NBER WORKING PAPER SERIES

OBLIVIOUS EQUILIBRIUM FOR CONCENTRATED INDUSTRIES

C. Lanier Benkard
Przemyslaw Jeziorski
Gabriel Y. Weintraub

Working Paper 19307
http://www.nber.org/papers/w19307

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2013

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 paper has also benefitted from conversations with Ariel Pakes, Ali Yurukoglu, and seminar participants at various conferences and institutions. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Oblivious Equilibrium for Concentrated Industries
C. Lanier Benkard, Przemyslaw Jeziorski, and Gabriel Y. Weintraub
NBER Working Paper No. 19307
August 2013
JEL No. L0,L1,L13

ABSTRACT

This paper explores the application of oblivious equilibrium to concentrated industries. We define 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. We perform computational experiments that show that POE are often close to MPE in concentrated industries with characteristics similar to real world industries even when OE are not. We derive error bounds for evaluating the performance of POE when MPE cannot be computed. Finally, we demonstrate an important trade-off facing empirical researchers between implementing an equilibrium concept that is computationally light in a richer economic model, and implementing MPE in a simpler one.

C. Lanier Benkard
Stanford Graduate School of Business
655 Knight Way
Stanford, CA 94305
and NBER
lanierb@stanford.edu

Przemyslaw Jeziorski
Haas School of Business
2220 Piedmont Avenue
University of California at Berkeley
Berkeley, CA 94720-1900
przemekj@haas.berkeley.edu

Gabriel Y. Weintraub
Columbia Business School, Uris 402
3022 Broadway
New York, NY 10027, U.S.A.
gweintraub@columbia.edu

1 Introduction

There has been much recent work in I.O. on empirical applications of dynamic oligopoly models (e.g., Collard-Wexler (2013), Fowlie, Reguant, and Ryan (2012), Goettler and Gordon (2011), Jeziorski (2013), Ryan (2012), and 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, i.e., the set of firms in the market 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 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 since 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 modeling details are heavily dictated by computational concerns, typically at some expense to the credibility of the economic analysis.

In a recent paper, Weintraub, Benkard, and Van Roy (2008) propose a method for analyzing Ericson and Pakes (1995)-style dynamic models of imperfect competition that is intended to address some of these concerns. In that paper, 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 its computation time is independent of the 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. However, while the OE model has some very appealing properties, it also has some weaknesses. One that has often been raised is the appropriateness of the model for highly concentrated industries. Since many of the most interesting policy questions in I.O. focus on highly concentrated industries, this is an important flaw.

In this paper 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 "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 (that only depend on a firm's own state), and our hope is that POE will provide a better model of more concentrated industries than OE does. Moreover, if entry and exit are modeled identically for dominant and (non-dominant) 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 mainly 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 sapce 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) remain independent of the number of firms in the overall industry. As a consequence, while 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. In many industries 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. A second motivation for POE as a behavioral model is the idea that, 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. However, note that we do not explicitly model an information acquisition process, and we believe that doing so would be quite complex from a game theoretic perspective.

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 industries 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 since 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 industries (industries with C2>0.90 – corresponding to approximately the top 1% of manufacturing industries in the US) we find that OE is not typically close to MPE. We explore the performance of POE in these cases, and find that it depends critically on the rate of turnover among leading firms. If turnover among leading firms is within the range of, or even slightly exceeds, levels typically observed in real world industries[1], then we find that POE is typically close to MPE in very concentrated industries with only one or two dominant firms. On the other hand, if turnover among leading firms is extremely high, higher than is realistic, then we find that it requires up to three or four dominant firms to obtain results close to MPE.

Similarly to Weintraub, Benkard, and Van Roy (2010), we are also able to 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. Since in most applications it would not be possible to compute MPE for comparison purposes, the error bounds may be used to guide 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). In our computational experiments we find the bounds computed to be quite loose, but we believe that the bounds may be more useful in empirical applications that typically have more firms. Moreover, tighter bounds could potentially be derived using application specific details.

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 known by all firms, while (non-dominant) fringe firms' states are ignored, and this difference in information leads to differences in behavior between the two types of firms. Because dominant firms' states are tracked by all other firms, 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' information sets in equilibrium calculations. While our information structure is somewhat natural, and leads to behavior that is not unrealistic, arbitrary restrictions of firms' information sets to facilitate computation could lead to unintended and unnatural firm behavior.

We also apply POE to the empirical model of Collard-Wexler (2013) (henceforth 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

[1]See Sutton (2007).

improves many statistics, 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. They can either compute a simple equilibrium concept such as POE in a richer model, or compute exact MPE in a simpler one, but cannot 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 modeling restrictions that requires.

A weakness of POE is that it is not as theoretically clean a concept as OE in two main respects. In an abstraction from reality, our implementation of POE does not allow for firms to transit from fringe to dominant or vice-versa. In reality over long horizons the identities of "dominant" firms would change, as new firms grow and become dominant while old firms decay and exit. In our model firms have incomplete information of the current industry state because they know only their own state and the states of the dominant firms. If the identities of the dominant firms were to change in response to equilibrium behavior then in the model firms' information sets would be endogenous. This would lead to non-Markovian behavior of dominant firms and would require a new approach to the problem. Thus, while we do think that this is a potentially interesting direction for future research, we opted to start with the more straightforward problem of modeling a fixed set of dominant firms. Additionally, we do not think that this assumption is all that unrealistic for many empirical problems, which tend to be interested in the effects of a policy over short to medium length time periods in which the set of dominant firms would be fairly stable.

Additionally, there is an issue that, since in POE firms have only incomplete knowledge of the current industry state, in principle the complete history of industry states becomes relevant for predicting the state variables of untracked competitor fringe firms. In our implementation of POE firms use finite, typically short, histories in their strategies. 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.

We are not the only researchers to recognize the methodological and computational hurdles in the empirical application of dynamic oligopoly models, and there have now been several other solutions proposed to this problem. For example, Pakes and McGuire (2001) proposes 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) suggests 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 (2013) for an application of this technique). Farias, Saure, and Weintraub (2012) introduces 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.

Additionally, most empirical applications necessarily make some modeling 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 (2012)). Many papers 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 modeling contexts.

Compared with these papers, our methods take a different approach that is best described as limiting the information available to firms (or alternatively restricting strategies of 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 complex 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 informational (or strategy) restrictions, computation time for OE and POE is independent of the 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 industries. Ifrach and Weintraub (2012) and Corbae and D'Erasmo (2012) also build on our idea of restricting firms' strategies in dynamic oligopoly models.

Another related paper by Fershtman and Pakes (2012) considers dynamic oligopoly models with asymmetric information. While their economic model is fundamentally different from ours due to the presence of asymmetric information, their paper tackles many similar issues to ours because OE and POE limit the information sets of firms and therefore our model could be cast as one of asymmetric information even if our underlying economic model is 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 next section introduces the model. Section 3 precisely defines POE. Section 4 provides computational experiments evaluating POE. Section 5 computes POE for an empirical example. Finally, 6 concludes the paper.

2 A Dynamic Model of Imperfect Competition

In this section we formulate a model of an industry in which firms compete in a single-good market. Our model is based on Weintraub, Benkard, and Van Roy (2008).

2.1 Model and Notation

The industry evolves over discrete time periods and an infinite horizon. We index time periods with non-negative integers $t \in \mathbb{N}$ ($\mathbb{N} = \{0, 1, 2, \ldots\}$). All random variables are defined on a probability space $(\Omega, \mathcal{F}, \mathcal{P})$ equipped with a filtration $\{\mathcal{F}_t : t \geq 0\}$. We adopt a convention of indexing by t variables that are $\mathcal{F}_t$-measurable.

Each firm that enters the industry is assigned a unique positive integer-valued index. The set of indices of incumbent firms at time t is denoted by S_t. At each time $t \in \mathbb{N}$, we denote the number of incumbent firms as n_t.

Firm heterogeneity is reflected through firm states. To fix an interpretation, we will refer to a firm's state as its quality level. However, firm states might more generally reflect productivity, capacity, the size of its consumer network, or any other aspect of the firm that affects its profits. At time t, the quality level of firm $i \in S_t$ is denoted by $x_{it} \in \mathbb{N}$.

We define the *industry state* s_t to be a vector over quality levels that specifies, for each quality level $x \in \mathbb{N}$, the number of incumbent firms at quality level x in period t. We define the state space $\overline{\mathcal{S}} = \left\{ s \in \mathbb{N}^\infty \middle| \sum_{x=0}^\infty s(x) < \infty \right\}$. Though in principle there are a countable number of industry states, we will also consider an extended state space $\mathcal{S} = \left\{ s \in \Re_+^\infty \middle| \sum_{x=0}^\infty s(x) < \infty \right\}$. This notion will be useful, for example, when considering the expected value of the industry state. For each $i \in S_t$, we define $s_{-i,t} \in \mathcal{S}$ to be the state of the *competitors* of firm i; that is, $s_{-i,t}(x) = s_t(x) - 1$ if $x_{it} = x$, and $s_{-i,t}(x) = s_t(x)$, otherwise. Similarly, $n_{-i,t}$ denotes to the number of competitors of firm i.

In each period, each incumbent firm earns profits on a spot market. A firm's single period expected profit $\pi(x_{it}, s_{-i,t})$ 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 $i \in S_t$ 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.

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 very general as h may take on either positive or negative values (e.g., allowing for positive depreciation). We denote the unit cost of investment by d.

Our model can accommodate a variety of different entry processes and we will use several different entry models in the remainder of the paper. Here we describe the entry model of Weintraub, Benkard, and Van Roy (2008). In that model in each period new firms can enter the industry by paying a setup cost κ. Entrants do not earn profits in the period that they enter. They appear in the following period at state $x^e \in \mathbb{N}$ and can earn profits thereafter. (It would not change any of the results to assume that the entry state was a random variable.)

