

The Browser War – Econometric Analysis of Markov Perfect Equilibrium in Markets with Network Effects¹

Mark Jenkins², Paul Liu², Rosa L. Matzkin³, Daniel L. McFadden⁴

April 22, 2004

PRELIMINARY, NOT FOR CITATION

ABSTRACT: When demands for heterogeneous goods in a concentrated market shift over time due to network effects, forward-looking firms consider the strategic impact of investment, pricing, and other conduct. 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 some “bad acts” by firms are eliminated. We 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. Our analysis is applied to a stylized description of the browser war between Netscape and Microsoft.

Keywords and Phrases: Oligopoly_Theory, Network_Externalities, Markov_Perfect_Equilibrium

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

¹We 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.

²The Brattle Group, Suite 1140, 353 Sacramento St., San Francisco CA 94111.

³Department of Economics, Northwestern University, 2003 Sheridan Road, Evanston, IL 60208.

⁴Corresponding author: Department of Economics, University of California, Berkeley, CA 94720-3880, mcfadden@econ.berkeley.edu.

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

I. Introduction

Concentrated markets for heterogeneous goods are often linked through time through demand or supply. Demand-side effects come from habit, inertia (switching costs), contagion (bandwagon) effects, or network (connectivity, interoperability) externalities. Supply-side effects come from investments in production capacity, product quality, and brand promotion. Forward-looking firms in such markets consider the strategic impact of their actions, and seek to maximize the expected present value of profits given the information they have and their beliefs about the behavior of other market participants. Quantitative analysis of firm conduct in such markets requires dynamic models 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 “bad acts” by a firm. This paper discusses the application of MPE to damage estimation in anti-trust litigation, and addresses some of the theoretical, computational, and econometric issues that arise. Our analysis is applied to the browser war between Netscape and Microsoft, and gives a framework for estimating the damages to Netscape caused by Microsoft “bad acts”.

The paper is organized as follows. Section 2 discusses the relationship between anti-trust litigation and dynamic models. Section 3 reviews the properties of MPE models, and discusses theoretical, computational, and econometric issues. Section 4 gives a capsule summary of the browser war between Netscape and Microsoft. Section 5 discusses criteria for a model designed to capture the most salient aspects of the browser war. Section 6 describes a stylized MPE model that meets most of these criteria. Section 7 describes the data used for estimation. Section 8 gives the results of model estimation, and provides estimates of the “but for” development of this market absent Microsoft “bad acts”. Section 9 gives conclusions.

2. Market Dynamics and Anti-Trust Law

The fundamental goal of anti-trust law, in its current interpretation by the courts, is to promote market efficiency and protect the welfare of market participants by limiting the acquisition and exercise of market power. This is accomplished through statutes 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 in which erecting entry barriers, or eliminating rival firms through acquisition or predation, play out through time. Then, anti-trust law, and assessments of liability and damages based upon it, must account consistently for the dynamic evolution of concentrated markets under alternative rules for firm conduct.

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 little attention to demand behavior. 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 now 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 behavior 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), Weiriden-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. The first of these developments has begun to influence anti-trust law through adoption of merger models that rely on analysis of consumer outcomes, rather than mechanical measures of concentration; see Federal Trade Commission (1997), Overstreet-Keyte-Gale (1998), Scheffman-Coleman (2003), Shapiro-Kovacic (2000), Starek-Stockum (1998), Reiffen-Vita (1995), Weiriden-Froeb (2002).

A more general reconsideration of the anti-trust law in light of modern economic analysis of demand behavior and market dynamics is timely. Demand-side dynamics is an important factor in most product markets, driving firm behavior on the positioning and launch of new products, and influencing their management of the product cycle. In the presence of such effects, firms conduct can have large strategic consequences. In some market circumstances, conduct that is illegal under current anti-trust law may increase market efficiency. For example, consumers may benefit in net from the ubiquity of proprietary product standards, even when they are promulgated to exclude competition. On the other hand, apparently minor announcement effects, exclusions, or delays that the law currently treats as inconsequential may be sufficient to tip a market with strong network demand or scale effects, and result in its monopolization. Even if the market remains contestable in principle, placing some limits on monopoly rents, the barriers created by these dynamic effects may be high enough to exclude entry by more innovative or potentially lower cost firms. Another example is the anti-trust charge of predatory pricing. Static analysis leads to the Areeda-Turner standard that pricing must be below short-run marginal cost (or, loosely, average variable cost) to be predatory. Further, the courts are skeptical of predation claims, on the grounds that consumers benefit from low prices in the short run, and that predation is implausible behavior 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, or supply-side scale or learning-by-doing effects, link the markets through time, neither the Areeda-Turner standard or the argument that predation is strategically unattractive necessarily hold, and the impacts of predation may be more dramatic, and the consequences for consumers more subtle, than the traditional analysis would suggest.

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, should take into account the incentives created by the market dynamics, and the strategic elements in firm response. Another reason for making market dynamics explicit in anti-trust litigation is that 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 projections of future market conditions, on the grounds that such projections are speculative. There is no question that such projections place an extra burden on the rectitude of 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.

For the anti-trust law to work effectively as a regulator of resource allocation, it must be able to discriminate between circumstances and conduct that in combination reduce market efficiency, and those that are neutral or beneficial. This point is not new, and is the motivation for the theory of contestable markets (Baumol-Panzer-Willig, 1982). However, it is particularly relevant for some issues of conduct and the dynamics of market structure. For example, suppose a firm with deep pockets uses predatory pricing and innovation to monopolize a market characterized by network effects. Consumers benefit from low prices in the short run, and in the longer run gain from the network benefit of ubiquity and lose from the higher prices and lower rates of innovation provided by the monopolist. To determine whether the predation was socially undesirable, one needs to obtain the “as is” and “but for” trajectories, and calculate the expected present value of the net welfare effect. MPE models of the type studied by Hall-Royer-Van Audenrode (2004) and by this paper provide templates for such analysis.

3. Markov Perfect Equilibrium Models of Concentrated Market Dynamics

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; for example, see Erickson-Pakes (1995), Fudenberg-Tirole (2000), 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 practical 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" 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 nuanced levels of 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 after some date T' , 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.⁵ 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.

(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 sufficiently rich to explain these observations, and sufficiently lean 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 are made. Note that 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

⁵Technically, 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 hidden Markov random fields is dense in the space of all stationary stochastic processes.

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. Note that 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.

With these considerations in mind, we write down a dynamic concentrated market model in fairly general terms, and later specialize it for application. Let $n = 1, \dots, N$ index firms, which may be incumbents or potential entrants. Let $t = 0, 1, \dots, T', \dots, T'', \dots$ index time periods, with $T' < T'' < +\infty$, where T' is a time after which the market fundamentals are time-invariant, and the econometrician observes the market in periods $t = 1, \dots, T''$, which include some when the system is time-invariant. The market is described by the following notation and accompanying assumptions:

- s_t a *state* vector in a compact subset S_t 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_t that is endowed with its Borel σ -field and a probability g_{z_t} .
- ε_t, η_t independent public exogenous shocks in a compact subset Ω of a finite-dimensional Euclidean space that is 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_n in a finite-dimensional Euclidean space with its Borel σ -field and a probability g_{ζ_n} that is absolutely continuous with respect to Lebesgue measure. Let $\zeta_t = (\zeta_{1t}, \dots, \zeta_{Nt})$ and $\Lambda = \Lambda_1 \times \dots \times \Lambda_N$. Also define $\zeta_{-nt} = (\zeta_{1t}, \dots, \zeta_{n-1,t}, \zeta_{n+1,t}, \dots, \zeta_{Nt})$.
- x_{nt} a policy for firm n , in a compact subset X_{nt} of a finite-dimensional Euclidean space, endowed with its Borel σ -field. Let $x_t = (x_{1t}, \dots, x_{Nt})$ denote a *profile of policies* of the firms, and $X_t = X_{1t} \times \dots \times X_{Nt}$. Define the profile of rivals' policies, $x_{-n,t} = (x_{1t}, \dots, x_{n-1,t}, x_{n+1,t}, \dots, x_{Nt}) \in X_{-n,t} = X_{1t} \times \dots \times X_{n-1,t} \times X_{n+1,t} \times \dots \times X_{Nt}$.
- Π_n a profit function for firm n , a function on $S_t \times X_t \times Z_t \times \Omega \times \Lambda_n$.
- h an equation of motion, a measurable map from $S_t \times X_t \times Z_t \times \Omega$ into S_{t+1} .
- β a common constant discount factor.

The state 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 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 probability g_{z_t} is interpreted as a *cumulative distribution function* or measure; this is the case for all the probabilities that appear in the model. 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 when policy decisions are made in period t , while η_t is assumed to be unknown to all firms. 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. Policy variables may include investment, pricing, product attributes, entry and exit from product lines, and actions such as bundling, tying, refusals to deal, and exclusionary contracts that are potentially “bad acts” under anti-trust law. Summarizing, the payoff-relevant information of firm n when it makes its policy decision in period t is a point $(s_t, z_t, \varepsilon_t, \zeta_{nt}) \in S_t \times Z_t \times \Omega \times \Lambda_n$. A *Markov mixed strategy* for the firm is a measurable mapping f_{nt} from this information into the space $P(X_{nt})$ of probabilities on X_{nt} . Note that this object is not the same as a *best-response* or *reaction* function, which maps the information above *and* the strategies of rivals into $P(X_{nt})$; 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 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.⁶

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 , and should be interpreted as the expectation of profit with respect to any such exogenous variables that are unknown to the firm. Note that it is time-invariant; with all time-varying effects on the profit function captured in the exogenous vector z_t . An important special case we will analyze will be profit functions in

⁶Some markets are oligoposonistic, with atomistic buyers who can be included as players in the game. 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.

