RAND Journal of Economics
Vol. 46, No. 4, Winter 2015
pp. 671–708

Oblivious equilibrium for concentrated industries

C. Lanier Benkard*

Przemyslaw Jeziorski**

and

Gabriel Y. Weintraub***

This article explores the application of oblivious equilibrium (OE) to highly concentrated markets. We define a natural extended notion of OE, called partially oblivious equilibrium (POE), that allows for there to be a set of strategically important firms (the "dominant" firms), whose firm states are always monitored by every other firm in the market. We perform computational experiments that explore the characteristics of POE, OE, and Markov perfect equilibrium (MPE), and find that POE generally performs well in highly concentrated markets. We also derive error bounds for evaluating the performance of POE for cases where MPE cannot be computed.

1. Introduction

■ There has been much recent work in industrial organization (IO) on empirical applications of dynamic oligopoly models (e.g., Benkard, 2004; Collard-Wexler, 2013 [henceforth CW]; Fowlie, Reguant, and Ryan, in press; Goettler and Gordon, 2011; Jeziorski, 2014; Ryan, 2012; Sweeting, 2013). The primary benefit of using a dynamic model is that such models allow us to study the effects of a government policy on technological progress and on industry structure, that is, the set of firms in the market, their technologies, and the products that they choose to offer. In many applications, such as merger analysis or environmental and energy regulation, these long-run effects can dwarf the short-run static effects, and thus analyzing them is of first-order importance. The cost of doing a dynamic analysis is that the models are inherently more complex. Indeed, this recent work was only made possible by the advent of new methods for estimating dynamic oligopoly models (Bajari, Benkard, and Levin, 2007; Pakes, Ostrovsky, and

* Stanford University and NBER; lanierb@stanford.edu.
** University of California at Berkeley; przemekj@haas.berkeley.edu.
*** Columbia University; gweintraub@columbia.edu.
We would like to thank Ben Van Roy for numerous fruitful conversations and valuable insights provided at the beginning stages of this project. We are grateful to Allan Collard-Wexler for generously sharing his data and for helpful input. This article has also benefitted from conversations with Ariel Pakes, Ali Yurukoglu, and seminar participants at various conferences and institutions.

Berry, 2007; Aguirregabiria and Mira, 2007; Pesendorfer and Schmidt-Dengler, 2003) that do not require the econometrician to compute *Markov perfect equilibria* (henceforth, MPE) of the underlying game being studied. This is important because MPE computation is subject to the curse of dimensionality.

Despite this recent progress, there remain substantial hurdles in the empirical application of dynamic oligopoly models. Even when it is possible to use these new methods to estimate the model parameters without computing an equilibrium, equilibrium computation is still required to analyze the effects of a counterfactual policy or environmental change. The result is that in applications many modelling details are still heavily dictated by computational concerns, typically at some expense to the credibility of the economic analysis.

In a recent article, Weintraub, Benkard, and Van Roy (2008) propose a method for analyzing Ericson and Pakes (1995; hereafter, EP) style dynamic models of imperfect competition that is intended to address some of these concerns. In that article, they defined a notion of equilibrium, *oblivious equilibrium* (henceforth, OE), in which each firm is assumed to make decisions based only on its own state and knowledge of the long-run industry state, but where firms ignore current information about competitors' states. The great advantage of OE is that their computation time is not systematically related to the overall number of firms in the industry, and thus they are much easier to program and compute than MPE, allowing researchers to analyze richer empirical models. Although OE have some very appealing properties, they also have some weaknesses. One that has often been raised is the appropriateness of the OE assumptions for highly concentrated markets. Because many of the most interesting policy questions in IO focus on highly concentrated markets, this is an important flaw.

In this article, we introduce an extension to OE designed to address this important case. We define an extended notion of oblivious equilibrium that we call *partially oblivious equilibrium* (POE) that allows for there to be a set of strategically important firms, that we call "dominant" firms, whose firm states are always monitored by every other firm in the market. Thus, for example, if there are two dominant firms, then in POE, each firm's strategy will be a function of its own state and also the states of both dominant firms. Such strategies allow for richer strategic interactions than do oblivious strategies, which depend only on a firm's own state, and our hope is that POE will provide a better model of more concentrated markets than OE does. Moreover, if entry and exit are modelled identically for dominant and (nondominant) "fringe" firms, then the POE model nests both OE and MPE, where OE is a POE model with no dominant firms, MPE is a POE model with all dominant firms, and where all other POE models represent intermediate cases.

The extension primarily trades off richer strategic interactions against increased computation time and memory requirements due to a larger state space. The state space for OE is of the order of a one firm problem; the state space for POE with one dominant firm is of the order of a two-firm problem; the state space for POE with two dominant firms is of the order of a three-firm problem, and so forth. In all cases, for a fixed number of dominant firms, compute time and memory requirements for POE (and OE) are not systematically related to the number of firms in the overall industry. As a consequence, although POE take substantially more time to compute than OE, it is still dramatically less than MPE.

POE can also be motivated as a behavioral model in its own right. It seems unrealistic that firms would be able to follow strategies that are functions of more than a handful of competitor firms. Strategies that are functions of all firms are high dimensional and highly complex, and it is hard to imagine how firms would obtain enough information to compute them and execute them. If firms follow strategies that are only functions of a few firms, it seems likely that their strategies would include only themselves and the leading firms, ignoring less important fringe firms. Additionally, if information is costly, then firms would only pay to learn the information that is most relevant to them. In the models we consider, the most valuable information would typically be the states of the leading firms. Finally, in many markets there are "leader" firms whose actions are followed more closely than the other firms in the industry. Indeed, there are older literatures (Von Stackelberg, 1934; Kydland, 1979) that model the dominant firm in an industry as moving first.

We perform a large number of computational experiments to evaluate how POE performs in practice and to compare POE with OE and MPE. We find that in markets with medium to high concentration, there is little difference between MPE and OE or POE with one, two, or three dominant firms. POE is generally closer to MPE than OE is, but the differences are small. In these cases, OE is clearly the best tool as it is the simplest to compute and would allow researchers to use the most robust economic model, while providing results that are the same as the other equilibrium concepts.

In very highly concentrated markets, markets with C2 $>$ 0.90, corresponding to approximately the top 1% of manufacturing industries in the United States, we find that OE sometimes does not work well. (It is not typically close to MPE, in ways that seem undesirable.) We explore the performance of POE in these cases and find that as long as turnover among leading firms is not too unrealistically high, within or even exceeding the range observed in real-world markets (see Sutton 2007), then POE with one or two dominant firms works well even in the most concentrated markets.

Another finding from our computational experiments is that the information structure of the model can be quite important in determining firm behavior. In the POE model, dominant firms' states are tracked by all firms, whereas individual (nondominant) fringe firms' states are not tracked. Because dominant firms' states are tracked, their actions have a direct impact on other firms' behavior, and they can more easily deter entry or investment by investing and becoming large. As a result, we find that in POE, dominant firms generally invest more than fringe firms, and this leads to them on average being larger and remaining large for longer. In essence, being labeled as "dominant" causes the firm to in fact be dominant most of the time. This asymmetry in behavior sets the POE model apart from the OE and MPE models, in which it is typically assumed that all firms would behave the same way at the same state of the world. These results also suggest using caution when simplifying firms' strategies in equilibrium calculations more generally. Although POE strategies are somewhat natural, and lead to behavior that is not unrealistic, arbitrary restrictions of firms' strategies to facilitate computation could lead to unintended and unnatural firm behavior.

We also apply POE to the empirical model of CW. We find that for the basic CW model OE is fairly close to MPE, though it is not exactly the same. Adding dominant firms makes many statistics closer to MPE, but we find that it takes four dominant firms to obtain results nearly the same as MPE. We also explore what happens in the CW model as we make the discretization of the state space finer. Coarse discretization of the state space is a common tool used in the empirical literature to make dynamic models more tractable (e.g., Benkard, 2004; Gowrisankaran and Town, 1997). The CW model is quite complex and thus, in order to compute MPE, the model discretizes individual firm size states to just three points. Because OE and POE are computationally light, we are able to solve the CW model on much finer state space grids. For our finest grid, the model has 66 billion state points.

We find that there are some differences in the results as we make the size grid finer. The main differences are that on the three point grid there are equal numbers of small and medium firms, but on the finer grids there are nearly twice as many medium firms as small ones. Furthermore, transition costs get much larger as firms are now changing size more often and thus paying more transition costs. The intuition for this finding is that the coarse grid has the same effect as an adjustment cost, so that on a coarse grid firms are unable to make fine adjustments to their size. When the grid is made finer, firms adjust size more often and there are more medium-sized firms and fewer small firms.

We believe that these results demonstrate an important trade-off facing empirical researchers in this area. Using a computationally simple equilibrium concept such as POE allows the researcher to use a richer economic model. Sometimes the economics of the model may be changed less by using OE or POE in place of MPE than they would by simplifying the model to facilitate computation of MPE. In other cases, it may be worth computing MPE even with the additional modelling restrictions that requires.

A weakness of POE is that it is not as theoretically clean a concept as OE, in two main respects. First, in an abstraction from reality, our implementation of POE does not allow for firms to transition from fringe to dominant or vice versa. In our model, firms' strategies are restricted to be functions only of their own state and the states of the dominant firms. If we were to allow transitions between fringe and dominant firms then, in the model, the set of variables that firms condition their strategies on would be endogenous. That is, not just the values of the state variables, but also which state variables are, in effect, known to firms, would be endogenous. In forming beliefs about the future, firms would then need to forecast one or more firms transitioning into dominance, in addition to the future values of the dominant firms' states. The probabilities of transitioning into dominance would depend not only on the dominant firms' states but also on all of the fringe firms' states. Thus, if firms could transition from fringe to dominant and vice versa, the transition process for dominant firms' states would be non-Markov, because its transition probabilities depend on something in addition to its own current value, namely, the value of the fringe firms' states. This non-Markov behavior would make the model substantially more complex and add modelling and computation costs.

At the same time, it is not clear to us that the benefits of such an extension would be worth the costs. Our goal is to capture strategic interactions between large firms in as simple and realistic a way as possible. Our subjective judgment is that ruling out dominant/fringe transitions would not be unrealistic for many empirical applications, which typically consider markets and time horizons over which the set of dominant firms is quite stable. In fact, in the past literature it has been common to model only the dominant firms in an industry, and then completely ignore some or all of the fringe firms, as if they did not exist. There is little hard empirical evidence on this issue, but the evidence that there is (Sutton, 2007) suggests that turnover among leading firms is typically quite low, happening once every few decades in the median industry. We conclude from this that the main consequence of ruling out transitions is that the model is best used for medium-term policy analysis, where medium term would typically measure at most a few decades in length. That said, we do also believe that the longer run dynamics of turnover among leading firms in a market is an interesting economic problem, and is a potentially interesting direction for future research.

The assumption also raises the issue of how researchers would identify which firms to label as "dominant." In most cases, these would simply be the top firms in the market in terms of market share, and the only issue would be how many dominant firms to model. To help with this question, we derive computationally tractable error bounds that bound the error that a firm would be making by playing a POE strategy rather than unilaterally deviating to a fully informed best response strategy (see also Weintraub Weintraub, Benkard, and Van Roy, 2010). The bounds can be used as a guide to empirical researchers in determining how well the POE model is doing at describing behavior in any particular application, and also in determining how many dominant firms are required, if any. We note that our bounds can be computed in large-scale applications for which it would be impossible to compute a fully informed best response, let alone MPE. In our computational experiments we find the bounds computed to be quite loose, but we believe that the bounds may be more useful in other empirical applications, which typically have more firms than our examples do. Moreover, we show that tighter bounds can be derived using application-specific details.

There is also a second technical issue that is related to the treatment of the dominant firms. In POE, firms' strategies depend on the values of the dominant firms' states, but firms have only incomplete knowledge of the current overall industry state because they only track the dominant firm states. Thus, the (unknown) current values of the fringe firm states are partially determined by past values of the dominant firms' states, which implies that past values of the dominant firms' states might potentially be useful in predicting the current fringe firm states. Furthermore, in principle an infinite history might be useful. In our implementation of POE, firms' strategies are assumed to be functions of finite, typically short, histories. We think that this restriction represents a reasonable compromise, and we have also explored altering the length of the history and found that it does not have much impact on the results (at least for the three models we utilize below), but we do recognize that the restriction is somewhat arbitrary.

□ **Related literature.** We are not the only researchers to recognize the methodological and computational hurdles in the empirical application of dynamic oligopoly models, and there have been several other solutions proposed to this problem. For example, Pakes and McGuire (2001) propose solving the game with a stochastic approximation algorithm. Their method reduces the size of the state space by solving the model only on a recurrent class of states and reduces the computation time at each state through simulation. Doraszelski and Judd (2012) suggest casting the game in continuous time, which similarly has the advantage of greatly reducing the computation cost at each state point by reducing the number of future state points that communicate with each current state (see also Jeziorski, 2015, for an application of this technique). Farias, Saure, and Weintraub (2012) introduce a method to approximate MPE based on approximate dynamic programming, in which the value function is approximated with a linear combination of basis functions. They show that a rich yet tractable set of basis functions works well for important classes of models. Santos (2012) introduces a state aggregation technique based on quantiles of the industry state distribution to try to break the curse of dimensionality.

Additionally, most empirical applications necessarily make some modelling simplifications that help reduce computation time. These simplifications range widely and include reducing the number of firms in the market to a workable number and/or coarsely discretizing the state space (Benkard, 2004; Gowrisankaran and Town, 1997; Collard-Wexler, 2013), or using functional approximations (such as polynomials or other functions) to the value function (Sweeting, 2013; Fowlie, Reguant, and Ryan, in press). Many articles also combine several of these approaches at once. These methods can be applied to greatly reduce computation time and memory requirements for exact MPE in a variety of modelling contexts.

Compared with these articles, our methods take a different approach that is best described as restricting strategies of firms (or alternatively, limiting the information available to firms), which reduces the effective state space of the model. The main cost to our approach relative to those above is that OE and POE are not as strategically rich as MPE. Thus, if full information MPE is desired and it is computationally feasible to compute it, researchers are likely better served by one of the other methods listed above. On the other hand, because of the strategy (or informational) restrictions, computation time for OE and POE is not systematically related to the overall number of firms in the industry, which facilitates equilibrium computation in rich economic models with extremely large state spaces, such as the example above that has 66 billion state points. It would not be possible to compute exact MPE in such models with present-day computing technology without simplifying the model. Moreover, we think that it is reasonably common to encounter situations like this one in empirical work on real-world markets.

Another related article by Fershtman and Pakes (2012) considers dynamic oligopoly models with asymmetric information. Although their economic model is fundamentally different from ours due to asymmetric information, their article tackles many similar issues to ours because OE and POE limit the information sets of firms, and therefore our equilibrium concepts have asymmetric information, even if our underlying economic model does not. Indeed, we believe that the simulated versions (see below) of OE, OE with aggregate shocks (Weintraub, Benkard, and Van Roy, 2010), and POE can all be viewed as special cases of their *restricted experience-based equilibrium*.

The most closely related article to ours is Ifrach and Weintraub (2014) who also consider a dynamic oligopoly market with few dominant firms and many fringe firms. Their approach is not built upon OE and is more complex than ours in two ways: (i) firms' strategies are functions of dominant firms' states plus summary statistics of the fringe firms' states, and (ii) in their model, the set of dominant firms arises endogenously in equilibrium. That is, they allow for firms to transition from fringe to dominant and vice versa. These generalizations come at a complexity cost in terms of modelling and computation. In their approach, the modeller needs to decide which fringe firms' statistics to track and how to model tier transitions. Additionally, the transition kernel of the state variables in the model is not Markov, and simulation is required to estimate it. In contrast, our approach is a natural extension of OE that is similar to MPE in its treatment of dominant firms. In our model, the primary modelling choice is the number of dominant firms. In addition, the dynamics in our model are Markov, making it cleaner and easier to implement and code. Overall, we think that a reasonable strategy for practitioners should be to study whether the simplest solution concept, namely OE, is enough for their applications, and if not, we suggest to add dominant firms and use POE. Ifrach and Weintraub (2014) provide an alternative if a more general model is required beyond POE.

The next section introduces the model. Section 3 precisely defines POE. Section 4 introduces the error bounds. Section 5 provides computational experiments evaluating POE. Section 6 computes POE for an empirical example. Finally, Section 7 concludes the article.

2. A dynamic model of imperfect competition

■ In this section, we formulate a model of an industry in which firms compete in a market over discrete time periods with an infinite horizon, and where firms can enter, exit, and invest. The model is similar to the EP model except that we allow for more general entry, exit, and investment processes, primarily to make these processes more realistic for the purposes of empirical work. The model closely follows that of Weintraub, Benkard, and Van Roy (2008).

□ **Model and notation.** We index time periods with nonnegative integers $t \in \{0, 1, 2, \ldots\}$. Each firm that enters the industry is assigned a unique positive integer-valued index, i. The set of indices of incumbent firms at time t is denoted by S_t, and the number of incumbent firms is denoted n_t.

As in the EP model, firm heterogeneity is reflected through firm states, x_{it}, and we will assume that the state space is discrete. Without loss of generality, we denote the values of the firm state using the nonnegative integers: $x_{it} \in \{0, 1, 2, \ldots\}$. To fix an interpretation of the model, we will refer to a firm's state as its quality level, and we would suggest that readers have in mind an example where single product firms compete in a logit demand model (and this will also be one of our examples below). However, firm states might more generally reflect productivity, capacity, the size of a firm's consumer network, or any other aspect of the firm that affects its profits.

We define the *industry state*, s_t, to be a vector that lists, for each possible quality level $x \in \{0, 1, 2, \ldots\}$, the number of incumbent firms at that quality level in period t. The state space $\overline{\mathcal{S}} = \{s \in \mathbb{N}^\infty | \sum_{x=0}^{\infty} s(x) < \infty\}$ is the set of all possible industry states.[1] For an incumbent firm i, we define $s_{-i,t}$ to be the state of the *competitors* of firm i; $s_{-i,t}$ is simply equal to s_t with firm i subtracted out. That is, $s_{-i,t}(x) = s_t(x) - 1$ if $x_{it} = x$, and $s_{-i,t}(x) = s_t(x)$, otherwise.

Again following EP, in each period, each incumbent firm earns profits on a spot market. A firm's single period expected profit is denoted by $\pi(x_{it}, s_{-i,t})$, and depends on its quality level x_{it} and its competitors' state, $s_{-i,t}$.

