An Instrumental Variable Approach to Dynamic Models

work in progress, comments welcome

Steven Berry and Giovanni Compiani
Yale University

September 26, 2017

Introduction

Empirical models of dynamic decision making play an important role in applied microeconomics. Much of the existing empirical literature models unobserved shocks as private information, independently distributed over time. This greatly simplifies the estimation methods, as outlined in a series of influential papers including Rust (1987), Hotz and Miller (1993) Bajari, Benkard and Levin (2007) (henceforth BBL), Pakes, Ostrovsky, and Berry (2007) and Pesendorfer and Schmidt-Dengler (2008).

We suggest an alternative approach, based on "generalized IV" methods, which maintains much of the structure of the earlier approaches, but emphasizes the econometric endogeneity of the dynamic states. We will focus on IO examples. In IO the endogenous states are often "market structure." We argue in favor of econometrically endogenous market structure.

Independent Private Shocks

Given the independent private information assumption, the current state does not reflect any (persistently) unobserved factors, only accumulated past luck. So, the current state is econometrically exogenous.

This leads naturally to 2-step Hotz-Miller style approaches:

1. first identify the dynamic policy function, "directly from the data" and then
2. use this plus Bellman's equation to identify structural parameters of the single-period return function.

For formal identification, see Magnac and Thesmar (2002) and related literature.

Problems with Exogenous States

But the treatment of states as econometrically exogenous may be problematic. It is greatly at odds with the focus, in associated static literatures, on the econometric endogeneity of market structure.

Think of possible states in an IO model: past entry, number of firms or outlets, capital, quality .... These are associated with various dynamic decisions: entry, store opening, investment in capital or quality. In each case, it seems likely that persistent unobservables are correlated with both the current decision and the current observed state.

This introduces a classic endogeneity issue into the identification/estimation of the dynamic policy function. We propose to deal with this via (generalized) IV methods.

Mixture / Panel Data Models

Since the endogeneity problem arises from the presence of unobservables, one approach is to control for these via "discrete types" model of unobserved heterogeneity. Keane and Wolpin (1997) is an early example.

Kasahara and Shimotsu (2009) consider finite mixture models. If we assume that all (true) transitions are first-order Markov, then longer dependence in the data is interpreted as due to unobservables. With discrete data, can identify some discrete heterogeneity, which can get more complicated as T increases (as long as the true dependence does not.) See application/estimation in Arcidiacono and Miller (2011), also Bonhomme, Lamadon and Manresa (2016).

IV vs. Mixture Models

We view our IV models as complementary to the finite mixture approach, as neither is strictly nested in the other.

As an approach to endogeneity, finite heterogeneity models have much in common dynamic panel data models, whereas our approach builds on classic IV intuition.

The Kasahara and Shimotsu (2009) approach requires $T > 3$, and then only for very simple heterogeneity. We can use $T = 2$, but only sometimes get point identification.

One could also blend the approaches.

An IV Alternative

At the most basic level, we adopt the classic two-step approach but identify (or at least restrict) the policy function via generalized instrumental variable ("GIV") methods following the exposition in Chesher and Rosen (2015).

Chesher Rosen '15 builds on a very large prior and on-going literature, including many early papers on incomplete models by Manski and/or Tamer, work on discrete entry models by Tamer and also work including Molchanov (2005), Beresteanu, Molinari, and Molchanov (2011) and Chesher, Rosen and Smolinski (2013).

Chesher Rosen '15 is useful to us both because of the generality of the results but also for the exposition that focuses on IV intuition in a broad class of models with discrete outcomes and possibly incomplete models.

IV Intuition

The IV intuition is that past exogenous variation will be correlated with current states. E.g. if Detroit was large and rich 30 years ago, it may have many Sears stores today. Past macro shocks may effect today's market structure, and these may interact with market-level characteristics. Past regulatory regimes may be correlated with market structure (consider hospitals).

Outline

Today

- Dynamic Model (examples)
- GIV restrictions
- Identified Set of Policies
- Worked Examples
  - monopoly entry / number of stores
  - duopoly number of stores (given enough time)
- Inference / Computed Example
- Beginning of a data example

Only IO examples but not obvious extensions to labor, etc.

Model

We begin with a single firm model. We see a large set of markets, each with a single firm, for a small fixed T.