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 unit costs associated with implementing the firm's policy, 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 is $\Pi_n(s_t, x_t, z_t, \varepsilon_t, \zeta_{nt}) \equiv R_n(s_t, x_t, z_t, \varepsilon_t) - C_n(s_t, x_{nt}, z_t, x_{nt} \cdot \zeta_{nt})$, 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_{nt} be a convex set or that the profit function be concave in x_{nt} . 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. It may have beliefs about its rivals' information and behavior, but these beliefs must in the end also depend only on the information that it has..

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

Assumption 1. The shocks $\varepsilon_t, \eta_t, \zeta_t \in \Omega \times \Omega \times \Lambda$ are independent and identically distributed across time, and independent of each other, and their probabilities $g_\varepsilon, g_\eta,$ and g_{ζ_n} are common knowledge. The observable vectors $z_t \in Z_t$ are also independent across time, with probabilities g_{z_t} that 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, η_t , and all future exogenous vectors and shocks, are not predictable by a firm. The econometrician observes z_t for $t = 1, \dots, T''$, but does not observe the shocks $\varepsilon_t, \eta_t,$ or ζ_t .⁷

Assumption 2. The market structure is time-homogeneous beginning at time T' . Specifically, the state space S_t , the exogenous vector space Z_t , and the policy spaces X_{1t}, \dots, X_{Nt} are independent of t for $t \geq T'$.

We will study the game under the following simplifying regularity assumption, chosen because it is easy to check and likely to be satisfied in applications.

Assumption 3. The profit functions $\Pi_n(s_t, x_t, z_t, \varepsilon_t, \zeta_{nt})$ and the equation of motion $h(s_t, x_t, z_t, \varepsilon_t, \eta_t)$ are Lipschitz in their arguments.⁸ The discount factor satisfies $0 < \beta < 1$.

A *Markov Perfect Equilibrium* (MPE) is a profile of Markov mixed strategies f_{nt}^* for all firms and periods with the subgame-perfect Nash property that for each period t and firm n , given the current strategies of rivals and all future strategies, firm n has no incentive to alter its assigned strategy f_{nt}^* . Note that $f_{nt}^*(x|s_t, z_t, \varepsilon_t, \zeta_{nt})$ is a cumulative distribution function conditioned on $s_t, z_t, \varepsilon_t, \zeta_{nt}$. It is a function from $X_t \times S_t \times Z_t \times \Omega \times \Lambda_n$ into the unit interval, and can also be interpreted as the univariate distribution of a stochastic process from $S_t \times Z_t \times \Omega \times \Lambda_n$

⁷The assumption that the exogenous variables and shocks are not predictable by the firms can be softened by interpreting these variables as *innovations* in stochastic processes, and building the *predictable* parts of these processes into the state vectors and laws of motion.

⁸If a state or exogenous vector has some discrete components, then the Lipschitz condition is satisfied trivially in these components.

into X_t . Let $V_{nt}(s_t, z_t, \varepsilon_t, \zeta_{nt})$ 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. 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 $t' > t$ and the contemporaneous shocks η_t and the private information of rivals, $\zeta_{-n,t} = (\zeta_{1t}, \dots, \zeta_{n-1,t}, \zeta_{n+1,t}, \dots, \zeta_{Nt})$. The valuation functions then satisfy Bellman's equations,

$$(1) \quad V_{nt}(s_t, z_t, \varepsilon_t, \zeta_{nt}) = \max_{x_{nt}} \{ \rho_{nt}(s_t, x_{nt}, z_t, \varepsilon_t, \zeta_{nt}) + \beta \mu_{nt}(s_t, x_{nt}, z_t, \varepsilon_t) \},$$

with the notation

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

for the expected value of current profit, and

$$(3) \quad \mu_{nt}(s_t, x_{nt}, z_t, \varepsilon_t) \equiv \mathbf{E}_{-n,t} \int_{X_{-n,t}} \prod_{j \neq n} f_{jt}^*(dx_{jt} | s_t, z_t, \varepsilon_t, \zeta_{jt}) V_{n,t+1}(h(s_t, x_{-n,t}, x_{nt}, z_t, \varepsilon_t, \eta_t), z_{t+1}, \varepsilon_{t+1}, \zeta_{n,t+1})$$

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 (1). The Nash property of MPE requires that the support of $f_{nt}^*(\cdot | s_t, z_t, \varepsilon_t, \zeta_{nt})$ be contained in $X_{nt}^*(s_t, z_t, \varepsilon_t, \zeta_{nt})$.

Existence 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. The result is a minor variation on theorems of Federgruen and Whitt. We include a proof for completeness, and also as a basis for some conjectures that follow. 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 $\{A_{1Kt}, \dots, A_{KKt}\}$ of $S_t \times Z_t \times \Omega$ for $K \rightarrow \infty$, 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 $(s_t, z_t, \varepsilon_t)$ 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 $(s_t, z_t, \varepsilon_t)$ 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 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 $n = 1, \dots, N$ that are time-invariant for $t \geq T'$ and satisfy condition (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 $k = 1, \dots, K$ and $K \rightarrow \infty$.

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 (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 $M/(1-\beta)$ 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 $A = A_x \times A_K \times A_z$ in this space,

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

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 $(V_1, \dots, V_N, f_1, \dots, f_n) \in W_1 \times \dots \times W_N \times F_1 \times \dots \times F_N$, the function $\rho_{nT}(s_T, x_{nT}, z_T, \varepsilon_T, \zeta_{nT})$ from (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 $(V_1, \dots, V_N, f_1, \dots, f_n) \in W_1 \times \dots \times W_N \times F_1 \times \dots \times F_N$, the function $\mu_{nT}(s_T, x_{nT}, z_T, \varepsilon_T)$ from (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 (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

$$\begin{aligned} & -M|\zeta_{nT}' - \zeta_{nT}''| \\ & \leq \rho_{nT}(s_T, x_{nT}(\zeta_{nT}''), z_T, \varepsilon_T, \zeta_{nT}'') - \rho_{nT}(s_T, x_{nT}(\zeta_{nT}'), z_T, \varepsilon_T, \zeta_{nT}'') \\ & \leq V_n'(s_T, z_T, \varepsilon_T, \zeta_{nT}') - V_n'(s_T, z_T, \varepsilon_T, \zeta_{nT}'') \\ & \leq \rho_{nT}(s_T, x_{nT}(\zeta_{nT}'), z_T, \varepsilon_T, \zeta_{nT}') - \rho_{nT}(s_T, x_{nT}(\zeta_{nT}'), z_T, \varepsilon_T, \zeta_{nT}'') \leq M|\zeta_{nT}' - \zeta_{nT}''|. \end{aligned}$$

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 $W_1 \times \dots \times W_N \times F_1 \times \dots \times F_N$ 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 $(V_{1T}, \dots, V_{NT}, f_{1T}^*, \dots, f_{NT}^*) \in W_1 \times \dots \times W_N \times F_1 \times \dots \times F_N$ that satisfies (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 (1) from period T' . Let T be a period satisfying $T \leq T'$, and suppose that strategies f_{nt}^* and valuation functions V_{nt} have been established satisfying (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 $F_1 \times \dots \times F_N$ into itself. Therefore, the Glicksburg-Kakutani fixed point theorem establishes the existence of strategies $(f_{1T}^*, \dots, f_{NT}^*) \in F_1 \times \dots \times F_N$ satisfying (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 $K \rightarrow \infty$. 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 (1) on the countable union of partition points B^* . \square

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

used 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 a future draft of this paper.

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; see Appendix 1. 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.

Assumption 4. For $n = 1, \dots, N$, the profit function of the firm can be written as a revenue function minus a cost function,

$$\Pi_n(s_t, x_t, z_t, \varepsilon_t, \zeta_{nt}) \equiv R_n(s_t, x_t, z_t, \varepsilon_t) - C_n(s_t, x_{nt}, z_t, x_{nt} \cdot \zeta_{nt})$$

where revenue is independent of the private information ζ_{nt} , and cost is independent of ε_t and of the policies $x_{-n,t}$ of rivals, and depends on the private information ζ_{nt} through a single index $x_{nt} \cdot \zeta_{nt}$, with C_n a strictly monotone, concave function of this index.⁹

Example: ζ_{nt} is a vector of unit costs, commensurate with x_{nt} , that gives the firm's private cost of policy x_{nt} , so that $C_n(s_t, x_{nt}, z_t, x_{nt} \cdot \zeta_{nt}) \equiv c_n(s_t, x_{nt}, z_t) + x_{nt} \cdot \zeta_{nt}$. 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 linear additive index is trivially concave and strictly monotone

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 $V_{nt}(s_t, z_t, \varepsilon_t, \zeta_{nt}) = \max_{x_{nt}} \{\rho_n(s_t, x_{nt}, z_t, \varepsilon_t, \zeta_{nt}) + \beta \mu_n(s_t, x_{nt}, z_t, \varepsilon_t)\}$ from

(1). We will argue that V_{nt} is a convex function of ζ_{nt} , and that its derivative with respect to

⁹For the purposes of Theorem 2, the separation of the profit function into revenue and cost, and the exclusion of $x_{-n,t}$ and ε_t from the cost function, are not needed.