The model also allows for entry and exit. In each period, each incumbent firm observes a positive real-valued sell-off value ϕ_{it} that is private information to the firm. If the sell-off value exceeds the value of continuing in the industry then the firm may choose to exit, in which case it earns the sell-off value and then ceases operations permanently.

[1] Though in principle there are a countable number of industry states, for some of the technical details we will also consider an extended state space $\mathcal{S} = \{s \in \Re_+^\infty | \sum_{x=0}^{\infty} s(x) < \infty\}$. This notion will be useful, for example, when considering the expected value of the industry state.

If the firm instead decides to remain in the industry, then it can invest to improve its quality level. If a firm invests $\iota_{it} \in \Re_+$, then the firm's state at time $t+1$ is given by,

$$x_{i,t+1} = x_{it} + h(x_{it}, \iota_{it}, \zeta_{i,t+1}),$$

where the function h captures the impact of investment on quality and $\zeta_{i,t+1}$ reflects uncertainty in the outcome of investment. Uncertainty may arise, for example, due to the risk associated with a research and development endeavor or a marketing campaign. This specification is quite general as h may take on either positive or negative values, for example, allowing for positive depreciation. This process generalizes the EP model so that the process can depend on the current value of the firm's state variable, and also so that the firm's state variable can in principle move up or down an arbitrary number of steps each period. We denote the unit cost of investment by d.

Our model can also accommodate a variety of different entry processes, and we will use several different entry models in the remainder of the article. Specifically, the number of entrants can be fixed or random, entry costs can be fixed or random, and the entry state can be predetermined, randomly determined, or even controlled (meaning the entrant can either choose the entry state or at least influence it). Here we describe the entry model of Weintraub et al. (2008). In that model, there is an infinite supply of potential entrants, and in each period potential entrants can enter the industry by paying a fixed setup cost κ. Entrants do not earn profits in the period that they enter. They appear in the following period at state x^e and can earn profits thereafter. We describe equilibrium behavior below, but it should be clear that, due to an infinite supply of entrants, with this entry process entrants will earn zero profits and there will be no symmetric pure strategy equilibrium for potential entrants.

Each firm aims to maximize expected net present value. The interest rate is assumed to be positive and constant over time, resulting in a constant discount factor of $\beta \in (0, 1)$ per time period.

In each period, events occur in the following order:

(i) Each incumbent firms observes its sell-off value and then makes exit and investment decisions.
(ii) The number of entering firms is determined and each entrant pays an entry cost of κ.
(iii) Incumbent firms compete in the spot market and receive profits.
(iv) Exiting firms exit and receive their sell-off values.
(v) Investment outcomes are determined, new entrants enter, and the industry takes on a new state s_{t+1}.

□ **Model primitives.** The model as specified is general enough to encompass numerous applied problems in economics. Indeed, a blossoming recent literature on EP-type models has applied similar models to advertising, auctions, collusion, consumer learning, environmental policy, international trade policy, learning-by-doing, limit order markets, mergers, network externalities, and other applied problems (see Doraszelski and Pakes, 2007). To study any particular applied problem it is necessary to further specify the primitives of the model, including:

profit function	π
sell-off value distribution	$\sim \phi_{it}$
investment impact function	h
investment uncertainty distribution	$\sim \zeta_{it}$
unit investment cost	d
entry cost	κ
discount factor	β

In most empirical applications the profit function would not be specified directly but would instead result from a deeper set of primitives that specify a demand function, a cost function, and a static equilibrium concept. A good example of such primitives is the model described above: logit demand with single product firms, constant marginal cost, and a static equilibrium that is Nash in prices. Another example would be the well-known model of Berry, Levinsohn, and Pakes (1995). Note also that though we call the agents in the model "firms" and our examples in this article are models of industries, there is no reason that the agents could not be individuals or other agents, and no reason that our ideas cannot be applied in other settings.

□ **Assumptions.** We make several assumptions about the model primitives, beginning with the profit function. The first assumption assumes that profits are increasing in the firm's own state (product quality), and that profits are positive and bounded. These are natural assumptions in our model.

Assumption 2.1.

1. For all $s \in \mathcal{S}$, $\pi(x, s)$ is increasing in x.
2. For all $x \in \mathbb{N}$ and $s \in \mathcal{S}$, $\pi(x, s) \geq 0$, and $\sup_{x,s} \pi(x, s) < \infty$.

We make similar assumptions about investment and the distributions of the private shocks. Again, these assumptions are natural. The economic implications of the assumptions are described below.

Assumption 2.2.

1. The variables $\{\phi_{it} | t \geq 0, i \geq 1\}$ are i.i.d. and have finite expectations and well-defined density functions with support $\Re_+$.
2. The random variables $\{\zeta_{it} | t \geq 0, i \geq 1\}$ are i.i.d. and independent of $\{\phi_{it} | t \geq 0, i \geq 1\}$.
3. For all (x, ζ), $h(x, \iota, \zeta)$ is nondecreasing in ι.
4. For all $\iota > 0$ and x, $\mathcal{P}[h(x, \iota, \zeta_{i,t+1}) > 0] > 0$.
5. There exists a positive constant $\overline{h} \in \mathbb{N}$ such that $|h(x, \iota, \zeta)| \leq \overline{h}$, for all (x, ι, ζ). There exists a positive constant $\bar{\iota}$ such that $\iota_{it} < \bar{\iota}$, $\forall i, \forall t$.
6. For all $k \in \{-\overline{h}, \ldots, \overline{h}\}$ and x, $\mathcal{P}[h(x, \iota, \zeta_{i,t+1}) = k]$ is continuous in ι.
7. The transitions generated by $h(x, \iota, \zeta)$ are unique investment choice admissible.

Assumptions 2.2.1 and 2.2.2 assume that investment and exit outcomes are random and idiosyncratic conditional on the state. Assumptions 2.2.3 and 2.2.4 imply that investment on average increases the value of the firm's state variable (product quality). Note that positive depreciation is neither required nor ruled out. Assumption 2.2.5 places a finite bound on how much progress can be made or lost in a single period through investment. In EP this bound is 1, but here we simply assume that it is finite. Assumption 2.2.6 ensures that the impact of investment on transition probabilities is continuous. Assumption 2.2.7 is an assumption introduced by Doraszelski and Satterthwaite (2010) that ensures a unique solution to the firms' investment decision problem. It is used to guarantee existence of an equilibrium in pure strategies and is satisfied by many of the commonly used investment specifications in the literature.

The next assumption is about the entry process. For our main results, we assume that there are a large number of potential entrants who play a symmetric mixed entry strategy. In that case, the number of actual entrants is well approximated by the Poisson distribution (see Weintraub, Benkard, and Van Roy, 2008, for a derivation of this result). This leads to the following assumptions:

Assumption 2.3.

1. The number of firms entering during period t is a Poisson random variable that is conditionally independent of $\{\phi_{it}, \zeta_{it} | t \geq 0, i \geq 1\}$, conditioned on s_t.
2. $\kappa > \beta \cdot \bar{\phi}$, where $\bar{\phi}$ is the expected net present value of entering the market, investing zero and earning zero profits each period, and then exiting at an optimal stopping time.

We denote the expected number of firms entering at industry state s_t, by $\lambda(s_t)$. This state-dependent entry rate will be endogenously determined in equilibrium, and our solution concept will require that it satisfies a zero expected profit condition. Modelling the number of entrants as a Poisson random variable has the advantage that it allows us to realistically model entry for a wide variety of markets with varying numbers of firms. However, as noted above, our results can accommodate other entry processes as well. Assumption 2.3.2 ensures that the sell-off value by itself is not so large that it provides sufficient reason alone to enter the industry.

□ **Equilibrium.** As a model of industry behavior we focus on pure strategy Markov perfect equilibrium (MPE), in the sense of Maskin and Tirole (1988). We further assume that equilibrium is symmetric, such that all firms use a common stationary investment/exit strategy.

In MPE, incumbent firms follow a Markov (investment and exit) strategy that is self-generating in the sense of being optimal when all competitor firms follow the same strategy. At the same time, the equilibrium mixed-entry strategy among potential entrants is an entry rate such that entrants are exactly indifferent between entering and not entering. We now introduce some notation required to state these two properties precisely.

We denote each incumbent firm's investment strategy as $\iota(x_{it}, s_{-i,t})$, which is a function of the firm's own state and its competitors' states. Incumbent firms also have an exit strategy, $\rho(x_{it}, s_{-i,t})$, that takes the form of a cutoff rule: firm $i \in S_t$ exits at time t if and only if the sell-off value is greater than ρ, that is, $\phi_{it} \geq \rho(x_{it}, s_{-i,t})$.[2] Let $\mathcal{M}$ denote the set of exit/investment strategies, and Λ denote the set of entry rate functions.

We define the value function $V(x, s|\mu', \mu, \lambda)$ to be the expected net present value for a firm at state x when its competitors' state is s, given that its competitors each follow a common strategy $\mu \in \mathcal{M}$, the entry rate function is $\lambda \in \Lambda$, and the firm itself follows strategy $\mu' \in \mathcal{M}$:

$$V(x, s|\mu', \mu, \lambda) = E_{\mu',\mu,\lambda}\left[\sum_{k=t}^{\tau_i} \beta^{k-t}\left(\pi(x_{ik}, s_{-i,k}) - d\iota_{ik}\right) + \beta^{\tau_i - t}\phi_{i,\tau_i}\middle| x_{it} = x, s_{-i,t} = s\right],$$

where i is taken to be the index of a firm at quality level x at time t, τ_i is a random variable representing the time at which firm i exits the industry, and the subscripts of the expectation indicate the strategy followed by firm i, the strategy followed by its competitors, and the entry rate function. In an abuse of notation, we will use the shorthand, $V(x, s|\mu, \lambda) \equiv V(x, s|\mu, \mu, \lambda)$, to refer to the expected discounted value of profits when firm i follows the same strategy μ as its competitors.

An equilibrium to our model comprises an investment/exit strategy $\mu = (\iota, \rho) \in \mathcal{M}$, and an entry rate function $\lambda \in \Lambda$ that satisfy the following conditions:

(i) Incumbent firm strategies represent a MPE:

$$\sup_{\mu' \in \mathcal{M}} V(x, s|\mu', \mu, \lambda) = V(x, s|\mu, \lambda) \qquad \forall x \in \mathbb{N},\ \forall s \in \overline{\mathcal{S}}. \tag{2.1}$$

(ii) At each state, either entrants have zero expected profits or the entry rate is zero (or both):

$$\begin{aligned} \textstyle\sum_{s\in\overline{\mathcal{S}}} \lambda(s)\left(\beta E_{\mu,\lambda}\left[V(x^e, s_{-i,t+1}|\mu, \lambda)|s_t = s\right] - \kappa\right) &= 0 \\ \beta E_{\mu,\lambda}\left[V(x^e, s_{-i,t+1}|\mu, \lambda)|s_t = s\right] - \kappa &\leq 0 \qquad \forall s \in \overline{\mathcal{S}} \\ \lambda(s) &\geq 0 \qquad \forall s \in \overline{\mathcal{S}}. \end{aligned}$$

Doraszelski and Satterthwaite (2010) establish existence of an equilibrium in pure strategies for a closely related model. We do not provide an existence proof here because it is long and cumbersome and would replicate this previous work. With respect to uniqueness, in general we presume that our model may have multiple equilibria.[3]

[2] In Weintraub et al. (2008) we show that there always exists an optimal exit strategy of this form, even among very general classes of exit strategies.

[3] Doraszelski and Satterthwaite (2010) also provide an example of multiple equilibria in their closely related model.

There are a wide variety of dynamic programming algorithms available that can be used to solve for MPE to our model. However, these algorithms require compute time and memory that grow at least proportionately with the number of relevant industry states, which in turn grows quickly (at least high-order polynomially) with the number of individual states and firms in the industry. In the empirical models we have in mind, the number of industry states is typically so large that it would not be possible to store the strategy function on a computer, let alone compute it. This difficulty motivates our alternative approach.

3. Partially oblivious equilibrium

■ Although compute time and memory requirements for OE are not systematically related to the overall number of firms in a market, the notion of OE abstracts away some of the strategic effects of dynamic competition that seem likely to be important in concentrated markets, and it would be useful to reintroduce some of these effects into the model. In an effort to do this, in this section we introduce a natural extension of OE that we call *partially oblivious equilibrium* (henceforth, POE). In POE, firms' information sets consist of their own firm state (quality) and also the firm states of a few strategically important firms, that we label *dominant firms*.

As discussed in the Introduction, in POE we assume that there are a fixed number n of dominant firms whose identities do not change over time; firms always keep track of the same set of dominant firms and, in the basic model, all new entrants become part of the fringe. In the basic model, dominant firms also never exit the industry. It is a straightforward extension, that we implement in some of the computational experiments below, to add entry and exit of dominant firms to the model (but still without allowing transitions between dominant and fringe firms). In that case, the POE model nests both OE and MPE as special cases, in which there are no dominant firms and all dominant firms, respectively. We omit this feature of the model here to keep the exposition cleaner.

Let $\overline{D}$ denote the set of (indices associated with) the dominant firms. In the POE model, each firm i has its quality state x_{it}, and additionally has a binary state that identifies whether or not firm i is dominant. Hence, the state of firm i at time t will now be denoted $\overline{x}_{it}$, where $\overline{x}_{it} = (x_{it}, 1)$ if $i \in \overline{D}$, and $\overline{x}_{it} = (x_{it}, 0)$ if $i \notin \overline{D}$. This specification is necessary to allow for equilibrium strategies to differ between dominant firms and fringe firms.

□ **Partially oblivious strategies.** In principle, we would like to model partially oblivious strategies as being functions of the firm's own state (which, as above, includes the firm's quality and also whether the firm is dominant or not), and also the states of the dominant firms. However, there is a technical complication with such an approach that we need to first address before precisely defining partially oblivious strategies.

In the POE model, because firms are assumed not to know (or not to keep track of) the values of the fringe firms' states, it is necessary to model firms' beliefs about the values of fringe firms' states. In OE, all firms are fringe firms and, instead of optimizing against the actual industry state, firms optimize against the long-run average industry state in equilibrium. If there are many firms, such behavior is close to optimal. We will do something analogous for POE, but there is a complication. In the POE model, fringe firms also know the state of the dominant firms and follow strategies that are functions of the dominant firms' states. This feature of the model causes the distribution of fringe firms to interact with the state of the dominant firms (through fringe firm strategies). As a result, the distribution of fringe firm states at any time t depends, in principle, on all past values of the dominant firms' states–through the past fringe firms' strategies. To make the most precise prediction about the fringe, then, firms should use a complete history of dominant firm states, which implies that firm strategies should be functions of the complete history of dominant firm states.

There are two problems with such an approach. First, computing expectations this way is unrealistic as a model of behavior. In reality, firms simply do not have infinite memories. Second, such a model would also be computationally impractical. Instead, we will restrict firms to predict the fringe firms' states using a finite set of statistics that depend on the past history of the dominant firms' state.[4] We thus restrict firms' strategies so that each firm's decisions depend only on the firm's state, the current state of the dominant firms, and a finite set of statistics that depend on the history of realizations of the dominant firms' states. We call such restricted strategies *partially oblivious strategies*.

Formally, let y_t be a vector that represents the states of the dominant firms at time t.[5] Let the sequence w_t represent a finite set of statistics that depend on the history of realizations of the dominant firms' states: $\{w_t \in \mathcal{W} = \mathcal{W}_1 \times \cdots \times \mathcal{W}_K : t \geq 0\}$ where $w_t(1) = y_t$, for all $t \geq 0$, and $\mathcal{W}_j$ are countable sets. Partially oblivious strategies are functions of the firm's own state, x_{it}, and the set of statistics of the dominant firm states, w_t. As $w_t(1) = y_t$, firm strategies are always functions of the current dominant firms' state. The state variables $w_t(2), \ldots, w_t(K)$ allow firms to incorporate additional information about the history of realizations of the dominant firms' state, such as recent past realizations of the dominant firms' state, into their strategies. As previously discussed, this information could be useful to better predict the current distribution of fringe firm states.

We define $\tilde{\mathcal{M}}_p$ and $\tilde{\Lambda}_p$ as the set of partially oblivious (investment and exit) strategies and the set of partially oblivious entry rate functions, respectively. If firm i uses investment and exit strategy $\mu \in \tilde{\mathcal{M}}_p$, then firm i takes action $\mu(\overline{x}_{it}, w_t)$ at time period t, where $\overline{x}_{it}$ is the state of firm i at time t (which indicates the firm's quality level and whether it is dominant or not), and w_t is the set of statistics of the dominant firms' states. Similarly, if the entry rate function is $\lambda \in \tilde{\Lambda}_p$, the entry rate is equal to $\lambda(w_t)$ at time period t.

In the following, we assume the stochastic process $\{w_t : t \geq 0\}$ is Markov and admits a unique invariant distribution that assigns positive mass to all states $w \in \mathcal{W}$.

Assumption 3.1. We assume that, for any partially oblivious strategy, $\{w_t : t \geq 0\}$ is a finite irreducible and aperiodic Markov chain adapted to the filtration generated by $\{y_t : t \geq 0\}$. For all $t \geq 0$, $w_t(1) = y_t$.

In words, the assumption means that every feasible strategy yields full support on the state space of the dominant firms' states, w_t. The assumption ensures that the expectations over fringe firms' states defined in the next subsection are well defined at every state.

The assumption is useful because when we define equilibrium below, we will require that firms' beliefs about the distribution of fringe firms be defined at every state, w_t, under all feasible strategies. Firms' beliefs, which we will make precise in the next subsection, are simply formed by rational expectations under the given strategies. If not all states were reachable under a particular set of strategies, then rational beliefs would be undefined for some states, which would undermine the definition of equilibrium.

Assumption 3.1 is somewhat strong, but it is satisfied in many empirical models. In discrete games such as the CW model or Pakes, Ostrovsky, and Berry (2007), for example, the i.i.d error term (logit in CW) provides full support on the set of available choices, which yields full support on the state space. In other models, such as Ericson and Pakes (1995) or Benkard (2004), Assumption 3.1 would not strictly be satisfied because there are strategies under which investment is always zero that would not yield full support. However, in most applications of interest (like those presented in Section 5), this would just be a minor technical problem because strategies of this form would not typically arise as best responses for real-world parameter values.

