DISCUSSION PAPER SERIES

No. 11069

VERY SIMPLE MARKOV-PERFECT INDUSTRY DYNAMICS

Jaap H. Abbring, Jeffrey R. Campbell, Jan Tilly and Nan Yang

INDUSTRIAL ORGANIZATION

ISSN 0265-8003

VERY SIMPLE MARKOV-PERFECT INDUSTRY DYNAMICS

Jaap H. Abbring, Jeffrey R. Campbell, Jan Tilly and Nan Yang

Discussion Paper No. 11069
January 2016
Submitted 16 January 2016

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: (44 20) 7183 8801
www.cepr.org

VERY SIMPLE MARKOV-PERFECT INDUSTRY DYNAMICS†

Abstract

This paper develops an econometric model of oligopoly dynamics that can be estimated very quickly from market-level observations of demand shifters and the number of producers. We show that the model has an essentially unique symmetric Markov-perfect equilibrium and provide an algorithm that calculates it quickly. We embed this algorithm in a nested fixed point estimation procedure and apply the result to U.S. local cinema markets. Estimates from County Business Patterns data point to very tough competition for film exhibition rights. Sunk costs make the industry's transition following a permanent demand shock last 10 to 15 years.

JEL Classification: C25, C73 and L13
Keywords: counterfactual policy analysis, demand uncertainty, dynamic oligopoly, firm entry and exit, nested fixed point estimator, sunk costs and toughness of competition

Jaap H. Abbring jaap@abbring.org
CentER, Department of Econometrics & OR, Tilburg University and CEPR

Jeffrey R. Campbell jcampbell@frbchi.org
Federal Reserve Bank of Chicago and CentER, Tilburg University

Jan Tilly jtilly@econ.upenn.edu
University of Pennsylvania

Nan Yang yangnan@nus.edu.sg
National University of Singapore

† We thank Ben Chabot for sharing his practitioner's knowledge of the Motion Picture Theater industry. The research of Jaap Abbring is financially supported by the Netherlands Organisation for Scientific Research (NWO) through Vici grant 453-11-002. Nan Yang gratefully acknowledges the financial support from the National University of Singapore Start-up Grant R-313-000-100-133. The views expressed herein are those of the authors, and they do not necessarily represent those of the Federal Reserve Bank of Chicago, the Federal Reserve System, or its Board of Governors.

1 Introduction

In this paper, we present an econometric model of firm entry, competition, and exit in oligopolistic markets. It features toughness of competition, sunk entry costs, and market-level demand and cost shocks. We allow firms to use mixed strategies and close the model by focusing on symmetric Markov-perfect equilibria. The model's key simplifying assumption is that producers' expected payoffs are identical when entry and survival decisions are made. Using this and the equilibrium implications of mixed strategies for payoffs, we create an algorithm for equilibrium computation that calculates the fixed points of a finite sequence of low-dimensional contraction mappings. Since it relies on contraction mappings, the algorithm is *guaranteed* to calculate an equilibrium. We prove that adding a competitor cannot increase incumbents' equilibrium continuation values. This result in turn ensures that the symmetric equilibrium calculated by our algorithm is essentially unique.

We use these results to develop a nested fixed point (NFXP) procedure that employs market-level panel data on demand shifters and the number of producers and extends Rust's (1987) algorithm for full information maximum likelihood estimation. One novel aspect of our NFXP procedure is that it accounts for the observable implications of mixing over survival actions. Because equilibrium calculation requires so little time, our procedure can (and in our empirical application does) calculate a separate equilibrium for each of several hundred heterogeneous markets at every trial value of the parameter vector.

We use the model for an empirical analysis of entry, competition, and exit in local markets of the U.S. Motion Picture Theaters industry (NAICS 512131). To this end, we first show that our model is suitable to study this industry. We quantify the influence of producer heterogeneity (measured with mid-March employment) on the evolution of the number of firms serving *local* markets using County Business Patterns data from 2000 through 2009 from 573 Micropolitan Statistical Areas (μSAs). Using Poisson regressions, we find that our favored measure of heterogeneity — the Herfindahl-Hirschman Index divided by the value it would attain if producers were homogeneous — has *statistically* significant but *economically* small effects. The average absolute change in the Poisson regressions' forecasts from omitting heterogeneity is 0.036 theaters compared with an average forecast of 2.46 theaters. In this way, the data make a strong case for abstracting from persistent heterogeneity

among theaters within a local market. This is fortunate because our model does so and because high concentration in these local industries precludes applying models that gain comparable simplicity from competition between many firms, like Hopenhayn's (1992) and Weintraub, Benkard, and Van Roy's (2008).

Our estimates imply that adding a single theater to a monopoly market lowers the producers' surplus per consumer by almost half. Adding two more theaters brings the per consumer surplus to 34 percent of its monopoly value. It follows that cinemas compete fiercely in their local markets, despite earlier evidence that local competition has only small effects on box office prices (Davis, 2005). We take this as evidence that movie theaters intensely compete for movie screening rights. The estimated model's sunk costs are substantial, so the initial number of incumbents influences the number of producers for 10 to 15 years following a very persistent demand shock. Without sunk costs, producers' dynamic considerations practically vanish and transition to the long run is almost instantaneous. The industry that is composed of *all* local markets adjusts to persistent demand reductions with both decreased entry and increased exit.

The movie industry is no stranger to large and persistent demand shocks: In the early 1950s, the expansion of television halved movie theater attendance (see e.g. Takahashi, 2015). Current developments like the advance of internet video streaming may pose a similar threat to movie theaters. Netflix, for example, plans to premiere big movies on its video-on-demand service on the same day that they open in cinemas (Kafka, 2013; Harwell, 2015); and Paramount Pictures intends to make some movies available for home viewing only two weeks after their initial theatrical releases (Schwartzel and Fritz, 2015). With this present policy relevance as motivation, we investigate whether a policy that limits competition for screening rights could undo the impact of such a change on the long-run average number of firms. This would be reminiscent of the 1970 Newspaper Preservation Act, which sought to maintain media variety by allowing local newspapers to collude under "joint operating agreements." We find that allowing all theaters to split the monopoly surplus would more than offset the effects of a 25 percent permanent reduction in demand on the number of theaters. A policy that restricts joint operating agreements to duopoly markets would still counter the effects of a 17 percent permanent demand decrease. Such large impacts from changing the toughness of local competition on the number of producers illustrate its economic

importance.

Methodologically, our analysis extends Bresnahan and Reiss's (1990; 1991b) approach to the measurement of the effects of entry on profitability to a dynamic setting. We begin with Abbring and Campbell's (2010) model of last-in first-out oligopoly dynamics. In their model, entry and continuation incur sunk and fixed costs, demand is stochastic, and incumbent firms make continuation decisions sequentially in the order of their entry. They restricted attention to Markov-perfect equilibria in which older firms' first-mover advantages allow them to outlive their otherwise identical younger rivals for sure. They showed that there exists a unique such equilibrium and that — with additional restrictions on the stochastic process for demand — firms use a pure strategy that prescribes entry and continuation if and only if demand exceeds, respectively, entry and exit thresholds on demand (see also Abbring and Campbell, 2015). That result gives a structural interpretation to the dynamic ordered probit estimated by Bresnahan and Reiss (1993). Abbring and Campbell (2010) motivate such a dynamic approach, even if one is only interested in static aspects of competition, by showing that option value considerations in markets with sunk costs and uncertainty bias estimates of Bresnahan and Reiss's (1991b) static entry model towards finding tough competition.

In this paper, we replace the (admittedly special) sequential continuation decisions of Abbring and Campbell (2010) with simultaneous ones. Our equilibrium characterization and computation do not rely on Abbring and Campbell's sequential timing assumptions, nor on their restriction to last-in first-out dynamics, but instead leverage the properties of mixed-strategy equilibria in a novel way. Also, we provide a more complete econometric development of our model. To this end, we add a market-level shock to both potential entrants' sunk costs of entry and incumbents' fixed costs of continuation. This is observed by market participants but not by the econometrician. As in Bresnahan and Reiss (1990, 1991b), this market-level cost shock serves as the model's econometric error.

Our model can be viewed as a special case of Ericson and Pakes's (1995) Markov-perfect industry dynamics framework. Ericson and Pakes focused on equilibria in pure strategies, but Doraszelski and Satterthwaite (2010) showed that such equilibria might not exist in their original framework. To address this problem, Gowrisankaran (1999) added privately-observed firm-specific shocks to the costs of continuation; and Doraszelski and Satterthwaite formally proved that such an

augmented framework has an equilibrium in pure strategies. Research following Ericson and Pakes (summarized by Doraszelski and Pakes, 2007) has generally adopted this augmented version of their framework.

This paper returns to Ericson and Pakes's original complete-information approach. We show that the firm-specific shocks that guarantee existence of an equilibrium in pure strategies in the augmented framework obscure a useful consequence of firms employing mixed strategies: Because, in equilibrium, firms earn the value of the outside option (zero) whenever they nontrivially randomize over exit and survival, *symmetric equilibrium payoffs to incumbents contemplating survival equal either zero or the value of all incumbents choosing certain continuation.* This insight allows us to calculate continuation values from some nodes of the game tree without knowing everything about the game's subsequent play. Combining this insight with a demonstration that continuation payoffs weakly decrease with the number of active firms also yields the contraction mappings we use both to calculate the equilibrium and to demonstrate its uniqueness. Although we apply the insight here to a model of homogeneous firms, it can be applied (with effort) to characterize the equilibria of models with persistent heterogeneity. Indeed, Abbring, Campbell, and Yang (2015) provide equilibrium existence, uniqueness, and algorithmic results for an extension of this paper's model with persistent firm-specific technology shocks for industries with at most two active firms.

Because our model has a unique equilibrium that is easy to compute, we can straightforwardly estimate it with maximum likelihood, which is asymptotically efficient. The maximum likelihood estimator has a standard asymptotic distribution that is easy to calculate. Moreover, it can incorporate market-specific covariates in a standard and transparent way by calculating a separate equilibrium for each market. This makes our model particularly useful when heterogeneity across markets dominates heterogeneity across firms within a market.

In contrast, there is no guarantee that the augmented Ericson and Pakes framework has a unique equilibrium; and computing even one of its equilibria can be onerous. Therefore, methods for its estimation — such as Bajari, Benkard, and Levin's (2007) and Pakes, Ostrovsky, and Berry's (2007) — have avoided equilibrium calculation altogether. Instead, they assume that producers in all sample markets employ the same equilibrium strategy, use this to estimate firms' expected behavior from their observed choices, and identify the structural parameters of interest from

the implied individual choice problems. This "two-step" approach cleverly solves the problem of estimation when there could be equilibrium multiplicity. However, unlike the maximum likelihood estimator of our model, the two-step estimators are generally not efficient. Moreover, their asymptotic distributions may be hard to compute; in fact, Bajari, Benkard, and Levin (p. 1349) "believe it will typically be easiest to use subsampling or the bootstrap to estimate standard errors." Finally, users of the two-step approach have had to coarsen covariates before estimation so that they could estimate transition probabilities by pooling data across markets (e.g. Dunne, Klimek, Roberts, and Xu, 2013).

Even with two-step estimates of the structural parameters of the augmented Ericson and Pakes framework in hand, equilibrium multiplicity seriously impedes its practical application. After all, structural parameters are rarely of interest *per se* but rather are intermediate inputs into the analysis of environmental changes and policy interventions. At best, equilibrium multiplicity substantially complicates these counterfactual equilibrium analyses. In contrast, conducting counterfactual policy analysis with our model is straightforward. For example, we repeatedly calculate the model's unique equilibrium (for each of the 573 markets in our sample) to determine the size of a permanent demand shock that a joint operating agreement exactly offsets. Of course, not every important question concerning dynamic oligopolies can be cast within an environment with a unique equilibrium; and further methodological developments for policy analysis with multiple equilibria (such as calculating bounds on counterfactuals as in Eizenberg, 2014, and Reguant, 2015) will undoubtedly help complete this final step towards industrial organization's quest to overcome the Lucas critique. Nevertheless, we believe that improving our theoretical understanding of the equilibrium set of Ericson and Pakes's framework, as we do in this paper and elsewhere, will also substantially contribute to this long-run research goal.

Existing empirical research sometimes uses models that are somewhat similar to ours but with less judicious assumptions. For example, the one-location model in Pakes, Ostrovsky, and Berry (2007), which was used by Dunne, Klimek, Roberts, and Xu (2013), is basically the model of this paper with idiosyncratic shocks replacing our aggregate shock. Using a high-cost model when a low-cost close substitute is available wastes research resources. Although our model will not fit every possible circumstance, having it ready for use when appropriate can only

increase the pace of empirical progress in this area.

Our computational approach is reminiscent of applying backward induction to compute the equilibria of dynamic directional games, as in Cabral and Riordan (1994) and Judd, Schmedders, and Yeltekin (2012). Iskhakov, Rust, and Schjerning (2015b) systemize this familiar procedure into an algorithm for computing *all* these games' equilibria. In the games considered, the state space can be partially ordered using primitive restrictions on the state's transitions: State A comes after state B if A can be reached from B but not the other way around. Iskhakov, Rust, and Schjerning's algorithm iterates backwards through this partially ordered set of states. Transitions from states considered in a given iteration to states considered in later iterations are impossible, so the algorithm can calculate equilibrium outcomes and continuation values recursively. Our algorithm similarly iterates over an ordered partition of our game's state space. However, our game is *not* directional and in each iteration transitions to states not yet visited by our algorithm *are* possible. Instead of exploiting the directionality of state transitions hard wired into the primitives of Iskhakov, Rust, and Schjerning's framework, we rely on the fact that the expected symmetric equilibrium payoff in any survival subgame in which firms exit with positive probability must be zero. This allows us to order one state after another state if the latter can be reached from the former but the opposite transition requires firms to choose exit with a positive probability.

The remainder of the paper proceeds as follows. The next section presents the model's primitives, and Section 3 discusses equilibrium existence, uniqueness, and computation. Section 4 develops the model's empirical implementation, which includes sampling, likelihood construction, and maximum likelihood estimation using the NFXP procedure. Section 5 applies our econometric model to the Motion Picture Theaters industry. Section 6 concludes with a perspective on extending this paper's approach to models with persistent firm-specific shocks, as in Ericson and Pakes's original framework, along the lines of Abbring, Campbell, and Yang (2015).

Five appendices provide supporting results. Appendices A, B, and C present a slightly more general version of the model that dispenses with assumptions motivated by econometric considerations and provides proofs of all our equilibrium characterization results for that encompassing framework. Appendix D provides conditions for our model's identification based on minimal and readily available panel data on the number of producers and demand shifters. The structure implied

by the firms' use of a mixed strategy identifies the scale of the econometric error's distribution without *a priori* restrictions on how demand or another observable regressor influences payoffs. This distinguishes our identification results from those for dynamic single agent models (Magnac and Thesmar, 2002), which require either a known distribution of the econometric error or further structure on the payoffs. Appendix E reports results from Monte Carlo experiments that demonstrate the NFXP estimator's small-sample accuracy and light computational demands.

2 The Model

Consider a market in discrete time indexed by $t \in \mathbb{N} \equiv \{1, 2, \ldots\}$. In period t, firms that have entered in the past and not yet exited serve the market. Each firm has a name $f \in \mathcal{F} \equiv \mathbb{N} \times \mathbb{N}$. The first component of a firm's name gives the date in which it has its only opportunity to enter the market, and the second component gives its position in that date's entry queue. Aside from the timing of their entry opportunities, the firms are identical.[1]

Figure 1 details the actions taken by firms in period t and their consequences for the game's state at the start of period $t+1$. We call this the game's *recursive extensive form.* For expositional purposes, we divide each period into two subperiods, the entry and survival subgames. Play in period t begins on the left with the entry subgame. If $t = 1$, nature sets N_1 (the number of firms serving the market in period 1) and C_1 (the initial demand state); if $t > 1$, these are inherited from the previous period. We use $\mathcal{C}$ to denote the support of C_t. Although we consistently refer to C_t as "demand," it can encompass any observed, relevant, and time-varying characteristics of the market.

Each incumbent firm earns a surplus $\pi(N_t, C_t)$ from serving the market, and all firms value future profits and costs with the discount factor $\rho \in [0, 1)$. We assume that

- $\exists \check{\pi} < \infty$ such that $\forall n \in \mathbb{N}$ and $\forall c \in \mathcal{C}$, $\mathbb{E}[\pi(n, C') | C = c] \leq \check{\pi}$;
- $\exists \check{n} \in \mathbb{N}$ such that $\forall n > \check{n}$ and $\forall c \in \mathcal{C}$, $\pi(n, c) = 0$; and
- $\forall n \in \mathbb{N}$ and $\forall c \in \mathcal{C}$, $\pi(n, c) \geq \pi(n + 1, c)$.

[1] As noted in the Introduction, we only require firms to have identical *expected* payoffs when making entry and survival decisions. See Sections 5.2 and 6 for further discussion.

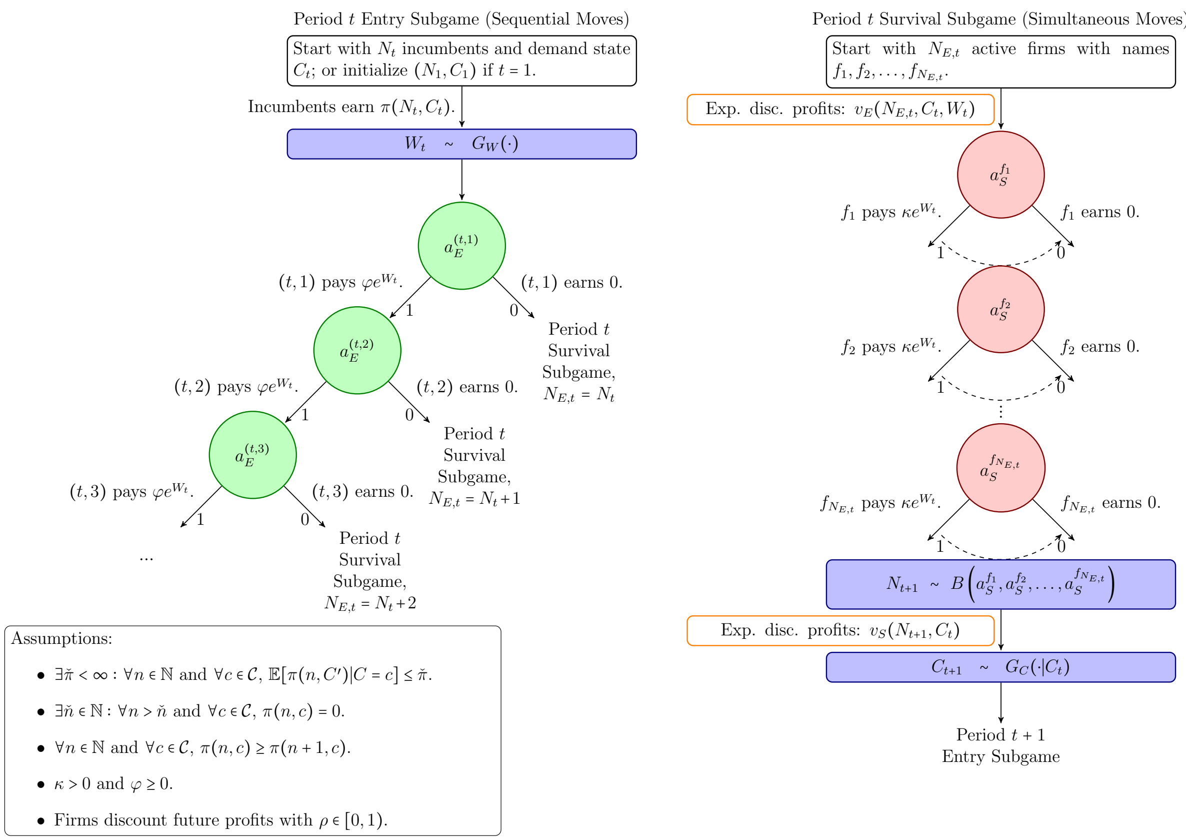


Figure 1: The Model's Recursive Extensive Form

Here and throughout; we denote the next period's value of a generic variable Z with Z', random variables with capital Roman letters, and their realizations with the corresponding small Roman letters. The first assumption is technical and allows us to restrict equilibrium values to the space of bounded functions. We will use the second assumption to bound the number of firms that will participate in the market simultaneously. It is not restrictive in empirical applications to oligopolistic markets. The third assumption requires the addition of a competitor to reduce weakly each incumbent's surplus. That is, what Sutton (1991) labelled the *toughness of competition* must dominate any complementarities between firms' activities like those considered by Honoré and De Paula (2010).

After incumbents earn their surpluses, nature draws the current period's shock to continuation and entry costs, W_t, from a distribution G_W with positive density everywhere on the real line. Then, entry *per se* can take place. The period t entry cohort consists of firms with names in $\{t\} \times \mathbb{N}$. These firms make their entry decisions

sequentially in the order of their names' second components. We denote firm f's entry decision with $a_E^f \in \{0,1\}$. An entrant ($a_E^f = 1$) pays the sunk cost $\varphi \exp(W_t)$, with $\varphi \geq 0$. A firm choosing not to enter ($a_E^f = 0$) earns a payoff of zero and never has another entry opportunity. Such a refusal to enter also ends the entry subgame, so firms remaining in this period's entry cohort that have not yet had an opportunity to enter *never* get to do so. Since the next firm in line faces exactly the same choice as did the firm that refused to enter, this convenient assumption does not affect any symmetric equilibrium outcome. Since every period has at least one firm refusing an available entry opportunity, the model is one of free entry.

The total number of firms in the market after the entry stage equals $N_{E,t}$, which sums the incumbents with the actual entrants. Denote their names with $f_1, \ldots, f_{N_{E,t}}$. In the survival subgame, these firms simultaneously choose probabilities of remaining active, $a_S^{f_1}, \ldots, a_S^{f_{N_{E,t}}} \in [0,1]$.[2] Subsequently, all survival outcomes are realized independently across firms according to the chosen Bernoulli distributions. Firms that survive pay a fixed cost $\kappa \exp(W_t)$, with $\kappa > 0$. Firms that exit earn 0 and never again participate in the market. The N_{t+1} surviving firms continue to the next period, $t+1$.[3] To end the period, nature draws a new demand state C_{t+1} from the conditional distribution $G_C(\cdot \mid C_t)$.