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 3 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 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 ζ_{nt}' and ζ_{nt}'' are distinct points, consider a convex combination $\theta\zeta_{nt}' + (1-\theta)\zeta_{nt}''$, and let x_{nt}^* be any maximizer that gives the function $V_{nt}(s_t, z_t, \varepsilon_t, \theta\zeta_{nt}' + (1-\theta)\zeta_{nt}'')$. Then,

$$\begin{aligned} V_{nt}(s_t, z_t, \varepsilon_t, \theta\zeta_{nt}' + (1-\theta)\zeta_{nt}'') &= \rho_n(s_t, x_{nt}^*, z_t, \varepsilon_t, \theta\zeta_{nt}' + (1-\theta)\zeta_{nt}'') + \beta\mu_n(s_t, x_{nt}^*, z_t, \varepsilon_t) \\ &\leq \theta\{\rho_n(s_t, x_{nt}^*, z_t, \varepsilon_t, \zeta_{nt}') \cdot x_{nt}^*\} + \beta\mu_n(s_t, x_{nt}^*, z_t, \varepsilon_t) \\ &\quad + (1-\theta)\{\rho_n(s_t, x_{nt}^*, z_t, \varepsilon_t, \zeta_{nt}'') \cdot x_{nt}^*\} + \beta\mu_n(s_t, x_{nt}^*, z_t, \varepsilon_t) \\ &\leq \theta V_{nt}(s_t, z_t, \varepsilon_t, \zeta_{nt}') + (1-\theta)V_{nt}(s_t, z_t, \varepsilon_t, \zeta_{nt}''). \end{aligned}$$

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,

$$\begin{aligned} V_{nt}(s_t, z_t, \varepsilon_t, \theta\zeta_{nt}' + (1-\theta)\zeta_{nt}'') &\geq \rho_n(s_t, x_{nt}'', z_t, \varepsilon_t, [\theta\zeta_{nt}' + (1-\theta)\zeta_{nt}''] \cdot x_{nt}'') + \beta\mu_n(s_t, x_{nt}'', z_t, \varepsilon_t) \\ &\geq V_{nt}(s_t, z_t, \varepsilon_t, \zeta_{nt}'') - \max\{\lambda_+ \theta(\zeta_{nt}'' - \zeta_{nt}') \cdot x_{nt}'', \lambda_- \theta(\zeta_{nt}'' - \zeta_{nt}') \cdot x_{nt}''\}, \end{aligned}$$

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 $\lambda_+(\zeta_{nt}'' - \zeta_{nt}') \cdot x_{nt}''$ and $\lambda_-(\zeta_{nt}'' - \zeta_{nt}') \cdot x_{nt}''$. 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 $\lambda_+(\zeta_{nt}'' - \zeta_{nt}') \cdot x_{nt}''$. Since this

holds for each ζ_{nt}' 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. \square

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 later 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.

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). Consider the long run, time-invariant structure of the model. As in the proof of existence, one can use the Bellman recursion as a mapping from the space of valuation functions and mappings and strategies into itself. A fixed point of this mapping will give the valuation function for the model once it is time-invariant. Given this function, the Nash strategies satisfying a fixed point and the valuation functions can be obtained by backward recursion. 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. If the setup satisfies Assumptions 1-4, then almost sure existence of a pure strategy solution is guaranteed by Theorems 1 and 2. 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.

Consider approximating a function $V(x)$ on the d -dimensional cube $[-1,1]^d$ by a linear combination

$$V(x) = \sum_{0 \leq i_1 + \dots + i_d \leq m} T_{i_1}(x_1) \cdots T_{i_d}(x_d) b_{i_1 \dots i_d} + \varepsilon(x),$$

where the $T_i(z)$ are univariate Chebyshev orthogonal polynomials of degree $j = 0, \dots, m$, and $\varepsilon(x)$ is a residual. These polynomials satisfy the recursion

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

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 $i = 1, \dots, t$. Further, for $j, k \leq m$, they satisfy the finite orthogonality condition

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

Suppose z_{mj} , $j = 1, \dots, m+1$, 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

$$(4) \quad Y_k = \sum_{j=0}^M A_{kj}B_j + \varepsilon_k, \quad k = 1, \dots, K$$

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 ε_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, (4) is $Y = AB + \varepsilon$ and the least squares solution is

$$(5) \quad B = [A'A]^{-1}A'Y.$$

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 (4).) 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 (4) which can be written in partitioned matrix form

$$(6) \quad \begin{bmatrix} Y \\ y \end{bmatrix} = \begin{bmatrix} A & F \\ G & H \end{bmatrix} \begin{bmatrix} B \\ b \end{bmatrix} + \begin{bmatrix} \varepsilon \\ v \end{bmatrix},$$

where Y is $K \times 1$, y is $k \times 1$, A is the $K \times M$ array of Chebyshev polynomials evaluated at the original grid points, F is the $K \times m$ array of augmented polynomials evaluated at the original grid points, G is the $k \times M$ array of Chebyshev polynomials evaluated at the augmented grid

points, H is the $k \times m$ array of augmented polynomials evaluated at the augmented grid points, B is $M \times 1$, and b is $m \times 1$. The least squares solution solves

$$(7) \quad \begin{bmatrix} A'Y + G'y \\ F'Y + H'y \end{bmatrix} = \begin{bmatrix} A'A + G'G & A'F + G'H \\ F'A + H'G & F'F + H'H \end{bmatrix} \begin{bmatrix} B \\ b \end{bmatrix}.$$

Define the $M \times M$ matrix

$$(8) \quad Q = D^{-1} - D^{-1}G'[I_k + GD^{-1}G']^{-1}GD^{-1}.$$

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 $k \times k$ inverse $[I_k + GD^{-1}G']^{-1}$. The formula for a partitioned inverse is

$$(9) \quad \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}^{-1} = \begin{bmatrix} (C_{11})^{-1} + (C_{11})^{-1}C_{12}R^{-1}C_{21}(C_{11})^{-1} & -(C_{11})^{-1}C_{12}R^{-1} \\ -R^{-1}C_{21}(C_{11})^{-1} & R^{-1} \end{bmatrix},$$

where $R = C_{22} - C_{21}(C_{11})^{-1}C_{12}$. Apply this formula to (7). 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 $m \times m$ matrix

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

Computation of the solution to (7) requires inversion of the $m \times m$ matrix R , as well as computation of the matrix Q which contains the $k \times k$ 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.

In application, we will have valuation functions $V_{jt}(s, z, \varepsilon, \zeta_n, \theta)$ for each player that depend on the vector θ of parameters of the problem, and we will have to approximate these functions in steady state, and in each of the observation periods where exogenous forcing variables are operating. However, the right-hand-side array in (6) does not depend on θ , so the weights to be applied to the $Y(\theta)$ in the augmented version of (5) to calculate $B(\theta)$ can be calculated once, before iterative search for θ .

The MPE solution algorithm for a given θ can be characterized as follows: First approximate the time-invariant valuation functions using the Chebyshev polynomial approximations and augmented grid just described. The givens of the computation include specification of (empirical) distributions for shocks and exogenous variables. Search for coefficients of these polynomials that are a fixed point of the Bellman recursion. We use a gradient search for the fixed point, but note that 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 be computationally demanding. Once the stationary valuation function is determined, then the valuation functions in the initial non-stationary part of the problem are calculated by backward recursion. Once this is complete, the model is run forward, with draws from the empirical distributions of the shocks, to estimate trajectories. The distance between the estimated and observed trajectories is then computed, and the process is iterated over trial values of the parameters θ to minimize this distance, following the econometric method described next.

Estimation Strategies

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.

We divide the econometric problem into three 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 disturbances of the system, and their estimation within the limits of identification determine these partially or fully. This substantially reduces the burden of estimation search within the computation of the MPE. Specifically, structure that is estimated consistently in the first stage can be held fixed during stage two while remaining parts of the system are estimated, and then all the parameters of the system 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, if the equation of motion and profit functions include both exogenous shocks and current policies that the firms choose knowing these shocks, then 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 disturbances, 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 should not be 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.

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

$$(10) \quad s_{t+1} = h(s_t, x_t, z_t, \varepsilon_t, \eta_t),$$

and the profit functions

$$(11) \quad \pi_{nt} = R_n(s_t, x_t, z_t, \varepsilon_t) - C_n(s_t, x_{nt}, z_t, x_{nt}, \zeta_{nt}).$$

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

$$(12) \quad x_{nt} = r_{nt}(s_t, z_t, \varepsilon_t, \zeta_{nt})$$

for $n = 1, \dots, N$. In the first stage of estimation, the structure of r_{nt} in (12) 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 (10) are independent of the remaining shock η_t . If this shock is additive, then (10) can be estimated by nonlinear least squares. If the shock is not additive, then Matzkin (2003) shows that either a dimension-reducing functional restriction, or 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.¹⁰ 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).