[4] This idea is similar to OE with aggregate shocks; see Weintraub, Benkard, and Van Roy (2010). Fershtman and Pakes (2012) also consider finite histories in a setting of asymmetric information.

[5] To be precise, $y_t = (x_{(i_1)t}, x_{(i_2)t}, \ldots, x_{(i_D)t})$, where $x_{(i_k)t}$ is the $k-$th order statistic of $(x_{i_1 t}, x_{i_2 t}, \ldots, x_{i_D t})$, i_j is the index of the j-th dominant firm, and D is the total number of dominant firms. Our equilibrium concept will assume anonymous strategies that are a function of the order statistics of $(x_{i_1 t}, x_{i_2 t}, \ldots, x_{i_D t})$ and will not depend on the identities of the dominant firms.

Still, there are also models where Assumption 3.1 may fail to hold in a way that matters and inhibits equilibrium existence and computation, such that at interesting parameter values we require the ability to compute beliefs at states that are never reached under some relevant alternative strategies. In such cases, there are two possible ways to proceed, and we believe that both would be viable alternatives in almost any model.

The first alternative is to drop Assumption 3.1 and then weaken the definition of equilibrium. Instead of defining equilibrium on the whole state space, as we do below, we could require the equilibrium conditions to hold only on the interior of the recurrent class of states (or alternatively, any set of states larger than the recurrent class but possibly smaller than the whole state space such that beliefs were well defined on this set). This is essentially what is done in a similar context in Fershtman and Pakes (2012). Under this weaker notion of equilibrium, because the equilibrium conditions need not hold outside the recurrent class, firms' beliefs need not be defined outside the recurrent class for an equilibrium to exist, and we do not need Assumption 3.1.

The second alternative is to drop Assumption 3.1 and then provide an alternative way of defining beliefs about the fringe firms at states that are never reached. Because rational beliefs are not defined at states that are never reached, there is little discipline placed on beliefs at these states by theory, and it would be possible to define POE under many different model-dependent specifications of beliefs. For example, we could define optimistic beliefs, such that a firm believes that if such a state is reached then all the fringe firms will have exited, or pessimistic beliefs, such that beliefs are that all the fringe firms will be at high states, or status quo beliefs, so that nothing changes from the current state. It would also be possible to compute equilibria under different assumptions about off-equilibrium beliefs and then see if it made any difference to the results.

Below, we maintain Assumption 3.1 because it allows us to always use rational beliefs, which we believe is a natural way to proceed.

□ **Expected fringe firms' state.** In the POE model, firms have only knowledge of their own state and the dominant firms' state. As noted above, we need to make an assumption about how firms form expectations over the fringe firms' state when the dominant firms' state statistic takes on a particular value $w_t = w$. To operationalize this, analogously to OE, we assume that firms believe that the fringe firms' state is equal to the *expected* fringe firms' state conditional on the current value of the dominant firm state statistics $w_t = w$ (that would result in the long run under a particular strategy). We now define this expectation formally, which requires some additional notation.

Formally, suppose that firms use investment/exit strategy $\mu \in \tilde{\mathcal{M}}_p$ and enter according to $\lambda \in \tilde{\Lambda}_p$. Let f_t represent the state of the fringe firms. That is, f_t is a vector over quality levels that specifies, for each quality level $x \in \mathbb{N}$, the *number of fringe firms* that are at quality level x in period t. We also define the state space for the fringe firms analogously to the industry state $(\overline{\mathcal{S}})$: $\overline{\mathcal{F}} = \{f \in \mathbb{N}^\infty | \sum_{x=0}^{\infty} f(x) < \infty\}$. Assumption 3.1 together with Assumptions 2.1, 2.2, and 2.3 imply that $\{(f_t, w_t) : t \geq 0\}$ is a Markov chain that admits a unique invariant distribution. We assume that (f_0, w_0) is distributed according to the invariant distribution of $\{(f_t, w_t) : t \geq 0\}$. Hence, (f_t, w_t) is a stationary process. We can now define the conditional expected fringe firm's state, which we will denote $\tilde{f}$. For all $w \in \mathcal{W}$, we define

$$\tilde{f}(w|\mu, \lambda) = E[f_t | w_t = w; \mu, \lambda].$$

In words, $\tilde{f}(w|\mu, \lambda)$ is the long-run expected fringe firm state when dynamics are governed by partially oblivious strategy μ and partially oblivious entry rate function λ, conditional on the current realization of the dominant firms' state statistics w_t being w.

□ **Partially oblivious value function.** Let $\pi(x_{it}, f_{-i,t}, y_{-i,t})$ be the single-period profits for a firm i in state x_{it}, when the fringe firms other than firm i are in state $f_{-i,t}$ and dominant firms other than firm i are in state $y_{-i,t}$. Recall that, in our model formulation, static period profits depend only on the firm states of all dominant and fringe firms, and not on which firms are dominant.

We now define a *partially oblivious value function* as follows:

$$\tilde{V}(\overline{x}, w|\mu', \mu, \lambda) = E_{\mu',\mu}\left[\sum_{k=t}^{\tau_i} \beta^{k-t}\left(\pi(x_{ik}, \tilde{f}_{-i}(w_k|\mu, \lambda), y_{-i,k}) - d\iota_{ik}\right) + \beta^{\tau_i - t}\phi_{i,\tau_i}\middle|\overline{x}_{it} = \overline{x}, w_t = w\right],$$

where i is the index of a firm in state $\overline{x}$ at time period t, recalling that $\overline{x}$ includes the firm's quality state and a binary state indicating whether the firm is dominant.[6]

This value function should be interpreted as the expected net present value of firm i at state $\overline{x}$ when the dominant firms' statistics have value w, firm i follows partially oblivious strategy μ', and competitors use strategy μ and enter according to λ. If firm i is in the fringe, then it must subtract itself when computing $\tilde{f}_{-i}(w_k|\mu, \lambda)$. Appendix A Subtracting Oneself in OE shows how to compute $\tilde{f}_{-i}(w_k|\mu, \lambda)$ for fringe firms in several commonly used models. Finally, in an abuse of notation, we define $\tilde{V}(\overline{x}, w|\mu, \lambda) = \tilde{V}(\overline{x}, w|\mu, \mu, \lambda)$.

□ **Partially oblivious equilibrium.** A *partially oblivious equilibrium* consists of an investment and exit strategy $\mu \in \tilde{\mathcal{M}}_p$ and an entry rate function $\lambda \in \tilde{\Lambda}_p$ that are self-generating in the sense that (i) the investment and entry strategy μ is optimal when all competitor firms follow the same strategy μ and the entry rate is λ, and (ii) such that when all firms follow strategy μ and the entry rate is λ, potential entrants are exactly indifferent between entering and not entering. Formally, these conditions are:

(i) Firm strategies optimize a partially oblivious value function.

$$\sup_{\mu' \in \tilde{\mathcal{M}}_p} \tilde{V}(\overline{x}, w|\mu', \mu, \lambda) = \tilde{V}(\overline{x}, w|\mu, \lambda), \qquad \forall(\overline{x}, w) \in \{(\overline{x}, w) : w \in \mathcal{W}, \overline{x} \in \overline{X}(w)\}. \quad (3.1)$$

(ii) The partially oblivious expected value of entry is zero or the entry rate is zero (or both) .

$$\sum_{w \in \mathcal{W}} \lambda(w)(\beta E[\tilde{V}((x^e, 0), w_{t+1}|\mu, \lambda)|w_t = w] - \kappa) = 0,$$
$$\beta E[\tilde{V}((x^e, 0), w_{t+1}|\mu, \lambda)|w_t = w] - \kappa \leq 0, \quad \forall w \in \mathcal{W},$$
$$\lambda(w) \geq 0, \quad \forall w \in \mathcal{W}.$$

To derive, a POE it is enough to consider one dominant firm and one fringe firm. New entrants become part of the fringe. Finally, if $n = 0$, so that there are no dominant firms, a POE is an OE.

In Section 3, we provide an algorithm for computing a POE. The state space of the firm's dynamic programming problem scales with the number of firm states and with the size of $\mathcal{W}$, the feasible set for the dominant firm statistics. As the set $\mathcal{W}$ becomes richer, for example, as we add more dominant firms, more computation time and memory is needed.

□ **Algorithms and computations.** In this section, we introduce an algorithm to compute POE.[7] The algorithm is an iterative algorithm. It is initialized with OE strategies, $(\tilde{\mu}, \tilde{\lambda})$ (lines 1–3), and the associated OE value function $\tilde{V}$. The algorithm then starts by computing the expected fringe firm state conditional on the dominant firm statistics, $\tilde{f}(w|\mu, \lambda)$ (line 5). This step can be computed via simulation, or more simply by solving systems of linear equations associated to "balance equations" that are typically used to compute expectations in Markov processes. We provide details of this procedure in Appendix A Balance Equations.

The algorithm then computes the strategy that maximizes the partially oblivious value function (line 6). In our implementation of the algorithm, we use Gauss-Seidel value iteration for this step. Next, the algorithm uses the zero-profit conditions to update the entry rates

[6] Note that, as a technical matter, because w includes a complete list of the quality states of the dominant firms, for dominant firms $\tilde{V}$ is only defined for quality levels x_{it} that are listed in w. For all $w \in \mathcal{W}$, we can define this set as $\overline{X}(w) = \{(x, 1) : x = w(1, k) \text{ for some } k = 1, \ldots, D\} \cup \{(x, 0) : x \in \mathbb{N}\}$, where $w(1, k)$ is the k-th component of $w(1)$.

[7] Throughout this section, we consider only states $(\overline{x}, w) \in \{(\overline{x}, w) : w \in \mathcal{W}, \overline{x} \in \overline{X}(w)\}$, where $\overline{X}(w)$ is defined in the previous footnote.

(line 8). If the value of entering is larger (smaller) than the entry cost, the entry rate is proportionally increased (decreased). Finally, strategies and entry rates are updated "smoothly" (lines 12 and 13). The smoothing parameters N_1, N_2, η_1, and η_2 are set after some experimentation to speed up convergence.[8] If the termination condition of the outer loop is satisfied with $\epsilon_1 = \epsilon_2 = 0$, we have a POE. Small values of ϵ_1 and ϵ_2 allow for small errors associated with limitations of numerical precision.

Algorithm 1 Partially Oblivious Equilibrium Solver

1: $\lambda(w) := \tilde{\lambda}$, for all w
2: $\mu(\overline{x}, w) := \tilde{\mu}(x)$, for all $\overline{x}, w$
3: $k := 0$
4: **repeat**
5: Compute $\tilde{f}(w|\mu, \lambda)$, for all w
6: Choose $\mu^* \in \tilde{\mathcal{M}}_p$ to maximize $\tilde{V}(\overline{x}, w|\mu^*, \mu, \lambda)$ simultaneously for all $\overline{x}, w$
7: **for all** w **do**
8: $\lambda^*(w) := \lambda(w)\left(\beta E\left[\tilde{V}((x^e, 0), w_{t+1}|\mu^*, \mu, \lambda)\middle| w_t = w\right]/\kappa\right)$
9: **end for**
10: $\Delta_1 :=\| \mu - \mu^* \|_\infty, \Delta_2 := \|\lambda - \lambda^*\|_\infty$
11: $k := k + 1$
12: $\mu := \mu + (\mu^* - \mu)/(k^{\eta_1} + N_1)$
13: $\lambda := \lambda + (\lambda^* - \lambda)/(k^{\eta_2} + N_2)$
14: **until** $\Delta_1 \leq \epsilon_1$ and $\Delta_2 \leq \epsilon_2$

□ **Extension: simulated OE.** In markets with a large number of firms, the actual state of i's competitors ($f_{-i,t}$) will be close to its expected state ($\tilde{f}_{-i}(\mu, \lambda)$) with high probability due to a law of large numbers. Hence, in our previous work on OE, we found that it was computationally easier and did not impact the results to replace $f_{-i,t}$ with its expectation in the firm's maximization problem. That is, firms were assumed to form beliefs such that the distribution of competitor firms was always exactly equal to its long-run average, and when there are many firms such beliefs are close to correct. Specifically, in our past work we defined an oblivious value function as:

$$\tilde{V}(x|\mu', \mu, \lambda) = E_{\mu'}\left[\sum_{k=t}^{\tau_i} \beta^{k-t}(\pi(x_{ik}, \tilde{f}_{-i}(\mu, \lambda)) - d\iota_{ik}) + \beta^{\tau_i - t}\phi_{i,\tau_i}\middle| x_{it} = x\right], \tag{3.2}$$

where, as above, $\tilde{f}_{-i}(\mu, \lambda)$ represents the expected state of i's competitors and profits are evaluated at this expected state instead of the actual state. POE is defined analogously above, but including dominant firms.

If there are only a small number of firms, variation in the industry state over time is greater and may matter more in the firm's optimization problem. This is particularly relevant for single-period profits that are not smooth as a function of the competitors' states. In these cases, bringing the expectation outside the profit function may improve the model. In words, firms may be able to do better by optimizing against the long-run distribution of competitor states, instead of always assuming that competitors were at the long-run average state.

Thus motivated, it is possible to make a slight modification to OE, which we will call "Simulated OE" or OE-SIM, in which firms instead integrate profits over the full equilibrium distribution of their competitors. The new OE value function would then be

$$\tilde{V}(x|\mu', \mu, \lambda) = E_{\mu',\mu,\lambda}\left[\sum_{k=t}^{\tau_i} \beta^{k-t}(\pi(x_{ik}, f_{-i,k}) - d\iota_{ik}) + \beta^{\tau_i - t}\phi_{i,\tau_i}\middle| x_{it} = x\right]. \tag{3.3}$$

[8] In the computational experiments below, we mostly used $N_1 = N_2 = 1$ and $\eta_1 = \eta_2 = 0.2$. In a few troublesome cases, we required $\eta_= \eta_2 = 0.05$.

Note that $f_{-i,k}$ is still distributed according to the *invariant distribution of competitors' states*. Hence, this is still an "oblivious" value function that does not keep track of the *actual industry state*. However, the expectation over competitors' states with strategies (μ, λ) has been moved outside of the profit function. In small markets, the added computational burden of this extra layer of integration (which derives primarily from having to recompute π at many states) is low, so for small markets OE-SIM is still easy to compute. For larger markets the computational burden could become high, but in those cases it is less helpful. As part of the proof of the main theorem in Weintraub, Benkard, and Van Roy (2008), we show that the difference between the value functions in (3.2) and (3.3), and hence the difference between OE and OE-SIM, goes to zero as the number of firms in the industry becomes large, under relatively weak smoothness conditions over the single-period profit function.

We explore these issues further in the computational experiments below. Note that it would also be possible to compute a simulated version of POE.

4. Error bounds

■ Because firms' strategies are restricted in POE, it may be useful to consider how POE strategies perform relative to a full information strategy. The way we will do this is to bound the amount by which a firm can improve its expected discounted value by unilaterally deviating to a full information Markov strategy. If the bound is found to be large, that would mean that POE strategies are potentially very restrictive. In that case, it seems likely that firms would use more complex strategies in practice. Thus, it may be useful to make the model richer, such as by increasing the number of dominant firms.

We derive error bounds only for fringe firms in the POE model. Because dominant firms' states are tracked by fringe firms, deviations by dominant firms lead to changes in the distribution of future industry states, so it is not straightforward to derive an error bound for dominant firms that is easily computed.

We define $\mathcal{M}_p$ and Λ_p as the set of *extended* Markov strategies and entry rate functions for the fringe. An extended Markov strategy assumes full information about the industry state, so it is a function of the firm's own state, the full industry state (including the dominant firms' state), and the dominant firms' state statistics. It is important to consider the dominant firms' state statistics in the extended Markov strategies, because other firms use this information when using POE strategies. If a fringe firm i uses strategy $\mu \in \mathcal{M}_p$, then at time period t, fringe firm i takes action $\mu(\bar{x}_{it}, f_{-i,t}, w_t)$. Because we will only consider unilateral deviations for fringe firms, we do not extend the information set of the dominant firms, still restricting their strategies to depend only on their own state, other dominant firms' state, and the dominant firms' state statistics. Similarly, if $\lambda \in \Lambda_p$, then the entry rate at time t is $\lambda(f_t, w_t)$. (Recall that $w_t(1) = y_t$, and hence strategies are a function of y_t, the current dominant firms' state.)

For extended Markov strategies $\mu', \mu \in \mathcal{M}_p$ and extended entry rate function $\lambda \in \Lambda_p$, with some abuse of notation, we define for $\bar{x} = (x, 0)$ the fringe extended value function,

$$\begin{aligned} &V(\bar{x}, f_{-i}, w|\mu', \mu, \lambda) \\ &= E_{\mu',\mu,\lambda}\left[\sum_{k=t}^{\tau_i} \beta^{k-t}(\pi(x_{ik}, f_{-i,k}, y_k) - d\iota_{ik}) + \beta^{\tau_i - t}\phi_{i,\tau_i} \middle| x_{it} = x, f_{-i,t} = f_{-i}, w_t = w\right], \end{aligned} \tag{4.1}$$

where i is taken to be the index of a firm at quality level x at time t and f_{-i} to be the initial state of fringe competitors of i. Because i is a fringe firm, we can write y_k instead of $y_{-i,k}$. The extended value function generalizes the value function defined in Section 2, allowing for dependence on extended strategies. We use this value function to evaluate the *actual* expected discounted profits garnered by a firm that uses an extended Markov strategy with full information of the industry state.

Consider a POE strategy and entry rate $(\tilde{\mu}, \tilde{\lambda}) \in \tilde{\mathcal{M}}_p \times \tilde{\Lambda}_p$. We assume the initial state (f_0, w_0) is sampled from the invariant distribution of $\{(f_t, w_t) : t \geq 0\}$. Hence, (f_t, w_t) is a stationary process, it is distributed according to its invariant distribution for all $t \geq 0$. To abbreviate, let $\tilde{f}_{-i}(w) \equiv \tilde{f}_{-i}(w|\tilde{\mu}, \tilde{\lambda})$. Let $\Delta(f_{-i}, w) = \{\pi(x, f_{-i}, w(1)) - \pi(x, \tilde{f}_{-i}(w), w(1))\}_{x \in \mathbb{N}}$. The vector $\Delta(f_{-i}, w)$ represents the profit differences when the single-period profit function is evaluated at the actual fringe firm state f_{-i} versus at the conditional expected state $\tilde{f}_{-i}(w)$. We have the following result.[9]

Theorem 4.1. Let Assumptions 2.1, 3.1, 2.2, and 2.3 hold. Then, for any POE $(\tilde{\mu}, \tilde{\lambda})$, and fringe firm i with state $\bar{x} = (x, 0)$, $x \in \mathbb{N}$,

$$\begin{aligned} E\Big[&\sup\nolimits_{\mu' \in \mathcal{M}_p} V(\bar{x}, f_t, w_t | \mu', \tilde{\mu}, \tilde{\lambda}) \\ &- V(\bar{x}, f_t, w_t | \tilde{\mu}, \tilde{\lambda})\Big] \leq E\left[E_{\tilde{\mu}, \tilde{\lambda}}\left[\sum_{k=t}^{\infty} \beta^{k-t} H_k\left(\Delta(f_{-i,k}, w_k)\right) \Big| x_{it} = x, f_{-i,t}, w_t \right]\right], \end{aligned} \tag{4.2}$$

for a sequence of random functions $\{H_k : k \geq t\}$.