The timing of our game is similar to that in Ericson and Pakes (1995). Like them, we allow for unlimited sequential entry.[4] Such free entry is a standard and convenient assumption in empirical applications in which we cannot identify a finite set of potentially active firms. We could straightforwardly adapt the analysis to a

[2] We do not explicitly model the firms' randomization devices. A more complete development would assign each active firm an independent uniformly-distributed random variable that is irrelevant for payoffs and have each firm choose a set of realizations that direct it to survive. In this extension, a survival probability equal to one could indicate either that the firm chooses to exit *never* or that it chooses to exit whenever its random variable falls into a particular non-empty set of measure zero. Throughout this paper, we will assume the former and interpret $a_S = 0$ and $a_S = 1$ as dictating *certain* exit and survival.

[3] The assumption that entrants immediately contemplate exit might seem strange, but exit immediately following entry never occurs in equilibrium. However, this timing assumption removes an unrealistic possibility. If entrants did not make these continuation decisions, then they could effectively commit to continuation. This would allow an entrant to displace an incumbent only by virtue of this commitment power.

[4] See their page 60:

> We assume that, in each period, ex ante identical firms decide to enter sequentially until the expected value of entry falls sufficiently to render further entry unprofitable.

finite set of potential entrants (each of which has exactly one entry opportunity) that make entry choices simultaneously, but this would introduce "entry mistakes" as in Cabral (2004). Like Ericson and Pakes, and unlike Abbring and Campbell (2010), we assume simultaneous survival decisions. Because we allow for mixed survival rules, this may lead to excessive exits: Fewer firms may end up surviving than could have continued profitably. Since entry precedes exit, potential entrants cannot take immediate advantage of such "exit mistakes" and thereby outmaneuver incumbents. This is not so relevant to Ericson and Pakes, who restrict strategies to be pure (at the expense of losing equilibrium existence; see Doraszelski and Satterthwaite, 2010).

Before continuing to the model's analysis, we review some of its key assumptions from the perspective of its econometric implementation using market-level panel data. In Section 4, we will assume that, for each market, the data contain information on N_t, C_t, and possibly some time-invariant market characteristics X that shift the market's primitives. The market-level cost shocks W_t are not observed by the econometrician and serve as the model's structural econometric errors. Because they are observed by all firms and affect their payoffs from entry and survival, they make the relation between the observed demand state C_t and the market structure N_t statistically nondegenerate. Bresnahan and Reiss (1991a) noted that equilibrium multiplicity can arise in static games with econometric errors that have complete support and are at least somewhat independent across both players and outcomes. Our specification of a single shock to all firms' continuation and entry costs avoids this difficulty. The assumptions on $\{C_t, W_t\}$ make it a first-order Markov chain satisfying Rust's (1987) conditional independence assumption.[5] This ensures that the distribution of (N_t, C_t) conditional on $(N_{t^\star}, C_{t^\star})$ for all $t^\star < t$ depends only on (N_{t-1}, C_{t-1}), so we require only the model's transition rules to calculate the conditional likelihood function.

3 Equilibrium

We assume that firms play a symmetric Markov-perfect equilibrium (Maskin and Tirole, 1988), a subgame-perfect equilibrium in which all firms use the same Markov

[5] Rust (1987) defines "conditional independence" for a *controlled* Markov process, but his definition specializes to our case of an externally specified process $\{C_t, W_t\}$ if we take the control to be trivial. Rust's conditional independence assumption allows both W_t and C_t to depend on C_{t-1}. Our analysis easily extends to this case.

strategy.

3.1 Markov Strategies

A Markov strategy maps *payoff relevant states* into actions. When a potential entrant (t, j) makes its entry decision in period t, the payoff-relevant states are the number of firms in the market plus the current period's potential entrants up to and including (t, j), $M_t^j \equiv N_t + j$, the current demand C_t, and the cost shock W_t. We collect these into the vector $(M_t^j, C_t, W_t) \in \mathcal{H} \equiv \mathbb{N} \times \mathcal{C} \times \mathbb{R}$. Similarly, we collect the payoff-relevant state variables of a firm f contemplating survival in period t in the $\mathcal{H}$-valued $(N_{E,t}, C_t, W_t)$. Since survival decisions are made simultaneously, this state is the same for all active firms. A Markov strategy is a pair of functions $a_E : \mathcal{H} \to \{0, 1\}$ and $a_S : \mathcal{H} \to [0, 1]$. The *entry rule* a_E assigns a binary indicator of entry to each possible state. Similarly, a_S gives a survival *probability* for each possible state. Since time and firms' names themselves are not payoff relevant, we henceforth drop the subscript t and the superscript j from the payoff-relevant states.

3.2 Symmetric Markov-Perfect Equilibrium

In a symmetric Markov-perfect equilibrium, a firm's expected continuation value at a particular node of the game can be written as a function of that node's payoff-relevant state variables. Two of these value functions are particularly useful for the model's equilibrium analysis: the *post-entry* value function, v_E, and the *post-survival* value function, v_S. The post-entry value $v_E(N_E, C, W)$ equals the expected discounted profits of a firm in a market with demand state C, cost shock W, and N_E firms just after all entry decisions are made. The post-survival value $v_S(N', C)$ equals the expected discounted profits from being active in the same market with N' firms just after the survival outcomes are realized. The post-survival value does not depend on W because that cost shock has no forecasting value and is not directly payoff relevant after survival decisions are made. Figure 1 shows the points in the survival subgame where these value functions apply.

Because the payoff from leaving the market is zero, a firm's post-entry value in a state (n_E, c, w) equals the probability that it survives, $a_S(n_E, c, w)$, times the expected payoff from surviving. The latter equals this firm's expected post-survival

value net of its fixed costs of survival. So, v_E and v_S satisfy

$$v_E(n_E, c, w) = a_S(n_E, c, w)\left(\mathbb{E}_{a_S}\left[v_S(N', c)\middle| N_E = n_E, C = c, W = w\right] - \kappa \exp(w)\right).$$

The expectation $\mathbb{E}_{a_S}$ over N' takes survival of the firm of interest as given. That is, it takes N' to be one plus the outcome of $n_E - 1$ independent Bernoulli (survival) trials with success probability $a_S(n_E, c, w)$. Its subscript makes its dependence on a_S explicit. It conditions on the current values of C and W because these influence the survival probability's value.

A firm's post-survival value equals the expected sum of the surplus and post-entry value that accrue to the firm in the next period, discounted to the current period with ρ:

$$v_S(n', c) = \rho \mathbb{E}_{a_E}\left[\pi(n', C') + v_E(N'_E, C', W')\middle| N' = n', C = c\right].$$

Here, $\mathbb{E}_{a_E}$ is an expectation over the next period's demand state C', cost shock W', and post-entry number of firms N'_E. This expectation operator's subscript indicates its dependence on a_E. In particular, given $N' = n'$, N'_E is a deterministic function of $a_E(\cdot, C', W')$.

A strategy (a_E, a_S) forms a symmetric Markov-perfect equilibrium with payoffs (v_E, v_S) if and only if no firm can gain from a one-shot deviation from its prescriptions. Thus, given the pair of payoff functions (v_E, v_S), their corresponding strategy must satisfy

$$\begin{aligned} a_E(m, c, w) &\in \arg\max_{a \in \{0,1\}} a\Big(\mathbb{E}_{a_E}\left[v_E(N_E, c, w) | M = m, C = c, W = w\right] - \varphi \exp(w)\Big), \\ a_S(n_E, c, w) &\in \arg\max_{a \in [0,1]} a\Big(\mathbb{E}_{a_S}\left[v_S(N', c) | N_E = n_E, C = c, W = w\right] - \kappa \exp(w)\Big). \end{aligned}$$

Before proceeding to the equilibrium analysis, we wish to note and dispense with an uninteresting source of equilibrium multiplicity. If a potential entrant is indifferent between its two choices, we can construct one equilibrium from another by varying only that choice. Similarly, an incumbent monopolist can be indifferent between continuation and exit, and we can construct one equilibrium from another by changing that choice alone. To avoid such uninteresting caveats to our results, we follow Abbring and Campbell (2010) by focusing on equilibria that *default to*

inactivity. In such an equilibrium, a potential entrant that is indifferent between entering or not stays out,

$$\mathbb{E}_{a_E}\left[v_E(N_E, c, w) | M = m, C = c, W = w\right] = \varphi \exp(w) \Rightarrow a_E(m, c, w) = 0,$$

and an active firm that is indifferent between *all* possible outcomes of the survival stage exits,

$$v_S(n_E, c) = \cdots = v_S(1, c) = \kappa \exp(w) \Rightarrow a_S(n_E, c, w) = 0.$$

The restriction to equilibria that default to inactivity does *not* restrict the game's strategy space. Hereafter, we require the strategy underlying a "symmetric Markov-perfect equilibrium" to default to inactivity. When W follows a continuous distribution, an exact indifference between activity and inactivity occurs with probability zero. For this reason, the restriction to equilibria that default to inactivity is very, very weak.

3.3 Existence, Uniqueness, and Computation

This subsection presents our analysis of equilibrium existence, uniqueness, and computation. For this, three features of the model are disposable: the serial independence of W_t, the additive separability of per-period surplus from the costs of continuation, and the invariance of sunk costs to the number of firms and the current demand state. Appendices A and B generalize the model by relaxing these assumptions and Appendix C provides proofs of all of this subsection's results extended to that more general model.

We start by noting that the assumption that per-period surplus equals zero if more than $\check{n}$ firms serve the market bounds the long-run number of firms in equilibrium.

Lemma 1 (Bounded number of firms) *In a symmetric Markov-perfect equilibrium that defaults to inactivity,* $\forall c \in \mathcal{C}$ *,* $\forall w \in \mathbb{R}$*, and* $\forall n > \check{n}$*;* $a_E(n, c, w) = 0$ *and* $a_S(n, c, w) < 1$*.*

Intuitively, in a symmetric equilibrium, firms cannot survive for sure with $n > \check{n}$ firms because this would give them negative payoffs (which firms could avoid by

exiting instead). To see this, note that if all firms continue for sure, each would incur a positive continuation cost and earn a zero surplus one or more times (due to our assumption $\pi(n,c) = 0$ for all $n > \check{n}$), and then collect a zero post-entry value in the first future period in which firms leave with positive probability. So, $a_S(n,c,w) < 1$ and $v_E(n,c,w) = 0$. Because no firm would be willing to pay a positive sunk cost to enter a survival subgame with zero expected payoff, $a_E(n,c,w) = 0$.[6]

In equilibrium, the market can have more than $\check{n}$ active firms only if $N_1 > \check{n}$. Because these firms exit with positive probability until there are $\check{n}$ or fewer of them, N_t must eventually enter $\{0,1,\ldots,\check{n}\}$ permanently. Consequently, the equilibrium analysis hereafter focuses on the restrictions of a_E, v_E, and a_S to $\{1,2,\ldots,\check{n}\}\times\mathcal{C}\times\mathbb{R} \subset \mathcal{H}$ and of v_S to $\{1,2,\ldots,\check{n}\}\times\mathcal{C}$. Extending an equilibrium strategy over this restricted state space to the full state space is straightforward.

The next step in the equilibrium analysis uses the assumption that *per-period* surplus weakly decreases with the number of competitors to show that the same monotonicity applies to the post-survival value functions.

Lemma 2 (Monotone equilibrium payoffs) *In a symmetric Markov-perfect equilibrium that defaults to inactivity,* $\forall c \in \mathcal{C}$, $v_S(n',c)$ *weakly decreases with* n'.

Lemma 2 says that no endogenous complementarity between firms arises in equilibrium. Although this is intuitive, it is not a trivial result. Indeed, in a more general model, one can construct a counter example in which two firms with *persistently* high productivity complement each other's presence by jointly deterring the entry of two or more low-productivity firms.

To appreciate the implications of Lemma 2, consider a one-shot simultaneous-moves survival game played by n_E active firms. In it, each of the n' survivors earns $-\kappa\exp(w) + v_S(n',c)$, where v_S is the post-survival value in a symmetric Markov-perfect equilibrium of our dynamic game, and each exiting firm earns zero. The Nash equilibria of this game are intimately connected to the Markov-perfect equilibria of our model. In particular, a survival rule from a symmetric Markov-perfect equilibrium forms a symmetric Nash equilibrium of the one-shot game, and vice versa.

This one-shot game has many equilibria in the trivial case that $v_S(n_E,c) = \cdots = v_S(1,c) = \kappa\exp(w)$. In this case, our restriction to equilibria that default to

[6] If $\phi = 0$, then the restriction that the equilibrium strategy defaults to inactivity also requires that $a_E(n,c,w) = 0$.

inactivity picks $a_S(n_E, c, w) = 0$. In the more interesting case where $v_S(n', c) \neq \kappa \exp(w)$ for at least one $n' \in \{1, \ldots, n_E\}$, Lemma 2 guarantees that the one-shot game has a *unique* symmetric Nash equilibrium. To show this, we distinguish three subcases.

- First, suppose that $v_S(1, c) \leq \kappa \exp(w)$. Lemma 2 implies that $v_S(n', c) \leq \kappa \exp(w)$ for all $n' \in \{1, \ldots, n_E\}$. Therefore, exiting for sure (setting $a_S(n_E, c, w) = 0$) is a weakly dominant strategy and forms one symmetric equilibrium. Furthermore, since $v_S(n', c) \neq \kappa \exp(w)$ for at least one $n' \in \{1, \ldots, n_E\}$, we know that $v_S(n_E, c) < \kappa \exp(w)$. Therefore, exiting for sure is also the unique best response to any positive symmetric continuation probability.

- Next, suppose that $v_S(n_E, c) \geq \kappa \exp(w)$. Lemma 2 implies that $v_S(n', c) \geq \kappa \exp(w)$ for $n' = 1, \ldots, n_E$. Therefore, continuing for sure (setting $a_S(n_E, c, w) = 1$) is a weakly dominant strategy and forms one symmetric equilibrium. Since $v_S(n', c) \neq \kappa \exp(w)$ for at least one $n' \in \{1, \ldots, n_E\}$, we also know that $v_S(1, c) > \kappa \exp(w)$. Therefore, continuing for sure is also the unique best response to any continuation probability less than one.

- For the last subcase, suppose that $v_S(1, c) > \kappa \exp(w) > v_S(n_E, c)$. No symmetric pure strategy equilibrium exists, because the best response to all other firms continuing for sure is to exit for sure, and vice versa. In a mixed strategy equilibrium, firms must be indifferent between continuation and exit:

$$\sum_{n'=1}^{n_E} \binom{n_E - 1}{n' - 1} a_S^{n'-1} (1 - a_S)^{n_E - n'} (v_S(n', c) - \kappa \exp(w)) = 0.$$

  By the intermediate value theorem, there is some $a_S \in (0, 1)$ that solves this indifference condition.[7] This establishes existence of a mixed strategy equilibrium. Lemma 2 and this case's preconditions together guarantee that the left hand side strictly decreases in a_S. Therefore, there is only one symmetric mixed strategy equilibrium.

For future reference, we state this equilibrium uniqueness result with

[7]Note that the left hand side continuously varies with a_S from a positive value when a_S is near 0 to a negative value when a_S is near 1.

Corollary 1 *Let v_S be the post-survival value function associated with a symmetric Markov-perfect equilibrium that defaults to inactivity. Consider the one-shot survival game in which n_E firms simultaneously choose between survival and exit (as in the survival subgame of Figure 1), each of the n' survivors earns $v_S(n', c) - \kappa \exp(w)$ with $v_S(n', c) - \kappa \exp(w) \neq 0$ for at least one $n' \in \{1, \dots, n_E\}$, and each exiting firm earns zero. This game has a unique symmetric Nash equilibrium, possibly in mixed strategies.*

When the individual payoff from joint continuation is positive, this unique Nash-equilibrium strategy from Corollary 1 guarantees that firms survive for sure and receive this payoff. In all other states, each firm is either indifferent between surviving and exiting or prefers to exit for sure; and following the strategy gives each of them an expected payoff of zero. The assumption that a Markov-perfect strategy defaults to inactivity ensures that the post entry payoff equals zero in the trivial case with $v_S(n_E, c) = \cdots = v_S(1, c) = \kappa \exp(w)$ excluded by Corollary 1. Thus, we also have

Corollary 2 *If v_E and v_S are the post-entry and post-survival value functions associated with a symmetric Markov-perfect equilibrium that defaults to inactivity, then*

$$v_E(n_E, c, w) = \max\{0, v_S(n_E, c) - \kappa \exp(w)\}.$$

Note that Corollary 2 in combination with Lemma 2 implies that $v_E(n_E, c, w)$ also weakly decreases with n_E.

With Corollaries 1 and 2 in hand, we proceed to demonstrate equilibrium existence constructively. Our equilibrium uniqueness result and algorithm for equilibrium calculation follow from the construction as byproducts. Begin with calculating $v_E(\check{n}, \cdot, \cdot)$ and $v_S(\check{n}, \cdot)$. From Lemma 1, there will be no entry in a period starting with $\check{n}$ firms, so

$$v_S(\check{n}, c) = \rho \mathbb{E}\Big[\pi(\check{n}, C') + v_E(\check{n}, C', W') \Big| C = c\Big].$$

This and Corollary 2 give us

$$v_E(\check{n}, c, w) = \max\left\{0, \rho \mathbb{E}\Big[\pi(\check{n}, C') + v_E(\check{n}, C', W') \Big| C = c\Big] - \kappa \exp(w)\right\}.$$

The right-hand side defines a contraction mapping on the complete space of bounded functions on $\mathcal{C}\times\mathbb{R}$, with a unique fixed point $v_E(\check{n},\cdot,\cdot)$. Although we are constructing a *candidate* equilibrium, the fixed point's uniqueness implies that this is the only possible equilibrium post-entry value. This fixed point immediately yields $v_S(\check{n},\cdot)$. Again, this is the only possible candidate value. Finally, any entry rule that is consistent with these payoffs and individual optimality that also defaults to inactivity must dictate entry into a market with $\check{n}-1$ incumbents if and only if the payoff from doing so is positive. That is

$$a_E(\check{n},c,w) = \mathbb{1}\left[v_E(\check{n},c,w) > \varphi\exp(w)\right].$$

Here, $\mathbb{1}[x] = 1$ if x is true and equals 0 otherwise.

With $v_E(\check{n},\cdot,\cdot)$ and $a_E(\check{n},\cdot,\cdot)$ calculated, the construction of the remaining candidate value functions and entry rules proceeds recursively. For given n, suppose that $v_E(n',\cdot,\cdot)$ and $a_E(n',\cdot,\cdot)$ for $n' = n+1, n+2, ..., \check{n}$ are in hand, and define

$$\mu(n,c,w) \equiv n + \sum_{n'=n+1}^{\check{n}} \prod_{m=n+1}^{n'} a_E(m,c,w) = n + \sum_{m=n+1}^{\check{n}} a_E(m,c,w).$$

This is the number of firms that will be active in a period that starts with n firms after all that period's potential entrants have followed the candidate entry rule. The second equality follows from the monotonicity of v_E (and hence a_E) and the fact that a_E takes values in $\{0,1\}$. Lemma 1 implies that there can be at most $\check{n}-n$ entrants when the period starts with n active firms.

For $n<\check{n}$, Corollary 2 implies

$$v_E(n,c,w) = \max\left\{0, \rho\mathbb{E}\Big[\pi(n,C') + v_E(\mu(n,C',W'),C',W')\Big|C=c\Big] - \kappa\exp(w)\right\}.$$

Given the values of $v_E(n',\cdot,\cdot)$ for $n' = n+1,\ldots,\check{n}$, the right-hand side defines a contraction mapping with $v_E(n,\cdot,\cdot)$ as its unique fixed point. This allows us to obtain $v_S(n,\cdot)$. Finally, a firm in state (n,c,w) enters if and only if

$$\mathbb{E}_{a_E}\left[v_E(N_E,c,w)|M=n,C=c,W=w\right] > \varphi\exp(w).$$

Since Lemma 2 implies that further entry cannot make an incumbent better off, this inequality implies that the firm would enter in the absence of further entry,

$v_E(n,c,w) > \varphi \exp(w)$. On the other hand, because later entrants pay the same entry costs, further entry will never take post-survival values below $\varphi \exp(w)$, so $v_E(n,c,w) > \varphi \exp(w)$ also implies that entry's expected payoff exceeds its cost. Therefore, any equilibrium entry rule consistent with $v_E(n,c,w)$ must satisfy [8]

$$a_E(n,c,w) = \mathbb{1}\left[v_E(n,c,w) > \varphi \exp(w)\right].$$

When this recursion is complete, we have the unique continuation values and entry rules that are consistent with an equilibrium. To determine a candidate survival rule a_S, we set $a_S(n_E,c,w) = 0$ for all (n_E,c,w) such that $v_S(n_E,c) = \cdots = v_S(1,c) = \kappa \exp(w)$ and find an equilibrium to Corollary 1's one-shot survival game for all other (n_E,c,w). *If* the candidate is actually an equilibrium, then Corollary 1 guarantees that these survival rules are unique. This is indeed the case.

Theorem 1 (Equilibrium existence and uniqueness) *There exists a unique symmetric Markov-perfect equilibrium that defaults to inactivity.*

4 Empirical Implementation

The previous section shows that there exists a unique symmetric Markov-perfect equilibrium for given primitives π, κ, φ, ρ, G_C, and G_W. Given (N_1, C_1), this equilibrium induces a distribution for the process $\{N_t, C_t\}$. This section studies how observations of this process from a panel of markets can be used to estimate the model's primitives.

4.1 Sampling

Suppose that we have data from $\check{r}$ markets indexed with $r = 1,\ldots,\check{r}$. For each market, we observe the number of active firms $N_{r,t}$ and the demand state $C_{r,t}$ in each period $t = 1,\ldots,\check{t}$; for some $\check{t} \geq 2$. We also observe some time-invariant characteristics of each market, which we store in a vector X_r. However, we have no observations of the cost shocks $W_{r,t}$. Our estimation utilizes no direct observations of firms' input choices, sales volumes, costs, revenues, or profits. Including such information (for

[8] In the more general model of the Appendix, we assume that the sunk costs of entry weakly increase with the number of firms already committed to production in the next period. The logic of this paragraph applies straightforwardly to that setting.

example, by equating profits to those from Cournot competition with an estimated demand curve and constant marginal costs) into the econometric analysis is feasible.