Then we consider the oligopoly case.

Single Period Return

Monopoly Model

Panel data for a (fixed) T periods, with the periods denoted $t = 1, \ldots, T$. The full set of variables are not necessarily available for any prior history of the firm, $t < 1$, although there may be some partial history.

Single Period Profits for market i: $\pi\left(a_{it}, x_{it}, w_{it}, u_{it}; \theta_\pi\right)$

- x_{it} endogenously chosen state(s)
- a_{it} choice (policy) variable(s) (action)
- w_{it} exogenous profit shifter(s)
- $u_{it} \in \mathbb{R}$, unobserved (to us) serially correlated unobservable

Any additional exogenous variables correlated with x_{it} (i.e. policy prior to $t = 1$) are denoted r_i. Could extend (at cost) to multiple unobservables (esp. with multiple actions).

Further notation

$$
\begin{aligned}
x_i &= (x_{i1}, \dots, x_{iT}) \in \mathbb{X} \\
a_i &= (a_{i1}, \dots, a_{iT}) \in \mathbb{A} \\
w_i &= (w_{i1}, \dots, w_{iT}) \in \mathbb{W} \\
u_i &= (u_{i1}, \dots, u_{iT}) \in \mathbb{U}
\end{aligned}
$$

Transitions

Endogenous States:

$$\Gamma(x_{it+1}|a_{it}, x_{it})$$

Could be deterministic (degenerate) or stochastic state transitions.

Exogenous Observed States

$$\Lambda(w_{it+1}|w_{it})$$

Unobserved (by us) States

$$\Phi(u_{it+1}|u_{it}; \theta_u)$$

Assume Γ, Λ are known and/or identified from data, but θ_u is unknown.

Examples

Table: Some Single Agent IO Examples

State, x_{it}	Action, a_{it}	$\mathcal{A}(x_{it})$	Transition
Capital	Investment	$\mathbb{R}^+$	$x_{it+1} = \lambda x_{it} + a_{it}$
Entry	Out/In	$\{0,1\}$	$x_{it+1} = a_{it}$
Retail	# of Stores	$\mathcal{I}^+$	$x_{it+1} = a_{it}$
Quality	R&D	$\mathbb{R}^+$	$x_{it+1} \sim f(x_{it}, a_{it})$

Identification

We focus first on identification.

For purposes of identification, we assume that we observe the true data generating process, across firms or agents, denoted

$$P(a_i, x_i, w_i, r_i).$$

This is equivalent to seeing T period panel on a very large (in fact, infinite) cross-section of firms or agents.

The unknowns are the parameters of profits θ_π and θ_u. Nothing in our general discussion of identification requires these to be finite dimensional, but in practice we consider only finite-dimensional parametric models (including fully flexible profit functions with discrete (a_i, x_i, w_i)).

Dynamic Problem

Bellman Equation:

$$V\left(x_{it}, w_{it}, u_{it}\right) =$$

$$\max_{a_{it} \in \mathcal{A}(x_{it})} \left(\pi\left(a_{it}, x_{it}, w_{it}, u_{it}, \theta_{\pi}\right) + \delta E\left[V\left(x_{it+1}, w_{it+1}, u_{it+1}\right) | a_{it}, x_{it}, w_{it}, u_{it}\right]\right.$$

where

$$E\left[V\left(x', w', u'\right) | a, x, w, u\right] =$$

$$\int \int \int V\left(x', w', u'\right) d\Gamma(x'|a, x) d\Lambda(w'|w) d\Phi(u'|u; \theta_u).$$

Policy function

In the true model (and therefore the data), the policy function takes the form

$$a_{it} = \sigma\left(x_{it}, w_{it}, u_{it}\right) \in \mathcal{F}.$$

There may be qualitative restrictions (monotonicity) embodied in the set of functions $\mathcal{F}$.

For a particular $\theta = (\theta_\pi, \theta_u)$, the policy function derived from known transitions and Bellman's equation is:

$$a_{it} = \sigma_\theta(x_{it}, w_{it}, u_{it})\,.$$

Endogenous and Exogenous Variables

Exogenous variables (available as instruments)

$$z_i = (w_i, r_i).$$

Scalar (per period) unobservable: u_{it}. Key assumption: strict exogeneity

$$z_i \perp u_i.$$

Endogenous variables:

$$y_{it} = (a_{it}, x_{it})$$

$$y_i = (y_{it}, \ldots, y_{iT}).$$

Endogeneity

Example: Continuous Investment

In the dynamic model, x_{it} will be econometrically endogenous because of the serially correlated u_{it}.

