Empirical Models of Industry Dynamics with Endogenous Market Structure

Steven Berry[1] and Giovanni Compiani[2]

[1]Yale University, New Haven, CT, 06520; email: steven.berry@yale.edu.
[2]University of Chicago, Chicago, IL, 60637; email: Giovanni.Compiani@chicagobooth.edu.

Xxxx. Xxx. Xxx. Xxx. YYYY. AA:1–28

This article's doi:
10.1146/((please add article doi))

Keywords

dynamic models, endogenous market structure, autocorrelation, state dependence, two-step methods

Abstract

This article reviews recent developments in the study of industry dynamics, with a special emphasis on the econometric endogeneity of market structure. Endogeneity of market structure follows from the presence of serially correlated unobservable shocks to the profitability of firms' dynamic decisions, a feature common to many empirical settings. We particularly focus on extensions of standard two-step methods that leverage instrumental variables to address endogeneity, in both single-agent and oligopoly models.

Contents

1. Introduction 2
2. Single-Firm Dynamics 4
   2.1. Model Setup 4
   2.2. The Identification Problem 6
   2.3. Single-Firm Entry/Exit Example 6
   2.4. The Initial Conditions Problem 7
3. Approaches With No Serial Correlation in u 7
   3.1. Full-Solution MLE 7
   3.2. Two-Step Methods 8
4. Approaches With Serial Correlation in u 12
   4.1. Full-Solution MLE 12
   4.2. Mixture Models 12
   4.3. Introduction to IV Methods in the Single-Firm Case 14
   4.4. An IV Special Case 14
   4.5. A GIV First Step 14
   4.6. Point-Identifying the Policy Function in a Continuous IV Example 16
   4.7. Set-Identifying the Policy Function Using GIV in the Entry Example 16
   4.8. The Second Step with Serial Correlation in u 18
5. Oligopoly 19
   5.1. Full Computational Methods 19
   5.2. Two-Step Methods applied to Oligopoly 20
   5.3. GIV Methods in Oligopoly 21
6. Conclusion 23
7. Tables 28

1. Introduction

The field of Industrial Organization (IO) studies firms and markets in equilibrium. Many classic IO models are static. This is not surprising, as it can be difficult enough to model interactions between firms without modeling how firms and markets change over time. Yet, we know that industries and firms evolve. IO models often speak of "market structure," which is the broad category of market primitives that are held fixed in a static model of oligopoly price or quantity. These primitives include features like the number of firms, the cost and demand characteristics of those firms, and so forth. While market structure might be held fixed in the short run, it is clearly an economic outcome that is built up over time, in a dynamic setting.

We think that market structure is dynamic partly because we think that sunk costs may be important. In the presence of sunk costs, we understand that markets may exhibit *hysteresis*, a dependence on past market conditions, as in Dixit (1992). Furthermore, in the presence of sunk costs, decisions about, e.g., firm entry, investment, or product development depend on firm's forward-looking beliefs about future market conditions. All of this points the study of market structure towards explicitly dynamic settings.

There are, however, many interesting static empirical models of market structure, for example those reviewed in Berry & Reiss (2007) and Berry & Tamer (2007). These models may be an appropriate approximation to reality in the a case where market fundamentals are (relatively) unchanging and firms have settled into a clear "best-response" Nash equilibrium to rivals' behavior. However, even in a relatively unchanging market, a static model will not be able to distinguish sunk from fixed costs, as the distinction is entirely dynamic. This matters since many counterfactual policies may depend critically on the nature of sunk costs.

To state, even informally, rough conditions for a credible use of static market structure models is to make a case for dynamic models. A strong counterargument, though, is that dynamic models of market structure face an extremely difficult set of challenges. In the end, we may worry that the attempt to introduce dynamics creates so many compromises that the result is not better than the static version.

This article reviews the difficulties and trade-offs that applied empirical researchers face in estimating dynamic models of market structure. It covers a set of possible solutions, with an emphasis on competing approaches that differ in their computational tractability and their (relative) realism.

In particular, this review focuses on the econometric endogeneity of market structure that follows from the presence of serially correlated unobservable shocks to the profitability of firms' dynamic decisions. It seems clearly preferable to allow for these serially correlated shocks, but the history of the literature shows that this leads to challenges for both identification and estimation/computation. Possible modeling solutions that allow for realistic serial correlation are an active area of research and the primary subject of this review.

The topic of serial correlation in unobservables sounds "technical," but it involves issues of first-order importance to the modeling of industry dynamics. If market structure is built up over time, then it depends on the past profitability of the industry. Since the data is unlikely to capture all drivers of firm profitability, past profitability typically includes the effect of past unobservables. But if unobservables are correlated over time, this implies that current market structure is correlated with current unobservables, creating an endogeneity problem.

This logic is famously emphasized, for example, in Olley & Pakes (1996). They consider capital stock as an element of market structure, which is built up over time and which then shifts the short-run marginal cost curves of oligopolistic firms. In the estimation of a production function, they emphasize that the firm's capital stock will be correlated with the current-period unobserved productivity shock precisely because that shock is correlated with the past shocks that influenced past investment decisions, which in turn led to the current capital stock. Thus, the "market structure" of the short-run fixed capital stock is econometrically endogenous in the production function even if the capital stock is not directly determined by the current productivity shock. Importantly, ignoring this endogeneity will lead to misleading economic conclusions about, for example, the role of capital, labor and unobserved productivity in explaining output changes over time.

Igami & Yang (2016) provide another example, where "market structure" is the number of firms in a market. Their dataset consists of Canadian hamburger chain outlets ("stores") and they provide some simple descriptive evidence that points to the importance of market structure endogeneity. Table 1 is adapted from Table 3 in their paper.[1] We see that adding fixed effects (for market and firm) greatly changes the coefficients in these descriptive regressions. If we gave the coefficients a causal interpretation, the first column results (without fixed effects) would appear to say that, if anything, the presence of a rival firm increases the

[1]Standard errors are in parentheses and a set of further controls (population and income) are omitted from the table. Further details are available in the original paper.

probability that the own-firm will enter the market.[2]

The literature on empirical dynamic models is vast and earlier reviews include Ackerberg et al. (2007), Aguirregabiria & Mira (2010) and Arcidiacono & Ellickson (2011), among others. Here, we focus primarily on the issue of econometrically endogenous market structure. We first consider single-agent problems in Section 2. Many methods for single-agent models extend easily to multiple-agent settings and have the added benefit of simpler notation, so we use the single-agent case to illustrate the different approaches. We start from the methods that rule out serial correlation in the unobservables (Section 3) and then move on to those that do allow for serial correlation and thus account for the econometric endogeneity of market structure (Section 4). We emphasize two approaches to serial correlation and endogeneity. The first is a mixture model approach, following on Kasahara & Shimotsu (2009), that models persistent unobserved heterogeneity via a limited number of discrete types. The second is an instrumental variables (IV) approach, presented in a general context by Berry & Compiani (2020) and in a clever special case by Kalouptsidi, Scott & Souza-Rodrigues (2020). In Section 5, we turn to oligopoly models, with a special emphasis on the unique challenges that are introduced there, notably the issue of multiple equilibria. However, many of the ideas of the single-firm case carry over to the oligopoly setting so that most of our work is done by that point.

Finally, it is worth noting that while we focus our discussion on examples of endogenous market structure, the issue of serial correlation in dynamic models is much broader and thus our discussion could be applied to a much wider range of empirical settings.

2. Single-Firm Dynamics

We start by considering single-firm settings. These models may be directly applicable to situations on both extremes of competition: either firms that are "market takers" and ignore the behavior of their rivals (on one hand) or strict monopoly firms (on the other).

We first introduce some general notation and explicitly state the identification problem. We then consider a simple entry/exit example, which is helpful to illustrate the alternative approaches described in Sections 3 and 4. We conclude this section with a discussion of the initial conditions problem in models with serially correlated errors.

2.1. Model Setup

The model setup and notation closely follow Berry & Compiani (2020). We consider identification of a model that generates data on a large set of markets indexed by i. Since, in this section, there is a single firm per market, we will often use the expression "firm i" to refer to "the firm in market i." Within-sample time periods are denoted by $t = 1, ..., T$. Firm i's current market structure is x_{it} and in period t firm i chooses an action a_{it} out of the set of feasible actions, denoted $\mathcal{A}(x_{it})$. Examples of market structure x_{it} include a continuous measure of capital stock, an indicator of whether a firm is operating in a market, and the current quality level of a firm's product. Actions a_{it} associated with those example states might be (respectively) investment, entry/exit, and R&D expenditure. The single-period profits of firm i are given by

$$\pi\left(a_{it}, x_{it}, w_{it}, u_{it}; \theta_\pi\right), \tag{1}$$

[2] Indeed, this interpretation is offered in the Toivanen & Waterson (2005) study of UK data.

where w_{it} is a vector of exogenous profit shifters that are observed by both the firm and the researcher while u_{it} is an exogenous profit shifter that is observed by the firm but not by the researcher.

The law of motion for the unobservables is

$$\Phi(u_{it+1}|\lambda_{it};\theta_u), \tag{2}$$

with θ_u a vector of parameters that govern the distribution of u_{it}. Note that (2) implicitly assumes that the unobservables follow a first-order Markov process. The full vector of unknown parameters of the dynamic model is $\theta \equiv (\theta_\pi, \theta_u)$. The term λ_{it} includes various possible sources of serial correlation. One leading special case is a simple first-order autocorrelation process, where we have

$$\lambda_{it} \equiv u_{it}. \tag{3}$$

A second important special case is time-invariant discrete heterogeneity. Due to time-invariance, we can drop the t subscript and write

$$\lambda_i \in (\bar{\lambda}_1, \bar{\lambda}_2, \ldots, \bar{\lambda}_M). \tag{4}$$

In this notation, the $\bar{\lambda}_k$ are the possible discrete values of the persistent heterogeneity. The parameters θ_u then include the $\bar{\lambda}$ vector plus the probabilities of each of those discrete values.

The endogenous market structure evolves over time according to the transition probabilities

$$\Gamma(x_{it+1}|a_{it}, x_{it}, w_{it}). \tag{5}$$

Our examples will focus on special cases involving deterministic transitions that are specified by the model, but the framework allows for any transitions that can be directly estimated from the data. The exogenous states w_{it} are assumed to evolve according to the law of motion

$$\psi(w_{it+1}|w_{it}). \tag{6}$$

As is typical in the literature, we assume that both (5) and (6) are directly observed or known by the researcher.