We assume that $(\{N_{r,t}, C_{r,t}; t = 1, \ldots, \check{t}\}, X_r)$ is distributed independently across markets.[9] The initial conditions $(N_{r,1}, C_{r,1}, X_r)$ are drawn from a distribution that we leave unspecified. Thereafter, industry dynamics follow the transition rules implied by Section 3's unique equilibrium, with primitives $\pi_r(\cdot,\cdot) = \pi(\cdot,\cdot\,|X_r, \theta_P)$, $\kappa_r = \kappa(X_r, \theta_P)$, $\varphi_r = \varphi(X_r, \theta_P)$, and $\rho_r = \rho(X_r, \theta_P)$ for some finite vector θ_P; $G_{C,r}(\cdot \mid \cdot) = G_C(\cdot \mid \cdot\,; X_r, \theta_C)$ for some finite vector θ_C; and $G_{W,r}(\cdot) = G_W(\cdot\,; X_r, \theta_W)$ for some finite vector θ_W.[10]

4.2 Likelihood

We focus on inferring the structural parameters $\theta \equiv (\theta_P, \theta_C, \theta_W)$ from the conditional likelihood $\mathcal{L}(\theta)$ of θ for data on market dynamics $\{N_{r,t}, C_{r,t}; t = 2, \ldots, \check{t}; r = 1, \ldots, \check{r}\}$ given the initial conditions $(N_{r,1}, C_{r,1}, X_r; r = 1, \ldots, \check{r})$.[11] Using the model's Markov structure and conditional independence, this likelihood can be written as $\mathcal{L}(\theta) = \mathcal{L}_C(\theta_C) \cdot \mathcal{L}_N(\theta)$, with

$$\mathcal{L}_C(\theta_C) \equiv \prod_{r=1}^{\check{r}} \prod_{t=1}^{\check{t}-1} g_C\left(C_{r,t+1} \mid C_{r,t}; X_r, \theta_C\right),$$

the marginal likelihood of θ_C for the demand state dynamics; and

$$\mathcal{L}_N(\theta) \equiv \prod_{r=1}^{\check{r}} \prod_{t=1}^{\check{t}-1} p\left(N_{r,t+1} \mid N_{r,t}, C_{r,t}; X_r, \theta\right),$$

[9] Our estimation procedure can be extended to allow for observed (to the econometrician) time-varying covariates that are common across markets, such as business cycle indicators, provided that firms can use the model's primitives to forecast their evolution.

[10] These assumptions rule out persistent unobserved heterogeneity in the primitives across markets. Relaxing this and appropriately extending our NFXP procedure is straightforward in principle, but it does require us to provide a model-based solution to the "initial conditions problem" that $(N_{r,1}, C_{r,1}, X_r)$ is not independent of the persistent unobservables.

[11] We neither specify nor estimate the initial conditions' distribution, because we want to be agnostic about their relation to the dynamic model. We could instead assume that the initial conditions are drawn from the model's ergodic distribution. This would allow us to develop a more efficient estimator, at the price of robustness. Moreover, it would allow us to deal with the initial conditions problems mentioned in Footnote 10.

the conditional likelihood of θ for the evolution of the market structures.[12] Here, $g_C(\cdot\,|\,\cdot\,; X_r, \theta_C)$ is the density of $G_{C,r}$ and $p(n'|n, c; X_r, \theta) = \Pr(N_{r,t+1} = n'|N_{r,t} = n, C_{r,t} = c; X_r, \theta)$ is the equilibrium probability that market r with n firms and in demand state c has n' firms next period.

Note that $\mathcal{L}_C(\theta_C)$ can be computed directly from the demand data, without ever solving the model. To calculate $\mathcal{L}_N(\theta)$ we need to compute the equilibrium transition probabilities $p(\cdot\,|\,\cdot\,; X_r, \theta)$ for each distinct value of X_r in the sample. To this end, we first compute the equilibrium post-survival values $v_{S,r}$ corresponding to the primitives implied by X_r and θ. From these, we obtain cost-shock thresholds for entry and *sure* survival, defined by $\overline{w}_{E,r}(n, c) \equiv \log v_{S,r}(n, c) - \log(\kappa_r + \varphi_r)$ and $\overline{w}_{S,r}(n, c) \equiv \log v_{S,r}(n, c) - \log \kappa_r$.

For $n' > n$, $p(n'|n, c; X_r, \theta)$ can easily be calculated as the probability that $W_{r,t}$ falls into $[\overline{w}_{E,r}(n'+1, c), \overline{w}_{E,r}(n', c))$. For $n' \leq n$, the computations are complicated by equilibrium mixing of survival decisions. For example, the number of firms can remain unchanged either because survival is a dominant action or because firms choose to exit with positive probability but (randomly) they all survive. Therefore, the probability that $n' = n$ sums the probability that $W_{r,t}$ falls into $[\overline{w}_{E,r}(n+1, c), \overline{w}_{S,r}(n, c)]$ (so that survival is a dominant action) with the probability that it instead equals some $w \in (\overline{w}_{S,r}(n, c), \overline{w}_{S,r}(1, c))$ (so that incumbents mix exit and survival) and that all n firms survive when they mix with probability $a_S(n, c, w)$. Similar complications arise for $n' \in \{1, \ldots, n-1\}$, which can only occur if firms nontrivially mix exit and survival, and for $n' = 0$, which can occur if either the incumbents all choose certain exit or they are mixing nontrivially and all exit by chance. Accounting for the influence of mixed strategies on $p(n'|n, c; X_r, \theta)$ in these last three cases is tedious but straightforward.

The model's specification allows us to include the econometric error's variance in θ_W and estimate it. Some of our predecessors who solved analogous single-agent decision problems fixed this distribution's scale (e.g. Rust, 1987), so the reader might justifiably wonder whether or not our model is identified from data on only N and C. In Appendix D we provide a complete demonstration that it is indeed identified. A novel aspect of our identification argument is the use of firms' mixed strategies

[12] As in Ericson and Pakes (1995), firms begin to earn profits in the period *after* their entry decisions. Since $N_{r,t+1}$ is determined before the realization of $C_{r,t+1}$, its conditional distribution depends only on $C_{r,t}$.

to identify the variance of the shock to sunk and fixed costs. The interested reader can consult that appendix for details.

4.3 Estimation

We have created C++ and Matlab code for computing the full information maximum likelihood estimator of θ. As in Rust (1987), computation proceeds in three steps:

1. Estimate θ_C with $\tilde{\theta}_C \equiv \arg\max_{\theta_C} \mathcal{L}_C(\theta_C)$;
2. estimate (θ_P, θ_W) with $(\tilde{\theta}_P, \tilde{\theta}_W) \equiv \arg\max_{(\theta_P,\theta_W)} \mathcal{L}_N(\theta_P, \tilde{\theta}_C, \theta_W)$; and
3. estimate θ by maximizing the full likelihood function $\hat{\theta} \equiv \arg\max_\theta \mathcal{L}(\theta)$, using $\tilde{\theta} \equiv (\tilde{\theta}_P, \tilde{\theta}_C, \tilde{\theta}_W)$ as a starting value for the chosen optimization routine.

Note that the partial likelihood estimator $\tilde{\theta}$ computed in the first two steps is consistent, but not efficient. The third step's estimator $\hat{\theta}$ is asymptotically efficient. To compute estimated standard errors, we use the outer-product-of-the-gradient estimator of the (full) information matrix. In particular, we assume that $\check{r}$ is large and $\check{t}$ is small and use the average over markets of the outer products of the market-specific gradients, evaluated at $\hat{\theta}$.

The C++ code provides a full implementation of this three-step NFXP procedure for specifications with and without covariates. We use a standard non-linear, gradient-based optimizer to perform the optimization, and we compute all gradients analytically. The Matlab code provides a more user friendly implementation of the NFXP procedure with neither covariates nor analytical gradients that can be used as a sandbox for experimentation and teaching.

In Appendix E, we report the results of Monte Carlo experiments that estimate the model's parameters given data generated by the model itself. Using these same "data," we also estimate our model using Su and Judd's (2012) MPEC procedure. Those experiments lead us to four conclusions. First, the NFXP estimator can distinguish between economically meaningful hypotheses with observations from as few as 250 markets over 10 years. Second, asymptotic distribution theory gives a good guide to standard confidence intervals' coverage probabilities with such small samples. Third, the NFXP estimator takes very little time to compute. The average estimation time across 1,000 experiments was only between 20 and 30 seconds depending on the specification. Thus, our estimator's computational burden is not

substantially greater than that of static models of long-run industry structure, a point emphasized by Pakes, Ostrovsky, and Berry (2007) regarding their two-step estimators. Fourth and finally, the MPEC procedure always calculated the same estimates as our NFXP estimator but was 25 times slower. Thus, it appears that our estimator passes the initial quality assurance test of being accurate and relatively easy to compute when applied to simulated data.

5 Motion Picture Theaters

In this section we apply the model and its estimator to an empirical analysis of the Motion Picture Theaters industry. Davis (2006) showed that theater locations substantially influence consumers' decisions about whether and where to attend film screenings. Indeed, one's probability of attending a given theater declines considerably when the travel distance moves from between zero and five miles to between five and ten miles.[13] Davis (2002) found that the concomitant low cross-price elasticities from such spatial preferences impact firms' pricing behavior. Using observations from a New Haven area theater that experimented with a temporary price cut, Davis (2002) established that rivals five to seven miles away responded with lower prices but those ten to twelve miles away did not. This spatial differentiation of Motion Picture Theaters has two implications for our empirical analysis. First, theaters that are sufficiently far apart plausibly operate in distinct markets. Consequently, in Section 5.1, we define such markets using readily available geographic data. Second, variation across markets in the within-market spatial structure of demand is likely to come with variation in their profits and toughness of competition. To capture this, we include in X_r a measure of the diversity of consumers' geographic preferences.

The empirical analysis proceeds in four steps. First, we describe the data we use for the estimation. Second, we conduct an empirical examination of whether the model's abstraction from persistent heterogeneity across producers makes it unsuitable for application to these data. We conclude that it does not. Third, we present estimates of the model's parameters and discuss their implications for the toughness of competition between theater owners for screening rights. Fourth, we present several simulations of the model that highlight the long-lasting impact of

[13] See the logit model estimates reported in Table 5 of Davis (2006).

initial conditions, illustrate the importance of sunk costs in determining the length of "short-run" transitions to the long run, and quantify how lenient antitrust policy can offset permanent negative demand shocks.

5.1 The Data

Our analysis equates a market with a Micropolitan Statistical Area (μSA) as defined by the Office of Management and Budget. Each one is based around an urban core of at least 10,000 but less than 50,000 inhabitants.[14] We dropped the μSA "The Villages, FL," because its population growth far exceeds that of any other μSA. The remaining 573 μSAs account for about ten percent of the United States population. We measured the diversity of the geographic preferences of each μSA's residents using the locations and populations of its constituent year 2000 Census tracts. For this, we supposed that each census tract is a circle with an area equal to that of the tract itself, that population is uniformly distributed over the area enclosed by the circle, that all travel within a tract must pass through its center, and that travel between tracts follows straightline roads that connect their centers.[15] We then measured *geographic preference diversity* with the average distance between two randomly-chosen residents of the μSA. Likewise, we can measure the average distance between two randomly chosen individuals from two distinct μSAs. By construction, μSAs are geographically isolated from larger Metropolitan Statistical Areas, so we measure a given μSA's *geographic market isolation* as the shortest such distance to another μSA.

For the 573 μSAs, Table 1 displays the five standard quantiles for population, median household income, geographic preference diversity, and geographic market isolation. Population varies by about a factor of four from the 10th to the 90th percentiles. For the United States as a whole, median household income equalled $\$41,990$ in 2000. This is considerably higher than the median value across the μSAs, $\$33,520$. More than 80 percent of the μSAs have median household incomes within

[14]We use the release of the "Annual Estimates of the Population of Metropolitan and Micropolitan Statistical Areas from April 1, 2000 to July 1, 2009" from the US Census Bureau, which includes information on 574 μSAs.

[15]For these calculations, we used the tract population and geographic location information from the National Census Tracts Gazetteer File for the 2000 Decennial Census. See http://www.census.gov/geo/maps-data/data/gazetteer2000.html for its documentation. We used each tract's latitude and longitude in this file as its center.

Table 1: Summary Statistics for μSAs

	Quantile				
	10	25	50	75	90
Population	23.51	32.57	42.67	62.32	87.71
Median Household Income	27.40	30.40	33.52	38.04	42.40
Geographic Preference Diversity	9.24	11.17	13.37	16.81	21.23
Geographic Market Isolation	23.94	28.77	37.61	51.83	72.70

Note: All variables are measured as of 2000 for the 573 μSAs in our sample. Population is expressed in thousands of people, Median household income is expressed in thousands of dollars per year, and the remaining variables are expressed in miles. Please see the text for further details.

$\$10,000$ of this central tendency. The median geographic preference diversity is 13.37 miles. Perhaps unsurprisingly, this variable is highly skewed to the right. The 10th percentile is 9.24 miles, while the 90th percentile is 21.23 miles. Given the evidence from Davis (2002, 2006) regarding urban consumers' transportation costs for attending movies, it is plausible that the least geographically diverse μSAs in our sample might form a single geographic market. On the other hand, those with the most geographic preference diversity might actually be collections of two or more "markets" with relatively low elasticities of substitution across them. In any case, the measures of geographic isolation indicate that the elasticities of substitution across locations within a μSA should be much larger than those across μSAs. Its median value across μSAs equals 37.61 miles. Indeed, there are only eight μSAs where this distance is less than twenty miles. We conclude that the μSAs are isolated enough from each other so that consumer substitution between them can be ignored.

The Motion Picture Theaters industry (NAICS code 512131) consists of all establishments that primarily display first-run and second-run motion pictures, except for drive-in theaters. Our estimation uses annual counts of the number of *theaters* (not firms) in each μSA from the County Business Patterns (CBP), beginning in 2000 and ending in 2009. The top panel of Table 2 reports the frequencies of the number of theaters across all of the μSA-year observations. No theaters serve the market in about twenty percent of the observations, a single theater serves about half of them, and about thirty percent of our observations have more than one theater. The maximum number of theaters observed is nine, but only 4.9 percent of the observations have four or more. Each row of Table 2's bottom

Table 2: Frequencies and Transition Rates from the County Business Patterns

% of μSA-Year Observations by Number of Movie Theaters				
0	1	2	3	≥ 4
19.3	50.6	19.4	5.8	4.9

	% of Transitions Given N_{t-1}				
$\downarrow N_{t-1}/N_t \rightarrow$	0	1	2	3	≥ 4
0	88.9	9.3	1.6	0.2	0.0
1	4.3	89.2	5.9	0.4	0.1
2	0.7	13.1	77.9	7.2	1.1
3	0.0	3.7	22.1	62.2	11.9
≥ 4	0.4	0.8	2.4	13.0	83.5

Note: The top panel gives the distribution of the number of movie theaters per μSA from 2000 to 2009 from the County Business Patterns for the 573 μSAs in our sample. The bottom panel displays the conditional probability of transitioning from N_{t-1} movie theaters in a μSA at time $t-1$ (row) to N_t theaters at time t (column).

panel reports the observed frequencies of the number of theaters conditional on its previous year's value. Regardless of the initial number of theaters, the most common outcome is for it to remain unchanged. Nevertheless, the number of theaters changes in about 15 percent of the observed annual transitions.

In addition to this panel of producer counts, our estimation requires repeated measurements of the demand indicator C and cross-sectional measurements of time-invariant market characteristics X. The time-invariant market characteristics we employ are median income, dummy variables indicating membership in the nine U.S. Census Divisions, and an indicator for geographic preference diversity above its median value. For C, we use annual population for each μSA as published by the Census Bureau. For our sample from 2000 to 2009, the mean and standard deviation of the annual population growth rate equal 0.34 percent and 1.11 percent. The Census Bureau estimates these for non-census years using the most recent decennial census as a baseline, so they have very large adjustments between 2009 and 2010. The mean and standard deviation of population growth between these two years equals 1.5 percent and 3.1 percent. Since the measured changes between 2009 and 2010 disproportionately arise from differences in measurement methodology rather than true population changes, we end our estimation sample in 2009.

5.2 Heterogeneity and Industry Dynamics

Since the pioneering empirical work of Dunne, Roberts, and Samuelson (1988), a rich literature has arisen that documents high producer turnover within narrowly-defined industries and measures its role in aggregate productivity growth. Our analysis focuses on a very different (and somewhat older) question: How does changing market size change the number of producers and the profits they earn? This question does not obviously require consideration of producer heterogeneity, and indeed both Bresnahan and Reiss (1990) and Campbell and Hopenhayn (2005) examine it in the context of models with homogeneous firms. However, just because it is convenient to treat all producers symmetrically does not mean it is appropriate to do so. Persistent differences between firms might substantially impact their entry and continuation decisions and thereby bring the results of applying our model into question. At the same time, the simple observation that producers differ on some observable dimensions does not invalidate an approach like ours that treats them symmetrically. Our model only requires that producers' *expected* profits are identical, because expected profits govern entry and exit decisions. This leaves a great deal of room for incorporating observable transitory heterogeneity across producers. For example, the theaters in our data might always have very different sizes (measured with e.g. sales) if the (randomly-chosen) theater with the highest-quality film (the "blockbuster") attracts all quality-sensitive consumers and the remaining quality-insensitive consumers split themselves equally across theaters. As long as blockbusters do not systematically go to one particular theater, such within-period heterogeneity is irrelevant for producers' entry and continuation decisions.

Whether abstracting from producer heterogeneity within a local market is a helpful simplification or a fatal error is ultimately an empirical question. The common observation of high producer turnover at the national level might seem to settle that question, since our model trivially predicts that entry and exit never occur simultaneously. However, only turnover *within* μSAs is relevant for the question of whether the model can be usefully applied to our sample of local markets.[16] We

[16] We estimate that the entry and exit rates of theaters between the 2002 and 2007 Economic Censuses were *at least* 9.2 percent and 11.3 percent. Therefore, a more thorough investigation would almost certainly find that the Motion Picture Theaters industry is no exception to the empirical rule of high producer turnover at the national level. We calculated these lower bounds using reports of the number of establishments in both years' Economic Censuses, 4979 and 4879, and the number of establishments in the 2002 Census that did not operate for all of

seek to settle this question for the theaters in our data set by examining directly how producer heterogeneity influences the evolution of the number of producers. Our homogeneity assumption requires measures of producer heterogeneity to have insubstantial effects in a forecasting model of the number of firms that accounts for time-invariant market characteristics, the realization of demand (population in our case), and the lagged values of the number of producers. To investigate this, we have estimated Poisson regression models with the number of firms in year t as the dependent variable. The independent variables include the logarithm of population in year $t-1$ and its square, dummy variables for the number of theaters in year $t-1$, calendar-year dummies, census-division dummies, linear and squared terms in median income measured in 2000, and the logarithms of geographic preference diversity and geographic isolation.

In one of the forecasting models, we also include a particular measure of heterogeneity. If we denote the size of the jth active producer in market r in year t with $Q_{j,r,t}$, then this can be written as

$$H_{r,t} = \frac{1}{N_{r,t}} \sum_{j=1}^{N_{r,t}} Q_{j,r,t}^2 \Big/ \left(\frac{1}{N_{r,t}} \sum_{j=1}^{N_{r,t}} Q_{j,r,t} \right)^2 .$$

This is the *uncentered* second moment of producer size divided by its first moment squared. It can be easily shown that $H_{r,t}$ equals the Herfindahl-Hirschman Index (HHI) multiplied by $N_{r,t}$. Since the HHI obtains its minimum value of $1/N_{r,t}$ when producers have equal sizes, $H_{r,t}$ equals the multiplicative "correction" one must apply to $1/N_{r,t}$ to get the true HHI. Its minimum value (obtained with identical producers) is one. By including this measure in our forecasting model, we give the data the opportunity to indicate whether or not producer heterogeneity substantially impacts the evolution of $N_{r,t}$. If our structural model were literally true, then we could exclude $H_{r,t-1}$ from the forecasting model without cost. On the other hand, we expect $H_{r,t-1}$ to substantially improve forecasts of $N_{r,t}$ generated from other models

2002. We counted these as exits. Since exits could also occur in any of the inter-census years, this is a lower bound. The 11.3 percent exit rate equals $100 \times 2 \times 555/(4979 + 4879)$ percent. To get the entry rate, we subtracted the number of establishments that were active for all of 2002 from the number of establishments in the 2007 Economic Census. The result, 455, is a lower bound on the number of entrants between the two censuses. The 9.2 percent entry rate equals $100 \times 2 \times 455/(4979 + 4879)$ percent. The data underlying these calculations can be found at http://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ECN_2007_US_51SSSZ1&prodType=table.

