0}(2, Y) = 0 = B^{i^0}(2, X)$; i.e., if and only if $B^i(2, Y) \geq 0$.
(2) If both firms announce F, both switch at time 1.
(3) If both firms announce A, neither switches at time 1 nor time 2.
(4) If one firm announces F and the other announces A, employ a bandwagon strategy $\{i', \bar{i}\}$, where $\bar{i}$ is as before and i' is defined by $B^{i'}(2, Y)i^0 = \bar{i}[B^{i'}(2, Y) - B^{i'}(1, Y)]$.

The only part of the description of equilibrium that requires explanation is part (4). We provide a discussion rather than a formal proof that would largely mimic the proofs of Propositions 4 and 5.

The major change from the no-communication case is in each firm's subjective probability assessment that it will be joined if it initiates a switch. Previously, this was merely the probability $(1 - \bar{i})$. Now, however, if the other firm has announced "A," this probability is given by Prob $\{i \geq \bar{i} | i < i^0\} = (i^0 - \bar{i})/i^0 = 1 - \bar{i}/i^0$. Since $i^0 < 1$, we have $(i^0 - \bar{i})/i^0 < (1 - \bar{i})$. This merely says that a firm is more pessimistic that it will be joined in a switch if the other firm has announced "A."

In showing that these strategies form an equilibrium, a typical calculation is the following: should a type $i > i'$ deviate to a strategy of switching at time 2, if the other firm switches at time 1, from its proposed strategy of switching at time 1? Under its current strategy it earns

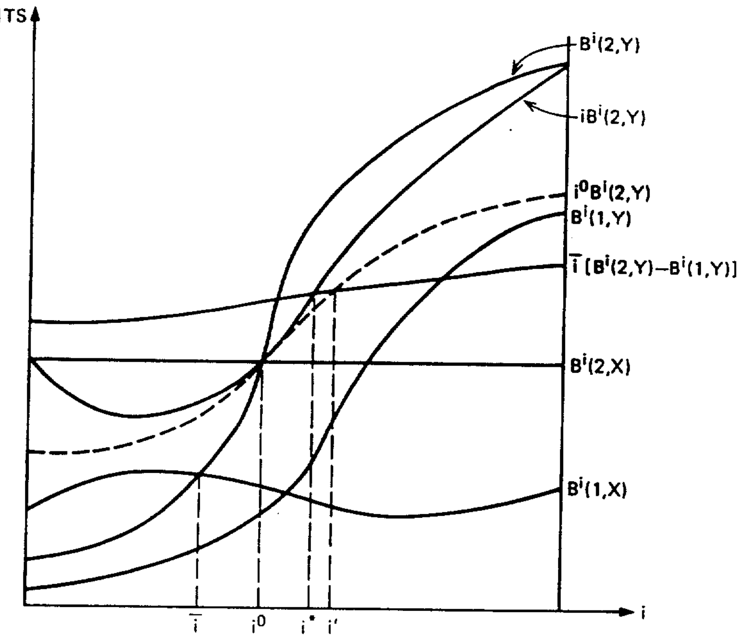
$$B^i(2, Y)\,\Pr\{j > \bar{i} | j < i^0\} + B^i(1, Y)[1 - \Pr\{j > \bar{i} | j < i^0\}]$$
$$= B^i(2, Y)(1 - \bar{i}/i^0) + B^i(1, Y)(\bar{i}/i^0)$$
$$= B^i(2, Y) - \bar{i}/i^0[B^i(2, Y) - B^i(1, Y)].$$

If it deviates to the alternative suggested strategy, it earns $B^i(2, X) \equiv 0$ with certainty (since the opponent has announced N). The deviation pays if and only if $B^i(2, Y)i^0 < \bar{i}[B^i(2, Y) - B^i(1, Y)]$. It is this that motivates the definition of i' given above. This is illustrated in Figure 3.

[13] A more elaborate—even a multistage—system of communication before play begins would reduce to this in effect. The reason is that each player either wants to encourage the other to switch, or wants to discourage him, and this preference depends only on the player's own type, not on the other's. Thus we get "bang-bang" communication strategies: one either chooses the most encouraging or the most discouraging communication strategy. Thus, there is no need to consider more than two communication strategies.

FIGURE 3

THE EFFECT OF COMMUNICATION ON THE BANDWAGON EQUILIBRIUM



Notice that for all $i > i^0$, $iB^i(2, Y) > i^0B^i(2, Y)$. Therefore $i' > i^*$. This means that a strictly smaller set of types will initiate the switch in the communication case when announcements differ than in the no-communication case. As a result, "excess momentum" and "excess inertia" of the kinds that arise when the preferences of the firms differ are respectively lower and higher than in the no-communication case. Thus, while communication in the form of a "straw vote" eliminates excess inertia where the preferences of the firms coincide, it increases inertia where the preferences differ.

