The Browser War – Econometric Analysis of Markov Perfect Equilibrium in Markets with Network Effects[1]

Mark Jenkins[2], Paul Liu[3], Rosa L. Matzkin[4], Daniel L. McFadden[5]
April 22, 2004; revised December 31, 2004

ABSTRACT: When demands for heterogeneous goods in a concentrated market shift over time due to network or contagion effects, forward-looking firms consider the strategic impact of investment, pricing, and other conduct. Network effects may be a substantial barrier to entry, giving both entrants and incumbents powerful strategic incentives to “tip” the market. A Markov perfect equilibrium model captures this strategic behavior, and permits the comparison of “as is” market trajectories with “but for” trajectories under counterfactuals where “bad acts” by some firms are eliminated. Our analysis is applied to a stylized description of the browser war between Netscape and Microsoft. Appendices give conditions for econometric identification and estimation of a Markov perfect equilibrium model from observations on partial trajectories, and discuss estimation of the impacts of firm conduct on consumers and rival firms.

Keywords and Phrases: Oligopoly_Theory, Network_Externalities, Markov_Perfect_Equilibrium

JEL Codes: C35, C51, C53, C73, L13, L15, L41, L63

This paper is posted at http://emlab.berkeley.edu/~mcfadden

---

[1] We are indebted to Peter Adams for research assistance, and to Ken Wise, Uirich Doraszelski, and Mark Satterthwaite for useful discussions. This research was supported in part by The Brattle Group in conjunction with litigation support projects for AOL/Time Warner and for Sun Microsystems. An early version of this paper was presented at the International Conference on Industrial Organization, Chicago, April 2004.

[2] Department of Economics, Stanford University, Stanford, CA 94305-6072

[3] The Brattle Group, 353 Sacramento Street, Suite 1140, San Francisco, CA 94111

[4] Department of Economics, Northwestern University, 2003 Sheridan Road, Evanston, IL 60208.

[5] Corresponding author: Department of Economics, University of California, Berkeley, CA 94720-3880, mcfadden@econ.berkeley.edu.

Contents

The Browser War – Econometric Analysis of Markov Perfect Equilibrium in Markets with Network Effects

Mark Jenkins, Paul Liu, Rosa L. Matzkin, Daniel L. McFadden

"In the long run, there is just another short run."
Abba Lerner

"That which is static and repetitive is boring.
That which is dynamic and random is confusing.
In between lies art."
John A. Locke

I. Introduction

The dynamics of demand and supply often link concentrated markets for heterogeneous goods through time. Demand-side effects come from inertia (e.g., habit, switching cost, and lock-in effects); contagion (e.g., bandwagon or addiction effects), and network externalities (e.g., connectivity or interoperability effects); we will refer to these collectively as "network effects". There are also supply-side effects that come from capital sunk in production capacity, product quality, and brand promotion. Forward-looking firms consider the strategic possibility of pursuing market share to "tip" markets in which network effects are important. An Incumbent that dominates a market through share and scope has an incentive to maintain share as a barrier that makes entry costly and risky. A potential entrant has an incentive to grab market share to overcome the incumbent advantage. These strategic incentives, and the prominent role that network effects give to market penetration, are features of dynamic concentrated markets that are not captured in traditional industrial organization models that picture the long run as a sequence of myopic short-run market decisions by participants.

Quantitative analysis of firm conduct in markets with network effects requires dynamic models of firm behavior that predict market trajectories. The theory of *Markov Perfect Equilibrium* (MPE) introduced by Maskin and Tirole (1987,1988a,1988b), and developed as an analytic tool for industrial organization applications by Ericson and Pakes (1995), Pakes and McGuire (1994), and others, captures strategic behavior in evolving markets, and provides a platform for econometric analysis of dynamic market structure. This theory can be used to quantify the impacts on market participants of specific “bad acts” by a firm. This paper discusses the application of MPE to damage estimation in anti-trust litigation, illustrates this discussion with an analysis of the browser war between Netscape and Microsoft, and offers a framework for estimating the damages to Netscape caused by Microsoft “bad acts”. Section 2 discusses the relationship between anti-trust litigation and dynamic models. Section 3 analyzes the browser war between Netscape and Microsoft, where network effects appear to have sharpened strategic incentives and driven a “no-holds-barred” battle for market share in which Microsoft “bad acts” may have been decisive in tipping the market. Section 4 gives conclusions. Appendices review the game theoretic properties of MPE models and issues of econometric identification and estimation.

2. Market Dynamics and Anti-Trust Law

The fundamental goal of anti-trust law, as currently interpreted by the courts, is to promote market efficiency and protect the welfare of market participants by limiting the acquisition and exercise of market power; see Kovacic and Shapiro (1999). This is accomplished through statues that provide *per se* limits on firm conduct, and give consumers and rival firms standing to sue for relief and damages from a firm’s “bad acts”. The acquisition and exercise of market power is a dynamic process that can drive firm behavior on the launch and positioning of new products, and influence their management of the product cycle, where network effects can reinforce barriers to entry. Then, anti-trust law, and assessments of liability and damages based upon it, need to account consistently for the dynamic evolution of concentrated markets under alternative rules for firm conduct. This is particularly true for the task of quantifying and estimating damages. The traditional focus of anti-trust cases on contemporaneous damages is inadequate when illegal conduct has strategic implications for extended market trajectories under “as is” and “but for” conditions. Courts have been reluctant to accept damages estimates based on future market projections that may be speculative. Assuring reliable projections does place heavy demands on the economic analysis, but a tightly specified and carefully estimated model of market dynamics may nevertheless be the best evidence on the harm caused by "bad acts" that have strategic consequences.

Anti-trust law is not rooted in a comprehensive economic theory of concentrated markets, but its evolution has been influenced by economic analysis. Historically, the economic theory of concentrated markets has focused on static short-run, or long-run steady state, market outcomes without explicit dynamics, and with modest attention to consumer behavior and the structure of demand. The classic "workable competition", Cournot, Bertrand, Stackelberg, and capacity competition models of market conduct fit this mold; see Areeda (1975), Tirole (2000), Kreps-Shenckman (1982).

Two major developments in industrial organization offer the prospect of putting the anti-trust law on a better dynamic foundation. The first is the adaptation from market research of models of demand that describe consumer behavior and welfare in response to new products and changing product attributes, and capture network externalities and other demand-side dynamics; see Anderson-de Palma-Thisse (1996), Weirden-Froeb (1994). The second is the continuing development of the theory of dynamic stochastic games, particularly the MPE solution concept, and an expanding literature on its application to concentrated markets. These developments have begun to influence anti-trust law through adoption of merger models that rely on analysis of consumer outcomes rather than mechanical measures of concentration, and consideration of the dynamics of market definition in high-technology industries; see Federal Trade Commission (1997), Overstreet-Keyte-Gale (1998), Scheffman-Coleman (2003), Shapiro-Kovacic (2000), Starek-Stockum (1998), Reiffen-Vita (1995), Weirden-Froeb (2002), Varian (), United States v. Sungard Data Systems (2001).

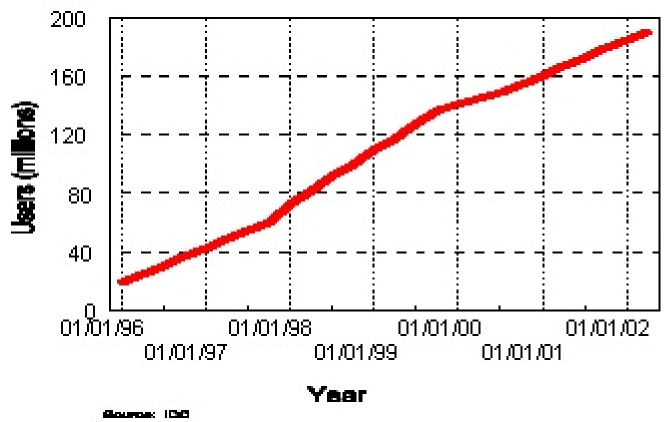
Another area where market dynamics is critical are anti-trust charges of predation through pricing, tying, bundling, and exclusionary arrangements. Static analysis of a single market leads to the Areeda-Turner standard that to be predatory, pricing must be below short-run marginal cost (or, loosely, average variable cost). Further, the courts are skeptical of price predation claims, on the grounds that consumers benefit from low prices in the short run, and predation must be self-defeating because recurring entry will continue to require predatory response, removing the prospect of long-run benefits that might offset short-term costs. However, when demand-side network effects link the markets through time, neither the Areeda-Turner standard or the argument that predation is strategically unattractive are necessarily accurate indicators for market consequences; see Doganoglu (2003). Suppose a firm, with deep pockets obtained from a monopoly in its primary market, faces a threat of primary-market entry by rivals in a secondary market characterized by network effects. Suppose this monopolist cross-subsidizes innovation in the secondary market, and through aggressive pricing, perhaps including bundling and tying to its primary product, succeeds in tipping the secondary market in its favor. Consumers benefit from low secondary market prices in the short run, and in the longer run gain the network benefit of ubiquity, but lose from the higher prices and lower rates of innovation the monopolist can eventually establish in both markets once the path to entry is closed. To determine whether the cross-subsidy was socially undesirable, one needs to obtain "as is" and "but for" dynamic trajectories for this market, and calculate the expected present value of the net welfare effect.

For the anti-trust law to work effectively as a regulator of resource allocation, it must be able to discriminate between circumstances and strategic conduct in dynamic markets that harm consumer welfare and market efficiency, and those that are in net neutral or beneficial. Apparently minor announcement effects, exclusions, or delays that the law currently treats as inconsequential may be sufficient to tip a market with strong network or scale effects, and result in its monopolization. On the other hand, some conduct that is illegal under current anti-trust law, such as an industry's promulgation of proprietary product standards in order to exclude entry, may increase market efficiency if the benefits of ubiquity offset the losses associated with increased market power. These examples suggest that the interpretation of firm behavior in the "as is" world, and projection of "but for" behavior when some acts are disallowed, needs to take into account the incentives created by the market dynamics, the strategic elements in firm response, and intertemporal consumer welfare. MPE models of the type studied by Hall-Royer-Van Audenrode (2004) and by this paper provide templates for such analysis.

3. The Browser War: Netscape v. Microsoft

3.1. Background. We study the market for web browsers, software programs that allow the user to view over the Internet a particular type of web content called Hyper Text Markup Language (HTML). In early 1993, Marc Andreessen and a group of fellow students developed the first graphical web browser, called Mosaic, while working for the National Center for Supercomputing Applications (NCSA) at the University of Illinois at Urbana-Champaign (UIUC).[6] Word about Mosaic spread rapidly and by the end of 1994, Mosaic had been downloaded about two million times.[7] The introduction into the marketplace of a graphical browser was pivotal in spurring interest in the Internet; Figure 1 shows that the number of internet users grew from less than 10 million in mid-1995 to over 190 million in the second quarter of 2002.

Figure 1. Number of Internet Users



In April 1994, Andreessen partnered with James Clark to form Netscape Communications.[8] The Netscape team developed a commercial Internet browser, called Netscape Navigator, which had a beta release in October 1994 and an official release in December 1994. Navigator was immediately successful. By February 1995, it had captured 60 percent of the fledgling market, and it was the leader of the "Internet Revolution".[9] By January, 1996, its share of the browser market was 90 percent.[10]

---

[6] Cusumano and Yoffie, p. 3.
[7] Cusumano and Yoffie, p. 95. The figures include other methods of obtaining Mosaic, but downloads were the dominant form of distribution.
[8] Cusumano and Yoffie, p. 7. The company was originally called Mosaic Communications.
[9] Cusumano and Yoffie, p. 7, 107.
[10] Cusumano and Yoffie, p. 9.

Navigator's quality was a large factor in its success[11]; another contributing factor was Netscape's innovative "free but not free" concept in which fully functional versions of Navigator could be downloaded for a free trial period and then users were asked but not required to pay $39 per copy.[12]

Microsoft was late in entering the internet software business. Although it did plan to introduce a rudimentary browser in its Windows 95 operating system, it apparently did not initially perceive internet software as a significant market or threat to its Windows operating system core business.[13] Nonetheless, in June 1995, Microsoft and Netscape had a meeting at which Microsoft offered Netscape the option of either entering into a "special relationship" with Microsoft and agreeing to develop only non-Windows 95 browsers, or being regarded as a competitor and treated as such.[14] Netscape declined, as the special relationship would have foreclosed them from developing for Microsoft's soon-to-be ubiquitous Windows platform. In December 1995, Microsoft CEO Bill Gates publicly declared that Netscape had "awakened a sleeping giant" and that Microsoft would reverse course and become "hard core" on "embracing and extending" the Internet.[15] Thus began the "browser wars." Over the years that followed, Netscape and Microsoft would repeatedly one-up each other with new and improved versions of their respective browsers; see Table 1. The initial versions of Microsoft's Internet Explorer (IE) lacked the full functionality of Netscape Navigator, and it was common for consumers to download Navigator in preference to Internet Explorer even when IE came pre-installed on their new computers. However, with Version 3.0 of IE released in August 1996, the two browsers were generally viewed by reviewers and the public as comparable in features and quality.

---

[11] Netscape was consistently rated as the best browser product then on the market. Cusumano and Yoffie, p. 96.

[12] Cusumano and Yoffie, p. 98.

[13] Cusumano and Yoffie, p. 9. Microsoft planned to incorporate a rudimentary web browser into the forthcoming release of Windows 95 but otherwise little apparent attention was paid to the Internet. Cusumano and Yoffie, pp. 107-108. Microsoft later changed its tune, realizing that the application programming interfaces (APIs) exposed by Netscape Navigator could potentially allow software developers to create programs that only required the presence of Navigator (and not Windows) to run.

[14] Findings of Fact, ¶¶82-89.

[15] Cusumano and Yoffie, pp. 109-110.

Table 1

Windows Internet Explorer Version	Final Release Date	Netscape Navigator Version	Final Release Date
		1.0	December 1994
		1.1	April 1995
1.0	August 1995		
2.0	November 1995		
		2.0	March 1996
3.0	August 1996	3.0	August 1996
4.0	October 1997		
		4.5	October 1998
5.0	March 1999		
5.5	July 2000		
		6.0	November 2000
6.0	October 2001		
		7.0	August 2002

Source: http://www.blooberry.com/indexdot/history.bro

In bundling Internet Explorer with every copy of Windows, making Internet Explorer separately available for free downloading, and investing substantial resources into improving the quality of its browser, Microsoft sought to "cut off Netscape's air supply"[16] and undermine Netscape's "free but not free" strategy.[17] Microsoft had good reasons to go after Netscape. Internet activities and commerce were emerging as a major component of computer use, offering profit opportunities to the firms controlling internet software development. In addition, internet software and user interfaces were a serious threat to Microsoft's dominance in the market for operating systems, raising the possibility of internet-based cross-platform middleware that could marginalize and commoditize the

[16] David A. Kaplan, *Silicon Boys*, p. 276.
[17] Cusumano and Yoffie, p. 141.

Windows operating system. If Microsoft could control the internet software market, then its operating system monopoly would be protected from entry through this channel.

Demand side network effects, in which familiarity, compatibility, and availability of connections solidify the customer base of a successful browser and make its market difficult to penetrate, raised the stakes for both Netscape and Microsoft. With deep pockets, Microsoft could engage in an R&D blitz that might leapfrog ahead of Netscape. Beyond this, Microsoft engaged in a combination of acts that the trial court in U.S. v. Microsoft found anticompetitive, We exclude acts where the trial court findings were reversed on appeal, but not those where the issue was remanded for further adjudication.[18] Microsoft bundled Internet Explorer with Windows, and required its presence on the desktop, making it unlikely that consumers and Original Equipment Manufacturers (OEM's) would seek alternative browsers.[19] It provided free downloads of Internet Explorer, essentially eliminating the possibility that Netscape Navigator could be sold at a profitable price. In January 1998, Netscape was finally driven to lower the official price of its browser to zero.[20] Both bundling and predatory pricing claims are difficult to pursue under the existing anti-trust law, but in a market with strong network effects, they may in tandem have been what was needed to tip the browser market in Microsoft's favor.

In addition, Microsoft attempted to cut off the primary channels through which Netscape distributed its browsers.[21] First, Microsoft entered into agreements with many Internet Service Providers (ISP's) that restricted their ability to distribute alternative browsers such as Netscape Navigator.[22] Second, Microsoft entered into agreements with Online Services,

---

[18] Evans (2001) lists the government's claims in the case of U.S. v. Microsoft, and summarizes the disposition of each claim. Extended discussion of the nature and legal foundation of the various claims can be found in Breshnahan (2004), Davis, MacCrisken, and Murphy (2001), Economides (2001), Evans, Nichols, and Schmalensee (2001), Fisher and Rubinfeld (2000), Gilbert and Katz (2001), Klien (2001), and Schmalensee (1999).

[19] Findings of Fact, ¶159. Microsoft also removed in Windows 98 the ability to uninstall Internet Explorer and prohibited modification of the Windows boot sequence to promote non-Microsoft software. Findings of Fact, ¶170, 204. The emergence and persistence of products with the full functionality of IE for both internet and desktop management demonstrate that tying IE to the operating system was not technologically necessary. We treat this tying as a "bad act", but note that this remained a contentious issue at the time the *U.S. v. Microsoft* suit settled.

[20] Cusumano and Yoffie, p. 144.

[21] Findings of Fact, ¶143.

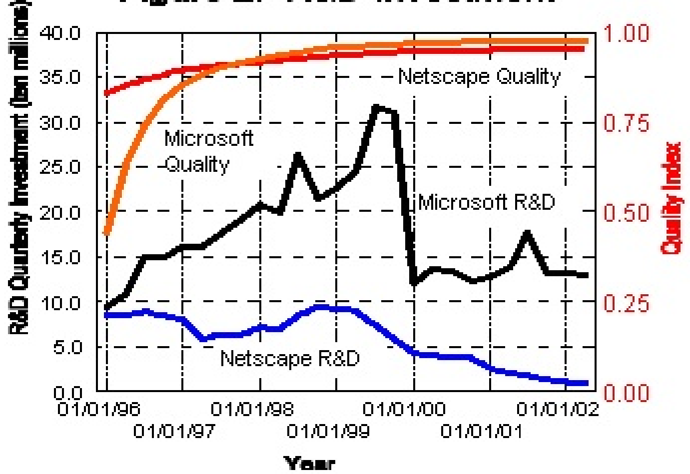
[22] Findings of Fact, ¶247.

Internet Content Providers, and Independent Software Vendors that induced greater use of Internet Explorer. Third, Microsoft contracts with OEM's restricted the placement of non-Microsoft content on the desktop, or gave price discounts to OEM's that provided exclusively Microsoft content. These acts individually affected only small parts of the market, but their impact during a period where there was intense competition for market share and the network benefits of ubiquity made them potentially quite important in determining whether the browser market would "tip" in Microsoft's favor.

3.2. Preliminary Analysis.[23] The *U.S. v. Microsoft* case and the *Netscape v. Microsoft case* produced considerable data on the browser war, but only a limited portion of this is in the public domain. In the analysis that follows, the authors have pieced together and squared off fragmentary information from the public record to reconstruct investment and market share series from the first quarter of 1996 through the second quarter of 2002. The accuracy of this reconstruction is problematic. Consequently, our analysis based on these data must be treated as illustrative. Future analyses with access to company records might reach substantially different conclusions.

An oft-cited figure in the *U.S. v. Microsoft* case held that Microsoft invested approximately $100 million a year in the development of Internet Explorer. Netscape's *total* annual investment (including its investment in server software) was at most a third of that amount, with its browser investment amounting to perhaps no more than a quarter of the total, or approximately $10

Figure 2. R&D Investment



[23] The data upon which this section and subsequent analysis are based are given in Appendix 3.

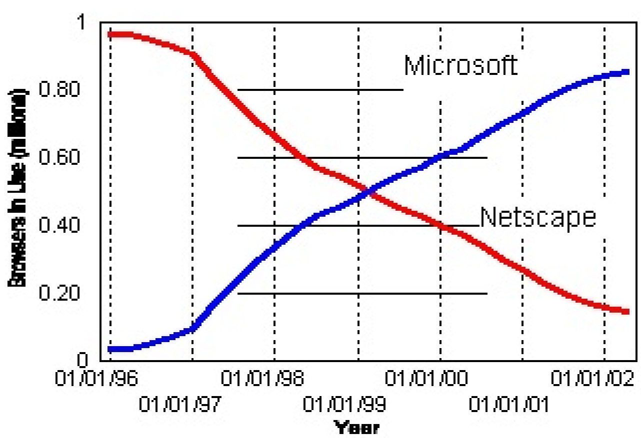
million.[24] Figure 2 shows quarterly browser R&D investment levels for the firms, pieced together by the authors from fragmentary public data; the resulting series should be treated as illustrative rather than definitive. This figure also shows "Quality" indices[25] defined by the following transformation that exhibits sharply diminishing returns to R&D for "mature" software and reflects in part the lag between investment and the release of a new version:

(Quality index in quarter) = 1 - 1/(Cumulative investment at beginning of quarter).

As a result of Microsoft's activities in the browser market, Microsoft's market share of installed browsers rose steadily; see Figure 3. The curve for Netscape in this figure includes all non-Microsoft browser; the "other" category was initially nearly ten percent of the market, but quickly fell to less than two percent.[26]

During this period, the total number of internet users and the installed customer base for each browser grew rapidly; see Figure 4. By the time Microsoft's restrictions on OEMs and ISP agreements were lifted in 1998, the browser war was effectively over. Even though Netscape still had 58 percent installed browser market share in the first quarter of 1998,[27] Microsoft dominated the distribution channels and had considerable momentum.

Figure 3. Browser Market Shares



---

[24]While Netscape initially focused almost exclusively on browser development, later efforts to enter various server software markets led to sharp declines in browser R&D as a percent of total investment.

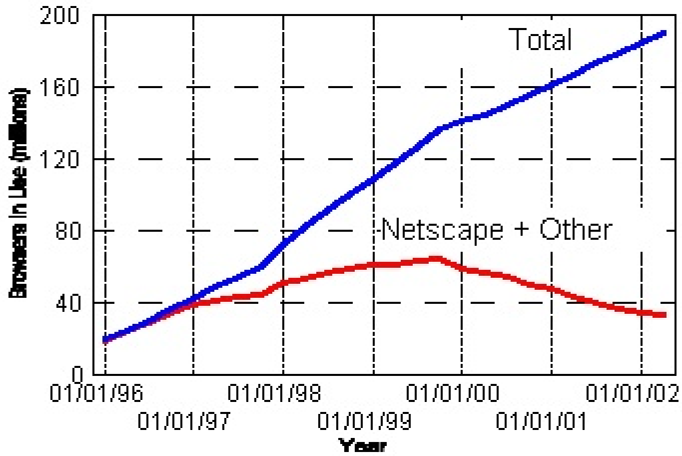
[25]The transformation chosen for "quality" is the result of a limited econometric search for functional forms that best explain demand response to investment.