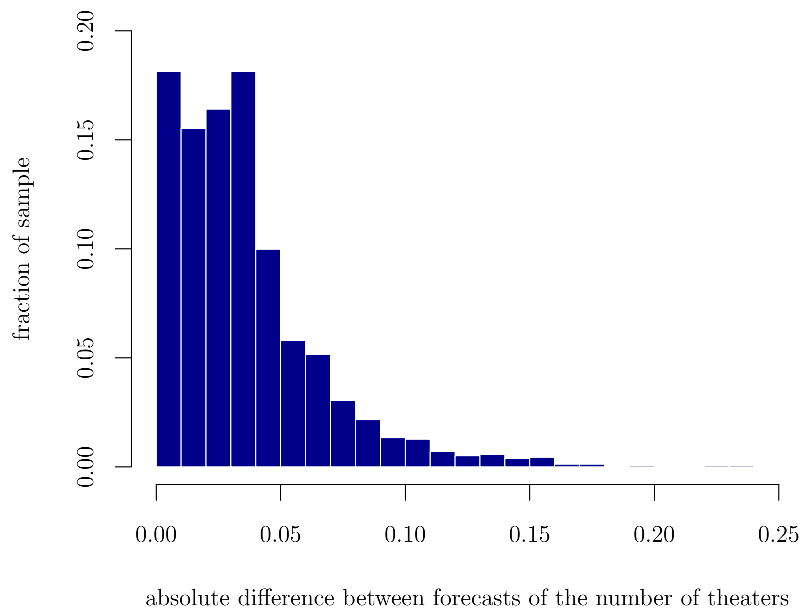


Figure 2: Histogram of the Absolute Differences between Model Forecasts

Note: The figure plots the fraction of market-year observations for which the absolute difference of the two models' predicted (expected) values of the number of theaters falls into bins with width equal to 0.01 theater. Both models are Poisson regressions that include calendar-year dummies, census-division dummies, linear and squared terms in the previous year's population, linear and squared terms in median income measured in 2000, the logarithms of geographic preference diversity and geographic isolation (as defined in the text), and indicators for the number of theaters serving the market in the previous year. One model also includes the logarithm of $H_{r,t-1}$, heterogeneity's contribution to the HHI in the previous year. Both models are estimated with the sample of $1,572$ observations from 2001 through 2009 with $N_{r,t-1} \geq 2$.

of industry dynamics. For example, in models of increasing market dominance, such as that in Cabral (2002), heterogeneity increases over time on average, which increases the incentive for lagging producers to exit. Similarly, the introduction of a Big-Box retailer to a market increases producer heterogeneity and induces smaller competitors to exit, so a high value of $H_{r,t-1}$ should predict a reduction in $N_{r,t}$.[17]

In our data, we measure each theater's size with the midpoint of the employment size category to which it belongs in the year's mid-March pay period. The County Business Patterns always reports this for each theater without identifying the theater itself. We include the logarithm of $H_{r,t-1}$ in our forecasting model, so its estimated coefficient can be interpreted as an elasticity. So that we do not bias the forecasting model towards finding that heterogeneity is unimportant, we include

[17]See Haltiwanger, Jarmin, and Krizan (2010) for evidence of this effect of Big-Box retailers on their smaller competitors.

only observations for which $N_{r,t-1} \geq 2$. Over our ten year sample, there are 1,572 such observations. Over them, the mean value of $H_{r,t-1}$ is 1.31, and its standard deviation is 0.30. The estimated coefficient on the logarithm of $H_{r,t-1}$ equals −0.096, and its standard error is 0.036. Therefore, $H_{r,t-1}$ has a statistically-significant and negative effect on $N_{r,t}$. However, this effect is economically small. We can see this in two ways. First, we note that the logarithm of $H_{r,t-1}$ has a standard deviation of 0.21 in our sample. Therefore, a one-standard deviation increase in this measure of heterogeneity decreases the predicted number of firms by only about 2 percent. Second, we estimated the same forecasting model but excluding the logarithm of $H_{r,t-1}$ and calculated forecasts using it and our complete model. Figure 2 gives the histogram of the absolute differences between the two models' forecasts. The median absolute difference between the two models' predictions is 0.030 theaters, and the mean absolute difference is 0.036 theaters. In comparison, the mean forecasted values for both models equal 2.46 theaters. Since there is no economically significant effect of producer heterogeneity on the evolution of $N_{r,t}$, we conclude that abstracting from producer heterogeneity is a helpful simplification rather than a fatal error.

5.3 Estimates

The NFXP procedure requires us to specify the demand process G_C, the distribution G_W of the cost shocks, and the per period surplus function π as functions of a finite vector of parameters. For the demand process, we follow Tauchen (1986). We restrict $C_{r,t}$ to a grid of 200 points equally spaced on a logarithmic scale with distance d: $c_{[1]}$, $c_{[2]} = c_{[1]} \exp(d)$,. . . , $c_{[200]} = c_{[1]} \exp(199d)$. So that the growth of C_t is approximately normally distributed with mean μ and variance σ^2, we specify the probability of transitioning to $c_{[i]}$ from $c_{[j]}$ for any $i = 2, ..., 199$ and $j = 1, ..., 200$ with

$$\Pr[C' = c_{[i]} | C = c_{[j]}] = \Phi\left(\frac{\log c_{[i]} + \frac{d}{2} - \log c_{[j]} - \mu}{\sigma}\right) - \Phi\left(\frac{\log c_{[i]} - \frac{d}{2} - \log c_{[j]} - \mu}{\sigma}\right).$$

The probabilities of transitioning to the grid's end points equal

$$\Pr[C' = c_{[1]} | C = c_{[j]}] = \Phi\left(\frac{\log c_{[1]} + \frac{d}{2} - \log c_{[j]} - \mu}{\sigma}\right)$$

and

$$\Pr[C' = c_{[200]} | C = c_{[j]}] = 1 - \Phi\left(\frac{\log c_{[200]} - \frac{d}{2} - \log c_{[j]} - \mu}{\sigma}\right),$$

respectively. The lower bound of the demand grid equals the minimum population observed in our data, 11,011, divided by 1.25. Analogously, the upper bound equals the maximum population, 197,912, multiplied by 1.25. For estimation, we replace each observation of μSA population with the closest grid point.

The maximum number of movie theaters sustainable, $\check{n}$, is fixed at the maximum number of theaters observed in the data, nine. We specify the distribution G_W of the cost shocks to be normal with standard deviation ω and mean $-\omega^2/2$, so that $\exp(W_{r,t})$ has a log-normal distribution with unit mean and scale parameter ω. We normalize κ to one and fix the discount factor ρ at $\frac{1}{1.05}$. The specification for the producers' surplus function is

$$\pi_r(n, c) = \exp(\beta' X_r^{(1)}) \frac{c}{n} k(n; X_r^{(2)}).$$

For this, we split the market characteristics in X_r into two sub-vectors, $X_r \equiv (X_r^{(1)}, X_r^{(2)})$. Those in $X_r^{(1)}$ affect the surplus log-linearly and include the logarithm of median income (expressed as a deviation from the logarithm of the average median income across our 573 μSAs) and dummies for all Census Divisions excluding New England. The remaining characteristics in $X_r^{(2)}$ interact with k and thereby affect the toughness of competition in a general way.[18] We both estimate the model without heterogeneity in k across markets (trivial $X_r^{(2)}$) and with k depending on whether a market's geographic preference diversity is above or below its median value, 13.4 miles ($X_r^{(2)}$ equal to an indicator for diversity exceeding 13.4 miles). We set $k(4) = k(5) = \cdots = k(9)$ to accommodate the paucity of observations with four or more theaters.

Table 3 reports the estimated parameters for two specifications, one that ignores geographic preference diversity and another that takes it into account.[19] Each

[18]The monotonicity assumption in Section 2 requires the flow surplus to weakly decrease with the number of firms. This assumption is satisfied at the maximum-likelihood estimates.

[19]The estimation results reported in this section differ somewhat from those reported in a version of the paper that we circulated in January 2014, because we have corrected an inconsistency between the model's parameterization in the paper and that in the code. The results have not changed qualitatively.

specification's estimation requires about 2 hours using 2 Xeon E5472 CPUs (each of which has 4 cores) on a single machine. Intel released this chip at the end of 2007; so it can hardly be considered state of the art. Since the calculation of the markets' equilibria easily parallelize, we expect substantial speed gains from using more recently-developed hardware.

In the first specification, the full-information maximum likelihood estimates of the demand process's drift and innovation standard deviation, 0.34 and 1.21 percent, are very close to the unconditional sample mean and standard deviation of population growth, 0.34 and 1.11 percent. The coefficients in β are jointly and (with the exceptions of those multiplying two division dummies) individually significant. The mean sunk cost of entry, φ, is over fifty times the mean fixed cost of continuation. However, one should *not* interpret this as a measure of the typical sunk cost paid because entry only occurs when the realization of the cost shock is low. To calculate more informative measures of fixed and sunk costs, we simulated the estimated model for the New England Census Division. In the simulation, the average fixed cost of continuation and sunk cost of entry *paid* were 0.47 and 0.92. The estimates of all these parameters from the specification that accounts for geographic preference diversity are similar to these baseline estimates.

Table 4 reports transformations of the estimates from Table 3 with a more straightforward economic interpretation. The first row reports $1/(k(1) \times 10^3)$, the population in thousands that sets a monopolist's current profit (the surplus earned minus the fixed continuation cost incurred in a period) to zero in a New England market with average median income when the fixed continuation cost equals one. The baseline specification's estimate of this is 26,360 people. We expect that concentrating customers' locations increases a monopolist's profit by making it easier to simultaneously satisfy their geographic preferences. The estimates from the model that accounts for geographic preference diversity support this prior. It takes $29,420$ people to support a monopolist in a μSA with geographic preference diversity above the median and $25,980$ to support a monopolist in a μSA with preference diversity below the median. A Wald test indicates that this difference is significant at the five percent level.

The remaining rows of Table 4 report estimates of $k(n+1)/k(n)$, the share of the surplus per consumer left after the addition of a competitor. These indicate very tough competition. In the first specification, duopolists' producers' surplus per

Table 3: Parameter Estimates

		Geographic Preference Diversity	
	All μSAs	Diversity > 13.4 miles	Diversity ≤ 13.4 miles
$k(1) \times 10^5$	0.38	0.34	0.38
	(0.05)	(0.05)	(0.05)
$k(2) \times 10^5$	0.21	0.20	0.18
	(0.03)	(0.03)	(0.03)
$k(3) \times 10^5$	0.17	0.17	0.14
	(0.03)	(0.03)	(0.03)
$k(4) \times 10^5$	0.13	0.14	0.10
	(0.02)	(0.02)	(0.02)
φ	51.23	49.06	
	(9.82)	(9.75)	
Median Income	1.75	1.74	
	(0.07)	(0.08)	
Mid Atlantic	0.88	0.85	
	(0.17)	(0.17)	
East North Central	-0.64	-0.60	
	(0.15)	(0.15)	
West North Central	-0.49	-0.45	
	(0.14)	(0.14)	
South Atlantic	0.05	0.09	
	(0.16)	(0.16)	
East South Central	-0.73	-0.71	
	(0.15)	(0.15)	
West South	-0.52	-0.49	
	(0.16)	(0.16)	
Mountain	-0.33	-0.29	
	(0.15)	(0.15)	
Pacific	-0.12	-0.07	
	(0.15)	(0.15)	
ω	1.75	1.74	
	(0.15)	(0.15)	
$\mu \times 10^2$	0.34	0.34	
	(0.00)	(0.00)	
$\sigma \times 10^2$	1.21	1.21	
	(0.00)	(0.00)	
$-\log \mathcal{L}$	9199.13	9192.24	
Number of Markets	573	287	286

Note: Standard errors are reported in parentheses. The data include 573 μSAs from 2000 to 2009, and $\check{n}$ equals nine, which is the maximum of the number of active firms observed in the data. The values of $k(5), \ldots, k(9)$ identically equal $k(4)$. The first column reports estimates of a specification in which $k(n)$ does not vary between markets. The second and third columns report estimates of a specification in which $k(n)$ differs between markets with geographic preference diversity below and above its median value of 13.4 miles ($X_r^{(2)}$). Both specifications include the logarithm of median income (in deviation from the logarithm of the average median income across μSAs) and Census Division dummies (excluding New England) as market characteristics in $X_r^{(1)}$. Please see the text for further details.

Table 4: Estimates of the Toughness of Competition

		Geographic Preference Diversity	
	All μSAs	Diversity > 13.4 miles	Diversity $\leq$ 13.4 miles
$1/\left(k(1) \times 10^3\right)$	26.35	29.42	25.98
	(3.50)	(4.00)	(3.64)
$k(2)/k(1)$	0.54	0.60	0.48
	(0.14)	(0.14)	(0.20)
$k(3)/k(2)$	0.82	0.84	0.78
	(0.06)	(0.06)	(0.10)
$k(4)/k(3)$	0.77	0.79	0.66
	(0.08)	(0.08)	(0.21)
Number of Markets	573	287	286

Note: This table is based on the model's estimates as reported in Table 3. Standard errors are reported in parentheses. The ratio $1/(k(1) \times 10^3)$ can be interpreted as the population (in thousands of people) that sets a monopolist's current profit to zero in a New England market with average median income when the fixed cost equals one. The ratio $k(n+1)/k(n)$ is an indicator of the toughness of competition. Please see the text for further details.

consumer equals 54 percent of a monopolist's. A third competitor decreases this surplus to 82 percent of the duopolists'. In markets with four or more competitors, the producers' surplus per consumer further decreases to 77 percent of that in a market with three firms. Altogether, adding three or more theaters to a market with a single incumbent brings the surplus per customer down to 34 percent of its monopoly value.

The theoretical literature on spatial differentiation overwhelmingly points to heterogeneity of consumers' locations as a source of market power. This leads us to expect producers' surplus to fall less rapidly with additional competition in the high-diversity markets. The estimates from the second specification support this conjecture. The ratios of the producers' surplus $k(n+1)/k(n)$ are indeed higher in high diversity than in low diversity markets. A Wald test indicates that these differences are jointly statistically significant at the one percent level. This suggests that entering theaters can lessen the toughness of competition with their location choices.

These estimates of tough competition contrast with other evidence from this industry. Davis (2005) provides evidence on competition for customers from regressions of theaters' admissions prices against indicators of the presence of other theaters at various distances using data from large (relative to μSAs) U.S. cities in the 1990s. Based on both across-market and within-market-over-time variation, he concludes that

> ... the magnitude of the price-reducing effect of local competition appears to be economically modest.

Prior research on the vertical relationships between theater owners and their upstream suppliers, film distributors, has emphasized formal and informal arrangements to manage the *popcorn conflict* over the final ticket price: Popcorn and other concession sales are complements with theater attendance, and theater owners keep all surplus from concession sales while splitting surplus from ticket sales with the film distributor. Therefore, theater owners prefer lower ticket prices than do distributors. The motion picture industry operates under a relatively unique legal regime, under which the producers of films are legally barred from directly influencing box-office pricing or vertically integrating with motion picture theaters. Nevertheless, repeated interactions between distributors and theater owners might give distributors indirect and extralegal control over box-office prices.[20] Supporting the view that film distributors constrain theaters' pricing choices, Davis (2006) finds that

> ... the average theater owner would prefer to actually lower admissions prices, if she could attract the same set of films.

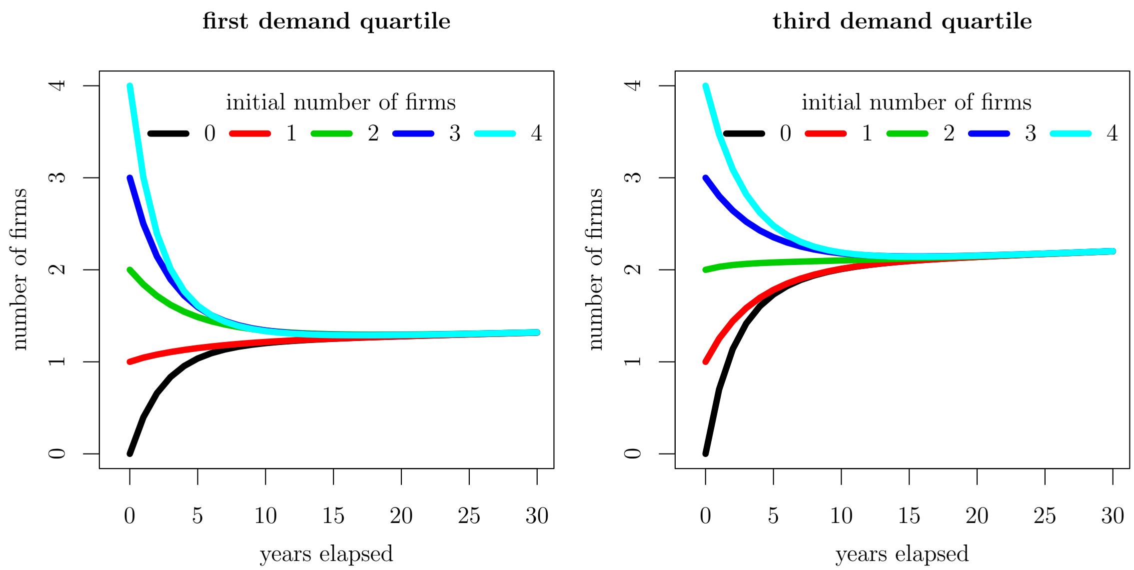
Accordingly, we find it implausible that our estimates reflect fierce competition for customers. Instead, we believe that adding theaters to a market increases competition for film exhibition rights. Indeed, Gil and LaFontaine (2012) find some evidence that owners of Spanish theaters with higher local market shares get better deals from film distributors.

5.4 Dynamic Implications

The estimates in Table 3 show that movie theaters face substantial sunk entry costs and uncertainty about future profits. Consequently, nontrivial dynamic considerations govern their entry and exit. This section explores the implied dynamics of local movie theater markets by simulating the estimated model and counterfactual versions of it. This is straightforward because its equilibrium is unique and easy to calculate.

[20] Orbach and Einav (2007) review the legal environment in which theater owners negotiate with film distributors and set admissions prices.

Figure 3: Initial Conditions and the Evolution of the Expected Number of Firms

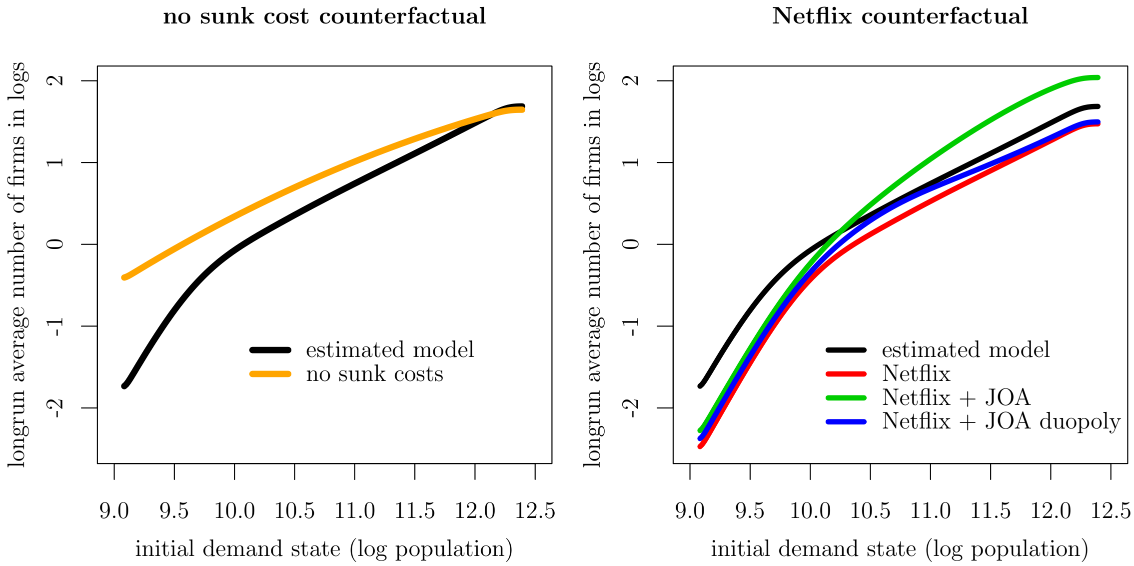


Note: These panels report the evolution of the expected number of active firms implied by the model. The model estimates are taken from the second specification reported in Table 3 and correspond to a high diversity market in New England with median income equal to \$34,417, the average median income across μSAs. The left panel sets the initial demand state to 32,558, the first quartile of the 2009 population distribution across μSAs in our data. The right panel sets it to 64,119, the third quartile. The initial values for the number of active firms are marked as such in both graphs.

We first consider the dynamics of a single local market. We find that a market's initial number of active firms and initial demand state (population) have long-lasting effects on its expected number of active firms. Figure 3 illustrates this for a New England market with average median income, high geographic preference diversity, and parameters equal to the estimates from the second specification in Table 3. Its left panel sets initial demand in this market equal to the first quartile of the 2009 population distribution across μSAs $(32,570)$ and plots the evolution of the expected number of firms from various initial numbers of firms. Its right panel sets initial demand to the third quartile of that same distribution $(65,119)$. In both panels, the dependence on the initial number of firms vanishes only after 10 to 15 years. Sunk costs are key to this gradual adjustment; if we were to set them to zero in our model, the expected number of firms would lose its relation to the initial number of firms right away.[21]

[21]Because incumbent firms can commit to serving the market in the next period before entry takes place, their survival decisions are nontrivially dynamic even in the absence of sunk costs.

Figure 4: Expected Number of Active Firms in the Long Run



Note: These panels show the log of the expected number of active firms after thirty years as a function of the log initial population. The black curves show this relationship for the estimated model. They are based on the second specification reported in Table 3 and correspond to market in New England with high geographic preference diversity and median income equal to $\$34,417$ (the average median income across μSAs). In the left panel, the orange curve corresponds to a counterfactual model without sunk costs. In this counterfactual, fixed costs are raised to $1 + \phi(1 - \rho)$. In the right panel, the red curve shows the relationship for the Netflix counterfactual, i.e. a 25 percent reduction in consumers' propensity to go to the movies. The green and blue curves correspond to the counterfactuals that add JOAs to the Netflix counterfactual for all markets and duopolies, respectively.

Figure 3 shows that the average number of firms serving the industry after 30 years depends on the market's initial demand state. If the market starts at the first quartile of the population distribution, it will have on average 1.32 active firms after 30 years (left panel); if it starts at the third population quartile, it will end up with 2.20 firms (right panel). By construction the demand process is stationary, so this dependence vanishes if given enough time. However, this process requires well over 1,000 years. (This is not surprising, since we designed G_C to approximate the short-run behavior of a random walk.) Therefore, the usual mathematically-convenient definition of the "long run" — the model's ergodic distribution — is not practically relevant for this application. Instead, we define the "long-run" with the distribution of the number of active firms 30 years from the present. Based on this definition and on the results in Figure 3, we conclude that the transition to the long run requires

Nevertheless, the equilibrium without sunk costs closely mimics repeated static Nash play of a one-shot entry game with no incumbency.