For example, suppose x_{it} is capital, a_{it} in investment and (w_{it}, u_{it}) shocks to the profitability of investment. Past shocks will effect current capital via the capital accumulation process $\Gamma(x_{it}, a_{it})$. With serially correlated shocks, the correlation between u_{t-1} and x_{it} will typically imply a correlation between x_{it} and u_{it}.

Luckily, past exogenous shifters $w_{i\tau}$, $\tau < t$, will also be correlated with x_{it}. Intuitively, these should be available to serve as instruments to identify the dynamic policy function. Other past shocks to the investment process (regulations, field experiments, one-time disasters, etc) are also possible instruments.

Incomplete Model

$$
\begin{aligned}
a_{iT} &= \sigma\left(x_{iT}, w_{iT}, u_{iT}\right) \\
x_{it} &\sim \Gamma\left(x_{iT-1}, a_{iT-1}\right), \\
a_{T-1} &= \sigma\left(x_{iT-1}, w_{iT-1}, u_{T-1}\right) \\
x_{T-1} &\sim \Gamma\left(x_{iT-2}, a_{T-2}\right), \\
&\vdots \\
a_1 &= \sigma\left(x_{i1}, w_{i1}, u_{i1}\right).
\end{aligned}
$$

But: there is no model of the endogenous x_{i1}, which is inherited from unobserved prior history. If period 1 was the "birth" of the agent, we might have a model for x_{i1}, which would complete the single agent model. (Of course, this still doesn't guarantee identification.) Oligopoly will introduce a second source of incompleteness: multiple equilibria.

Chesher and Rosen GIV

Idea: (set) identify the policy function from classic instrumental variables conditions, extended to "Generalized IV" (GIV) to deal with

1. incomplete model (we only have necessary or sufficient conditions on the unobservables, not necessary and sufficient) and
2. discrete variables, as in entry/exit models
3. lack of point identification of the parameters, even in the absence of problems 1 and 2.

In the single-agent case, we may have all or none of these problems. Discrete actions naturally lead to conditions on sets of unobservables that give a particular policy a_{it}.

Chesher and Rosen Identified Set of Policy Functions

Building on many others, CR consider sets of u_{it} that are necessary for values of the endogenous observables.
Dropping exogenous covariates w_i, if the sequence (x_i, a_i) occurs, then u_i is in the set

$$\mathcal{U}(a_i, x_i, \sigma) = \{u_i : \sigma(x_{it}, u_{it}) = a_{it}, \forall t\}$$

The condition $\{u_i \in \mathcal{U}(a_i, x_i, \sigma)\}$ is then a necessary condition for event (x_i, a_i).

Policy Functions consistent with IV conditions

A pair $(\sigma(x_{it}, w_{it}, u_{it}), \theta_u)$ is then in the CR identified set iff for all closed sets $\mathcal{S} \in \mathbb{U}$ and $\forall z$

$$\Pr\big(\mathcal{U}(a_i, x_i, w_i, \sigma) \subseteq \mathcal{S} \;|z\big) \leq \Phi(\mathcal{S}; \theta_u) \tag{1}$$

For any S, the LHS is the probability, conditional on z_i, of the outcomes $y_i = (a_i, x_i)$ that have $\{u_i : u_i \in S\}$ as a necessary condition. The RHS is the probability of that necessary condition wrt the distribution of u_i. The LHS varies with the observed distribution of the data, with σ, and with z. The RHS is determined by the distribution of u_i, which by assumption does not depend on z.

Chesher and Rosen Identified Policy Functions

cont

In fact, CR show that to obtain the sharply identified θ set we only need to check sets $\mathcal{S} \in Q(\sigma, w_i)$, which for discrete data consists of the sets $\mathcal{U}(a_i, x_i, w_i, \sigma)$ plus certain intersections of those "elemental" sets.