[26]The sources for Figures 3 and 4 are public records in the U.S. v. Microsoft case and internet activity tracking services. The final series are adjusted by the authors, and should be treated as illustrative rather than definitive.,

[27] Figures from http://www.cen.uiuc.edu/bstats/latest.html.

The actions by Microsoft against Netscape were the focus in the U.S. Department of Justice's antitrust suit against Microsoft in 1997.[28] However, even though the District Court and the Court of Appeals found many of the Microsoft actions to be illegal,[29] the final settlement between the U.S. Department of Justice and Microsoft in November 2002 provided little relief to Netscape.[30] Netscape filed a separate civil antitrust suit against Microsoft in January 2002.[31] That suit was settled in May 2003 for $750 million.[32] By July 2002, Microsoft had a 87 percent of the browsers in use. However, the barriers to entry in the browser market are not impermeable. By December 2004, fueled by security problems with IE, the Mozilla family of browsers, which include Firefox and the latest release of Netscape, had whittled IE's share of the market down to 71.7 percent.

Figure 4. Installed Base



3.3 A Dynamic Stochastic Game Model. The question of how to estimate the impact of Microsoft's illegal actions raises the question as to what constitutes a proper model for analysing firm behavior in this market.[33] Traditional static industrial organization models fail to capture network effects and dynamic features of the market such as the importance of investment on future profits; as a result, a dynamic market model is likely to be necessary. In this paper, we have developed and estimated a dynamic Markov Perfect

---

[28] [Cite DOJ Complaint.]
[29] [Cite Findings of Fact, Appeal Court Opinion]
[30] See *Final Judgment*, viewed at http://www.usdoj.gov/atr/cases/ms_index.htm on March 4, 2004.
[31] See *Complaint* in *Netscape v. Microsoft*, viewed at http://legal.web.aol.com/decisions/dlother/Netscape%20v.%20Microsoft.pdf on March 4, 2004.
[32] See http://money.cnn.com/2003/05/29/technology/microsoft/.
[33] We were engaged by Netscape to calculate damages in *Netscape v. Microsoft*.

Equilibrium (MPE) model which is motivated by the following browser market characteristics:

1) Firms. The browser market consisted primarily of Netscape's Navigator and Microsoft's Internet Explorer for most of its history. While there have been other browsers in the market (e.g., Opera, Mozilla), these were not significant during the 1996-2002 period. However, these minor players constituted a fringe that would have limited the ability of Netscape and Microsoft to raise prices substantially. We simplify our analysis by treating Netscape and Microsoft as duopolists in the browser market, and lumping the fringe browsers together with Netscape.

2) Browser Acquisition and Use. An important distinction in browsers is between acquisition by either pre-installation or initial download, or subsequent download of new versions, and use. This is potentially a complex consumer decision problem, as acquisition choices may be made with the option value of future choices and expectations of future environments in mind. However, it seems fairly realistic to simplify this drastically, and assume that consumers make a single initial browser choice at the time a new PC or new operating system is purchased, and then use this chosen browser (with updates) for the life of the PC. We will assume further that the consumer makes the initial browser choice myopically on the basis of costs and qualities prevailing at the moment. (Note that this assumption is consistent with an MPE model of consumers as *active* market participants whose policy decisions are functions only of current state. However, we will not exploit this modelling option, and will instead treat consumer behaviour as myopic and non-strategic.) Assume that when both browsers are available pre-installed, the consumer simply chooses the one that will be used. Similarly, when neither are available pre-installed, the consumer simply downloads the preferred browser. Finally, when only one browser is available pre-installed, the consumer then decides whether to use this browser, or to abandon it and download the other browser (at an added time cost).

3) Demand Equation – Inclusion of Network Effect. The browser market is characterized by strong network effects. As the Findings of Fact stated, "[Microsoft] believed that a comparable browser product offered at no charge would still not be compelling enough to consumers to detract substantially from Navigator's existing share of browser usage. This belief was due, at least in part, to the fact that Navigator already enjoyed a very large installed based and had become nearly synonymous with the Web in the public's consciousness."[34]

4) Demand Equation – Exclusion of Price. Price was not a primary consideration in choice of browser. Published browser price did not reflect the actual price typically paid by a user, as users mostly did not pay the voluntary license fee. Netscape received only a small revenue stream from browser licenses. Once Microsoft entered the market and priced its browser at zero, Netscape found it had to respond by doing similarly, thus extinguishing browser licenses as a revenue source. One aspect of browser pricing is that the effective price to the consumer exceeded the license price by the cost of downloading and installing the software (if required). This is potentially relevant in assessing consumer response to exclusion of Netscape from distribution on new computers. It would be useful to include prices in the model if predation were an issue and estimates of market prices under "but for" conditions were needed. If prices are policy variables, then the distinction between browser acquisition and use becomes important. However, pricing was not a focus of the legal case, and we do not include prices in the model analysed in this paper.

5) Demand Equation – Inclusion of Quality. Beyond the network effects, consumers cared primarily about the quality of their Internet experience. Newspaper and magazines regularly reviewed the then-current versions of Netscape Navigator and Internet Explorer in head-to-head matchups. Technologically-savvy computer users promoted the browser they viewed most favourably. Quality is determined to a large extent by the amount of investment placed in upgrading each incremental version of the browser. We were unable to obtain consistent historical direct measures of browser quality, and instead have chosen an empirical transformation of cumulative investment that explains well the evolution of market shares.

6) Demand Equation – Symmetry. Fully informed, rational consumers will choose among products on the basis of their generic hedonic attributes (including network

[34] Findings of Fact, ¶143.

attributes), independently of non-generic brand effects.[35] Then, when the hedonic attributes of the products are captured in the model, the demand function will be generic, with symmetry across products in its dependence on hedonic attributes.

7) Profit Equation – Inclusion of Derivative Revenues per Browser. Browser revenues were always a relatively small portion of Netscape's business plan. Netscape intended to leverage its success in the browser business into other sources of revenue such as server software sales and advertising revenue from its Internet portal.[36] Microsoft could use Internet Explorer to solidify its Windows monopoly and establish its own Internet portal. Thus Internet browsers were viewed as software with insignificant direct revenues but with substantial revenues via linkages to other markets. These indirect revenues would likely be in direct proportion to the number of browsers, and could potentially be larger for Microsoft than for Netscape (to reflect the benefits to Microsoft's Windows monopoly).

The following paragraphs give the model specification; a summary listing of the variables is given in Table 2. The exogenous variable N(t) represents the total installed base of browsers. We assume that each Internet user utilizes exactly one browser to access the Internet, so that N(t) also represents the number of internet users, and that the only choice made by the user is which browser to adopt. We simplify the model by assuming that there was no feedback from changes in browser quality to the total number of computers purchased or upgraded, and the frequencies with which browsers were installed or replaced. This is accurate only if the initial version 1.1 of Navigator would have been sufficient to fuel the full internet revolution without further quality improvements in browsers. This may be an acceptable first approximation, but it would be interesting to elaborate our simplified model to allow some overall internet use and browser demand response to quality.

---

[35] Rational but incompletely informed consumers, or irrational ones, may use brand identification as a token for unobserved quality. We assume as a working hypothesis that browser users are rational and informed.

[36] Cusumano and Yoffie, p. 25, state "The key was to build market share and create the standard. Profits would eventually follow." They also state, at p. 7, that "Initially, Netscape's business model called for developing two sets of products—the browser, which would catapult Netscape to fame, and Web servers, which would pay the company's bills."

The Microsoft "bad acts" variables are $a_0(t)$ = IAP, defined as the proportion of new Internet Access Provider accounts that were subject to Microsoft agreements prohibiting distribution of Navigator, and $a_1(t)$ = WINDOW, defined as the proportion of all new computers that were delivered with IE, but not Navigator, preloaded. Once IE had been bundled into the Windows 98 operating system, which was introduced in July 1998 and within three quarters had displaced earlier operating systems including Windows 95, the variable WINDOW was one. In earlier periods, WINDOW is the proportion of computers delivered by OEM's who had signed restrictive agreements with Microsoft that precluded supplying Navigator. The policy variables determined within the model are the investments in development, $i_0(t)$ and $i_1(t)$.

Table 2. List of Variables		
Symbol	Type	Description
$j = 0,1$		firms, $j = 0$ for Netscape, $j = 1$ for Microsoft
$t = 0,1,..,T$		observation time periods, $T = 25$ (1996 Q1 through 2002 Q2)
$t = T+1,...,H$		forecast time periods from T+1 to horizon $H \leq \infty$
T^*		a time period, $T^* \leq T$, following which all exogenous variables are stationary
H^*		a finite terminal time period, $T \leq H^* \leq H$, chosen as an (artificial) computational horizon
$B_0(t)$	SO	Netscape installed base of browsers at <u>beginning</u> of period t
$N(t)$	XO	Total installed base of browsers at <u>beginning</u> of period t
$S_0(t)$	DO	Netscape share of installed base at <u>beginning</u> of period t
$K_j(t)$	SO	accumulated browser investment for firm j at <u>beginning</u> of period t
$Q_j(t)$	DO	browser quality for firm j at <u>beginning</u> of period t
$i_j(t)$	PO	current browser R&D investment for firm j
$n(t)$	DO	number of consumers choosing a new browser in period t
$s_0(t)$	DO	share of new browsers in period t going to Netscape
$r_j(t)$	XO	current revenue in period t per firm j browser used
$a_k(t)$	XO	intensity of Microsoft's bad acts in period t, where k = 0 = IAP and k = 1 = Windows
$\varepsilon(t)$	XU	preference-side, market-level demand shifter
$\zeta_j(t)$	XU	firm j private investment cost

Types: State (S), Policy (P), Exogenous (X), Observed (O), Unobserved (U), (D) Derived

The variable $r_j(t)$ represents the indirect revenue that a firm derives from each browser user in a given period; e.g., from advertising, fees for directing users to e-commerce sites, and the marginal indirect profit from sales of enterprise and server software and operating systems. Indirect revenue per user need not be the same for both firms. The observed exogenous variables N(t), $r_j(t)$, and $a_k(t)$, and the unobserved variable ε(t), are known to both firms when they make their investment policy decisions in period t. The unobserved exogenous variable $\zeta_j(t)$ is private information known only to firm j, and not to its rival, at the time its investment decision is made; this are interpreted as a component of the firm's unit cost of R&D investment. When the model is considered in counterfactual circumstances with the "bad acts" turned off, we assume that all remaining exogenous variables take on the same values that they did in the "as is" world, and follow stationary Markov processes that hold for at least part of the observation period and all periods beyond. We assume that the firms know these Markov processes and use them in forming expectations. We will assume that the shocks ε(t), $\zeta_0(t)$, and $\zeta_1(t)$ are independent of each other and independent across time. (It complicates the predictions of firms if the shocks instead follow Markov processes, so that some information on the distribution of current and future shocks may be embedded in past equilibria. Our independence assumption avoids having to deal with this issue.)

This model determines three state variables: the Netscape installed base $B_0(t)$ and the accumulated R&D investment of each firm, $K_0(t)$ and $K_1(t)$. In addition, the observed and unobserved exogenous variables N(t), $R_0(t)$, $r_1(t)$, $a_0(t)$,$a_1(t)$, ε(t), $\zeta_0(t)$, and $\zeta_1(t)$ can be interpreted as state variables with equations of motion characterized by Markov transition probabilities that are determined outside the model. From these variables, one can derive the variables $S_0(t)$, $Q_0(t)$, $Q_1(t)$, $s_0(t)$, and n(t). In the following analysis, it will be convenient to formulate the model using some of the derived variables. In the following analysis, we will repeatedly use *ramp* functions ρ(x,y,z) defined by

(1) $$\rho(x,y,z) = \max\{x,\min\{y,z\}\}.$$

Accumulated R&D investment and the total number of browser users satisfy the equations of motion

(2) $$K_j(t+1) = i_j(t) + K_j(t),$$

(3) $$N(t+1) = n(t) + (1-d_B)N(t),$$

where d_B denotes the rate at which existing browser users reconsider their browser selection, and n(t) is exogenous. "Quality" is defined by

(4) $$Q_j(t) = 1 - 1/K_j(t).$$

Firm 0's share of the browser installed base is

(5) $$S_0(t) = B_0(t)/N(t).$$

The heart of the model is the share $s_0(t)$ of new users choosing browser 0, which must satisfy $0 \leq s_0(t) \leq 1$. This will determine the equation of motion for installed base,

(6) $$B_0(t+1) = (1-d_B)B_0(t) + n(t)s_0(t).$$

These equations combine to give an equation of motion for firm 0's share of installed base,

(7) $$S_0^*(t+1) = (1-d_B)S_0(t)N(t)/N(t+1) + s_0(t)[1-(1-d_B)N(t)/N(t+1)]$$
$$S_0(t+1) = \rho(0,1,S_0^*(t+1)),$$

where the "*" denotes a latent variable and the ramp mapping restricts the share to the unit interval. We assume that consumer choice of a new browser is influenced by a network effect $S_0(t)$, the quality difference between the alternatives, $Q_0(t)-Q_1(t)$, the "bad acts" variables, and the public shock $\varepsilon(t)$. We assume a simple linear specification with the share restricted to the unit interval,

(8) $$s_0^*(t) - ½ = \alpha_1(S_0(t) - ½) + \alpha_2(Q_0(t)-Q_1(t)) + \alpha_3 IAP(t) + \alpha_4 WINDOW(t) + \varepsilon(t)$$
$$s_0(t) = \rho(0,1,s_0^*(t)).$$

We write (8) without a constant term, consistent with the principle that informed, rational consumers will choose between the two browser brands symmetrically, once network effects, relative quality, and other hedonic attributes of the market are included in the model. We anticipate that this demand model will show a network effect ($\alpha_1 > 0$), a response to relative quality ($\alpha_2 > 0$), and an impact of the Microsoft "bad acts" (α_3, $\alpha_4 < 0$).

The profit of each firm in period t is

$$(9) \qquad \pi_j = r_j(t)B_j(t) - (1+\zeta_j(t))i_j(t) - \gamma i_{jj}(t)^2 .$$

The first term in the profit function reflects the advertising and other revenue earned as a result of having an installed base of browser users, the second reflects the firm's investment in browser development, modulated by the private costs associated with investment. The third term is a cost penalty for crash R&D, with a coefficient γ_j determining its significance. The investment policy variable in this equation is bounded below by zero. It is also possible to bound it above when n(t) and r(t) are stationary by noting that investment will never exceed the present value of maximum possible industry profit.

The firms play the market game defined by the payoff functions (9). Firm j seeks to maximize its expected discounted profit stream at each time t,

$$(10) \qquad EPV_j(t) = \mathbf{E}_{-j,t} \sum_{m=t}^{H} \pi_j \beta^{m-t},$$

where $\beta < 1$ is a discount factor and the expectation $\mathbf{E}_{-j,t}$ is taken with respect to the firm's beliefs about all future and unobserved variables and the strategy of its rival, and is conditioned on the information known to the firm when the investment choice must be made, including the state of the system, the history of play, the public information $\varepsilon(t)$, and its private information $\zeta_j(t)$. This is then a stochastic dynamic game with the following properties:

1) The payoff functions and equations of motion are continuously differentiable in the policy and state variables, and time-invariant. It the variables n(t), $r_j(t)$, $a_k(t)$, $\varepsilon(t)$, $\zeta_j(t)$ are

stationary for $t \geq T^*$ and $H = +\infty$, the discount factor β is less than one, and the depreciation rate d_B is positive, then the game is stationary for $t \geq T^*$.

2) If the horizon H is finite,and the exogenous observed and unobserved variables n(t) $r_j(t)$, $a_k(t)$, $\varepsilon(t)$, $\zeta_j(t)$ have compact support, and there is an upper bound on investment, then the policy space is compact, and the existence of (mixed) open-loop Bayes-Perfect-Nash (BPN) solutions is a standard game theory result; see Federgruen (1976), Whitt (1980). Further, these open-loop solutions have closed-loop feedback representations. If the horizon is infinite and the game is compact and stationary with a discount factor $\beta < 1$, the existence of (mixed) closed-loop feedback BPN solutions is obtained under some restrictions on how players form expectations, and with the possible use of random events extrinsic to the model as coordination mechanisms; see Duffie, Geanakoplos, Mas-Colell, McLennan (1994), Mertens, Parthasarathy (2003), and Appendix 1.

3) The expected present value of profit $EPV_j(t)$, optimized in $i_j(t)$ given beliefs regarding future variables and conjectures on the response of its rival, is a convex function of $\zeta_j(t)$, and consequently is almost everywhere twice continuously differentiable in $\zeta_j(t)$. Since $EPV_j(t)$ is linear decreasing in $\zeta_j(t)i_j(t)$, the derivative of its optimized value with respect to $\zeta_j(t)$, when it exists, equals the negative of a unique value of $i_j(t)$ that achieves this optimum; see McFadden (1978). Therefore, for almost all $\zeta_j(t)$, a BPN strategy is pure. If the $\zeta_j(t)$ have continuous CDF's, this implies that a BPN solution to the game is pure with probability one. A more general version of this result is stated formally in Appendix 1.

4) In general, the structure of the game does not guarantee that a closed-loop feedback Bayes-Nash solution is unique, or that there are not open-loop Bayes-Nash solutions that cannot be represented as closed-loop feedback solutions. For some symmetric initial conditions, there can clearly be multiple solutions in which one firm or the other captures the market.

The specification (4) of the mapping from cumulative investment to product quality implies that the marginal impact of additional investment on $EPV_j(t)$ is bounded. From (7), for $t' > t$,

$$(11) \qquad K_0(t')^2 \frac{\partial s_0(t')}{\partial i_0(t)} = -K_1(t')^2 \frac{\partial s_0(t')}{\partial i_1(t)} = \begin{cases} \alpha_3 & \text{if } s_0^*(t') \in (0,1) \\ 0 & \text{if } s_0^*(t') \notin (0,1) \end{cases} .$$

Then, if $t \geq T^*$,

$$(12) \qquad \frac{\partial EPV_0(t)}{\partial i_0(t)} \leq -1 - \zeta_0(t) - 2\gamma_0 i_0(t) + \alpha_3 \frac{\beta}{1-\beta} \frac{r_0(t)N(t)}{K_0(t)^2}.$$

A similar inequality holds for firm 1. If the $\zeta_j(t)$ are bounded above -1, then these inequalities establish upper bounds on Kj(t), above which R&D investment is unprofitable and further investment is zero.

Suppose the cumulative R&D investments of both firms approach their upper bounds. Then relative quality approaches a limit and optimal investment levels approach zero. Suppose, in addition, that the bad acts variables are turned off for $t \geq T^*$ and the number of installed browsers N(t) reaches a finite limit. Then the market will eventually approach stationarity, with a limiting distribution whose characteristics depend on α_1 and the limiting value κ of $\alpha_2(Q_0(t) - Q_1(t))$. Assume positive network effects, $\alpha_1 > 0$. Substituting (8) into (7) yields the limiting equation of motion for installed base share,

$$(13) \qquad S_0(t+1) = (1-d_B)S_0(t) + d_B\rho(0,1,½+\kappa+\alpha_1(S_0(t)-½)+\varepsilon(t)\}\}.$$

When $0 < ½+\kappa+\alpha_1(S_0(t)-½)+\varepsilon(t) < 1$, this is a linear difference equation in which the coefficient of $S_0(t)$ is $1-d_B(1-\alpha_1)$. If $0 < \alpha_1 < 1$, then this coefficient is positive and less than one, and in the absence of the barriers at 0 and 1, this linear difference equation has a stationary solution centered at $½ + \kappa/(1-\alpha_1)$. When the barriers are present, the support of a stationary solution will be centered around $½ + \kappa/(1-\alpha_1)$ if it is between the barriers, and otherwise will include the barrier nearest $½ + \kappa/(1-\alpha_1)$. If $\alpha_1 \geq 1$, then the coefficient of $S_0(t)$ in the linear difference equation is at least one, and in the absence of barriers, the equation has no stationary solution. The equation with barriers will have a stationary solution whose support includes one barrier or the other, or both. To analyze the behavior of (13) further, let G denote the CDF of $\varepsilon(t)$, which we assume to be continuous with a support $[-c_1,+c_2]$. Then (13) defines a Markov chain with a transition CDF $P(q'|q) = \text{Prob}(S_0(t+1) \le q'|S_0(t)=q)$ satisfying

$$(14)\qquad P(q'|q) = \begin{cases} 0 & \text{if } q' < (1-d_B)q \\ G\left(\dfrac{q'-(1-d_B)q}{d_B} - \dfrac{1-\alpha_1}{2} - \kappa - \alpha_1 q\right) & \text{if } (1-d_B)q \le q' < (1-d_B)q+d_B \\ 1 & \text{if } q' \ge (1-d_B)q+d_B \end{cases}$$

The probability of a transition with $S_0(t+1) \le S_0(t) = q < 1$ is $G((1-\alpha_1)(q-\tfrac{1}{2})-\kappa)$. In the case $0 < \alpha_1 < 1$, the probability of a non-increasing transition rises with q, while if $\alpha_1 > 1$, the probability of a non-increasing transition falls with q. If $(1-\alpha_1)/2+\kappa < -c_2$, then $S_0(t) = 0$ is an absorbing state; and if $(1-\alpha_1)/2-\kappa < -c_1$, then $S_0(t) = 1$ is an absorbing state. It is possible for both extremes to be absorbing states (e.g., $c_1 = c_2 = \tfrac{1}{4}$, $\kappa = 0$, $\alpha_1 = 2$), in which case the process is non-ergodic, with the early history of the market leading it to "tip" permanently to one firm or the other. However, if the support of $\varepsilon(t)$ is sufficiently broad, the chain is irreducible, and will have a unique invariant distribution F satisfying

$$(15)\qquad F(q') \equiv \int_0^1 P(q'|q)F(dq)$$

$$= F\left(\max\left(0,\frac{q'-d_B}{1-d_B}\right)\right) + \int_{\max(0,\frac{q'-d_B}{1-d_B})}^{\min(1,\frac{q'}{1-d_B})} G\left(\frac{q'-(1-d_B)q}{d_B} - \frac{1-\alpha_1}{2} - \kappa - \alpha_1 q\right)F(dq).$$

The distribution F can be computed as the limit of repeated iterations of (15) from any initial distribution, or alternately by solving the system of linear equations obtained by taking for q' and q a sufficiently fine grid on [0,1]. In the limiting case where the model is stationary and the effect of quality is negligible so that (15) characterizes the limiting distribution of shares, the expected present values of profit at an artificial termination period H' are $EPV_0(H') = r_0(H')N(H')[1 - \int_0^1 F(q)dq]/(1-\beta)$ and $EPV_1(H') = r_1(H')N(H')\int_0^1 F(q)dq /(1-\beta)$.