The result is an upper bound on the extra profit a firm could make by deviating from a POE and instead playing a fully informed Markov best response. Note that the left-hand side of the bound cannot be computed in general because it would require solving a dynamic programming problem that is subject to the curse of dimensionality. On the other hand, the right-hand side is easy to compute using simulation. It depends on the functions H_k, which are functions of the profit differences $\Delta(f_{-i,k}, w_k)$. For example, in one version of the bound, these functions can be specified as:

$$H_k\left(\Delta(f_{-i,k}, w_k)\right) = 2 \sup_{y \in \{0, \ldots, x + (k-t)\overline{h}\}} |\Delta(f_{-i,k}, w_k)|,$$

where $|\cdot|$ denotes the absolute value of each component of the vector and the supremum is taken over the components of the vector y specified. The bound can be computed easily via simulation because (f_k, w_k) is distributed according to the invariant distribution under POE, which is known and is easy to sample from. Notably, the bound does *not* depend on the optimal Markov unilateral deviation strategy.

The bound is constructed by comparing the expected discounted sum of differences of single-period profits evaluated at the actual fringe state, and the expected fringe state. Note that POE strategies are derived using single-period profits evaluated at the expected fringe state, whereas the fully informed Markov best response strategy uses the actual fringe state. The functions H_k consider the maximum difference of these profits among all individual states y reachable by time period k when starting from state x. The number 2 appears in the function H because the comparison of value functions is done twice, once for each value function on the left-hand side.

The specification for H_k given above is among the simplest possible examples. However, in Appendix A Error Bounds and Proofs, we provide other specifications for the functions H_k that are more involved, but that provide tighter bounds and that can also be computed via simulation.

We also note that our bounds are also quite general and do not take advantage of many of the detailed modelling assumptions. In particular, they are valid for a variety of entry processes. Allowing the bound to depend on specific details of the model would allow us to achieve tighter expressions (see Weintraub, Benkard, and Van Roy, 2010, for an example).

5. Computational experiments

■ In this section, we present a set of computational experiments that examine the dynamics generated by the POE and OE models in highly concentrated markets. We begin by analyzing two

[9] We note that the derivations of these error bounds are similar to those of our earlier article in which we introduced OE with aggregate shocks Weintraub, Benkard, and Van Roy (2010).

structurally different models: a differentiated products logit demand model with price setting and investment in product quality, and a homogeneous products Cournot model with investment that reduces marginal cost.

Price-quality competition. Our first model attempts to mimic standard IO models commonly used in empirical work. We consider an industry with differentiated products where each firm's state variable represents the quality of its product. There are m consumers in the market. In each period t, consumer j receives utility u_{ijt} from consuming the good produced by firm i given by:

$$u_{ijt} = \theta_1 \ln(x_{it} + 1) + \theta_2 \ln(Y - p_{it}) + \nu_{ijt},\ i \in S_t,\ j = 1, \ldots, m,$$

where Y is the consumer's income, and p_{it} is the price of the good produced by firm i. ν_{ijt} are i.i.d. random variables distributed Gumbel that represent unobserved characteristics for each consumer–good pair. There is also an outside good that provides consumers zero utility. We assume consumers buy at most one product each period and that they choose the product that maximizes utility. Under these Assumptions, our demand system is a classic logit model.

Let $N(x_{it}, p_{it}) = \exp(\theta_1 \ln(x_{it} + 1) + \theta_2 \ln(Y - p_{it}))$. Then, the expected market share of each firm is given by:

$$\sigma(x_{it}, s_{-i,t}, p_t) = \frac{N(x_{it}, p_{it})}{1 + \sum_{j \in S_t} N(x_{jt}, p_{jt})},\ \forall i \in S_t.$$

We assume that firms set prices in the spot market. If there is a constant marginal cost c, the Nash equilibrium of the pricing game satisfies the first-order conditions,

$$Y - p_{it} + \theta_2 (p_{it} - c)(\sigma(x_{it}, s_{-i,t}, p_t) - 1) = 0\ ,\ \forall i \in S_t. \tag{5.1}$$

There is a unique Nash equilibrium in pure strategies, denoted p_t^* (Caplin and Nalebuff, 1991). Expected profits are given by:

$$\pi_m(x_{it}, s_{-i,t}) = m\sigma(x_{it}, s_{-i,t}, p_t^*)(p_{it}^* - c)\ ,\ \forall i \in S_t.$$

Firms can also invest $\iota \geq 0$ to improve their product quality over time. A firm's investment is successful with probability $\frac{a\iota}{1+a\iota}$, in which case the quality of its product increases by one level. The firm's product may also depreciate one quality level at random with probability δ, independently each period. Our model differs from Pakes and McGuire (1994) here because the depreciation shocks in our model are idiosyncratic. Combining the investment and depreciation processes, it follows that the transition probabilities for a firm in state x that does not exit and invests ι are given by:

$$\mathcal{P}[x_{i,t+1} | x_{it} = x, \iota] = \begin{cases} \frac{(1-\delta)a\iota}{1+a\iota} & \text{if } x_{i,t+1} = x + 1 \\ \frac{(1-\delta)+\delta a\iota}{1+a\iota} & \text{if } x_{i,t+1} = x \\ \frac{\delta}{1+a\iota} & \text{if } x_{i,t+1} = x - 1\ . \end{cases}$$

Although it would be straightforward to implement entry and exit in the model, to simplify the model, we omit them from our computations.

In our computational experiments, we found that the logit model above was not capable of generating markets with one-firm concentration ratios above about 0.3. Thus, in order to test the POE model for extremely highly concentrated markets, we use a second quantity-cost (Cournot) model.

Quantity-cost competition. Our second model is carefully designed to be a "worst case" for OE. It will generate very highly concentrated near-monopoly markets that also have a fringe of smaller firms and can also generate high turnover among the leader firms. This will allow us to explore the performance of OE and POE in cases more extreme than would be found in any empirical application.

We consider an industry with a homogeneous product and quantity setting, where the state of each firm represents its marginal cost of production. The industry has a linear inverse demand function,

$$P(q_t) = m_1 - \sigma \sum_i q_{it},$$

The marginal cost for firm i in state x_{it} is

$$\mathrm{MC}(x_{it}) = \gamma_0 \exp(-(\gamma_1 x_{it} - \gamma_2)).$$

With this specification we are able to generate highly concentrated markets, but the model also has the property that in periods of high concentration lagging firms stop producing altogether, leading to a pure monopoly in many periods. It turns out that it is somewhat difficult to endogenously generate a market in which small firms continue to produce when the industry is very highly concentrated. Thus, to accomplish this goal we take a shortcut and assume that there is also a second market that always yields m_2 total surplus, that is split equally among the active firms. One can think of the first market as the "national" market, in which all firms compete in quantities, and the second market as a set of "local" markets, where each firm captures their local market completely regardless. This second market provides the small firms some market share even when there is a very efficient large firm.

Period profits for firm i in state x_{it} are given by:

$$\Pi(q_{-i,t}, x_{it}|s_{-i,t}) = \max_{q_{it}} \left\{ P(q_t)q_{it} - \mathrm{MC}(x_{it})q_{it} + m_2/n_t \right\},$$

where n_t is the number of active firms. Firms set quantities simultaneously conditioning on the observed state, and the spot market is assumed to be in static Nash equilibrium.

In this model, investment improves the state and reduces marginal cost. Transitions are modelled the same as in the logit model above. Also, as in the logit model, allowing for entry and exit would be straightforward, but we omit them in order to simplify the computations.

□ **Comparison of OE, POE, and MPE.** We first investigate the relationship between OE, POE, and MPE in the two models above. Our intention in this section is to identify the circumstances under which the three equilibrium concepts tend to be similar, as well as the circumstances under which the three tend to differ and, in those cases, explore how they differ.

In both models, we can obtain a wide range of industry structures by moving only two parameters: investment cost and depreciation. When investment is cheap, all firms invest a lot and the industry is not very concentrated. As the investment parameter increases, industry concentration increases, up to a point where investment is so expensive that firms stop investing altogether. The depreciation rate influences the rate of churn in the industry—when depreciation is low, leader firms retain their advantage for longer (and vice versa). We mention this here because the rate of turnover of leader firms is one aspect of an industry that we find to be important to the results.

The remaining parameter values were chosen to reflect reasonable economic fundamentals and then fixed for all the experiments. They are listed in Table 1. In each case, we model eight active firms and eight individual firm states. We use such a small number of firms and firm states because it reduces the burden of computing MPE to a level at which we can easily compute equilibria for many different parameter values. Finally, when implementing POE, we include only the current values of dominant firms' states.

For all three types of equilibria (OE, POE, and MPE) there are potentially multiple equilibria, though we would typically expect there to be fewer OE and POE than MPE (among other things, asymptotically every OE has an MPE nearby—Weintraub, Benkard, and Van Roy, 2008). We have made no attempt to compute all the equilibria for each model, but we have found that our

TABLE 1 Parameters for Numerical Simulations

	Cournot	Logit
Market size (M)	200.0	100.0
Demand slope (σ)	10.0	-
Marginal cost (γ_0)	1.0	-
Marginal cost (γ_1)	0.50	-
Marginal cost (γ_2)	5.0	-
Fixed cost (F)	0.0	-
Second demand (m_2)	5.0	-
Income (Y)	-	1.0
Quality sensitivity (θ_1)	-	0.8
Price sensitivity (θ_1)	-	0.5
Marginal cost	-	0.5
Investment effectiveness (a)	0.8	0.7
Investment cost	[5.0–40.0]	[0.1–4.3]
Discount factor	0.9	0.9
States per firm	8	8

computational algorithms consistently select the same equilibrium each time.[10] All of the results below apply to this particular equilibrium.

Logit model. Figures 1 and 2 provide a comparison of basic industry statistics between OE, POE, and MPE under a range of investment cost parameters for low and high depreciation, respectively. The top left panel in each figure shows that, for the logit model, as we move the investment cost parameter we obtain one-firm concentration ratios ranging from about 0.1 (representing eight equally sized firms) to a high of only 0.3. The top right panel shows the probability that a firm at state 1 will end up in the top half of the state space within the next 100 periods under the MPE strategies, a measure of the rate of turnover among the leading firms. Turnover among leading firms is always extremely high for this model (though leading firms are not much larger than laggards).

The next three panels compare long-run averages of investment, consumer surplus, and producer surplus under OE and POE with that under MPE. We compute POE with one, two, and three dominant firms (labeled POE(D), where D is the number of dominant firms). The remaining three panels show the error bounds for fringe firms, as well as the actual maximal Markov best response improvement in firm values for fringe and dominant firms, respectively. The bounds are presented in percentage terms relative to the value function, where industry states are averaged out according to their invariant distribution.

Several features of the results stand out. First, in the logit model, despite the fact that these markets have no more than eight firms, for almost all parameter values the differences are small between OE, POE, and MPE. The only exception to this rule is the large percentage difference in investment for the highest investment costs. However, this difference is explained by the fact that investment is close to zero for these cases, so a high percentage difference corresponds to a small absolute difference. Consistently with what we have found in previous work (Weintraub, Benkard, and Van Roy, 2010) consumer and producer surplus in OE are typically within 1%–2% of their values in MPE, and firms can only obtain at most a 1.5% (fringe) or 0.4% (dominant) improvement by using full Markov strategies.

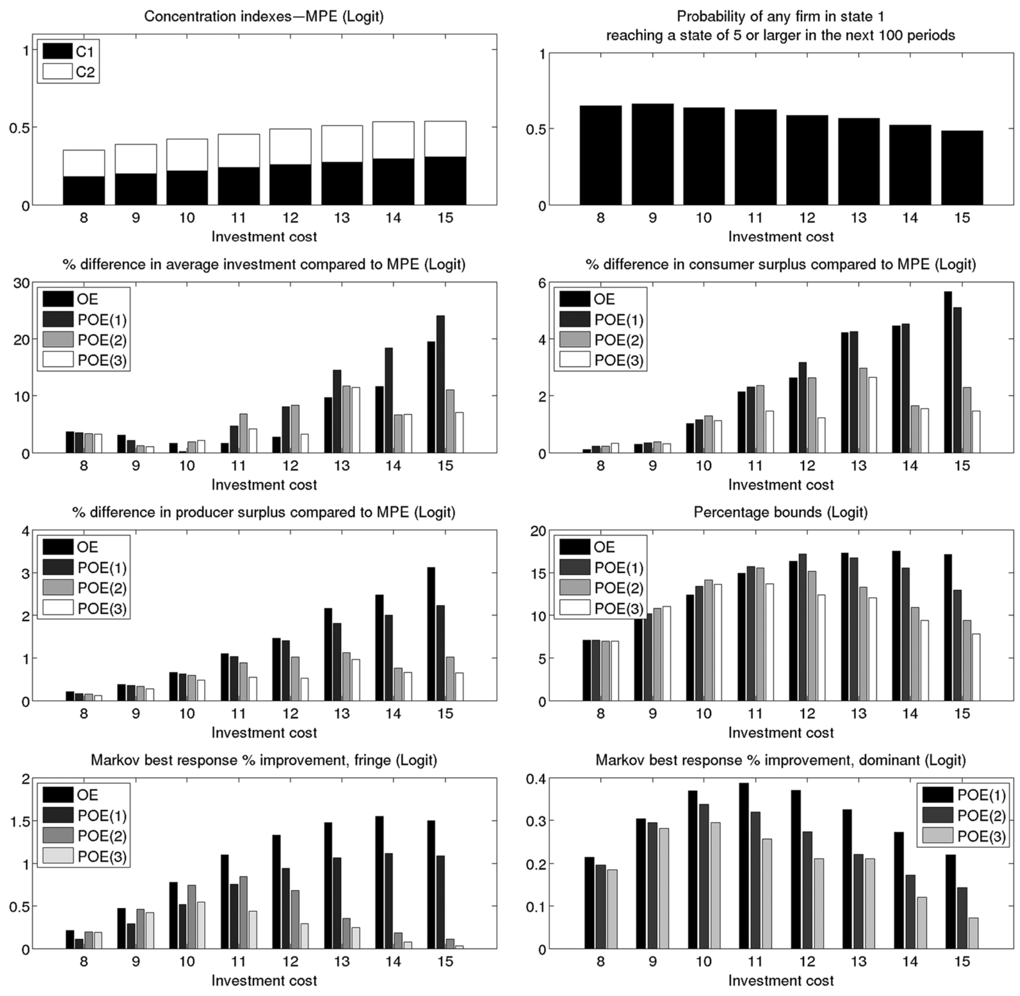
The second result is that adding dominant firms in the logit model almost always leads to results closer to MPE. In many cases the improvement is dramatic, though again even OE is close to MPE for this model, so in fact the overall differences between the three models remain small.

Because there are so few firms in this model, and because there is a difference in profits between fringe and leader firms, the error bounds are fairly large, ranging from a few percent to

[10] Pakes and McGuire (1994) found the same thing and Besanko et al. (2010) provide a detailed explanation of why this happens.

FIGURE 1

LONG-RUN INDUSTRY STATISTICS FOR OE AND POE VERSUS MPE, LOGIT MODEL $\delta = 0.2$



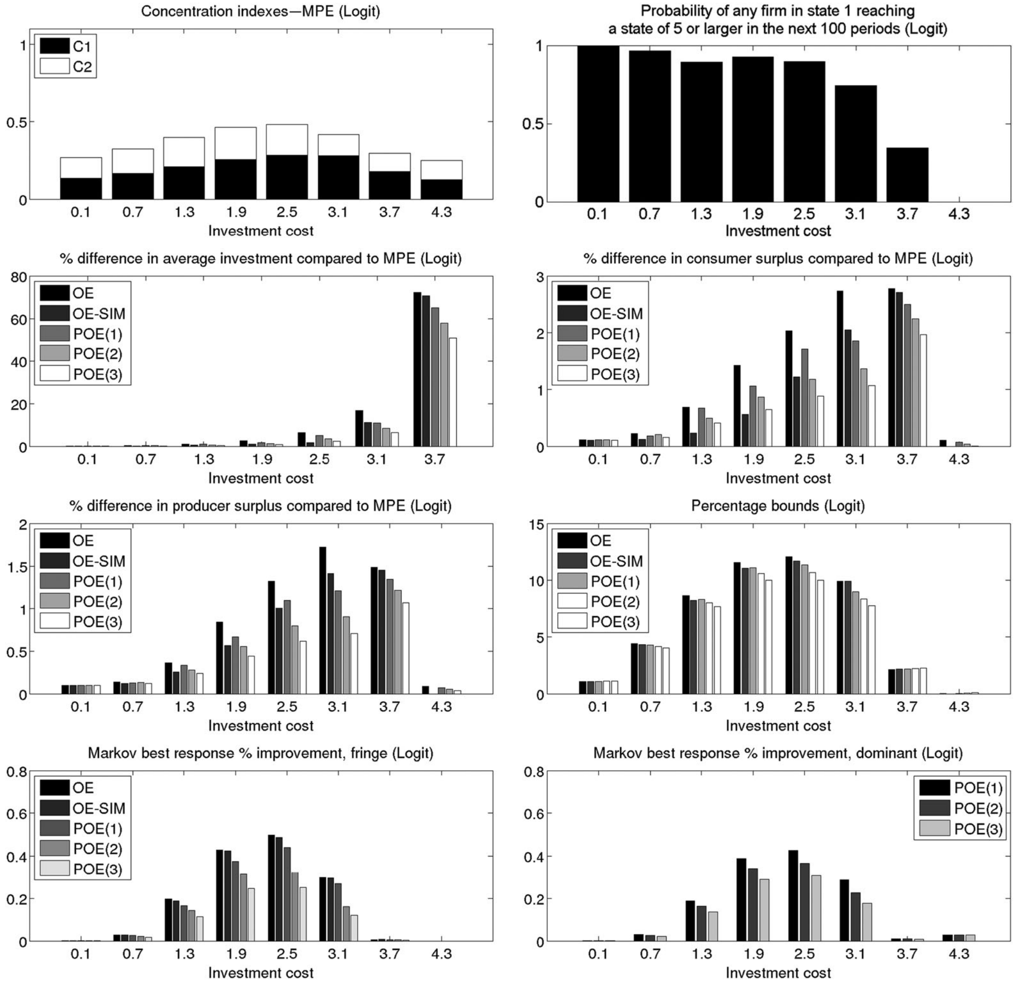
over 15%. However, as can be seen in the next two panels, the bounds are very loose because the fully informed Markov best response differences are in actuality quite small, typically on the order of less than 1%.