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 disturbances 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 disturbances 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 ε_t' is uniformly distributed, and then

¹⁰ 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.

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 $n = 2, \dots, N$ 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

$$(13) \quad r_{1t}(s_t, z_t, \varepsilon_t, \zeta_{1t}) = g_{x_1|s,z,\zeta_1}^{-1}(g_\varepsilon(\varepsilon_t)),$$

where $g_{x_1|s,z,\zeta_1}$ 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 ,

$$(14) \quad \varepsilon_t = g_\varepsilon^{-1}(g_{x_1|s,z,\zeta_1}(x_{1t})).$$

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 (12) for $n > 1$, we can now treat ε_t as observable. Applying again Matzkin (1999, 2003), this time to (12) for $n > 1$, we get

$$(15) \quad r_{nt}(s_t, z_t, \varepsilon_t, \zeta_{nt}) = g_{x_n|s,z,\varepsilon}^{-1}(g_{\zeta_2}(\zeta_{nt})),$$

where $g_{x_n|s,z,\varepsilon}$ 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 ,

$$(16) \quad \zeta_{nt} = g_{\zeta 2}^{-1}(g_{x_n|s,z,\varepsilon}(x_{nt}))$$

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 (14), 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

$$(17) \quad h(s_t, z_t, x_t, \varepsilon_t, \eta_t) = g_{s^+|s,z,x,\varepsilon}^{-1}(g_\eta(\eta_t))$$

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 . Moreover,

$$(18) \quad h(s_t, z_t, x_t, \varepsilon_t, \eta_t) = g_\eta^{-1}(g_{s^+|s,z,x,\varepsilon}(s_{t+1})).$$

This completes the proof. \square

Nonparametric estimators for the identified functions can be obtained by replacing the conditional distributions by nonparametric estimators of them.

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 MPE dynamic computations using the first stage parameter estimates plus a finite grid for the remaining parameters. The second stage then uses simulations from the MPE computations at each grid point in the parameter space, and uses a minimum distance criterion to pick the best grid point. 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, which we will call θ , to minimize a criterion defined as a sum of squared deviations of the predicted “as is” policies and states from their observed values for $t = 1, \dots, T$. The search is facilitated by the smooth behavior of the valuation functions and policy functions with small changes in θ , provided the solutions to Bellman’s equations are regular. Schematically, we have

$$(19) \quad V_{jt}(s, z, \theta) \approx A(s, z)B_{jt}(s, z, \theta),$$

where $A(s, z)$ is the augmented polynomial approximation (6). Suppose one starts from a solution to the MPE problem at an initial θ_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 θ , and plug these derivatives into the expression for the derivative of the estimation criterion with respect to θ . 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. Nothing is essentially different if instead of using a fixed grid for stage 2 parameters, one iterates between the MPE computation and minimum distance steps using some search algorithm in parameter space. 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, 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 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.

Issues in Damage Estimation

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 likely to be more robust.

4. The Browser Wars: Netscape versus Microsoft

In this paper we examine the market for web browsers. Web browsers are 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).¹¹ Word about Mosaic spread rapidly and by the end of 1994, Mosaic had been downloaded about two million times.¹² The introduction into the marketplace of a graphical browser was pivotal in spurring interest in the Internet; see Figure 1 below (the periods are quarters, with Period 0 representing the first quarter of 1996, and Period 25 representing the second quarter of 2002). By the second quarter of 2002 there were over 190 million Internet users as measured by IDC.

In April 1994, Andreessen partnered with James Clark to form Netscape Communications.¹³ 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”.¹⁴ By January, 1996, its share of the browser market was 90 percent.¹⁵ Navigator’s quality was a large factor in its success¹⁶; 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.¹⁷

Figure 1

¹¹ Cusumano and Yoffie, p. 3.

¹² Cusumano and Yoffie, p. 95. The figures include other methods of obtaining Mosaic, but downloads were the dominant form of distribution.

¹³ Cusumano and Yoffie, p. 7. The company was originally called Mosaic Communications.

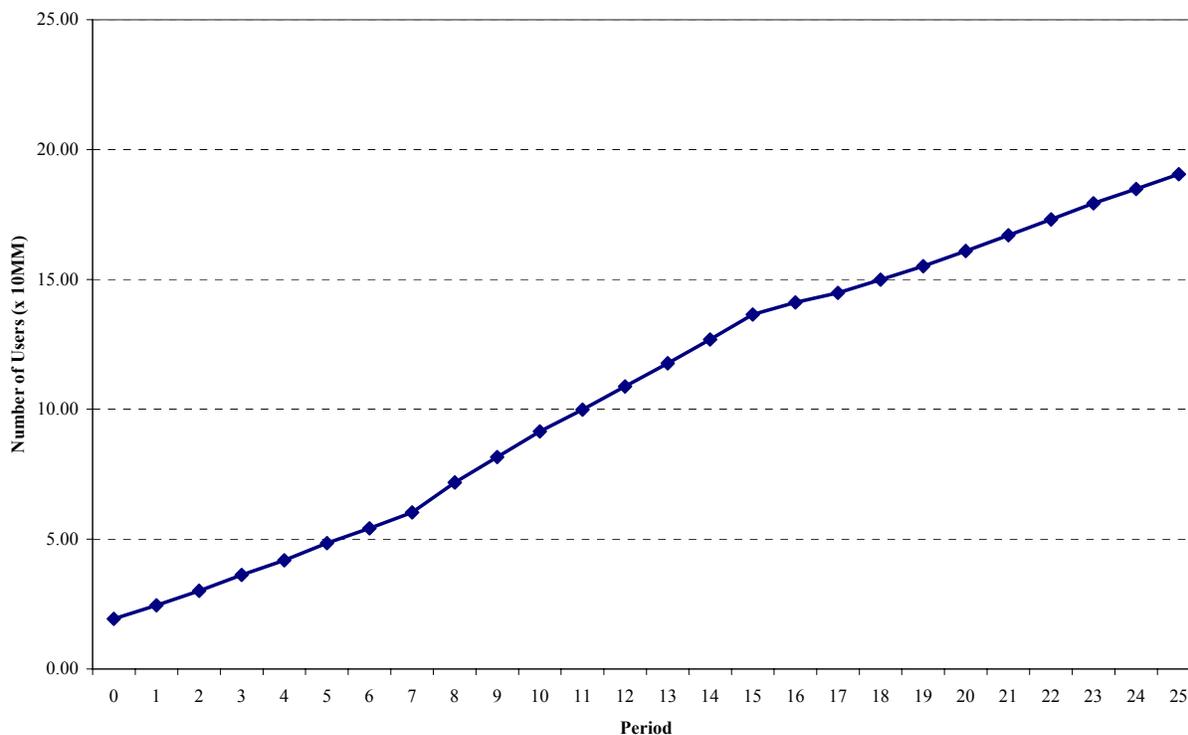
¹⁴ Cusumano and Yoffie, p. 7, 107.

¹⁵ Cusumano and Yoffie, p. 9.

¹⁶ Netscape was consistently rated as the best browser product then on the market. Cusumano and Yoffie, p. 96.

¹⁷ Cusumano and Yoffie, p. 98.

Number of Internet Users



Source: IDC.

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.¹⁸ Nonetheless, in June 1995, Microsoft and Netscape had a meeting

¹⁸ 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. [Cite]

at which Microsoft offered Netscape the option to either enter into a “special relationship” with Microsoft and agree to develop only non-Windows 95 browsers, or be regarded as a competitor and treated as such.¹⁹ Netscape declined to enter into an agreement with Microsoft, as doing so 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.²⁰ Thus began the “browser wars.”²¹ 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.

Table 1

Windows IE		Netscape	
<u>Version</u>	<u>Final Release</u>	<u>Version</u>	<u>Final Release</u>
		1.0	Dec 1994
		1.1	Apr. 1995
1.0	Aug. 1995		
2.0	Nov. 1995		
		2.0	Mar. 1996
3.0	Aug. 1996	3.0	Aug. 1996
		4.0	Jun. 1997
4.0	Oct. 1997		
		4.5	Oct. 1998
5.0	Mar. 1999		
5.5	Jul. 2000		
		6.0	Nov. 2000
6.0	Oct. 2001		
		7.0	Aug. 2002

Source: <http://www.blooberry.com/indexdot/history/brc>

¹⁹ Findings of Fact, ¶¶82-89.

²⁰ Cusumano and Yoffie, pp. 109-110.

²¹ [Cite Cusumano and Yoffie]

In bundling Internet Explorer with every copy of Windows, in making Internet Explorer separately available for free downloading, and in investing substantial resources into improving the quality of its browser (Figure 2), Microsoft sought to “cut off Netscape’s air supply”²² and undermine Netscape’s “free but not free” strategy.²³ An oft-cited figure in the DOJ’s antitrust case against Microsoft held that Microsoft invested approximately \$100 million a year.²⁴ 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 million.²⁵ The initial versions of Internet Explorer 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.