3.4. Markov Perfect Equilibrium. A one-shot stochastic game is defined by policy spaces for the players, and payoff functions of a profile of policies from these spaces and random factors that are unknown to some or all of the players when they choose their policies. A (mixed strategy) *Bayes-Nash* solution to a one-shot stochastic game is defined by a profile of conditional probabilities on the policy spaces of each player, where the conditioning is on the information available to the player, with the property that each player's probability maximizes his expected payoff, given the probabilities of his rivals and given his beliefs about random factors in the game. A mixed strategy solution is *pure* if the support of each player's probability is a singleton.

A dynamic stochastic game is one played in a series of stages, with equations of motion that transform the state of the system and policy choices into updated states. At each stage a player has information on the state of the system and the history of play. Random factors can enter current and future payoffs, and the equations of motion. A Bayes-Nash solution is *perfect*, and is termed a *Bayes-Perfect-Nash* (BPN) solution, if at every stage of the game, the profile of conditional probabilities given the information then available continues to be a Bayes-Nash solution. A BPN solution to a dynamic stochastic game is *Markov*, and is termed a *Markov Perfect Equilibrium* (MPE), if the conditional probabilities at each stage have an closed-loop representation as functions solely of "payoff relevant" history encapsulated in state variables for the system over a finite period. In finite horizon games, every BPN can be formally written in closed-loop form, but this is not true of infinite-horizon games. However, the class of BPN solutions that are of MPE form with specified, low-order "payoff-relevant" histories is quite rich and behaviorally appealing, with substantial computational advantages, justifying their adoption in applications to markets with network effects.

Let $Z(t) = (n(t),r_0(t),r_1(t),a_0(t),a_1(t))$ denote the vector of exogenous variables in the model, let $S(t) = (S_0(t),B_0(t),B_1(t),K_0(t),K_1(t),Q_0(t),Q_1(t),N(t),\min(t,T^*))$ denote the vector of state variables, including the number of periods of market operation min(t,T*) where the exogenous variables were not all stationary, and let $x(t) = (i_0(t),i_1(t))$ denote the vector of policy variables. Let

(16) $$S(t+1) = h(S(t),Z(t),x(t),\varepsilon(t))$$

denote the vector of equations of motion for the state variables. Let

(17) $$\pi_j(t) = g_j(S(t),Z(t),x(t)) - (1+\zeta_j(t))i_j(t)$$

denote the profit of firm j in period t, from (9). Let f_j denote an open-loop mixed strategy for firm j; it is a conditional CDF, depending on S(t), Z(t), ε(t), and $\zeta_j(t)$, whose support is contained in the range of possible investment levels $i_j(t)$. A MPE is a profile of strategies (f_0,f_1) with the property that for firm 0, at each time t every point in the support of $f_0(\cdot|S(t),Z(t),\varepsilon(t),\zeta_0(t))$ is a maximizer of $\mathbf{E}_{-0,t} \sum_{m=t}^{H} \pi_0 \beta^{m-t}$, where the expectation is taken with respect to the strategy f_1 of its rival and the rival's private shock $\zeta_1(t)$ in the current and future periods, and all variables in future periods, including exogenous variables and public and private shocks; and a symmetric condition holds for firm 1. We will assume that the private shocks $\zeta_0(t)$ and $\zeta_1(t)$ have continuous CDF's. Then, as noted above, the mixed strategies f_j will be almost surely pure, and can be represented as functions from S(t), Z(t), ε(t), and $\zeta_j(t)$ into the interval of possible investment levels. The maximized value of expected present value, $\mathbf{E}_{-0,t} \sum_{m=t}^{H} \pi_0 \beta^{m-t}$, is denoted $V_0(S(t),Z(t),\varepsilon(t),\zeta_0(t))$ and is termed the *valuation function* of this firm in period t. Similarly, firm 1 has a valuation function $V_1(S(t),Z(t),\varepsilon(t),\zeta_1(t))$.

The MPE and the valuation functions of the firms are characterized by Bellman's backward recursion. Define

(18) $\lambda_0(S(t),Z(t),i_0(t),\varepsilon(t),\zeta_0(t))$

$$\equiv \mathbf{E}_{-0,t}\int_{X_{-0}} f_1(di_1(t)|S(t),Z(t),\varepsilon(t),\zeta_1(t))[g_0(S(t),Z(t),x(t)) - (1+\zeta_0(t))i_j(t)],$$

where X_{-0} is the domain of the policy $i_1(t)$ of firm 0's rival. This is the expected value of firm 0's current profit, given the information it has on the current state and its current policy. Define

(19) $\mu_0(S(t),Z(t),i_0(t),\varepsilon(t))$

$$\equiv \mathbf{E}_{-0,t}\int_{X_{-0}} f_1(di_1(t)|S(t),Z(t),\varepsilon(t),\zeta_1(t))V_0(h(S(t),Z(t),x(t),\varepsilon(t)),Z(t+1),\varepsilon(t+1),\zeta_0(t+1)).$$

This is the expectation as of period t of the expected present value in period t+1 of firm 0's optimal profits from period t+1 on. Then, the valuation function V_0 satisfies the Bellman equation

(20) $V_0(S(t),Z(t),\varepsilon(t),\zeta_0(t))$

$$= \quad \mathbf{max}_{i_0(t)}\{\lambda_0(S(t),Z(t),i_0(t),\varepsilon(t),\zeta_0(t)) + \beta\mu_0(S(t),Z(t),i_0(t),\varepsilon(t))\}.$$

Analogous definitions hold for firm 1.

Let θ denote the vector of parameters of the model. These include the parameters of the share equation, α_1 to α_5, and the profit functions, γ_0 and γ_1, the evaluation rate d_B on installed browsers, the discount rate β, and the parameters characterizing the Markov processes for the exogenous variables $Z(t) = (n(t),r_0(t),r_1(t),a_0(t),a_1(t))$ and shocks $\varepsilon(t)$, $\zeta_1(t)$, and $\zeta_2(t)$. We first discuss computation of the model for given parameter values, and after that estimation of the parameters by matching observed and computed features of the market trajectories.

3.5. Computation of MPE Solutions and Trajectories. To compute MPE, we follow the program developed by Pakes and McGuire (1994,2001), utilizing the adaptation by Judd (1999) of classical functional approximation by orthogonal polynomials; see Jackson

(1941), Lorentz (1966), Press *et al* (1986). The first step constructs the terminal (for H finite) or asymptotic (for H infinite) valuation function for each firm. The second step uses the Bellman equation (15) and backward recursion to find the Bayes-Nash optimal policy functions, and recursively the earlier period valuation functions. The third stage rolls the model forward one or more times, given these functions, the initial state, and draws of the random variables ε(t) and ζj(t), to compute simulated trajectories under "as is" or "but for" conditions.

The first stage is trivial when the market closes after period $H < +\infty$, since then $V_j(S(H),Z(H),\varepsilon(H),\zeta_0(H)) = 0$. However, if $H = +\infty$, or if for computation one chooses $H^* < H$, it is necessary to construct terminal or asymptotic stationary state valuation functions. One can interpret the Bellman recursion as a mapping from the space of valuation functions and strategies into itself. A fixed point of this mapping will give the valuation function for the stationary model. The computational technique is to approximate the family of valuation functions using Chebyshev polynomials, and approximate the state space using a finite grid, and then to solve the fixed point problem in the finite-dimensional space spanned by these approximations. The following paragraphs outline the basic properties of Chebyshev approximations, and go on to describe a method for augmenting the grid that is computationally straightforward and permits the approximation error to be controlled on the observed trajectory of the game.

Univariate Chebyshev orthogonal polynomials of degree j = 0,...,*m* are defined by the recursion

(21) $$T_0(x) = 1,\ T_1(x) = x, \text{ and } T_{t+1}(x) = 2xT_t(x) - T_{t-1}(x) \text{ for } t = 2,3,...$$

They satisfy the orthogonality condition $\int_{-1}^{1} T_i(x)T_j(x)(1-x^2)^{1/2}dx = \pi\delta_{ij}(1+\delta_{i1})/2$, and the roots of $T_t(x)$ are $x_{ti} = \cos(\pi(I-1/2)/t)$ for $t > 0$ and $i = 1,...,t$. Further, for $j,k \le m$, they satisfy the finite orthogonality condition

(22) $$\sum_{i=1}^{m} T_i(x_{mi})T_k(x_{mi}) = m\delta_{ik}(1+\delta_{j0})/2.$$

Approximate a valuation function V(x) on the d-dimensional cube $[-1,1]^d$ by a linear combination of these polynomials

$$(23) \qquad V(x) = \sum_{0 \le i_1 + \ldots + i_d \le m} T_{i_1}(x_1) \cdot \ldots \cdot T_{i_d}(x_d) b_{i_1 \ldots i_d} + v(x),$$

where v(t) is the approximation error. Suppose z_{mj}, j = 1,...,m+1, are the roots of the polynomial $T_{m+1}(z)$, and are used to form a grid on the cube $[-1.1]^d$ in which the coordinates in each dimension are these roots. Reindex the polynomials and coefficients with a single running index *j*, and grid points with a single running index *k*. Then, the system above evaluated at the grid points can be written

$$(24) \qquad Y_k = \sum_{j=0}^{M} A_{kj} B_j + v_k, \; k = 1,\ldots,K$$

where A_{kj} is the value of the j^{th} polynomial at the k^{th} grid point, Y_k is the value of V at the k^{th} grid point, and v_k is a residual. We allow the possibility that some polynomials are omitted, so that the number of grid points $K = (m+1)^d$ exceeds the number of polynomials *M*. If $K = M$, the system can be solved exactly, while if $K > M$, we assume the solution that minimizes the sum of squared residuals. Written in matrix notation, (24) is Y = AB + v and the least squares solution is

$$(25) \qquad B = [A'A]^{-1}A'Y.$$

The orthogonality property of Chebyshev polynomials implies that A′A is a diagonal matrix D, so that this solution can be obtained without matrix inversion. Note that D has a simple form, with diagonal elements $m^d/2^p$, where m is the order of the Chebyshev polynomials and p is the number of univariate polynomials in the product for a given term in the approximation that are of degree greater than zero.

Suppose now that we start from a Chebyshev system with K evaluation points and M polynomials, with $K \geq M$, and add k additional evaluation points and m additional polynomials, with $k \geq m$. (Note that the notation k and m for the augmentation dimensions are not the same as the running index for grid points and the order of the Chebyshev polynomials that appeared in (23).) The additional evaluation points will be data-driven rather than polynomial roots, chosen so that observed states are in the grid. Consequently, there is no advantage to taking the additional polynomials to be orthogonal or Chebyshev. For example, they may be simple powers higher than the maximum degree of the included Chebyshev polynomials.

The system we seek to solve is an augmented version of (24) which can be written in partitioned matrix form

$$(26)\qquad \begin{bmatrix} Y \\ y \end{bmatrix} = \begin{bmatrix} A & F \\ G & H \end{bmatrix}\begin{bmatrix} B \\ b \end{bmatrix} + \begin{bmatrix} \mathrm{v} \\ \mathrm{v}' \end{bmatrix},$$

where Y is K×1, y is k×1, A is the K×M array of Chebyshev polynomials evaluated at the original grid points, F is the K×m array of augmented polynomials evaluated at the original grid points, G is the k×M array of Chebyshev polynomials evaluated at the augmented grid points, H is the k×m array of augmented polynomials evaluated at the augmented grid points, B is M×1, and b is m×1. The least squares solution solves

$$(27)\qquad \begin{bmatrix} A'Y+G'y \\ F'Y+H'y \end{bmatrix} = \begin{bmatrix} A'A+G'G & A'F+G'H \\ F'A+H'G & F'F+H'H \end{bmatrix}\begin{bmatrix} B \\ b \end{bmatrix}.$$

Define the M×M matrix

$$(28)\qquad Q = D^{-1} - D^{-1}G'[I_k+GD^{-1}G']^{-1}GD^{-1}.$$

Note that Q is the inverse of $A'A + G'G \equiv D + G'G$. Since D is diagonal, the computation of Q requires only the k×k inverse $[I_k+GD^{-1}G']^{-1}$. The formula for a partitioned inverse is

(29) $$\begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}^{-1} = \begin{bmatrix} (C_{11})^{-1} + (C_{11})^{-1} C_{12} R^{-1} C_{21} (C_{11})^{-1} & -(C_{11})^{-1} C_{12} R^{-1} \\ -R^{-1} C_{21} (C_{11})^{-1} & R^{-1} \end{bmatrix},$$

where $R = C_{22} - C_{21}(C_{11})^{-1}C_{12}$. Apply the formula (29) to solve for the Chebyshev coefficients in (26). Note that

$$C_{11} = A'A + G'G, \text{ implying } (C_{11})^{-1} = Q,$$
$$C_{12} = A'F + G'H$$
$$C_{21} = F'A + H'G$$
$$C_{22} = F'F + H'H$$

Then R is the m×m matrix

$$R = F'F + H'H - (F'A + H'G)Q(A'F + G'H).$$

Computation of the solution to (26) requires inversion of the m×m matrix R, as well as computation of the matrix Q which contains the k×k inverse $[I_k+GD^{-1}G']^{-1}$. The matrices to be inverted are symmetric positive definite, so a fast Cholesky inversion method can be used. When m = k = 1, no matrix inversion is required.

The first stage computation of the valuation functions can now be summarized for the case that $H = +\infty$. For the finite grid of evaluation points in the augmented Chebyshev approximation, trial values Y for the valuation functions and trial values for optimal policy functions map via the Bellman recursion (15) into new trial values for Y and the optimal policy responses for each firm in (15). A fixed point of this mapping is then an approximation to the stationary state valuation function and policy functions. The policy functions and valuation functions are uniformly bounded, so that the domain of this mapping is compact and convex. Then, a fixed point exists, and may be approximated using a gradient search, although convergence is not guaranteed and in practice can be problematic. An alternative is to use a pivot method of the Scarf type, which would guarantee convergence, but would be computationally demanding. In our application, the fixed point approximation will be simplified first because the exogenous variables n(t), r_0(t), r_1(t), a_0(t), and a_1(t) are assumed to be constant in the stationary state, and do not become added dimensions of the grid, and second because at the critical grid points where cumulative investment levels are large, the optimal policy response of both firms is zero.

The second stage of the MPE calculation is to recurse the model backward, from a given finite H with terminal valuation functions that are identically zero, or a computational terminal period H* at which the stationary valuation function is calculated from the first stage, using the Bellman recursion (15) and finding the fixed point for the profile of optimal policy functions. These will be modified only through dependence on new levels of exogenous variables that did not appear in later periods.

The third stage of the MPE calculation is to roll the model forward, starting from the initial state of the system, with draws from the distributions of shocks, using the optimal policy functions obtained in the second stage. This produces trajectories for the state and policy variables in the model. These calculations, done in parallel for “as is” and “but for” settings for exogenous variables, provide comparative trajectories that then can be used as a basis for approximating the expected present value of lost profit due to “bad acts” by one firm.

3.6. Estimation Strategies. The MPE model developed in Section 3.3 has a vector of parameters $\theta = (\alpha_1.\alpha_2,\alpha_3,\alpha_4,\alpha_5,\gamma_0,\gamma_1)$ that appear in the specification of the equations of motion and the profit functions of the firms. In addition, it contains the parameters d_B and β that can be calibrated from external sources, and the parameters characterizing the Markov processes for the exogenous variables $Z(t) = (n(t),r_0(t),r_1(t),a_0(t),a_1(t))$ and shocks ε(t), ζ_1(t), and ζ_2(t).

We divide the econometric problem into <u>three</u> stages, as is done by Bajari, Benkard, and Levin (2003). The first stage is structural estimation of the equation of motion and profit functions. This is done prior to computation of the MPE, so that at this stage the firms’ policy functions are not specified structurally. The second stage is computation of the MPE, incorporating estimation of parameters not identified in the first stage, using a minimum distance criterion to match observed and computed trajectories of states and policies. The third stage is a one-step iteration of both first and second stage estimates to obtain minimum distance estimators for the complete system.

The motivation for this estimation sequence is that the equation of motion and profit functions contain most of the parameters and shocks of the system, and their estimation within the limits of identification determine these partially or fully. The parameters estimated consistently in the first stage can be held fixed during stage two, substantially reducing the burden of estimation search within the computation of the MPE. Finally, all the parameters of the system are updated in a single stage three iteration that does not require extensive recomputation of the MPE. A one-step theorem (Newey-McFadden, 1994) will assure that the three-stage procedure is asymptotically equivalent to direct minimum distance estimation of the full system.

In the first stage estimation, the equations of motion and profit functions in general include both exogenous shocks, and current policies that the firms choose knowing these shocks, so that the current policy variables are endogenous. To deal with this, reduced form estimates of the policy functions may be usable as instrumental variables, but several econometric issues are involved. First, the policy variable reduced form equations are in general not parametrically specified at this stage, even if the equation of motion and profit functions are fully parametric, since their form depends on the solution of the game. Second, natural specifications of the structural equations may involve non-additive shocks, making it necessary to use econometric methods appropriate to this situation. Third, identification will depend on functional restrictions or the availability of exogenous variables that influence the determination of policy choices, but are otherwise excluded from some of the structural relationships. Suitable instruments will be variables observed retrospectively by the econometrician that are related to the private information of the firms, but independent of market-level shocks. Note that leads and lags of observed exogenous variables, and nonlinear transformations of predetermined variables, are not in general potential instruments for this problem. Under the Markov assumptions, lagged exogenous variables are not predictive given current values. Leading exogenous variables should also not be predictive, since otherwise the players would form these predictions and incorporate them into their behavior. Finally, non-linear transformations of predetermined variables are valid instruments only if one can be confident of the functional specification and their exclusion within this specification. It the equations of motion and profit functions are fully identified, then the empirical distributions formed from these equation residuals are consistent estimator of the distributions for the shocks $\varepsilon(t)$, $\zeta_1(t)$, and $\zeta_2(t)$.

Consider the first-stage identification and estimation of the equation of motion (8) and the profit functions (9) in the model in Section 3.3. In this equation of motion,

$$(30) \qquad \begin{aligned} s_0{}^*(t) - \tfrac{1}{2} &= \alpha_1(S_0(t) - \tfrac{1}{2}) + \alpha_2(Q_0(t)-Q_1(t)) + \alpha_3 IAP(t) + \alpha_4 WINDOW(t) + \varepsilon(t) \\ s_0(t) &= \rho(0,1,s_0{}^*(t)), \end{aligned}$$

the explanatory variables $S_0(t)$ and $Q_0(t)-Q_1(t)$ are predetermined, and IAP(t) and WINDOW(t) are exogenous. If our assumption is correct that $\varepsilon(t)$ is independent across time, and the support of $\varepsilon(t)$ is sufficiently narrow so that at the observed configurations of explanatory variables, the dependent variable satisfies $0 < s_0(t) < 1$ with probability one, then (30) can be estimated consistently by ordinary least squares. Otherwise, if $\varepsilon(t)$ is serially correlated, lagged exogenous variables in the system can be used as instruments for $S_0(t)$ and $Q_0(t)-Q_1(t)$. If at some observations there is a positive probability that $s_0{}^*(t)$ lies outside (0,1), then (30) has a two-sided Tobit form. Letting $w(t) \equiv \tfrac{1}{2} + \alpha_1(S_0(t) - \tfrac{1}{2}) + \alpha_2(Q_0(t)-Q_1(t)) + \alpha_3 IAP(t) + \alpha_4 WINDOW(t)$ as a shorthand, the observations not at the boundaries satisfy

$$(31) \qquad s_0(t) = w(t) - w(t)[G(1-w(t)) - G(-w(t)] - \int_{-w(t)}^{1-w(t)} G(\varepsilon)d\varepsilon + \xi(t),$$

where G is the CDF of $\varepsilon(t)$ and $\xi(t)$ is a disturbance that is orthogonal to the remaining terms in (31). If G is fully parameterized, then (31) can be estimated by nonlinear least squares, or if instruments are required, by GMM. If G is not parametric, consistent estimation requires the methods of Appendix 2.

Next consider estimation of the profit functions $\pi_j = r_j(t)B_j(t) - (1+\zeta_j(t))i_j(t) - \gamma i_{jj}(t)^2$. Rewrite these as

$$(32) \qquad [\pi_j - r_j(t)B_j(t) - i_j(t)]/i_j(t) = -\gamma i_{jj}(t) - \zeta_j(t).$$

This equation has an additive disturbance, but the policy variable $i_j(t)$ is determined in the MPE solution of the model as a function $i_j(t) = I_j(S(t),Z(t),\varepsilon(t),\zeta_j(t))$ that depends on $\zeta_j(t)$, and hence is endogenous. Potential instruments for $i_j(t)$ include the investment policies of rivals, which are independent of $\zeta_j(t)$. Appendix 2 discusses estimation of generalizations of equations like (32) using the methods of Matzkin (2004).

Let ψ denote the subvector of parameters that were not identified in stage one. In stage two, with estimates of the identified stage one parameters, trial values for ψ, and draws from the empirical distributions of the shocks, the MPE is calculated using the algorithms of Section 3.5, and trajectories are computed. The parameters ψ are chosen to minimize a distance criterion defined in terms of differences of observed and calculated states and policies over observed partial trajectories. (Alternately, we could have used the distance of the calculated MPE first-order conditions from zero when evaluated at the observed policy variables.) This is a form of Method of Simulated Moments estimation (McFadden, 1989), and the protocols of that method are followed to ensure consistent estimates of the identifiable ψ parameters. (Unidentified ψ parameters have no effect on model trajectories, and can be arbitrarily normalized without loss of generality.) In practice, we carry out the MPE dynamic computations using an iterative search for the distance-minimizing parameters. Such a search is facilitated by the smooth behavior of the valuation functions and policy functions with small changes in ψ, provided the solutions to Bellman's equations are regular. Schematically, we have

$$(33) \qquad V_{jt}(s,z,\theta) \approx A(s,z)B_{jt}(s,z,\theta),$$

where A(s,z) is the augmented polynomial approximation (26). Suppose one starts from a solution to the MPE problem at an initial ψ^1. Differentiate the first-order conditions for the

Bellman equations and the equations of motion to determine the derivatives of the B's and x's with respect to ψ, and plug these derivatives into the expression for the derivative of the estimation criterion with respect to ψ. For convenience, the derivatives can be done numerically. Then, do a gradient search to minimize the criterion. In principle, it is not necessary to re-solve the original problem, but in practice, periodic resolving should be done to correct the cumulative drift in the approximation from the MPE solution.

One major difference between our formulation of the estimation problem and that of Bajari *et al (2003)* is that they concentrate on models where the current payoff functions and equations of motion are linear in unknown parameters. That is not necessarily the case for the model in our application, due to the boundaries, adding to the computational burden of finding the Nash fixed points. Also, we do not require symmetry for the Nash solution, as that is appropriate only if all the players are similarly situated. In the application to Netscape and Microsoft, there were substantial differences in the firms that can not be conveniently described by model variables.