The firm's dynamic problem is given by the Bellman equation:

$$\begin{gathered} V\left(x_{it}, w_{it}, u_{it}\right) = \\ \max_{a_{it} \in \mathcal{A}(x_{it})} \left(\pi\left(a_{it}, x_{it}, w_{it}, u_{it}; \theta_\pi\right) + \beta E\left[V\left(x_{it+1}, w_{it+1}, u_{it+1}\right) | a_{it}, x_{it}, w_{it}, u_{it}; \theta_u\right] \right). \end{gathered} \tag{7}$$

where β denotes the discount factor and V the value function. Following much of the literature, we assume throughout that the discount factor β is known.

The expected value function on the right-hand side of this expression is determined by the laws of motion of the different variables, i.e.

$$\begin{gathered} E\left[V\left(x_{it+1}, w_{it+1}, u_{it+1}\right) | a_{it}, x_{it}, w_{it}, u_{it}; \theta_u\right] = \\ \int\int\int V\left(x_{it+1}, w_{it+1}, u_{it+1}\right) d\Gamma(x_{it+1}|a_{it}, x_{it}, w_{it}) d\psi(w_{it+1}|w_{it}) d\Phi(u_{it+1}|\lambda_{it};\theta_u) \end{gathered} \tag{8}$$

Note that the expectation of the future value function in (7) and (8) depends on θ_u because that parameter governs the serial correlation of the unobservables, which influences future expected profits conditional on u_{it}.

Associated with the true Bellman equation is then the "policy function" that gives the optimal action for each state,

$$a_{it} = \sigma(x_{it}, w_{it}, u_{it}). \tag{9}$$

It is important to distinguish the true policy function, generated by the Bellman equation evaluated at the true value of the parameter, from the policy function that would result from the Bellman equation evaluated at arbitrary guesses for the parameter θ. We denote the policy function consistent with an arbitrary parameter θ as $\hat{\sigma}(x_{it}, w_{it}, u_{it}; \theta)$. Obviously, if the true parameter is θ_0, then

$$\sigma(x_{it}, w_{it}, u_{it}) = \hat{\sigma}(x_{it}, w_{it}, u_{it}; \theta_0). \tag{10}$$

2.2. The Identification Problem

For purposes of identification, we assume that we observe the true data generating process for the observable variables (a_i, x_i, w_i), but not for u. The underlying parameters to be identified are $\theta = (\theta_\pi, \theta_u)$.[3] It many applications, it will also be useful to think separately about the identification of the policy function, $\sigma(x_{it}, w_{it}, u_{it})$.

2.3. Single-Firm Entry/Exit Example

We now introduce a simple single firm entry/exit example which will illustrate different approaches in Sections 3 and 4. Consider a monopolist entry/exit example where the endogenous state $x_{it} \in \{0, 1\}$ indicates whether the firm was active in the market in the prior period and $a_{it} \in \{0, 1\}$ is the decision to operate in the current period. The single-period profit from being active in the market is

$$\pi(x, w, u) = \bar{\pi}(x, w) - u, \tag{11}$$

where $\bar{\pi}(x, w)$ is the "variable profit" of operations and the scalar u is a random fixed cost. The exogenous profits shifters are discrete, taking on one of K_w possible values. The sunk cost of entry, $\bar{\pi}(1, w) - \bar{\pi}(0, w)$, is here allowed to depend on the exogenous profit shifters w. In every period when the firm is inactive, it earns a single-period profit of zero. However, the firm retains the ability to re-enter the market. The value function is then

$$V(x, w, u) =$$

$$\max(\bar{\pi}(x, w) - u + \beta E_{w', u'}[V(1, w', u') | w, u; \theta_u], \beta E_{w', u'}[V(0, w', u') | w, u; \theta_u]), \tag{12}$$

Under well-understood conditions,[4] the value function is strictly decreasing in u and so the policy function $\sigma(x, w, u)$ involves a cut-off rule where the firm enters if and only if

$$u < \delta(x, w), \tag{13}$$

where $\delta(x, w)$ is the value of u that sets the expected dynamic return of being in the market equal to the value of being out. From (12), this is defined implicitly by

$$\delta(x, w) = \bar{\pi}(x, w) + \beta E_{w', u'} \left[V(1, w', u') - V(0, w', u') | w, u = \delta(x, w); \theta_u \right]. \tag{14}$$

[3]Because u enters the single-period profit function, there is a somewhat arbitrary distinction between the parameters of the single-period profit function, θ_π, and the parameters of the distribution of unobservables, θ_u. However, in many cases it is clear how to define θ_u so that it contains the parameters that govern serial correlation.

[4]See Stokey et al. (1989) and, for examples close to the present context, Bajari et al. (2007).

We denote the true cutoffs in the data as $\delta(x, w)$ and the cutoffs that result from computation of the firm's fixed point at an arbitrary parameter vector θ as $\hat{\delta}(x, w, \theta)$. In this special example and related cases with cutoff rules, $\delta(x, w)$ is the most general description of the policy function. Moreover, when x and w are discrete (as in our example), this involves a finite parametrization.

In the most general case, the unknown profit parameters, θ_π, are the $2K_w$ variable profit terms $\bar{\pi}(x, w)$. We assume that the normalized marginal density of u, $\phi_0(u)$, is known and that the unknown parameter θ_u controls the serial correlation of u.

2.4. The Initial Conditions Problem

When the data tracks each firm or market from the beginning of their potential life, the distribution of the first-period unobservables, u_{i1} can be considered an additional primitive of the model. However, if we first observe firms in the middle of their existence, serially correlated unobservables will likely be selected by past history. Specifically, the distribution of u_{i1} will not be equal to the unconditional marginal distribution of the unobservables. This creates a well-known "initial conditions problem," as discussed in many classic papers, including Heckman (1981), Chamberlain (1985), Blundell & Bond (1998), and Wooldridge (2005). These papers emphasize that structural parameters may not be identified without placing restrictive assumptions on the distribution of initial conditions.

Honoré & Tamer (2006) note that an alternative is to look for estimators that allow for unspecified initial conditions. In the context of dynamic panel data models, they show that leaving initial conditions unspecified may result in set-identified parameters. They also show that in many cases the identified set is quite small and thus useful for economic analysis. As discussed below, Berry & Compiani (2020) take a similar approach to initial conditions in the context of dynamic models of endogenous market structure.

3. Approaches With No Serial Correlation in u

This section discusses approaches to recovering the primitives of dynamic models under the assumption that the unobservables are not serially correlated. While, as discussed above, this restriction effectively amounts to assuming away the econometric endogeneity of market structure, it greatly simplifies the analysis and is thus maintained in the much of the empirical literature to date. Lessons from this literature prove to be very useful once serial correlation is introduced, and we will refer back to these lessons below.

While the focus throughout is primarily on model identification, we do discuss selected important computation and estimation issues that can influence the choice of methods.

3.1. Full-Solution MLE

Dating back at least to Rust (1987), one popular approach to dynamic models uses the structure of the Bellman equation to write the likelihood of the data as a function of the structural parameters. We illustrate with the simple entry/exit model of Section 2.3, providing a useful starting point for discussion.

The cutoff rule in (13) defines a set of intervals in $\mathbb{R}^T$ giving the set of $(u_{i1}, \ldots, u_{iT})$ values that are consistent with the data. For example, if firm i chooses $a_{it} = 1$ in period t, then we know that $u_{it} < \delta(x_{it}, w_{it})$, and if it chooses $a_{it} = 0$, then $u_{it} > \delta(x_{it}, w_{it})$. Without serial correlation, the two-period likelihood when the firm is not active in the first period but is active in the second period is

$$\mathcal{L}_i(\theta) = \int_{\hat{\delta}(x_{i1},w_{i1},\theta)}^{\infty} \int_{-\infty}^{\hat{\delta}(x_{i2},w_{i2},\theta)} \phi_0(u_2)\phi_0(u_1)du_2du_1. \tag{15}$$

Critically, it is possible to use the unconditional density ϕ_0 for the first-period unobservable u_1 in (15) only under the assumption of no serial correlation in the unobservables, unless we observe the firm or market from the beginning of its existence.

A full-computation MLE method proceeds by evaluating the likelihood function in (15) at trial values of the parameters θ, which in the general case requires computational techniques (such as value-function iteration) to solve for the value function and the policy cutoffs $\hat{\delta}(x,w,\theta)$. Rust (1987) and Rust (1994) refer to this method as a "nested fixed point," since the Bellman equation must be solved for each trial value of θ. Rust and later authors find computational short-cuts that apply to special cases, whereas Dubé et al. (2012) develop a different computational approach based on more modern advances.

3.2. Two-Step Methods

Motivated by a desire to avoid the computational burden inherent in full-solution methods, Hotz & Miller (1993) propose a two-step alternative that does not require solving the model for each candidate parameter value.[5] In the first step, the policy function is recovered from the data. For example, when actions are discrete, the policy function is identified from observed conditional choice probabilities, leading to the moniker of "CCP" methods. In the second step, the policies are combined with restrictions from Bellman's equation to recover the structural profit parameters.

When specialized to the entry/exit model of Section 2.3, the first step involves estimating the probabilities of entry for each value of (x,w), $p(x,w)$. This works because when u is not serially correlated, (x,w) are econometrically exogenous in the policy function equation (9) and thus $p(x,w)$ capture the true "causal" effect of (x,w). Specifically,

$$p(x,w) \equiv \Pr(u < \delta(x,w)) = \Phi_0(\delta(x,w)), \tag{16}$$

where Φ_0 is the CDF of u. Assuming that Φ_0 is strictly increasing, $\delta(x,w)$ is then recovered as

$$\delta(x,w) = \Phi_0^{-1}(p(x,w)). \tag{17}$$

Thus, in the example, knowing $p\,(x,w)$ is equivalent to knowing the entry cutoffs $\delta\,(x,w)$. This first-step identification of $\delta(x,w)$ depends entirely on the observed data and the assumed distribution for u, with no use of the dynamic model other than the existence of a cutoff rule. While our example involves a binary action, the original Hotz & Miller (1993) paper derives a vector equivalent of $\delta(x,w)$ by inverting the action probabilities in a multinomial discrete choice problem. The idea is the same. Further, the Hotz-Miller insight of uncovering the policy function in a first step can be extended to many other cases lacking serial correlation, including the wide range of discrete choice problems considered in Berry (1994) and Berry, Gandhi & Haile (2013).