Finally, we also computed OE-SIM for the $\delta = 0.7$ case. For $\delta = 0.2$, computation of OE-SIM took too long to make its computation practical for many different parameter values, so we omitted it from the graphs. In general, OE-SIM does provide results closer to MPE than OE does, but typically the improvements are small. However, in rare cases it provides results closer to MPE than even POE does. Recall that when we compute POE we are not simulating out the integral over the fringe firms, so POE need not improve upon OE-SIM. As a reminder, OE-SIM provides an improvement only for markets with small numbers of firms, so what we are seeing is that in these particular markets there is enough variance in the market structure over time to make a small but noticeable difference in the equilibrium calculations.

Cournot model. Figures 3–5 provide the same set of comparisons for the Cournot model for three different depreciation rates, respectively ($\delta \in \{0.2, 0.5, 0.8\}$). The first thing to note about the results is that, in using this special modified Cournot model and these parameter values, we are

FIGURE 2

LONG-RUN INDUSTRY STATISTICS FOR OE AND POE VERSUS MPE, LOGIT MODEL, $\delta = 0.7$

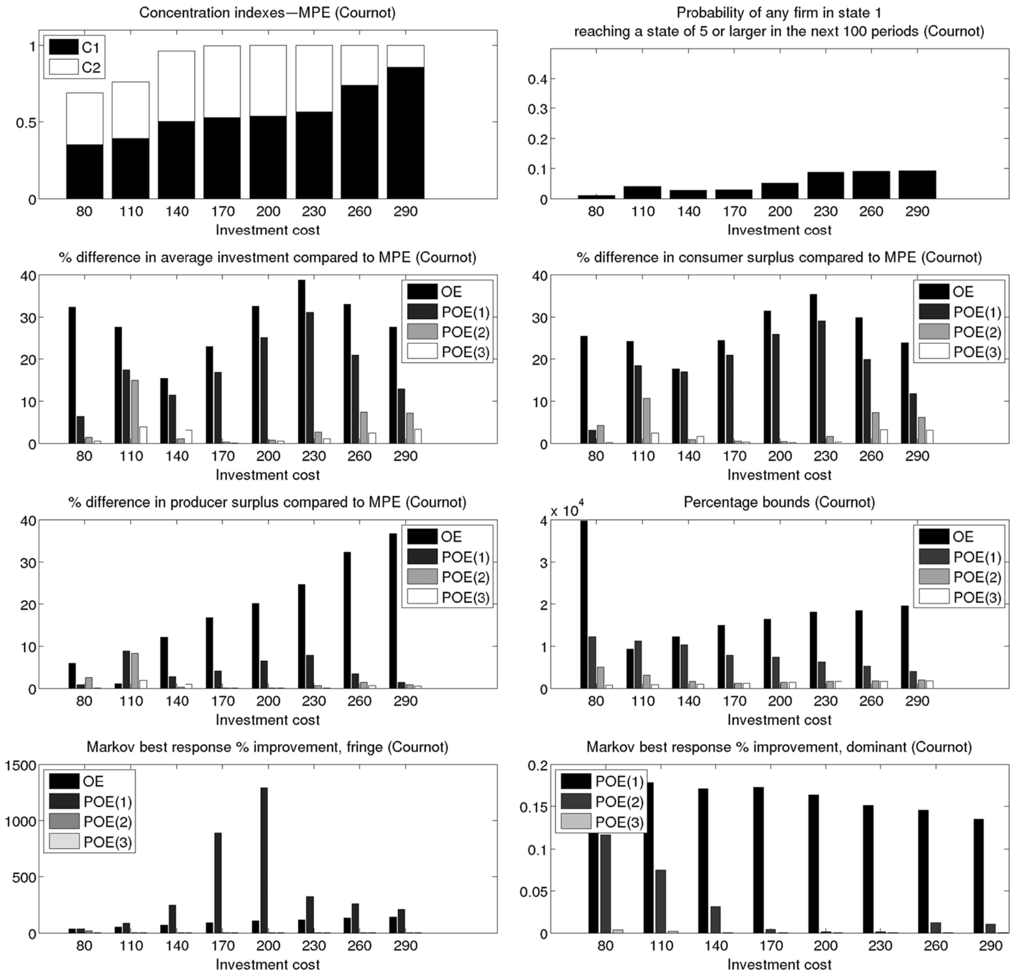


looking at very extreme examples of concentration. The top left panel of the figures shows that the markets in this model have one firm concentration ranging from 0.3 all the way up to near 1.0 as we increase the investment cost parameter. Meanwhile, C2 is never less than 0.5 and is typically larger than 0.9. These statistics would place these markets in the extreme tail of census industries. For comparison, in the 2007 US Census of Manufacturers, 98% of six digit NAICS manufacturing industries have C4 less than 0.90, and 99% have C4 less than 0.95. This extreme concentration results in OE failing to be close to MPE in many cases. For example, producer and consumer surplus can be as much as almost 40% different in the worst cases.

Nevertheless, even in these extreme cases, adding dominant firms generates equilibria that look much closer to MPE than OE does, in almost all cases. POE(3), for example, generates investment, consumer and producer surplus within 5% of MPE for all parameter values when $\delta = 0.2$ or $\delta = 0.5$. Best response improvements for dominant firms are always small for POE. Best response improvements for fringe firms are small as long as there are at least two dominant firms. These results suggest that firms playing POE(2) or POE(3) would not perceive any benefit from deviating to a fully informed strategy.

FIGURE 3

LONG-RUN INDUSTRY STATISTICS FOR OE AND POE VERSUS MPE, COURNOT MODEL, $\delta = 0.2$

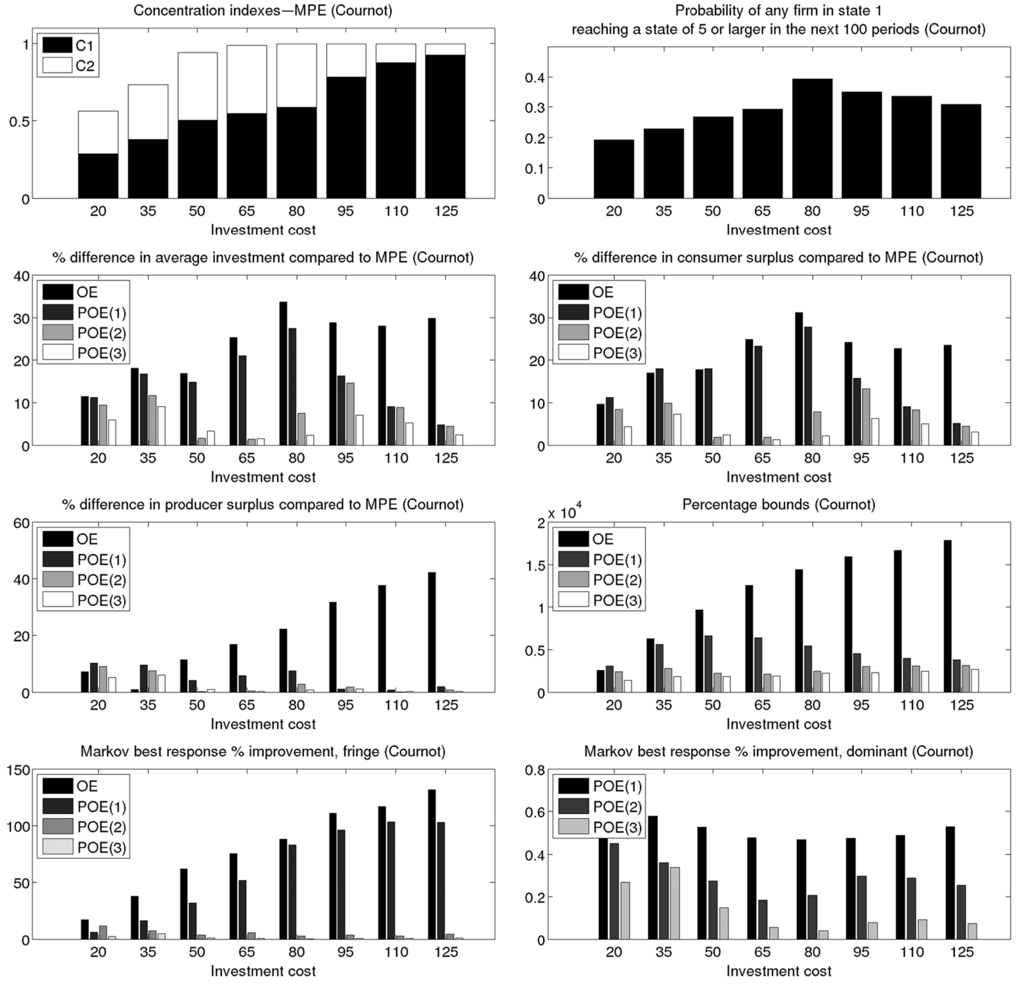


POE is not as close to MPE when $\delta = 0.8$. However, we view this level of depreciation as unrealistic. In addition to being unrealistic on its own, high depreciation also implies an unrealistically high rate of turnover among the leading firms. The top right panel of Figure 5 shows that when $\delta = 0.8$, the probability that a firm in state 1 will become a "leader" firm (defined as operating in the top half of the state space) within 100 periods is between 0.35 and 0.5. What this means is that more than one in every three of the very smallest firms in the industry will become a leader firm every 100 periods. At the same time, leader firms capture virtually all the market (top left panel). For comparison, in the data collected by Sutton (2007), in a 23-year period, only 18 of 45 industries saw the leader and second-place firm change places, whereas in the remaining 27 industries the leader held its place throughout. Our measure of turnover instead considers very small fringe firms becoming leader firms, a much rarer event.

We also have an intuition for these results. When there is some stability among leader firms, as in real-world industries, knowing only the leading firms' states (as is the case in POE) is sufficient for all firms to optimize well. However, when leader firms are very unstable it is helpful for firms to have more detailed information about the industry. The more stable the leading firms

FIGURE 4

LONG-RUN INDUSTRY STATISTICS FOR OE AND POE VERSUS MPE, COURNOT MODEL, $\delta = 0.5$



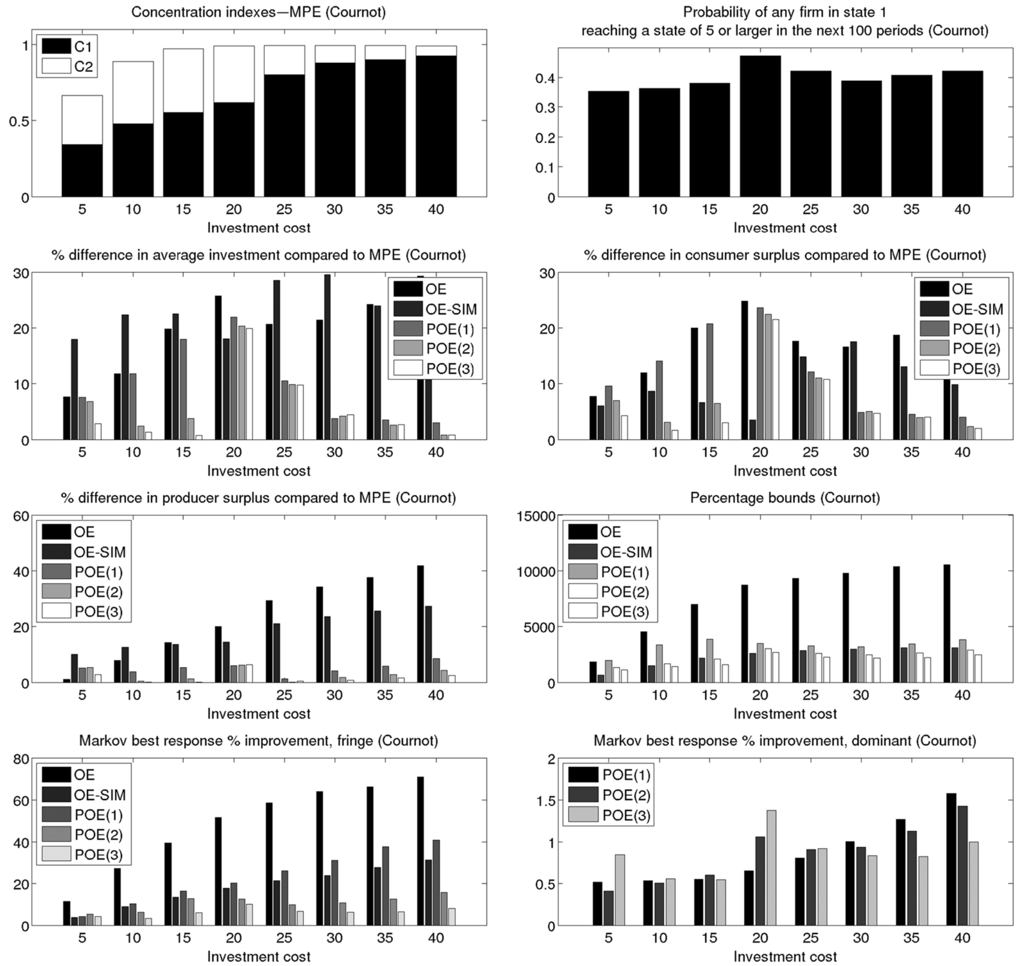
are, the closer are the POE markets to MPE, though the results in Figures 3–5 suggest that POE is close to MPE for a pretty wide range of markets.

Finally, the error bounds are again very high for the basic OE, and substantially lower but still high for the OE-SIM and POE models. Also, because there are such a small number of active firms in this model and firm turnover is so high, the bounds continue to be very loose in this model. Our expectation is that the error bounds would be more useful in less extreme examples.

Summary of findings. Summing up our findings: for the logit model all the equilibrium concepts are close to MPE, but OE-SIM and POE are closer to MPE than OE is. For the Cournot model, OE is not always close to MPE. However, adding dominant firms brings the equilibrium closer to MPE, and POE(3) is fairly close to MPE for cases where turnover among leading firms is similar to (or even well above) that of real-world industries. POE is not close to MPE when turnover among leading firms is very high, at levels that we view as strongly atypical.

FIGURE 5

LONG-RUN INDUSTRY STATISTICS FOR OE AND POE VERSUS MPE, COURNOT MODEL, $\delta = 0.8$



For these simple examples, MPE took more than 10 minutes to compute, whereas OE took less than one second, and POE with three dominant firms took about 10 seconds, demonstrating the trade-off between computational tractability and choice of equilibrium concept.[11]

□ **Comparison of fringe and dominant firms' dynamics.** Above, we compared the overall industry dynamics in OE and POE to those of MPE. We now look deeper into the underlying dynamics of individual firms in POE. For this section, we include entry and exit of fringe firms in the model to show that there is entry deterrence.

Figure 6 shows the long-run distribution of firm states for the Cournot model for OE, POE(1), POE(2), and POE(3), as well as the distribution of states broken down by firm type for dominant and fringe firms. We show only the results for one parameterization and one model here, but the results are qualitatively the same for all parameter values in both models.

What we can see in the figure is that even though both dominant firms and fringe firms have identical profit and cost parameters, in the POE model, dominant and fringe firms follow different dynamics. Dominant firms are in fact almost always dominant, and fringe firms are almost always

[11] Reported computation times are per core on a 2.5GHz AMD Opteron machine with 128GB of RAM using C++.

FIGURE 6

LONG-RUN DOMINANT/FRINGE STATE DISTRIBUTION, COURNOT, INVESTMENT COST=5

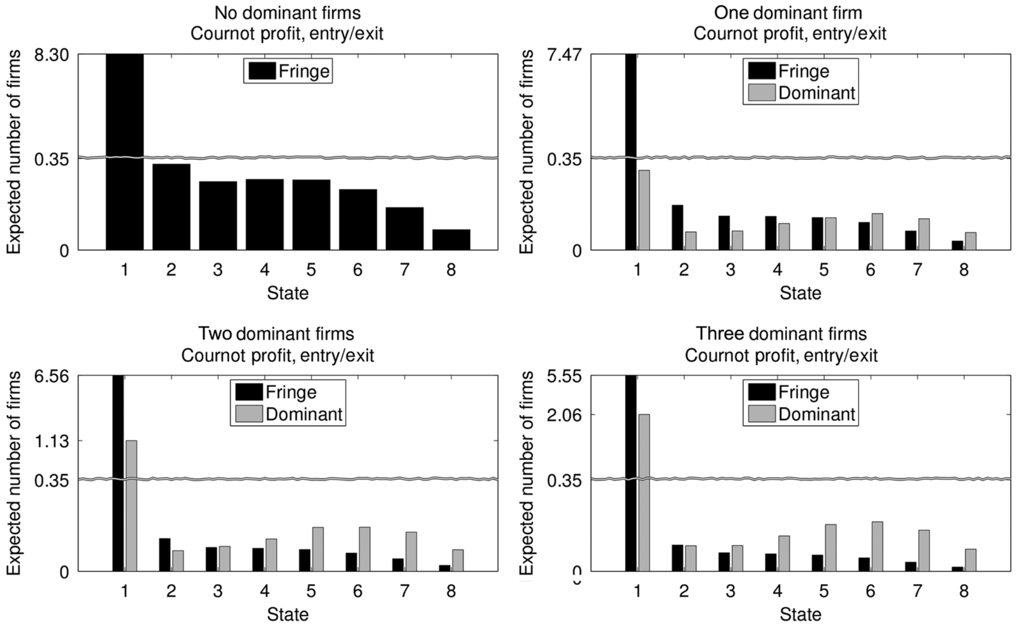
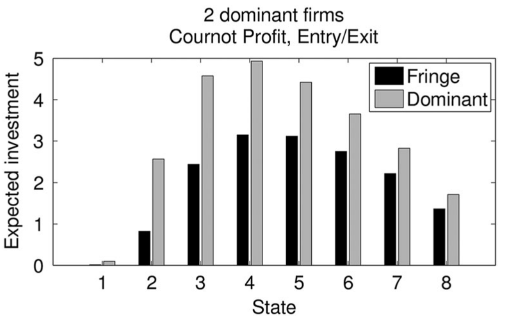


FIGURE 7

LONG-RUN DOMINANT/FRINGE INVESTMENT, POE(2), COURNOT, INVESTMENT COST = 5



small. Figure 7 further shows that this comes about because dominant firms invest more in every state.