The collection of sets $Q(\sigma, w_i)$ is the "core determining set" as defined in CR (2015) and earlier work, the minimal collection of closed sets $\mathcal{S} \in \mathbb{U}$ that yields the sharp identified set for θ. These include the overlapping sets of $\mathcal{U}(a_i, x_i, w_i, \sigma)$, excluding cases of strict subsets. For simple low dimensional discrete problems, $Q(\sigma, z_i)$ can be easy to compute and not "too big", but it can otherwise grow very (indeed infinitely) large.

This result tells us the sets we need to check to get sharp identification.

Equalities and Complete Models

In some cases, some of the inequalities in (1) are equalities, because the necessary conditions are *necessary and sufficient* for particular (sets of) actions.

In a complete model, all of the conditions would be equalities. However, as usual, this does not guarantee that the parameters are point identified.

Policies identified by GIV Alone

For a given θ_u, denote by

$$\Sigma^{IV}(\theta_u) \subseteq \mathcal{F}.$$

the set of σ functions identified from the data and the IV restrictions—*i.e.* those that satisfy condition (1) $\forall \mathcal{S} \in Q(\sigma, z)$ and $\forall z$. (Note when $\Sigma(\theta_u) = \varnothing$, for some θ_u, then that value of θ_u is excluded from the identified set without any use of the Bellman equation.)

Note that the sets Σ^{IV} are determined exclusively by the IV conditions and the data, with no use of the dynamic model. Making use of Bellman's equation will further shrink the set of identified policies (see discussion below).

Identified Set of Structural Parameters

For any $\theta = (\theta_\pi, \theta_u)$, we can use the Bellman equation to compute the implied policy $\sigma_\theta(x_{it}, w_{it}, u_{it})$. For $\theta = (\theta_\pi, \theta_u)$, this policy is

$$\sigma_\theta(x_{it}, w_{it}, u_{it}) \equiv$$

$$\underset{a_{it} \in \mathcal{A}(x_{it})}{\operatorname{argmax}} \Big(\pi(a_{it}, x_{it}, w_{it}, u_{it}) + \delta E\big[V(x_{it+1}, w_{it+1}, u_{it+1}) | a_{it}, x_{it}, w_{it}, u_{it} \big] \Big).$$

The sharply identified set of parameters is then

$$\Theta_{ID} \equiv \{\theta = (\theta_\pi, \theta_u) \colon \sigma_\theta(x_{it}, w_{it}, u_{it}) \in \Sigma^{IV}(\theta_u)\} \tag{2}$$

This imposes both the dynamic model and the GIV restrictions. This is the sharply identified set because any θ in this set generates a policy function that cannot be rejected by the data plus the IV condition.

Connection to Two Step Models

The prior exposition suggests that we could

1. identify a (set of) policy function(s) that are consistent with the data and the IV restrictions and then
2. see which structural parameters are consistent with the identified policy function(s). These are the identified parameters.

These are the same steps as in the 2-step dynamic literature with independent shocks (following on Hotz-Miller). The complications here are the presence of parameters for the unobservables (necessary at least to model the degree of serial correlation) and the IV methods in the first step, as opposed to directly "fitting the policy to data." The methods are the same when the distribution of u_t is known (or normalized) and the states are exogenous (independent errors.)

Connection to Two Step Models

Continued

Under appropriate assumptions, one might point identify the policy function in step 1. In this case, our procedure is just like BBL, except that one runs GIV first stage (vs assuming exogenous covariates.)

For example, when the choice variable is continuous, point identification of the policy is attained under the assumptions in Chernozhukov and Hansen (2005).

When the choice variables is discrete, often one only gets partial identification of the policy (Chesher (2010); Chesher, Rosen and Smolinski (2013)).

Connection to Two Step Models

Continued

However, as in the i.i.d. literature, there is no reason to proceed in 2 steps, in that order.

As a computational alternative, one could search over the space of θ's, for each possible θ

1. computing $\sigma_\theta(x_{it}, w_{it}, u_{it})$ via the contraction mapping and then
2. testing whether σ_θ survives the IV restrictions applied to the data. If so, that particular θ is in the identified set, otherwise not.

Example "First-Step" Policies Identified by GIV Alone

Binary Marginal Example *a la* Chesher (2010)

Consider the "binary-binary" model with $a \in (0,1)$ and $x \in (0,1)$.

Policy function: $a = \sigma(x,u)$, $u \sim \text{unif}(0,1)$.