[5] The broad idea of the two-step method is reviewed in many places, including Aguirregabiria & Mira (2010).

Bajari, Benkard & Levin (2007), henceforth "BBL," consider a case with continuous actions where the policy function takes the form

$$a = \sigma(x, w, u), \tag{18}$$

with a and u continuously distributed. Under appropriate assumptions, the methods of Stokey et al. (1989) can be used to establish the strict monotonicity of σ in u, so that the equation can be inverted to obtain

$$u = \sigma^{-1}(x, w, a) \tag{19}$$

This is a non-separable regression of the form in Matzkin (2003) and can be identified from the inverse distribution function of a conditional on (x, w). As in step one of Hotz-Miller, this then gives us the policy function directly from the data, without reference to the dynamic model. More complicated versions, with a mix of discrete and continuous variables are also possible. In each of these "extended CCP" examples, the policy function is point-identified in the first step without reference to Bellman's equation. This first step identifies the data-generating process without recovering the underlying structural parameters that necessary for many interesting counterfactuals.

The second step of a CCP-style method conditions on the policy function from the first step and imposes Bellman's equation to recover the single-period profit parameters. There are several alternative approaches to this step. In this review, we focus on the forward simulation method of Hotz, Miller, Sanders & Smith (1994) ("HMSS"). This method is broadly applicable to the class of models considered in the CCP literature and BBL also emphasize forward simulation. The approach is useful for our purposes because Berry & Compiani (2020) extend the idea to the case of serially correlated unobservables. We discuss that extension in Section 4.8.

To review the forward simulation procedure as applied to the entry/exit model, recall the cutoff defined in equation (14). Without serial correlation, we drop the conditioning on u, giving

$$\delta(x, w) = \bar{\pi}(x, w) + \beta E_{w', u'} \left[V(1, w', u') - V(0, w', u') | w \right]. \tag{20}$$

HMSS show how to use first-step policy functions, together with a guess for the profit parameters, to "forward simulate" the value functions in (20). Intuitively, starting from a state (x, w), draw u from its assumed distribution, use the known policy function to obtain the action $a = \sigma(x, w, u)$ and then assign the profit $\pi(a, x, w, u)$ to that action. The known state transitions then predict new states (x', w'), which are used to obtain next period's profits via the same steps, and so forth. The sum of discounted profits computed in this way can be used to construct an unbiased estimate of the value function and the average of many such simulations will provide a more precise estimate. Denote such a simulated value function by $\tilde{V}(x, w, u; \sigma, \theta)$.

Furthermore, HMSS note that if the single-period profit function is linear in a set of parameters, then the forward-simulated version of the expected value function will also be linear in those parameters. This yields a system of linear-in-parameters equations of the form (20), which we write as

$$\delta(x, w) = h_0(x, w; \sigma) + h_1(x, w; \sigma)\theta_\pi. \tag{21}$$

The δ on the left-hand side of this equation is known from the first CCP step and the h_0, h_1 functions on the right-hand side are known from the forward simulation, given the σ uncovered in the first step. We then have a set of linear equations (one for each (x, w)) in the unknown θ_π. The parameter is point identified if the equations have a unique solution in θ_π, which is easy to check. The argument in HMSS is applied to multinomial choice, as opposed to this binary example, but the logic is exactly the same.

BBL note that the HMSS examples is exposited only for the dynamic discrete-choice example. They propose a more general strategy of forward simulating the value function under alternative policies, $\sigma'(x, w, u)$. Since the true policy maximizes the value function, at the true θ_π it must be that for all possible policies $\sigma'(x, w, u)$,

$$\tilde{V}(x, w, u; \sigma, \theta_\pi) \geq \tilde{V}(x, w, u; \sigma', \theta_\pi), \tag{22}$$

This yields very many inequality constraints. A finite set of such constraints may not point-identify θ_π, so BBL consider set-identification of θ_π, even when σ is point-identified in the first step.

Berry & Compiani (2020) propose two alternatives to the BBL inequalities. Each generalizes to the case of serially correlated unobservables. The first, more general, approach relies on a single "policy function" iteration of the Bellman equation. The second approach, applicable in a very wide range of cases, generalizes HMSS to a broader class of problems while retaining computational simplicity. In this second case, we note that HMSS is implicitly using an "indifference" condition that applies to a much broader class of models. We now review each of the two approaches.

As a first alternative to the BBL inequalities, Berry & Compiani (2020) propose that a guess θ_π be rejected if a single policy iteration on the forward-simulated Bellman equation does not return the first-step $\sigma(x, w, u)$. In the entry/exit example this is

$$\tilde{\sigma}(x, w, u; \sigma, \theta_\pi) \equiv$$

$$1\left\{\bar{\pi}(x, w) - u + \beta E_{w',u'}[\tilde{V}(1, w', u'; \sigma, \theta_\pi) - \tilde{V}(0, w', u'; \sigma, \theta_\pi)|w] > 0\right\}, \tag{23}$$

where $1\{\cdot\}$ is the indicator function. We then exclude a candidate θ_π from the identified set if

$$\tilde{\sigma}(x, w, u; \sigma, \theta) \neq \sigma(x, w, u) \tag{24}$$

for any value of (x, w, u). This amounts to checking whether the firm's static "best response" to its future self playing σ is to also play σ. This is one iteration on the policy-function fixed point implied by Bellman's equation. Beyond the entry/exit example, the general method is to reject a given θ_π if it does not solve the policy-function problem in one interation, an idea general to any problem where Bellman's equation generates a unique policy function.

However, the policy-function iteration still requires some computational effort since it involves searching over candidate values of θ_π. As a second alternative to BBL style inequalities, Berry & Compiani (2020) show that the HMSS approach implicitly uses a set of indifference conditions in the unobservables. To see this in the simple entry/exit example, note that the expected discounted values of taking action a, denoted by $v(a, x, w, u)$, are

$$\begin{aligned} v(0, x, w, u) &= \beta E_{w',u'}\left[V(0, w', u')|w\right] \\ v(1, x, w, u) &= \bar{\pi}(x, w) - u + \beta E_{w',u'}\left[V(1, w', u')|w\right] \end{aligned}$$

At $u = \delta(x, w)$, these two equations imply equation (20). That is, setting $u = \delta(x, w)$ equates the values of being in and out of the market and this results in the HMSS style condition in (20). A similar indifference condition, across the action-specific values of all the choices, holds in the multinomial analysis of HMSS.

Berry & Compiani (2020) go further and show formally that policy functions in problems with discrete actions are generally defined by indifference conditions, as long as payoffs are continuous in the unobservables. Under mild conditions, for every pair of actions a and a' there is an unobservable $\tilde{u}(a, a', x, w)$ such that the firm is indifferent between actions a and a' when the firm is at the state $(x, w, \tilde{u}(a, a', x, w))$. That is, letting $\tilde{v}\left(a, x, w, \tilde{u}(a, a', x, w); \sigma, \theta_\pi\right)$ denote the forward-simulated version of the action-specific value functions, we have

$$\tilde{v}\left(a, x, w, \tilde{u}(a, a', x, w); \sigma, \theta_\pi\right) = \tilde{v}\left(a', x, w, \tilde{u}(a, a', x, w); \sigma, \theta_\pi\right). \tag{25}$$

If, as in the CCP literature, the first step of the identification procedure uniquely identifies σ, then we can treat σ as known when we get to the second step. The values $\tilde{v}(a, w, u; \sigma, \theta_\pi)$ can be forward simulated and they are linear in θ_π when the single-period profit function is linear in θ_π. In that case, then, for a given (a, a', x, w), (25) defines one linear equation in θ_π.[6] Berry & Compiani (2020) note that there will typically be at least one equation of the form of (25) for each combination of (a, a', x, w). In many discrete examples, this is a sufficient number of equations to potentially identify a θ_π of length equal to the number of distinct combinations of (a, x, w). In our entry/exit model where a, x and w each take on discrete values, this implies that we could consider the identification of a model with the most flexible ("natural") profit parameterization, i.e. one that treats the value of $\bar{\pi}$ at each combination of (x, w) as a separate parameter. Whether the implied equations actually invert is directly verifiable from a given data generating process.[7]

We can also consider continuous actions, or a mix of continuous and discrete actions. With continuous actions, the analog of the indifference conditions may be found in first-order conditions. Stokey et al. (1989) provide sufficient conditions for the differentiability of the value function, under which the optimal continuous actions satisfy

$$\frac{\partial \tilde{v}(a, x, w, u; \sigma, \theta_\pi)}{\partial a} = 0. \tag{26}$$

Note that the derivative $\partial \tilde{v} / \partial a$ can often be forward simulated and again will typically be linear in θ_π if the single-period profits are linear in θ_π. The first-order conditions then provide a large number (likely a continuum) of equations that restrict the values of θ_π. Again, the point-identification of θ_π via these conditions is verifiable. Berry & Compiani (2020) provide a particularly easy differentiable example based on a "stochastic accumulation" model.

In summary, then, our review of second-step CCP style methods is weighted toward ideas that extend to the case of serially correlated errors. The Berry-Compiani explication/extension of HMSS implies that the second step can be quite easy and that the inequality approach of BBL may be unnecessary in most cases, including examples with continuous actions. However, we do find that BBL's suggested use of forward simulation is quite useful and will extend nicely to the case of serial correlation.

[6] Note that we do not require that $\tilde{u}(a, a', x, w)$ be unique. Indeed, the original HMSS "indifference" conditions use a vector u at which the values of *all* actions, including the outside choice, are equal. There are other planes in the u space that equate the value of two actions, as in Ichimura & Thompson (1998). However, these are not necessary for identification in this example.

[7] Berry & Haile (2018) formally define "verifiable" as the identification of the binary truth or falsehood of the hypothesis that the given condition holds.