Microsoft apparently felt that the actions just described would not be sufficient to dethrone Netscape from its dominant position in the browser market. As a result, Microsoft also attempted to cut off the primary distribution channels through which Netscape distributed its browsers.²⁶ First, Microsoft refused to unbundle Internet Explorer from Windows, making it unlikely that consumers and OEMs would seek alternative browsers.²⁷ Second, Microsoft entered into agreements with many Internet Service Providers that restricted their ability to distribute alternative browsers such as Netscape Navigator.²⁸ Third, Microsoft entered into agreements with Online Services, Internet Content Providers, and

²² David A. Kaplan, *Silicon Boys*, p. 276. In January 1998, Netscape was finally forced to lower the price of its browser to nothing. Cusumano and Yoffie, p. 144.

²³ Cusumano and Yoffie, p. 141.

²⁴[CITE]

²⁵ Authors estimates, based on readings of Cusumano and Yoffie. The authors believe that while Netscape initially focused almost exclusively on browser development, later attempts to enter various server software markets led to sharp declines in browser investment as a percent of total investment.

²⁶ Findings of Fact, ¶143.

²⁷ 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.

²⁸ Findings of Fact, ¶247.

Independent Software Vendors that induced greater use of Internet Explorer. As a result of Microsoft's activities in the browser market, Netscape's market share slid steadily downward, while Microsoft's market share grew steadily upward. Figure 3 shows the two firm's market shares. Figure 4 shows that during this period, the total number of computers and the installed customer base for each browser grew rapidly, with Microsoft benefitting from its control of the distribution channels. In January 1998, Netscape was finally forced to lower the official price of its browser to zero.²⁹ By the time Microsoft's restrictions on OEMs and ISP agreements were lifted in 1998, the browser war was effectively over.³⁰ Even though Netscape still had 58 percent market share to Microsoft's 40 percent,³¹ Microsoft dominated the distribution channels and had gained all the momentum. The actions by Microsoft against Netscape were the focus in the U.S. Department of Justice's antitrust action against Microsoft in 1997.³² However, even though the District Court and the Court of Appeals found many of the Microsoft actions to be illegal,³³ the final settlement between the U.S. Department of Justice and Microsoft in November 2002 provided little relief to Netscape.³⁴ Netscape filed a separate civil antitrust suit against Microsoft in January 2002.³⁵ That suit was settled in May 2003 for \$750 million.³⁶ Today, Microsoft has 86 percent of the browser market and Netscape has only 9 percent.³⁷

²⁹ Cusumano and Yoffie, p. 144.

³⁰ [CITE source]

³¹ Figures from <http://www.cen.uiuc.edu/bstats/latest.html>.

³² [Cite DOJ Complaint.]

³³ [Cite Findings of Fact, Appeal Court Opinion]

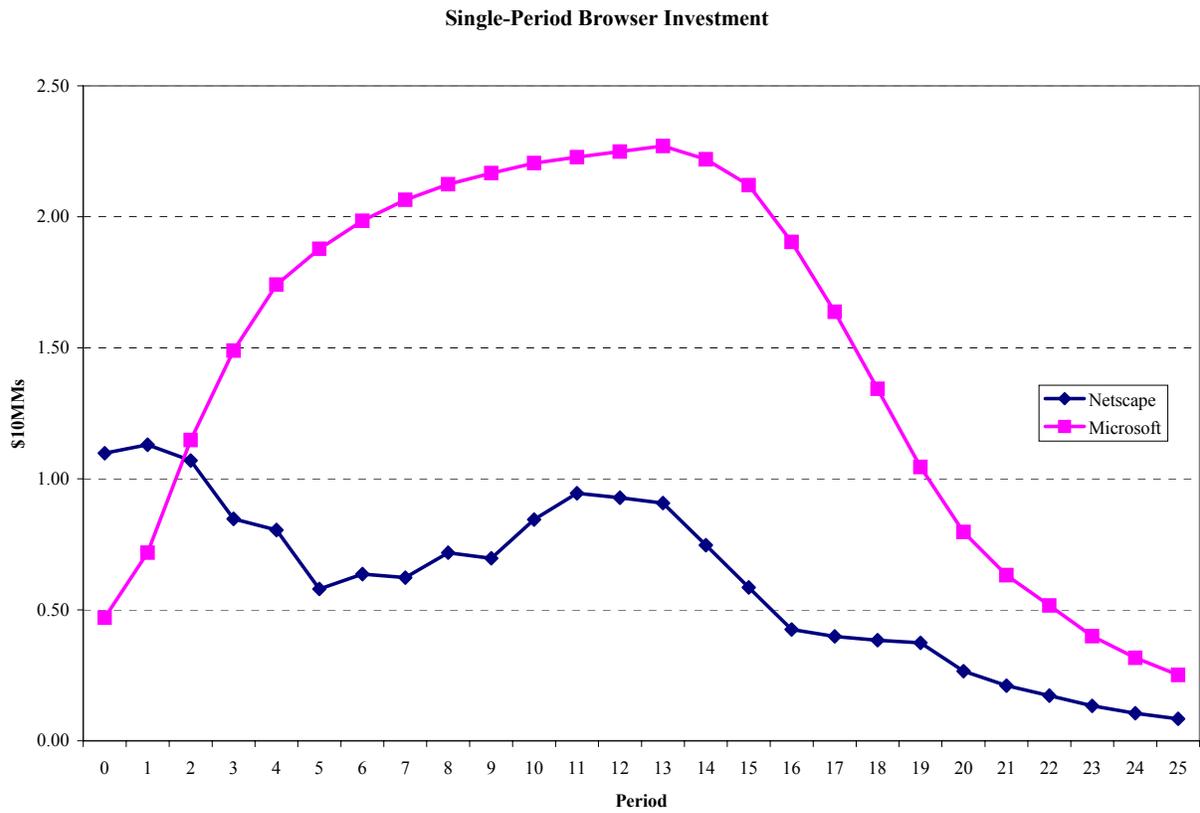
³⁴ See *Final Judgment*, viewed at http://www.usdoj.gov/atr/cases/ms_index.htm on March 4, 2004.

³⁵ See *Complaint in Netscape v. Microsoft*, viewed at <http://legal.web.aol.com/decisions/dlother/Netscape%20v.%20Microsoft.pdf> on March 4, 2004.

³⁶ See <http://money.cnn.com/2003/05/29/technology/microsoft/>.

³⁷ Figures from <http://www.cen.uiuc.edu/bstats/latest.html>.

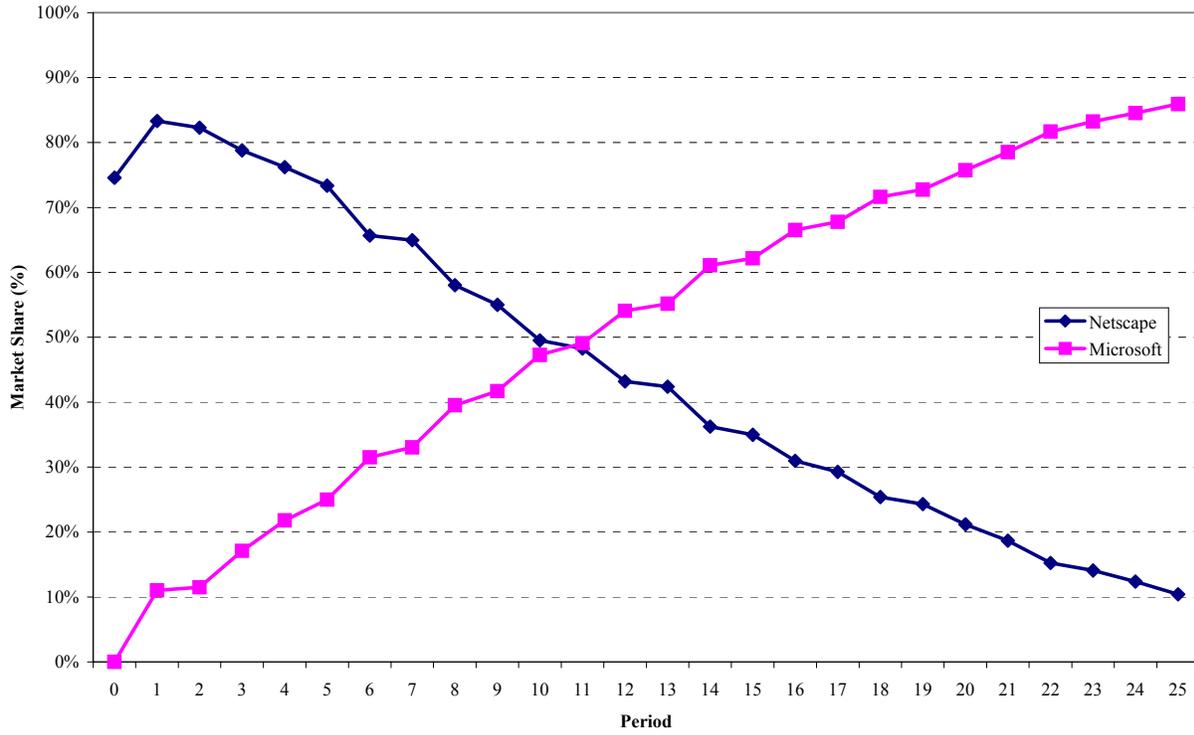
Figure 2



Source: Various, authors' estimates. This figure should not be construed as having been based in large part on actual data. Do not cite.