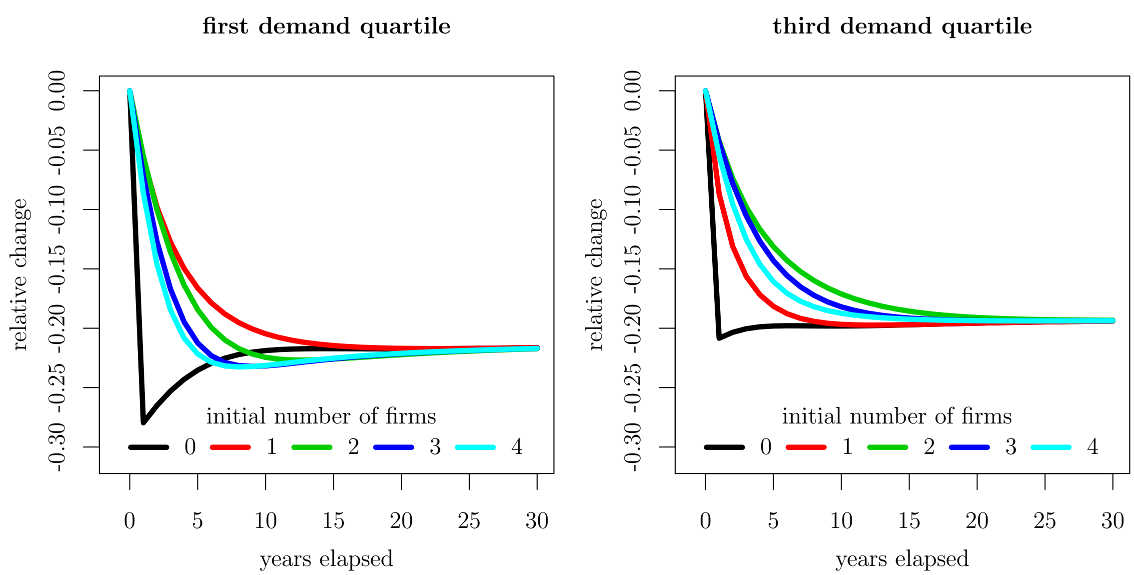
between 10 and 15 years.

Figure 4 further explores how the long-run average number of firms depends on initial demand. In its left panel, we plot the log of the market's initial population against the log of its expected number of active firms 30 years later. The black and orange curves show this function for our estimated model and for a counterfactual variant without sunk costs.[22] Their slopes give the percentage change in the number of active firms for a one percent change in initial population. In both cases, for large enough initial demand (and thus a high enough expected number of firms), it takes more than a 1 percent increase in initial demand for a 1 percent increase in the expected number of firms. This reflects our estimates of tough competition. At very high initial demand states, the expected number of firms is close to the maximum number of firms, even large increases in demand cannot entice further entry, and the slopes taper off to zero. At low enough initial demand states, the expected number of firms is close to zero and its elasticity with respect to demand is larger. Comparing both curves shows that sunk costs magnify this effect; this is consistent with Abbring and Campbell's (2010) finding that the corresponding option values shift the population thresholds at which firms exit down.

Next, we consider the market's dynamic responses to a permanent 25 percent fall in its consumers' propensity to patronize its theaters. Since this demand reduction could follow the screening of new movies by an internet streaming platform, we will refer to it as the "Netflix shock." Figure 5 plots the implied differences between the counterfactual outcomes and Figure 3's baseline outcomes as a share of those baseline outcomes for the same initial conditions as in Figure 3. Like the baseline outcomes themselves, the short-run responses to the Netflix shock depend strongly on the initial conditions, with only their dependence on initial demand persisting in the long run. With low initial demand, the market adjusts more quickly to the Netflix shock if it starts with a large number of firms. Similarly, with high initial demand, it adjusts more quickly if it starts with few firms. Eventually, the market sheds around 20 percent of its theaters. This long-run loss is larger if initial demand is low. Figure 4's right panel confirms this: The gap between the log long-run expected number of active firms in the baseline (black) and in the Netflix

[22]In the counterfactual, the sunk costs are annuitized and added to the fixed costs (see the note to Figure 4). Increasing the fixed costs further (so that the implied number of firms is more in line with that of the estimated model) shifts the orange line down without substantially changing its shape.

Figure 5: Relative Changes in the Expected Number of Firms after a Netflix Shock



Note: These panels report responses to the Netflix shock, a permanent 25 percent reduction in demand, for various initial conditions. Specifically, they plot the differences between the counterfactual expected numbers of firms following the Netflix shock and Figure 3's baseline expected numbers of firms as a share of those same baseline numbers. The model estimates are taken from the second specification reported in Table 3, and the simulations and correspond to a market in New England with high geographic preference diversity and median income equal to $34,417$ (the average median income across μSAs). The left panel sets the initial demand state to $32,558$, the first quartile of the 2009 population distribution across μSAs. The right panel sets it to $64,119$, the third quartile. The initial values for the number of active firms are marked as such in both graphs.

counterfactual (red) decreases with initial demand. One possible explanation for this is that the reduction of competition dampens the negative effects of the Netflix shock on the profits of surviving incumbents in, in particular, high demand markets, which are more likely to support two or more firms.

Our estimates contain no information on how consumers value variety, so we cannot evaluate the social optimality of the number of active theaters or its adjustment. However, our model does allow us to calculate the *positive* responses of the industry to a policy intervention.[23] To demonstrate this capability, we consider one possible policy response to the Netflix shock: an antitrust exemption that allows competing theaters to sign joint operating agreements (JOAs) that centralize the acquisition of screening rights. Such a policy would be reminiscent of the 1970 Newspaper Preservation Act, which allowed newspapers, which have long been in

[23]If external information on the benefits of variety was available, appending it to this empirical analysis would be straightforward.

decline, to centralize the choices of advertising rates. We consider two variants of this intervention. One allows agreements in all markets and the other restricts them to duopolies.[24] The latter would make sense if the benefits of variety are exhausted with two competitors.

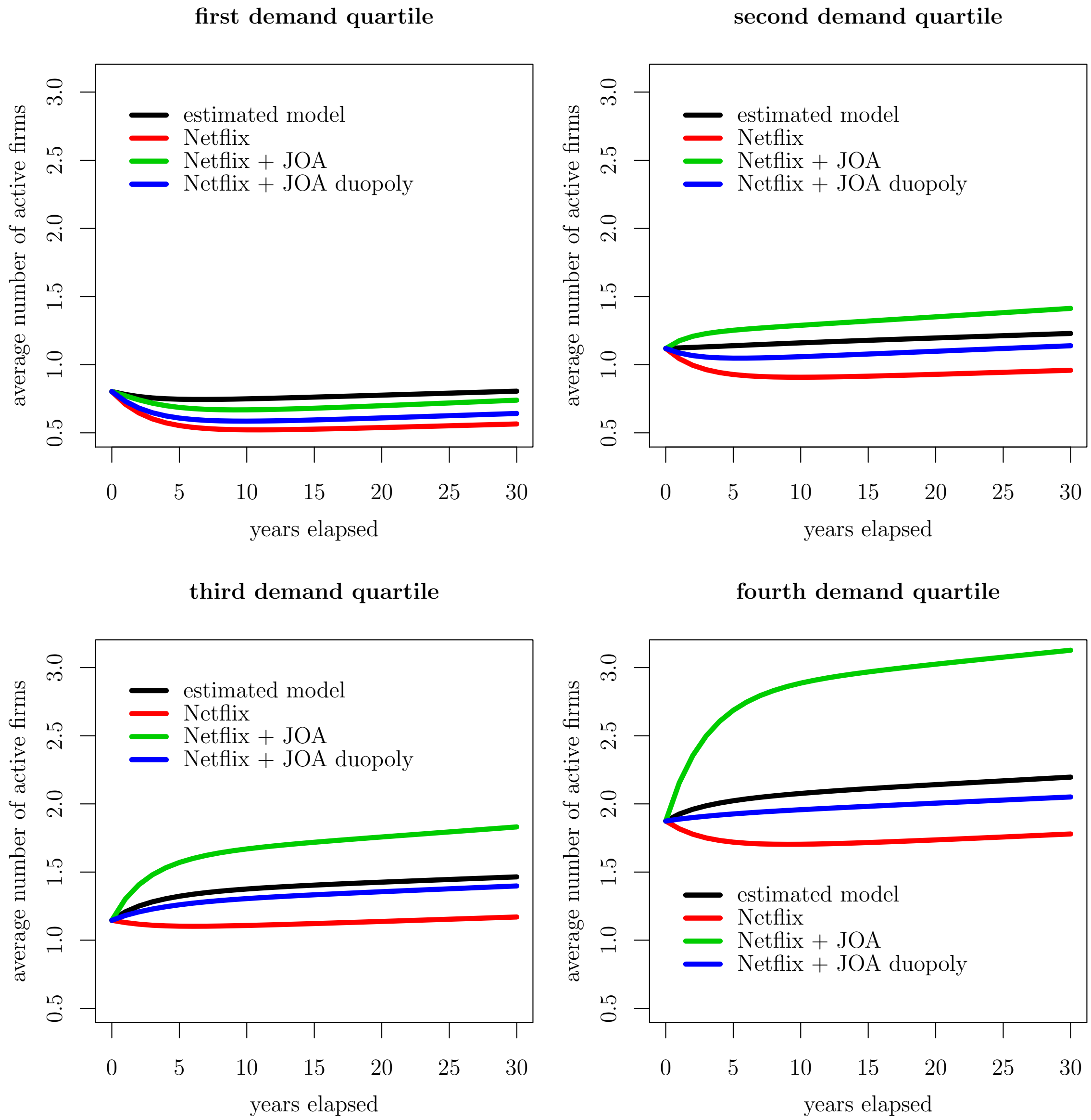
Figure 4's right panel plots the long run log expected number of firms if the Netflix shock is compensated with an all markets JOA (green) and if it is compensated with a duopoly JOA (blue). Neither JOA has much impact if initial demand is very low and markets are unlikely to support more than one theater. The baseline JOA policy is increasingly effective as initial demand, and thus the expected number of firms, increases. For large enough initial demand, it more than compensates the Netflix shock. On the other hand, the duopoly JOA only just compensates for the Netflix shock at intermediate levels of demand.

So far, we have focused on the dynamics of a single local market under various initial conditions. We finish this section by exploring the (counterfactual) evolution of our sample's actual markets following a Netflix shock with and without a JOA at the end of the sample period. To this end, we compute each market's equilibrium outcome paths starting from its actual 2009 demand state and market structure under various combinations of a Netflix shock and a JOA intervention in 2010. Figure 6 summarizes the results by partitioning the markets into four subsamples, one for each quartile of their 2009 population distribution, and plotting the average expected number of firms in each such subsample against the years elapsed since 2010. As our findings for the long run effects on a single market suggest, a duopoly JOA falls short of compensating the 25 percent Netflix shock in all years following it and all four subsamples. The baseline JOA more than offsets the Netflix shocks in all but the smallest markets, which benefit little from agreements among three or more theaters.

To further quantify the JOAs' effects, we have computed the sizes of the negative demand shocks that are exactly offset by each JOA in terms of the resulting average (across markets) long-run (30 year) expected number of firms. The baseline JOA exactly offsets a 36 percent permanent reduction in demand. This makes sense, given that it more than compensates for the 25 percent Netflix shock in at least

[24] We operationalize the JOAs for all markets by setting $k(n)$ to the estimated value of the monopolist's surplus $k(1)$ for all $n \geq 2$. We operationalize the JOAs for duopoly markets by setting $k(2)$ equal to the estimated value of $k(1)$ while keeping $k(n)$ unchanged and equal to their estimated values for all $n \geq 3$.

Figure 6: Expected Number of Firms following a Netflix Shock (Averaged over Markets by Demand Quartile)



Note: These panels report the evolution of the expected number of active firms implied by the estimated model and three counterfactual versions of it from 2010, averaged over the sampled markets by 2009 demand quartile. All three counterfactual models involve a Netflix shock, a permanent 25 percent reduction in demand, in 2010. The second and third counterfactual models add, respectively, an all markets and a duopoly JOA from 2010. The model estimates are taken from the second specification reported in Table 3.

75 percent of the markets. In contrast, a duopoly JOA only offsets a 17 percent demand reduction.

6 Conclusion

We have demonstrated uniqueness of our model's symmetric Markov-perfect equilibrium, provided an algorithm for its fast calculation, provided a nested fixed-point algorithm for its maximum-likelihood estimation, and applied these tools to quantify the toughness of competition between Motion Picture Theaters and the nature of short-run transitions following demand shocks in U.S. μSAs. This relatively complete development and application of a dynamic oligopoly model validates our title's assertion that our model's Markov-perfect industry dynamics are "very simple."

The model's key simplifying assumption is that firms have identical expected profits period-by-period when making their entry and continuation decisions. The data indicate that this is a helpful simplification rather than a fatal error for our application, but it would be very surprising if this conclusion characterized *all* industries. Accordingly, Abbring, Campbell, and Yang (2015) extended this paper's theoretical and computational analysis to a duopoly model ($\check{n} = 2$) with persistent and publicly-observed firm-specific shocks to profitability. These satisfy Ericson and Pakes's (1995) assumption that idiosyncratic shocks can only improve a firm's type. They use a generalization of this paper's algorithm to demonstrate equilibrium existence, uniqueness, and ease of computation. This facilitates a more thorough characterization of creative destruction in local oligopolies (like local markets of the Motion Picture Theaters industry) by the entry of global competitors (like Netflix and Hulu) that improve their operations over time. This would substantially generalize perfectly competitive models of technology-driven creative destruction, like that in Campbell (1998). Although this approach cannot mimic industries in which leading-edge entrants interact strategically with incumbents operating vintage technologies, the practical relevance of one large oligopolistic industry competing with a fringe of smaller oligopolistic industries seems high enough to us to merit further exploration.

References

Abbring, J. H. and J. R. Campbell (2010): "Last-In First-Out Oligopoly Dynamics," *Econometrica*, 78, 1491–1527. 3, 10, 12, 37

——— (2015): "Last-In First-Out Oligopoly Dynamics: Corrigendum," *Econometrica web site for supplementary material*, Forthcoming. 3

Abbring, J. H., J. R. Campbell, and N. Yang (2015): "Entry, Exit, and Technological Progress in Markov-Perfect Duopoly," Tech. rep., CentER, Tilburg University. 4, 6, 41

Bajari, P. L., C. L. Benkard, and J. Levin (2007): "Estimating Dynamic Models of Imperfect Competition," *Econometrica*, 75, 1331–1370. 4, 5, 64

Bresnahan, T. F. and P. C. Reiss (1990): "Entry in Monopoly Markets," *Review of Economic Studies*, 57, 531–553. 3, 26, 64

——— (1991a): "Empirical models of discrete games," *Journal of Econometrics*, 48, 57–81. 10

——— (1991b): "Entry and Competition in Concentrated Markets," *Journal of Political Economy*, 99, 977–1009. 3, 64

——— (1993): "Measuring the Importance of Sunk Costs," *Annales d'Economie et de Statistique*, 31, 181–217. 3

Cabral, L. M. B. (2002): "Increasing Dominance with No Efficiency Effect," *Journal of Economic Theory*, 102, 471–479. 28

——— (2004): "Simultaneous Entry and Welfare," *European Economic Review*, 48, 943–957. 10

Cabral, L. M. B. and M. H. Riordan (1994): "The Learning Curve, Market Dominance, and Predatory Pricing," *Econometrica*, 62, 1115–1140. 6

Campbell, J. R. (1998): "Entry, Exit, Embodied Technology, and Business Cycles," *Review of Economic Dynamics*, 1, 371–408. 41

Campbell, J. R. and H. A. Hopenhayn (2005): "Market Size Matters," *Journal of Industrial Economics*, 53, 1–25. 26

Davis, P. (2002): "Estimating Multi-Way Error Components Models with Unbalanced Data Structures," *Journal of Econometrics*, 106, 67–95. 22, 24

——— (2005): "The Effect of Local Competition on Admission Prices in the U.S. Motion Picture Exhibition Market," *The Journal of Law and Economics*, 48, 677–707. 2, 33

——— (2006): "Spatial Competition in Retail Markets: Movie Theaters," *RAND Journal of Economics*, 37, 964–982. 22, 24, 34

Doraszelski, U. and A. Pakes (2007): "A Framework for Applied Dynamic Analysis in IO," in *Handbook of Industrial Organization*, ed. by M. Armstrong and R. H. Porter, Amsterdam: Elsevier Science, vol. 3, chap. 4. 4

Doraszelski, U. and M. Satterthwaite (2010): "Computable Markov-Perfect Industry Dynamics," *RAND Journal of Economics*, 41, 215–243. 3, 10

Dubé, J.-P., J. T. Fox, and C.-L. Su (2012): "Improving the Numerical Performance of Static and Dynamic Aggregate Discrete Choice Random Coefficients Demand Estimation," *Econometrica*, 80, 2231–2267. 65

Dunne, T., S. Klimek, M. J. Roberts, and D. Y. Xu (2013): "Entry, Exit, and the Determinants of Market Structure," *RAND Journal of Economics*, 44, 462–487. 5

Dunne, T., M. J. Roberts, and L. Samuelson (1988): "Patterns of Firm Entry and Exit in U.S. Manufacturing Industries," *RAND Journal of Economics*, 19, 495–515. 26

Eizenberg, A. (2014): "Upstream Innovation and Product Variety in the U.S. Home PC Market," *Review of Economic Studies*, 81, 1003–1045. 5

Ericson, R. and A. Pakes (1995): "Markov-Perfect Industry Dynamics: A Framework for Empirical Work," *Review of Economic Studies*, 62, 53–82. 3, 4, 5, 6, 9, 10, 20, 41

Gil, R. and F. LaFontaine (2012): "Using Revenue Sharing to Implement Flexible Prices: Evidence from Movie Exhibition Contracts," *Journal of Industrial Economics*, 60, 187–219. 34

Gowrisankaran, G. (1999): "A Dynamic Model of Endogenous Horizontal Mergers," *RAND Journal of Economics*, 30, 56–83. 3

Haltiwanger, J., R. Jarmin, and C. J. Krizan (2010): "Mom-and-Pop meet Big-Box: Complements or Substitutes?" *Journal of Urban Economics*, 67, 116–134. 28

Harwell, D. (2015): "Netflix Plan: Bring Big-Screen Premieres to Subscribers' Small Screens," *Washington Post*, https://www.washingtonpost.com/business/economy/netflix-plan-bring-big-screen-premiers-to-subscribers-small-screens/2015/10/14/cf197154-71d0-11e5-8248-98e0f5a2e830_ story.html; accessed on December 4, 2015. 2

Honoré, B. E. and Á. De Paula (2010): "Interdependent Durations," *Review of Economic Studies*, 77, 1138–1163. 8

Hopenhayn, H. A. (1992): "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," *Econometrica*, 60, 1127–1150. 2

Hotz, V. J. and R. A. Miller (1993): "Conditional Choice Probabilities and the Estimation of Dynamic Models," *Review of Economic Studies*, 60, 497–529. 64

Iskhakov, F., J. Lee, J. Rust, K. Seo, and B. Schjerning (2015a): "Constrained Optimization Approaches to Estimation of Structural Models: Comment," *Econometrica*, Forthcoming. 69

Iskhakov, F., J. Rust, and B. Schjerning (2015b): "Recursive Lexicographical Search: Finding All Markov Perfect Equilibria of Finite State Directional Dynamic Games," *Review of Economic Studies*, Forthcoming. 6

Judd, K. L., K. Schmedders, and Ş. Yeltekin (2012): "Optimal Rules for Patent Races," *International Economic Review*, 53, 23–52. 6

Kafka, P. (2013): "Netflix Flirts with a New Idea: "Big" Movies at Your House, the Same Day They're in Theaters," *AllThingsD.com*, http://allthingsd.com/20131028/netflix-flirts-with-a-new-idea-big-movies-at-your-house-the-same-day-theyre-in-theaters/; accessed on May 31, 2015. 2

Magnac, T. and D. Thesmar (2002): "Identifying Dynamic Discrete Decision Processes," *Econometrica*, 70, 801–816. 7, 60, 63

Maskin, E. and J. Tirole (1988): "A Theory of Dynamic Oligopoly, I: Overview and Quantity Competition with Large Fixed Costs," *Econometrica*, 56, 549–569. 10

McAfee, R. P., H. M. Mialon, and M. A. Williams (2004): "What Is a Barrier to Entry?" *American Economic Review*, 94, 461–465. 49

Newey, W. K. and J. L. Powell (2003): "Instrumental Variable Estimation of Nonparametric Models," *Econometrica*, 71, 1565–1578. 64

Orbach, B. Y. and L. Einav (2007): "Uniform Prices for Differentiated Goods: The Case of the Movie-Theater Industry," *International Review of Law and Economics*, 27, 129–153. 34

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, 373–399. 4, 5, 22

Reguant, M. (2015): "A Bounds Approach to Counterfactual Analysis," Mimeo, Department of Economics, Northwestern University, Evanston, IL. 5

Rust, J. (1987): "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55, 999–1033. 1, 10, 20, 21, 69, 70

Schwartzel, E. and B. Fritz (2015): "Paramount to Speed Up Home Release of Movies," *Wall Street Journal*, http://www.wsj.com/articles/paramount-to-break-hollywoods-home-video-window-1436377631; accessed on July 10, 2015. 2

Su, C.-L. and K. L. Judd (2012): "Constrained Optimization Approaches to Estimation of Structural Models," *Econometrica*, 80, 2213–2230. 21, 68, 69, 70

Sutton, J. (1991): *Sunk Costs and Market Structure*, MIT Press. 8

Takahashi, Y. (2015): "Estimating a War of Attrition: The Case of the U.S. Movie Theater Industry," *American Economic Review*, 105, 2204–2241. 2

Tauchen, G. (1986): "Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions," *Economics Letters*, 20, 177–181. 29

Weintraub, G. Y., C. L. Benkard, and B. Van Roy (2008): "Markov Perfect Industry Dynamics with Many Firms," *Econometrica*, 76, 1375–1411. 2

Appendices

The model in the main text embodies structure on the stochastic processes for the state variables and their effects on profits that simplifies its maximum likelihood estimation but contributes nothing to its theoretical analysis. Appendices A, B, and C present a *general model* without this structure that encompasses the *special model* of the main text, define equilibrium in that model, and prove the appropriate generalizations of Lemmas 1 and 2, Corollaries 1 and 2, and Theorem 1. The remaining two appendices support the model's empirical application. Appendix D contains our identification analysis, while Appendix E presents the results of Monte Carlo experiments.