This behavior is a consequence of the information structure of the POE model. Because dominant firm states are tracked by all firms whereas fringe firm states are not, dominant firms have stronger investment incentives than fringe firms. Dominant firms know that, if they invest heavily, all other firms will know this and will react to it. This allows them to deter investment and deter entry, and both are a frequent outcome in the POE model. Fringe firms cannot deter entry (at least not directly) because if they were to invest more at a particular state, no other firm would know about this (except indirectly through the equilibrium expected distribution of industry states).

These results highlight the importance of the information structure used when computing equilibria, not just in our context but more generally. Even if it is the case that the true MPE strategies can be closely approximated by a simpler strategy such as the one used here, in the sense that if you projected MPE strategies on a simplification you would obtain a close relationship, imposing that simpler structure in equilibria computation can have unforeseen implications. When the simpler information structure is enforced in equilibrium, firms know for sure that their competitors must act according to this information structure, and their equilibrium strategies may exploit this in ways that were not present in the original model.

In the case of POE, we think that the information structure in the model is somewhat realistic for many concentrated markets, and the behavior of dominant firms in the model may also be realistic. For example, if Microsoft were to enter with a new type of software (say, a browser), it makes sense that competing firms would pay much more attention than if a new entrant were to enter with an otherwise identical product. However, we do want to emphasize that the individual firm dynamics are different from those in MPE, even though the aggregate behavior of the industry can be quite similar as we showed above in our numerical experiments. Furthermore, these results suggest that one should view arbitrary simplifying assumptions with caution in the context of equilibrium computation, as arbitrary simplifying assumptions can generate behavior that is different from MPE behavior in unexpected, unknown, and perhaps undesirable ways.

6. An empirical model

■ In this section, we investigate the properties of POE in a more complex empirical model due to CW. CW examines the effect of demand uncertainty on industry structure and sunk costs in the ready-mix concrete industry. A primary feature of his model is an aggregate demand shifter that follows a Markov process. Also, many of his markets are highly concentrated. In his counterfactual experiments he computes MPE using the stochastic approximation algorithm due to Pakes and McGuire (2001).

□ **Model overview.** In the CW model, there are a fixed set of N firms, with each firm having a controlled state variable x_{it} that represents whether it is active and, if so, its current and past size. In the article firm size is discretized to three values, $\{$small, medium, large$\}$, and x_{it} can take on seven possible values: $\{$small, small/medium, small/large, medium, medium/large, large, $\varnothing\}$. The notation "small/large" represents a firm that is currently small but was large at some point in the past. Similarly, the state "small/medium" represents a firm that is currently small but was previously medium but never large, and the state "small" represents a firm that is currently small and has never been either medium or large. The reason for tracking past firm size is that the sunk costs of changing size are allowed to differ depending on past size, reflecting the fact that it may be easier for a firm to grow if it was large previously, than if it was not. The state $\varnothing$ corresponds to being out of the market. In our version of the model we will also experiment with finer levels of discretization, allowing for more than three values for current firm size. We discuss that extension further below.

There is also an aggregate demand state M_t that follows an exogenous Markov process. The level of the aggregate state and the firm size states are assumed to be known by all firms. Each firm also has a set of private information logit state variables that are outlined below.

In each period, each incumbent firm chooses its next period's size from the set $\{$small, medium, large, $\varnothing\}$. Given past size, this is equivalent to choosing $x_{i,t+1}$ from the appropriate feasible set. Choosing $\varnothing$ is choosing to exit the market, where exit is irreversible.

Payoffs are given by

$$R(x_{i,t+1}, x_{-i,t+1}, M_{t+1}) - \tau(x_{i,t+1}, x_{it}) + \sigma\epsilon^{t}_{ia},$$

where $x_{-i,t+1}$ represents the vector of size states of firm i's competitors, $R(\cdot)$ represents the firm's current revenue function and is defined below, ϵ^{t}_{ia} is an i.i.d choice-specific preference shock that is private information to the firm, the parameter σ scales the variance of the error term, and τ

represents the sunk adjustment cost of changing from size x_{it} to $x_{i,t+1}$. (τ is assumed to be zero if $x_{i,t} = x_{i,t+1}$.) Note that the different actions/choices a correspond to the next state $x_{i,t+1}$. Current revenues are zero if the firm exits the market.

The model also allows for firm entry. Each period, for every inactive firm (there are always N firms including incumbents and potential entrants) a potential entrant appears who has the same payoff function as incumbent firms above, with sunk costs of entry captured by the τ function. Potential entrants are short lived, so that if they choose not to enter they cease to exist and a new potential entrant appears in their place the next period.

In the CW model, the revenue function is assumed to take the following form:

$$R(x_{i,t+1}, x_{-i,t+1}, M_{t+1}) = \sum_{\alpha \in \{sm,med,lg\}} 1\{x_{i,t+1} = \alpha\}\big(\theta_1^\alpha + \theta_2^\alpha M_{t+1} + \theta_{31}^\alpha \{NComp >= 1\} + \theta_{32}^\alpha * \{NComp > 1\} * \log(NComp - 1)\big), \tag{6.1}$$

where

$$NComp = \sum_k (x_{-i,t+1}(k) \neq \emptyset)$$

is the number of active competitors. The parameter θ_1^α is a fixed cost parameter, θ_2^α is the coefficient on aggregate demand, and θ_{31}^α and θ_{32}^α measure the effect of competition on firm profits.

There are two main differences in defining POE for this problem relative to those above: private information and aggregate shocks. In Section 3, we define a partially oblivious strategy as a function of the firm's own extended state $\bar{x}_{it}$ and a set of statistics about the dominant firms w_t. In the CW model, we need to extend this definition to include private information and the aggregate shocks. We define a *partially oblivious strategy* for the CW model as $\mu(\bar{x}_{it}, w_t, z_t, \epsilon_i^t)$, where z_t contains the current state of an aggregate shock as well as a finite set of statistics that summarize its history. A finite history may be useful in improving firms' beliefs about the current distribution of fringe firms because recent past values of the aggregate state may strongly influence this distribution (Weintraub, Benkard, and Van Roy, 2010).

For all $(\bar{x}, w, z)$ we define a *partially oblivious value function* as

$$\tilde{V}(\bar{x}, w, z|\mu', \mu) = E_{\mu',\mu}\left[\sum_{k=t}^{\tau_i} \beta^{k-t}[R(x_{i,k+1}, \tilde{f}_{-i}(w_{k+1}, z_{k+1}|\mu), y_{-i,k+1}, M_{k+1}) - \tau(x_{i,k+1}, x_{ik}) + \sigma\epsilon_{ia_{ik}}^k] | \bar{x}_{it} = \bar{x}, w_t = w, z_t = z\right], \tag{6.2}$$

where $y_{-i,k+1}$ represents the state of dominant competitors of firm i and we slightly abused notation to write R as a function of this vector and the conditional expected fringe state. Note that this formulation automatically handles entry and exit of both dominant and fringe firms through transitioning in and out of the inactive state.

For any w and z, the expected industry state $\tilde{f}(w, z|\mu)$ can be obtained by using a modified set of balance equations, similar to those discussed in Section 3 and described in detail in Appendix A Balance Equations, but that incorporate the aggregate shocks. These modified balance equations are described in Appendix A Balance Equations.

In practice, it is convenient to work with *choice-specific partially oblivious value functions* defined as:

$$\tilde{v}(a, \bar{x}, w, z|\mu', \mu) = E\big[R(a, \tilde{f}_{-i}(w_{t+1}, z_{t+1}|\mu), y_{-i,t+1}, M_{t+1}) - \tau(a, x_{it}) + \beta\tilde{V}(\bar{x}_{i,t+1}, w_{t+1}, z_{t+1}|\mu', \mu)|\bar{x}_{it} = \bar{x}, w_t = w, z_t = z\big]. \tag{6.3}$$

TABLE 2 Parameters of the Revenue Function in the CW Model

	Current Size Small	Medium	Large
Fixed cost (θ_1^α)	−139.00	−244.00	−285.00
Demand shifter (θ_2^α)	20	35	45
First competitor (θ_{31}^α)	−48.00	−58.00	−63.00
More competitors (θ_{32}^α)	−17.00	−44.00	−48.00

Note: The Standard Deviation of the Preference Shock σ is set to 133.

In the notation in equation (6.3), the firm takes action a in the first period and then follows strategy μ' in all later periods. In case of exit, or if the potential entrant decides to stay out, $\tilde{v}(a^E, \bar{x}, w, z|\mu', \mu) = 0$, where a^E is the action to go to an inactive state.

The POE *ex ante* value function and POE strategy μ solve the following fixed point equation:

$$\tilde{V}(\bar{x}, w, z|\mu, \mu) = E_{\epsilon_i^t}\left\{\max_a \tilde{v}(a, \bar{x}, w, z|\mu, \mu) + \sigma\epsilon_{i,a}^t\right\}, \ \forall(\bar{x}, w, z).$$

If private information ϵ is distributed as type-1 extreme value, one can obtain an *ex ante* best response value function solving the following Bellman equation:

$$\tilde{V}(\bar{x}, w, z|\mu', \mu) = \sigma \log\left[\sum_a \exp\left(\sigma^{-1}\tilde{v}(a, \bar{x}, w, z|\mu', \mu)\right)\right] + \sigma\gamma, \tag{6.4}$$

where γ is Euler's constant. The transition probabilities of firms take a simple logit form as well:

$$P[a_{it} = a|\bar{x}_{it}, w_t, z_t; \mu', \mu] = \frac{\exp\left(\sigma^{-1}\tilde{v}(a, \bar{x}_{it}, w_t, z_t|\mu', \mu)\right)}{\sum_{a'} \exp\left(\sigma^{-1}\tilde{v}(a', \bar{x}_{it}, w_t, z_t|\mu', \mu)\right)} \tag{6.5}$$

In this case, the maximization step of the POE algorithm is equivalent to iterating on equations (6.3) and (6.4) until convergence for a given $\tilde{f}(w, z|\mu)$ and μ. In the limit, the procedure produces choice-specific optimal partially oblivious value functions as well as optimal partially oblivious strategies μ'. The expected industry state can be subsequently updated using equation (6.5) and balance conditions (C2).

□ **Results.** We first consider an exact replica of the CW model used in his article.[12] Parameter values for the revenue function are listed in Table 2 and the sunk adjustment costs parameters are listed in Table 3. CW uses four different categories of markets based upon market size. In this exercise, we use the estimated parameters for market category three, which has an average of 4.65 firms in long-run equilibrium. We also solved the model for market category two, with qualitatively identical results.

Table 4 compares OE, POE, and MPE for the CW model with three size levels and 10 potential firms. As above, in implementing POE we use only the current values of dominant firms' states and of the aggregate shocks. We experimented with using longer histories, but it made no difference to the results. It is not possible to compute welfare for this model, so we instead report long-run average statistics that describe the equilibrium industry dynamics, including the firm size distribution, firm turnover and growth rates, and also several measures of sunk costs incurred by firms in the equilibrium, including sunk costs of changing size and sunk costs of entry.

[12] Our reported statistics vary from those reported in CW because we report long-run averages whereas the statistics reported in the article represent the average across many markets (of different sizes) of short-run simulations starting from a (different) particular observed state for each market. The starting points are confidential data that we cannot access. We have verified separately with Allan Collard-Wexler that our MPE policy and value functions are identical to the ones that he computed for the article.

TABLE 3 Parameters of the Adjustment Cost in the CW Model

	Target Size		
Past state	Small	Medium	Large
Small	0.00	−332.00	−1809.00
Small, past medium	0.00	−772.00	−608.00
Small, past large	0.00	−325.00	−343.00
Medium	−107.00	0.00	101.00
Medium, past large	−314.00	0.00	43.00
Large	−254.00	−403.00	0.00
Inactive	−1002.00	−2000.00	−1771.00

TABLE 4 Industry Statistics Across Different Equilibria, CW Model

	MPE	OE	POE(1)	POE(2)	POE(3)	POE(4)	OE-SIM
Number of active firms	4.65	4.37	4.37	4.37	4.39	4.43	4.92
Number of small firms	2.67	2.51	2.52	2.53	2.55	2.57	2.83
Number of medium firms	0.74	0.69	0.69	0.69	0.69	0.70	0.77
Number of large firms	1.24	1.17	1.16	1.15	1.15	1.16	1.32
Number of entrants/exitors	0.13	0.12	0.12	0.13	0.13	0.13	0.12
Entry costs	127.26	121.49	124.59	126.30	127.18	127.28	123.36
Transition costs	187.28	175.61	174.08	173.69	176.71	176.90	200.18
Growth out of small	0.30	0.28	0.28	0.28	0.29	0.29	0.32
Growth into large	0.31	0.29	0.29	0.29	0.30	0.30	0.33

We find that, for this model, all statistics for OE are within 6% of those for MPE, so that OE is close to MPE but not exactly the same. In general, there are about 6% fewer firms and most other statistics reflect this same difference. The exceptions are that OE entry and exit rates and rates of size transition are nearly identical to those statistics in MPE.

When we add a single dominant firm, entry costs move closer to MPE levels, but other statistics do not change much relative to OE. Adding a second dominant firm has a similar effect. With four dominant firms, all statistics move closer to MPE, and entry costs are now nearly identical to MPE. Summarizing the results, OE matches MPE fairly closely in general, but is not exactly the same, and POE further improves the match for all statistics, but some small differences remain even with four dominant firms.

The simulated version of OE is substantially different from OE and actually lies on the other side of MPE from OE. For example, though OE has fewer firms and fewer transition costs than MPE, OE-SIM has more firms and more transition costs than MPE. The net result is that OE-SIM is different from OE but about the same distance away from MPE. To us, the differences do not seem worth the extra computation cost associated with OE-SIM.

Of course, the benefit of using POE over MPE in a model like this one is that it is much easier to program and compute. Using the specification above, for any number of active firms OE takes a few seconds to compute whereas POE(1) and POE(2) take a few minutes and POE(3) and POE(4) take a few hours. Meanwhile, 7-firm MPE takes days, 10-firm MPE takes months, whereas 12-firm MPE is infeasible to compute.[13]

□ **Exploring the level of discretization.** Because MPE is so difficult to compute, in implementing MPE it is often necessary to make many simplifications for computational tractability. For example, one simplification that is commonly used, and that is used in the CW model, is to coarsely discretize the state space. Because OE and POE are much lighter computationally than

[13] Reported computation times are per core on a 2.5GHz AMD Opteron machine with 128GB of RAM using C++. We were able to compute a 10-firm MPE in a reasonable time using massive parallelization.

TABLE 5 Parameters of the Adjustment Cost in the CW Model—Interpolation

	Target Size		
Past State	Small	Medium	Large
Small	0.00	−712.58	−1453.2
Small, past medium	0.00	−414.1	−844.51
Small, past large	0.00	−103.85	−211.79
Medium	−107.00	0.00	−430.42
Medium, past large	−200.48	0.00	−107.94
Large	−408.86	−208.38	0.00
Inactive	−1212.9	−1586.1	−1974

MPE, it is possible to explore discretizing the model into much finer grids. Such a generalization may be particularly interesting in the CW model because a finer grid allows firms to optimize their size more closely, and also because in reality, firms are not constrained to a three-point grid so a finer discretization may potentially be a better match to the observed data.

There are two complications in making the discretization of the model finer in the CW model. First, we require adjustment cost and entry cost parameters for each level of the finer grid, but we do not have access to the confidential US Census data used to estimate the CW model so we cannot estimate these new parameter values directly. To solve this problem, we instead approximate the estimated parameters obtained in CW (listed in Table 3) using a linear function of the state, and then interpolate the parameter values for the finer grid.[14] Table 5 lists the implied values of the resulting interpolated parameters for a grid of size four (with three firm size states plus out) for comparison. The interpolated parameters are similar to the estimated parameters for a grid of size four. We use the interpolated parameters for all grid sizes in this section to avoid introducing differences due to different adjustment costs.

The second complication in making the discretization of the model finer is that, as there is an idiosyncratic shock for each possible size choice in the model, there is no way to make the grid finer without adding more logit error terms to the firm's decision problem. These additional error terms alter the underlying economics of the model because, in a finer discretization, incumbent firms have more size choices available to them and hence more logit draws to choose from each period. Of course, the maximum of a larger set of draws is going to be larger, which would imply that the value of being active should go up as we make the grid finer. Thus, if we hold all parameters constant and simply make the grid finer, we find that there are more and more active firms. For our purposes, we would like to hold the firm payoffs approximately constant when we change the grid size to keep the economics of the model similar. Thus, in order to counteract the effect of the additional error terms, we make a downward adjustment to the revenue function for active firms (equally for all sizes) so as to keep the number of active firms the same for all grid sizes.

Figure 8 and Table 6 show statistics for POE(1) for five grid sizes: $\{3,6,9,12,15\}$. For these grid sizes, there are approximately 350,000, 38 million, 890 million, 9.6 billion, and 66 billion states, respectively.

We find that there are some large differences in the results as we make the size grid finer. The main differences are that on the three-point grid there are equal numbers of small and medium firms, but on the finer grids there are nearly twice as many medium firms as small ones. Furthermore, transition costs get much larger as firms are now changing size more often and thus paying more transition costs. This last effect happens because the coarse grid size is acting like a large adjustment cost that prevents small movements in size. On the finer grid, firms make more small changes to their size. Entry and exit rates also increase on the finer grids, by a factor of about 50%. The number of firms is being held constant so the higher entry and transition costs can be interpreted as higher costs per firm.

[14] Appendix B provides the details of this interpolation.