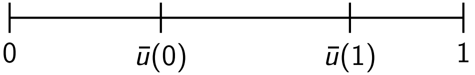


Figure: Policy Cut-offs in the Example

As an example of identification, we look at restrictions only involving the marginal distribution of u, next consider IV restrictions on the joint distribution across periods.

Table: Generalized Inverse Sets for the Binary-Binary Marginal Example

a	x	$\mathcal{U}(a,x)$
1	1	$(0, \bar{u}(1))$
1	0	$(0, \bar{u}(0))$
0	1	$(\bar{u}(1), 1)$
0	0	$(\bar{u}(0), 1)$

Table: Restrictions via Elemental Sets for the Binary-Binary Marginal Example

$\mathcal{S}$	$\Pr\big(\mathcal{U}(a_i, x_i, \sigma) \subseteq \mathcal{S} \mid z\big)$	$\leq \Phi(\mathcal{S}; \theta_u)$
$\mathcal{U}(1,1)$	$\Pr((1,1)\|z) + \Pr((1,0)\|z)$	$\leq \bar{u}(1)$
$\mathcal{U}(1,0)$	$\Pr((1,0)\|z)$	$\leq \bar{u}(0)$
$\mathcal{U}(0,1)$	$\Pr((0,1)\|z)$	$\leq 1 - \bar{u}(1)$
$\mathcal{U}(0,0)$	$\Pr((0,0)\|z) + \Pr((0,1)\|z)$	$\leq 1 - \bar{u}(0)$

GIV Restrictions in the Example Policy Function

From the restrictions in the table,

$$\Pr((1,1)|z) + \Pr((1,0)|z) \leq \bar{u}(1) \leq (1 - \Pr((0,1)|z)) \quad \Rightarrow$$

$$\max_z \left[\Pr((1,1)|z){+}\Pr((1,0)\ |z)\right] \leq \bar{u}(1), \quad \text{and}$$

$$\bar{u}(1) \leq \min_z \left[\Pr((1,1)|z){+}\Pr((1,0)\ |z){+}\Pr((0,0)|z)\right].$$

Similarly for $\bar{u}(0)$

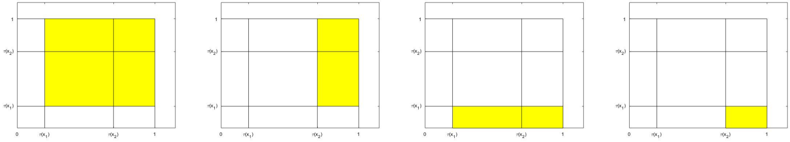
$$\Pr((1,0)|z) \leq \bar{u}(0) \leq (1 - \Pr(0,1|z) + \Pr((0,0)|z)) \quad \Rightarrow$$

$$\max_z \left[\Pr((1,0)|z)\right] \leq \bar{u}(0), \quad \text{and}$$

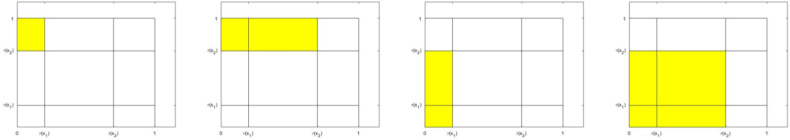
$$\bar{u}(0) \leq \min_z \left[\Pr((1,0)|z) + \Pr((1,1)|z)\right].$$

Next: elemental and core sets using the bivariate distribution of u, with data on 2 transitions.

Sets $\mathcal{U}(a_i, x_i, w_i, \sigma)$ for fixed w with $(x_{i1}, a_{i1}, a_{i2}) =$



(a) (0,0,0) (b) (1,0,0) (c) (0,0,1) (d) (1,0,1)



(e) (0,1,0) (f) (1,1,0) (g) (0,1,1) (h) (1,1,1)

Figure: Elemental Core Sets

Core Determining Sets

cont

The core determining set includes unions of the “elemental” sets $\mathcal{U}(a_i, x_i, w_i, \sigma)$. The relevant unions are of “partially overlapping” sets.