4. Approaches With Serial Correlation in u

After briefly covering two established approaches that allow for serial correlation—full-solution MLE and methods based on mixture models—we focus on the more recent generalized instrumental variable (GIV) approach in Berry & Compiani (2020).

4.1. Full-Solution MLE

It is possible to adapt the full-solution MLE approach described in Section 3.1 to the case with serial correlation in the unobservables. Again, unless one observes firms from the beginning of their existence, this requires modeling the dependence of the distribution of the first-period u_{i1} on (x_{i1}, w_{i1}). This conditional distribution then replaces the unconditional $\phi_0(u_1)$ in (15) and the distribution of u_2 is conditioned on u_1 and parameterized by θ_u. One approach to initial conditions is to flexibly parameterize the distribution of u_0 as a function of (x_{i1}, w_{i1}); another is to assume that it is equal to the stationary distribution generated by the model (see, e.g., Collard-Wexler (2014)).

To our knowledge, the identification properties of the fixed point MLE method are not well-explored in the general case with serial correlation. To build some intuition, consider our entry/exit example where the policy cutoffs δ, which enter the likelihood, depend on a limited amount of data. Specifically, the model implies that past data is excluded from these cutoffs. As a consequence, if the degree of hysteresis in the data cannot be entirely explained by the cutoffs, the likelihood method may find evidence of serial correlation. The methods below further clarify the role of exclusion restrictions and make this intuition more precise.

4.2. Mixture Models

The problem of serially correlated unobservables can be reframed as a problem of "unobserved heterogeneity." The challenge involves "controlling for" the persistent aspects of firms or markets that we don't see. One suggestion is to posit discrete unobserved heterogeneity, such as the time-invariant discrete heterogeneity in equation (4).

In labor economics, beginning at least with Heckman & Singer (1984), discrete heterogeneity is a popular approach to disentangling persistent heterogeneity from "state dependence." In our context, state dependence would follow, for example, from sunk costs that make a firm more likely to be active in a market if it was active in the prior period. Dynamic labor supply models often employ low-dimensional time-persistent discrete unobserved heterogeneity, as in Wolpin & Keane (1994) and a large related literature.

In an important contribution, Kasahara & Shimotsu (2009) discuss the identification of finite mixture models in the context of two-step methods. In our entry/exit example, we could specify the single-period profits from entry as

$$\pi(x, w, u) = \bar{\pi}(x, w, \lambda) - \epsilon, \tag{27}$$

where the unobservables are now $u = (\lambda, \epsilon)$. In the simplest case, λ would take two possible values, $\lambda \in \{0, 1\}$, that are time-invariant. The spirit of Kasahara & Shimotsu (2009) is that all of the time-persistent heterogeneity is in λ, so ϵ is assumed independent over time.

While Magnac & Thesmar (2002) obtain negative results for mixture approaches with two periods of data, Kasahara & Shimotsu (2009) consider the advantages of longer periods of data. One reason for the longer time series is to deal with the initial conditions problem. With discrete heterogeneity, fully flexible initial conditions add only a finite number of extra parameters. The additional restrictions coming from more periods of data can then achieve point-identification.

For some intuition, suppose we see the joint distribution of three periods of discrete data, $\Pr(a_3, x_3, a_2, x_2, a_1, x_1)$, where we suppressed the notation for w. According to a first-order Markov model, we should be able to predict this distribution exactly via the Markov representation. Say that $\tilde{p}(a_t, x_t|a_{t-1}, x_{t-1})$ is the first-order Markov transition function, constant across time. If correct, this model should fit the data for every two-period transition. One can also test longer and shorter sequences, constrained only by the length of the data. For example, for every observed data sequence we should have

$$\Pr(a_3, x_3, a_2, x_2, a_1, x_1) = \tilde{p}(a_3, x_3|a_2, x_2)\tilde{p}(a_2, x_2|a_1, x_1)p^*(a_1, x_1), \tag{28}$$

where $p^*(a_1, x_1)$ is an "initial condition." If the restrictions are rejected, there are two possible conclusions. First, the underlying data process may not actually be first-order Markov. Second, the apparent long dependence in the data might be explained by persistent hidden states. If these states are indexed by m, then there are hidden probabilities $\tilde{p}^m(a_t, x_t|a_{t-1}, x_{t-1})$ and hidden initial conditions $p^{*,m}(a_1, x_1)$ for types $m = (1, \ldots, M)$.

Kasahara & Shimotsu (2009) consider all the possible sequences and subsequences of the data and form all the possible restrictions. Variation in x and w will help greatly with identification. If d is the number of covariates and T the number of time periods, then Kasahara & Shimotsu (2009) show that there are on the order of d^T restrictions. With sufficiently long time series ($T \geq 3$) and sufficiently rich variation of the data, they show it is possible to use the restrictions to identify a limited number of different hidden types and, with even larger T, to identify more types and/or types that can change over time. The identification problem is, as usual, made more complicated by the initial conditions problem. As mentioned above, the discrete heterogeneity literature deals with this, first, by restricting the heterogeneity to depend on a small number of types and, second, by using longer periods of data.

That the number of types is limited by the time-periods and variability of the data is not surprising. A great advantage of the method, however, is that once the type probabilities are identified, all of the classic first and second step CCP approaches come into play. In terms of the first step, once we known the action (choice) probabilities conditional on λ, we can use them to identify the λ-specific policy functions. Because ϵ in (27) is independent over time, all of the classic CCP second step methods work as well. This includes not only the HMSS style forward simulation methods, but also the original second step method of Hotz & Miller (1993) as well as the "finite dependence" approaches that are well-summarized in Arcidiacono & Ellickson (2011).

To the degree that the empirical curse of dimensionality (i.e. the statistical problem of estimating many choice probabilities) is a problem for the original CCP models, it is an even larger problem for the multiple-type mixture model, as we are trying to identify a larger number of probabilities and are cutting the data into smaller bins to do so. To gain possible efficiencies, Arcidiacono & Miller (2011) develop an MLE approach.

We can see some similarities between the instrumental variables intuition and the mixture-model approach. In mixture models, the exclusion of sufficiently past history from the "causal" policy function is critical. Furthermore, there has to be sufficient variation in this excluded history. Finally, once we condition on the discrete heterogeneity, past history is "exogenous" in the sense that it is independent of the current unobservables. This combination of exclusion, variation and exogeneity is familar from IV methods. The next subsection will push this idea further.

4.3. Introduction to IV Methods in the Single-Firm Case

We now turn to formal instrumental variable methods introduced in this context by Berry & Compiani (2020) and, for an interesting special case, by Kalouptsidi, Scott & Souza-Rodrigues (2020). We next discuss a modified two-step method motivated by Berry & Compiani (2020). In the first step, identification of the policy functions is modified to use "Generalized Instrumental Variable" (GIV) methods, as discussed in Chesher & Rosen (2017) and elsewhere.[8] The GIV approach can handle both the initial conditions problem and endogenous market structure by leveraging instrumental variables and the structure of the model. The approach may result in point-identification of the policy functions, but it also allows for set-identification. In either case, the second-step forward simulation approach of subsection 3.2 carries over easily. When the policy function is set-identified, the second step is applied to each policy in the identified set. This results in an identified set for the single-period profit parameters θ_π.

4.4. An IV Special Case

Kalouptsidi, Scott & Souza-Rodrigues (2020) discuss the problem of endogenous states and propose an IV approach for a special case. They call their method an Euler conditional choice probability, or "ECCP", approach. In their model, there are many firms within each market and oligopoly behavior is assumed away. Serially correlated shocks are modeled at the market level and the form of the serial correlation can be quite general, in contrast to the mixture model approach. At the individual firm level, additive time-independent shocks allow for techniques to be adapted from the CCP literature, including from the "finite dependence" literature that starts with Hotz & Miller (1993) and is extended in Arcidiacono & Miller (2011) and elsewhere.

The model treats market-level terms as fixed effects that can be differenced out across firms within market and finite dependence creates a kind of multi-period indifference condition related to that in Section 3.2. The result is an equation that is linear in the parameters and is amenable to IV approaches. The paper provides a nice set of empirical examples with endogenous states (durable goods, land use, technology adoption and labor supply).

The ECCP method point-identifies firm-specific profit parameters, but not parameters on market-level effects. The authors note the potential complementarity between ECCP and GIV methods. Under the appropriate conditions, the ECCP approach could be used to identify some parameters, with remaining parameters identified (possibly set-identified) by GIV methods. We turn to those methods next.

4.5. A GIV First Step

The idea of the GIV first step is to set-identify policy functions from the data, using some structure from the model together with instrumental variables. An appealing feature of the

[8] That paper considers a broad class of models with nonseparable error structures, develops an approach explicitly based on the instrumental variables logic and provides a sharp characterization of the identified set. The results build on the work of Galichon & Henry (2011) and Beresteanu et al. (2011), while the broad approach to set-identification is informed by a vast literature that includes Manski & Tamer (2002), Tamer (2003), Manski (2003), Chernozhukov et al. (2007), Berry & Tamer (2007), Ciliberto & Tamer (2009), Beresteanu et al. (2011), Galichon & Henry (2011), Chesher (2010), and Andrews & Shi (2013).

GIV approach is that it accommodates features relevant to dynamic settings, notably discreteness of states and outcomes, set-identification, and incompleteness of the model, as discussed (for example) in Tamer (2003). In dynamic settings, incompleteness will often arise in the case of an unknown initial condition. In the absence of incompleteness, the GIV approach will often be equivalent to MLE.

To be useful, potential instrumental variables should be correlated with current-period endogenous states and yet excluded from the current period policy function and independent of u. One class of potential IVs in our model consists of past values of the exogenous w. In many specifications, past values of w do not enter the current period policy function and so are "excluded exogenous" variables, available as instruments as long as they shift current states (which is typically guaranteed by the dynamic nature of the model). Exogenous variables from the pre-sample period may be particularly useful in dealing with the initial conditions problem if they are correlated with the initial state. An example of such variables might be past demand shifters, such as market size, that are correlated with current market structure (conditional on current market size). Some such variables may be available from the pre-sample period even though the full set of variables is not.