5. Conclusion and further directions

■ In this article we have analyzed the problem of coordinating innovation or a change of standard in an industry in which products not compatible with others are at a substantial disadvantage. We have shown that there can be inefficient inertia, or inefficient innovation, and that these problems cannot be entirely resolved by communication among firms.

Some important topics which we have left untouched, but which would be appropriate for further work, are the following.

(1) In reality, a standard is often a more complex object than we have implicitly assumed by supposing that a firm either "adopts" or "does not adopt." In particular, compatibility need not be symmetric: for example, a computer company can try to arrange that the software written for its competitors' machines will run on its machines, but not *vice versa.* A somewhat similar contest produced the peculiarly shaped holes in old fashioned safety razor blades. This, of course, represents an attempt to get network externalities for oneself while denying them to competitors.

(2) The literature on optimal product diversity (Salop, 1979) assumes that the benefits from standardization come from production economies of scale. It would be interesting to analyze the tradeoff with variety if the benefit from reduction in variety came from consumer-side network externalities.

(3) All our models above are timeless in the sense that, in the end, payoffs are determined only by *who* has adopted a standard, not by *when* the standard was adopted. In some cases there may be benefits to early adoption of what later becomes an industrywide standard: the first-mover advantage. On the other hand, it may be costly to be incompatible with the majority of firms in the industry for the length of time it takes for them to follow; and, of course, there is a possibility that they may not follow. Thus, even apart from bandwagon effects, timing becomes an interesting issue. (For some related work, see Wilson (1984)). To address these issues of timing, Rohlfs (1974) considered an adjustment process in which (in contrast to the present work) consumers choosing whether to subscribe to a communications service with network externalities make their decisions on the basis of *current* payoffs. He exhibits multiple equilibria and critical-mass phenomena, analogies to which could also be drawn here. Dybvig and Spatt (1983) develop and analyze government incentive schemes to deal with the externalities that arise in a model like that of Rohlfs.

(4) It is widely believed that "large" firms have a great deal of strategic power in the kind of *de facto* standard-setting we analyze here. This can be examined in the context of our model: a large firm's customers experience relatively little change in their payoff when other firms decide whether to be compatible with the large firm. By contrast, the large firm's decision substantially affects the payoffs to buyers of other firms' products. It is an open question whether this concentration of power leads to distortions in the industry's choice of technology.

We are studying these and related issues, and we believe there are many other interesting questions to be investigated in the area.

Appendix

■ The proofs of Propositions 2 and 3 follow.

Proof of Proposition 2. We begin by proving three lemmas to get Proposition 2.

Lemma A. If $n = 2$, each firm (nonstrictly) prefers to go first.

Proof. Call the firms A and B, and their decisions (X or Y) k_A and k_B. Let (k_A, k_B) be an equilibrium when A goes first. Then, as pointed out in the discussion following the statement of Proposition 2, k_A is also A's best response to k_B. Therefore, B can achieve his payoff from (k_A, k_B), if he moves first, simply by choosing k_B as before. Of course, B may be able to do better by making another choice when he goes first.

Lemma B. Whatever n may be, any firm would (nonstrictly) rather be #1 than #2.

Proof. This follows from Lemma A, if we collapse the responses of firms $3, 4, \ldots, n$ into the payoffs for firms A and B, which are trading places 1 and 2. All that needs to be checked is that the reduced game continues to satisfy Assumption 1, and that is clear.

Lemma C. For any n, and any $j = 1, 2, \ldots, (j - 1)$ a firm in position $(j + 1)$ would (nonstrictly) like to trade places with the firm in position j.

Proof. Lemma B assures us that this would be true if we could think of the actions of $1, 2, \ldots, (j - 1)$ as not responding to the change. We then must show that any response by the early players will be favorable to the firm (call it B), which has switched from $(j + 1)$ to j.

The reason this is true is that the switch has made the consolidated response function of players $j, j + 1, \ldots, n$ (considered together) more in line with B's preferences (Lemma 1). Therefore, players $1, \ldots, (j - 1)$, considered as playing a game with the responses of $j, j + 1, \ldots, n$ collapsed into the payoff functions, have had their preferences shifted in the direction of B's desires.