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:

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

2.2 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 (e.g. Berry, Levinsohn, and Pakes 1995).

2.3 Assumptions

We make several assumptions about the model primitives, beginning with the profit function.

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 $\sup_{x,s} \pi(x, s) < \infty$.*

The assumptions are natural. Assumption 2.1.1 ensures that increases in quality lead to increases in profit. Assumption 2.1.2 ensures that profits are positive and bounded. We also make assumptions about investment and the distributions of the private shocks:

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 $\overline{\iota}$ such that $\iota_{it} < \overline{\iota}$, $\forall i, \forall t$.*

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 .*

Again the assumptions are natural and fairly weak. Assumptions 2.2.1 and 2.2.2 imply that investment and exit outcomes are idiosyncratic conditional on the state. Assumption 2.2.3 and 2.2.4 imply that investment is productive. 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. 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 specifications in the literature.

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, and our solution concept will require that it satisfies a zero expected profit condition. Modeling 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 industries 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 sufficient reason to enter the industry.

2.4 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 particular, there is a function ι such that at each time t, each incumbent firm $i \in S_t$ invests an amount $\iota_{it} = \iota(x_{it}, s_{-i,t})$. Similarly, each firm follows an exit strategy that takes the form of a cutoff rule: there is a real-valued function ρ such that an incumbent firm $i \in S_t$ exits at time t if and only if $\phi_{it} \geq \rho(x_{it}, s_{-i,t})$. In Weintraub, Benkard, and Van Roy (2008) we show that there always exists an optimal exit strategy of this form even among very general classes of exit strategies. Let $\mathcal{M}$ denote the set of exit/investment strategies such that an element $\mu \in \mathcal{M}$ is a pair of functions $\mu = (\iota, \rho)$, where $\iota : \mathbb{N} \times \mathcal{S} \to \Re_+$ is an investment strategy and $\rho : \mathbb{N} \times \mathcal{S} \to \Re_+$ is an exit strategy. Similarly, we denote the set of entry rate functions by Λ, where an element of Λ is a function $\lambda : \mathcal{S} \to \Re_+$.

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 follows a common strategy $\mu \in \mathcal{M}$, the entry rate function is $\lambda \in \Lambda$, and the firm itself follows strategy $\mu' \in \mathcal{M}$. In particular,

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

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

1. Incumbent firm strategies represent a MPE:

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

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

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

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.[2]

Dynamic programming algorithms can be used to optimize firm strategies, and equilibria to our model can be computed via their iterative application. However, these algorithms require compute time and memory that grow proportionately with the number of relevant industry states, which is often intractable in contexts of practical interest. This difficulty motivates our alternative approach.

3 Partially Oblivious Equilibrium

In this section we extend the concept of oblivious equilibrium to allow firms to track the states of a few firms that we call *dominant firms*, and have beliefs about the states of all other firms, which we call *fringe firms*. We call this new concept *partially oblivious equilibrium*, hereafter POE.

As discussed above, 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 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, in which case the POE model nests both OE and MPE as special cases in which there are no dominant firms and all dominant firms, respectively.

Let $\overline{D} = \{i_1, i_2, ..., i_n\}$ be the set of indices associated with the dominant firms. We extend the state

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

of firms to include a binary variable that identifies dominant firms. Hence, the state of firm i at time t is $\overline{x}_{it} = (x_{it}, 1)$ if $i \in \overline{D}$, and $\overline{x}_{it} = (x_{it}, 0)$ if $i \notin \overline{D}$. This specification will allow for equilibrium strategies to differ between dominant firms and fringe firms.

3.1 Partially Oblivious Strategies

A partially oblivious strategy is a function of the firm's state (which in this case also indicates whether the firm is dominant or not) and the state of firms in $\overline{D}$. Let y_t be a vector that represents the state of the dominant firms at time t. Formally, $y_t = \left(x_{(i_1)t}, x_{(i_2)t}, ..., x_{(i_n)t}\right)$, where $x_{(i_k)t}$ is the $k-$th order statistic of $(x_{i_1 t}, x_{i_2 t}, ..., x_{i_n t})$. Our equilibrium concept will assume anonymous strategies that are a function of the order statistics of $(x_{i_1 t}, x_{i_2 t}, ..., x_{i_n t})$ and will not depend on the identities of the dominant firms.

While there are typically at most a few dominant firms in a given industry, so that it is computationally feasible to optimize over strategies that are a function of the dominant firms' state, typically the number of fringe firms is large, so that it would not typically be computationally feasible to optimize over strategies that are a function of the fringe firms' state. Instead, we will assume that firms make estimates of the fringe firms' state based on averages. If there are many firms, because of averaging effects, firms should be able to accurately predict the fringe firms' state for a given time period. However, since fringe firms' strategies are functions of the dominant firms' states, in order to make a precise prediction about the fringe, firms would require knowledge of the entire history of dominant firms' states. Computing expectations this way would be computationally impractical; instead, we will allow firms to predict the fringe firms' state based on a finite set of statistics that depend on the past history of the dominant firms' state.[3]

Based on this motivation, we will 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' state. We call such restricted strategies *partially oblivious strategies*. To convey this dependence, we define the sequence $\{w_t \in \mathcal{W} = \mathcal{W}_1 \times ... \times \mathcal{W}_K : t \geq 0\}$ where $w_t(1) = y_t$, for all $t \geq 0$, and $\mathcal{W}_j$ are countable sets.

We define $\tilde{\mathcal{M}}_p$ and $\tilde{\Lambda}_p$ as the set of partially oblivious strategies and the set of partially oblivious entry rate functions, respectively. If firm i uses 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). 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.

Since $w_t(1) = y_t$, firms keep track of the current dominant firms' state when making decisions with

[3]This idea is similar to OE with aggregate shocks; see Weintraub, Benkard, and Van Roy (2010).

partially oblivious strategies. The state variables $w_t(2), ..., w_t(K)$ allow firms to incorporate additional information about the history of realizations of the dominant firms' state into the strategies. This information could be useful to better predict the average fringe firm state conditional on observing w_t.

We make the following assumption over the state statistics w_t.

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 $t \geq 0$, $w_t(1) = y_t$.*

The assumption implies that the Markov chain $\{w_t : t \geq 0\}$ admits a unique invariant distribution that assigns positive mass to all states $w \in \mathcal{W}$. The assumption is useful because it implies that firms' beliefs over the distribution of fringe firms are well defined off the equilibrium path. While the above assumption imposes a requirement over all strategies, a weaker assumption that only requires that $\{w_t : t \geq 0\}$ is a finite irreducible and aperiodic Markov chain under strategies that could potentially be best responses suffices. Moreover, even this weaker assumption is not absolutely necessary since it could be replaced with an alternative assumption about how firms form beliefs at states that are not reachable in equilibrium. Fershtman and Pakes (2012) discuss this issue in more detail in the context of their model.

Different partially oblivious strategies can be defined depending on the specification of w_t. For example, suppose that $w_t = \{y_t, y_{t-1}, ..., y_{t-K+1}\}$ so that the statistics correspond to the last realizations of the dominant firms' state. In this case $\mathcal{W}$ is the set of feasible K-tuples of consecutive dominant firms' state. If, in addition, for any set of partially oblivious strategies that dominant firms use, $\{y_t : t \geq 0\}$ is a finite irreducible and aperiodic Markov chain, then this specification satisfies Assumption 3.1.

3.2 Fringe Firms' Expected State

In a POE firms make decisions assuming that the fringe firms' state is the *expected* fringe firms' state conditional on the current value of the dominant firm state statistics w_t. Formally, suppose that firms use 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 $\overline{\mathcal{S}}$: $\overline{\mathcal{F}} = \left\{ f \in \mathbb{N}^\infty \middle| \sum_{x=0}^{\infty} f(x) < \infty \right\}$. 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.

For given strategies, firms predict the fringe firm state based on the current realization of the dominant firm statistics, w_t. Accordingly, 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 w_t being w.

3.3 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}, if the fringe firms other than firm i are in state $f_{-i,t}$ and dominant firms other than firm i are in state $y_{-i,t}$. Since vectors $f_{-i,t}$ and $y_{-i,t}$ do not contain the firm i, in order to obtain them we need to subtract the firm i from f_t if i is in the fringe or from y_t if i is dominant. Note, however, that in our model formulation period profits depend only on the firm states x of all dominant and fringe firms, and not on which firms are dominant.

Because y_t represents the dominant firms' state, note that if $i \in \overline{D}$, then $x_{it} = x$ only if $y_t(k) = x$, for some $k = 1,...,n$. Consequently, for all $w \in \mathcal{W}$, we define the set $\overline{X}(w) = \{(x,1) : x = w(1,k) \text{ for some } k = 1,...,n\} \cup \{(x,0) : x \in \mathbb{N}\}$, where $w(1,k)$ is the $k-$th component of $w(1)$.[4]

We define a *partially oblivious value function* for all $(\overline{x}, w) \in \{(\overline{x}, w) : w \in \mathcal{W}, \overline{x} \in \overline{X}(w)\}$:

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

where i is the index of a firm in state $\overline{x}$ at time period t. 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 B 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 $\tilde{V}(\overline{x}, w|\mu,\lambda) = \tilde{V}(\overline{x}, w|\mu,\mu,\lambda)$.

3.4 Partially Oblivious Equilibrium

We now define a new solution concept: a *partially oblivious equilibrium* consists of a strategy $\mu \in \tilde{\mathcal{M}}_p$ and an entry rate function $\lambda \in \tilde{\Lambda}_p$ that satisfy the following conditions:

1. Firm strategies optimize a partially oblivious value function.

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

[4] Recall that for all $t \geq 0$, $w_t(1) = y_t$.