More formally, the potential instruments are

$$z_i = (r_i, w_i), \tag{29}$$

where the vector r_i consists of information prior to the sample period. To motivate the econometric use of these instruments, we assume independence of the instrument and the unobservables:[9]

$$z_i \perp u_i.$$

Table 2, taken directly from Berry & Compiani (2020), gives some ideas of possible instruments in different contexts. As in all applied situations, the independence assumption may be better motivated in some examples than in others and, as with all IV methods, this discussion will be a key component of applied work. One advantage of GIV methods is that they bring this discussion to the forefront of the identification approach.

Given these IVs, we now sketch the use of GIV methods to set-identify the policy function. Let $\vec{a}_i \equiv [a_{i1}, \ldots, a_{iT}]$ and similarly for $\vec{x}_i, \vec{w}_i$ and $\vec{u}_i$. If the sequence $(\vec{a}_i, \vec{x}_i, \vec{w}_i)$ occurs, then $\vec{u}_i$ must be in the inverse image set

$$\mathcal{U}(\vec{a}_i, \vec{x}_i, \vec{w}_i, \sigma) = \{\vec{u}_i \colon \sigma(x_{it}, w_{it}, u_{it}) = a_{it}, \forall t\}.$$

Then, the Chesher-Rosen GIV conditions for identification are as follows. A pair $(\sigma(x_{it}, w_{it}, u_{it}), \theta_u)$ is in the identified set if and only if

$$\Pr(u_i \in \mathcal{S}; \theta_u) \geq \Pr\big(\mathcal{U}(a_i, x_i, w_i, \sigma) \subseteq \mathcal{S} \,|z\big) \tag{30}$$

for all closed sets $\mathcal{S}$ in the space of unobservables, and for all instrument values z. There are obviously very many "test sets" $\mathcal{S}$ that one could check. Chesher and Rosen show how to find the "core-determining" subset of these sets, i.e. the minimal collection of sets that one needs to check in order to characterize the sharp identified set. This collection includes all the "elemental sets," comprising the list of $\mathcal{U}(a_i, x_i, w_i, \sigma)$ across all the possible values of

[9] While we focus on this restriction throughout the paper, Chesher & Rosen (2017) show that the GIV approach may also be applied under weaker assumptions, such as mean or quantile independence.

actions and states. However, the core-determining set also includes the unions of partially overlapping elemental sets, excluding cases of strict subsets. We denote the resulting sharp identified set as

$$\Sigma^{IV}(\theta_u) \subseteq \mathcal{F}, \tag{31}$$

where $\mathcal{F}$ is the set of possible σ functions. The set $\mathcal{F}$ can be restricted to include, for example, only those σ functions that satisfy natural monotonicity restrictions grounded in the model. Note that $\Sigma(\theta_u)$ depends on θ_u since the left-hand side of (30) depends on the joint distribution of $\vec{u}_i$. Further, if $\Sigma(\theta_u)$ is the null set, then that value of θ_u is rejected by the data and the GIV conditions.

Before turning to the second step of the Berry-Compiani approach, we illustrate the first step via two examples: first, an extension of the continuous investment problem discussed in the context of (19), and, second, our single-firm entry/exit model. The first example might plausibly provide a point-identified policy function, while the second example seems likely to lead to set-identification.

4.6. Point-Identifying the Policy Function in a Continuous IV Example

Consider a continuous choice problem, such as an investment problem with convex costs of investment, that leads to a strictly positive investment level, a_{it}, in each period. Here, the state x_{it} is the current capital stock and w_{it} could be within-sample cost shifters. The unobservable could represent a shock to the profitability of investment. A formal version of this model is given in Olley & Pakes (1996). Under appropriate monotonicity conditions, we can invert the policy function as in (19) and write

$$u_{it} = \sigma^{-1}(x_{it}, w_{it}, a_{it}), \quad u_i \perp z_i. \tag{32}$$

This differs from a similar example in BBL only because we need to use an IV strategy to deal with the potential correlation of u and x. Luckily, equation (32) takes exactly the form of the quantile IV regression in Chernozhukov & Hansen (2005). That paper provides conditions for the point-identification of σ. Under those conditions, we have completed step one of the analog to the CCP two-step method. Further, note that equation (32) also yields identification of all u_{it}, which implies that its distribution, including the serial correlation parameter θ_u, is identified.

4.7. Set-Identifying the Policy Function Using GIV in the Entry Example

With discrete variables, it is less likely that IV conditions point-identify the policy function. Chesher (2010) considers set-identification of discrete-outcome models via instrumental variables. This subsumes the problem of recovering the policy function for our entry/exit example in the especially challenging case where we only see one period of data on (a_{i1}, x_{i1}) and we do not place any restrictions on the initial condition other than the availability of an exogenous instrument z_i that predicts x_{i1}.

We illustrate with our simple entry/exit model, for simplicity dropping variation in w. First consider the extreme example of data on just one transition: all we see for each firm is (a_{i1}, x_{i1}, z_i). The data give us the observed probabilities, $p(x_{i1}, z_{i1})$, of being active in the market but with serially correlated errors and an initial conditions problem these do not give the causal effects of x on entry. Therefore, we cannot invert these choice probabilities, as in (16), to find the cutoffs $\delta(x)$ characterizing the policy function.

As an alternative, drawing on the bounds estimation literature, Chesher (2010) works with the *necessary conditions* for actions, i.e. imposes the restriction that the probability of a necessary condition for an event be greater than or equal to the observed probability of that event. For instance, the necessary condition for $a_{i1} = 1$ is the cutoff rule $u_{i1} < \delta(x_{i1})$ and the necessary condition for $a_{i1} = 0$ is $u_{i1} > \delta(x_{i1})$. In this extreme case, then, we have four necessary conditions for the outcomes of the endogenous variables a_{i1} and x_{i1}. With sunk costs of entry, entry should be more likely when $x_{i1} = 1$ and so we expect that $\delta(1) > \delta(0)$. Given this monotonicity restriction, we can note that $u_{i1} < \delta(1)$ is a necessary condition not just for the event $(a_{i1}, x_{i1}) = (1, 1)$ but also for the event $(1, 0)$, i.e. when costs are low the firm is active no matter whether it was in or out last period. Similarly, $u_{i1} > \delta(0)$ is a necessary condition for both the event $(a_{i1}, x_{i1}) = (0, 1)$ and the event $(a_{i1}, x_{i1}) = (0, 0)$. For every value of z, this gives a set of straightforward bounds on the policy parameters $\delta(0)$ and $\delta(1)$. These bounds come from the model, the instruments and the entry probabilities:

$$\begin{aligned}
&\Pr(u_1 < \delta(1)) \geq \Pr(a = 1, x_1 = 1|z) + \Pr(a = 1, x_1 = 0|z) \\
&\Pr(u_1 < \delta(0)) \geq \Pr(a = 1, x_1 = 0|z) \\
&\Pr(u_1 > \delta(1)) \geq \Pr(a = 0, x_1 = 1|z) \\
&\Pr(u_1 > \delta(0)) \geq \Pr(a = 0, x_1 = 1|z)) + \Pr(a = 0, x_1 = 0|z).
\end{aligned}$$

Note that the probabilities on the left-hand side are not conditioned on z because u is independent of z by assumption. Even if there is only one value of z (i.e. there is no instrument), the structure of the model yields nontrivial upper and lower bounds. However, Chesher (2010) emphasizes that variation in the instrument is helpful because, e.g., some values of z might be predictive of $x_1 = 1$ and this will increase the conditional probabilities involving $x_1 = 1$, tightening those inequality constraints. Other values of z might predict $x_1 = 0$, increasing those probabilities. In the limit, if some value of z perfectly predicts $x_1 = 1$, then those bounds collapse to a point, possibly leading to point-identification of $\delta(1)$. If we also had variation in w, this could further tighten the bounds.

With only one period of data, there is no hope of learning about any parameter characterizing the serial correlation in the unobservables. With $T = 2$, however, we can make progress. Table 3 displays probabilities of necessary conditions associated with the eight combinations of (x_1, a_1, a_2) that are possible in our example.[10] In the first column are the probabilities of necessary conditions for the events, calculated via the bivariate distribution of u, which depends on θ_u. In the second column are probabilities of events in the data. At the true values of δ and θ_u, the probabilities in the first column must be greater than those in the second column. The inequalities based on Table 3 are special cases of (30) where the sets $\mathcal{S}$ are taken to be the elemental sets corresponding to the eight possible sequences (x_1, a_1, a_2). As mentioned above, characterizing the sharp identified set requires also considering unions of partially overlapping elemental sets. In our example, this would expand the number of restrictions from eight to a total of thirteen. Adding w back into the model would further increase the number of sets.

Note that, unlike the $T = 1$ case, the probabilities of the necessary conditions in Table 3 depend on θ_u. If, for example, u_{i1} and u_{i2} were perfectly correlated, the event $(x_1, a_1, a_2) = (1, 0, 1)$ would not be possible; similarly, a serial correlation parameter close to one would make that event unlikely. Thus, imposing inequalities based on two or more

[10] Similar information is displayed in two-dimensional graphs in Berry & Compiani (2020).

time periods places restrictions on θ_u and the number of restrictions increases in the number of time periods in the data. Berry & Compiani (2020) illustrate the advantages of more time periods, and better instruments, via computed examples.

The discussion so far has focused on (set-)identification of the policy functions and θ_u. In practice, with finite samples, one typically wants to go one step further and obtain confidence regions. Given that the model restrictions take the form (30), the large literature on moment inequalities provides approaches to conduct inference (e.g., Chernozhukov et al. (2007), Andrews & Soares (2010), Beresteanu et al. (2011), Galichon & Henry (2011), Andrews & Shi (2013), and Chernozhukov et al. (2013)). Within this literature, of particular importance are the papers that focus on the case where the number of inequalities is large relative to the sample size (e.g., Menzel (2014), Andrews & Shi (2017), and Chernozhukov et al. (2018)), since this scenario is likely to arise in the GIV framework, especially when the number of time periods in the data is large.

4.8. The Second Step with Serial Correlation in u

The Berry-Compiani first step results in an identified set for the policy functions—in our entry/exit example, the thresholds $\delta(x, w)$—plus the θ_u parameters. To map this into the space of θ_π, Berry & Compiani (2020) note that the identified set for the structural parameters is

$$\Theta^{ID} \equiv \{\theta = (\theta_\pi, \theta_u) \colon \hat{\sigma}(x_{it}, w_{it}, u_{it}; \theta) \in \Sigma^{IV}(\theta_u)\}, \tag{33}$$

where $\hat{\sigma}(\cdot, \cdot, \cdot; \theta)$ is again the policy function that results from Bellman's equation evaluated at θ. The identification condition says that the solution to the dynamic model at the parameter θ must satisfy the GIV conditions in the data. This defines the sharply identified set.