FIGURE 8

FIRM SIZE DISTRIBUTION AS WE CHANGE GRID SIZE, CW MODEL

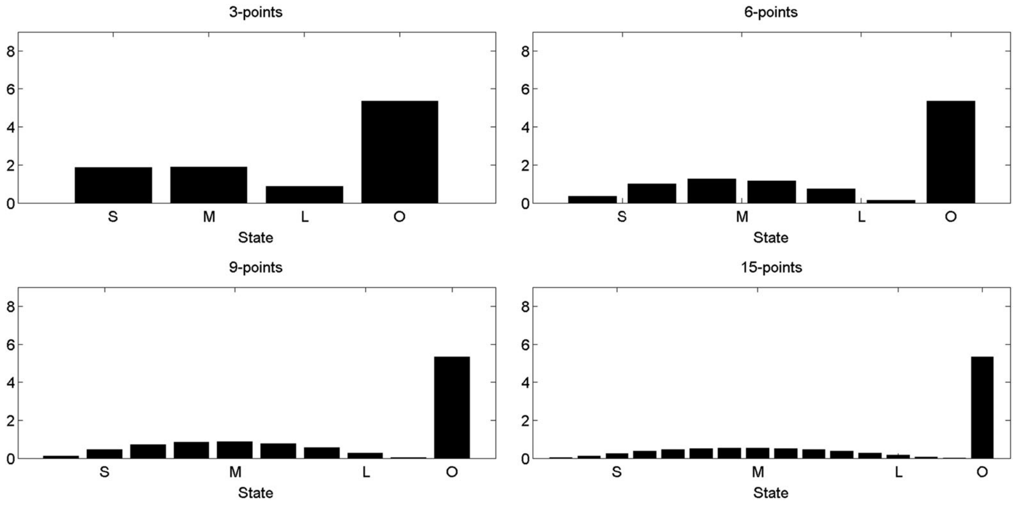


TABLE 6 Industry Statistics as we Change Grid Size, CW Model

	3 Points	6 Points	9 Points	12 Points	15 Points
Number of active	4.65	4.64	4.65	4.65	4.65
Number of small firms	1.87	1.34	1.28	1.26	1.26
Number of medium firms	1.90	2.42	2.48	2.50	2.50
Number of large firms	0.88	0.88	0.89	0.89	0.89
Number of entrants/exitors	0.022	0.024	0.028	0.031	0.034
Entry Costs	31.43	34.63	37.00	38.93	40.55
Transition costs	255.51	327.95	340.50	345.71	348.21
Avg. firm size	12.44	12.92	12.90	12.86	12.85

Once the grid size hits about nine points, however, the results begin to stabilize, suggesting that nine size points may be a fine enough grid to capture the dynamics of firm-size movements in this model. Note that the state space for the nine-point grid is approximately 2500 times as large as that for the three-point grid, making computation of MPE around 2500 times more difficult. For POE, the nine-point grid is only seven times more difficult to compute than the three-point grid.

Importantly, our point is not to say whether the finer grid is more or less reflective of the concrete industry. We could not determine this without reestimating the model for different grid sizes and evaluating the fit of the model, something that we cannot do without access to the original data. Moreover, the effect of the additional logit shocks as we make the grid finer makes such a comparison difficult. Our point is instead one about the empirical application of dynamic oligopoly models in general. In empirical applications, it is often necessary to make modelling simplifications that are purely for computational reasons. These results demonstrate that such simplifications can have an impact on the economics of the model, and therefore there is a benefit to having the ability to explore richer versions of the economic model.

We believe that empirical researchers are thus faced with a trade-off. They can either compute a simple equilibrium concept such as POE in a richer model, or compute exact MPE in a simpler one, but could not compute MPE in the richer model. In either case, the results will likely not be exactly equal to MPE in the richer model. Sometimes the economics of the model may be changed less by using OE or POE in place of MPE than they would by simplifying the model to facilitate computation of MPE. In other cases, it may be worth computing MPE even with the additional modelling restrictions that requires.

7. Conclusions

■ In this article we considered the application of oblivious equilibria to highly concentrated markets. We defined an extended notion of oblivious equilibrium that we call *partially oblivious equilibrium* (POE) that allows for there to be a set of "dominant firms," whose firm states are always monitored by every other firm in the market. Such a model can be motivated in two different ways: either as an approximation to a full information model, or as an appealing behavioral model on its own.

We then explored the behavior of POE relative to OE and MPE in a wide variety of markets through computational experiments. Summarizing the results, we found that OE was surprisingly close to MPE even in many highly concentrated markets. In extremely concentrated markets OE was not close to MPE, but POE generally was close to MPE as long as turnover among the leading firms was not too high. The results suggest that these tools could be useful in a wide variety of empirical applications.

We also applied POE to the empirical model of CW and demonstrated a trade-off between implementing an equilibrium concept that is computationally light (such as POE) in a richer economic model, and implementing MPE in a simpler model. Empirical work in this area often pushes the boundaries of what is possible computationally, and thus we think that this is a trade-off that would often be faced in applied work.

We also caution that we found that the behavior of dominant firms and fringe firms differ in the POE model. In some markets, this difference may be realistic but in others it may not be. The finding also suggests that researchers should use caution in specifying the information structure of their models more generally. Reducing the state space of the model through arbitrary informational assumptions could lead to unrealistic outcomes in the model.

Appendix A

□ **Error bounds and proofs.** In this section, we provide several specifications for the functions H_k in our error bounds. For this purpose, with some abuse of notation, let $\Delta_A(f_{-i}, w) = \sup_{x\in A}(\pi(x, f_{-i}, w(1)) - \pi(x, \tilde{f}_{-i}(w), w(1)))$ and let $\Delta(x, f_{-i}, w) = \pi(x, f_{-i}, w(1)) - \pi(x, \tilde{f}_{-i}(w), w(1))$. Also, let $\underline{x}(k,t) = [x - (k-t)\overline{h}]^+$. We have the following result that we prove at the end of this section.

Theorem A1. Let Assumptions 2.1, 3.1, 2.2, and 2.3 hold. Then, for any POE $(\tilde{\mu}, \tilde{\lambda})$, and fringe firm i with state $\bar{x} = (x, 0), x \in \mathbb{N}$,

$$E\left[\sup_{\mu'\in\mathcal{M}_p} V(\bar{x}, f_t, w_t|\mu', \tilde{\mu}, \tilde{\lambda}) - V(\bar{x}, f_t, w_t|\tilde{\mu}, \tilde{\lambda})\right] \leq \sum_{k=t}^{\infty}\beta^{k-t}E\left[\left[\Delta_{\{\underline{x}(k,t),\ldots,x+(k-t)\overline{h}\}}(f_{-i,k}, w_k)\right]^+\right]$$
$$+ E\left[E_{\tilde{\mu},\tilde{\lambda}}\left[\sum_{k=t}^{\tau_i}\beta^{k-t}\left(\pi(x_{ik}, \tilde{f}_{-i}(w_k), y_k) - \pi(x_{ik}, f_{-i,k}, y_k)\right)\bigg|x_{it} = x, f_{-i,t}, w_t\right]\right]. \quad \text{(A1)}$$

Suppose that, for all $f \in \overline{\mathcal{F}}$ and $w \in \mathcal{W}$, the function $\Delta(x, f_{-i}, w)^+$ is nondecreasing in x. Then, for any POE $(\tilde{\mu}, \tilde{\lambda})$, and fringe firm state $\bar{x} = (x, 0), x \in \mathbb{N}$,

$$E\left[\sup_{\mu'\in\mathcal{M}_p} V(\bar{x}, f_t, w_t|\mu', \tilde{\mu}, \tilde{\lambda}) - V(\bar{x}, f_t, w_t|\tilde{\mu}, \tilde{\lambda})\right] \leq \sum_{k=t}^{\infty}\beta^{k-t}E_{\hat{\mu},\tilde{\mu},\tilde{\lambda}}\left[\Delta(x_{i,k}, f_{-i,k}, w_k)^+\Big|x_{i,t} = x\right]$$
$$+ E\left[E_{\tilde{\mu},\tilde{\lambda}}\left[\sum_{k=t}^{\tau_i}\beta^{k-t}\left(\pi(x_{ik}, \tilde{f}_{-i}(w_k), y_k) - \pi(x_{ik}, f_{-i,k}, y_k)\right)\bigg|x_{it} = x, f_{-i,t}, w_t\right]\right], \quad \text{(A2)}$$

where $\hat{\mu}$ denotes a strategy in which the fringe firm never exits the industry and invests an infinite amount at every state.

The strategy $\hat{\mu}$ in the second bound is used only to generate a stochastic process for the firm state that increases at the maximum possible rate. In our model, under $\hat{\mu}$ the firm would invest $\bar{\iota}$, the maximum possibly amount (see Assumption

2.2 every period. Under the assumption that Δ is nondecreasing in x, this stochastic process generates the largest possible difference in discounted profits. The second bound is tighter than the first when Δ is nondecreasing in x because rather than compute the bound at the maximum difference in profits in all periods (as in the first bound), the second bound recognizes that it may be technologically infeasible for the firm to achieve the maximum difference in profits until some time in the future, and therefore due to discounting, the bound may be made smaller after accounting for this delay.

Note that relative to the notation introduced in Section 4, the functions H_k for the first bound correspond to:

$$H_k\left(\Delta(f_{-i,k}, w_k)\right) = \left[\Delta_{\{\underline{x}(k,t),\ldots,x+(k-t)\overline{h}\}}(f_{-i,k}, w_k)\right]^+ + \left(\pi(x_{ik}, \tilde{f}_{-i}(w_k), y_k) - \pi(x_{ik}, f_{-i,k}, y_k)\right) .$$

The functions are random because of their implicit dependence on the random variable x_{ik}. For the second bound, the functions correspond to:

$$H_k\left(\Delta(f_{-i,k}, w_k)\right) = \Delta(x_{i,k}, f_{-i,k}, w_k)^+ + \left(\pi(x_{ik}, \tilde{f}_{-i}(w_k), y_k) - \pi(x_{ik}, f_{-i,k}, y_k)\right) .$$

Again, the functions are random because of their implicit dependence on the random variables x_{ik}; in this case, in the first term this process is controlled by strategy $\hat{\mu}$ and in the second term by strategy $\tilde{\mu}$.

These bounds can be easily computed via simulation by recalling that (f_k, w_k) is distributed according to the invariant distribution for all $k \geq 0$. Hence, they require sampling states $(f_{-i,t}, w_t)$ according to this invariant distribution (for fixed $x_{it} = x$), and then simulating industry trajectories starting from that state.

Proof of Theorem A.1. Let

$$\mu^*(\bar{x}) = \begin{cases} \tilde{\mu}(\bar{x}) & \text{if } \bar{x} = (x, 1) \\ \mu^*(\bar{x}) & \text{if } \bar{x} = (x, 0), \end{cases}$$

be an optimal Markovian (nonoblivious) best response μ^* to POE $(\tilde{\mu}, \tilde{\lambda})$ for fringe firms, keeping the POE strategy of the dominant firm unchanged.

Hence, $\mu^* \in \mathcal{M}_p$ is such that

$$\sup_{\mu' \in \mathcal{M}_p} V(\bar{x}, f, w|\mu', \tilde{\mu}, \tilde{\lambda}) = V(\bar{x}, f, w|\mu^*, \tilde{\mu}, \tilde{\lambda}),\ \forall \bar{x} = (x, 0), x \in \mathbb{N}, f, w,$$

where $\mathcal{M}_p$ was defined in Section 4 as the set of extended Markov strategies.

Take any state of the fringe $\bar{x} = (x, 0)$, $x \in \mathbb{N}$. We have that:

$$\begin{aligned} E[V(\bar{x}, f_t, w_t|\mu^*, \tilde{\mu}, \tilde{\lambda}) - V(\bar{x}, f_t, w_t|\tilde{\mu}, \tilde{\lambda})] &= E[V(\bar{x}, f_t, w_t|\mu^*, \tilde{\mu}, \tilde{\lambda}) - \tilde{V}(\bar{x}, w_t|\tilde{\mu}, \tilde{\lambda})] \\ &+ E[\tilde{V}(\bar{x}, w_t|\tilde{\mu}, \tilde{\lambda}) - V(\bar{x}, f_t, w_t|\tilde{\mu}, \tilde{\lambda})]. \end{aligned} \tag{A3}$$

First, let us bound the first term in the right-hand side of the previous equation.

Because $\tilde{\mu}$ and $\tilde{\lambda}$ attain a POE, for all $\bar{x} = (x, 0)$, w,

$$\tilde{V}(\bar{x}, w|\tilde{\mu}, \tilde{\lambda}) = \sup_{\mu' \in \tilde{\mathcal{M}}_p} \tilde{V}(\bar{x}, w|\mu', \tilde{\mu}, \tilde{\lambda}) = \sup_{\mu' \in \mathcal{M}_p} \tilde{V}(\bar{x}, w|\mu', \tilde{\mu}, \tilde{\lambda}),$$

where the last equation follows because there will always be an optimal POE strategy when optimizing a partial oblivious value function even if we consider Markovian strategies that keep track of the full industry state. It follows that,

$$\begin{aligned} &V(\bar{x}, f, w|\mu^*, \tilde{\mu}, \tilde{\lambda}) - \tilde{V}(\bar{x}, w|\tilde{\mu}, \tilde{\lambda}) \leq \\ &E_{\mu^*,\tilde{\mu},\tilde{\lambda}}\left[\sum_{k=t}^{\tau_i} \beta^{k-t}\left(\pi(x_{ik}, f_{-i,k}, y_k) - \pi(x_{ik}, \tilde{f}_{-i}(w_k), y_k)\right)\middle| x_{it} = x, f_{-i,t} = f_{-i}, w_t = w\right] \\ &= \sum_{k=t}^{\infty} \beta^{k-t} \sum_{\substack{x', \in \mathbb{N} \\ f'_{-i} \in \overline{\mathcal{S}}, w' \in \tilde{\mathcal{W}}}} P_{\mu^*,\tilde{\mu},\tilde{\lambda}}[x_{ik} = x', f_{-i,k} = f'_{-i}, w_k = w' \mid x_{it} = x, f_{-i,t} = f_{-i}, w_t = w] \\ &\times\tilde{}\left(\pi(x', f'_{-i}, y') - \pi(x', \tilde{f}_{-i}(w'), y')\right), \end{aligned}$$

where we abbreviated $\tilde{f}_{-i} = \tilde{f}_{-i}(\mu, \lambda)$. We can write:

$$\begin{aligned} &P_{\mu^*,\tilde{\mu},\tilde{\lambda}}[x_{ik} = x', f_{-i,k} = f'_{-i}, w_k = w' \mid x_{it} = x, f_{-i,t} = f_{-i}, w_t = w] \\ &= P_{\mu^*,\tilde{\mu},\tilde{\lambda}}[x_{ik} = x' \mid f_{-i,k} = f'_{-i}, w_k = w', x_{it} = x, f_{-i,t} = f_{-i}, w_t = w] \\ &\times P_{\mu^*,\tilde{\mu},\tilde{\lambda}}[f_{-i,k} = f'_{-i}, w_k = w' \mid x_{it} = x, f_{-i,t} = f_{-i}, w_t = w]. \end{aligned}$$

Additionally,

$$\begin{aligned} &P_{\mu^*,\tilde{\mu},\tilde{\lambda}}[f_{-i,k} = f'_{-i}, w_k = w' \mid x_{it} = x, f_{-i,t} = f_{-i}, w_t = w] \\ &\qquad = P_{\tilde{\mu},\tilde{\lambda}}[f_{-i,k} = f'_{-i}, w_k = w' \mid f_{-i,t} = f_{-i}, w_t = w], \end{aligned}$$

because under POE strategies, $(f_{-i,k}, w_k)$ is independent of x_{it}, conditional on $(f_{-i,t}, w_t)$. Replacing and using Fubini's Theorem, we obtain:

$$V(\bar{x}, f, w|\mu^*, \tilde{\mu}, \tilde{\lambda}) - \tilde{V}(\bar{x}, w|\tilde{\mu}, \tilde{\lambda}) \leq$$
$$\sum_{k=t}^{\infty} \beta^{k-t} \sum_{\substack{f'_{-i} \in \overline{S} \\ w' \in \hat{\mathcal{W}}}} P_{\tilde{\mu},\tilde{\lambda}}[f_{-i,k} = f'_{-i}, w_k = w' \mid f_{-i,t} = f_{-i}, w_t = w]$$
$$\times \left[\max_{z \in \{\underline{x}(k,t), \ldots, x+(k-t)\overline{h}\}} \left(\pi(z, f'_{-i}, y') - \pi(z, \tilde{f}_{-i}(w'), y') \right) \right]^+ .$$

Finally, multiplying by $q(f, w)$, the invariant distribution of $\{(f_t, w_t) : t \geq 0\}$, summing over all (f, w), and using Fubini we get:

$$E[V(\bar{x}, f_t, w_t|\mu^*, \tilde{\mu}, \tilde{\lambda}) - \tilde{V}(\bar{x}, w_t|\tilde{\mu}, \tilde{\lambda})] \leq \sum_{k=t}^{\infty} \beta^{k-t} E\left[\left[\Delta_{\{\underline{x}(k,t), \ldots, x+(k-t)\overline{h}\}}(f_{-i,k}, w_k) \right]^+ \right] . \tag{A4}$$

Now, let us bound the second term in equation (A3). We have that,

$$\tilde{V}(\bar{x}, w|\tilde{\mu}, \tilde{\lambda}) - V(\bar{x}, f, w|\tilde{\mu}, \tilde{\lambda})$$
$$= E_{\tilde{\mu},\tilde{\lambda}} \left[\sum_{k=t}^{\tau_i} \beta^{k-t} \left(\pi(x_{ik}, \tilde{f}_{-i}(w_k), y_k) - \pi(x_{ik}, f_{-i,k}, y_k) \right) \middle| x_{it} = x, f_{-i,t} = f_{-i}, w_t = w \right].$$

Hence,

$$E[\tilde{V}(\bar{x}, w_t|\tilde{\mu}, \tilde{\lambda}) - V(\bar{x}, f_t, w_t|\tilde{\mu}, \tilde{\lambda})]$$
$$= E\left[E_{\tilde{\mu},\tilde{\lambda}} \left[\sum_{k=t}^{\tau_i} \beta^{k-t} \left(\pi(x_{ik}, \tilde{f}_{-i}(w_k), y_k) - \pi(x_{ik}, f_{-i,k}, y_k) \right) \middle| x_{it} = x, f_{-i,t}, w_t \right] \right] . \tag{A5}$$

The result follows by equations (A3), (A4), and (A5). The second bound follows by a similar argument, so we omit its proof. □

□ **Subtracting oneself in OE.** In this section, we derive $\tilde{f}_{-i}(\mu, \lambda)$ for two leading cases:

(i) a fixed maximum number of firms with short-lived entrants
(ii) unbounded number of firms with Poisson entry process

In case (i), we address a version of the Ericson and Pakes (1995) model in which there is a maximum number of firms N, and whenever the actual number of firms n is less than N, there is one potential short-lived entrant for each open "position," that is, each period there are $N - n$ entrants. This version of the model has been commonly used in recent empirical work such as CW and in numerous articles by Doraszelski. Case (ii) is the standard version of the OE model that allows for many potential entrants. In this section, we consider OE with no dominant firms and therefore suppress the conditioning on w. There are natural extensions of the two cases we consider to POE that we use on our numerical experiments.