The final stage of estimation is a one-step Gauss-Newton iteration in all parameters to reduce the generalized distance. Asymptotic statistical theory implies that a single step from initially consistent estimators, utilizing estimation of second derivatives of the generalized distance with respect to the parameters, yields estimators that are efficient within the class of minimum distance estimators, see Newey-McFadden (1994). A side benefit of this procedure is that it gives ready estimates of asymptotic standard errors of the parameter estimates. (If the number of simulation repetitions does not grow with sample size, then there is some loss of asymptotic efficiency.) In practice, linear search within the final one-step iteration may be needed to guarantee an improvement in minimum distance in a finite sample.

Some studies using MPE models have found that it is computationally advantageous to use polynomial approximations for the policy functions as well as the valuation functions. We have not done so in this paper, but note that for problems where convergence of the numerical search to a fixed point is difficult, such approximations may facilitate solution. Bajari *et al* (2003) estimate policy functions as well as the equation of motion in the first stage, assuming that these functions are linear in parameters. Because these policy functions depend structurally on the valuation functions derived in the second stage, all one can hope to get at the first stage are (nonparametric) reduced forms. The Bajari estimators may nonetheless be useful, as their estimated values may be proper instruments for the policy variables that appear endogenously in the equation of motion, and they may provide good starting values for second-stage computations.

3.7 Model Parameters and Trajectories. We estimate the equation of motion (30) by ordinary least squares. The results are given in Table 3. We note that further econometric analysis is required to obtain consistent estimators if either serial correlation is present or there is a positive probability of hitting a boundary. The parameters γ_0 and γ_1 in the profit functions are not estimated in stage 1 because profits are not observed in our database. The parameter $d_B = 0.3$ is calibrated from data on the rate at which operating systems are upgraded or computers replaced. The parameter $\beta = 0.9$ is selected near the upper limit of stability for our current Chebyshev approximation, and is below Netscape's actual quarterly rate of return to financial capital

Table 3. OLS Estimates of Equation of Motion Parameters

Variable	Variable		Coeff.	Std.	T-Stat	P-Value
	Mean	Std. Dev.		Error		
Netscape new browser share (depend. variable, $s_0(t)$ - ½)	-0.05	0.2715	---	---	---	---
Base share ($S_0(t)$ - ½)	0.035	0.275	0.7725	0.044	17.58	0.00
Quality difference (IINVDIF)	0.017	0.0935	0.2855	0.1049	2.72	0.01
IAP	0.0032	0.0062	-4.9276	1.1544	-4.27	0.00
WINDOW	0.5852	0.4566	-0.1161	0.011	-11.00	0.00
Std. Dev. of Residuals	0.0326					
R-Squared	0.9873					

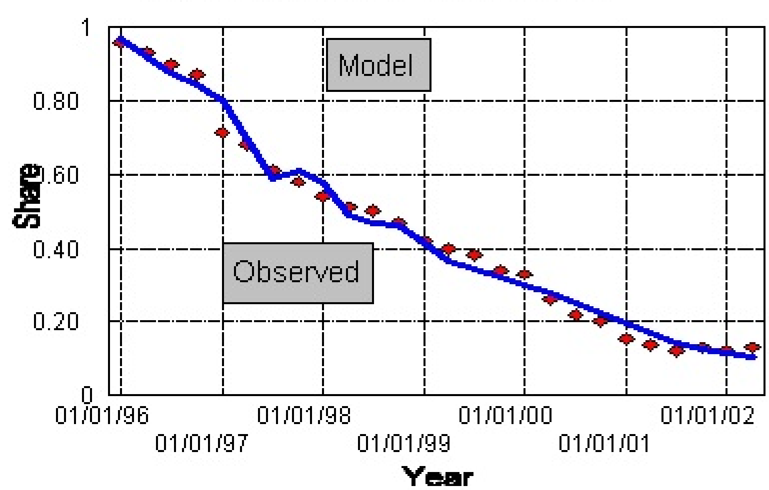
The observed values of $s_0(t)$ range from a minimum of 0.12 to a maximum of 0.96. The empirical distribution of the residuals from the Table 3 model is contained in the interval

[-0.085,0.038]. This supports informally the hypothesis that the support of ε(t) is sufficiently narrow so that the correction (31) to the OLS regression is unnecessary. A Breusch-Godfrey test for serial independence is accepted. We also test for the presence of a constant term in the model in Table 3, and accept the hypothesis that it is zero. We note however that in the presence of a constant term, the WINDOW coefficient becomes insignificant. Consequently, it is difficult to distinguish econometrically between the hypothesis that Microsoft's penetration of the market was due to the tying of IE to the operating system, and the alternative hypothesis that consumers have a "brand preference" for Microsoft.

The coefficients of the model are of expected sign, with a significant positive network effect that is less than one, so that the model tends to an equilibrium share. If the "bad acts" variables are eventually turned off, and both browsers converge to the same quality level, then the Netscape share will be centered around ½. However, If the tying of the operating system and IE continues indefinitely, so that WINDOW remains one, then the Netscape share will tend toward $½ + \alpha_4/(1-\alpha_1) = -0.0104$, which corresponds to a complete tipping of the market to Microsoft, aside from sufficiently positive ε(t) shocks that would keep Netscape in the market at a small expected share.

Figure 5 plots the "as is" predicted model market shares and the observed shares for new browsers; the model with stage one estimated parameters tracks well the observed trajectory of sales share, but does not at this stage impose consistency requirements between actual and predicted investment.

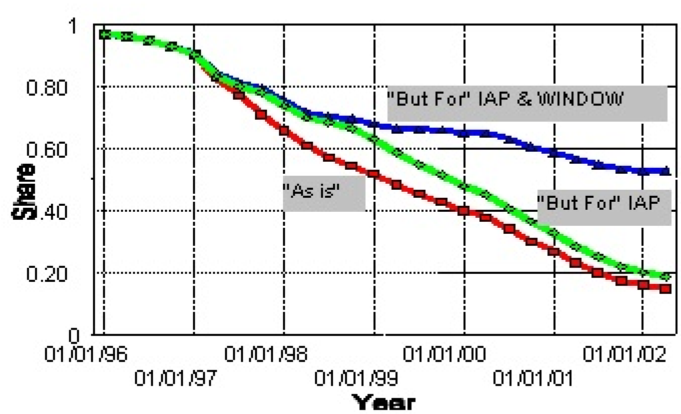
Figure 5. Netscape Market Share
Observed and Estimated



It is premature to consider “but for” scenarios, as we have not yet done the MPE model calculations required to predict firm policies in these alternative environments. Nevertheless, it is somewhat useful to calculate trajectories for Netscape market share under “but for” alternatives, holding the firm’s investment levels fixed at their “as is” levels. This gives some indication of the partial effect of “bad acts” before the firms adjust their investment policies to their absence. Figure 6 plots these “as is” and “but for” predicted models. In these trajectories, the “but for” predictions assume that investment and quality levels remain at “as is” levels. This misstates the “but for” scenario obtained from the MPE solution, in which Microsoft will optimize its R&D investment taking into account the additional difficulty of overcoming the incumbent advantage without using “bad acts”. On the basis of this partial analysis, one would conclude that the effect of IAP was relatively temporary, causing Netscape’s share to fall two to four quarters earlier than it would have otherwise, but having no effect on asymptotic shares. On the other hand, WINDOW was critical to tipping the market to Microsoft; otherwise as a consequence of the modeling assumption that the firms were symmetrically positioned in the browser market once all quality and network effects are accounted for, shares would have tended asymptotically to equality. Of course, in addition to these conclusions being based on a partial analysis, there are two other reasons to interpret them with caution. First, the data are interpolated from partial public records, and are not consistently reliable. Second, only first stage model estimates are used, and the estimated models are not very robust.

Figure 6. Netscape Share

"As is" and "But For" (without MPE adjustments)



We next compute a MPE solution to the model, using the steps described in Section 3.5, and carry out the stage 2 and 3 estimation steps to obtain model parameters. Since we do not have data on the browser profits of the firms, we are unable to recover empirical distributions for the private information variables ζ_0 and ζ_1. Table 4 gives the parameter estimates obtained from the third estimation stage.

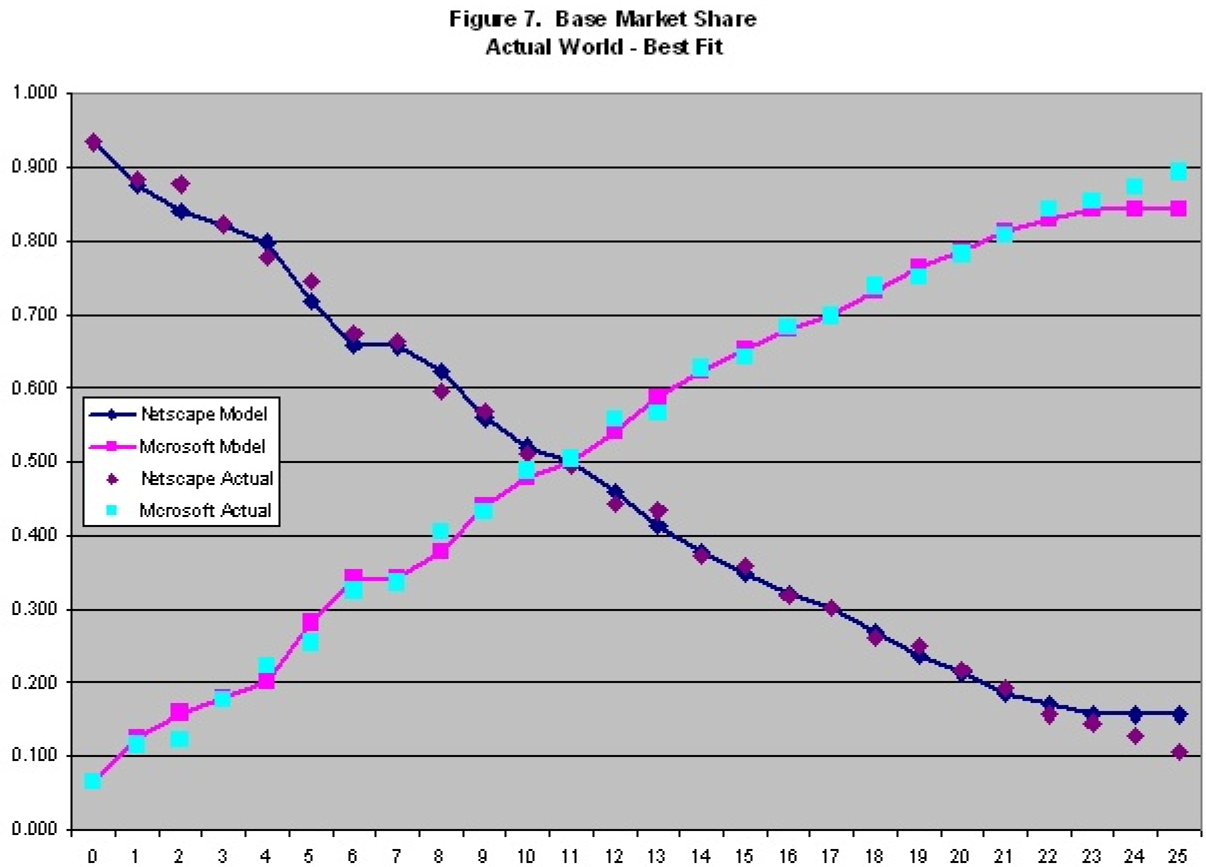
Table 4. Parameter Estimates from Stages 1 and 3

Parameter	Symbol	Stage 1	Stage 3
Base share, $S_0(t)$ - ½	α_1	0.7725	0.4761
Quality difference, $Q_0(t)$ - $Q_1(t)$	α_2	0.2855	0.3563
IAP	α_3	-4.9276	-5.8218
WINDOW	α_4	-0.1161	-0.1399
Netscape Invest. Cost	γ_0	----	1
Microsoft Invest. Cost	γ_1	----	0.05
Browser replacement rate (calibrated)	d_B	0.3	0.3
Discount rate (calibrated)	β	0.9	0.9

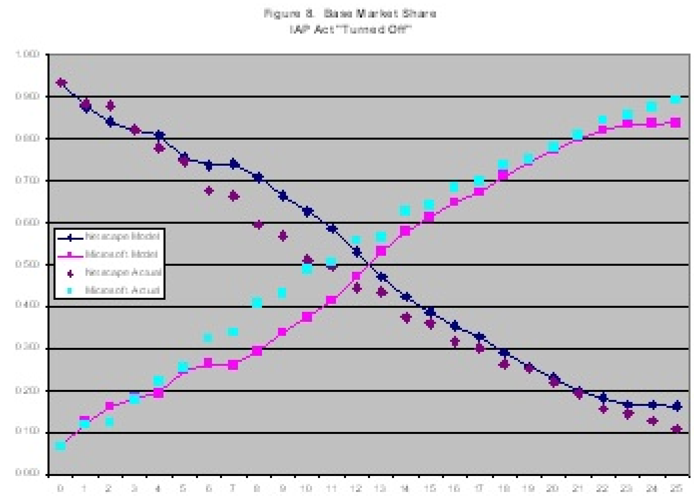
Estimation from fitted trajectories imposing the consistency requirements of MPE equilibrium investment policy has an economically important impact on parameter estimates. The network effect is weaker, and the effect of quality is stronger. The effects of “bad acts” are stronger. We have not yet computed standard error estimates for the stage 3 parameter estimates, but note that for several reasons they are not necessarily smaller than those associated with the stage 1 estimates. First, stage 3 standard errors will be obtained by a bootstrap procedure, and will incorporate simulation error and to some extent, finite sample error. Second, the stage 3 estimates impose additional information, from investment, and additional restrictions, from the requirement that the observed trajectory approximate a MPE solution.

Figure 7 gives the observed and fitted firm installed base shares from the model fitted to the “as is” trajectory, and the stage 3 parameter estimates. The first quarters of 1996, 1998, 2000, and 2002 are numbered 0, 8, 16, and 24, respectively in this chart. With the exception of some deviation of the model from actual shares in the final two quarters, the model fits well the relatively smooth decline in Netscape share over the observation period.

3.8. Market "Tipping" and Damages from Microsoft "Bad Acts". Once the MPE model has been estimated, so that it fits baseline "as is" trajectories over the observed history of



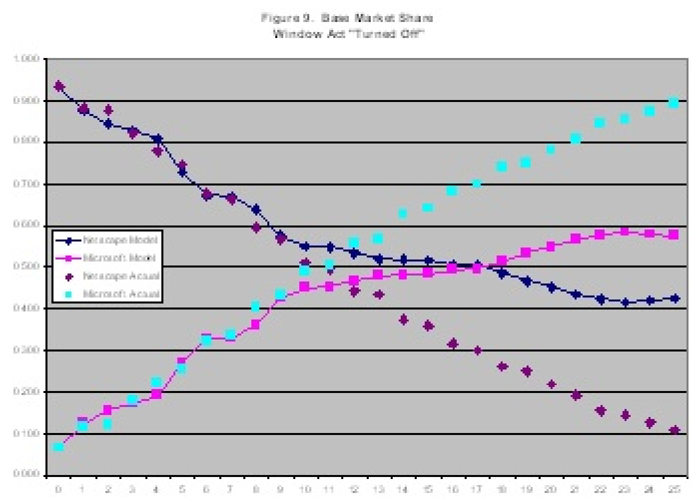
the market, it can be used to estimate trajectories under alternative "but for" conditions where "bad acts" are excluded, and from this estimate damages attributable to these "bad acts". A few precautions, applicable to many damage studies, apply here. First, prediction of trajectories in these models generally requires simulation of unobserved shocks. For example, the empirical distribution of the fitted shocks obtained in estimation of the model may be the basis for simulation draws. To reduce the variance of damage estimates, common draws should be used under baseline and "but for" conditions. Second, if the natural experiment provided by the observed market trajectory is insufficient to identify reliably the demand response to changes in "bad acts", then it may be necessary to use external studies, such as market research studies, to estimate these effects. Third, the statistical reliability of damage estimates should be provided as part of the analysis. In principle, the delta-method can be applied to linearize the model in its parameters around their fitted values, and this linearization can be combined with the asymptotic covariance estimate for the parameter estimates from the estimation phase to produce standard errors for the damage estimates. In practice, a bootstrapping procedure, with a common resample used for both baseline and "but for" scenarios, is likely to do a better job of capturing the second-order non-linearities of model, and is thus likely to be more robust.



We next solve the model for an MPE with one or both of the Microsoft "bad acts" eliminated. Figure 8 shows the effect of removing the IAP restriction, where Microsoft contracts with internet access providers required that they not provide Navigator. The figure compares the actual base shares trajectories with the "but for" prediction when IAP is always zero, but the WINDOW variable, reflecting the tying of Internet Explorer to the Windows operating system, remaining in place. The figure shows that the effect of IAP was to accelerate the decline of Netscape's share from mid-1997 until the end of 1999 by one to three quarters, but that even in the absence of IAP, the market would have eventually tipped to Microsoft as a result of the WINDOW tying of Internet Explorer to the operating system.

Next consider the effect of removing the WINDOW tying arrangement, but leaving the IAP exclusionary contracts in place. The predicted MPE trajectories in this scenario are given in Figure 9. Here, the market tends in the long run to equal shares, but the effect of IAP and the Microsoft investment program to improve the quality of IE would have been to make Microsoft the browser market leader starting in mid-1999, with a base share rising to nearly 60 percent.

Finally, Figure 10 gives the predicted market share trajectories with both IAP and WINDOW removed. In this case, the two firms tend to long run equal market shares, but with Microsoft overtaking Netscape in mid-2000 and attaining a share near 58 percent due to a quality difference between Internet Explorer and Navigator. The investment pattern is related to the model parameters for the cost of "crash R&D", with Netscape's cost γ_0 = 1.00 much higher than Microsoft's cost γ_1 = 0.05. The MPE model's overall ability to predict investment trajectories is modest, and the investment data are problematic, so that this parameter difference must be interpreted with caution. However, it seems to indicate that the Microsoft organization was able to take on R&D projects at a substantially larger scale than Netscape without substantial diminishing returns. This implication is consistent with the relative sizes of the two companies, and the presumable ability of Microsoft to move software developers between browser and operating system projects. We conclude that the effect of the WINDOW tying of Internet Explorer to the Windows operating system "tipped" the browser market to Microsoft, and that in the absence of this "bad act" and the IAP exclusionary contracts, Netscape would have lost browser share more slowly, retaining the largest share until mid-2000, and retaining a market share above 40 percent, with a very long run tendency toward equal shares.



We have indicated that the MPE model is only modestly successful in fitting the observed "as is" investment trajectories. A significant issue is the difficulty in determining both R&D investment levels and timing for Netscape and Microsoft. Public records do not give a reliable breakdown of total investment for either firm between browser development and development of other products. This is a particular problem for Microsoft, where the lines of operating system and browser development are blurred, and for both firms, where the development of browsers occurred in parallel with development of enterprise server software products and web developer tools. Also, because we have no public profit center data for browsers, we do not observe profits for the two firms from browser operations, and hence are unable to refine the models for investment cost for the firms, or identify the effects of private investment costs. Finally, lack of consistent independent data on product quality does not allow us to determine empirically the relationship between cumulative R&D investment and browser quality as perceived by consumers. With these caveats, we give in Figures 11-14 the observed and model trajectories for investment for the two firms, first in the "as is" case, and then in the respective "but for" scenarios without IAP, without WINDOW, and without both.

Figure 10. Base Market Share
Both Acts "Turned Off"

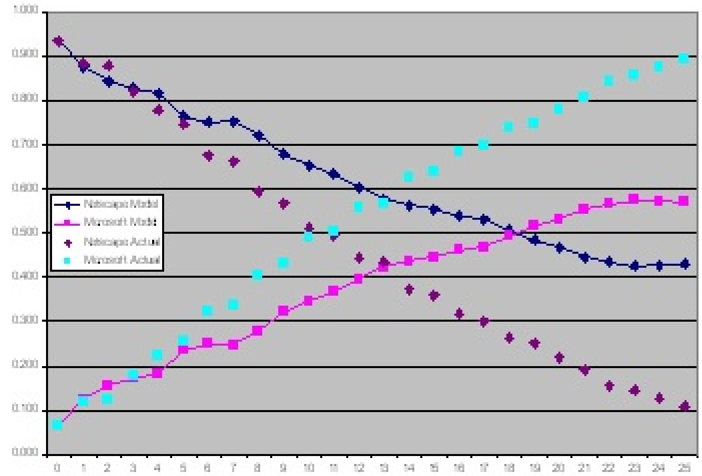
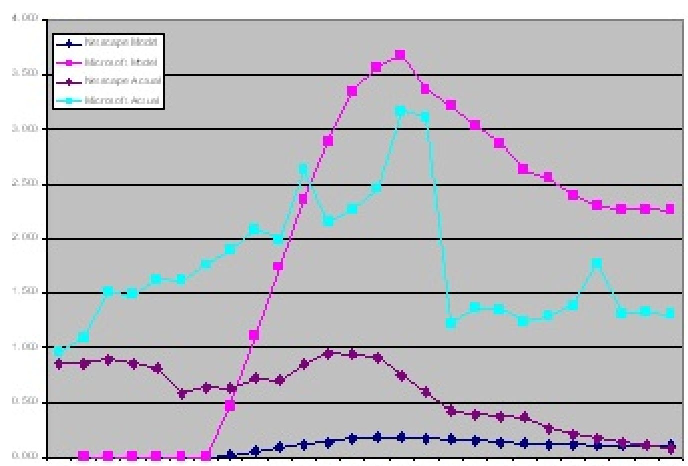


Figure 11. Quarterly Investment
Actual World



The optimal investment trajectories from the MPE solution to the model have the gross qualitative features of the observed trajectories, with a rapid runup in Microsoft R&D investment, followed by a substantial decline once Navigator was no longer a factor in the market. However, the MPE solution shows a delay in investment by both firms beyond initial levels that is not observed, and a lower level of Netscape investment than observed. This presumably occurs in the MPE solution because with a relatively steep discount rate and the prospect that Microsoft could tip the market in its favor, it was in the interest of Netscape in the "as is" world to abandon the market to Microsoft rather than contesting it, extracting as much profit as possible in the short run from its existing base.