A Primitives

In the general model, firms' profits in period t depend on the $\mathcal{Y}$-valued vector Y_t. Figure 7 gives the model's recursive extensive form. Period t starts in state (N_t, Y_t), and each incumbent earns profits $\tilde{\pi}(N_t, Y_t)$.[25] As in the special model, all players have names giving the date of their entry opportunity and their position in that date's entry queue. In the entry subgame of period t, firm (t, j) pays the sunk cost $\tilde{\varphi}(M_t^j, Y_t)$ upon entry, where again $M_t^j \equiv N_t + j$. As before, a potential entrant's payoff from choosing inactivity equals zero. Progressing to the period t survival subgame, an active firm choosing survival incurs no cost *during period t*. The expected profits from operating in period $t+1$ subsume the special model's costs of continuation. At the end of the period, nature draws Y_{t+1} from the Markov transition distribution $\tilde{G}(\cdot\,|Y_t)$.

The restrictions we place on the payoffs are

A1. $\mathbb{E}\left[\tilde{\pi}(n, Y')|Y = y\right]$ exists for all $y \in \mathcal{Y}$ and $\exists \check{\pi} < \infty$ such that $\forall n \in \mathbb{N}$ and $\forall y \in \mathcal{Y}$, $\mathbb{E}\left[\tilde{\pi}(n, Y')|Y = y\right] < \check{\pi}$;

A2. $\exists \check{n} \in \mathbb{N} : \forall n > \check{n}$ and $\forall y \in \mathcal{Y}$, $\tilde{\pi}(n, y) < 0$;

A3. $\forall n \in \mathbb{N}$ and $\forall y \in \mathcal{Y}$, $\tilde{\pi}(n, y) \geq \tilde{\pi}(n+1, y)$; and

[25] Here and throughout this appendix, we place a tilde over any primitive, strategy, or value function with a similar name in the special model.

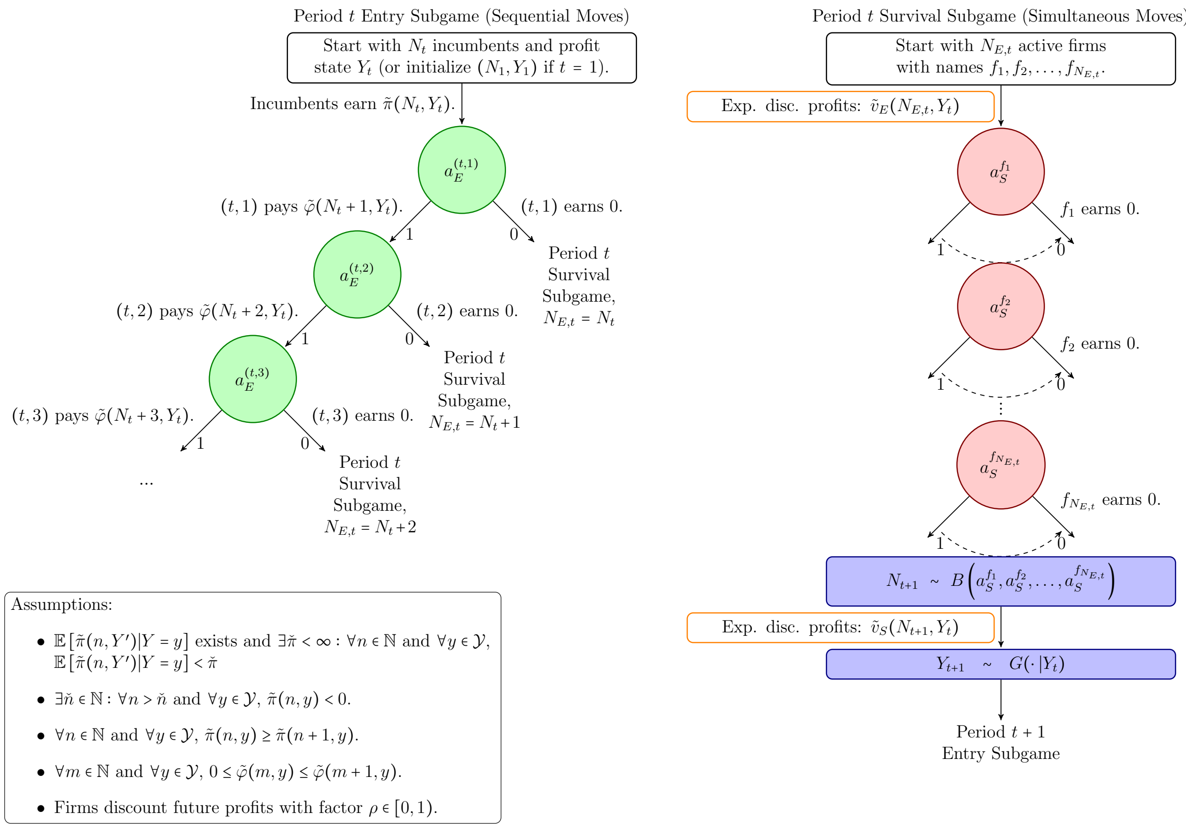


Figure 7: The General Model's Recursive Extensive Form

A4. $\forall m \in \mathbb{N}$ and $\forall y \in \mathcal{Y}$, $0 \leq \tilde{\varphi}(m, y) \leq \tilde{\varphi}(m + 1, y)$.

To cast the special model within this more general framework, set

$$\begin{aligned}
Y_t &\equiv (C_t, W_t, W_{t-1}), \\
\tilde{\pi}(n; c, w, w_{-1}) &\equiv \pi(n, c) - \rho^{-1}\kappa \exp(w_{-1}), \\
\tilde{\varphi}(m; c, w, w_{-1}) &\equiv \varphi \exp(w), \text{ and} \\
\tilde{G}(c, w, w_{-1}|C_{t-1}, W_{t-1}, W_{t-2}) &\equiv \begin{cases} G_C(c|C_{t-1})G_W(w) & \text{if } W_{t-1} \leq w_{-1} \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}$$

Note that in the special model,

$$\mathbb{E}[\tilde{\pi}(n, Y')|Y = (c, w, w_{-1})] = \mathbb{E}[\pi(n, C')|C = c] - \rho^{-1}\kappa \exp(w).$$

Thus, the assumptions that $\mathbb{E}[\pi(n, C')|C = c] \leq \breve{\pi}$, $\pi(n, c) \geq 0$, and $\kappa > 0$ ensure

that $\mathbb{E}[\tilde{\pi}(n, Y')|Y = y]$ exists and is bounded above by $\breve{\pi}$, so Assumption A1 holds. This allows us to restrict equilibrium post-entry values to the space of bounded non-negative functions. The special model's assumptions that $\kappa > 0$ and that there exists $\check{n}$ such that $\pi(n, c) = 0$ for all $n > \check{n}$ imply Assumption A2, which is key to bounding the number of firms in equilibrium. The assumption that $\pi(n, c) \geq \pi(n + 1, c)$ in the special model directly gives Assumption A3. Finally, Assumption A4 generalizes the special model's assumption of a constant $\varphi \geq 0$. Since $\tilde{\varphi}(m, \cdot)$ may increase in m, the general model can have an *economic barrier to entry* (McAfee, Mialon, and Williams, 2004).[26]

B Equilibrium

For both a potential entrant and a firm contemplating survival, the payoff-relevant variables are the number of firms committed to play that period's survival subgame and the current profit state. A Markovian strategy is a pair of functions, $\tilde{a}_E : \mathbb{N} \times \mathcal{Y} \to \{0, 1\}$ and $\tilde{a}_S : \mathbb{N} \times \mathcal{Y} \to [0, 1]$. As in the special model (see Footnote 2), we interpret $\tilde{a}_S(n_E, y) = 0$ and $\tilde{a}_S(n_E, y) = 1$ as dictating certain exit and survival in state (n_E, y). An equilibrium strategy and its associated continuation values $\tilde{v}_E$ and $\tilde{v}_S$ satisfy

$$\begin{aligned}
\tilde{v}_E(n_E, y) &= \max_{a \in [0,1]} a \mathbb{E}_{\tilde{a}_S} [\tilde{v}_S(N', y)|N_E = n_E, Y = y], && (1)\\
\tilde{v}_S(n', y) &= \rho \mathbb{E}_{\tilde{a}_E} [\tilde{\pi}(n', Y') + \tilde{v}_E(N'_E, Y')|N' = n', Y = y], && (2)\\
\tilde{a}_E(m, y) &\in \arg\max_{a \in \{0,1\}} a\big(\mathbb{E}_{\tilde{a}_E} [\tilde{v}_E(N_E, y)|M = m, Y = y] - \tilde{\varphi}(m, y)\big), \text{ and} && (3)\\
\tilde{a}_S(n_E, y) &\in \arg\max_{a \in [0,1]} a \mathbb{E}_{\tilde{a}_S} [\tilde{v}_S(N', y)|N_E = n_E, Y = y]. && (4)
\end{aligned}$$

The expectation operators condition on the deciding firm choosing to be active in the next period and on all other firms using the entry or exit rule in the operator's subscript. Note that in the special model $\tilde{v}_E$, $\tilde{a}_E$ and $\tilde{a}_S$ equal v_E, a_E and a_S; but $\tilde{v}_S(n'; c, w, w_{-1}) = v_S(n', c) - \kappa \exp(w)$ because the general model incorporates the continuation costs into $\tilde{\pi}(n, y)$. As in the text, we restrict attention to equilibria

[26] Although adding a barrier to entry to the model's theoretical analysis is straightforward, our identification proof does rely on the special model's constant specification for $\tilde{\varphi}(m, \cdot)$. Thus, the identification of barriers to entry remains an important open area of inquiry.

in strategies that default to inactivity. In such an equilibrium, a potential entrant that is indifferent between entering or not stays out,

$$\mathbb{E}_{\tilde{a}_E}\left[\tilde{v}_E(N_E, y) | M = m, Y = y\right] = \tilde{\varphi}(m, y) \Rightarrow \tilde{a}_E(m, y) = 0,$$

and an active firm that is indifferent between *all* possible outcomes of the survival stage exits,

$$\tilde{v}_S(n_E, y) = \cdots = \tilde{v}_S(1, y) = 0 \Rightarrow \tilde{a}_S(n_E, y) = 0.$$

The general model's equilibrium characterization begins with the appropriate analogues to Lemmas 1 and 2 and Corollaries 1 and 2. Appendix C contains their proofs.

Lemma 1* (Bounded number of firms in the general model) *In a symmetric Markov-perfect equilibrium that defaults to inactivity, $\forall y \in \mathcal{Y}$ and $\forall n > \check{n}$, $\tilde{a}_E(n, y) = 0$ and $\tilde{a}_S(n, y) < 1$.*

Lemma 2* (Monotone equilibrium payoffs in the general model) *In a symmetric Markov-perfect equilibrium that defaults to inactivity, $\forall y \in \mathcal{Y}$, $\tilde{v}_S(n', y)$ weakly decreases with n'.*

Corollary 1* *Let $\tilde{v}_S$ be the post-survival value function associated with a symmetric Markov-perfect equilibrium that defaults to inactivity. Consider the one-shot survival game in which n_E firms simultaneously choose between survival and exit (as in the survival subgame of Figure 7), each of the n' survivors earns $\tilde{v}_S(n', y)$, with $\tilde{v}_S(n', y) \neq 0$ for at least one $n' \in \{1, \ldots, n_E\}$, and each exiting firm earns zero. This game has a unique symmetric Nash equilibrium, possibly in mixed strategies.*

Corollary 2* *If $\tilde{v}_E$ and $\tilde{v}_S$ are the post-entry and post-survival value functions associated with a symmetric Markov-perfect equilibrium that defaults to inactivity, then*

$$\tilde{v}_E(n_E, y) = \max\{0, \tilde{v}_S(n_E, y)\}.$$

Just as with the special model, we constructively demonstrate equilibrium existence and uniqueness. We present the algorithm for this equilibrium calculation

in Algorithm 1. It begins by initializing the number of firms under consideration n to $\check{n}$ and both the candidate equilibrium entry rule $\tilde{\alpha}_E : \mathbb{N} \times \mathcal{Y} \to \{0,1\}$ and a dummy function $f^\star : \mathcal{Y} \to [0, \frac{\rho\check{\pi}}{1-\rho}]$ to zero. We will denote the candidate equilibrium post-entry and post-survival value functions by $\tilde{\nu}_E$ and $\tilde{\nu}_S$, respectively. The algorithm then enters its main loop, which begins by constructing the number of firms that will be active if we begin with n firms and potential entrants follow the rule in the current value of $\tilde{\alpha}_E$. For a given $y \in \mathcal{Y}$, this is

$$\tilde{\mu}(n,y) \equiv n + \sum_{n'=n+1}^{\infty} \prod_{m=n+1}^{n'} \tilde{\alpha}_E(m,y).$$

In the first pass through the loop, $\tilde{\mu}(\check{n}, y) = \check{n}$.

After initialization, the algorithm uses Bellman equation iteration (on the dummy function $f^\star$) to solve the general model's dynamic programming problem. The relevant Bellman operator is

$$\begin{aligned} T_n(f^\star)(y) = \max\{0, \rho\mathbb{E}[\tilde{\pi}(n,Y') + \mathbb{1}[\tilde{\mu}(n,Y') = n]f^\star(Y') \\ + \mathbb{1}[\tilde{\mu}(n,Y') > n]\tilde{\nu}_E(\tilde{\mu}(n,Y'),Y') \,|\, Y = y]\}. \end{aligned} \tag{5}$$

The next two steps assign the fixed point of T_n stored in $f^\star(\cdot)$ to $\tilde{\nu}_E(n,\cdot)$ and use (2) to construct $\tilde{\nu}_S(n,\cdot)$. In the loop's final step, the initial value of $\tilde{\alpha}_E(n,y)$ is replaced with $\mathbb{1}[\tilde{\nu}_E(n,y) > \tilde{\varphi}(n,y)]$. If the current value of n exceeds one, it is decremented and the algorithm returns to the top of the main loop. If instead n equals one, then the algorithm proceeds to its final task, setting the candidate equilibrium survival rule to zero in the trivial case that $\tilde{\nu}_S(n,\cdot) = \cdots = \tilde{\nu}_S(1,\cdot) = 0$, and to the highest probability consistent with equilibrium in the survival subgame otherwise.[27] Algorithm 1 only computes candidate post-entry and post-survival values and a candidate survival rule on $\{1,\ldots,\check{n}\} \times \mathcal{Y}$. As noted in the main text, it is straightforward to extend them to the full state space $\mathbb{N} \times \mathcal{Y}$. Because this extension is not of much applied interest, we keep it implicit and simply refer to Algorithm 1 as computing a candidate equilibrium.

The appropriate generalization of Theorem 1 to the general framework states that the candidate equilibrium strategies and payoffs arising from Algorithm 1 correspond to the unique Markov-perfect Nash equilibrium that defaults to inactivity.

[27] Because these survival rule calculations do not build on each other, they can be parallelized.

Theorem 1* (Equilibrium existence and uniqueness in the general model) *There exists a unique symmetric Markov-perfect equilibrium that defaults to inactivity. The equilibrium strategy and corresponding equilibrium payoffs are those computed by Algorithm 1.*

C Proofs

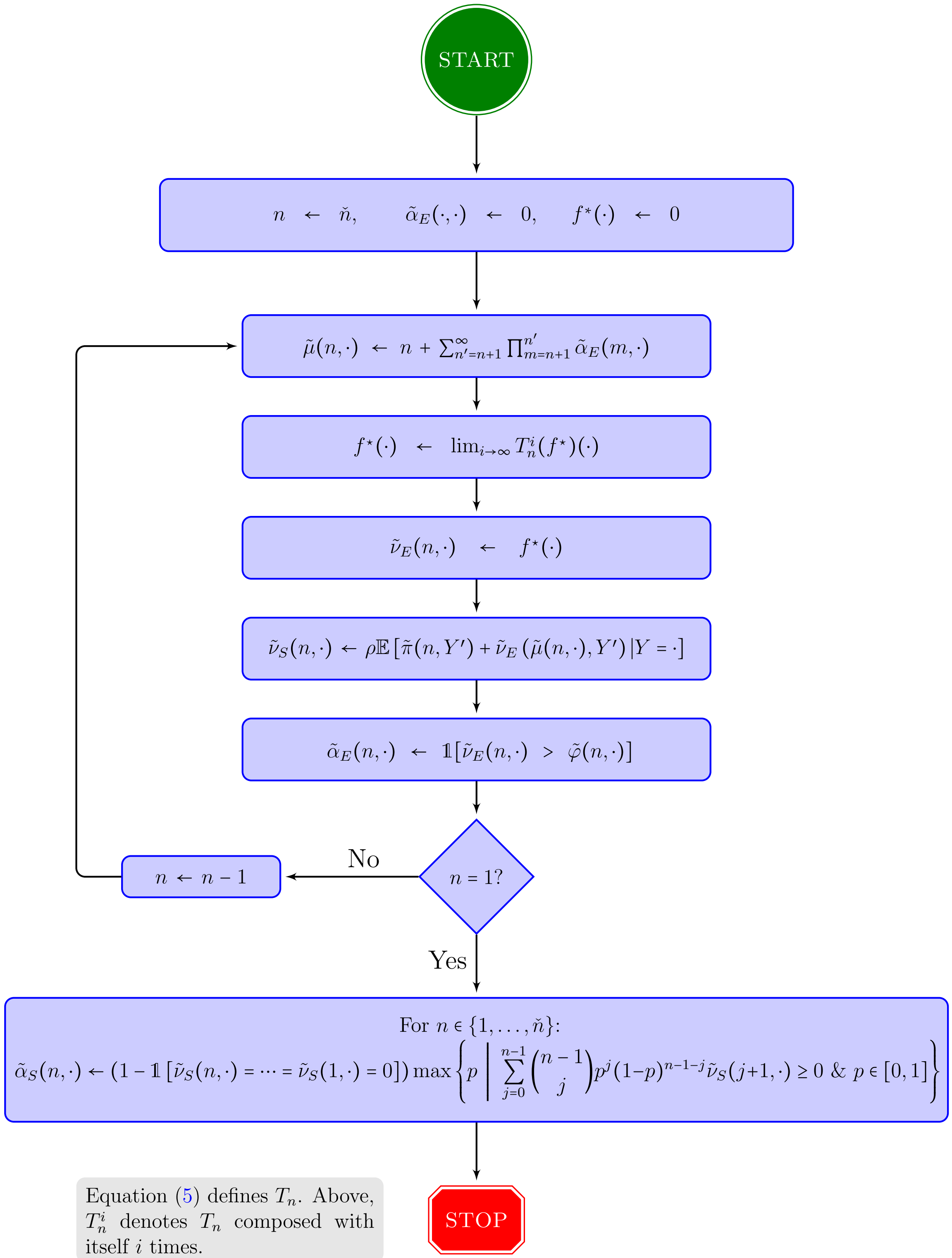
This appendix contains the formal proofs of the numbered results in Appendices A and B.

Proof of Lemma 1*. First, we will prove that $\tilde{a}_S(n,y) < 1$ for all $y \in \mathcal{Y}$ and $n > \check{n}$. Consider a period-$t^\star$ survival subgame with $N_{E,t^\star} = n > \check{n}$ firms and profit state $Y_{t^\star} = y$. Define the random time τ as the first period weakly after $t^\star$ in which the firms choose exit with a positive probability, with $\tau \equiv \infty$ if they never do:

$$\tau \equiv \min\{\{t \geq t^\star : \tilde{a}_S(N_{E,t}, Y_t) < 1\} \cup \{\infty\}\}.$$

We need to show that $\tilde{a}_S(n,y) < 1$. Suppose to the contrary that $\tilde{a}_S(n,y) = 1$. Then, $\tau > t^\star$. By definition, exit occurs only in or after period τ, so we know that $N_t = N_{E,t-1} \geq n$ for $t \in \{t^\star + 1, \ldots, \tau\}$. (Recall that we take $\tilde{a}_S(\cdot) = 1$ to dictate sure survival, not merely almost-sure survival.) Since $n > \check{n}$, this together with Assumption A2 implies that $\tilde{\pi}(N_t, Y_t) < 0$ for $t \in \{t^\star + 1, \ldots, \tau\}$. If $\tau = \infty$, then the incumbent firms receive an infinite sequence of strictly negative payoffs. If instead $\tau < \infty$, then the incumbent firms receive a finite sequence of strictly negative payoffs followed by the expected continuation value from playing the period-τ survival subgame $\tilde{v}_E(N_{E,\tau}, Y_\tau)$, which equals zero by (1), (4), and the definition of τ. Therefore, the period-$t^\star$ post-survival payoff $\tilde{v}_S(n,y) < 0$. Since any of the n period-$t^\star$ incumbent firms can raise its payoff to zero by choosing certain exit, the supposition that $\tilde{a}_S(n,y) = 1$ must be incorrect.

Next, we will prove that $\tilde{a}_E(n,y) = 0$ for all $y \in \mathcal{Y}$ and $n > \check{n}$. Consider the decision of the first potential entrant, firm $(t^\star, 1)$, in a period-$t^\star$ entry subgame that starts with $N_{t^\star} = n - 1 > \check{n} - 1$ incumbents and profit state $Y_{t^\star} = y$. (This value of n need not equal that used above to initialize an arbitrary period-$t^\star$ survival subgame.) Note that this firm pays $\tilde{\varphi}(n,y)$ upon entry. In return, it receives a



Algorithm 1: Equilibrium Calculation for the General Model

continuation value of zero (as proven above). If $\tilde{\varphi}(n,y) > 0$, then it maximizes its payoff by staying out of the industry and earning zero. If instead $\tilde{\varphi}(n,y) = 0$, then the assumption that the equilibrium strategy defaults to inactivity dictates the same action. In either case, $\tilde{a}_E(n,y) = 0$ as asserted. ∎

Proof of Lemma 2⋆. It suffices to prove that $\tilde{v}_S(n',y) \geq \tilde{v}_S(n'+1,y)$ for all post-survival number of firms $n' \geq 1$ and profit states $y \in \mathcal{Y}$. To this end, consider a period-$(t^\star+1)$ entry subgame with $N_{t^\star+1} = n'$. We refer to this as the *original* subgame and denote its equilibrium outcomes for $t \geq t^\star + 1$ with N_t and $N_{E,t}$. Now consider a second period-$(t^\star+1)$ entry subgame with the same profit state process, but one additional firm. We refer to this as the *perturbed* subgame and denote its initial number of firms with $N^+_{t^\star+1} = n'+1$ and its outcomes for $t \geq t^\star+1$ with N^+_t and $N^+_{E,t}$.