Proposition 2 now follows by repeated application of Lemma C. That is, to show that, given the order of the other $(n - 1)$ firms, a firm prefers to be earlier in that sequence rather than later, one simply imagines the firm's repeatedly moving up one place and bumping its predecessor one place down (as in progress up a squash ladder). This proves Proposition 2.

Proof of Proposition 3. We actually prove a stronger version of Proposition 3:

(i) Consider the game with two fixed orderings of moves. Let e_1 and e_2 be perfect equilibria of the games corresponding to those orderings. Let $S(e_1)$ be the set of firms that switch in equilibrium e_1 and let $S(e_2)$ be those that switch in e_2. Then there exists another order of moves with its perfect equilibrium e, such that $S(e_1) \cup S(e_2) \subseteq S(e)$.

(ii) There exists an order of moves giving a perfect equilibrium e^* such that $S(e^*)$ is the union of all sets $S(e)$ for equilibria e.

(iii) If moves are in endogenous order, then the set $S(e^*)$ of firms will switch to Y.

Proof. Begin with the equilibrium e_1. Preserving the order of moves within $S(e_1)$ and $N \setminus S(e_1)$, move the members of $S(e_1)$ to the front. (So, for instance, if $n = 5$ and $S(e_1) = \{2,4\}$, then we would have a new order of moves 2, 4, 1, 3, 5.) It is clear that, in this new order, *at least* all the firms in $S(e_1)$ will switch. Now, leaving fixed the order of $S(e_1)$, rearrange the members of $N \setminus S(e_1)$ so that the members of $S(e_2) \setminus S(e_1)$ come immediately after the members of $S(e_1)$, and come in the order they took in e_2. It should now be clear that with the moves in that order all the members of $S(e_1) \setminus S(e_2)$ will choose Y. This proves the first part of Proposition 3. To clarify the somewhat involved rearrangement, we now give an example to illustrate.

Let $n = 5$, $S(e_1) = \{2,4\}$ when the order is 1,2,3,4,5 (e_1) and $S(e_2) = \{3,4,5\}$ when the order is 1,4,5,3,2 (e_2). Then we first change 1,2,3,4,5 to 2,4,1,3,5 (bringing the elements of $S(e_1)$ to the front). Next, keeping 2,4 fixed at the front, we rearrange 1,3,5 so that 5 and 3 are brought forward, and placed in that order because that is how they appear in e_2. Thus we have 2,4,5,3,1. In this order, 2, 3, 4, and 5 will all switch.

The second claim of Proposition 3 follows by repeated application of the first part.

The final claim, that all firms in $S(e^*)$ will switch if the timing of moves is endogenous, can be shown by induction on n as follows: the first firm to move in e^* can move rapidly and choose Y. (If it were not the first to move, it would be because another firm had committed to Y, since "moves" X do not really count, as they are reversible.) He can then rely on the (inductively assumed) proposition for the remaining firms to ensure that the maximal set, i.e., $S(e^*)$ less himself, of the others will choose Y. This puts everyone in the same position as in e^* itself, so the outcome is that $S(e^*)$ will switch. This proves Proposition 3.

References

ARTHUR, W.B. "On Competing Technologies and Historical Small Events: The Dynamics of Choice under Increasing Returns." Mimeo, Stanford University, 1983.

DAVID, P.A. "Understanding the Economics of QWERTY, or Is History Necessary?" Mimeo, Stanford University, 1984.

DYBVIG, P.H. AND SPATT, C. "Adoption Externalities as Public Goods." *Journal of Public Economics*, Vol. 20 (1983), pp. 231–248.

FARRELL, J. "Communication in Games." Mimeo, Massachusetts Institute of Technology, 1982.

HEMENWAY, D. *Industrywide Voluntary Product Standards*. Cambridge: Ballinger, 1975.

KATZ, M. AND SHAPIRO, C. "Network Externalities, Competition, and Compatibility." Woodrow Wilson School Discussion Paper #54, Princeton University, 1983. Forthcoming *American Economic Review*.

KINDLEBERGER, C.P. "Standards as Public, Collective and Private Goods." *Kyklos*, Vol. 36 (1983), pp. 377–397.

ROHLFS, J. "A Theory of Interdependent Demand for a Communications Service." *Bell Journal of Economics and Management Science*, Vol. 5 (1974), pp. 16–37.

SALOP, S.C. "Monopolistic Competition with Outside Goods." *Bell Journal of Economics*, Vol. 10 (1979), pp. 141–156.

U.S. FEDERAL TRADE COMMISSION. "Standards and Certification: Final Staff Report." Washington, D.C.: U.S. Government Printing Office, 1983.

WILSON, C. "Games of Timing with Incomplete Information." Mimeo, New York University, 1984.