In Figure 12, with WINDOW in place but IAP removed, the MPE optimal investment strategies are virtually unchanged from the "as is" world. However, Figures 13 and 14, with WINDOW removed, show greater Microsoft investment from mid-1998 to the end of 2000, the result of an

Figure 13. Quarterly Investment
Window Act "Turned Off"

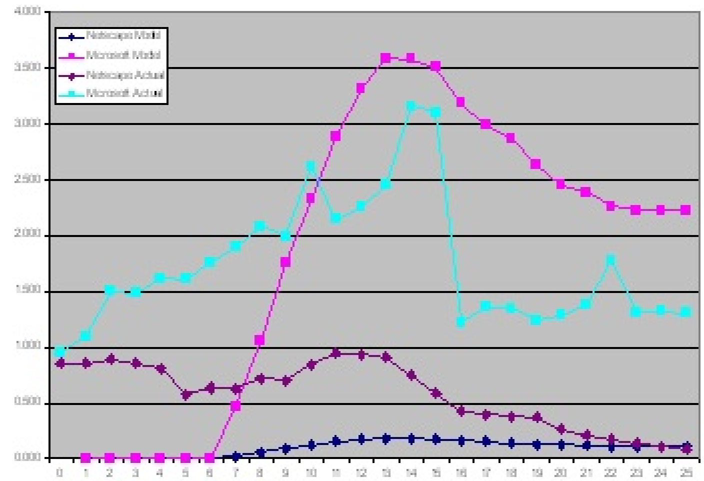


Figure 12. Quarterly Investment
IAP Act "Turned Off"

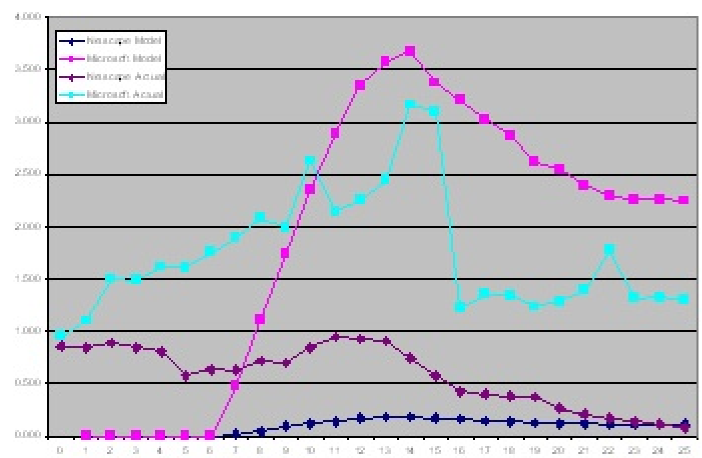
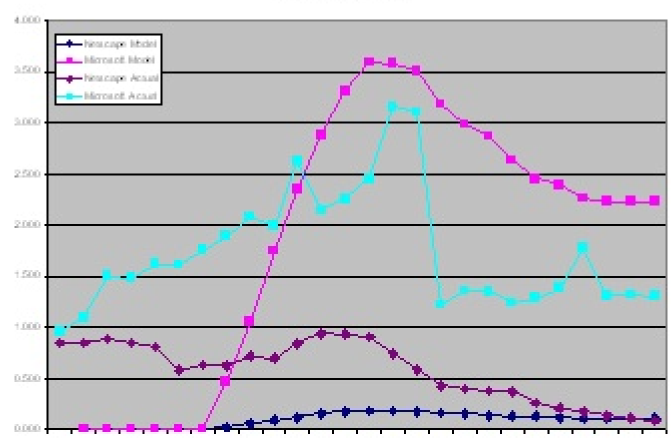


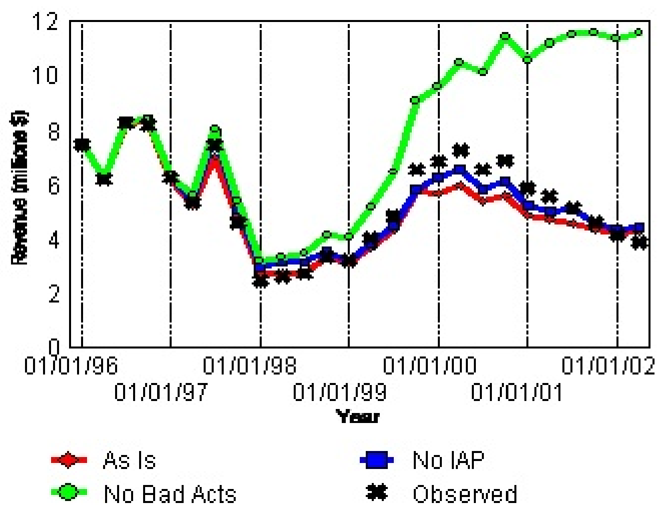
Figure 14. Quarterly Investment
Both Acts "Turned Off"



increased effort by Microsoft to penetrate the market by improving the quality of Internet Explorer when the option of tying it to the operating system is removed. This points out a possible impact on consumer welfare of “bad acts” by a firm in a concentrated market – these acts may substitute for quality competition that benefits consumers.

It is possible to use the projected trajectories for base shares and investment under “as is” and “but for” conditions to estimate trajectories of Netscape revenues, assuming that the revenue per browser used remains at its observed trajectory in the “as is” world, and from these trajectories estimate the lost revenue to Netscape from these acts. We have done this, and in Figure 15 plot the Netscape revenue estimates “as is” and “but for” IAP, or both IAP and WINDOW, for the observation period, the first quarter of 1996 through the second quarter of 2002. From these trajectories and their stationary state extension into the future, the expected present value, as of the second quarter 2002, of Netscape browser revenue calculated at a quarterly discount rate of 2.5 percent is $355 million under “as is” conditions. The present value of lost revenue from IAP, with WINDOW remaining in place, is $21 million, and the present value of lost revenue from both IAP and WINDOW is $365 million. These are then rough estimates of damages to Netscape in the browser market from Microsoft “bad acts”. A number of caveats are necessary. These estimates do not include losses to Netscape’s enterprise software and server software businesses, the potentially large but difficult to measure loss in option value to Netscape from its potential as a middleware platform that would have been directly competitive to the architecture of personal computing and the Windows operating system monopoly, or losses associated with the eventual sale and restructuring of the diminished Netscape business. Further, these numbers come from problematic data pieced together from public sources, and from a model of limited robustness. With these cautions in mind,

Figure 15. Netscape Browser Revenues



we note that the Netscape-Microsoft settlement in June 2002 for approximately $750 million was about three-forths of the trebled Netscape lost revenue in the browser market, and perhaps about one-half of trebled damages if other provable Netscape losses were accounted for. Thus, this settlement appears to be at the upper end of the typical range for civil anti-trust settlements.

4. Conclusions

The results of this paper show that it is feasible to use the tools of Markov-perfect equilibrium models combined with observed trajectories for firms in a concentrated market to estimate models and approximate Markov perfect equilibria to assess the impact of firm conduct and estimate damages from “bad acts”. Employing an MPE model that was motivated by observed characteristics of the browser market, and piecing together from public sources data on the browser war between Netscape and Microsoft, we have estimated the effects of network externalities and product quality on market trajectories, and used it to provide predictions of market outcomes under “but for” scenarios in which particular “bad acts” are eliminated. We conclude that Microsoft’s exclusionary contracts with Internet Access Providers, found illegal by the Appeals Court in U.S. v. Microsoft, was by itself a modest source of Netscape lost revenue, about $21 million before trebling. However, Microsoft’s tying of Internet Explorer to the Windows operating system, and the arrangements under which it was difficult or inconvenient for OEM’s to preinstall another browser, was sufficient to tip the market to Microsoft, and the source of substantial Netscape lost revenue, about $369 million before trebling. The issue of the legality of Microsoft’s tying conduct was remanded by the Appeals Court, and DOJ’s settlement with Microsoft left it unresolved. Had the Netscape civil anti-trust case gone to trial, this would have been the critical issue in dispute.

APPENDICES ON MARKOV PERFECT EQUILIBRIUM

Appendix 1. Game Theoretic Foundations

A1.1. *Dynamic Stochastic Games with Private Information.* A time-linked concentrated market for heterogeneous goods can be modeled as a dynamic stochastic game, whose elements are the market participants and their policy spaces, information sets, and payoff functions, the state of the system, and its equation of motion. Both stochastic games and their application to dynamic markets have been the subject of extensive literatures; see Erickson-Pakes (1995), Fudenberg-Tirole (2000), Kydland (1975), Maskin-Tirole (1988a,b), Pakes-McGuire (1994, 2001), Tirole (2003). We re-examine with an econometrician's eye the role of various assumptions on information and structure that give equilibria in pure strategies, with the objective of specializing the general stochastic game model to a form that is useful for computation and econometric analysis. Several basic theoretical, econometric, and practical considerations enter our model specification:

(1) The Markov Perfect Equilibrium modeling approach that we adopt restricts the information of firms to the "payoff-relevant" history of the market, summarized in state variables that follow a first-order Markov process, and restricts response to "closed-loop" or "feedback" functions of each firm's information. Important motivations for the MPE assumption are econometric and computational tractability. However, the MPE approach is also behaviorally plausible in its relatively limited requirements for commitment, coordination, and rationality. Within the MPE framework, it is possible to obtain a rich variety of models by varying assumptions on what firms know about the structure of the market and about exogenous shocks, what information is public or private, and what firms believe about their rivals' information and behavior. We will assume below that the structure of the market game is stationary given observed external driving variables that themselves become stationary some time before observations on market trajectories end, and justify this as a necessary condition for obtaining an identifiable model that is useful for applications. An implication of this restriction is that the Markov assumption, which ensures that a finite-dimensional state vector is a sufficient statistic for the relevant history of the game, is not a substantial additional restriction.[37] However, the restriction of Markov states to "payoff-relevant" history, and the restricted dimensionality of states in any practical model, significantly limit the possible solutions. In principle, one can test a particular MPE specification against more general solutions within the GMM estimation framework discussed later, and if necessary augment the state variables in the Markov model to accommodate observed strategic behavior.

---

[37]Technically, our model is a *continuous-state hidden Markov process* when time is discrete, and a *hidden Markov random field* when time is continuous. Kunsch-Geman-Kehagias (1995) show that the class of stationary hidden Markov random fields is dense in the space of all stationary stochastic processes.

(2) In applications, one will have more or less complete observations on market state and on firms' policies over a specific time interval. This may be augmented with data from outside the market that is informative on structural features of the market, such as market research studies on consumer preferences and engineering studies of production technology, and external factors that influence the market and potentially provide natural experiments, such as changing levels of foreign trade and other demand shifters. A critical criterion for MPE model specification is that the model be rich enough to encompass these observations, and lean enough so that these observations are sufficient to identify the model structure and enable estimation of "but for" market outcomes. The first of these requirements will be met by introducing exogenous shocks that correspond to the non-systematic statistical variations present in market data. In this connection, it is necessary to detail assumptions on what firms know and when they know it, and what the econometrician knows from the observed historical record. Note that when policy choices are made, shocks may be public information to all firms, private information to a single firm, or unknown to all firms. Further, the econometrician may be unable to observe some information that was known to the firms during historical market operation, and in other cases may be able to observe retrospectively information that was unknown to the firms at the time decisions were made. In general, shocks that are known to firms will enter their policy response functions, while shocks that are unknown are omitted from their policy functions but enter their outcome functions. Further, the form of policy functions will be influenced by beliefs about unknown shocks. A classic literature in econometrics provides a template for these considerations; see Mundlak and Hoch (1965), and Fuss-McFadden-Mundlak (1978). The second requirement, that the structure of the MPE model be identified from observations on the market over a specified time interval, places conditions on what firms can know that the econometrician does not. Events subsequent to a firm's policy choice can influence those choices only if they are predictable by the firm at the time of the choice. A successful econometric model of firm behavior must incorporate the predictions of events that do influence choice. Thus, it is critical that the econometrician have all the information that is available to firms and predictive of future events. Put the other way around, information that is available to the firm and not to the econometrician must be non-predictive, or in econometric terminology, Granger non-causal, given the information that is common to the firm and the econometrician. With this assumption, events beyond the observation period that are not predictable by the econometrician cannot influence observed behavior, and by Occam's razor should be excluded from the model. A natural assumption consistent with this conclusion is that the market becomes stationary before the end of the observation period, once all predictable effects are built into the state vector.

(3) Private information for each firm that enters the cost of its policy will, under quite general conditions, be sufficient to guarantee that MPE are almost surely pure. This property will, in turn, guarantee reasonable approximation properties for restricted classes of continuous policy functions; e.g., polynomials.

(4) MPE are unique only under extremely restrictive assumptions that are difficult to meet in applications. When there are multiple MPE, there are substantive issues of how the players could use tacit coordination to select among them, and what econometric analysis can identify regarding selection and coordination rules. We seek to resolve the indeterminancies of selection and coordination econometrically.

With these considerations in mind, we write down a dynamic concentrated market model in fairly general terms that can be specialized to the browser war model in Section 3.3. All factors that might make the model time-variant are incorporated into vectors of exogenous variables, so that the equation of motion and payoffs in the game, given these exogenous variables, are time-invariant. Let n = 1,...,N index firms, which may be incumbents or potential entrants. Let t = 0,1,...,T,...,H index time periods, with $T < H \le +\infty$, where the econometrician observes the market in periods t = 1,...,T and H is a horizon after which the market is closed. Let $T^* \le H$ denote a period after which all exogenous variables are stationary. The market is described by the following notation and accompanying assumptions:

- s_t — a *state* vector in a compact subset S of a finite-dimensional Euclidean space that is endowed with its Borel σ-field and relative Lebesgue measure
- z_t — a vector of observable exogenous variables in a compact subset of a finite-dimensional Euclidean space Z that is endowed with its Borel σ-field and a probability g_{zt}. The vector z_t will incorporate all time-varying elements in the dynamic process, so that the model conditioned on z_t is time-invariant.
- ε_t,η_t — public exogenous shocks in a compact subset Ω of a finite-dimensional Euclidean space endowed with its Borel σ-field and probabilities g_ε and g_η.
- ζ_{nt} — a private firm n exogenous shock in a non-degenerate compact rectangle Λ in a finite-dimensional Euclidean space with its Borel σ-field and a probability $g_{\zeta n}$ that is absolutely continuous with respect to Lebesgue measure. Let $\zeta_t = (\zeta_{1t},\dots,\zeta_{Nt}) \in \Lambda^N$. Also define $\zeta_{-n,t} = (\zeta_{1t},\dots,\zeta_{n-1,t},\zeta_{n+1,t},\dots,\zeta_{Nt})$.
- x_{nt} — a policy for firm n, in a compact subset X_n of a finite-dimensional Euclidean space, endowed with its relative Borel σ-field. Let $x_t = (x_{1t},\dots,x_{Nt})$ denote a *profile of policies* of the firms, and $X = X_1\times\dots\times X_N$. Define the profile of rivals' policies, $x_{-n,t} = (x_{1t},\dots x_{n-1,t},x_{n+1,t},\dots,x_{Nt}) \in X_{-n} = X_1\times\dots\times X_{n-1}\times X_{n+1}\times\dots\times X_N$.
- Π_n — a profit function for firm n, a function on $S\times X\times Z\times\Omega\times\Lambda$.
- h — an equation of motion, a measurable map from $S\times X\times Z\times\Omega$ into S.
- β — a common constant discount factor.

The state s_t may include both discrete and continuous payoff-relevant information such as entry status, market size, cumulative investment in product quality, and firm market shares. The observable exogenous variables z_t may include periods remaining until the horizon H, demand shifters, prices of firm inputs, and rules of conduct that are determined outside the model. In particular, we treat firms' beliefs about the applicability and consequences of the anti-trust law, and engagement in "bad acts" that violate this law, as determined exogenously by the legal system and a "firm culture" that promotes or discourages compliance. We assume that z_t is public knowledge to the firms when period t decisions are made. The exogenous shocks ε_t and η_t will represent demand effects. The distinction is that ε_t is assumed to be public knowledge to the firms, while η_t is assumed unknown to all firms, when policy decisions are made in period t, but η_t is assumed known to all firms before policy decisions are made in t+1. The exogenous variables z_t and shocks ε_t differ only in what the *econometrician* observes, and are treated as similar contemporaneously known public information by the firms. The exogenous shock ζ_{nt} will represent cost effects that are private knowledge to firm n when decisions are made in period t; we assume that ζ_{nt} enters the profit function for firm n, but not the profit functions of rivals or the equation of motion. We will assume that the exogenous vector z_t and the shocks ε_t, η_t, ζ_{1t},..., ζ_{Nt} follow independent first-order time-homogeneous Markov processes, and that the private cost shocks are time-independent. Then, the conditional probability of these variables, given history, has the form $g_z(z_t|z_{t-1})\cdot g_\varepsilon(\varepsilon_t|\varepsilon_{t-1})\cdot g_\eta(\eta_t|\eta_{t-1})\cdot g_{\zeta 1}(\zeta_{1t})\cdot\ldots\cdot g_{\zeta N}(\zeta_{Nt})$. The motivation for this assumption on the exogenous and stochastic elements is both theoretical and econometric. It implies that z_t, ε_t, and η_t can be treated as components of an extended state vector, along with s_t in the conventional MPE model structure. It limits firms' ability to predict these variables, and determines the structure of statistical dependence between observed variables and shocks. In particular, the private information of a firm is not predictable by other firms from history. In general, we interpret z_t as including social policy instruments that are time-invariant and are changed in the "but for" world as interventions to suppress "bad acts".

Policy variables may include investment, pricing, product attributes, entry and exit from product lines. In principle, the list of policy variables could include actions such as bundling, tying, refusals to deal, and exclusionary contracts that are potentially "bad acts" under anti-trust law, and observed firm behavior in these dimensions could be modeled explicitly as profit-maximizing given firm beliefs about the likelihood and potential costs of anti-trust litigation. In practice, the difficulty of identifying firm beliefs regarding litigation make this impractical. Instead, we interpret "bad acts" policies as predetermined to the market model, thus part of z_t, and analyze "but for" scenarios that result from interventions that change these predetermined policies.

Summarizing, the payoff-relevant information of firm n when it makes its policy decision in period t is a point $(s_t,z_t,\varepsilon_t,\zeta_{nt}) \in S\times Z\times\Omega\times\Lambda$. Firms are assumed to know the stochastic laws governing exogenous variables and shocks. A *Markov mixed strategy* for the firm is a measurable mapping f_n from this information into the space $P(X_n)$ of probabilities on X_n. This strategy is *almost surely pure* if, given (s_t,z_t,ε_t), the set of ζ_{nt} for which the support of the image of f_n is a singleton has probability one. When a Markov strategy is almost surely pure, f_n can be interpreted as a measurable mapping from $S\times Z\times\Omega\times\Lambda$ into X_n. Note that f_n is not the same as a *best-response* or *reaction* function, which maps the information above *and* the strategies of rivals into $P(X_n)$; the latter is the mapping for which the former is a fixed point in a MPE.

We will assume a non-atomic population of consumers who are price-takers and do *not* respond strategically. In particular, we assume that contagion or network effects influence demand solely through the previous period market state, and that consumers do not form expectations about future market states based on market history or firm conduct.[38]

The per-period profit function $\Pi_n(s_t,x_t,z_t,\varepsilon_t,\zeta_{nt})$ will reflect current production costs and revenues associated with its technology and the structure of consumer demand with its embedded network effects, and investment and marketing expenses. Note that Π_n is a function solely of information the firm knows and the realized policies of all the firms; it does not depend on unknown shocks such as η_t, Note that Π_n is time-invariant, with all time-varying effects captured in the exogenous vector z_t

The objective of the firm will be to maximize the expected present value of its stream of per-period profits, taking into account the information it has. The firm may have conjectures about its rivals' information and behavior, but these beliefs must in the end depend only on the policy-relevant information it has available. We will analyze the useful case of profit functions in which the private information has the same dimension as the policy of the firm and appears in a single index $x_{nt}\cdot\zeta_{nt}$. In this form, the ζ_{nt} can be interpreted as random policy costs associated with implementing the firm's action, such as a setup cost on entry or scrap cost on exit, or a random surcharge on investment expenditure reflecting administrative efficiency. Letting R_n denote the revenue function of the firm, and C_n its cost function, the profit function in this case can be written $\Pi_n(s_t,x_t,z_t,\varepsilon_t,\zeta_{nt}) \equiv R_n(s_t,x_t,z_t,\varepsilon_t) - C_n(s_t,x_{nt},z_t,x_{nt}\cdot\zeta_{nt})$, with revenue not depending on ζ_{nt} and cost not depending on ε_t or the policies of rivals, $x_{-n,t}$. We will not require that X_n be a convex set or that the profit function be concave in x_{nt}. However, we will assume that $C_n(s_t,x_{nt},z_t,x_{nt}\cdot\zeta_{nt})$ is concave and increasing in the index, and that ζ_{nt} has a continuous density. The concavity condition is satisfied, in particular, if C_n is additive in $x_{nt}\cdot\zeta_{nt}$, as will be the case if the ζ_{nt} are unit costs of policy actions. These assumptions on the structure of private information will guarantee that MPE solutions will be almost surely pure. The use of private information as a device to purify mixed strategies has been noted in the literature (Dor and Sattherwhite, 2004; Athey, 2004); our formulation gives a very simple proof.

The equation of motion $s_{t+1} = h(s_t,x_t,z_t,\varepsilon_t,\eta_t)$ maps the current state, exogenous variables, shocks, and policy realizations into the state in the following period. We assume that this equation does not depend on the private information of the individual firms, which influences the market solely through their realized policies. We absorb all time-varying effects into z_t, so that h is time-invariant. In the general theory of stochastic games, the law of motion is usually described in terms of Markov transition probabilities. In our notation, these would be obtained, for Borel sets $A \subseteq S$, from the relationship $P(A|s_t,x_t,z_t,\varepsilon_t) = g_\eta(\{\eta|h(s_t,x_t,z_t,\varepsilon_t,\eta) \in A\})$. For example, in the special case that the state spaces are finite, the law of motion is a (controlled) Markov chain, and

[38]Some markets are oligoposonistic, with atomistic buyers who can be included as active players in the game. There is also an interesting class of stochastic game extensions of the "Coase conjecture" problem, where consumers strategically time purchases based on expectations regarding future firm conduct. We will not analyze the possibility that firm announcements could become "bad acts" to promote market power when consumers respond strategically, but note that this is a problem of interest.

$P(s_{t+1}|s_t,x_t,z_t,\varepsilon_t)$ denotes the transition probabilities. With this basic model setup and the earlier discussion as motivation, we make the following assumptions:

Assumption 1. The observable vectors $z_t \in Z$ follow a first-order Markov process with probability $g_z(z_t|z_{t-1})$. The shocks $(\varepsilon_t,\eta_t,\zeta_{1t},...,\zeta_{Nt}) \in \Omega\times\Omega\times\Lambda^N$ follow an independent first-order Markov process with probability $g_\varepsilon(\varepsilon_t|\varepsilon_{t-1})\cdot g_\eta(\eta_t|\eta_{t-1})\cdot g_{\zeta 1}(\zeta_{1t})\cdot...\cdot g_{\zeta N}(\zeta_{Nt})$. The probabilities g_z, g_ε, g_η, and $g_{\zeta n}$ for n = 1,...,N are common knowledge. At each time t when policies x_{nt} are chosen, all firms know z_t and ε_t, and each firm n knows its private shock ζ_{nt}. Contemporaneous private shocks of rivals ζ_{-nt}, η_t, and all future exogenous vectors and shocks, are predictable by a firm only to the extent of knowing their Markov process. The econometrician observes z_t for t = 1,...,T, but does not observe the shocks ε_t, η_t, or ζ_t.