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

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

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 $n = 0$, a POE is an OE.

In Section 3.6 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.

3.5 Error Bounds

Because firms' strategies are restricted in POE, it is useful to measure the performance of POE strategies relative to a full information strategy. For this purpose, we derive error bounds for fringe firms in this model. Because dominant firms' states are tracked by fringe firms, and thus a deviation by a dominant firm may lead to a change in the distribution of future industry states, it is not straightforward to derive an error bound for dominant firms that is easily computed. Specifically we want to bound the amount by which a fringe firm at state $\bar{x} = (x, 0), x \in \mathbb{N}$ can improve its expected net present value by unilaterally deviating from the POE strategy, and instead following an optimal Markovian best response.

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 is a function of the firm's own state, the full industry state (including the dominant firms' state), and the dominant firms' state statistics. 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 firm, still restricting it's strategy to depend only on it's own state, other dominant firms' state, and the dominant firms' statistics. Similarly, if $\lambda \in \Lambda_p$, then the entry rate at time t is $\lambda(f_t, w_t)$.[5]

For extended Markov strategy $\mu', \mu \in \mathcal{M}_p$ and extended entry rate function $\lambda \in \Lambda_p$, with some abuse

[5]Recall that $w_t(1) = y_t$, hence, strategies are a function of y_t.

of notation, we define for $\bar{x} = (x, 0)$ the fringe extended value function,

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

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. Since 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 2.4 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. The extension of Markov strategies is useful so that firms can predict the future evolution of the industry playing against competitors that use POE strategies.

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 $\tilde{f}_{-i}(w) \equiv \tilde{f}_{-i}(w|\tilde{\mu}, \tilde{\lambda})$. With some abuse of notation, let $\Delta_A(f_{-i}, w) = \sup_{x \in A}\left(\pi(x, f_{-i}, w(1)) - \pi(x, \tilde{f}_{-i}(w), w(1))\right)$ 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 in the Appendix.

Theorem 3.1. *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 $\bar{x} = (x, 0), x \in \mathbb{N}$,*

$$(3.3) \quad E\left[\sup_{\mu' \in \mathcal{M}_p} V(\bar{x}, f_t, w_t|\mu', \tilde{\mu}, \tilde{\lambda}) - V(\bar{x}, f_t, w_t|\tilde{\mu}, \tilde{\lambda})\right] \leq \sum_{k=t}^{\infty} \beta^{k-t} E\left[\left[\Delta_{\{\underline{x}(k,t),\ldots,x+(k-t)\overline{h}\}}(f_{-i,k}, w_k)\right]^+\right] + E\left[E_{\tilde{\mu},\tilde{\lambda}}\left[\sum_{k=t}^{\tau_i} \beta^{k-t}\left(\pi(x_{ik}, \tilde{f}_{-i}(w_k), y_k) - \pi(x_{ik}, f_{-i,k}, y_k)\right)\middle| x_{it} = x, f_{-i,t}, w_t\right]\right].$$

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* $\bar{x} = (x, 0), x \in \mathbb{N}$*,*

$$(3.4) \quad E\left[\sup_{\mu' \in \mathcal{M}_p} V(\bar{x}, f_t, w_t | \mu', \tilde{\mu}, \tilde{\lambda}) - V(\bar{x}, f_t, w_t | \tilde{\mu}, \tilde{\lambda})\right]$$

$$\leq \sum_{k=t}^{\infty} \beta^{k-t} E_{\hat{\mu}, \tilde{\mu}, \tilde{\lambda}} \left[\Delta(x_{i,k}, f_{-i,k}, w_k)^+ \Big| x_{i,t} = x\right]$$

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

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. 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.

Recall that (f_k, w_k) is distributed according to the invariant distribution for all $k \geq 0$, so both bounds can be computed using simulation. Both bounds are also quite general and do not take advantage of many of the detailed modeling assumptions. In particular, they are valid for various entry processes. Allowing the bounds to depend on more details of the model would allow us to achieve tighter expressions (see Weintraub, Benkard, and Van Roy 2010 for an example).

3.6 Algorithms and Computations

In this section we introduce an algorithm to compute POE. Throughout this section we only consider states $(\overline{x}, w) \in \{(\overline{x}, w) : w \in \mathcal{W}, \overline{x} \in \overline{X}(w)\}$.

The algorithm is an iterative algorithm. It starts by computing the expected fringe firm state conditional on the dominant firm statistics, $\tilde{f}(w|\mu, \lambda)$ (step 5). It then computes the strategy that maximizes the partially oblivious value function (step 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 (step 8). Finally, strategies and entry rates are updated "smoothly" (steps 12 and 13). The parameters N_1, N_2, γ_1, and γ_2 are set after some experimentation to speed up convergence. 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.

For initialization, let $(\tilde{\mu}, \tilde{\lambda})$ be an OE. Let $\tilde{V}$ be the respective value function.

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: $n := 0$
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:     $n := n + 1$
12:     $\mu := \mu + (\mu^* - \mu)/(n^{\gamma_1} + N_1)$
13:     $\lambda := \lambda + (\lambda^* - \lambda)/(n^{\gamma_2} + N_2)$
14: **until** $\Delta_1 \leq \epsilon_1$ and $\Delta_2 \leq \epsilon_2$

We finish by suggesting a way of computing $\tilde{f}(w|\mu, \lambda)$ (step 5 in the algorithm). 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 $i \notin \overline{D}$. The probability that the firm exits from a state (x, w) is one minus the sum of the transition probabilities from that state. 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, $q(w)$. Then, $r(x, w)$ satisfies the balance equations:

$$r(x, w) = \sum_{(y,w')} r(y, w')p(y, w', x, w) + \mathbf{1}(x = x^e) \sum_{w'} \lambda(w')q(w')p(w', w), \tag{3.5}$$

where $p(w', w) = P[w_{t+1} = w \mid w_t = w']$ and $\mathbf{1}$ is the indicator function. 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 by solving another set of balance equations. From these two objects, we obtain $\tilde{f}(w|\mu, \lambda)$.

3.7 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. Specifically, in our past work we defined an oblivious value function as:

$$\tilde{V}(x|\mu', \mu, \lambda) = E_{\mu'}\left[\sum_{k=t}^{\tau_i} \beta^{k-t}(\pi(x_{ik}, \tilde{f}_{-i}(\mu, \lambda)) - d\iota_{ik}) + \beta^{\tau_i - t}\phi_{i,\tau_i}\middle| x_{it} = x\right] \tag{3.6}$$

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.

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 approximation. 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

$$\tilde{V}(x|\mu', \mu, \lambda) = E_{\mu',\mu,\lambda}\left[\sum_{k=t}^{\tau_i} \beta^{k-t}(\pi(x_{ik}, f_{-i,k}) - d\iota_{ik}) + \beta^{\tau_i - t}\phi_{i,\tau_i}\middle| x_{it} = x\right] \tag{3.7}$$

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 industries 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 industries OE-SIM is still easy to compute. For larger industries 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 (3.6) and (3.7), 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 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 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 I.O. 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:

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

where Y is the consumer's income, p_{it} is the price of the good produced by firm i, and ψ is a scaling factor. 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 $N(x_{it}, p_{it}) = \exp(\theta_1 \ln(\frac{x_{it}}{\psi} + 1) + \theta_2 \ln(Y - p_{it}))$. Then, the expected market share of each firm is given by:

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

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,

$$Y - p_{it} + \theta_2 (p_{it} - c)(\sigma(x_{it}, f_{-i,t}, p_t) - 1) = 0 \; , \; \forall i \in S_t \; . \qquad (4.1)$$

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

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

Firms can also invest $\iota \geq 0$ in order 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:

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

While it would be straightforward to implement entry and exit in the model, in order 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 industries we use a second quantity-cost (Cournot) model.

QUANTITY-COST COMPETITION. Our second model is designed to be a "worst case" for OE in that it will generate very highly concentrated near monopoly markets with a fringe of smaller firms, and can also generate high turnover among the leader firms. 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 industries, but the model also has the property that in periods of high concentration lagging firms stop producing altogether and exit the market, leading to a pure monopoly. This case is not particularly interesting for our purposes.

In order to ensure that small firms continue to exist when the industry is highly concentrated, we 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 on the basis of price, and the second market as a set of "local" markets, where each firm captures their local market completely regardless of price.

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

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

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. The models of transitions is the same as in the logit model above. Also, as in the logit model, though allowing for entry and exit would be straightforward in this model, in order to simplify the computations we do not model these features of the industry.

4.1 Comparison with MPE

We first investigate the relationship between OE, POE and MPE in the two models above. 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). 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 2. In each case we model 8 active firms and 8 firm states. We use such a small number of 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. We experimented with using longer histories but it made no difference to the results.

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 computational algorithms consistently select the same equilibrium each time.[6] All of the results below apply to this particular equilibrium.

[6] Pakes and McGuire (1994) found the same thing and Besanko, Doraszelski, Kryukov, and Satterthwaite (2005) provide a detailed explanation of why this happens.

LOGIT MODEL. Figures 2 and 3 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 this 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 as high as 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 fairly high for this model. 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. 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 industries 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 over 15%. However, as can be seen in the next two panels, the bounds are very loose since the fully informed Markov best response differences are in actuality quite small, typically on the order of less than one percent.

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.

Summing up our findings for the logit model, all the equilibrium concepts are quite close to MPE, but OE-SIM and POE are closer to MPE than OE is. For this simple example MPE took more than 10 minutes to compute, while 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.

COURNOT-MODEL. Figures 4-6 provide the same set of comparisons for the Cournot model for three different depreciation rates respectively ($\delta = 0.2, 0.5, 0.8$). The first thing to note about the results is that in using this special modified Cournot model and these parameter values, we are looking at very extreme examples of concentration. The top left panel of the figures shows that the industries 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 industries in the extreme tail of census industries.[7] 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.