Figure 3

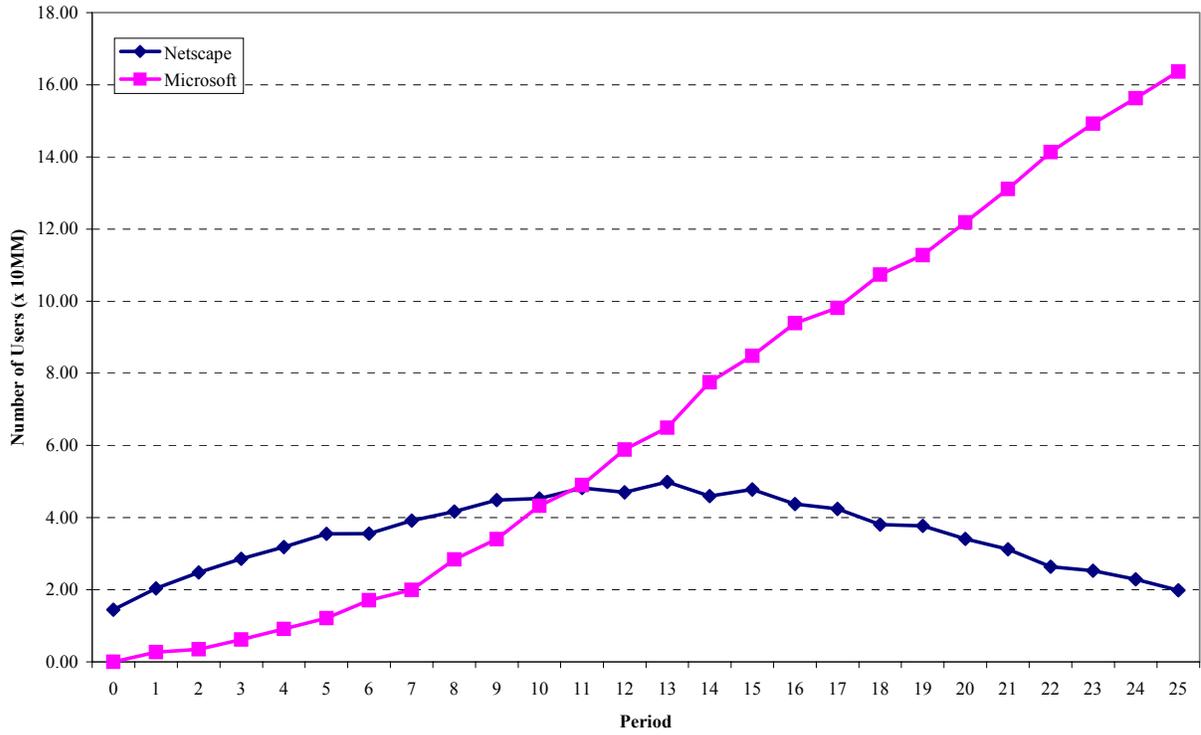
Base Market Share



Source: UIUC data (<http://www.cen.uiuc.edu/bstats/latest.html>)

Figure 4

Installed Base



Source: Calculated from number of Internet users and browser market share.

5. Motivating the Econometric Modeling

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.³⁸ 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 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 players in the market (e.g., Opera, Mozilla), these players tend to be minor. However, these minor players constituted a fringe that would have limited the ability of Netscape and Microsoft to raise prices substantially.

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 rather 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 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

³⁸ We were engaged by Netscape to calculate damages in *Netscape v. Microsoft*.

that this assumption is also consistent with an MPE model of the consumer, making the consumer's policy decision a function only of current state. However, we will not exploit this interpretation.) 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 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.”³⁹

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. However, that

³⁹ Findings of Fact, ¶143.

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 Investment.** 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.⁴⁰ Quality is determined to a large extent on the amount of investment placed in upgrading each incremental version of the browser.

6) **Profit Equation – Inclusion of Derivative Revenues per Browser.** Browser revenues were always a relatively small portion of the 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.⁴¹ Microsoft could use Internet Explorer to solidify its Windows monopoly and establish its own Internet portal. Thus Internet browsers can perhaps be viewed as a software with no direct revenues but with substantial potential 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).

⁴⁰For example, XXXX

⁴¹ 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.”

6. The Model

The above discussion motivates a model in which (1) market share for browsers is a function of investment and previous market share, reflecting quality and network effects; (2) profits are a function of fixed revenues per user (with different values for Microsoft and Netscape) multiplied by market share for browsers multiplied by market size, less investment, and (3) each firm's choice variable is investment. The following paragraphs give the details of the model specification. Table 2 summarizes the variables that appear.

Table 2. List of Variables		
Symbol	Type	Description
$j = 0, 1$		firms
$t = 0, 1, 2, \dots$		time periods (quarters beginning Q1 1996)
$B_j(t)$	S	installed base of browsers for firm j , $j=0, 1$
$I_j(t)$	S	accumulated browser investment for firm j , $j=0, 1$
$q_j(t)$	—	browser quality for firm j , $j=0, 1$
$i_j(t)$	PO	current investment in product development for firm j , $j=0, 1$
$N(t)$	XO	number of installed browsers
$n(t)$	XO	number of consumers choosing a new browser
$s_j(t)$	XO	share of consumers choosing firm j 's browser, $j=0, 1$
$r_j(t)$	XO	revenue per browser used
$a_k(t)$	XO	intensity of Microsoft's bad acts, $k=0, 1, 2$; 0=IAP, 1=OEM, 2=Tying
$\varepsilon(t)$	XU	preference-side, market-level demand shifter

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

This model has four state variables: installed base of browsers for each firm $B_j(t)$ and the accumulated investment for each firm $I_j(t)$. Accumulated investment is determined through an accounting relationship by the competing investments in development by the two firms,

and installed base is determined by consumer choice. Quality, which appears in the demand function, is determined through an logistic transformation of accumulated investment. The policy variables determined within the model are the investments in development, $i_0(t)$ and $i_1(t)$.

The model has observed exogenous variables, $N(t)$, $r_j(t)$, and $a_k(t)$, and one unobserved variable $\varepsilon(t)$, that are known to the firms when they make their investment policy decisions in a period. Past the period of observation, all exogenous variables are assumed to be i.i.d. with distributions that reflect their characteristics during the period of observation. We assume that they are always treated by the firms as i.i.d. random variables over time, so that firms form expectations with respect to these variables when valuing the future.

The exogenous variable $N(t)$ represents the number of Internet users in each period. Since we assume that each Internet user utilizes exactly one browser to access the Internet, $N(t)$ also represents the total installed base of browsers. The variable $r_j(t)$ represents the indirect revenue that a firm derives from each browser user in a given period. Indirect revenue per user need not be the same for both firms.

The bad act intensity variables, $a_k(t)$, represent the percentage of consumers that were affected by each act in a given period. The intensity of the IAP act in a given period was calculated as the number of new subscribers to affected IAPs (i.e. IAPs that were involved in restrictive agreements with Microsoft) divided by the total number of new IAP subscribers. The intensity of the OEM act in a given period was calculated as the number of computers shipped by affected OEMs divided by the total number of OEM shipments, and the intensity of the tying act was calculated as the number of new Windows licenses sold divided by the total number of new computers.

Consumers' decisions on acquisition and retirement of browsers are exogenous to the browser market. The number of consumers choosing a browser in each period is

$$(19) \quad n(t) = N(t) - (1-d_b)N(t-1),$$

where d_b denotes the rate at which existing browser users reconsider their browser selection.

The share of new users choosing browser 0, $s_0(t)$, is determined by the logistic

$$(20) \quad s_0(t) = 1 / (1 + \exp(-x)),$$

where

$$x = \exp[\theta_0 + \theta_1(B_0(t-1) - B_1(t-1)) + (\theta_2 q_0(t-1) - \theta_2 q_1(t-1)) - \theta_3 a_0(t) - \theta_4 a_1(t) - \theta_5 a_2(t) + \varepsilon(t)].$$

The share of new users choosing browser 1, $s_1(t)$, is then

$$s_1(t) = 1 - s_0(t).$$

The installed base for each firm satisfies the equation of motion

$$(21) \quad B_j(t) = (1 - d_B)B_j(t-1) + n(t)s_j(t).$$

The second state variable, accumulated investment, satisfies the equation of motion

$$(22) \quad I_j(t) = (1 - d_I)I_j(t-1) + i_j(t),$$

The profit of each firm in period t is

$$(23) \quad \pi_j = r_j(t)B_j(t) - i_j(t) - \lambda i_j(t)^2.$$

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 dollar investment in browser development, and the third reflects the frictional costs associated with crash development.

The valuation functions of the firms are the expected (with respect to r, a, ε) present value (with an observed discount factor β) of the stream of period profits. These expected present

values at time $t-1$ of profits from time t forward are functions of the state variables $B_j(t-1)$, and $I_j(t-1)$.

7. Data

Estimation of the model requires the following data:

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 previous section. For the IAP act, quarterly data on IAP subscriptions were obtained from

AdKnowledge. For the OEM act, quarterly data on OEM shipments were obtained from IDC. For the Tying act, data on the installed base of computers were obtained from IDC, and data on the number of Windows licenses sold were based on authors' estimates.

6) **Depreciation Rates** - The depreciation rate of installed base is assumed to be 0.30, based on authors' estimates. The depreciation of accumulated investment is assumed to be 0.25 through December 2000 and 0.10 thereafter.