Assumption 2. The market structure is time-homogeneous. Specifically, the state space S, the exogenous vector space Z, the policy spaces $X_1,...,X_N$, the equation of motion h, and the profit functions Π_n for n = 1,...,N are independent of t, with any time dependence contained in the exogenous vector z_t.

Assumption 3. The profit function $\Pi_n(s_t,x_t,z_t,\varepsilon_t,x_{nt}\cdot\zeta_{nt})$ and the equation of motion $h(s_t,x_t,z_t,\varepsilon_t,\eta_t)$ are continuously differentiable in their arguments, with Lipschitz derivatives.[39] The discount factor satisfies $0 < \beta < 1$.

Assumption 4. For firms n = 1,...,N, the profit function Π_n depends on private information solely through the single index $x_{nt}\cdot\zeta_{nt}$ and is a convex decreasing function of this index. The private information ζ_{nt} has a probability $g_{\zeta n}$ that is absolutely continuous with respect to Lebesgue measure.

A *Markov Perfect Equilibrium* (MPE) is a profile of Markov mixed strategies f_n^* for all firms and periods with the subgame-perfect Nash property that for each period t and firm n, given the strategies of rivals, firm n has no incentive to alter its assigned strategy f_n^*. Note that $f_n^*(x|s_t,z_t,\varepsilon_t,\zeta_{nt})$ can be interpreted as a conditional cumulative distribution function, a measurable function from $X\times S\times Z\times\Omega\times\Lambda_n$ into the unit interval. Let $V_n(s_t,z_t,\varepsilon_t,\zeta_{nt})$ denote the *valuation function* of firm n, the maximum expected present value of its stream of profits from period t forward, given its payoff-relevant information. Note that we have made V_n time-invariant by incorporating all time dependence into z_t. The MPE and the valuation functions of the firms are characterized by Bellman's backward recursion. Let $\mathbf{E}_{-n,t}$ denote the expectation operator with respect to variables that are *unknown* to firm n at the time of its policy decision in period t: $z_{t'}$ and the shocks $\varepsilon_{t'},\eta_{t'},\zeta_{t'}$ for

[39] If a state or exogenous vector has some discrete components, then the Lipschitz condition is satisfied trivially in these components.

$t' > t$, the contemporaneous private information of rivals, $\zeta_{-n,t} = (\zeta_{1t},\ldots,\zeta_{n-1,t},\zeta_{n+1,t},\ldots,\zeta_{Nt})$, and the contemporaneous shock η_t. The valuation functions then satisfy Bellman's equations,

$$(A.1) \qquad V_n(s_t,z_t,\varepsilon_t,\zeta_{nt}) = \max_{x_{nt}}\{\lambda_n(s_t,x_{nt},z_t,\varepsilon_t,\zeta_{nt}) + \beta\mu_n(s_t,x_{nt},z_t,\varepsilon_t)\},$$

with the notation

$$(A.2) \qquad \lambda_n(s_t,x_{nt},z_t,\varepsilon_t,\zeta_{nt}) \equiv \mathbf{E}_{-n,t}\int_{X_{-n,t}} \prod_{j\neq n} f_j^*(dx_{jt}|s_t,z_t,\varepsilon_t,\zeta_{jt})\Pi_n(s_t,x_{-n,t},x_{nt},z_t,\varepsilon_t,\zeta_{nt})$$

for the expected value of current profit, and

$$(A.3) \qquad \mu_n(s_t,x_{nt},z_t,\varepsilon_t)$$
$$\equiv \mathbf{E}_{-n,t}\int_{X_{-n,t}} \prod_{j\neq n} f_j^*(dx_{jt}|s_t,z_t,\varepsilon_t,\zeta_{jt})V_n(h(s_t,x_{-n,t},x_{nt},z_t,\varepsilon_t,\eta_t),z_{t+1},\varepsilon_{t+1},\zeta_{n,t+1})$$

for the expected value of future profit. Let $X_{nt}^*(s_t,z_t,\varepsilon_t,\zeta_{nt})$ denote the *maximizer correspondence*, the set of points x_{nt} that achieve the maximum in (A.1). The Nash property of MPE requires that the support of $f_n^*(\cdot|s_t,z_t,\varepsilon_t,\zeta_{nt})$ be contained in $X_{nt}^*(s_t,z_t,\varepsilon_t,\zeta_{nt})$.

A1.2. Existence and Purity of MPE. Existence of subgame perfect Nash or MPE solutions to dynamic stochastic games is the subject of an extensive and deep literature; see for example Beggs-Klemperer (1992), Curtat (1996), Doraszelski-Satterthwaite (2003), Duffie-Geanakoplos-Mas Colell-McLennan (1994), Fudenberg-Tirole (2000), Haller-Lagunoff (2000), Majumdar and Sundaram (1991), Maskin-Tirole (1988a,b), Mertens (2003), Mertens-Parthasarathy (2003), Nowak (2003), Reider (1979), Rosenberg (1998), Solon (1998), Sorin (2003). In overview, existence in infinite-horizon discounted stochastic games is well-established for finite state spaces, and these results are easily extended to countable state spaces; see Federgruen (1976) and Whitt (1980). In state spaces that are not countable, existence has been proved only under very restrictive structural assumptions; see Amir (2001), Curtat (1996). There are however, proofs of the existence of ε-equilibrium under quite general conditions; see Duffie *et al* (1994), Mertens-Parthasarathy (2003), Solon (1998).

We will provide a relatively simple, self-contained existence result that assumes the state space is countable except in the special dimension of private firm information. This result is a minor variation on theorems of Federgruen and Whitt, and is sufficient to support our computations of MPE, which are of course restricted to the rational numbers. We include a proof for completeness. We will also show, with an added assumption, that an MPE in pure strategies exists almost surely. We do not deal with the coordination problem in the case of multiple MPE, or consider a supergame in which coordinating communication between players is possible.

To connect our fundamental assumption that $S_t \times Z_t \times \Omega$ is a compact metric space to an existence theory that requires countability, we will assume a nested sequence of finite partitions $\{A_{1Kt},\ldots,A_{KKt}\}$ of $S_t \times Z_t \times \Omega$ for $K \to \infty$, and an associated set B_K of *partition points* $(s_{tkK},z_{tkK},\varepsilon_{tkK}) \in A_{kKt}$ with the property that each of these points remains in all succeeding refinements, and the union B* of these points is dense in $S_t \times Z_t \times \Omega$. For each K, we will assume that the firm's knowledge of the state (s_t,z_t,ε_t) is limited to the partition set in which appears, and thus indistinguishable from the partition point $(s_{tkK},z_{tkK},\varepsilon_{tkK})$ in this partition set. As a consequence, given partition K, the strategy of firm n, $f_{nt}(x|s_t,z_t,\varepsilon_t,\zeta_{nt})$ is piecewise constant for (s_t,z_t,ε_t) in a partition set A_{kKt}, and can be written as $f_{nt}(x|s_{tkK},z_{tkK},\varepsilon_{tkK},\zeta_{nt})$ for the associated partition point. One can interpret this setup as one in which the firm has incomplete information on the state of the system, and chooses a strategy that is Nash given this information. Because the information set is finite, the firm has no scope for selecting a strategy that chatters excessively with changes in the state, precluding the possibility of non-measurable limits. Refining the partition increases the information, so that at least from an applied perspective, the information loss from partitioning becomes negligible. Note that equilibria established for a given partition K are exact for B_K, but not for the fundamental underlying state space. Consequently they can be interpreted as ε-equilibria, different in construction but similar in spirit to the ε-equilibria studied by Whitt (1980), Duffie *et al* (1994), and Mertens-Parthasarathy (2003).

Theorem 1. Suppose Assumptions 1-3. Then there exist valuation functions V_{nt} and a profile of strategies f_{nt}^* for n = 1,...,N that are time-invariant for $t \geq T^*$ and satisfy condition (A1.1) for a MPE on the countable dense subset of $S_t \times Z_t \times \Omega$ given by the union of the partition points $(s_{tkK},z_{tkK},\varepsilon_{tkK})$ for k = 1,...,K and $K \to \infty$.

Proof: Fix K, and consider the state space $B_K \times \Lambda$. We first establish the existence of a time-invariant valuation function and strategy profile for the time-invariant subgame starting in period T*. Let T denote any time greater than T*. The functions $\Pi_n(s_T,x_T,z_T,\varepsilon_T,\zeta_{nT})$ and $h(s_t,x_t,z_t,\varepsilon_t,\eta_t)$ are Lipschitz; let M denote a uniform bound on these function and their Lipschitz constants, which exists since all domains are compact. We will use a fixed point argument that there exist valuation functions that are reproduced by, and strategies for the firms that are consistent with, the Bellman recursion (A.1). We begin by defining a series of sets and mappings, and obtaining their properties.

[1] Let W_n denote the set of real-valued functions V on $B_K \times \Lambda_n$ that are Lipschitz in Λ_n, with a bound M/(1-β) on the function and on the Lipschitz constant. Then, W_n is a convex compact subset of the Banach space of bounded measurable functions on $B_K \times \Lambda_n$; see Dunford and Schwartz (1965, Theorem IV.5.6).

[2] Let F_n denote the set of cumulative distribution functions f_n on X_{nT} that are conditioned on and measurable with respect to $B_K \times \Lambda_n$. This is a subset of the Banach space $L_1(X_{nT} \times B_K \times \Lambda_n)$. Further, for each rectangle $A = A_x \times A_K \times A_\zeta$ in this space,

$$\int_A f_n(x|s,z,\varepsilon,\zeta) g_\zeta(d\zeta)\mu_K(A_K)\mu_x(dx) \le g_\zeta(A_\zeta)\mu_K(A_K)\mu_x(A_x),$$

where μ_K is counting measure on B_K and μ_x is relative Lebesgue measures on X_{nT}. Then, F_n is weakly sequentially compact and convex; see Dunford and Schwartz (1965, Theorem IV.8.9).

[3] For each $(V_1,\ldots,V_N,f_1,\ldots,f_n) \in W_1 \times \ldots \times W_N \times F_1 \times \ldots \times F_N$, the function $\rho_{nT}(s_T,x_{nT},z_T,\varepsilon_T,\zeta_{nT})$ from (A.2) is an expectation (with respect to $\zeta_{-n,T}$, whose distribution is independent of x_{nT},ζ_{nT}) of the function Π_n which is uniformly Lipschitz in (x_{nT},ζ_{nT}), and is therefore again Lipschitz in these arguments with the same constant.

[4] For each $(V_1,\ldots,V_N,f_1,\ldots,f_n) \in W_1 \times \ldots \times W_N \times F_1 \times \ldots \times F_N$, the function $\mu_{nT}(s_T,x_{nT},z_T,\varepsilon_T)$ from (A.3) is an expectation (with respect to $\zeta_{-n,T}$, $\zeta_{n,T+1}$, and ε_{T+1} whose distributions do not depend on x_{nT}, and with respect to s_{T+1} whose distribution has a Lipschitz dependence on x_{nT} and is independent of $\zeta_{n,T+1}$) of the function $V_{n,T+1}$ which is independent of x_{nT}. Consequently, μ_{nT} inherits the bound $M/(1-\beta)$ on $V_{n,T+1}$, and is Lipschitz in x_{nT}.

[5] The function $\rho_{nT}(s_T,x_{nT},z_T,\varepsilon_T,\zeta_{nT}) + \beta\,\mu_{nT}(s_T,x_{nT},z_T,\varepsilon_T)$ in (A.1) is uniformly bounded by $M + \beta M/(1-\beta) = M/(1-\beta)$, is Lipschitz in ζ_{nT} with constant M, and Lipschitz in x_{nT} on its compact domain. One implication is that the maximizer correspondence $X_{nT}^*(s_T,z_T,\varepsilon_T,\zeta_{nT})$ is non-empty and upper hemicontinuous on $B_K \times \Lambda_n$; see Hildenbrand (1974). A second implication is that $V_n'(s_T,z_T,\varepsilon_T,\zeta_{nT}) = \max_{x_{nt}}\{\rho_{nT}(s_T,x_{nT},z_T,\varepsilon_T,\zeta_{nT}) + \beta\mu_{nT}(s_T,x_{nT},z_T,\varepsilon_T)\}$ is bounded with constant $M/(1-\beta)$. Also, let $x_{nT}(\zeta_{nT}) \in X_{nT}^*(s_T,z_T,\varepsilon_T,\zeta_{nT})$ be any selection, and note that maximization implies

$$\begin{aligned}
&-M|\zeta_n T^* - \zeta_n T^*| \\
&\quad \le \rho_{nT}(s_T,x_{nT}(\zeta_{nT}''),z_T,\varepsilon_T,\zeta_{nT}'') - \rho_{nT}(s_T,x_{nT}(\zeta_n T^*),z_T,\varepsilon_T,\zeta_{nT}'') \\
&\qquad \le V_n'(s_T,z_T,\varepsilon_T,\zeta_n T^*) - V_n'(s_T,z_T,\varepsilon_T,\zeta_{nT}'') \\
&\qquad\quad \le \rho_{nT}(s_T,x_{nT}(\zeta_n T^*),z_T,\varepsilon_T,\zeta_n T^*) - \rho_{nT}(s_T,x_{nT}(\zeta_n T^*),z_T,\varepsilon_T,\zeta_{nT}'') \le M|\zeta_n T^* - \zeta_n T^*|.
\end{aligned}$$

Therefore, V_n' is Lipschitz in ζ_{nT} with constant M. Therefore, $V_n' \in W_n$.

[6] Let ψ_n denote the mapping from $B_K \times \Lambda_n$ into subsets of F_n defined by $\psi_n(s_T,z_T,\varepsilon_T,\zeta_{nT}) = P(X_{nT}^*(s_T,z_T,\varepsilon_T,\zeta_{nT}))$; i.e., the set of all cumulative distribution functions with support contained in the maximizer correspondence. The correspondence ψ_n is non-empty and convex-valued. The upper hemicontinuity of the maximizer correspondence implies that ψ_n is upper hemicontinuous.

The preceding results establish that $W_1 \times \ldots \times W_N \times F_1 \times \ldots \times F_N$ is convex and compact, and that the mapping given by [5] and [6] is a upper hemicontinuous convex-valued correspondence from this space into itself. Then, by the Glicksburg-Fan fixed point theorem (Saveliev, 1999), there exists $(V_{1T},\ldots,V_{NT},f_{1T}^*,\ldots,f_{NT}^*) \in W_1 \times \ldots \times W_N \times F_1 \times \ldots \times F_N$ that satisfies (A.1) and returns the same valuation functions. This completes the proof that a stationary MPE exists in the subgame starting in period T*.

Now proceed by backward recursion of (A.1) from period T*. Let T be a period satisfying $T \le T^*$, and suppose that strategies f_{nt}^* and valuation functions V_{nt} have been established satisfying (A.1) for t > T. Define F_n and ψ_n as in [2] and [6] above, and note that again by the same construction F_n is convex and compact, and ψ_n is a upper hemicontinuous convex-valued correspondence from $F_1 \times ... \times F_N$ into itself. Therefore, the Glicksburg-Kakutani fixed point theorem establishes the existence of strategies $(f_{1T}^*,...,f_{NT}^*) \in F_1 \times ... \times F_N$ satisfying (A.1). This completes the proof that a MPE exists for the full game with finite partition K.

Now consider the sequence of MPE obtained by the proof above as $K \to \infty$. For each point in the union of the partition points, the sequence of MPE has a convergent subsequence at this point. This follows from pointwise convergence of the sequence of uniformly bounded valuation functions, the compactness of the strategies (which are cumulative distribution functions) at this point, and the inequalities implied by the maximization in the Bellman recursion. Use the Cantor diagonal process to obtain a limiting MPE that satisfies (A.1) on the countable union of partition points B*. □

It is mathematically difficult to extend Theorem 1 to the fundamental compact metric state space without countability. To do so requires that the valuation functions and firm strategies not chatter with changes in state. We conjecture that empirical process methods, particularly the metric entropy conditions that give uniform stochastic equicontinuity, can be applied to admissible mixed strategies to obtain results with relatively weak conditions; see Pollard (1984), Shorak and Wellner (1986), McFadden (1989). We believe that the required bounds on entropy can be met through monotonicity or supermodularity conditions that are not as restrictive as those used by Curtat (1996) and Amir (1996). Resolution of this conjecture is left for the future.

For some purposes, it may be useful to identify invariant distributions for the state of the model after time T* when the model is time-invariant, and estimate these distributions from the empirical distribution of realized states. With the MPE strategies given by Theorem 1, the law of motion of the model is a time-invariant Markov process, and there are some straightforward (high-level) assumptions under which it will have a unique invariant distribution and the ergodic property that the empirical distribution from any initial state will converge weakly to the invariant distribution. Define a Markov transition kernel

$$P(A|s_t,x_t,z_t,\varepsilon_t) = g_\eta(\{\eta|h(s_t,x_t,z_t,\varepsilon_t,\eta) \in A\}).$$

For this discussion, assume that the state spaces are finite. Suppose for any pair of states s'',s', there is a set of z,ε,ζ occurring with positive probability on which $P(s''|s',z,\varepsilon,\zeta) < 1$, and there exists a finite sequence of states s^k with $s^0 = s'$ and $s^K = s''$ with the property that there are sets of z,ε,ζ occurring with positive probability on which $P(s^{k+1}|s^k,z,\varepsilon,\zeta) > 0$. Then, the process is irreducible and acyclic, has an invariant distribution which is unique, and has the strong ergodic property that for any continuous function on $S_T \times Z_T \times \Omega \times \Lambda$, a time average of this function on any realized trajectory converges to its expectation at an exponential rate. Related results that are much more general are given by Duffie, Geanakoplos, Mas-Colell, McLennan (1994).

In computational approximation of MPE, mixed strategies create difficulties, so that the analysis is greatly aided if pure strategy MPE exist. This has been accomplished in previous literature through strong assumptions on the profit functions, or through the introduction of private information that acts as a mixing device. The first approach requires that the firm's policy set X_{nt} be convex, and the profit functions be strictly concave on this set, or that it satisfy some strong separability and super-modularity conditions; see Curtat (1996). It is difficult to use the first approach in our application in which network effects influence the determination of market shares. The second approach has been used in specific applications by Pakes and McGuire (1994) and Doraszelski, and Satterthwaite (2003). We adapt the second approach to our model, introducing a functional specification that gives the desired result.

Theorem 2. Suppose Assumptions 1-4. Then, with probability one, the MPE established in Theorem 1 will assign only pure strategies for all firms in all time periods.

Proof: Consider the function $V_{nt}(s_t,z_t,\varepsilon_t,\zeta_{nt}) = \max_{x_{nt}}\{\lambda_n(s_t,x_{nt},z_t,\varepsilon_t,\zeta_{nt}) + \beta\mu_n(s_t,x_{nt},z_t,\varepsilon_t)\}$ from (A.1). We will argue that V_{nt} is a convex function of ζ_{nt}, and that its derivative with respect to ζ_{nt}, if it exists, is proportional to a unique pure strategy for the firm. But a convex function is twice continuously differentiable almost everywhere; see Alexandroff (1939). This result, combined with the Assumption 4 condition that the probability g_ζ is absolutely continuous with respect to Lebesgue measure, ensures that the firm's strategy is almost surely pure and continuous in ζ_{nt}. We complete the proof by demonstrating that V_{nt} is convex, and that when its derivative exists, it is almost everywhere a non-zero multiple of a unique optimal policy. First note that $\rho_n(s_t,x_{nt},z_t,\varepsilon_t,\zeta_{nt})$ is an expectation (independent of ζ_{nt}) of of the negative of a cost function that is concave and monotone in ζ_{nt}, and hence ρ_n is convex and monotone in ζ_{nt}. Suppose that $\zeta_n T^*$ and ζ_{nt}'' are distinct points, consider a convex combination $\theta\zeta_n T^*+(1-\theta)\zeta_{nt}''$, and let x_{nt}^* be any maximizer that gives the function $V_{nt}(s_t,z_t,\varepsilon_t,\theta\zeta_n T^*+(1-\theta)\zeta_{nt}'')$. Then,

$$\begin{aligned}V_{nt}(s_t,z_t,\varepsilon_t,\theta\zeta_n T^*+(1-\theta)\zeta_{nt}'') &= \rho_n(s_t,x_{nt}^*,z_t,\varepsilon_t,\theta\zeta_n T^*+(1-\theta)\zeta_{nt}'') + \beta\mu_n(s_t,x_{nt}^*,z_t,\varepsilon_t)\\ &\le \theta\{\rho_n(s_t,x_{nt}^*,z_t,\varepsilon_t,\zeta_n T^*\cdot x_{nt}^*) + \beta\mu_{nt}(s_t,x_{nt}^*,z_t,\varepsilon_t)\}\\ &\quad + (1-\theta)\{\rho_n(s_t,x_{nt}^*,z_t,\varepsilon_t,\zeta_{nt}''\cdot x_{nt}^*) + \beta\mu_{nt}(s_t,x_{nt}^*,z_t,\varepsilon_t)\}\\ &\le \theta V_{nt}(s_t,z_t,\varepsilon_t,\zeta_n T^*) + (1-\theta)V_{nt}(s_t,z_t,\varepsilon_t,\zeta_{nt}'').\end{aligned}$$

This completes the proof of concavity. Finally, let x_{nt}'' denote a maximizer that gives the function $V_{nt}(s_t,z_t,\varepsilon_t,\zeta_{nt}'')$, and note that by its definition,

$$\begin{aligned}V_{nt}(s_t,z_t,\varepsilon_t,\theta\zeta_n T^*+(1-\theta)\zeta_{nt}'') &\ge \rho_n(s_t,x_{nt}'',z_t,\varepsilon_t,[\theta\zeta_n T^*+(1-\theta)\zeta_{nt}'']\cdot x_{nt}'') + \beta\mu_{nt}(s_t,x_{nt}'',z_t,\varepsilon_t)\\ &\ge V_{nt}(s_t,z_t,\varepsilon_t,\zeta_{nt}'') - \max\{\lambda_+\theta(\zeta_{nt}'' - \zeta_n T^*)\cdot x_{nt}'',\lambda_-\theta(\zeta_{nt}'' - \zeta_n T^*)\cdot x_{nt}''\},\end{aligned}$$

for any scalar θ, positive or negative, that leaves the argument in the domain Λ_n, and λ_+ and λ_- are the left and right hand side derivatives of ρ_n with respect to the index, which always exist and by strict monotonicity and convexity are almost surely signed and non-zero. Letting θ go to zero from above and below, the left and right derivatives of V_{nt} bracket the expressions $\lambda_+(\zeta_{nt}'' - \zeta_n T^*)\cdot x_{nt}''$ and $\lambda_-(\zeta_{nt}'' - \zeta_n T^*)\cdot x_{nt}''$. Then, if ζ_{nt}'' is a point where V_{nt} is differentiable, these right and left hand derivatives coincide, implying that $\lambda_+ = \lambda_-$ so that ρ_n is also differentiable at this point, and the differential of V_{nt} is $\lambda_+(\zeta_{nt}'' - \zeta_n T^*)\cdot x_{nt}''$. Since this holds for each $\zeta_n T^*$ in a small ball around ζ_{nt}'', it follows that the derivative of V_{nt} is x_{nt}''. Since x_{nt}'' is any maximizer at ζ_{nt}'', it is unique. This completes the proof of the theorem. □

Example: Let $R_n(s_t,x_t,z_t,\varepsilon_t)$ denote the revenue function of the firm, and $C_n(s_t,x_{nt},z_t,x_{nt}\cdot\zeta_{nt}) \equiv c_n(s_t,x_{nt},z_t) + x_{nt}\cdot\zeta_{nt}$ its cost function, with revenue not depending on ζ_{nt} and cost not depending on ε_t or the policies of rivals, $x_{-n,t}$. In this setup, ζ_{nt} is a vector of unit costs, commensurate with x_{nt}, that gives the firm's private cost of policy x_{nt}. The profit function is then $\Pi_n(s_t,x_t,z_t,\varepsilon_t,\zeta_{nt}) \equiv R_n(s_t,x_t,z_t,\varepsilon_t) - c_n(s_t,x_{nt},z_t) - x_{nt}\cdot\zeta_{nt}$, and is trivially concave and strictly monotone in $x_{nt}\cdot\zeta_{nt}$. Components of ζ_{nt} may include entry setup and exit shutdown costs that do not otherwise enter the technology or the determinants of firm revenue.