We have found that the extent to which adding dominant firms improves these statistics depends critically on the rate of turnover among leader firms in the MPE. So long as turnover among the leading firms is not too high ($\delta = 0.2$ or $\delta = 0.5$), adding dominant firms generates equilibria that look much closer to MPE than OE does. POE3, for example, generates investment, consumer and producer surplus within 5% of MPE for all parameter values when $\delta = 0.2$ or $\delta = 0.5$.

On the other hand, POE is not as close to MPE when $\delta = 0.8$. We have found that this difference is caused by the very high rate of turnover among the leading firms when depreciation is high. The top right panel of figure 6 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). What we are looking at, then, is an industry where firms bounce around a lot in the size distribution, spending a lot of time as fringe firms, but also are frequently near monopolists, though this does not last very long.

[7] For comparison, in the 2007 U.S. Census of Manufacturers 98% of six digit NAICS manufacturing industries have C4 less than 0.90, and 99% have C4 less than 0.95.

We view this case as somewhat unrealistic compared with most real world industries,[8] but nevertheless it is useful to study because it exhibits the case where POE is furthest from MPE.

Figure 6 shows that when firm turnover is high adding dominant firms does still typically improve upon OE, though not by as much as before. In addition, there is one case (investment cost=20) where consumer surplus remains over 20% different from that in MPE even in POE3. Best response improvements for dominant firms are always small but can be large for fringe firms in OE. POE, on the other hand, generates producer surplus that is nearly identical to MPE, as well as small unilateral deviations from the POE strategy for both dominant and fringe firms even in the worst cases.

We also have an intuition for these results. When there is some stability among leader firms, 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 becomes critical for a fringe firm to have more detailed information about the industry, for example so it can know exactly the moment when the industry is such that it is poised to become a leader firm for a few periods or a new leader competitor will arise. The more stable the leading firms are, the closer are the POE industries to MPE, though the results in figures 4-6 suggest that POE is close to MPE for a pretty wide range of industries.

Moreover, when turnover among leaders becomes even higher than this, such that each period all firm's states are close to being an iid random draw from the long run distribution, then the results reverse and even OE closely matches MPE. The reason for this reversal is that in that case the current states of competitor firms have little bearing on future profits (since competitor states are close to iid), and thus the firm's optimal policy in MPE is a function only of its own state. Thus, the worst case scenario for OE and POE is the one exhibited, in which there is enough serial correlation in the state that the current state matters a lot for the near future, but the serial correlation is low enough that the industry remains quite unstable over time.

We also computed OE-SIM for the $(\delta = 0.8)$ case. As before, the results for OE-SIM are somewhat mixed. For low values of the investment cost, OE-SIM actually makes things worse relative to OE, while for medium to high values it somewhat improves things. Oddly, the one case where OE-SIM seems to help a lot (investment cost equals 20) corresponds to the one case where POE does not improve things much, but we do not know why this is the case.

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

[8] 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, while 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.

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 than the two model used in this section.

4.2 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 in order to show that there is entry deterrence.

Figure 7 shows the overall distribution of the long-run distribution of firm states for the Cournot model for both OE, POE1, POE2, and POE3, as well as the distribution of states broken down by firm type for dominant and fringe firms. We only show the results for one parameterization here, but the results are qualitatively the same for all parameter values in the Cournot model, and indeed are similar for all models that we have tried including the logit model above and the CW model outlined below.

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 small. Figure 8 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 while 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 industries, 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.

5 An Empirical Model

In this section we investigate the properties of POE in a more complex empirical model due to Collard-Wexler (2013) (henceforth CW). Collard-Wexler (2013) 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 counter-factual experiments he computes MPE using the stochastic approximation algorithm due to Pakes and McGuire (2001).

5.1 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 paper 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 iid choice specific preference shock that is private information to the firm, 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:

$$\begin{aligned} (5.1)\quad R(x_{i,t+1}, x_{-i,t+1}, M_{t+1}) = \sum_{\alpha\in\{sm,med,lg\}} \mathbb{1}\{x_{i,t+1} = \alpha\}\Big(&\theta^\alpha_1 + \theta^\alpha_2 M_{t+1} + \theta^\alpha_{31}\{NComp >= 1\} \\ &+ \theta^\alpha_{32} * \{NComp > 1\} * \log(NComp - 1)\Big) \end{aligned}$$

where

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

is the number of active competitors. The parameter θ^α_1 is a fixed cost parameter, θ^α_2 is the coefficient on aggregate demand, and θ^α_{31} and θ^α_{32} measure the effect of competition on firm profits. The parameter σ scales the variance of the error term.

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 both 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 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 since 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

$$
\begin{aligned}
&\tilde{V}(\bar{x}, w, z|\mu, \mu') = \\
&\qquad E_{\mu,\mu'}\left[\sum_{k=t}^{\tau_i} \beta^{k-t}[R(x_{i,k+1}, \tilde{f}_{-i}(w_{k+1}, z_{k+1}|\mu'), y_{-i,k+1}, M_{k+1}) \right. \\
&\qquad\qquad \left. - \tau(x_{i,k+1}, x_{ik}) + \sigma\epsilon_{ia}^k]|\bar{x}_{it} = \bar{x}, w_t = w, z_t = z\right]
\end{aligned}
\tag{5.2}
$$

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 balance equation (3.5) that incorporates the aggregate shocks

$$
r(x, w, z) = \sum_{(y,w',z')} r(y, w', z')p(y, w', z', x, w, z), \tag{5.3}
$$

There is no λ term here because entry is already incorporated in p through transitioning from the inactive state. We can also obtain a steady state for the joint distribution of dominant firms and aggregate shocks from another set of balance equations. From these two objects we obtain $\tilde{f}(w, z|\mu)$.

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

$$
\begin{aligned}
&\tilde{v}(a, \bar{x}, w, z|\mu, \mu') = \\
&\qquad E\left[R(a, \tilde{f}_{-i}(w_{t+1}, z_{t+1}|\mu'), y_{-i,t+1}, M_{t+1}) - \tau(a, x_{it}) \right. \\
&\qquad\qquad \left. +\beta\tilde{V}(\bar{x}_{i,t+1}, w_{t+1}, z_{t+1}|\mu, \mu')|\bar{x}_{it} = \bar{x}, w_t = w, z_t = z\right]
\end{aligned}
\tag{5.4}
$$

In the notation in equation (5.4) 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:

$$\tilde{V}(\bar{x}, w, z|\mu, \mu') = \sigma \log\left[\sum_a \exp\left(\sigma^{-1}\tilde{v}(a, \bar{x}, w, z|\mu, \mu')\right)\right] + \sigma\gamma \tag{5.5}$$

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

$$P[a_{it} = a|\bar{x}_{it}, w_t, z_t; \mu, \mu'] = \frac{\exp\left(\sigma^{-1}\tilde{v}(a, \bar{x}_{it}, w_t, z_t|\mu, \mu')\right)}{\sum_{a'} \exp\left(\sigma^{-1}\tilde{v}(a', \bar{x}_{it}, w_t, z_t|\mu, \mu')\right)} \tag{5.6}$$

In this case, the maximization step of the POE algorithm is equivalent to iterating on equations (5.4) and (5.5) 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 (5.6) and balance conditions (5.3).

5.2 Results

We first consider an exact replica of the CW model used in his paper.[9] Parameter values for the revenue function are listed in Table 3 and the sunk adjustment costs parameters are listed in Table 4. Collard-Wexler (2013) 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 5 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

[9]Our reported statistics vary from those reported in Collard-Wexler (2013) because we report long run averages while the statistics reported in the paper represent the average across many markets (of different sizes) of short run simulations starting from a (different) particular observed state for each market. 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 paper.

and sunk costs of entry.

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, while 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 while POE1 and POE2 take a few minutes and POE3 and POE4 take a few hours. At the same 7-firm MPE takes days, 10-firm MPE takes months, while 12 firm MPE is infeasible to compute.[10]

5.3 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. Since OE and POE are much lighter computationally than 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,

[10] All reported computation times we obtained using 1-core 2.5GHz AMD Opteron machine with 128GB of RAM. The code was written in C and C++, and compiled using a highly optimized compiler. We were able to compute a 10-firm MPE in a reasonable time using massive parallelization.

we require adjustment cost and entry cost parameters for each level of the finer grid, but we do not have access to the confidential U.S. 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 4) using a linear function of the state, and then interpolate the parameter values for the finer grid.[11] Table 6 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, since 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 9 and Table 7 show statistics for POE-1 for five grid sizes for the firm size state: $\{3,6,9,12,15\}$. For these grid sizes there are approximately 350 thousand, 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.

Once the grid size hits about nine points, however, the results begin to stabilize, suggesting that nine size

[11] Appendix C provides the details of this interpolation.

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.

Importantly, our point is not to say whether the finer grid is more or less reflective of the cement industry. We could not determine this without re-estimating 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 modeling simplifications that are purely for computational reasons. These results demonstrate that such simplifications can have an impact on economics of the model, and that there is potentially some 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 modeling restrictions that requires.

6 Conclusions

In this paper 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. We then explored the behavior of POE relative to OE and MPE in a wide variety of industries through computational experiments.

Summarizing the results, we found that OE was surprisingly close to MPE even in many highly concentrated industries. In extremely concentrated industries 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 Collard-Wexler (2013), 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 empirical work.

We also caution that we found that the behavior of dominant firms and fringe firms differ in the POE model. In some industries 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.

A Proof of error bound

Proof of Theorem 3.1. 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 (non-oblivious) 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 3.5 as the set of extended Markov strategies.