8. Model Results

The initial least squares estimate of the demand function generated the results in Table 2. Most of the coefficients in this estimation are not statistically significant. However, the coefficients for base and quality are highly statistically significant, and of the expected sign: In other words, a greater base is likely to lead to a greater market share, *ceteris paribus*. The quality variable also has the expected sign that an increase in the quality of a firm's product is predicted to lead to an expected increase in market share for that firm.

In Figure 5 below, we have plotted the predicted model market shares versus the data used for market share (these are the installed base market shares, not the incremental market shares). As one can see, the MPE model using the parameters estimated above does not precisely track actual market shares. At this time, we are in the process of refining our coefficient estimates by fitting the predictions of the model to the actual data.

Of the "bad act" variables, the tying and OEM acts do not have the expected positive sign (recall that the coefficient for the "bad acts" enters into Microsoft's demand function: thus, a positive coefficient will in expectation *increase* Microsoft's market share.) However, the IAP bad acts do have the expected sign, albeit hardly significant. These (highly) preliminary results do not show a statistically significant effect of bad acts on the relative shares of Netscape and Microsoft. Our "bad acts" variables are currently crudely defined and may not be capturing the true impact of a properly measured "bad acts" variable. Further work to more precisely quantify the "bad acts" is needed.

Table 3

<i>Regression Statistics</i>				
Multiple R	0.980401			
R Square	0.961187			
Adjusted R Square	0.950973			
Standard Error	0.276896			
Observations	25			

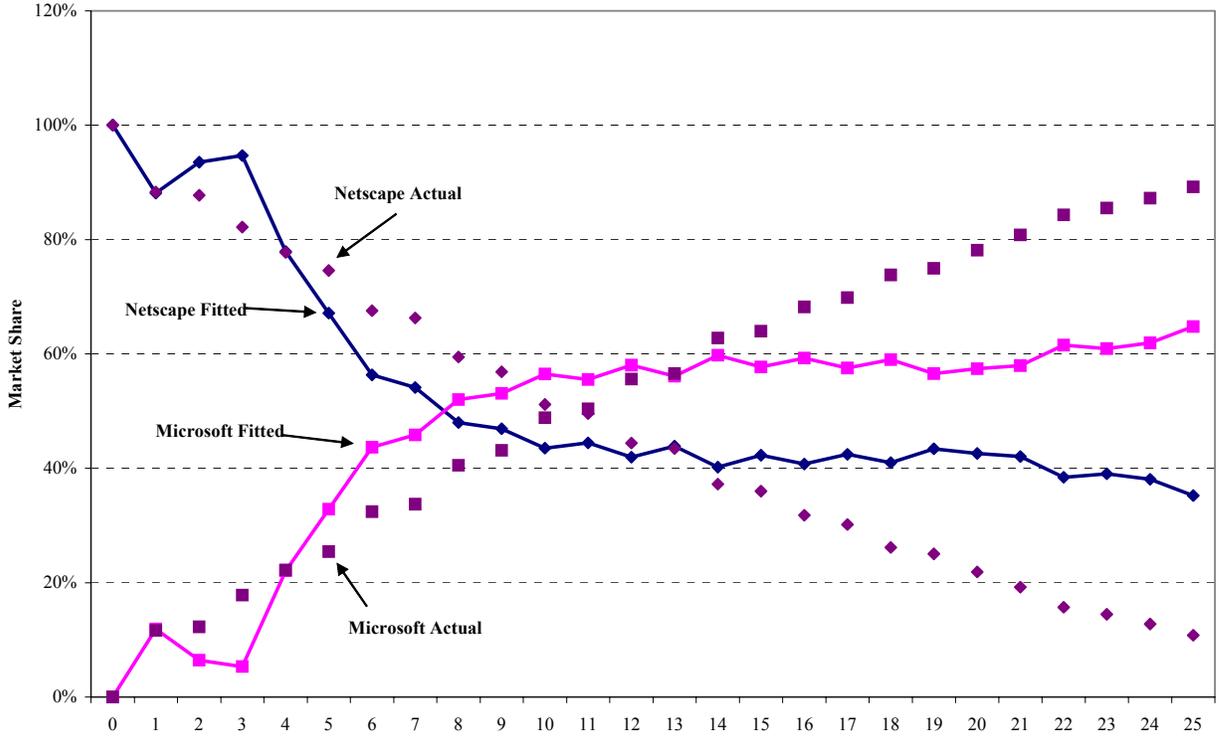
ANOVA				
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>
Regression	5	36.07586	7.215173	94.10512
Residual	19	1.456757	0.076671	
Total	24	37.53262		

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>
Intercept	-0.57014	0.813648	-0.70072	0.491974
Base	0.125579	0.027399	4.583383	0.000203
Quality	5.204694	0.673127	7.732116	2.77E-07
IAP Act	0.378175	0.24061	1.571732	0.132517
OEM Act	-0.16884	0.221222	-0.76322	0.454708
Tying Act	-0.66097	1.368629	-0.48295	0.634652

Note: Preliminary result. Do not cite. This table should not be construed as having been based in large part on actual data.

Figure 5

Installed Base Market Share, Actual Data vs. Fitted



Note: Preliminary result. Do not cite.

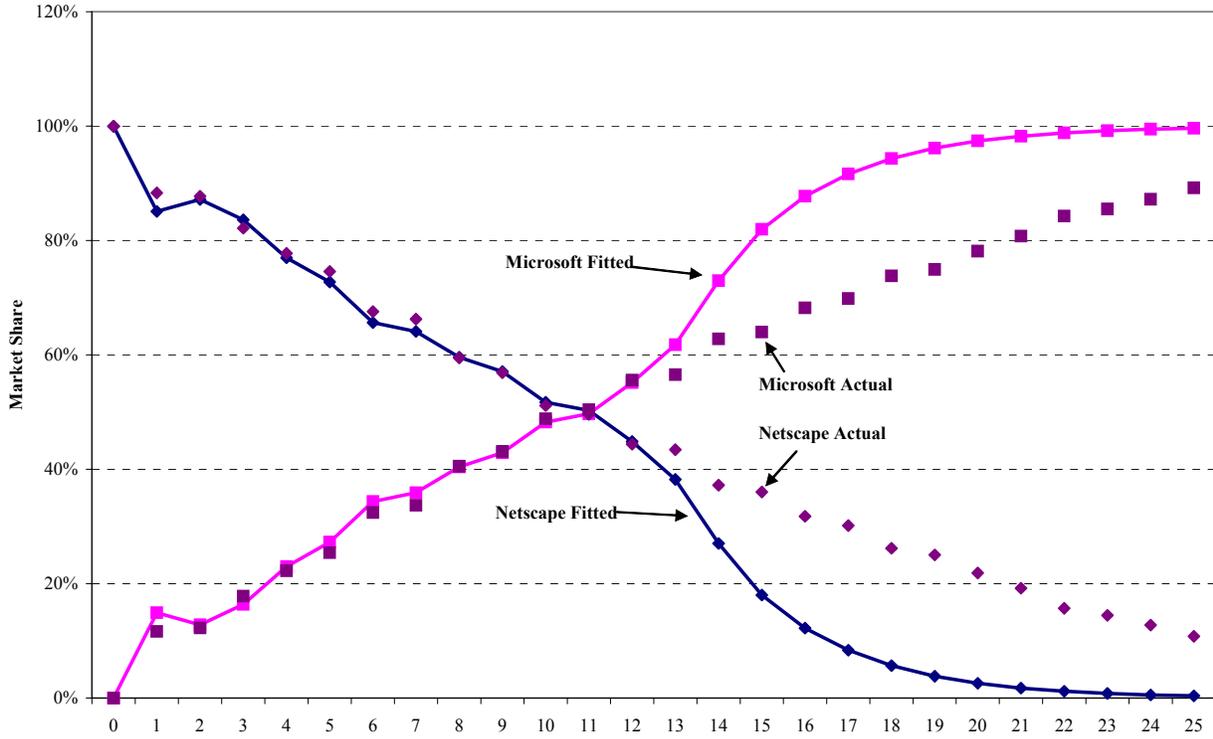
For purposes of illustrating the use of this model for predicting market share in the absence of bad acts, we have turned off the IAP variable, leaving the OEM and tying variables intact (as well as the rest of the variables). Figure 6 below compares the predicted market shares with the IAP acts turned off to the predicted market shares with the IAP acts turned on. As one can see, the difference in the predicted market shares is not large. These results suggest that the IAP bad acts had relatively little impact on Netscape's and Microsoft's relative market shares, subject to all the caveats thus raised.

Though at this time we do not observe any substantial change in market share resulting from Microsoft bad acts (subject to the caveats raised), the model does not preclude the finding of a "market tipping." To demonstrate this result, we adjust the demand function parameters and again turn on and off the IAP "bad act". Figure 7 below compares actual installed base market shares to market shares fitted using a set of alternative parameters.

Had these parameters been what we had found from the estimation, we would have found that the effect of the IAP "bad acts" was to tilt the market to Microsoft's benefit. As seen in Figure 8 below, without the IAP "bad act" in this scenario, Netscape would have been able to rely on the network effects deriving from its dominant installed base and would have been able to maintain its dominant market share through the time period observed. Thus with a somewhat different set of parameters, we could find a substantial effect of Microsoft's bad acts on the relative market shares of Netscape and Microsoft.