Given the set of policies identified by the GIV first step, the second step method of subsection 3.2 carries over easily. First, note that it is still trivial to forward simulate value functions. For the purposes of forward simulation, the serially correlated u are just like a serially correlated w. In addition, policy functions will still be typically defined by boundaries in u space leading to indifference conditions.

llustrating the second-step indifference equations for the entry/exit example with serial correlation, note that the action-specific value functions are now

$$\begin{aligned}
\tilde{v}(1, x, w, u; \sigma, \theta) &= \bar{\pi}(x, w) - u + \beta E_{w', u'}\left[\tilde{V}(1, w', u') | w, u; \sigma, \theta\right] \\
\tilde{v}(0, x, w, u; \sigma, \theta) &= \beta E_{w', u'}\left[\tilde{V}(0, w', u') | w, u; \sigma, \theta\right].
\end{aligned}$$

From equation (14),

$$\tilde{v}(1, x, w, u = \delta(x, w); \sigma, \theta) = \tilde{v}(0, x, w, u = \delta(x, w); \sigma, \theta). \tag{34}$$

These are the equations used in the Berry-Compiani second-step procedure. Once again, they will be linear in θ_π when $\bar{\pi}(x, w)$ is linear in θ_π.

Note that some other CCP style second-step methods, such as in the original Hotz & Miller (1993) paper, cannot be directly employed in this example of serially correlated unobservables. That is because these methods use "tricks" that are specific to models with additive independent errors. In particular, they do not account for the conditioning on u in future expectations, as in equation (14). However, it is possible to choose specifications (as in subsection 4.2) that include both a serially correlated unobserved component and an additive time-independent unobserved component. The original Hotz-Miller second step will work in this case.

In the case of serially correlated errors, the forward simulation will depend on θ_u as well as θ_π, as θ_u is necessary to simulate future values of u. If the GIV first step produces an identified set of (σ, θ_u) pairs, then the Berry-Compiani second step based on (34) needs to be applied to each (σ, θ_u) in the set. This second step produces an identified set for θ_π. If the GIV first step produces a point-identified (σ, θ_u), then the second step will similarly procedure a point-identified θ_π (as long as there is a unique solution in θ_π to the indifference conditions in (34)). Similarly, a confidence region for θ can be produced by applying the second step to each element in the confidence region for (σ, θ_u).

Note that if for some reason the second step forward-simulated indifference condition method fails, Berry and Compiani's first idea for the second step, outlined in (23) and (24), is still available.

5. Oligopoly

Moving from single-firm to oligopoly problems adds realism and greatly increases the scope for interesting policy counterfactuals. However, the dynamic estimation problem becomes more complicated as the full computation approach becomes a "doubly nested fixed point." Given rivals' strategies, each firm is solving a "best reply" Bellman fixed-point equation that defines its own behavior as a function of rivals' strategies. In a dynamic Nash (or Bayes-Nash) equilibrium, these strategies themselves must solve a second fixed point: the mapping between strategies and the dynamic best-replies to those strategies. This raises questions of both existence and uniqueness of equilibria that make full computational methods particularly difficult. As a result, much of the oligopoly literature eventually followed the single-agent literature into models without serial correlation.

A more recent approach has been to tackle serial correlation and endogenous market structure by combining the insights of the Hotz-Miller style oligopoly literature with the insights of the discrete heterogeneity literature and/or the GIV approach of Berry & Compiani (2020). We first discuss full computational methods and then turn to the more recent advances.

5.1. Full Computational Methods

One motive for the a full computation approach in market structure models has been precisely to account for serial correlation. The work of Ericson & Pakes (1995), Pakes & McGuire (1994) and Pakes & Ericson (1996) emphasized the idea of rich oligopoly models with a mix of discrete and continuous variables together with serially correlated unobservables. They suggest an empirical strategy of fitting the ergodic market structure distribution computed from long-run simulations on the computed model to observed transitions in the data. This deals with the initial conditions problem by assuming that the industry has settled into its long-run distribution of transitions. Gowrisankaran & Town (1997) provide one of the rare full empirical applications of this approach.

The Ericson & Pakes (1995) method faces problems of both existence and uniqueness of equilibria, as discussed in Doraszelski & Pakes (2007), Doraszelski & Satterthwaite (2010), and Pesendorfer & Schmidt-Dengler (2010). It is hard to guarantee uniqueness of equilibrium in the general case and it can be hard or impossible to find all of the equilibria that may exist

(Borkovsky et al. (2012)).

In a series of paper, Igami considers the problem of dynamic market structure estimation in the context of industries that are not in a stationary equilibrium, but rather in the process of rising and/or falling, as in Igami (2017), Igami (2018), and Igami & Uetake (2020). In the first two of those papers, Igami takes a full computation approach that ensures a unique equilibrium, which aids both estimation and countefactual analysis. Specifically, Igami ensures uniqueness by (a) assuming sequential moves (with either deterministic or random order) and (b) modeling a long but finite horizon. Under thee conditions, the oligopoly game can be uniquely solved backwards from the end. Igami assumes serially uncorrelated errors, and therefore has no initial conditions problem. However, in one case he also traces the industry from its birth, which would solve the initial conditions problem even in the presence of serially correlated unobservables. One interesting extension would be to apply the Igami sequential-move approach to the case with unknown initial conditions, either in a GIV or mixture-model framework.

However, in many cases the problems of multiple equilibria led the oligopoly literature back to two-step methods with serially uncorrelated errors, as we discuss next.

5.2. Two-Step Methods applied to Oligopoly

The similar starting point of several dynamic oligopoly papers is to assume that, even when the model admits multiple equilibria, the industry plays the same equilibrium every time it reaches the same state (Aguirregabiria & Mira (2007), Bajari et al. (2007), Pakes et al. (2007), and Pesendorfer & Schmidt-Dengler (2008)).[11] In addition, each paper assumes that the unobservables are [i] independent over time and [ii] pure private information. Under these assumptions, a firm can treat its rivals' behavior just like "plays of nature." That is, the evolution of rivals' behavior is just like the evolution of the exogenous w profit shifters in the single-firm case. Further, private information means that firms cannot take current period rival shocks into account and independence over time means that neither own-firm nor rival states are correlated with current-period unobservables. Thus, there is no endogeneity problem. Under these assumptions, then, the computational simplicity of the pre-existing CCP methods can be brought to oligopoly dynamics. This includes both first and second step methods.

In this literature, Pesendorfer & Schmidt-Dengler (2008) is the closest to Hotz-Miller, taking particular care to make the formal connection between the dynamic oligopoly model and the Hotz-Miller framework. Pakes et al. (2007) "flips" Hotz-Miller, arguing that it is the distribution of unobservables that should be identified from knowledge of the single-period return, instead of the other way round. They argue that the elements of single-period variable profits can in many cases be identified from "static" data on prices and quantities (and perhaps variable cost or input data), whereas fixed and sunk costs are only revealed by dynamic behavior. We have already discussed BBL.

Empirical applications of these methods to market structure include Ryan (2012) and Fowlie et al. (2016) on environmental policy, Holmes (2011) on the entry of Walmart, Collard-Wexler (2013) on responses to demand fluctuations, Aguirregabiria & Ho (2012) on airlines, Sweeting (2013) and Jeziorski (2014) on product positioning and entry in radio mergers,

[11] An early version of the idea is in Rust (1994) and a related insight (in the context of dynamic auctions) can be found in Jofre-Bonet & Pesendorfer (2003).

Dunne et al. (2013) on entry subsidies for health care providers. Many such papers make explicit use of the empirical strategy developed by Bajari et al. (2007). Dunne et al. (2013) follow the Pakes et al. (2007) suggestion of estimating the variable profit function prior to the dynamic estimation of sunk and fixed costs.

Given the same assumption of "one equilibrium in the data," the mixture model approach of section 4.2 also carries over to the oligopoly context. In this case, we would assume that persistent unobservables are known to the firms, but that the single-period profits are further shocked by an independent and private information term, as in (27). In that equation, the extension to oligopoly involves adding the rivals' states to a firm's own state in the x vector.

As noted in the introduction, Igami & Yang (2016) provide an empirical example of mixture models applied to entry in fast food markets. The empirical strategy in that paper follows the likelihood approach of Arcidiacono & Miller (2011). The results of Kasahara & Shimotsu (2009) are used to identify the minimum number of discrete profit levels that would explain the serial correlation in the empirical transition, and that number is used in the empirical work. As noted, the paper emphasizes the incorrect inferences that would result from entirely ignoring persistent heterogeneity.

In two-step applications to oligopoly, the curse of dimensionality can be particularly severe because the states of rival firms enter the own-firm state space. One approach that is applicable to cases with a large number of small firms (and perhaps a small number of large firms) is the "oblivious equilibrium" concept of Weintraub et al. (2008). Another strategy is to consider continuous time models, as in Doraszelski & Judd (2012). Arcidiacono et al. (2016) discuss an appropriate two-step estimation approach and provide an empirical application that considers the effect of entry by Walmart on existing competitors. They model perfectly persistent heterogeneity, for each type of store, via a mixture model method in the first step.

5.3. GIV Methods in Oligopoly

As with full computational and two-step methods, the work done in the single-agent case carries over to GIV methods applied to oligopoly. As in the single-agent case, we let i index markets and t index time. In addition, we introduce $j = 1, \ldots, J$ to index firms that coexist in a market.

Firm j's profits depend on its own action a_{ijt} as well as its rivals' actions. Thus, letting $a_{it} = \left(a_{i1t}, \ldots, a_{iJt}\right)$, firm j's profit is now

$$\pi_j \left(a_{it}, x_{it}, w_{it}, u_{it}; \theta_\pi\right)$$

In equilibrium, each firm's policy is the single-agent best reply to its rivals' equilibrium strategies. The firm still solves a value function problem similar to (7), but now its expectations of the future evolution of endogenous market states depend on its action as well as the equilibrium actions of its rivals.