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

$$\begin{aligned} E[V(\bar{x}, f_t, w_t | \mu^*, \tilde{\mu}, \tilde{\lambda}) - V(\bar{x}, f_t, w_t | \tilde{\mu}, \tilde{\lambda})] &= E[V(\bar{x}, f_t, w_t | \mu^*, \tilde{\mu}, \tilde{\lambda}) - \tilde{V}(\bar{x}, w_t | \tilde{\mu}, \tilde{\lambda})] \\ &+ E[\tilde{V}(\bar{x}, w_t | \tilde{\mu}, \tilde{\lambda}) - V(\bar{x}, f_t, w_t | \tilde{\mu}, \tilde{\lambda})] \end{aligned} \tag{A.1}$$

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,

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

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

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

Additionally,

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

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:

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

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:

$$
\text{(A.2)} \quad E[V(\bar{x}, f_t, w_t|\mu^*, \tilde{\mu}, \tilde{\lambda}) - \tilde{V}(\bar{x}, w_t|\tilde{\mu}, \tilde{\lambda})] \leq \sum_{k=t}^{\infty} \beta^{k-t} E\left[\left[\Delta_{\{\underline{x}(k,t), \ldots, x + (k-t)\overline{h}\}}(f_{-i,k}, w_k) \right]^+ \right] .
$$

Now, let us bound the second term in equation (A.1). We have that,

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

Hence,

$$\text{(A.3)} \quad E[\tilde{V}(\bar{x}, w_t|\tilde{\mu}, \tilde{\lambda}) - V(\bar{x}, f_t, w_t|\tilde{\mu}, \tilde{\lambda})]$$
$$= E\left[E_{\tilde{\mu},\tilde{\lambda}}\left[\sum_{k=t}^{\tau_i} \beta^{k-t}\left(\pi(x_{ik}, \tilde{f}_{-i}(w_k), y_k) - \pi(x_{ik}, f_{-i,k}, y_k)\right) \Big| x_{it} = x, f_{-i,t}, w_t\right]\right] .$$

The result follows by equations (A.1), (A.2), and (A.3). The second bound follows by a similar argument so we omit its proof. $\square$

B 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". I.e., each period there are $N - n$ entrants. This version of the model has been commonly used in recent empirical work such as Collard-Wexler (2013) and in numerous papers 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.

B.1 Case (i)

In case (i) one can imagine numbering each "position" from $1, ..., N$. 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, since firm i takes up one slot and there are $N - 1$ slots remaining, we have that

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

B.2 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, Benkard, and Van Roy (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 $y \neq x_0$. Since the own firm i is at state x_0 there must be at least one firm at state x_0. I.e., 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 $y \neq x_0$

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

C Changing the Grid Size for the ACW Model

The ACW paper 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 5 moments given in the ACW paper; that is, the 33rd percentile is 8 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 1. 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 1 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 $B \in \{3, 6, 9, 12, 15\}$.

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 non-linear least squares. To prevent the state space from exploding we keep track of only Small, Medium and Large largest past size; that is, even for $B > 3$ we allow only $h \in \{\text{Sm}, \text{Med}, \text{Lg}\}$ and categorize intermediate past sizes as Sm, Med, Lg according to Table 1. At the

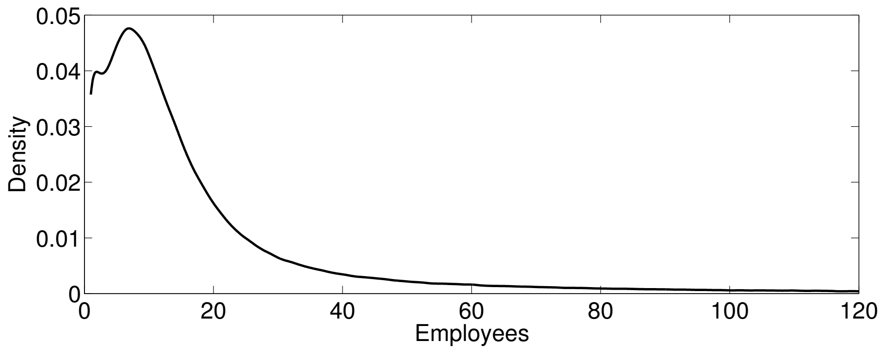

Figure 1: Distribution of the number of employees

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.

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

Table 1: Size bins

References

Aguirregabiria, V. and P. Mira (2007). Sequential estimation of dynamic discrete games. *Econometrica 75*(1), 1 – 54.

Bajari, P., C. L. Benkard, and J. Levin (2007). Estimating dynamic models of imperfect competition. *Econometrica 75*(5), 1331 – 1370.

Benkard, C. L. (2004). A dynamic analysis of the market for wide-bodied commercial aircraft. *The Review of Economic Studies 71*(3), 581–611.

Berry, S., J. Levinsohn, and A. Pakes (1995). Automobile prices in market equilibrium. *Econometrica 63*(4), 841–890.

Besanko, D., U. Doraszelski, Y. Kryukov, and M. Satterthwaite (2005). Learning-by-doing, organizational forgetting, and industry dynamics. Working Paper Northwestern University.

Caplin, A. and B. Nalebuff (1991). Aggregation and imperfect competition - on the existence of equilibrium. *Econometrica 59*(1), 25 – 59.

Collard-Wexler, A. (2013). Demand fluctuations in the ready-mix concrete industry. *Econometrica*, forthcoming.

Corbae, D. and P. D'Erasmo (2012). Capital requirements in a quantitative model of banking industry dynamics. Working Paper, Maryland U.

Doraszelski, U. and K. L. Judd (2012). Avoiding the curse of dimensionality in dynamic stochastic games. *Quantitative Economics 3*(1), 53–93.

Doraszelski, U. and A. Pakes (2007). A framework for applied dynamic analysis in IO. In *Handbook of Industrial Organization, Volume 3*. North-Holland, Amsterdam.

Doraszelski, U. and M. Satterthwaite (2010). Computable markov-perfect industry dynamics. *RAND Journal of Economics 41*(2), 215 – 243. Working Paper, Harvard University.

Ericson, R. and A. Pakes (1995). Markov-perfect industry dynamics: A framework for empirical work. *Review of Economic Studies 62*(1), 53 – 82.

Farias, V., D. Saure, and G. Weintraub (2012). An approximate dynamic programming approach to solving dynamic oligopoly models. *The RAND Journal of Economics 43*(2), 253–282.

Fershtman, C. and A. Pakes (2012). Dynamic games with asymmetric information: A framework for empirical work*. *The Quarterly Journal of Economics 127*(4), 1611–1661.

Fowlie, M., M. Reguant, and S. P. Ryan (2012, December). Market-based emissions regulation and industry dynamics. NBER Working Papers 18645, National Bureau of Economic Research, Inc.

Goettler, R. L. and B. R. Gordon (2011). Does amd spur intel to innovate more? *Journal of Political Economy 119*(6), 1141–1200.

Gowrisankaran, G. and R. Town (1997). Dynamic equilibrium in the hospital industry. *Journal of Economics and Management Strategy 6*(1), 45 – 74.

Ifrach, B. and G. Y. Weintraub (2012). A framework for dynamic oligopoly in concentrated industries. Working Paper, Columbia U.

Jeziorski, P. (2013). Estimation of cost synergies from mergers: Application to us radio. Working Paper, U.C. Berkeley.

Kydland, F. (1979). A dynamic dominant firm model of industry structure. *The Scandinavian Journal of Economics*, 355–366.

Maskin, E. and J. Tirole (1988). A theory of dynamic oligopoly, I and II. *Econometrica 56*(3), 549 – 570.

Pakes, A. and P. McGuire (1994). Computing Markov-perfect Nash equilibria: Numerical implications of a dynamic differentiated product model. *RAND Journal of Economics 25*(4), 555 – 589.

Pakes, A. and P. McGuire (2001). Stochastic algorithms, symmetric Markov perfect equilibrium, and the ‘curse’ of dimensionality. *Econometrica 69*(5), 1261 – 1281.

Pakes, A., M. Ostrovsky, and S. Berry (2007). Simple estimators for the parameters of discrete dynamic games, with entry/exit examples. *RAND Journal of Economics 38*(2), 373 – 399.

Pesendorfer, M. and P. Schmidt-Dengler (2003). Identification and estimation of dynamic games. Mimeo, London School of Economics.

Ryan, S. (2012). The costs of environmental regulation in a concentrated industry. *Econometrica 80*(3), 1019–1061.

Sutton, J. (2007). Market share dynamics and the ”persistence of leadership” debate. *American Economic Review 97*(1), 222–241.

Sweeting, A. (2013). Dynamic product repositioning in differentiated product markets: The case of format switching in the commercial radio industry. *Econometrica*, forthcoming.

Von Stackelberg, H. (1934). *Marktform und Gleichgewicht.-Wien & Berlin: Springer 1934. VI, 138 S. 8*. J. Springer.

Weintraub, G. Y., C. L. Benkard, and B. Van Roy (2008). Markov perfect industry dynamics with many firms. *Econometrica 76*(6), 1375–1411.

Weintraub, G. Y., C. L. Benkard, and B. Van Roy (2010). Computational methods for oblivious equilibrium. *Operations Research 58*(4), 1247–1265.