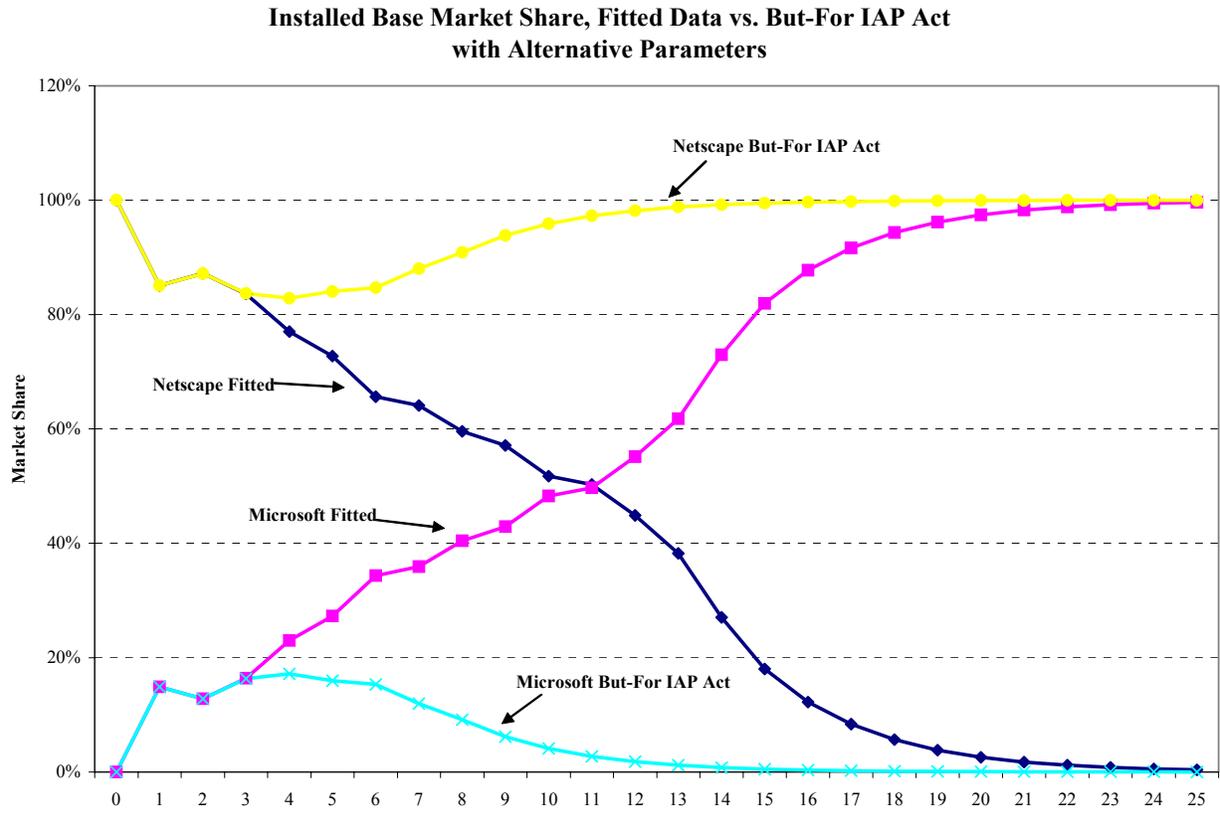
Figure 6

Installed Base Market Share, Actual Data vs. Fitted with Alternative Parameters



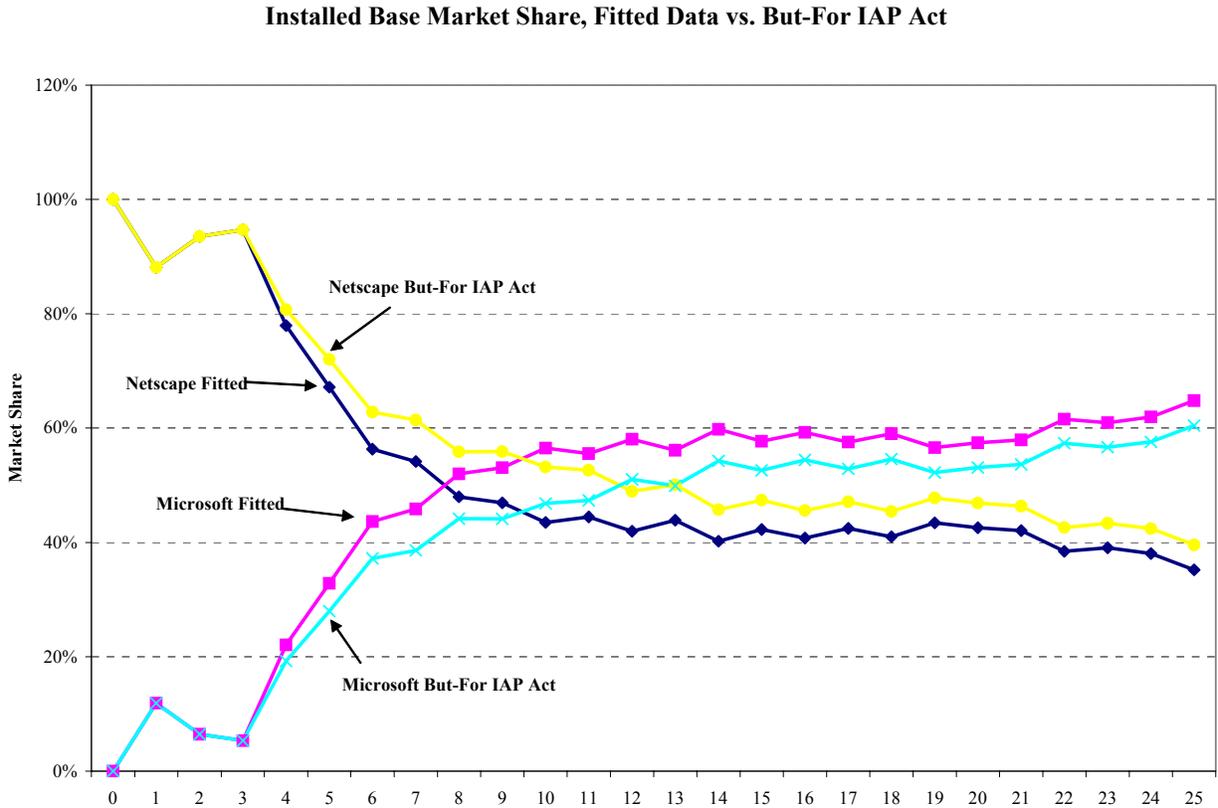
Note: Preliminary result. Do not cite.

Figure 7



Note: Preliminary result. Do not cite.

Figure 8



Note: Preliminary result. Do not cite.

9. Conclusions

The results of this paper show that it is possible, and even feasible, to use the tools of Markov-perfect equilibrium models combined with data in order to arrive at a unique Markov perfect equilibrium. Using an MPE model that was motivated by observed characteristics of the browser market, and using data relating to Netscape and Microsoft, we successfully estimated an MPE model and used it to provide predictions of alternate hypotheses. In

particular, we used the model to predict what Netscape's and Microsoft's market shares would have been "but for" the IAP acts that Microsoft committed. Despite the finding by the Appeals Court in U.S. v. Microsoft of illegal acts, our model predicts that the IAP acts had relatively little impact on Netscape's but-for market share. However, this is a highly preliminary result and subject to many caveats which are discussed in the paper; work on more precisely quantifying the illegal acts and the other variables is ongoing. We also show that with a different set of demand function parameters, absent the IAP "bad acts", the model predicts that Netscape would have been able to maintain its initial dominant position. Thus the full span of possible outcomes is observed in this model, and results will depend critically on the quality of the data and the specific magnitudes of the parameter estimates.

References

- Alexandrov, A.D., (1939) "Almost Everywhere Existence of Second Differentials of Convex Functions" (Russian) *Leningrad 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.
- Bajari, Patrick; Lanier Benkard; Jonathan Levin (2003) "Estimating Dynamic Models of Imperfect Competition," Stanford University Working Paper.
- Balder, Erik (2001) "On Equilibria for Discontinuous Games: Continuous Nash Approximation Schemes" University of Utrecht Working Paper.
- Baumol, William; J.C. Panzar; R.D. Willig (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 Perturbations of Quadratic Maps as Markov Processes," Cornell University Working Paper.
- Bhattacharya, Rabi; Mukul 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*.

- Brander, James; Anming Zhang (1993) "Dynamic Oligopoly Behavior in the Airline Industry," *International Journal of Industrial Organization*, 11, 407-433.
- 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; 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.
- Chen, Yongmin; Robert Rosenthal (1996) "Dynamic Duopolies with Slowly Changing Customers Loyalties," *International Journal of Industrial Economics*, 14, 269-296.
- 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.
- 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.
- 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.
- Erickson, Richard; Ariel Pakes (1995) "Markov Perfect Industry Dynamics: A Framework for Empirical Work," *The Review of Economic Studies*, 62, 53-82.
- Farrell, Joseph; Carl Shapiro (1988) "Dynamic Competition with Switching Costs," *The RAND Journal of Economics*, 19, 123-137.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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 Perfect Equilibria in Oligopolistic Price Wars," University of New South Wales 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 Vetterling (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 Chains," 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.
- 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.