In the oligopoly case, Berry & Compiani (2020) assume that the serially correlated unobservables are complete information to all of the firms. Private serially correlated information raises very difficult issues of signaling behavior, which would be a large additional complication. However, mixture models of the private information may help, as in Hodgson (2019).

Denoting the equilibrium policies of firm j's rivals by the function σ_{-j}, the firm expects the states to evolve in equilibrium according to transition probabilities of the form

$$\Gamma_j \left(x_{it+1} | a_{ijt}, x_{it}, w_{it}, \sigma_{-j}(x_{it}, w_{it}, u_{it})\right). \tag{35}$$

Thus, in equilibrium, the Bellman equation for firm j depends on the strategies played by its rivals, although we drop this dependence from the notation.

$$V_j\left(x_{it}, w_{it}, u_{it}\right) = \tag{36}$$

$$\max_{a_{ijt} \in \mathcal{A}(x_{ijt})} \left(\pi_j\left(a_{it}, x_{it}, w_{it}, u_{it}; \theta_\pi\right) + \beta E\left[V_j\left(x_{it+1}, w_{it+1}, u_{it+1}\right) | a_{it}, x_{it}, w_{it}, u_{it}; \theta_u \right] \right).$$

This dynamic program yields a "best response" strategy for firm j, which we assume is unique and denote by $\bar{\sigma}_j(\sigma_{-j}, \theta)$. We stack the best response function into a $J-$vector

$$\bar{\sigma}(\sigma, \theta) = (\bar{\sigma}_1(\sigma_{-1}, \theta), \ldots, \bar{\sigma}_J(\sigma_{-J}, \theta)).$$

Any equilibrium strategy, σ^*, then must satisfy the fixed point

$$\sigma^* = \bar{\sigma}(\sigma^*, \theta). \tag{37}$$

Given this, the set of possible equilibrium policy functions associated with a candidate parameter θ is given by

$$\Sigma^{EQ}(\theta) = \{\sigma^* : \sigma^* = \bar{\sigma}(\sigma^*, \theta)\}.$$

Following the discussion of earlier papers, we maintain the "one equilibrium in the data" assumption. The set $\Sigma^{EQ}(\theta_0)$, where θ_0 is the true parameter that generates our data, then contains the true policy function.

As in the single-agent case, we define the sharply identified set for the structural parameters as the set of values of θ that simultaneously satisfy the GIV restrictions and solve the equilibrium Bellman equation, i.e.

$$\Theta_{ID} \equiv \{\theta = (\theta_\pi, \theta_u) \colon \text{ there exists } \sigma^* \in \Sigma^{EQ}(\theta) \text{ such that } \sigma^* \in \Sigma^{IV}(\theta_u)\}. \tag{38}$$

In other words, a parameter vector θ belongs to the sharp identified set if there is a policy that both [i] is not rejected by the GIV restrictions and the data (given θ_u) and [ii] is an equilibrium strategy given θ.

Again, the first step consists of characterizing the set $\Sigma^{IV}(\theta_u)$ of policies that survive the GIV restrictions. However, this step will be complicated by a possibly large state space and by the presence of multiple firm unobservables in the policy functions. The large state space may lead to the use of parameterized and simplified policy functions, which is already common in existing CCP applications.

The first step can be illustrated through a simple extension of the Olley & Pakes (1996) style capital accumulation model of section 4.6 to the duopoly case. Denote the equilibrium policy functions of the two firms by

$$a_{ijt} = \sigma(x_{it}, w_{it}, u_{it}), \tag{39}$$

where the capital stocks are $x_{it} = (x_{i1t}, x_{i2t})$ and investments are $a_{it} = (a_{i1t}, a_{i2t})$. Similarly, (u_{i1t}, u_{i2t}) is the vector of serially correlated unobservables and w_{it} are exogenous shifters of the profitability of investment. Under the assumption that x_{ijt} and a_{ijt} are continuous variables and that the policy functions are continuous and injective in u_{it}, we can write

$$u_{ijt} = \sigma_j^{-1}(x_{it}, w_{it}, a_{it}), \tag{40}$$

This is now a two-equation version of the "quantile IV model" of Chernozhukov & Hansen (2005). Following on that paper, the policies may then be point-identified. Importantly, this would yield identification of each u_{ijt} and thus of its distribution, including the serial correlation parameter θ_u.

This example shows how multiple unobservables can naturally show up in the policy functions of oligopoly firms. In the case of discrete actions, this may pose particular problems that are yet to be fully explored in the literature. Once again, the problems might be dealt with in part by parsimoniously parameterizing the policy functions, while leaving the single-period profit functions as free as possible.

Up to (difficult) issues involving the dimensions of the observed and unobserved states, then, the GIV first step is the same in the oligopoly and single-firm cases.

The second step also follows through quite easily. Recall that Berry & Compiani (2020) proposed two approaches that accommodate serially correlated unobservables. Adapted to the oligopoly case, the first idea now amounts to calculating a best reply to (a) one's rivals' future behavior plus (b) one's own optimal future behavior. This is much easier than computing (i) the full best-response to rivals' behavior and especially (ii) the fixed-point of the dynamic oligopoly.

The "indifference" approach carries forward to the oligopoly case with even less modification. Recall again that the second step only employs the structure of the model and that whether a given variable was observed by us (or not) in the first step plays no role. Thus, in this step the states of rival firms, whether initially observed by us or not, simply become additional (x, w) terms in (25).

As an empirical oligopoly example, Collard-Wexler (2014) studies entry and exit in the concrete industry, modeling a parametric policy function and serially correlated market-level shocks. He considers a restrictive (although not unreasonable) initial condition assumption that allows him to point-identify and estimate the policy function parameters (as well as a serial correlation parameter) by MLE. His work is guided by an full-computation oligopoly framework found in Abbring & Campbell (2010). Berry & Compiani (2020) use a simplied version of the same data to illustrate how their approach allows one to drop the restrictive initial conditions and use a GIV first step. They also employ the linear indifference equations in (25), with different degrees of parametrization, to produce a confidence region for the single-period profit function that is valid given the set-identified policy functions. They show that the GIV method can easily reject the model with serially uncorrelated unobservables and that the presence of serial correlation greatly alters counterfactuals involving changes in the sunk cost of entry, as might be caused by changes in regulation.

This empirical application serves as a proof of concept for further empirical work. That work would ideally explore additional policy questions as it grapples with the issues of finding good instruments and dealing with the traditional problems of high-dimensional state spaces in dynamic modeling.

6. Conclusion

Two-step "CCP" methods without serially correlated errors helped the empirical analysis of theoretically endogenous market structure to overcome various problems with fully computed equilibrium oligopoly models. However, the initial gains in the literature came at the expense of econometrically exogenous market structure, with associated large likely biases in counterfactual predictions There are now at least two approaches to including serially correlated errors in such models: mixture model of discrete persistent heterogeneity and generalized instrumental variables methods that allow for general forms of serial correlation. At a practical level, GIV methods can allow for shorter time periods, unrestricted initial conditions, a mix of continuous and discrete actions, and different kinds of serial correlation. This comes at the cost of potentially set-identified parameters and counterfactuals. Theoretical concerns and existing empirical results show the importance of further developing this research agenda by applying and refining methods that allow for serial correlation in models of industry dynamics.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

Acknowledgements, general annotations, funding.

LITERATURE CITED

Abbring JH, Campbell JR. 2010. Last-in first-out oligopoly dynamics. *Econometrica* 78:1491–1527

Ackerberg D, Benkard L, Berry S, Pakes A. 2007. Econometric tools for analyzing market outcomes. In *Handbook of Econometrics*, eds. JJ Heckman, E Leamer, vol. 6A, chap. 63. North-Holland

Aguirregabiria V, Ho CY. 2012. A dynamic oligopoly game of the us airline industry: Estimation and policy experiments. *Journal of Econometrics* 168:156 – 173. The Econometrics of Auctions and Games

Aguirregabiria V, Mira P. 2007. Sequential estimation of dynamic discrete games. *Econometrica* 75:1–53

Aguirregabiria V, Mira P. 2010. Dynamic discrete choice structural models: A survey. *Journal of Econometrics* 156:38 – 67. Structural Models of Optimization Behavior in Labor, Aging, and Health

Andrews DW, Shi X. 2013. Inference based on conditional moment inequalities. *Econometrica* 81:609–666

Andrews DW, Shi X. 2017. Inference based on many conditional moment inequalities. *Journal of Econometrics* 196:275–287

Andrews DW, Soares G. 2010. Inference for parameters defined by moment inequalities using generalized moment selection. *Econometrica* 78:119–157

Arcidiacono P, Bayer P, Blevins JR, Ellickson PB. 2016. Estimation of Dynamic Discrete Choice Models in Continuous Time with an Application to Retail Competition. *The Review of Economic Studies* 83:889–931

Arcidiacono P, Ellickson PB. 2011. Practical methods for estimation of dynamic discrete choice models. *Annual Review of Economics* 3:363–394

Arcidiacono P, Miller R. 2011. Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity. *Econometrica* 7:1823–1868

Bajari P, Benkard CL, Levin J. 2007. Estimating dynamic models of imperfect competition. *Econometrica* 75:1331–1370

Beresteanu A, Molinari F, Molchanov I. 2011. Sharp identification regions in models with convex moment predictions. *Econometrica* 79:1785–1821

Berry S. 1994. Estimating discrete choice models of product differentiation. *RAND Journal of Economics* 23:242–262

Berry S, Compiani G. 2020. An instrumental variable approach to dynamic models. Tech. rep., Yale

Berry S, Tamer E. 2007. Identification in models of oligopoly entry. In *Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress*, eds. WN R. Blundell, T Persson, vol. 2. Cambridge University Press

Berry ST, Gandhi A, Haile PA. 2013. Connected substitutes and invertibility of demand. *Econometrica* 81:2087–2111

Berry ST, Haile PA. 2018. Nonparametric identification of simultaneous equations models with a residual index structure. *Econometrica* 86:289–315

Berry ST, Reiss P. 2007. Empirical models of entry and market structure. In *Handbook of Industrial Organization, vol. 3*, eds. M Armstrong, R Porter, chap. 29. North Holland

Blundell R, Bond S. 1998. Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics* 87:115–143