D Tables and Figures

	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
State multiplier (ψ)	-	1.0
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

Table 2: Parameters for numerical simulations

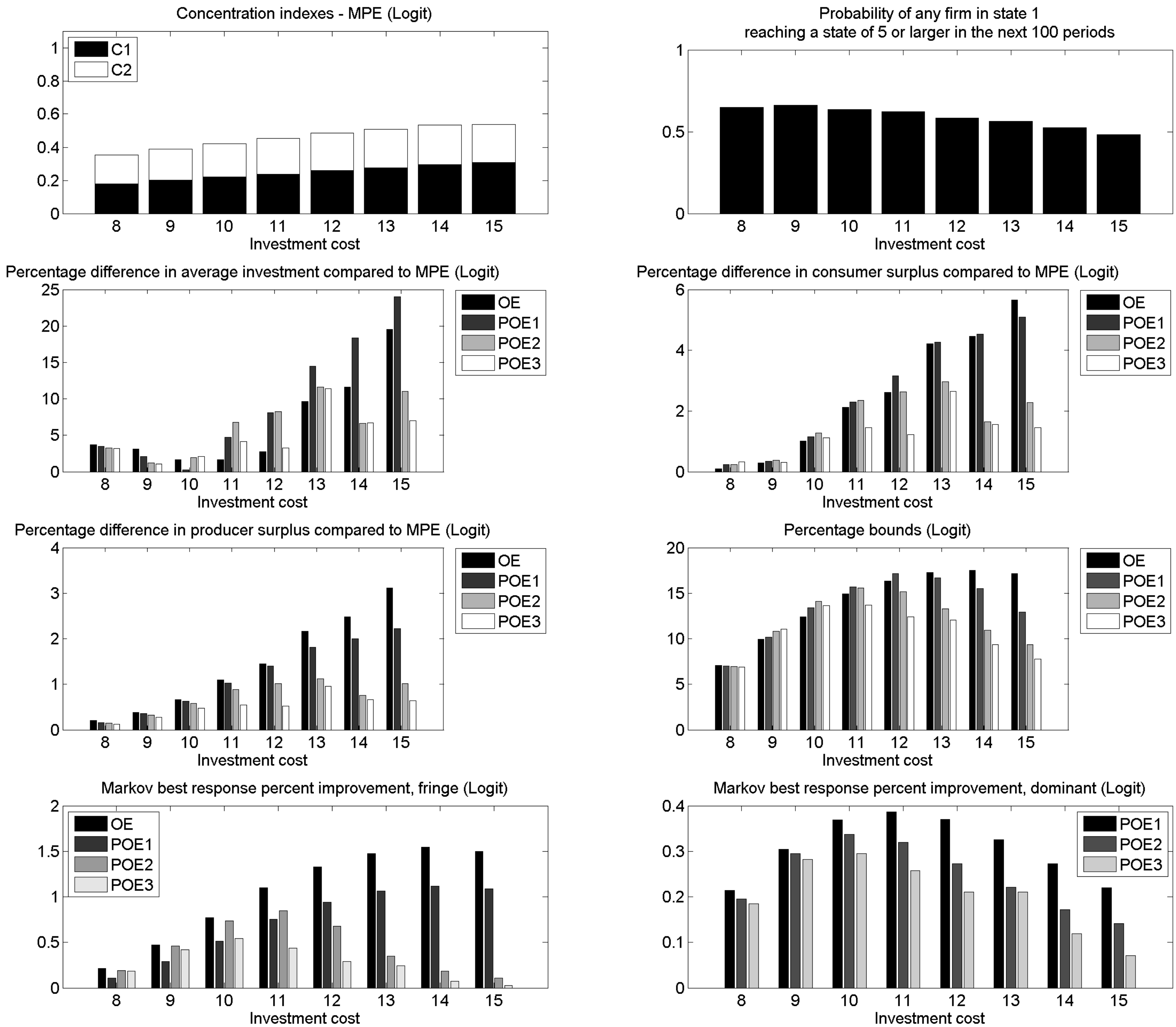

Figure 2: Long-run industry statistics for OE and POE versus MPE, logit model $\delta = 0.2$

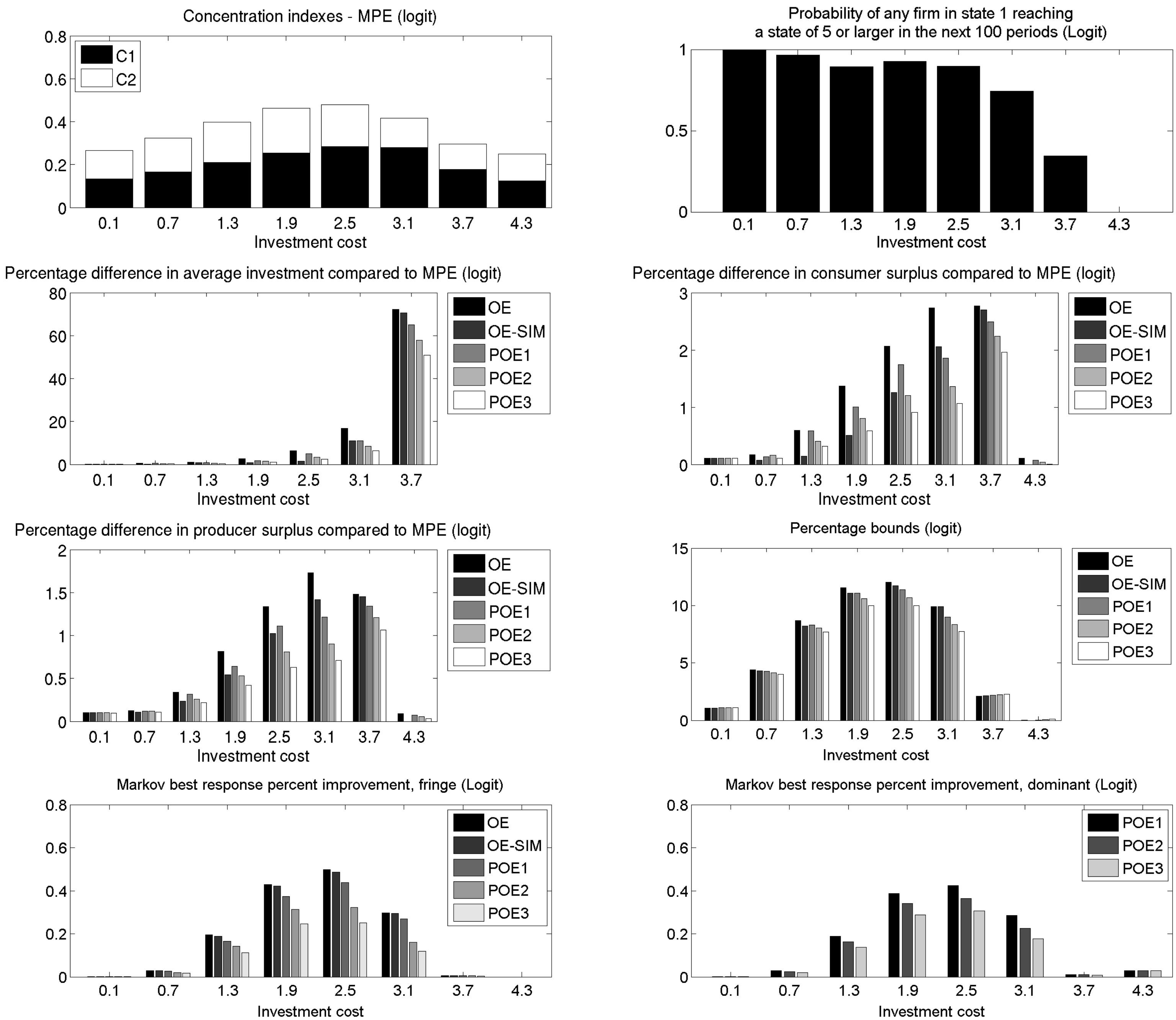

Figure 3: Long-run industry statistics for OE and POE versus MPE, logit model, $\delta = 0.7$