The preceding results do not establish a unique MPE, nor address the coordination problem associated with multiple MPE. Duffie, Geanakoplos, Mas-Colell, McLennan (1994) discuss extending the market game to include public information on "sunspots" that can be used as a coordination device. Uniqueness results based on game structure have been obtained only under stringent sufficient conditions; see Milgrom and Roberts (1990). The model we introduce for our application does not meet these conditions, and we are left with the possibility that computational solutions in applications may find only some of multiple MPE solutions.

Appendix 2. Econometric Identification and Estimation

In our discussion of econometric estimation of the structure of an MPE model, we will use the language of parametric estimation, treating the equation of motion and the profit functions of the firms as fully specified up to unknown parameter vectors. This is the leading case for applications, where limited observations will support only sparsely parameterized models. However, it is also useful to think of these parametric models as method of sieves approximations to the true structural equations which are nonparametric, and to consider questions of identification and estimation in this more general context. We will impose Assumptions 1-4, so that firm strategies are almost surely pure, and a firm's private information enters only its cost function as a concave function of a single index.

As described in Section 3.6, we divide the econometric problem into three stages: (1) structural estimation of the equation of motion and profit functions prior to the MPE calculations, (2) computation of the MPE with minimum distance estimation of unidentified parameters from the first stage, and (3) one or more iterations of the complete vector of parameters from the MPE trajectories using a generalized minimum distance criterion

Consider in more detail the first-stage identification and estimation of the equation of motion,

(A2.1) $s_{t+1} = h(s_t,x_t,z_t,\varepsilon_t,\eta_t)$,

and the profit functions

(A2.2) $\pi_{nt} = R_n(s_t,x_t,z_t,\varepsilon_t) - C_n(s_t,x_{nt},z_t,x_{nt}\cdot\zeta_{nt})$.

By Theorem 2, the MPE strategies of the firms are almost surely pure, and hence can be written as functions

(A2.3) $x_{nt} = r_{nt}(s_t,z_t,\varepsilon_t,\zeta_{nt})$

for n = 1,...,N. In the first stage of estimation, the structure of r_{nt} in (A2.3) is not yet determined. Note however that the r_{nt} are time-invariant for $t \geq T^*$.

We consider identification and estimation of these equations in several cases. The first and most straightforward occurs when there are no shocks in the equation of motion that are known to the firms when they choose their policies; i.e., ε_t is absent. Then, the observed policies x_{nt} in (A1.1) are independent of the remaining shock η_t. If this shock is additive, then (A1.1) can be estimated by nonlinear least squares. This is also the case if current policy choices enter the equation of motion with a lag, as is the case in our application. If the shock is not additive, then Matzkin (2003) shows that either a dimension-reducing functional restriction, normalization of r_{nt} (at some point s_{T^*},z_{T^*}), or normalization of the distribution of η_t, is needed for identification. This paper also gives applicable estimators when the equation is identified.[40] Estimation of the profit functions is not simple because even though ε_t is absent, the policy x_{nt} will depend on the shock ζ_{nt}. Identification can be achieved if there are variables in z_t such as observed input prices of rivals that are excluded from Π_{nt} but influence x_{nt} via their impact on the equation of motion and consequently on the valuation function of firm n. Estimation of the equation of motion may provide estimates of the structure of demand, determining the revenue portion of the profit functions and leaving only the costs which are functions solely of the firm's own policy choice. Using this, one can estimate the cost function, which depends on the endogenous policy choice, using the nonparametric instrumental variable methods in Matzkin (2004). In particular, the policies of rivals, which depend on their own private information, may be used to construct "unobserved instruments".

---

[40] An example of a functional restriction is an "index" model in which s_t,x_t,z_t enter h through a one-dimensional linear transformation with parametric weights.

Consider next the case where the publically known shock ε_t is present and correlated with the policy functions of the firms. We will concentrate on estimation of the equation of motion h in circumstances where this function is not necessarily additive in the shocks ε_t and η_t. Suppose there are exogenous variables in the profit functions that influence policy choices but are excluded from the equation of motion. Typical examples are input prices. If the shocks in the equation of motion are additive, the estimation problem is of standard GMM form, with orthogonality conditions between transformations of excluded exogenous variables and excluded instruments providing identification. If the shocks are not additive, then the methods of Matzkin (2003) can again be used. Assume that h is strictly increasing in both ε_t and η_t. Note that in the model $s_{t+1} = h(s_t,x_t,z_t,\varepsilon_t,\eta_t)$, one could make a transformation of variables $\varepsilon_t = \psi(\varepsilon T^*)$, where εT^* is uniformly distributed, and then absorb ψ into the definition of h. Since this is observationally equivalent, it is clear that a normalization is needed, either on the distribution of ε_t or on the structure of the function h. The following result normalizes the distributions of these unobserved variables.

Theorem 3. Suppose under Assumptions 1-4 that the private cost vector of the first firm, ζ_{1t}, is observed retrospectively by the econometrician. Suppose the equation of motion and profit functions are non-parametric, time-invariant, and in general non-additive in shocks. Suppose the distributions of ε_t, η_t, ζ_{nt} for n = 2,...,N are normalized. Then the equation of motion (10), the reduced-form policy functions (12), and the realized values of ε_t, η_t, and ζ_{nt} for n > 1 are identified.

Proof: Since (s_t,z_t,ζ_{1t}) is observable, and ε_t is independent of (s_t,z_t,ζ_{1t}), it follows from Matzkin (1999, 2003) that

(A2.4) $$r_{1t}(s_t,z_t,\varepsilon_t,\zeta_{1t}) = g_{x1|s,z,\zeta 1}{}^{-1}(g_\varepsilon(\varepsilon_t)),$$

where $g_{x1|s,z,\zeta 1}$ is the conditional distribution of x_{1t} given s_t,z_t,ζ_{1t}. Normalizing the marginal distribution of ε_t, this shows that r_{1t} is identified. Moreover, for any t,

(A2.5) $$\varepsilon_t = g_\varepsilon{}^{-1}(g_{x1|s,z,\zeta 1}(x_{1t})).$$

This expresses the value of ε_t as the value of a known function of $(x_{1t},s_t,z_t,\zeta_{1t})$. Hence, in (A2.3) for n > 1, we can now treat ε_t as observable. Applying again Matzkin (1999, 2003), this time to (A2.3) for n > 1, we get

(A2.6) $$r_{nt}(s_t,z_t,\varepsilon_t,\zeta_{nt}) = g_{xn|s,z,\varepsilon}{}^{-1}(g_{\zeta 2}(\zeta_{nt})),$$

where $g_{xn|s,z,\varepsilon}$ is the conditional distribution of x_{nt} given s_t,z_t,ε_t, which shows that r_{nt} is identified up to a normalization on the distribution of ζ_{2t}. Moreover, for any t, (A2.7) $\zeta_{nt} = g_{\zeta 2}^{-1}(\, g_{xn|s,z,\varepsilon}(x_{nt}))$

This expresses the value of the private information of firm n as a known function of $x_{nt},s_t,z_t,\varepsilon_t$. Since, from above, ε_t is a known function of $(x_{1t},s_t,z_t,\zeta_{1t})$, this can be interpreted as a known function of $x_{1t},x_{nt},s_t,z_t,\zeta_{1t}$.

Since, from (A2.5), ε_t is a known number, and, by assumption, η_t is independent of $s_t,z_t,x_t,\varepsilon_t$, it follows, again from Matzkin (1999,2003), that

(A2.8) $h(s_t,z_t,x_t,\varepsilon_t,\eta_t) = g_{s+|s,z,x,\varepsilon}^{-1}(g_\eta(\eta_t))$

where $g_{s+|s,z,x,\varepsilon}$ is the conditional distribution of s_{t+1} given $s_t,z_t,x_t,\varepsilon_t,\eta_t$, which shows that h is identified up to a normalization on g_η. The proof is completed by noting that

(A2.9) $h(s_t,z_t,x_t,\varepsilon_t,\eta_t) = g_\eta^{-1}(g_{s+|s,z,x,\varepsilon}(s_{t+1}))$. □

Nonparametric estimators for the identified functions can be obtained by substituting nonparametric estimators for the conditional distributions in the expressions above.

Consider second-stage estimation. Given parameter estimates (or nonparametric estimates of the equation of motion and profit functions), and given trial values for any parameters or functions not identified from the first-stage estimation, one can follow the computational algorithm outlined earlier to determine the stationary state fixed point valuation functions, then the fixed point valuation functions in earlier periods by backward recursion, then the realized strategies and trajectory by rolling the model forward from a given starting state. A calculated trajectory for states and policies can be interpreted as a simulation draw, given shocks drawn from their estimated empirical distribution. Repeated simulations gives an estimate of the expected trajectories. A generalized distance of these simulated trajectories from observed trajectories can then be computed. One can then iterate this process at alternative trial values for parameters to minimize this distance. Because of the computational burden, and lack of a guarantee of smoothness or convexity that would ensure easy convergence of the iteration, it is extremely helpful that most parameters are estimated consistently from stage 1, and a relatively small number have to be estimated initially in stage 2. Note that in general any parameter that influences the trajectories or observed outcomes of the firms is identified from the observed trajectories, within the limits of empirical identification in finite data sets.

In practice, we carry out the second stage MPE dynamic computations using the first stage parameter estimates plus gradient search for the remaining parameters. The minimum distance criterion in our application is defined in terms of differences of observed and calculated states over the observed partial trajectory. Alternately, we could have used the distance of the calculated MPE first-order conditions from zero when evaluated at the observed policy variables. One method for carrying out the stage 2 minimum distance estimation is to embed the Chebyshev approximation procedure for the valuation functions in an iterative algorithm to estimate the previously unidentified parameters ψ to minimize a criterion defined as a sum of squared deviations of the predicted “as is” policies and states from their observed values for t = 1,...,T″. The search is facilitated by the smooth behavior of the valuation functions and policy functions with small changes in ψ, provided the solutions to Bellman’s equations are regular. Schematically, we have

$$V_{jt}(s,z,\psi) \approx A(s,z)B_{jt}(s,z,\psi), \tag{A2.10}$$

where A(s,z) is the augmented polynomial approximation (26). Suppose one starts from a solution to the MPE problem at an initial ψ_1. Differentiate the first-order conditions for the Bellman equations and the equations of motion to determine the derivatives of the B’s and x’s with respect to ψ, and plug these derivatives into the expression for the derivative of the estimation criterion with respect to ψ. For convenience, the derivatives can be done numerically. Then, do a gradient search to minimize the criterion. In principle, it is not necessary to re-solve the original problem, but in practice, periodic resolving should be done to correct the cumulative drift in the approximation from the MPE solution. One major difference between our formulation and that of Bajari *et al (2003)* is that they concentrate on models where the current payoff functions are linear in unknown parameters. That is not generally the case for the models in our application, adding to the computational burden of finding the Nash fixed points. Also, we do not require symmetry for the Nash solution.

The final stage of estimation is one or more Gauss-Newton iterations in all parameters to minimize the generalized distance between observed and simulated trajectories. Asymptotic statistical theory implies that this single step, which requires estimation of second derivatives of the generalized distance with respect to the parameters, yields estimators that are efficient within the class of minimum distance estimators, see Newey-McFadden (1994). A side benefit of this procedure is that it gives ready estimates of asymptotic standard errors of the parameter estimates. (If the number of simulation repetitions does not grow with sample size, then there is some loss of asymptotic efficiency.) In practice, linear search within the final one-step iteration may be needed to guarantee an improvement in minimum distance in a finite sample.

Some studies using MPE models have found that it is computationally advantageous to use polynomial approximations for the policy functions as well as the valuation functions. We have not done so in this paper, but note that for problems where convergence of the numerical search to a fixed point is difficult, such approximations may facilitate solution. Bajari *et al* (2003) estimate policy functions as well as the equation of motion in the first stage, assuming that these functions are linear in parameters. Because these policy functions depend structurally on the valuation functions derived in the second stage, all one can hope to get at the first stage are (nonparametric) reduced forms. These may nonetheless be useful, as their estimated values may be proper instruments for the policy variables that appear endogenously in the equation of motion, and they may provide good starting values for second-stage computations.

Appendix 3. Data

VARIABLE	DEFINITION
date	First day of Quarter (mm/dd/yyyy)
prd	Period (quarters 1 to 26)
sbasens	Netscape share of all browsers in use, beginning of quarter
sbasems	Microsoft share of all browsers in use, beginning of quarter
sbns	Netscape share of Netscape and Microsoft browsers in use, beginning of quarter
usetot	Total number of computers with browsers, beginning of quarter (millions)
usenew	Total number of new computers with browsers in quarter (millions)
basens	Netscape total browsers installed, beginning of quarter (millions)
basems	Microsoft total browsers installed, beginning of quarter (millions)
basedif	Netscape - Microsoft browsers installed, beginning of quarter (millions)
salens	Netscape new browsers installed in quarter (millions)
salems	Microsoft new browsers installed in quarter (millions)
saledif	Netscape - Microsoft new browsers installed in quarter (millions)
ssalens	Netscape share of Netscape + Microsoft new browsers installed in quarter (millions)
ssalems	Microsoft share of Netscape + Microsoft new browsers installed in quarter (millions)
invns	Netscape browser R&D investment in quarter ($millions)
invms	Microsoft browser R&D investment in quarter ($millions)
invdif	Netscape - Microsoft R&D investment difference in quarter ($millions)
cinvns	Netscape cumulative browser R&D investment, beginning of quarter($millions)
cinvms	Microsoft cumulative browser R&D investment, beginning of quarter($millions)
cinvdif	Netscape - Microsoft cumulative browser R&D investment, beginning of quarter ($millions)
qns	Netscape browser quality = 1 - 1/cinvns
qms	Microsoft browser quality = 1 - 1/cinvms
dqdif	Netscape - Microsoft browser quality = qns - qms
verns	Netscape version number
verms	Microsoft version number
iap	Proportion of IAP accounts covered by Microsoft exclusivity agreements
window	Proportion of new computers sold with IE exclusively on desktop
revns	Netscape browser-related revenue ($millions)
revms	Microsoft browser-related revenue ($millions)

Sources:

1) **Number of Internet Users** - Quarterly data on number of Internet users were obtained from IDC. It is assumed that every Internet user utilizes exactly one browser, meaning the number of Internet users is equal to total installed base of browser users.

2) **Browser Market Shares** - Browser market shares were obtained from UIUC. UIUC measures market share by counting the number of hits registered by each browser type on UIUC's servers. For this model, it is assumed that these shares are representative of the actual shares of Internet users using each browser.

3) **Browser Investment** - Browser investment for Netscape was estimated based on total firm-wide investment from analyst reports and on the authors' assumptions about the percentage of firm-wide investment devoted to browser development. These assumptions were based in part on anecdotal evidence from Cusumano and Yoffie. Browser investment for Microsoft was estimated based on anecdotal evidence from the Findings of Fact in the DOJ antitrust case.

4) **Revenue per browser** - Revenue per browser was based on Internet advertising revenue trends, and for Microsoft, a percentage of operating system revenues which may have been at risk if another browser gained ubiquity.

5) **Bad acts** - The bad act intensities were calculated as described in the Section 3.2. For the IAP act, quarterly data on IAP subscriptions were obtained from AdKnowledge. For the WINDOW act, quarterly data on OEM shipments were obtained from IDC. The tying of IE to the Windows operating system began in the third quarter of 1998 with the release of Windows 98 SE; we assume that this operating system was phased in over three quarters.

date	prd	sbasens	sbasems	sbns	usetot	usenew	basens	basems	basedif
01/01/96	1	0.967	0.033	0.967	1.933	0.925	1.796	0.062	1.734
04/01/96	2	0.964	0.036	0.964	2.446	1.093	2.145	0.080	2.065
07/01/96	3	0.950	0.050	0.950	3.010	1.298	2.518	0.133	2.386
10/01/96	4	0.929	0.071	0.929	3.624	1.517	2.931	0.223	2.708
01/01/97	5	0.905	0.095	0.905	4.177	1.640	3.371	0.353	3.018
04/01/97	6	0.831	0.169	0.831	4.840	1.917	3.529	0.718	2.811
07/01/97	7	0.772	0.228	0.772	5.413	2.025	3.774	1.116	2.658
10/01/97	8	0.712	0.288	0.712	6.028	2.239	3.877	1.571	2.306
01/01/98	9	0.663	0.337	0.663	7.182	2.962	4.012	2.040	1.972
04/01/98	10	0.612	0.388	0.612	8.159	3.132	4.408	2.791	1.618
07/01/98	11	0.573	0.427	0.573	9.140	3.428	4.683	3.488	1.195
10/01/98	12	0.546	0.454	0.546	9.987	3.589	4.992	4.156	0.837
01/01/99	13	0.519	0.481	0.519	10.880	3.889	5.181	4.811	0.370
04/01/99	14	0.483	0.517	0.483	11.766	4.150	5.261	5.624	-0.363
07/01/99	15	0.454	0.546	0.454	12.690	4.453	5.342	6.427	-1.084
10/01/99	16	0.428	0.572	0.428	13.651	4.768	5.432	7.260	-1.828
01/01/00	17	0.397	0.603	0.397	14.118	4.562	5.424	8.229	-2.805
04/01/00	18	0.376	0.624	0.376	14.479	4.597	5.302	8.817	-3.515
07/01/00	19	0.339	0.661	0.339	14.988	4.853	4.907	9.574	-4.667
10/01/00	20	0.300	0.700	0.300	15.505	5.013	4.502	10.487	-5.984
01/01/01	21	0.268	0.732	0.268	16.098	5.244	4.154	11.351	-7.197
04/01/01	22	0.230	0.770	0.230	16.699	5.430	3.708	12.390	-8.682
07/01/01	23	0.200	0.800	0.200	17.308	5.619	3.343	13.356	-10.013
10/01/01	24	0.174	0.826	0.174	17.926	5.811	3.014	14.294	-11.280
01/01/02	25	0.160	0.840	0.160	18.485	5.936	2.865	15.061	-12.196
04/01/02	26	0.147	0.853	0.147	19.049	6.110	2.718	15.766	-13.048

prd	salens	salems	saledif	ssalens	ssalems	invns	invms	invdif	cinvns	cinvms
1	0.888	0.037	0.851	0.960	0.040	0.854	0.950	-0.097	5.850	1.800
2	1.017	0.077	0.940	0.930	0.070	0.848	1.097	-0.249	6.704	2.750
3	1.168	0.130	1.038	0.900	0.100	0.891	1.500	-0.609	7.551	3.847
4	1.320	0.197	1.123	0.870	0.130	0.847	1.489	-0.641	8.443	5.347
5	1.169	0.471	0.699	0.713	0.287	0.805	1.616	-0.811	9.290	6.836
6	1.303	0.613	0.690	0.680	0.320	0.580	1.613	-1.033	10.095	8.452
7	1.235	0.790	0.445	0.610	0.390	0.637	1.755	-1.118	10.674	10.064
8	1.299	0.940	0.358	0.580	0.420	0.623	1.890	-1.267	11.311	11.819
9	1.600	1.363	0.237	0.540	0.460	0.718	2.084	-1.366	11.934	13.709
10	1.597	1.535	0.063	0.510	0.490	0.697	1.993	-1.296	12.652	15.793
11	1.714	1.714	0.000	0.500	0.500	0.845	2.623	-1.778	13.349	17.786
12	1.687	1.902	-0.215	0.470	0.530	0.945	2.145	-1.201	14.194	20.409
13	1.634	2.256	-0.622	0.420	0.580	0.928	2.255	-1.327	15.139	22.554
14	1.660	2.490	-0.830	0.400	0.600	0.907	2.448	-1.541	16.066	24.809
15	1.692	2.761	-1.069	0.380	0.620	0.747	3.160	-2.413	16.974	27.257
16	1.621	3.147	-1.526	0.340	0.660	0.586	3.100	-2.514	17.720	30.417
17	1.506	3.057	-1.551	0.330	0.670	0.425	1.218	-0.793	18.306	33.517
18	1.195	3.402	-2.207	0.260	0.740	0.399	1.359	-0.960	18.731	34.735
19	1.068	3.785	-2.718	0.220	0.780	0.384	1.349	-0.965	19.130	36.093
20	1.003	4.010	-3.008	0.200	0.800	0.374	1.240	-0.866	19.514	37.442
21	0.800	4.444	-3.644	0.153	0.847	0.266	1.284	-1.019	19.889	38.682
22	0.747	4.683	-3.936	0.138	0.862	0.211	1.387	-1.176	20.154	39.967
23	0.674	4.945	-4.271	0.120	0.880	0.172	1.769	-1.597	20.365	41.353
24	0.755	5.055	-4.300	0.130	0.870	0.133	1.314	-1.181	20.537	43.122
25	0.712	5.224	-4.511	0.120	0.880	0.106	1.324	-1.218	20.670	44.436
26	0.794	5.316	-4.522	0.130	0.870	0.084	1.304	-1.220	20.776	45.761

prd	cinvdif	qns	qms	qdif	verns	verms	iap	window	revns	revms
1	4.050	0.829	0.444	0.385	1.1	2.0	0.000	0.008	7.434	0.000
2	3.953	0.851	0.636	0.214	2.0	2.0	0.000	0.008	6.148	9.063
3	3.704	0.868	0.740	0.127	2.0	2.0	0.000	0.064	8.228	9.262
4	3.095	0.882	0.813	0.069	3.0	3.0	0.000	0.067	8.155	9.518
5	2.454	0.892	0.854	0.039	3.0	3.0	0.004	0.058	6.231	9.556
6	1.643	0.901	0.882	0.019	3.0	3.0	0.012	0.058	5.297	10.211
7	0.610	0.906	0.901	0.006	4.0	3.0	0.023	0.067	7.414	10.097
8	-0.508	0.912	0.915	-0.004	4.0	4.0	0.009	0.069	4.561	10.787
9	-1.775	0.916	0.927	-0.011	4.0	4.0	0.008	0.060	2.444	10.444
10	-3.141	0.921	0.937	-0.016	4.0	4.0	0.017	0.058	2.592	10.592
11	-4.437	0.925	0.944	-0.019	4.0	4.0	0.010	0.315	2.686	10.686
12	-6.215	0.930	0.951	-0.021	4.5	4.0	0.000	0.570	3.284	11.284
13	-7.415	0.934	0.956	-0.022	4.5	4.0	0.000	0.812	3.185	11.185
14	-8.743	0.938	0.960	-0.022	4.5	5.0	0.000	1.000	3.969	11.969
15	-10.283	0.941	0.963	-0.022	4.5	5.0	0.000	1.000	4.795	12.795
16	-12.696	0.944	0.967	-0.024	4.5	5.0	0.000	1.000	6.509	14.509
17	-15.211	0.945	0.970	-0.025	4.5	5.0	0.000	1.000	6.807	14.807
18	-16.003	0.947	0.971	-0.025	4.5	5.0	0.000	1.000	7.221	15.221
19	-16.963	0.948	0.972	-0.025	4.5	5.5	0.000	1.000	6.508	14.508
20	-17.928	0.949	0.973	-0.025	6.0	5.5	0.000	1.000	6.846	14.846
21	-18.794	0.950	0.974	-0.024	6.0	5.5	0.000	1.000	5.815	13.815
22	-19.812	0.950	0.975	-0.025	6.0	5.5	0.000	1.000	5.533	13.533
23	-20.988	0.951	0.976	-0.025	6.0	5.5	0.000	1.000	5.122	13.122
24	-22.585	0.951	0.977	-0.026	6.0	6.0	0.000	1.000	4.577	12.577
25	-23.766	0.952	0.977	-0.026	6.0	6.0	0.000	1.000	4.112	12.112
26	-24.985	0.952	0.978	-0.026	6.0	6.0	0.000	1.000	3.827	11.827

References

Alexandrov, A.D., (1939) "Almost Everywhere Existence of Second Differentials of Convex Functions" (Russian) *Lenningrad State University, Annals*, Math Series 6, 3-35.