Case (i). In case (i), one can imagine numbering each "position" from $1, \ldots, N$. Suppose that all incumbent firms use a common oblivious strategy μ and let λ represent the common oblivious entry strategy. (In this model, an oblivious entry strategy λ simply represents a probability of entry that is played by every potential entrant at every state of the world.)

In the position model, these two strategies together generate a Markov chain for the evolution of each firm in each position, where in any given period the position contains either an active firm or a potential entrant. The probability that the position will move from containing an active firm to containing a potential entrant is given by the probability that the active firm exits the market. The probability that it moves back to having an active firm is given by the probability of entry. If there is no entry, the position stays in the "potential entrant/inactive" state. Recall that under OE strategies, this Markov chain will be ergodic (Weintraub, Benkard, and Van Roy, 2008).

Denote the invariant distribution under this Markov chain by q. Then, because firm i takes up one slot and there are $N - 1$ slots remaining, we have that

$$\tilde{f}_{-i}(\mu, \lambda) = (N - 1) * q.$$

Case (ii). The Poisson entry model is more complicated, but the intuition is similar. In this model, there is an infinite pool of entrants and entry follows a Poisson process. Exit is binomial at each state and state transitions are Markov. Let $\tilde{f}(x|\mu, \lambda)$ be the expected number of firms in the industry at state x under oblivious policies μ and λ. Weintraub et al. 2008 show that the long-run invariant distribution of the industry state $f_t(x)$ is given by independent Poisson random variables with means equal to $\tilde{f}(x|\mu, \lambda)$ for each state x.

We use this result to derive the expected competitors' state $\tilde{f}_{-i}(\mu, \lambda)$ for a firm i that is located at state x_0. First, consider the expected number of competitors at state x_0 (i.e., the number of competitors at the same state as the own firm), denoted $\tilde{f}_{-i}(x_0|\mu, \lambda)$. In the invariant distribution $f(x_0|\mu, \lambda)$ is Poisson with mean $\tilde{f}(x_0|\mu, \lambda)$, and is independent of $f(y|\mu, \lambda)$ for all $y \neq x_0$. Because the own firm i is at state x_0, there must be at least one firm at state x_0. That is, we cannot simply subtract one from $\tilde{f}(x_0|\mu, \lambda)$ because we have to account for the fact that we know that there is at least one firm. Instead, we have that

$$\tilde{f}_{-i}(x_0, \mu, \lambda) = E[(f(x_0|\mu, \lambda) - 1)|f(x_0|\mu, \lambda) \geq 1],$$

where, as above, $f(x_0|\mu, \lambda)$ is a Poisson random variable with mean $\tilde{f}(x_0|\mu, \lambda)$. Note that when $E(f(x_0|\mu, \lambda))$ is large, the probability that there are zero firms at x_0 goes to zero and the conditioning makes no difference. In that case, we would have that $\tilde{f}_{-i}(x_0|\mu, \lambda) \approx \tilde{f}(x_0|\mu, \lambda) - 1$.

By a similar argument, independence implies that for $y \neq x_0$,

$$\tilde{f}_{-i}(y|\mu, \lambda) = E[f(y|\mu, \lambda)] = \tilde{f}(y|\mu, \lambda).$$

□ **Balance equations.** We suggest a way of computing $\tilde{f}(w|\mu, \lambda)$ (step 5 in Algorithm 1). Let $p(x, w, y, w') = P_{\mu,\lambda}[x_{i,t+1} = y, w_{t+1} = w' \mid x_{it} = x, w_t = w]$, where $i \notin \overline{D}$. Note that the probability that a firm in individual state x exits when dominant firm statistics are w is given by $1 - \sum_{y,w'} p(x, w, y, w')$.

Let $\tilde{f}(x, w|\mu, \lambda)$ be the $x-$th component of $\tilde{f}(w|\mu, \lambda)$. Let $r(x, w)$ be the product of $\tilde{f}(x, w|\mu, \lambda)$ and the steady state probability that the dominant firm statistics process is in state w (under strategies (μ, λ)), $q(w)$. That is, $r(x, w) = E[f_t(x)|w_t = w; \mu, \lambda]q(w)$. Then, it is not hard to show that $r(x, w)$ satisfies the balance equations:

$$r(x, w) = \sum_{(y,w')} r(y, w')p(y, w', x, w) + \mathbf{1}(x = x^e) \sum_{w'} \lambda(w')q(w')p(w', w), \tag{C1}$$

where $p(w', w) = P[w_{t+1} = w \mid w_t = w']$ and $\mathbf{1}$ is the indicator function. To see this, note that (for $x \neq x^e$ to simplify the exposition), $r(x, w) = E[f_t(x)\mathbf{1}(w_t = w); \mu, \lambda]$. Moreover, by basic accounting and with some abuse of notation to index firms, we get the following expression:

$$f_{t+1}(x)\mathbf{1}(w_{t+1} = w) = \sum_{y,w'} \mathbf{1}(w_t = w') \sum_{j=1}^{f_t(y)} \mathbf{1}(x_{j,t+1} = x, w_{t+1} = w, x_{jt} = y, w_t = w')\,.$$

Taking expectations at both sides and considering that $\{(f_t, w_t) : t \geq 0\}$ is a stationary process:

$$\begin{aligned} r(x, w) &= \sum_{y,w'} E\left[E\left[\mathbf{1}(w_t = w') \sum_{j=1}^{f_t(y)} \mathbf{1}(x_{j,t+1} = x, w_{t+1} = w, x_{jt} = y, w_t = w')\middle| f_t, w_t\right]\right] \\ &= \sum_{y,w'} E\left[\mathbf{1}(w_t = w')f_t(y)p(y, w', x, w)\right] \\ &= \sum_{y,w'} r(y, w')p(y, w', x, w), \end{aligned}$$

where the first equation follows by the law of iterated expectations and the second because transitions are conditionally independent, conditioned on the industry state. The balance equations follow.

We can obtain $r(x, w)$ by solving this set of balance equations. We can also obtain steady state probabilities of the dominant firm process, $q(w)$, by solving another set of balance equations. From these two objects, we obtain $\tilde{f}(w|\mu, \lambda)$.

Balance equations for CW model. We modify the balance equations above to incorporate aggregate shocks:

$$r(x, w, z) = \sum_{(y,w',z')} r(y, w', z')p(y, w', z', x, w, z). \tag{C2}$$

There is no λ term here because entry is already incorporated in p through transitioning from the inactive state. For this reason, to solve the system above we also need to add a normalization equation:

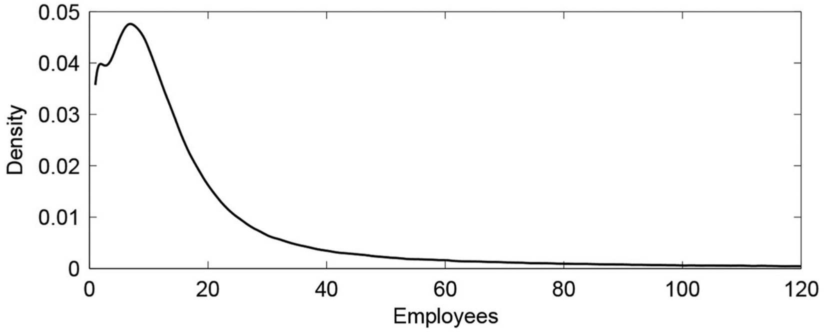
$$\sum_{x,w,z} r(x, w, z) = N - D,$$

where $N - D$ is the number of slots for fringe firms (including incumbents and potential entrants).

We can also obtain the steady state joint distribution of dominant firms and aggregate shocks from another set of balance equations. From these two objects, we obtain $\tilde{f}(w, z|\mu)$.

FIGURE D1

DISTRIBUTION OF THE NUMBER OF EMPLOYEES



Appendix B

The CW article categorizes firms into three size groups based on the number of employees. A firm is called small, medium, or large if its employment falls in the bottom, middle, or top third of the industry employment distribution, respectively. We extend this discretization by considering smaller percentile bins. First, we calibrate the distribution of employment using five moments given in the CW article; that is, the 33rd percentile is eight employees, the 66th percentile is 18 employees, 95th percentile is 110 employees, the mean is 27.24 employees, and the standard deviation is 79.03 employees. We parametrize the distribution using a mixture of two log-normal random variables and choose their parameters to minimize the square of the percentage deviation from the above 5 moments. We obtain the following specification of the number of employees L,

$$L = \begin{cases} \exp(2.51 + 1.57X_1) & \text{with Prob } 0.4, \\ \exp(2.45 + 0.57X_2) & \text{with Prob } 0.6, \end{cases}$$

where X_1 and X_2 are independent standard normals. The density of L is depicted at Figure D1. Using the calibrated distribution, we obtain $B \in \{3, 6, 9, 12, 15\}$ equal bins of firms by size. Let $L^{\alpha,B}$ be a number of employees at the median of the bin α in the discretization with B bins. Table D1 reports values of $L^{\alpha,B}$. Note that $L^{\alpha,3}$ corresponds to ACW discretization.

Using the $L^{\alpha,3}$ grid, we calibrate a linear interpolation of the profit parameters θ_1^α, θ_2^α, θ_{31}^α, and θ_{32}^α. Specifically, we obtain coefficients by running four least squares regressions,

$$\theta_z^\alpha \approx A_z^0 + A_z^1 \log(L^{\alpha,3}),$$

for $z \in \{1, 2, 31, 32\}$. Calibrated parameters A_z^0 and A_z^1 provide interpolants $\hat{\theta}_z^{\alpha,B}$ for $B \in \{3, 6, 9, 12, 15\}$.

Table D1 Size Bins

Small	Medium	Large
4.3	11.8	33.4
2.2 6.2	9.7 14.3	22.9 66.7
1.5 4.3 6.8	9.1 11.8 15.3	20.8 33.4 96.7
1.2 3.3 5.3 7.0	8.8 10.7 13.0 15.8	19.9 27.0 44.5 123.3
1.0 2.7 4.3 5.8 7.2	8.6 10.1 11.8 13.7 16.2	19.4 24.3 33.4 55.7 147.1

The interpolation of transition cost $\tau(x', x)$ is more complicated than the profit function because the current state $x = (\alpha, h)$ contains the information about the current size α as well as the largest past size h. We use the following linear transition cost specification:

$$\tau(x', x) \approx \begin{cases} C^1[\log(L^{\alpha',3}) - \log(L^{\alpha,3})] + C^2[\log(L^{\alpha',3}) - \log(L^{h,3})] & \text{if } \alpha' > \alpha \\ C^3[\log(L^{\alpha',3}) - \log(L^{\alpha,3})] + C^4[\log(L^{\alpha',3}) - \log(L^{h,3})] & \text{if } \alpha' < \alpha. \end{cases}$$

The above equation is calibrated using nonlinear least squares. To prevent the state space from exploding, we keep track of only small, medium and large past size; that is, even for $B > 3$, we allow only $h \in \{\text{Sm, Med, Lg}\}$ and categorize intermediate past sizes as small, medium, large according to Table D1 At the same time, we allow α' and α to take B

values. Formally, the transition cost is obtained using

$$\hat{\tau}^B(x', x) = \begin{cases} C^1[\log(L^{\alpha',B}) - \log(L^{\alpha,B})] + C^2[\log(L^{\alpha',B}) - \log(L^{h,3})] & \text{if } \alpha' > \alpha \\ C^3[\log(L^{\alpha',B}) - \log(L^{\alpha,B})] + C^4[\log(L^{\alpha',B}) - \log(L^{h,3})] & \text{if } \alpha' < \alpha. \end{cases}$$

Note that we use $L^{h,3}$ instead of $L^{h,B}$ because h takes on only three values.

References

AGUIRREGABIRIA, V. AND MIRA, P. "Sequential Estimation of Dynamic Discrete Games." *Econometrica*, Vol. 75(1) (2007), pp. 1–54.

BAJARI, P., BENKARD, C.L., AND LEVIN, J. "Estimating Dynamic Models of Imperfect Competition." *Econometrica*, Vol. 75(5) (2007), pp. 1331–1370.

BENKARD, C.L. "A Dynamic Analysis of the Market for Wide-Bodied Commercial Aircraft." *Review of Economic Studies*, Vol. 71(3) (2004), pp. 581–611.

BERRY, S., LEVINSOHN, J., AND PAKES, A. "Automobile Prices in Market Equilibrium." *Econometrica*, Vol. 63(4) (1995), pp. 841–890.

BESANKO, D., DORASZELSKI, U., KRYUKOV, Y., AND SATTERTHWAITE, M. "Learning-by-doing, Organizational Forgetting, and Industry Dynamics." *Econometrica*, Vol. 78(2) (2010), pp. 453–508.

CAPLIN, A. AND NALEBUFF, B. "Aggregation and Imperfect Competition–on the Existence of Equilibrium." *Econometrica*, Vol. 59(1) (1991), pp. 25–59.

COLLARD-WEXLER, A. "Demand Fluctuations in the Ready-Mix Concrete Industry." *Econometrica*, Vol. 81(3) (2013), pp. 1003–1037.

DORASZELSKI, U. AND JUDD, K.L. "Avoiding the Curse of Dimensionality in Dynamic Stochastic Games." *Quantitative Economics*, Vol. 3(1) (2012), pp. 53–93.

DORASZELSKI, U. AND PAKES, A. "A Framework for Applied Dynamic Analysis in IO." In *Handbook of Industrial Organization, Volume 3*. Amsterdam: North-Holland, 2007.

DORASZELSKI, U. AND SATTERTHWAITE, M. "Computable Markov-Perfect Industry Dynamics." *RAND Journal of Economics*, Vol. 41(2) (2010), pp. 215–243.

ERICSON, R. AND PAKES, A. "Markov-Perfect Industry Dynamics: A Framework for Empirical Work." *Review of Economic Studies*, Vol. 62(1) (1995), pp. 53–82.

FARIAS, V., SAURE, D., AND WEINTRAUB, G. "An Approximate Dynamic Programming Approach to Solving Dynamic Oligopoly Models." *RAND Journal of Economics*, Vol. 43(2) (2012), pp. 253–282.

FERSHTMAN, C. AND PAKES, A. "Dynamic Games With Asymmetric Information: A Framework For Empirical Work." *Quarterly Journal of Economics*, Vol. 127(4) (2012), pp. 1611–1661.

FOWLIE, M., REGUANT, M., AND RYAN, S. P. "Market-Based Emissions Regulation and Industry Dynamics." *Journal of Political Economy*, in press.

GOETTLER, R.L. AND GORDON, B.R. "Does AMD Spur Intel to Innovate More" *Journal of Political Economy*, Vol. 119(6) (2011), pp. 1141–1200.

GOWRISANKARAN, G. AND TOWN, R. "Dynamic Equilibrium in the Hospital Industry." *Journal of Economics and Management Strategy*, Vol. 6(1) (1997), pp. 45–74.

IFRACH, B. AND WEINTRAUB G.Y. "A Framework for Dynamic Oligopoly in Concentrated Industries." Working paper, Columbia University, 2014.

JEZIORSKI, P. "Estimation of Cost Synergies From Mergers: Application to U.S. Radio." *RAND Journal of Economics*, Vol. 45(4) (2014), pp. 816–846.

———. "Empirical Model of Dynamic Merger Enforcement – Choosing Ownership Caps in U.S. Radio" Working Paper, UC Berkeley, 2015.

KYDLAND, F. "A Dynamic Dominant Firm Model of Industry Structure." *Scandinavian Journal of Economics*, Vol. 81(3) (1979), pp. 355–366.

MASKIN, E. AND TIROLE, J. "A Theory of Dynamic Oligopoly, I and II." *Econometrica*, Vol. 56(3) (1988), pp. 549–570.

PAKES, A. AND MCGUIRE, P. "Computing Markov-Perfect Nash Equilibria: Numerical Implications of a Dynamic Differentiated Product Model." *RAND Journal of Economics*, Vol. 25(4) (1994), pp. 555–589.

———. "Stochastic Algorithms, Symmetric Markov Perfect Equilibrium, and the 'Curse' of Dimensionality." *Econometrica*, Vol. 69(5) (2001), pp. 1261–1281.

PAKES, A., OSTROVSKY, M., AND BERRY S. "Simple Estimators for the Parameters of Discrete Dynamic Games, With Entry/Exit Examples." *RAND Journal of Economics*, Vol. 38(2) (2007), pp. 373–399.

PESENDORFER, M. AND SCHMIDT-DENGLER, P. "Identification and Estimation of Dynamic Games." Mimeo, London School of Economics, 2003.

RYAN, S. "The Costs of Environmental Regulation in a Concentrated Industry." *Econometrica*, Vol. 80(3) (2012), pp. 1019–1061.

SANTOS, C. "An Aggregation Method to Solve Dynamic Games." Working Paper, Tilburg University, 2012.

SUTTON, J. "Market Share Dynamics and the 'Persistence of Leadership' Debate." *American Economic Review*, Vol. 97(1) (2007), pp. 222–241.

SWEETING, A. "Dynamic Product Positioning in Differentiated Product Markets: The Effect of Fees for Musical Performance Rights on the Commercial Radio Industry." *Econometrica*, Vol. 81(5) (2013), pp. 1763–1803.

VON STACKELBERG, H. *Marktform Und Gleichgewicht*. Berlin: Springer, 1934.

WEINTRAUB, G.Y., BENKARD, C.L., AND VAN ROY, B. "Markov Perfect Industry Dynamics With Many Firms." *Econometrica*, Vol. 76(6) (2008), pp. 1375–1411.

———. "Computational Methods For Oblivious Equilibrium." *Operations Research*, Vol. 58(4) (2010), pp. 1247–1265.