Borkovsky R, Doraszelski U, Kryukov Y. 2012. A dynamic quality ladder model with entry and exit: Exploring the equilibrium correspondence using the homotopy method. *Quantitative Marketing and Economics* 10:197

Chamberlain G. 1985. Heterogeneity, omitted variable bias, and duration dependence,in longitudinal analysis of labor market data. In *Longitudinal Analysis of Labor Market Data*, eds. JJ Heckman, B Singer. Cambridge UK: Cambridge Univ. Press

Chernozhukov V, Chetverikov D, Kato K. 2018. Inference on Causal and Structural Parameters using Many Moment Inequalities. *The Review of Economic Studies* 86:1867–1900

Chernozhukov V, Hansen C. 2005. An IV model of quantile treatment effects. *Econometrica* 73:245–261

Chernozhukov V, Hong H, Tamer E. 2007. Estimation and confidence regions for parameter sets in econometric models1. *Econometrica* 75:1243–1284

Chernozhukov V, Lee S, Rosen A. 2013. Intersection bounds: Estimation and inference. *Econometrica* 81:667–737

Chesher A. 2010. Instrumental variables models for discrete outcomes. *Econometrica* 78:575–601

Chesher A, Rosen A. 2017. Generalized instrumental variable models. *Econometrica* 83:959–989

Ciliberto F, Tamer E. 2009. Market structure and multiple equilibria in airline markets. *Econometrica* 77:1791–1828

Collard-Wexler A. 2013. Demand fluctuations in the ready-mix concrete industry. *Econometrica* 81:1003–1037

Collard-Wexler A. 2014. Mergers and sunk costs: An application to the ready-mix concrete industry. *American Economic Journal: Microeconomics* 6:407–447

Dixit A. 1992. Investment and hysteresis. *Journal of Economic Perspectives* 6:107–132

Doraszelski U, Judd KL. 2012. Avoiding the curse of dimensionality in dynamic stochastic games. *Quantitative Economics* 3:53–93

Doraszelski U, Pakes A. 2007. Chapter 30 a framework for applied dynamic analysis in io. In *Handbook of Industrial Organization*, eds. M Armstrong, R Porter, vol. 3. Elsevier, 1887 – 1966

Doraszelski U, Satterthwaite M. 2010. Computable markov-perfect industry dynamics. *The RAND Journal of Economics* 41:215–243

Dubé J, Fox JT, Su C. 2012. Improving the Numerical Performance of Static and Dynamic Aggregate Discrete Choice Random Coefficients Demand Estimation. *Econometrica* 80:2231–2267

Dunne T, Klimek SD, Roberts MJ, Xu DY. 2013. Entry, exit, and the determinants of market structure. *The RAND Journal of Economics* 44:462–487

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

Fowlie M, Reguant M, Ryan SP. 2016. Market-based emissions regulation and industry dynamics. *Journal of Political Economy* 124:249–302

Galichon A, Henry M. 2011. Set identification in models with multiple equilibria. *Review of Economic Studies* 78:1264–1298

Gowrisankaran G, Town RJ. 1997. Dynamic equilibrium in the hospital industry. *Journal of Economics & Management Strategy* 6:45–74

Heckman J. 1981. The incidental parameters problem and the problem of initial conditions in estimating a discrete time - discrete data stochastic process. In *Structural Analysis of Discrete Data with Econometric Applications*, eds. C Manski, D McFadden. Cambridge MA: MIT Press

Heckman J, Singer B. 1984. A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica* 52:271–320

Hodgson C. 2019. Information externalities, free riding, and optimal exploration in the uk oil industry. Tech. rep., Yale

Holmes TJ. 2011. The diffusion of wal-mart and economies of density. *Econometrica* 79:253–302

Honoré BE, Tamer E. 2006. Bounds on parameters in panel dynamic discrete choice models. *Econometrica* 74:611–629

Hotz J, Miller RA. 1993. Conditional choice probabilites and the estimation of dynamic models. *Review of Economic Studies* 60:497–529

Hotz VJ, Miller RA, Sanders S, Smith J. 1994. A simulation estimator for dynamic models of discrete choice. *The Review of Economic Studies* 61:pp. 265–289

Ichimura H, Thompson TS. 1998. Maximum likelihood estimation of a binary choice model with random coefficients of unknown distribution. *Journal of Econometrics* 86:269–95

Igami M. 2017. Estimating the innovator's dilemma: Structural analysis of creative destruction in the hard disk drive industry, 1981—1998. *Journal of Political Economy* 125:798–847

Igami M. 2018. Industry dynamics of offshoring: The case of hard disk drives. *American Economic Journal: Microeconomics* 10:67–101

Igami M, Uetake K. 2020. Mergers, innovation, and entry-exit dynamics: Consolidation of the hard disk drive industry, 1996–2015. *Review of Economic Studies* Forthcoming

Igami M, Yang N. 2016. Unobserved heterogeneity in dynamic games: Cannibalization and preemptive entry of hamburger chains in canada. *Quantitative Economics* 7:483–521

Jeziorski P. 2014. Estimation of cost efficiencies from mergers: application to us radio. *The RAND Journal of Economics* 45:816–846

Jofre-Bonet M, Pesendorfer M. 2003. Estimation of a dynamic auction game. *Econometrica* 71:1443–1489

Kalouptsidi M, Scott P, Souza-Rodrigues E. 2020. Linear IV regression estimators for single-agent dynamic discrete choice models. Tech. rep., Harvard. Forthcoming Journal of Econometrics

Kasahara H, Shimotsu K. 2009. Nonparametric identification of finite mixture models of dynamic discrete choices. *Econometrica* 77:pp. 135–175

Magnac T, Thesmar D. 2002. Identifying dynamic discrete decision processes. *Econometrica* 70:801–816

Manski CF. 2003. Partial identification of probability distributions. New York: Springer

Manski CF, Tamer E. 2002. Inference on regressions with interval data on a regressor or outcome. *Econometrica* 70:519–546

Matzkin RL. 2003. Nonparametric estimation of nonadditive random functions. *Econometrica* 71:1339–1375

Menzel K. 2014. Consistent estimation with many moment inequalities. *Journal of Econometrics* 182:329–350

Olley SG, Pakes A. 1996. The dynamics of productivity in the telecommunications equipment industry. *Econometrica* 64:1263–97

Pakes A, Ericson R. 1996. Empirical implications of alternative models of firm dynamics. *Journal of Economic Theory*

Pakes A, McGuire P. 1994. Computing markov perfect nash equilibrium: Numerical implications of a dynamic differentiated product model. *RAND Journal of Economics* :555–589

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

Pesendorfer M, Schmidt-Dengler P. 2008. Asymptotic least squares estimators for dynamic games1. *Review of Economic Studies* 75:901–928

Pesendorfer M, Schmidt-Dengler P. 2010. Sequential estimation of dynamic discrete games: A comment. *Econometrica* 78:833–842

Rust J. 1987. Optimal replacement of gmc bus engines: An empirical model of harold zurcher. *Econometrica* 55:999–1033

Rust J. 1994. Estimation of dynamic structural models, problems and prospects: Discrete decision processes. In *Advances in Econometrics*, ed. C Sims, chap. 4. Cambridge: Cambridge Univeristy, 119–170

Ryan SP. 2012. The costs of environmental regulation in a concentrated industry. *Econometrica* 80:1019–1061

Stokey NL, Lucas RE, Prescott EC. 1989. Recursive methods in economic dynamics. Cambridge MA: Harvard Univ.

Sweeting A. 2013. Dynamic product positioning in differentiated product markets: The effect of fees for musical performance rights on the commercial radio industry. *Econometrica* 81:1763–1803

Tamer E. 2003. Incomplete simultaneous discrete response model with multiple equilibria. *The Review of Economic Studies* 70:pp. 147–165

Toivanen O, Waterson M. 2005. Market structure and entry: Where's the beef? *The RAND Journal of Economics* 36:680–699

Weintraub GY, Benkard CL, Van Roy B. 2008. Markov perfect industry dynamics with many firms. *Econometrica* 76:1375–1411

Wolpin K, Keane M. 1994. The solution and estimation of discrete choice dynamic programming models by simulation and interpolation: Monte carlo evidence. *Review of Economics and Statistics* 76:648–672

Wooldridge JM. 2005. Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. *Journal of Applied Econometrics* 20:39–54

7. Tables

Table 1 Market Structure Coefficients with and without Fixed Effects

Ordered Probit Regressions of Entry/Exit
Igami and Yang (2016), Table 3

	Fixed Effects	
Profit Shifter	No	Yes
Own-store presence	-0.31	-0.78
	(0.02)	(0.03)
Rival-store presence	0.02	-0.23
	(0.01)	(0.02)

Table 2 Examples of Possible Instruments from Berry and Compiani (2020)

State	Example Instruments
Capital	Past investment cost
Out/In of Market	Past market population, past regulation
# of Stores	Distance from headquarters, interacted with time
Quality	Past R&D shocks, age of firm

Table 3 Probabilities of Necessary Conditions for Elemental Events with $T = 2$

Probability of Necessary Condition	Prob of Events in Data
$\Pr(u_1 < \delta(1), u_2 < \delta(1); \theta_u)$	$\Pr(1,1,1\|z) + \Pr(0,1,1\|z)$
$\Pr(u_1 < \delta(1), u_2 > \delta(1); \theta_u)$	$\Pr(1,1,0\|z) + \Pr(0,1,0\|z)$
$\Pr(u_1 > \delta(1), u_2 < \delta(0); \theta_u)$	$\Pr(1,0,1\|z)$
$\Pr(u_1 > \delta(1), u_2 > \delta(0); \theta_u)$	$\Pr(1,0,0\|z)$
$\Pr(u_1 < \delta(0), u_2 < \delta(1); \theta_u)$	$\Pr(0,1,1\|z)$
$\Pr(u_1 < \delta(0), u_2 > \delta(1); \theta_u)$	$\Pr(0,1,0\|z)$
$\Pr(u_1 > \delta(0), u_2 < \delta(0); \theta_u)$	$\Pr(0,0,1\|z) + \Pr(1,0,1\|z)$
$\Pr(u_1 > \delta(0), u_2 > \delta(0); \theta_u)$	$\Pr(0,0,0\|z) + \Pr(1,0,0\|z)$