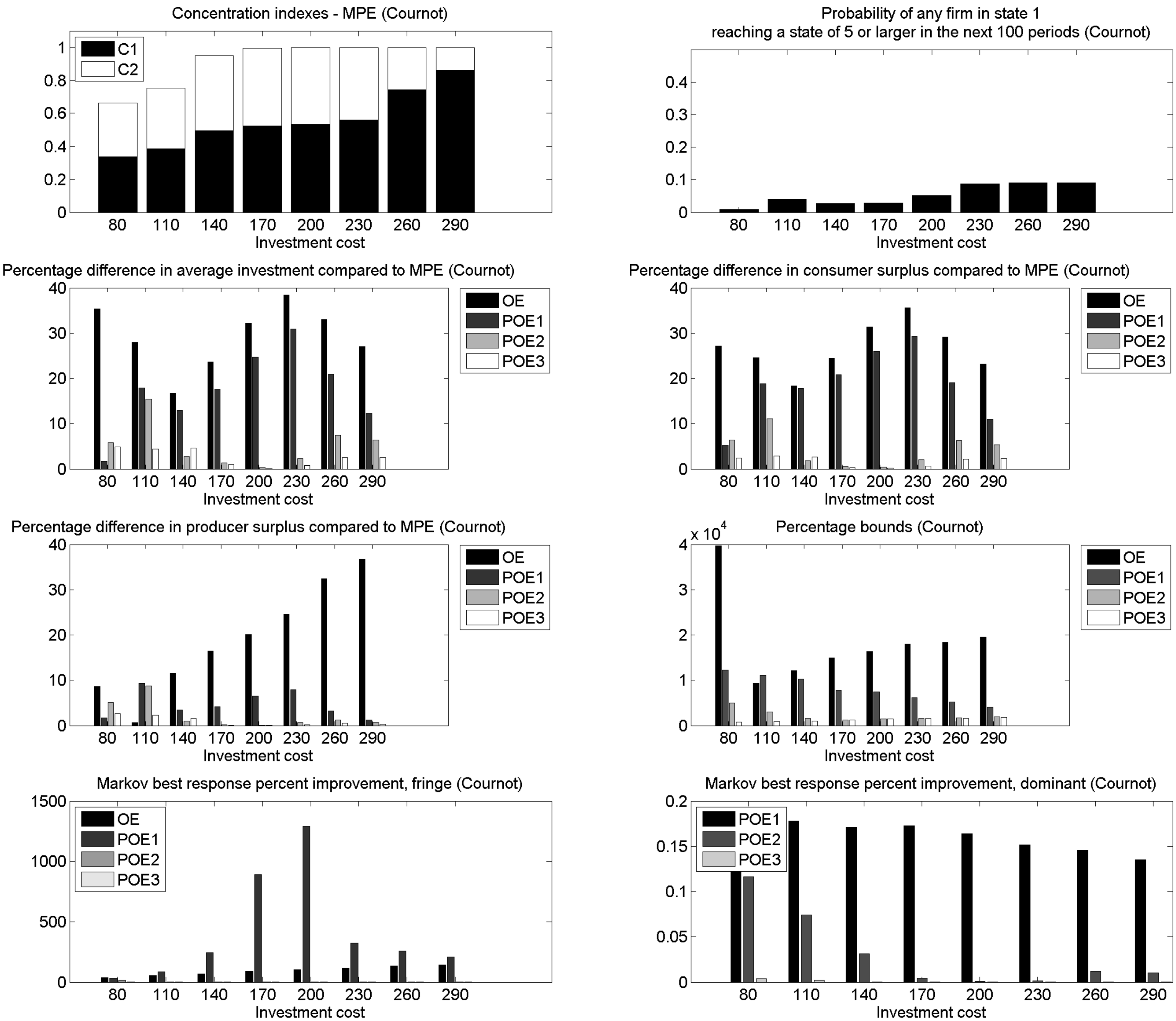

Figure 4: Long-run industry statistics for OE and POE versus MPE, Cournot model, $\delta = 0.2$

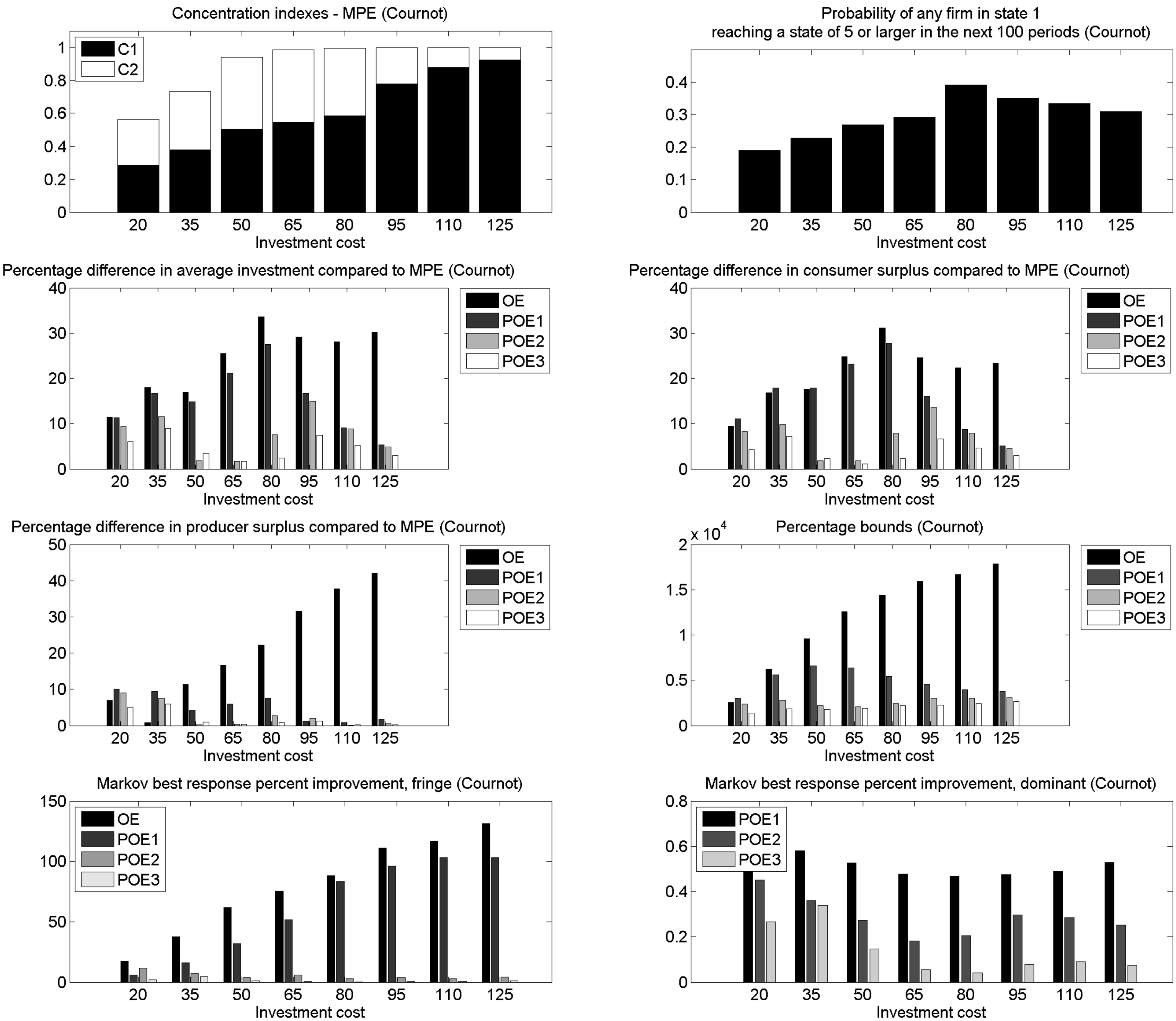

Figure 5: Long-run industry statistics for OE and POE versus MPE, Cournot model, $\delta = 0.5$

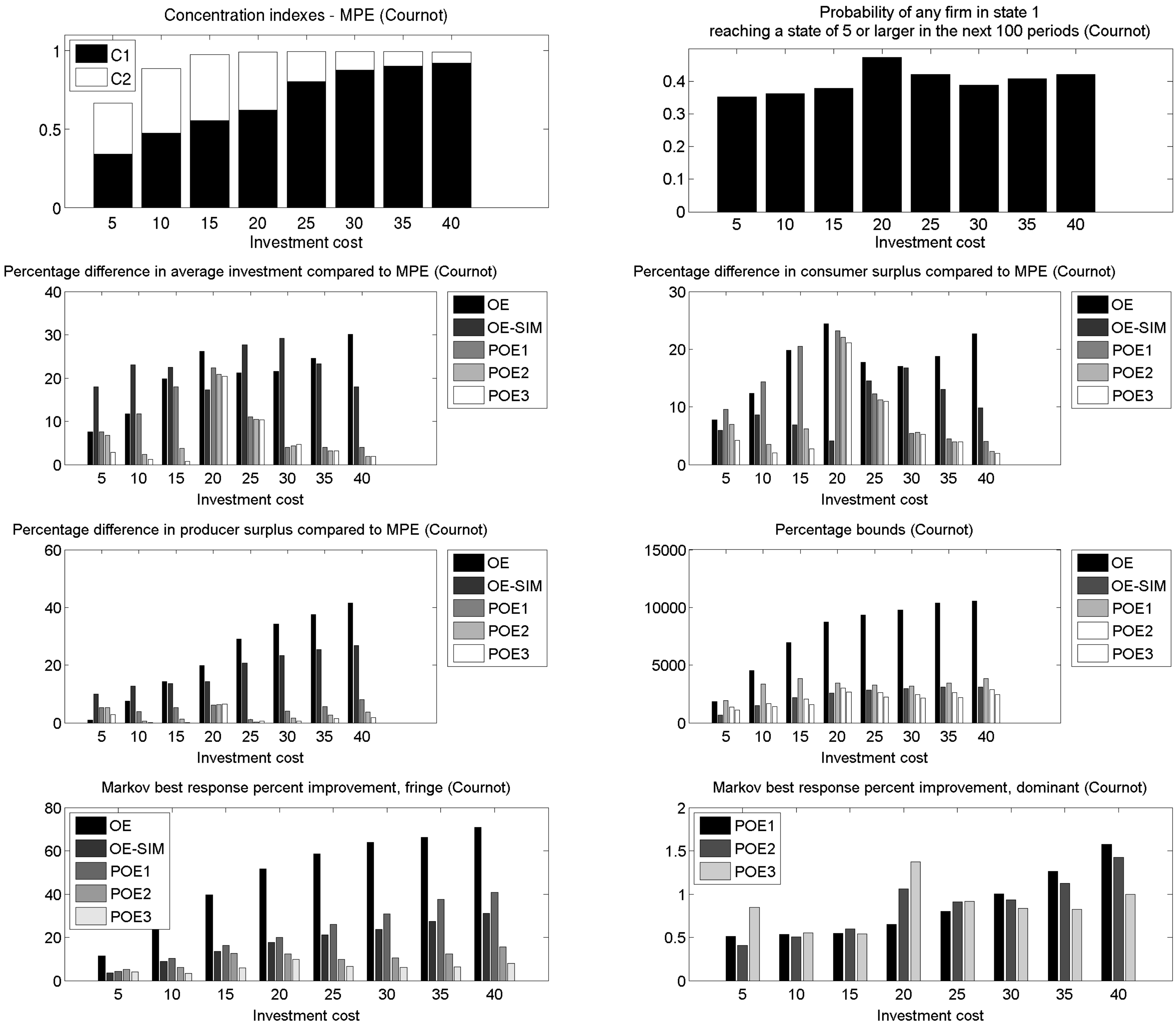

Figure 6: Long-run industry statistics for OE and POE versus MPE, Cournot model, $\delta = 0.8$