Define the random times τ and τ^+ with

$$
\begin{aligned}
\tau &\equiv \min\{\{t \geq t^\star+1 : \tilde{a}_S(N_{E,t}, Y_t) < 1\} \cup \{\infty\}\} \text{ and} \\
\tau^+ &\equiv \min\{\{t \geq t^\star+1 : \tilde{a}_S(N^+_{E,t}, Y_t) < 1\} \cup \{\infty\}\}.
\end{aligned}
$$

Then, there is no exit before period τ and τ^+ in the original and perturbed subgames, respectively. Consequently, $N_t \leq N^+_t$ and $N_{E,t} \leq N^+_{E,t}$ for all $t = t^\star+1, \ldots, \min\{\tau, \tau^+\}$. (Otherwise, the two subgames would have potential entrants in the same payoff-relevant states making different entry decisions, which would violate the assumption that the equilibrium strategy is Markovian.) This and Assumption A3 guarantee that

$$
\tilde{\pi}(N_t, Y_t) \geq \tilde{\pi}(N^+_t, Y_t) \tag{6}
$$

for all $t = t^\star+1, \ldots, \min\{\tau, \tau^+\}$.

If $\tau = \tau^+ = \infty$, then the desired result follows directly from (6):

$$
\begin{aligned}
\tilde{v}_S(n',y) &= \mathbb{E}\left[\sum_{t=t^\star+1}^{\infty} \rho^{t-t^\star} \tilde{\pi}(N_t, Y_t) \,\middle|\, Y_{t^\star} = y\right] \\
&\geq \mathbb{E}\left[\sum_{t=t^\star+1}^{\infty} \rho^{t-t^\star} \tilde{\pi}(N^+_t, Y_t) \,\middle|\, Y_{t^\star} = y\right] = \tilde{v}_S(n'+1, y).
\end{aligned}
$$

If instead $\min\{\tau,\tau^+\} < \infty$, then

$$\begin{aligned}\tilde{v}_S(n',y) =& \mathbb{E}\left[\sum_{t=t^\star+1}^{\min\{\tau,\tau^+\}} \rho^{t-t^\star}\tilde{\pi}(N_t,Y_t)\middle| Y_{t^\star}=y\right] \\ &+ \mathbb{E}\left[\rho^{\min\{\tau,\tau^+\}-t^\star}\tilde{v}_E\left(N_{E,\min\{\tau,\tau^+\}},Y_{\min\{\tau,\tau^+\}}\right)\middle| Y_{t^\star}=y\right]\end{aligned}$$

and

$$\begin{aligned}\tilde{v}_S(n'+1,y) =& \mathbb{E}\left[\sum_{t=t^\star+1}^{\min\{\tau,\tau^+\}} \rho^{t-t^\star}\tilde{\pi}(N_t^+,Y_t)\middle| Y_{t^\star}=y\right] \\ &+ \mathbb{E}\left[\rho^{\min\{\tau,\tau^+\}-t^\star}\tilde{v}_E\left(N^+_{E,\min\{\tau,\tau^+\}},Y_{\min\{\tau,\tau^+\}}\right)\middle| Y_{t^\star}=y\right].\end{aligned}$$

Inequality (6) implies that

$$\mathbb{E}\left[\sum_{t=t^\star+1}^{\min\{\tau,\tau^+\}} \rho^{t-t^\star}\tilde{\pi}(N_t,Y_t)\middle| Y_{t^\star}=y\right] \geq \mathbb{E}\left[\sum_{t=t^\star+1}^{\min\{\tau,\tau^+\}} \rho^{t-t^\star}\tilde{\pi}(N_t^+,Y_t)\middle| Y_{t^\star}=y\right],$$

so it suffices to prove that

$$\tilde{v}_E\left(N_{E,\min\{\tau,\tau^+\}},Y_{\min\{\tau,\tau^+\}}\right) \geq \tilde{v}_E\left(N^+_{E,\min\{\tau,\tau^+\}},Y_{\min\{\tau,\tau^+\}}\right). \tag{7}$$

To establish (7), we consider two subcases of $\min\{\tau,\tau^+\} < \infty$.

- If $\tau^+ = \min\{\tau,\tau^+\}$, then $\tilde{a}_S(N^+_{E,\tau^+},Y_{\tau^+}) < 1$ and $\tilde{v}_E(N^+_{E,\tau^+},Y_{\tau^+}) = 0$. Because post-entry values are non-negative, this implies (7) for $\tau^+ = \min\{\tau,\tau^+\}$.
- If $\tau = \min\{\tau,\tau^+\} < \tau^+$, then $\tilde{a}_S(N_{E,\tau},Y_\tau) < 1$ and there exists an $n \leq N_{E,\tau}$ such that $\tilde{v}_S(n,Y_\tau) \leq 0$ (otherwise, choosing certain survival would give a strictly higher payoff than choosing any positive probability of exit). Now consider a third, period-$(\tau+1)$ entry subgame with the same profit state process, but n firms initially. We refer to this as the n-subgame. Suppose that one of these firms, possibly in deviation from the equilibrium survival rule, never exits if $\tau^+ = \infty$, and does not exit before period τ^+ and follows the equilibrium survival rule from period τ^+ onwards if $\tau^+ < \infty$. Suppose that all other firms use the equilibrium strategy and denote the corresponding outcomes for $t \geq \tau+1$ with N_t^n and $N_{E,t}^n$. In the argument leading to (6), we have already established that $N_{E,\tau} \leq N^+_{E,\tau}$, so that $N^n_{\tau+1} = n \leq N_{E,\tau} \leq N^+_{E,\tau} = N^+_{\tau+1}$ (the last equality

follows because there is no exit in period $\tau < \tau^+$ of the perturbed game). This implies that $N_t^n \le N_t^+$ for $t = \tau+1, \ldots, \tau^+$ (otherwise, potential entrants facing the same payoff-relevant state would make different decisions, which would violate the assumption that the equilibrium strategy is Markovian).

If $\tau^+ = \infty$, we have that

$$\begin{aligned}\tilde{v}_S(n, Y_\tau) &\ge \mathbb{E}\left[\sum_{t=\tau+1}^{\infty} \rho^{t-\tau}\pi(N_t^n, Y_t)\,\middle|\, Y_\tau\right] \\ &\ge \mathbb{E}\left[\sum_{t=\tau+1}^{\infty} \rho^{t-\tau}\pi(N_t^+, Y_t)\,\middle|\, N_{\tau+1}^+, Y_\tau\right] = \tilde{v}_S(N_{\tau+1}^+, Y_\tau).\end{aligned} \tag{8}$$

Note that $\mathbb{E}\left[\sum_{t=\tau+1}^{\infty} \rho^{t-\tau}\pi(N_t^n, Y_t)\,|\, Y_\tau\right]$ equals the expected discounted profits that accrue to the (possibly) deviating firm in the n-subgame. The first inequality in (8) arises because this firm can only weakly increase its expected discounted payoffs by following the equilibrium survival rule, which would give it $\tilde{v}_S(n, Y_\tau)$. The second inequality in (8) comes from $N_t^n \le N_t^+$ and Assumption A3.

Similarly, if $\tau^+ < \infty$, we have that

$$\begin{aligned}\tilde{v}_S(n, Y_\tau) &\ge \mathbb{E}\left[\sum_{t=\tau+1}^{\tau^+} \rho^{t-\tau}\pi(N_t^n, Y_t) + \rho^{\tau^+-\tau} v_E(N_{E,\tau^+}^n, Y_{\tau^+})\,\middle|\, Y_\tau\right] \\ &\ge \mathbb{E}\left[\sum_{t=\tau+1}^{\tau^+} \rho^{t-\tau}\pi(N_t^n, Y_t)\,\middle|\, Y_\tau\right] \\ &\ge \mathbb{E}\left[\sum_{t=\tau+1}^{\tau^+} \rho^{t-\tau}\pi(N_t^+, Y_t)\,\middle|\, N_{\tau+1}^+, Y_\tau\right] = \tilde{v}_S(N_{\tau+1}^+, Y_\tau).\end{aligned} \tag{9}$$

The first and third inequalities follow as in the case with $\tau^+ = \infty$. The second inequality uses that post-entry values are non-negative. The final equality follows from the definition of τ^+, which implies that $\tilde{v}_E(N_{E,\tau^+}^+, Y_{\tau^+}) = 0$.

Finally, it follows from the definition of τ^+ and $\tau < \tau^+$ that $\tilde{a}_S(N_{E,\tau}^+, Y_\tau) = 1$. Thus, $\tilde{v}_E(N_{E,\tau}^+, Y_\tau) = \tilde{v}_S(N_{\tau+1}^+, Y_\tau) \ge 0$. Together with (8), (9), and $\tilde{v}_S(n, Y_\tau) \le 0$, this implies that $\tilde{v}_E(N_{E,\tau}^+, Y_\tau) = 0$. The non-negativity of post-entry values therefore guarantees that (7) holds for $\tau = \min\{\tau, \tau^+\} < \tau^+$.

Together, the results from these two subcases establish (7). ■

Proof of Corollary 1*. Corollary 1*'s one-shot survival game falls into one of three mutually-exclusive cases.

- $\tilde{v}_S(1, y) \leq 0$. Lemma 2* implies that $\tilde{v}_S(n', y) \leq 0$ for all $n' > 1$. Therefore, exiting for sure (setting $\tilde{a}_S(n_E, y) = 0$) is a weakly dominant strategy and forms one symmetric equilibrium. Furthermore, since $\tilde{v}_S(n', y) \neq 0$ for at least one $n' \in \{1, \ldots, n_E\}$, we know that $\tilde{v}_S(n_E, y) < 0$. Therefore, exiting for sure is also the unique best response to any positive symmetric continuation probability.

- $\tilde{v}_S(n_E, y) \geq 0$. Lemma 2* implies that $\tilde{v}_S(n', y) \geq 0$ for all $n' \in \{1, \ldots, n_E\}$. Therefore, continuing for sure (setting $\tilde{a}_S(n_E, y) = 1$) is a weakly dominant strategy and forms a symmetric equilibrium. Furthermore, since $\tilde{v}_S(n', y) \neq 0$ for at least one $n' \in \{1, \ldots, n_E\}$, we know that $\tilde{v}_S(1, y) > 0$. Therefore, continuing for sure is also the unique best response to any symmetric continuation probability less than one.

- $\tilde{v}_S(1, y) > 0$ and $\tilde{v}_S(n_E, y) < 0$. No symmetric pure strategy equilibrium exists, because the best response to all other firms continuing for sure is to exit for sure, and vice versa. In a mixed strategy equilibrium, firms must be indifferent between continuation and exit. By the intermediate value theorem, there is some $a_S \in (0, 1)$ that solves

$$\sum_{n'=1}^{n_E} \binom{n_E - 1}{n' - 1} a_S^{n'-1} (1 - a_S)^{n_E - n'} \tilde{v}_S(n', y) = 0. \tag{10}$$

  Lemma 2* guarantees that the left hand side of (10) weakly decreases with a_S, and the subcase's conditions strengthen that conclusion so that the left hand side of (10) strictly decreases in a_S. So, there is only one symmetric mixed strategy equilibrium.

This establishes the equilibrium uniqueness asserted by Corollary 1*. ■

Proof of Corollary 2*. The survival rule in a symmetric Markov-perfect equilibrium takes values equal to Corollary 1*'s symmetric Nash equilibrium strategies $\tilde{a}_S(n_E, y)$ in states (n_E, y) such that $\tilde{v}_S(n', y) \neq 0$ for at least one $n' \in \{1, \ldots, n_E\}$. The analysis in Corollary 1*'s proof implies that in these states $\mathbb{E}_{\tilde{a}_S}[\tilde{v}_S(N', y) | N_E = n_E, Y = y] = \tilde{v}_S(n_E, y)$ if $\tilde{v}_S(n_E, y) \geq 0$ (in which case

$\tilde{a}_S(n_E, y) = 1$) and $\mathbb{E}_{\tilde{a}_S}[\tilde{v}_S(N', y)|N_E = n_E, Y = y] = 0$ if $\tilde{v}_S(n_E, y) < 0$ ($\tilde{a}_S(n_E, y) < 1$). Moreover, in states (n_E, y) such that $\tilde{v}_S(1, y) = \cdots = v_S(n_E, y) = 0$, $\mathbb{E}_{\tilde{a}_S}[\tilde{v}_S(N', y)|N_E = n_E, Y = y] = \tilde{v}_S(n_E, y) = 0$. Either way, (1) reduces to $\tilde{v}_E(n_E, y) = \max\{0, \tilde{v}_S(n_E, y)\}$. ■

Proof of Theorem 1⋆. The proof is divided into three parts. First, we show that the candidate continuation values from Algorithm 1 satisfy the monotonicity requirements of Lemma 2⋆. Second, we use this to demonstrate that the candidate continuation values, survival rule, and entry rule satisfy (1)-(4). This establishes that the candidate strategy indeed forms an equilibrium. Third, we demonstrate that this equilibrium's existence implies its uniqueness.

Let $\tilde{\mu}$ be the value of the transition rule at the conclusion of Algorithm 1. (Because $\tilde{\mu}(n', \cdot)$ is set at the beginning of the loop iteration with $n = n'$ and is never altered again, it has the same value as that was used to construct the Bellman operator $T_{n'}$.) Fix $n \in \{1, 2, \ldots, \check{n}\}$ and suppose that for all n' such that $n+1 \leq n' \leq \check{n}$, we know that $\tilde{\nu}_E(n', \cdot) \geq \tilde{\nu}_E(n'+1, \cdot)$. (This condition is trivially true for $n = \check{n}$.) Consider evaluating T_n at the value of $f^\star$ in memory *after* the completion of the $n+1$-indexed dynamic programming problem. For all $y \in \mathcal{Y}$, this gives

$$\begin{aligned}
T_n(f^\star)(y) = &\max\{0, \rho\mathbb{E}[\tilde{\pi}(n, Y') + f^\star(Y') \\
&\quad + \mathbb{1}[\tilde{\mu}(n, Y') > n](\tilde{\nu}_E(\tilde{\mu}(n, Y'), Y') - f^\star(Y'))|Y = y]\} \\
\geq &\max\{0, \rho\mathbb{E}[\tilde{\pi}(n+1, Y') + f^\star(Y') \qquad (11)\\
&\quad + \mathbb{1}[\tilde{\mu}(n, Y') > n](\tilde{\nu}_E(\tilde{\mu}(n, Y'), Y') - f^\star(Y'))|Y = y]\} \\
= &\max\{0, \rho\mathbb{E}[\tilde{\pi}(n+1, Y') + f^\star(Y') \qquad (12)\\
&\quad + \mathbb{1}[\tilde{\mu}(n, Y') > n+1](\tilde{\nu}_E(\tilde{\mu}(n, Y'), Y') - f^\star(Y'))|Y = y]\} \\
= &\max\{0, \rho\mathbb{E}[\tilde{\pi}(n+1, Y') + f^\star(Y') \qquad (13)\\
&\quad + \mathbb{1}[\tilde{\mu}(n+1, Y') > n+1](\tilde{\nu}_E(\tilde{\mu}(n+1, Y'), Y') - f^\star(Y'))|Y = y]\} \\
= &\tilde{\nu}_E(n+1, y).
\end{aligned}$$

The inequality in (11) follows from Assumption A3, and the equality in (12) from the equivalence of $\tilde{\nu}_E(n+1, Y')$ with $f^\star(Y')$. Since $\tilde{\nu}_E(n', Y)$ weakly decreases in n' for $n' > n$, so does $\tilde{\alpha}_E(n', Y')$. Therefore, $\tilde{\mu}(n, Y') = \tilde{\mu}(n+1, Y')$ whenever $\tilde{\mu}(n, Y') > n+1$. This gives us (13). The final equality follows again from the equivalence of $\tilde{\nu}_E(n+1, Y')$ with $f^\star(Y')$. The operator T_n is a monotone contraction

mapping, so $T_n(f^\star)(n,\cdot) \geq \tilde{\nu}_E(n+1,\cdot)$ implies that its fixed point, $\tilde{\nu}_E(n,\cdot)$, weakly exceeds $\tilde{\nu}_E(n+1,\cdot)$. This is the desired monotonicity result for the proof's first part.

For the second part, begin with (4). For states (n,y) such that $\tilde{\nu}_S(n,y) = \cdots = \tilde{\nu}_S(1,y) = 0$, (4) imposes only the trivial requirement that $\tilde{\alpha}_S(n,y) \in [0,1]$. Algorithm 1's selection of $\tilde{\alpha}_S(n,y) = 0$ satisfies this. For all other states (n,y), since $\tilde{\nu}_S(n,y)$ is weakly decreasing in n, Algorithm 1 sets $\tilde{\alpha}_S(n,y)$ to one if $\tilde{\nu}_S(n,y) \geq 0$, to zero if $\tilde{\nu}_S(1,y) \leq 0$, and to the highest symmetric survival probability that leaves all firms indifferent between continuation and exit if both $\tilde{\nu}_S(n,y) < 0$ and $\tilde{\nu}_S(1,y) > 0$. In all three cases, $\tilde{\alpha}_S(n,y)$ forms a Nash equilibrium to the one-shot survival game with survival payoffs $\tilde{\nu}_S(n',y)$ for $n' = 1,\ldots,n$, so it satisfies (4). Equation (1) requires $\tilde{\nu}_E(n,y)$ to equal the expected payoff to this game, $\max\{0,\tilde{\nu}_S(n,y)\}$. To establish this equality, recall that Algorithm 1 sets $\tilde{\nu}_E(n,y)$ to the fixed point of (5). Therefore, we have

$$\tilde{\nu}_E(n,y) = \max\{0, \rho\mathbb{E}[\tilde{\pi}(n,Y') + \tilde{\nu}_E(\tilde{\mu}(n,Y'),Y') \,|\, Y = y]\},$$

where $\tilde{\mu}(n,Y') = n + \sum_{n'=n+1}^{\infty} \prod_{m=n+1}^{n'} \tilde{\alpha}_E(m,Y')$ is the number of firms active at the start of the next period's survival game when all potential entrants use the candidate equilibrium entry rule. Algorithm 1 assigns this maximum operator's second argument to $\tilde{\nu}_S(n,y)$. This directly ensures that the continuation values satisfy (2), and it sets $\tilde{\nu}_E(n,y)$ to $\max\{0,\tilde{\nu}_S(n,y)\}$, as required to satisfy (1). The only condition remaining unverified is (3). Since $\tilde{\nu}_E(n,y)$ weakly decreases in n, $\tilde{\alpha}_E(n,y)$ weakly decreases in n. Therefore, $\mathbb{E}_{\tilde{\alpha}_E}[\tilde{\nu}_E(N_E,y)|M = n, Y = y] > \tilde{\varphi}(n,y)$ if and only if $\tilde{\nu}_E(n,y) > \tilde{\varphi}(n,y)$. (The "if" part of this statement relies upon the assumption above that $\tilde{\varphi}(m,y)$ is weakly increasing with m.) Since $\tilde{\alpha}_E(n,y)$ is an indicator for this inequality's truth, it prescribes entry if and only if entry gives a positive expected payoff. Thus, the candidate strategy and continuation values satisfy (3). We conclude that Algorithm 1's candidate strategy indeed forms an equilibrium.

The remainder of this proof demonstrates equilibrium uniqueness. Corollary $1^\star$ and the requirement that the strategy defaults to inactivity together imply that there is a unique equilibrium survival rule corresponding to every equilibrium post-survival value function. Lemma $2^\star$ and the requirement that the strategy defaults to inactivity imply that an equilibrium entry rule prescribes entry if and only if

the continuation value from entering and immediately proceeding to the survival subgame is strictly positive. Therefore, each pair of equilibrium continuation value functions $\tilde{v}_S$ and $\tilde{v}_E$ has exactly one corresponding equilibrium strategy. Corollary $2^\star$ requires any equilibrium continuation values to be fixed-points of the Bellman operators used in Algorithm 1. Since these operators are contractions, the continuation values constructed by Algorithm 1 are the only possible equilibrium continuation values. Their corresponding strategy forms the unique symmetric Markov-perfect Nash equilibrium that defaults to inactivity. ■

D Identification

For the special model of the text, this appendix analyzes the extent to which we can determine θ uniquely when we observe the population $(\{N_t, C_t; t = 1, \ldots, \check{t}\}, X)$ underlying our data. Specifically, suppose that we know the distribution of (N', C') conditional on $(N, C, X) = (n, c, x)$ for all $n \in \mathbb{N}_0 \equiv \{0\} \cup \mathbb{N}$, $c \in \mathcal{C}$, and a specific value x of the market characteristics.[28] Throughout the remainder of this section, we keep conditioning on $X = x$ implicit, so the results demonstrate identification of the model's primitives as nonparametric functions of the market characteristics.

To begin, note that the population information directly identifies G_C.[29] The remaining primitives of interest are the model's fixed cost, κ, sunk cost φ, surplus function π, and the distribution G_W of the econometric error. Our identification argument for these parameters builds upon that of Magnac and Thesmar (2002), who retrieve value functions by applying the inverse cumulative distribution function of the econometric error to observed choice probabilities. Since this strategy requires knowledge of G_W, we assume that this belongs to the parametric family

$$G_W(w) = \Phi\left(\frac{w + \omega^2/2}{\omega}\right), \tag{14}$$

where Φ is the cumulative distribution function of a standard normally distributed

[28] For x fixed, the hypothetical data scenario that is informative about this distribution involves the number of transitions from (N, C) to (N', C') approaching infinity. Whether such transitions are coming from the same market or many different markets all with characteristics x plays no role in the identification argument.