Amir, Rabah (1996) "Continuous Stochastic Games of Capital Accumulation with Convex Transitions," *Games and Economic Theory*, 15, 111-151.

Amir, Rabah (2001) "Stochastic Games: A Lattice-Theoretic Approach," Core Discussion Paper.

Anderson, Simon; Andre de Palma; Jacques-Francois Thisse (1996) *Discrete Choice Theory of Product Differentiation*, MIT Press, Cambridge.

Areeda, Philip; Donald Turner (1975) "Predatory Pricing and Related Practices under Section 2 of the Sherman Act," *Harvard Law Review*, 88, 697-733.

Athey, Susan (1999) "Single Crossing Properties and the Existence of Pure Strategy Equilibria in Games of Incomplete Information," Stanford University working paper.

Bajari, Patrick; Lanier Benkard; Jonathan Levin (2003) "Estimating Dynamic Models of Imperfect Competition," Stanford University Working Paper.

Balder, Erik (2001) "On Equiibria for Discontinuous Games: Continuous Nash Approximation Schemes" University of Utrecht Working Paper.

Bhattacharya, Rabi; Mukul Majumdar (2004) "Dynamical Systems Subject to Random Shocks: An Introduction," *Economic Theory*, 23, 1-12.

Baumol, William; J.C. Panzar; R.D. Wilig (1982) *Contestable Markets and the Theory of Industry Structure*.

Beggs, Alan; Paul Klemperer (1992) "Multi-Period Competition with Switching Cost," *Econometrica*, 60, 651-666.

Benabou, Roland (1989) "Optimal Price Dynamics and Speculation with a Storable Good," *Econometrica*, 57, 41-80.

Benavides, Tomas; Genaro Acedo; Hong-Kun Xu (1996) "Random Fixed Points of Set-Valued Operations," *Proceedings of the American Mathematical Society*, 124, 831-838.

Benkard, Lanier (2003) "A Dynamic Analysis of the Market for Wide-Bodied Aircraft," Stanford University Working Paper.

Bhasker, V.; Fernando Vega-Redondo (1998) "Asynchronous Choice and Markov Equilibria," University of Essex Working Paper.

Bhattacharya, Rabi; Mukul Majumdar (2004) "Stability in Distribution of Random Pertubations of Quadratic Maps as Markov Processes," Cornell University Working Paper.

Bhattacharya, Rabi; Mukil Majumdar (2004) "Dynamical Systems Subject to Random Shocks: An Introduction," *Economic Theory*, 23, 1-12.

Billingsley, Patrick (1968) *Convergence of Probability Measures*, Wiley, New York.

Billingsley, Patrick (1986) *Probability and Measure*, Wiley, New York.

Billingsley, Patrick (1969) “Weak Convergence of Measures,” *Conference Board of the Mathematical Sciences*.

Bittanit, Sergio; Alan Laub; Jan Willems (1991) *The Riccati Equation*, Springer-Verlag: Berlin.

Brander, James; Anming Zhang (1993) “Dynamic Oligopoly Behavior in the Airline Industry,” *International Journal of Industrial Organization*, 11, 407-433.

Bresnahan, Timothy; Shane Greenstein (1997) “Technological Competition and the Structure of the Computer Industry,” Stanford University working paper.

Bresnahan, Timothy (2000) “Competition, Cooperation, and Predation in Innovative Industries,” Stanford University working paper.

Bresnahan, Timothy (2001) “Network Effects and Microsoft,” Stanford University working paper.

Bresnahan, Timothy (2001) “The Economics of the Microsoft Case,” Stanford University Working Paper.

Brynjolfsson, Erik; Chris Kemerer (1996) “Network Externalities in Microcomputer Software,” *Management Science*, 42, 1627-1647.

Bueler, Stefan (2002) “Network Competition and Supermodularity,” University of Zurich working paper.

Cabral, Luis; M. Riordan (1994) “The Learning Curve, Market Dominance, and Predatory Pricing,” *Econometrica*, 62, 1115-1140.

Cahuc, Pierre; Hubert Kempf (1997) “Alternative Time Patterns of Decisions and Dynamic Strategic Interactions,” *The Economic Journal*, 107, 1728-1741.

Caplin, Andrew; Barry Nalebuff (1991) “Aggregation and Imperfect Competition: On the Existence of Equilibrium,” *Econometrica*, 59, 25-59.

Carlton, Dennis; Michael Waldman (2000) “The Strategic Use of Tying to Preserve and Create Market Power in Evolving Industries,” University of Chicago working paper.

Carlton, Dennis; Robert Germer; Andrew Rosenfield (1997) “Communication Among Competitors, Game Theory, and Antitrust,” *George Mason Law Review*, 5, 423-.

Carlton, Dennis; Jeffrey Perloff (1994) *Modern Industrial Organization*, Addison-Wesley, New York.

Cartigny, P.’ P. Michel 2003) “On the Selection of One Feedback Nash Equilibrium in Discounted Linear-Quadratic Games,” *Journal of Optimization Theory and Applications*, 117, 231-243.

Chen, Yongmin; Robert Rosenthal (1996) “Dynamic Duopolies with Slowly Changing Customers Loyalties,” International Journal of Industrial Economics, 14, 269-296.

Clark, Richard; Kevin Hassett; Zoya Ivanova; Laurence Kotlikoff (2004) “Assessing the Economic Gains from Telecom Competition,” NBER working paper.

Curtat, Laurent (1996) “Markov Equilibria of Stochastic Games with Complementarities,” *Games and Economic Behavior*, 17, 177-199.

Cusumano, Michael; David Yoffie (1998) *Competing on Internet Time*, The Free Press, New York.

Davis, Steven; Kevin Murphy (2000) “A Competitive Perspective on Internet Explorer,” NBER working paper.

Deneckere, Raymond; Dan Kovenock (1992) "Price Leadership," *The Review of Economic Studies*, 59, 143-162.

Ding, Xiu (1997), "A Coincidence Theorem Involving Contractable Spaces" in Contractible Spaces, *Applied Math Letters*,10, 53-56.

Doganoglu, Toker (2003) "Dynamic Price Competition with Consumption Externalities," *Netnomics*, 5, 43-69.

Doraszelski, Ulrich; Mark Satterthwaite (2003) "Foundations of Markov-Perfect Industry Dynamics: Existence, Purification, and Multiplicity," Stanford University Working Paper.

Doyen, Luc (2000) "Lipschitz Kernel of a Closed Set-Valued Map," *Set-Valued Analysis*, 8, 101-109.

Duffie, D.; J. Geanakoplos; A. Mas-Colell; A. McLennan (1994) "Stationary Markov Equilibria," *Econometrica*, 62, 745-781.

Dunford, Nelson; Jacob Schwartz (1964) *Linear Operators*, Vol 1, Interscience: New York.

Dynkin, Ergenii; Aleksandr Yushkevich (1969) *Markov Processes*, Plenum Press, New York.

Economides, Nicholas (2001) "The Microsoft Antitrust Case," *Journal of Industry, Competition and Trade*.

Erickson, Richard; Ariel Pakes (1995) "Markov Perfect Industry Dynamics: A Framework for Empirical Work," *The Review of Economic Studies*, 62, 53-82.

Evans, David; Richard Schmalensee (2000) "The Economics of the Microsoft Antitrust Case: A Post-Trial Primer,", NERA working paper.

Evans, David; Albert Nichols; Richard Schmalensee (2001) "An Analysis of the Government's Economic Case in U.S. v. Microsoft," *Antitrust Bulletin*.

Evans, David (2001) "U.S. v. Microsoft: The Governement's Claims and the Court's Rulings," Washington Legal Foundation working paper.

Farrell, Joseph; Carl Shapiro (1988) "Dynamic Competition with Switching Costs," *The RAND Journal of Economics*, 19, 123-137.

Farrell, Joseph; Michael Katz (1999) Predation in Network Markets," University of California Berkeley working paper.

Federal Trade Commission (1997) *Horizontal Merger Guidelines*, http://www.ftc.gov/opa/1997/04/effpress.htm

Federgruen, A. (1978) "On N-Person Stochastic Games with Denumerable State Space," *Advances in Applied Probability*, 10, 452-471.

Fershtman, Chaim; Ariel Pakes (2000) "A Dynamic Oligopoly with Collusion and Price Wars," *The RAND Journal of Economics*, 31.

Fisher, Franklin; Daniel Rubinfeld (2000) "*U.S. v. Microsoft*: An Economic Analysis," in *Did Microsoft Harm Consumers: Two Opposing Views*, AEI-Brookings Joint Center for Regulatory Studies: Washington.

Froeb, Luke; Steven Tschantz; Phillip Cook (2003) "Bertrand Competition with Capacity Constraints: Mergers Among Parking Lots," *Journal of Econometrics*, 113, 49-67.

Fudenberg, Drew; Jean Tirole (2000) *Game Theory*, MIT Press, Cambridge.

Fudenberg, Drew; Jean Tirole (2000) "Pricing a Network Good to Deter Entry," *Journal of Industrial Economics*, 48, 373-390.

Garela, Alfredo; Robert Smity (1998) "Existence of Markov Perfect Equilibria (MPE) in Undiscounted Infinite Horizon Games," University of Michigan Working Paper.

Giannitsarou, Chryssi; Flavio Toxvaerd (2003) "Recursive Global Games," Hebrew University working paper.

Gilbert, Richard (1998) "Networks, Standards, and the Use of Market Dominance," in J. Kwoka and J. White (eds) The Antitrust Revolution: The Role of Economics, 3rd Edition, Oxford University Press: Oxford.

Gilbert, Richard; Michael Katz (2001) "An Economist's Guide to *U.S. v. Microsoft*," University of California, Berkeley working paper.

Gorniewicz, L.; S. Marano; M. Slosarski (1996) "Fixed Points of Contractive Multivalued Maps," *Proceedings of the American Mathematical Society*, 124, 2675-2683.

Gutev, Valentin; Stoyan Nedev (2003 "Continuous Selections and Reflexive Banach Spaces," University of Natal Working Paper.

Hall, Robert; Jimmy Royer; March Van Audenrode (2004) Potential Competition and the Prices of Network Goods: Desktop Software," Stanford University Working Paper.

Haller, Hans; Roger Lagunoff (2000) "Genericity and Markovian Behavior in Stochastic Games," *Econometrica*, 68, 1231-1248.

Hildenbrand, Werner (1974) *Core and Equilibria of a Large Economy*, Princeton University Press: Princeton.

Hopenhayn, Hugo; Edward Prescott (1992) "Stochastic Monotonicity and Stationary Distributions for Dynamic Economies," *Econometrica*, 60, 1387-1406.

Horniacek, Milan (2003) "Selection of Markov equilibria in a Dynamic Oligopoly with Production to Order," CERGE-EO Working Paper.

Horst, Ulrich (2003) "Stationary Equilibria in Discounted Stochastic Games with Weakly Interacting Players," Humboldt University working paper.

Ibragimov, I.; R. Has'minskii (1981) *Statistical Estimation*, Springer, Berlin.

Jackson, Dunham (1941) *Fourier Series and Orthogonal Polynomials*, Mathematical Association of America.

Judd, Kenneth (1999) *Numerical Methods in Economics*, MIT Press, Cambridge.

Judd, Kenneth (1997) "Closed-Loop Equilibrium in a Multi-Stage Innovation Race," Stanford University working paper.

Karp, Larry; Jeffrey Perloff (1996) "The Optimal Suppression of a Low-Cost Technology by a Durable-Good Monopoly," *The RAND Journal of Economics*, 27, 346-364.

Katz, Michael; Carl Shapiro (1985) "Network Externalities: Competition and Compatibility," *American Economic Review*, 75, 424-440.

Katz, Michael; Carl Shapiro (1994) "Systems Competition and Network Effects," *Journal of Economic Perspectives*.

Katz, Michael; Carl Shapiro (1999) "Antitrust in the Software Market," in *Competition, Convergence, and the Microsoft Monopoly*, Kluwer.

Keene, Michael; Kenneth Wolpin (1994) "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpretation," The Review of Economics and Statistics, 76, 648-672.

Kemperer, Paul (1995) "Competition when Consumers have Switching Costs: An Overview with Applications to Industrial Organization, Macroeconomics, and International Trade," *The Review of Economic Studies*, 62, 515-539.

Klien, Benjamin (2001) "The Microsoft Case: What Can a Dominant Firm Do to Defend Its Market Position?" Journal of Economic Perspectives, 15, 45-62.

Kreps, David; Jose Scheinkman (1983) "Quantity Precommitment and Bertrand Competition Yield Cournot Outcomes," *Bell Journal of Economics*, 14, 453-463.

Krueger, Dirk; Felix Kubler (2003) "Computing Equilibrium in OLG Models with Stochastic Production," Stanford University Working Paper

Kydland, Finn (1975) "Noncooperative and Dominant Player Solutions in Discrete Dynamic Games," International economic Review, 16, 321-335.

Levin, Jonathan (2003) "Supermodular Games," Stanford University Working Paper.

Lorentz, G. (1966), *Approximation of Functions*, Holt, Rinehart and Winston.

Marks, Robert; David Midgeley; Lee Cooper (2000) "Searching for Markov Perfet Equilibria in Oligopolistic Price Wars," University of New South Whales Working Paper.

Maskin, Eric; Jean Tirole (1988a) "A Theory of Dynamic Oligopoly. 1: Overview and Quantity Competition with Large Fixed Costs," *Econometrica*, 56, 549-569.

Maskin, Eric; Jean Tirole (1988b) "A Theory of Dynamic Oligopoly, II: Price Competition, Kinked Demand Curves, and Edgeworth Cycles," *Econometrica*, 56, 571-599.

Matzkin, Rosa L. (1999): "Nonparametric Estimation of Nonadditive Random Functions," mimeo, first version, Northwestern University.

Matzkin, Rosa L. (2003): "Nonparametric Estimation of Nonadditive Random Functions," *Econometrica*, 71, 1339-1375.

Matzkin, Rosa L. (2004) "Unobservable Instruments," Northwestern University Working Paper.

McFadden, Daniel (1978) "Existence of Non-Cooperative (Cournot-Nash) Equilibrium in an Economic Market Game," MIT Lecture Notes.

McFadden, Daniel (1989) "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration," *Econometrica*, 57, 995-1026.

Milgrom, Paul; John Roberts (1990) "Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities," *Econometrica*, 58, 1255-1277.

Mertens, Jean-Francois (2003) "A Measurable "Measurable Choice" Theorem," in A.Neyman and S. Sorin (eds) *Stochastic Games and Applications*, Kluwer: Dordrecht, 107-130.

Mertens, Jean-Francois; T. Parthasarathy (2003) "Equilibria for Discounted Stochastic Games,"in A. Neyman and S. Sorin (eds) *Stochastic Games and Applications*, Kluwer: Dordrecht, 130-172.

Miranda, Mario; Paul Fackler (2002) *Applied Computational Economics and Finance*, MIT Press, Cambridge.

Nelson, Phillip (2001) "Consumer Savings from Merger Enforcement," *Antitrust*, 83.

Neveu, Jacques (1965) *The Calculus of Probability*, Holden-Day, San Francisco.

Newey, Whitney; Daniel McFadden (1994) "Large Sample Estimation and Hypothesis Testing," in R. Engle and D. McFadden, (eds.) *Handbook of Econometrics*, Vol. IV, 2111-2245, North Holland: Amsterdam.

Neyman, Abraham; Sylvain Sorin (2003) *Stochastic Games and Applications*, Kluwer: Dordrecht.

Nowak, Andrezi (2003) "N-Person Stochastic Games: Extensions of the Finite State Space Case and Correlation," in A. Neyman and S. Sorin (eds) *Stochastic Games and Applications*, Kluwer: Dordrecht, 93-106.

Overstreet, Thomas; James Keyte; John Gale (1996) "Understanding Econometric Analysis of the Price Effects of Mergers Involving Differentiated Products," *Antitrust*, 30.

Pakes, Ariel; Paul McGuire (1994) "Computing Markov-Perfect Nash Equilibria: Numerical Implications of a Dynamic Differentiated Product Market," *The RAND Journal of Economics*, 25, 55-589.

Pakes, Ariel; Paul McGuire (2001) "Stochastic Algorithms, Symmetric Markov Perfect Equilibrium, and the "Curse of Dimensionality"," *Econometrica*, 69, 1261-1281.

Parthasarathy, K. (1967) *Probability Measures on Metric Spaces*, Academic Press, New York.

Pollard, David (1984) *Convergence of Stochastic Processes*, Springer, Berlin.

Press, William; Brian Flannery; Saul Teukolsky; William Vettering (1986) *Numerical Recipes*, Cambridge University Press, Cambridge, Sections 5.6-5.8.

Rao, Prakasa (1983) *Nonparametric Functional Estimation*, Academic Press, New York, Chapter 6.

Reiffen, Davie; Michael Vita (1995) "Is There New Thinking on Vertical Mergers?" *Antitrust Law Journal*, 63, 917-.

Rosenberg, D. (1998) "Duality and Markovian Strategies," *International Journal of Game Theory*, 27, 577-597.

Rust, John (1997) "Using Randomization to Break the "Curse of Dimensionality"," *Econometrica*, 65, 487-516.

Saveliev, Peter (1999) "Fixed Points and Selections of Set-Valued Maps on Spaces with Convexity," *International Journal of Mathematics and Mathematical Science*, 24, 595-612.

Scheffman, David; Mary Coleman (2003) "Quantitative Analysis of Potential Competitive Effects from a Merger," Federal Trade Commission Working Paper.

Shapiro, Carl; William Kovacic (2000) "Competition Policy: A Century of Economic and Legal Thinking," *Journal of Economic Perspectives*.

Shorack, Galen; Jon Wellner (1986) *Empirical Processes with Applications in Statistics*, Wiley, New York.

Solon, Eilon; Nicolas Vieille (2002) "Perturbed Markov Chaings," Northwestern University Working Paper.

Solon. Eilon (1998) “Discounted Stochastic Games,” *Mathematics of Operations Research*,” 23, 1010-1021.

Sorin, Sylvain (2003) “Discounted Stochastic Games,” in A. Neyman and S. Sorin (eds) *Stochastic Games and Applications*, Kluwer: Dordrecht, 51-55.

Starek, Roscoe; Stephen Stockum (1995) “What Makes Mergers Anticompetitive?: Unilateral Effects Analysis Under the 1992 Merger Guidelines,” *Antitrust Journal*, 801.

Tirole, Jean (2003) *The Theory of Industrial Organization*, MIT Press, Cambridge.

Topkis, Donald (1978) “Minimizing a Submodular Function on a Lattice,” *Operations Research*, 26, 305-321.

U.S. Department of Justice and the Federal Trade Commission (1992, 1997) Horizontal Merger Guidelines.

U. S. v. SunGard Data Systems, Inc. (2001) (Civ. No. 01-02196 (ESH)) 2001 WL 1474734, also available at http://www.dcd.uscourts.gov/01-2196a.pdf.

U. S. V. Microsoft (2000) Civil Action No. 98-1232 (TPJ), *Plaintiffs’ Revised Proposed Final Judgment; Court’s Finding of Facts; Conclusions of Law; Final Judgment*, District Court for the District of Columbia.

Varian, Hal (2001) “High Technology Industries and Market Structure,” University of California Berkeley working paper.

Vedenov, Dmitry; Mario Miranda (2001) “Numerical Solutions of Dynamic Oligopoly Games with Capital Investment,” *Economic Theory*, 18, 237-261.

Werden, Gregory (1988) “Merger Policy for the 21st Century: Charles D. Weller’s Guidelines are Not Up to the Task,” in Andrew Daugherty (ed) *Cournot Oligopoly*, 73.

Werden, Gregory; Luke Froeb, (1994) "The Effects of Mergers in Differentiated Products Industries: Structural Merger Policy and the Logit Model," *Journal of Law, Economics, & Organization*, 10, 407-426.

Werden, Gregory; Luke Froeb, (1998) "The Entry-Inducing Effects of Horizontal Mergers," *Journal of Industrial Economics*, 46, 525-543.

Werden, Gregory; Luke Froeb, (2002) "The Antitrust Logit Model For Predicting Unilateral Competitive Effects," *Antitrust Law Journal*, 70.

Whitt, W. (1980) “Representation and Approximation of Noncooperative Sequential Games,” *SIAM Journal of Control and Optimization*, 18, 33-48.