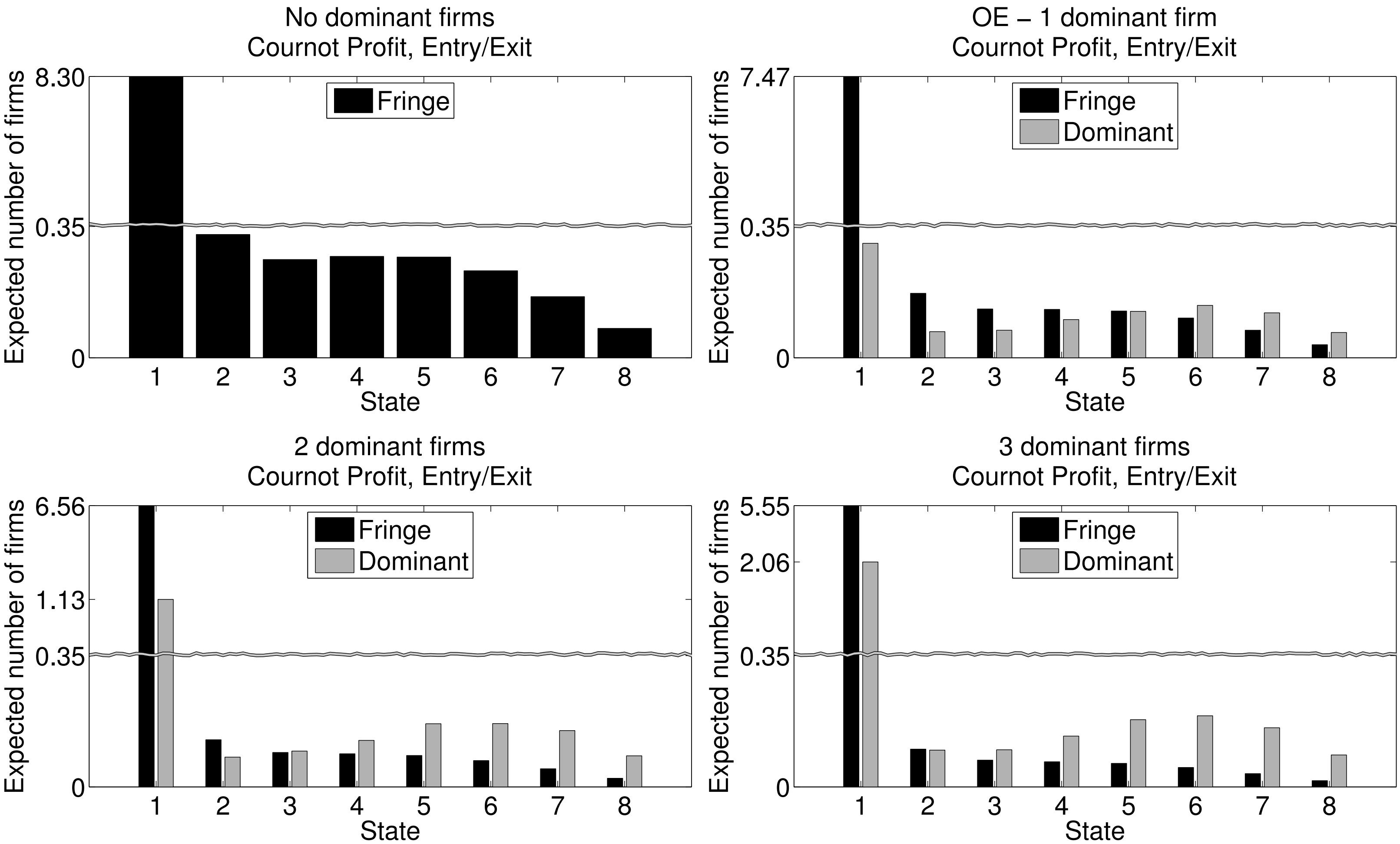

Figure 7: Long-run dominant/fringe state distribution, Cournot, investment cost=5

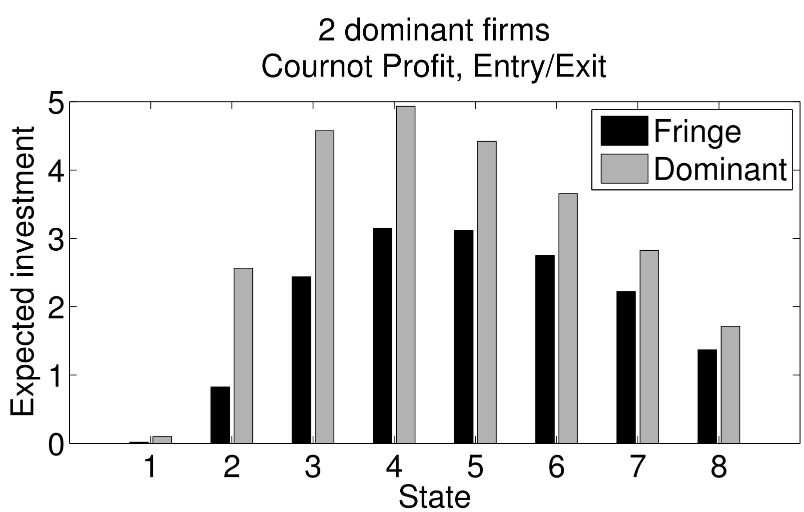

Figure 8: Long-run dominant/fringe investment, POE2, Cournot, investment cost=5

	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

Table 3: Parameters of the revenue function in the CW model. The standard deviation of the preference shock σ is set to 133

Past state	Target size		
	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: Parameters of the adjustment cost in the CW model

	MPE	OE	POE-1	POE-2	POE-3	POE-4	OE-SIM
# Active Firms	4.65	4.37	4.37	4.37	4.39	4.43	4.92
# Sm Firms	2.67	2.51	2.52	2.53	2.55	2.57	2.83
# Med Firms	0.74	0.69	0.69	0.69	0.69	0.70	0.77
# Lg Firms	1.24	1.17	1.16	1.15	1.15	1.16	1.32
# 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 Sm	0.30	0.28	0.28	0.28	0.29	0.29	0.32
Growth into Lg	0.31	0.29	0.29	0.29	0.30	0.30	0.33

Table 5: Industry statistics across different equilibria, CW model

Past state	Target size		
	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

Table 6: Parameters of the adjustment cost in the CW model – interpolation

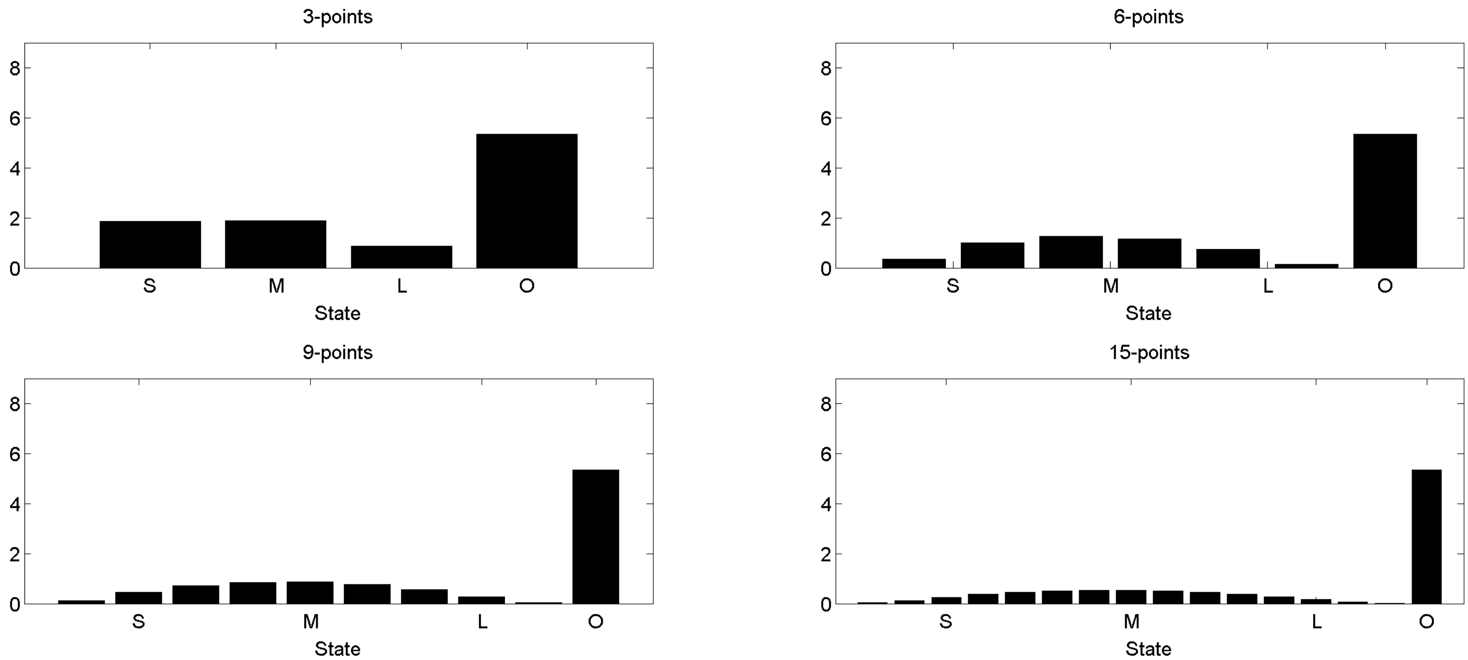

Figure 9: Firm size distribution as we change grid size, CW model

	3 points	6 points	9 points	12 points	15 points
# Active	4.65	4.64	4.65	4.65	4.65
# Sm Firms	1.87	1.34	1.28	1.26	1.26
# Med Firms	1.90	2.42	2.48	2.50	2.50
# Lg Firms	0.88	0.88	0.89	0.89	0.89
# 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

Table 7: Industry statistics as we change grid size, CW model