[29] Above, we specified this distribution as a function of a vector of parameters, θ_C. Such a parametric restriction might be of use when estimating using a finite sample, but it is not necessary for identification.

random variable, with density ϕ. That is, $\exp(W)$ has a log-normal distribution with unit mean and scale parameter ω. Since observations of the number of producers give us no information on the level of profits, we also normalize the mean per-period fixed cost κ to one.[30]

Continue by retrieving $\overline{w}_S(1,c)$ (a monopolist's survival threshold), up to the unknown scale and shift in G_W, from the probability of a monopolist surviving:

$$\frac{\overline{w}_S(1,c) + \omega^2/2}{\omega} = \Phi^{-1}\left(\Pr[N' \geq 1 | N = 1, C = c]\right).$$

Similarly, we can recover $\overline{w}_E(n,c)$ (the threshold from the probability of at least n firms entering a previously empty market):

$$\frac{\overline{w}_E(n,c) + \omega^2/2}{\omega} = \Phi^{-1}\left(\Pr[N' \geq n | N = 0, C = c]\right).$$

These and the definitions of $\overline{w}_S(1,c)$ and $\overline{w}_E(1,c)$ can be used to identify the sunk cost of entry up to the scale parameter ω:

$$\begin{aligned}\frac{\log(\varphi + 1)}{\omega} &= \frac{\overline{w}_S(1,c) - \overline{w}_E(1,c)}{\omega} \\ &= \Phi^{-1}\left(\Pr[N' \geq 1 | N = 1, C = c]\right) - \Phi^{-1}\left(\Pr[N' \geq 1 | N = 0, C = c]\right).\end{aligned}$$

In turn, this allows us to retrieve

$$\frac{\overline{w}_S(n,c) + \omega^2/2}{\omega} = \frac{\overline{w}_E(n,c) + \omega^2/2}{\omega} + \frac{\log(\varphi + 1)}{\omega}.$$

The argument's next step identifies the scale parameter ω. In a simple probit model, the analogous parameter is *not* identified unless one places an *a priori* restriction on the regressors' coefficients. For the present model, the mixing sometimes employed by exiting oligopolists provides information on the scale of payoffs relative to the econometric error. This information identifies ω without the use of auxiliary restrictions on payoffs.

To proceed, suppose that, for some $c^\star \in \mathcal{C}$ and $n^\star \in \{2, \ldots, \check{n}\}$,

$$\overline{w}_S(1, c^\star) = \cdots = \overline{w}_S(n^\star - 1, c^\star) > \overline{w}_S(n^\star, c^\star).$$

[30] This normalization implies that we do not identify cross-market differences in the *scale* of producers' surplus, fixed costs, and sunk costs. Rather, we identify producers' surplus and sunk costs relative to fixed costs for each market.

This is equivalent to requiring that

$$v_S(1, c^\star) = \cdots = v_S(n^\star - 1, c^\star) > v_S(n^\star, c^\star)$$

for some $c^\star$ and $n^\star$. This is a very weak condition, particularly in light of Lemma 2's result that $v_S(n, \cdot)$ *always* weakly decreases in n. Moreover, it can be verified in data, because we have already determined the sure survival thresholds up to a common scale and location shift.

Now, consider the probability of $n^\star$ incumbents simultaneously exiting:

$$\begin{aligned}
&\Pr[N' = 0 | N = n^\star, C = c^\star] \\
&\quad = \Pr[W \geq \overline{w}_S(1, c^\star)] + \int_{\overline{w}_S(n^\star, c^\star)}^{\overline{w}_S(1, c^\star)} [1 - a_S(n^\star, c^\star, w)]^{n^\star} g_W(w) dw \\
&\quad = \Pr[N' = 0 | N = 1, C = c^\star] + \int_{\overline{w}_S(n^\star, c^\star)}^{\overline{w}_S(1, c^\star)} [1 - a_S(n^\star, c^\star, w)]^{n^\star} g_W(w) dw. \quad (15)
\end{aligned}$$

Because the two transition probabilities in (15) are known, so is the integral on its right-hand side.

We will now show that this integral can be written as a known monotone function of ω, so that it identifies ω. Using $v_S(1, c^\star) = \cdots = v_S(n^\star - 1, c^\star)$, we can explicitly solve for the mixing probability $a_S(n^\star, c^\star, w)$:

$$a_S(n^\star, c^\star, w) = \left(\frac{v_S(1, c^\star) - \exp(w)}{v_S(1, c^\star) - v_S(n^\star, c^\star)} \right)^{\frac{1}{n^\star - 1}}.$$

Rewrite the integral on the right-hand side of (15) by substituting this expression for $a_S(n^\star, c^\star, w)$, replace post-survival values with sure survival thresholds, and change the variable of integration from w to $\varepsilon = (w + \omega^2/2)/\omega$. This gives

$$\int_{k_{n^\star}}^{k_1} \left[1 - \left(\frac{\exp(\omega k_1) - \exp(\omega \varepsilon)}{\exp(\omega k_1) - \exp(\omega k_{n^\star})} \right)^{\frac{1}{n^\star - 1}} \right]^{n^\star} \phi(\varepsilon) d\varepsilon, \quad (16)$$

with

$$k_1 \equiv \frac{\bar{w}_S(1, c^\star) + \omega^2/2}{\omega} \text{ and } k_{n^\star} \equiv \frac{\bar{w}_S(n^\star, c^\star) + \omega^2/2}{\omega}.$$

Because k_1 and $k_{n^\star}$ are known,

$$\frac{\exp(\omega k_1) - \exp(\omega\varepsilon)}{\exp(\omega k_1) - \exp(\omega k_{n^\star})} = \frac{1 - \exp(-\omega(k_1 - \varepsilon))}{1 - \exp(-\omega(k_1 - k_{n^\star}))} \tag{17}$$

is a known function of ω. Moreover, it is straightforward to verify that it is strictly increasing in ω for $\varepsilon \in (k_{n^\star}, k_1)$. Hence, the integrand in (16) is a known, strictly decreasing function of ω. Because the domain of integration of the integral in (16) is also known, this establishes that the integral itself is a known strictly decreasing function of ω, so that ω can be uniquely determined from the integral's known value.

With ω identified, we immediately recover φ, $\bar{w}_S = \log v_S$, $\bar{w}_E$, and v_E (and therewith a_S and a_E). The discount factor and per-period surplus function remain to be identified. For the discount factor, we can follow one of two approaches. First, we can assume that auxiliary information like the average borrowing rate for small businesses identifies ρ. Alternatively, we can follow Magnac and Thesmar (2002) and suppose that there exists an observable component of C that impacts next period's expected continuation values but *not* next period's expected surplus. More precisely, suppose that there exist two values $c_1 \neq c_2$ such that

$$\mathbb{E}_{a_E}[v_E(N'_E, C', W')|N' = n', C = c_1] \neq \mathbb{E}_{a_E}[v_E(N'_E, C', W')|N' = n', C = c_2]$$

but $\mathbb{E}[\pi(n', C')|C = c_1] = \mathbb{E}[\pi(n', C')|C = c_2]$. The former condition can be verified from data because v_E, a_E, G_C and G_W are identified, but the latter is an *a priori* exclusion restriction. Under this assumption, we can show that

$$\rho = \frac{v_S(n', c_1) - v_S(n', c_2)}{\mathbb{E}_{a_E}[v_E(N'_E, C', W')|N' = n', C = c_1] - \mathbb{E}_{a_E}[v_E(N'_E, C', W')|N' = n', C = c_2]}.$$

Of course, which of these approaches is most appropriate depends on the application at hand. In either case, given ρ we can recover $\mathbb{E}[\pi(n', C')|C = c]$ from the relevant Bellman equation.

We summarize these results in a theorem.

Theorem 2 *Suppose that ρ and κ are known and that G_W is specified up to scale as in (14). Furthermore, suppose that, for some $c^\star \in \mathcal{C}$ and $n^\star \in \{2, \ldots, \check{n}\}$,*

$$\Pr[N' = 0|N = 1, C = c^\star] = \cdots = \Pr[N' = 0|N = n^\star - 1, C = c^\star] < \Pr[N' = 0|N = n^\star, C = c^\star].$$

Then, the distribution of (N', C') *given* $(N, C) = (n, c)$ *for* $n \in \mathbb{N}_0$ *and* $c \in \mathcal{C}$ *uniquely determines* G_C*,* G_W*,* φ*, and* $\mathbb{E}[\pi(\cdot, C')|C = c]$ *for* $c \in \mathcal{C}$*.*

To emphasize that it can be verified in data, we have rewritten the condition that there exist $c^\star$ and $n^\star$ such that $\overline{w}_S(1, c^\star) = \cdots = \overline{w}_S(n^\star - 1, c^\star) > \overline{w}_S(n^\star, c^\star)$ in terms of known probabilities. The equivalence between the two sets of conditions follows from fact that the integral on the right-hand side of (15) has a positive integrand and so equals zero if and only if its limits of integration equal each other. That is, if and only if $w_S(n^\star, c^\star) = w_S(1, c^\star)$.

Theorem 2 only establishes identification of the expected surplus $\mathbb{E}[\pi(\cdot, C')|C = c]$, not of the surplus function π itself. This makes sense, because entry and exit decisions are taken after a period's surplus is earned and before next period's demand state C' is realized, so that observed market transitions only depend on π through the expected surplus. Nevertheless, in some applications, for example those involving counterfactual specifications of G_C, it may be useful to separately identify π. In these cases, π can be uniquely determined from the expected surplus provided that G_C satisfies a completeness condition of the type now routinely used in nonparametric identification analysis (see e.g. Newey and Powell, 2003).

We take three lessons away from this identification analysis. First, it is possible to identify the model's parameters without examining the cross-sectional relationship between N and C that Bresnahan and Reiss (1990, 1991b) use in their estimation. Second, estimation of our model need not follow the nested fixed point approach that we adopt. In the spirit of Hotz and Miller (1993), we could instead estimate the equilibrium value functions directly from observed transition probabilities and from these deduce the underlying primitives. Given the absence of firm-specific idiosyncratic shocks from our model, this deduction would substantially differ from that pioneered by Bajari, Benkard, and Levin (2007). We choose to estimate our primitives using NFXP maximum likelihood because this avoids the nonparametric specification of the decision rules that is required for a Hotz and Miller style estimator (which does not calculate the decision rules from primitives) and achieves statistical efficiency. Third, the use of nontrivial mixed strategies can identify the scale of the econometric error without imposing restrictions on players' payoffs. Identifying the analogous parameter in static discrete choice models always requires restricting the non-stochastic portion of payoffs in some way. Although we can impose similar restrictions on $\pi(n, c)$ (for example by requiring linearity in c),

these need not create straightforward restrictions on the equilibrium continuation values useful for extending static identification arguments to this dynamic setting. Nevertheless, we do not doubt the practical usefulness of such restrictions on payoffs for estimation with a finite sample.

E Monte Carlo Experiments

In this appendix, we investigate the statistical properties and computational performance of our estimation procedure with Monte Carlo experiments. For these, we set the maximum number of firms entering any market to $\check{n} = 5$, let the cost shocks be log-normally distributed with unit mean and scale parameter ω, normalize the mean per period fixed costs κ to one, fix the discount factor ρ at $\frac{1}{1.05}$, and interpret C_t as the number of consumers in the market. The statistical process governing C_t has support on 200 grid points that are equally spaced on the logarithmic scale. We use the same discretization procedure as in Section 5. We set $\mu = 0$ and $\sigma = 0.02$.

Each Monte Carlo experiment consists of 1,000 synthetic samples. We use four different sample sizes, each of them with ten time periods and between 100 and 1,000 ex ante identical markets. We compute the equilibrium and generate each sample market by simulating the evolution of (N, C), beginning with a draw from the model's ergodic distribution. We then use each sample to estimate the model's parameters with the three step procedure presented in Section 4. (Since this specification excludes variation in market characteristics, a single equilibrium calculation can support the likelihood function calculations for all of a sample's observations.) We use the mean and standard deviation of the log innovations of the generated C as starting values for the first step's likelihood function maximization. The starting parameter vector used for the second step equals a vector of ones multiplied by one random variable uniformly distributed on $[1, 10]$. Dubé, Fox, and Su (2012) caution that a nested fixed point algorithm can falsely converge when the tolerance criterion for the inner loop (which calculates the equilibrium) is set too loosely relative to that of the outer loop (which maximizes the likelihood function). We fix the convergence tolerance for the value function iteration at a value that is two orders of magnitude smaller than that for the likelihood maximization to avoid this potential pitfall.[31]

[31] We set the tolerance value to 10^{-10} for the inner loop and to 10^{-8} for the outer loop.

Table 5: Monte Carlo Results with Constant Surplus Per Consumer

	ř = 100	**ř = 250**	**ř = 500**	**ř = 1,000**
Averages of Estimates				
k	1.502	1.501	1.500	1.500
φ	10.219	10.064	10.054	10.017
ω	0.993	0.997	0.999	1.000
$\mu \times 10^2$	0.003	0.001	0.000	0.000
$\sigma \times 10^2$	1.999	2.001	2.001	2.001
Averages of Estimated Standard Errors				
k	0.050	0.031	0.022	0.015
φ	2.930	1.789	1.258	0.883
ω	0.070	0.044	0.031	0.022
$\mu \times 10^2$	0.068	0.043	0.030	0.021
$\sigma \times 10^2$	0.049	0.031	0.022	0.015
Monte Carlo Estimates of 95% Confidence Interval Coverage				
k	0.954	0.948	0.959	0.960
φ	0.930	0.932	0.944	0.938
ω	0.943	0.952	0.942	0.947
μ	0.948	0.944	0.951	0.949
σ	0.950	0.952	0.948	0.947

Note: Results of a Monte Carlo experiment using the three step NFXP estimator to estimate the model with one profit parameter (k) using 1,000 synthetic samples. The true value of k equals 1.5 and the true value of φ equals 10. The true value of the standard deviation of the log costs (ω) equals 1. Demand is discretized into 200 states. The demand process is governed by the drift parameter μ, which is set to zero, and the innovation standard deviation σ, which equals 0.02. The bottom-most panel displays the fraction of samples for which the estimated 95 percent confidence interval contained the parameter's true value.

We first simulate data from a model where the surplus function is parameterized as $\pi(c, n) = (c/n)k$ for $n \leq \check{n}$ and some fixed $k > 0$. This means that per consumer surplus is constant in the number of active firms n for $n \leq \check{n}$. We set the true values of k, φ, and ω, to 1.5, 10, and 1, respectively. The lower and upper bounds of the demand grid equal 0.5 and 5. Table 5 reports the corresponding Monte Carlo experiments' results. Its first panel gives the averages of the 1,000 estimates for each parameter, and it shows that the NFXP estimator is essentially without bias, even for the sample with only 100 markets. The second panel reports the averages of the estimated standard errors. For the sample with 100 markets, the average estimated standard error for the estimate of the sunk cost is 2.958. Therefore, we would expect a 95 percent confidence interval to approximately correspond to $(4, 16)$. This is possibly too wide for empirical usefulness, but the other estimates' standard

Table 6: Monte Carlo Results with Decreasing Surplus Per Consumer

	ř = 100	**ř = 250**	**ř = 500**	**ř = 1,000**
Averages of Estimates				
$k(1)$	1.813	1.806	1.802	1.802
$k(2)$	1.398	1.399	1.400	1.400
$k(3)$	1.196	1.198	1.198	1.200
$k(4)$	0.997	0.998	0.998	0.999
$k(5)$	0.892	0.899	0.899	0.898
φ	10.004	9.958	9.976	10.026
ω	0.986	0.993	0.997	1.000
$\mu \times 10^2$	0.003	0.002	0.000	-0.000
$\sigma \times 10^2$	2.000	2.000	2.000	2.001
Averages of Estimated Standard Errors				
$k(1)$	0.095	0.057	0.040	0.028
$k(2)$	0.094	0.058	0.041	0.029
$k(3)$	0.079	0.049	0.034	0.024
$k(4)$	0.075	0.047	0.033	0.023
$k(5)$	0.088	0.054	0.038	0.027
φ	3.433	2.082	1.457	1.031
ω	0.086	0.053	0.037	0.026
$\mu \times 10^2$	0.068	0.043	0.030	0.021
$\sigma \times 10^2$	0.049	0.031	0.022	0.015
Monte Carlo Estimates of 95% Confidence Interval Coverage				
$k(1)$	0.955	0.954	0.961	0.946
$k(2)$	0.945	0.942	0.948	0.939
$k(3)$	0.945	0.941	0.949	0.956
$k(4)$	0.951	0.945	0.943	0.950
$k(5)$	0.940	0.940	0.943	0.935
φ	0.884	0.901	0.926	0.944
ω	0.932	0.929	0.931	0.945
μ	0.941	0.951	0.965	0.956
σ	0.948	0.950	0.950	0.930

Note: Results of a Monte Carlo experiment using the three step NFXP estimator to estimate the model with five profit parameters $(k(1), k(2), ..., k(5))$ using 1,000 synthetic samples. The true value of $(k(1), k(2), ..., k(5))$ equals $(1.8, 1.4, 1.2, 1.0, 0.9)$ and the true value of φ equals 10. The true value of the standard deviation of the log costs (ω) equals 1. Demand is discretized into 200 states. The demand process is governed by the drift parameter μ, which is set to zero, and the innovation standard deviation σ, which equals 0.02. The bottom-most panel displays the fraction of samples for which the estimated 95 percent confidence interval contained the parameter's true value.

errors are relatively small. As expected, increasing the sample size decreases the standard errors approximately at the rate $\sqrt{\check{r}}$. So for $\check{r} = 500$ the standard error on φ is only 1.254. The table's final panel reports the Monte Carlo estimates of 95 percent confidence intervals' coverage probabilities. With the exception of those for $\check{r} = 100$, these are all within 2.0 probability points of their common nominal value. Apparently, the estimated standard errors provide accurate inference.

For our second set of simulations we use the same parameterization as before except that we define the flow surplus function as $\pi(c, n) = (c/n)k(n)$, where $(k(1), k(2), k(3), k(4), k(5))$ is set to $(1.8, 1.4, 1.2, 1.0, 0.9)$. This specification has the average surplus per consumer decrease in the number of active firms.[32] Table 6 reports the results of the corresponding Monte Carlo experiments. Again, all parameter estimates are essentially without bias, the estimated standard errors are small enough to be empirically useful, and the 95 percent confidence intervals have coverage probabilities close to their common nominal value. To check whether the estimator is able to distinguish a model with a decreasing per consumer surplus from a model with a constant surplus, we compute a likelihood ratio test for each sample. We can reject the null hypothesis $k(1) = \ldots = k(5)$ at the 95 percent confidence level in all of our synthetic samples regardless of the sample size. Overall, we conclude that the NFXP procedure has the potential to be empirically useful.

Since our equilibrium computation algorithm finds fixed points to relatively low dimensional contraction mappings, one would expect the estimation procedure to be relatively fast. Table 7 shows that this in fact is the case. On average, the computation of the three step maximum likelihood estimator takes between 20 and 30 seconds depending on the specification. This is much faster than the estimation time reported above for our application. The speed increase is entirely attributable to the absence of covariates, which requires computing an equilibrium for each market and each trial value of the parameters. One unexpected feature of Table 7 is that computation time can *decrease* as the number of markets grows. This happens because larger sample sizes smooth the objective function and thereby reduce the number of steps required by the optimization algorithm.

[32]These values generate a realistic distribution of firms per market. No firm is active in about 5 percent of the markets, a monopolist serves about 30 percent of the markets, and five firms serve about 5 percent of the markets. In contrast, the previous specification with constant per consumer surplus generates a distribution with an additional local mode at five firms (the value of $\check{n}$), because competition does not get "tougher" when the number of firms increases.

Table 7: Computational Performance

	$\check{r}=100$	$\check{r}=250$	$\check{r}=500$	$\check{r}=1,000$
one entry cost parameter, one profit parameter				
time per run (in seconds)	21.70	19.87	21.34	20.68
one entry cost parameter, five profit parameters				
time per run (in seconds)	31.04	27.21	29.64	31.07

Note: Average computational performance of the NFXP estimators in the synthetic samples. The estimator is implemented in C++ and runs as a single thread on an Intel Xeon E5472 at 3.00GHz.

Su and Judd's (2012) results suggest that we might be able to improve on the already rapid performance of our estimation procedure by using a mathematical programming with equilibrium constraints (MPEC) procedure in lieu of a nested fixed point algorithm. The MPEC estimator treats the value functions as a vector of nuisance parameters to be estimated subject to the equilibrium constraints implied by the sequence of Bellman equations. Thereby, MPEC omits the NFXP procedure's inner loop. Su and Judd compare the MPEC and NFXP procedures in a simulation study of Rust's (1987) bus engine renewal problem and conclude that MPEC is much faster. However, Iskhakov, Lee, Rust, Seo, and Schjerning (2015a) point out that a more efficient version of the NFXP procedure, as originally proposed by Rust, is as fast as the MPEC procedure. Because their results for Rust's single-agent renewal problem do not necessarily carry over to our game, it is instructive to compare our NFXP results to those from an application of the MPEC procedure.

We only use the MPEC method in the second step of the three-step procedure, since the first step is independent of the estimation procedure used in the second step and the third step essentially takes no time. Thus, if there are substantial differences between MPEC and the NFXP, these differences will be most visible in the second step of the three-step-procedure.

We implemented the MPEC estimator of our model in C++. We provided analytical derivatives for the objective function and the constraint Jacobian, and we explicitly accounted for the sparsity pattern of the constraint Jacobian. The Hessian was approximated using the Quasi-Newton BFGS method. The objective function was optimized using the commercial optimizer Knitro.

In contrast to Su and Judd, we found the MPEC estimator to be about 25 times

slower than the NFXP estimator. We obtained this result under very favorable starting values that fall within 10 percent of the truth. For these starting values, the MPEC estimator would always converge to estimates very close to those found by the NFXP estimator. We can only speculate as to why the MPEC estimator performs relatively poorly compared to the NFXP estimator. Su and Judd emphasize the usefulness of passing "sparsity patterns" to the optimizer, which indicate which derivatives of the constraints with respect to the nuisance parameters are identically zero. In their application to Rust's bus-engine replacement problem the constraint Jacobian is relatively sparse. Only 7 percent of its entries are non-zero. In our application, the constraint Jacobian is about 60 percent dense.[33] This is partly driven by the fact that the transition probability matrix for the demand state is fully dense. MPEC's relatively poor performance in our application appears to arise from the computation of the objective function's gradients with respect to the nuisance parameters, which requires repeatedly retrieving information from large and relatively dense matrices. These computational challenges might not be insurmountable, but our NFXP estimator seems to balance the costs of programmer time and execution time well.

[33] For the specification that corresponds to the Monte Carlo simulation reported in Table 6, there are $603,000$ nonzeros in the Jacobian (out of $1,007,000$ entries).