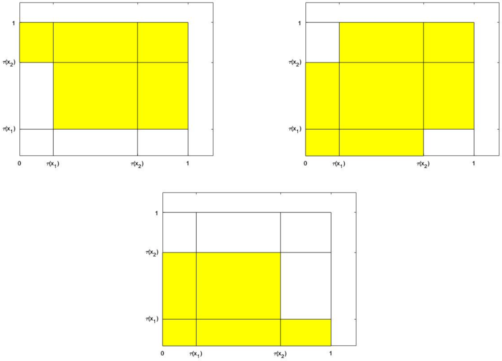


Figure: Unions of two $\mathcal{U}(a_i, x_i, w_i, \sigma)$

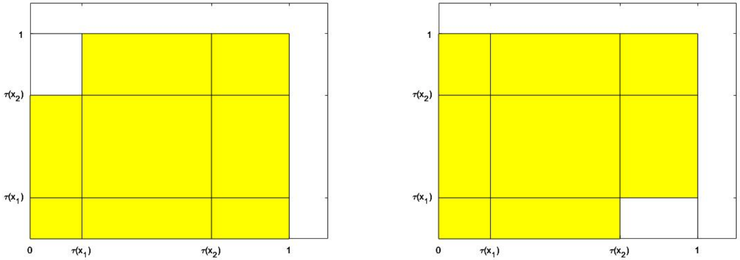


Figure: Additional Core Sets

Dynamic Games

more preliminary, just a sketch here

Our approach can be extended to models of dynamic strategic interaction. Specifically, consider dynamic games of complete information and allow for serially correlated unobservables. Let J be the number of players. Then agent j's policy in market i at time t is given by

$$a_{jit} = \sigma_j \left(w_{it}, x_{it}, u_{it} \right) \tag{3}$$

with $w_{it} = (w_{1it}, \ldots, w_{Jit})$, etc.

Policies now depend on the unobservables of all players.

With serial correlation, complete information is much easier than private information, but we can add iid private information on top of serially correlated public info, which makes computation (and existence) easier.

Broad Idea of the Games Approach

As in the single agent case, identification comes from combining the GIV restrictions with the Bellman equation that defines the "best response" of each firm to the actions of the other.

Computationally, it is often very difficult to solve for the set of equilibrium strategies. In this case, we can restrict that computation to those strategies that survive the GIV conditions. In a favorable case, this would be a small set.

Additional assumptions may simplify the task, e.g. ensuring the game is symmetric, that strategies are monotonic in some arguments, etc.

Computation: One Retail Firm Choosing Number of Stores

Consider the problem faced by a monopolist who needs to choose how many stores to open in a given market. For identification, we observe the joint distribution of observables in an panel of independent markets, with $T \geq 2$.

- $x_{it} \in \{0, 1, 2\}$ is the number of open stores,
- w_{it} is market size
- r_i (an excluded instrument) is market size at some point in the farther distant past.
- the action is tomorrow's number of stores: $x_{it+1} = a_{it}$

Computation: One Retail Firm Choosing Number of Stores

cont

- there is an unobserved (to us) fixed cost of $\beta\epsilon_{it}$ per each open store.
- ϵ is serially correlated: $\epsilon_{it} = e^{u_{it}}$, with

$$u_{it} = \rho u_{t-1} + \sigma \nu_{it}$$

- u_{i1} and the innovations ν_{it} are iid $N(0, 1)$.
- opening a new store requires a sunk cost of γ.

Flow Profit in the Example

Flow profit is variable profit minus the fixed costs of stores minus the sunk cost of any entry:

$$\pi_{it} = \alpha w_{it} \cdot \sqrt{x_{it}} - \beta \epsilon_{it} x_{it} - \gamma \left(a_{it} - x_{it} \right) \mathbb{I} \left\{ a_{it} > x_{it} \right\}$$

We set units via $\alpha = 1$ and so $\theta_\pi = (\beta, \gamma)$, $\theta_u = (\sigma, \rho)$.

Alternatively (and preferably) variable profit has been estimated from "price and quantity" data, outside of the dynamic model.

Policies and States in the Example

With $T = 2$, the policies and state transitions are

$$a_{i1} = \sigma(x_{i1}, w_{i1}, u_{i1})$$
$$a_{i2} = \sigma(x_{i2}, w_{i2}, u_{i2})$$

with $x_{i2} = a_{i1}$. The endogenous variables are $y_i = (a_{i1}, a_{i1}, x_{i1})$. We have two equations, but only one function to learn (stationarity), and we make a *joint* independence assumption that (u_{i1}, u_{i2}) is independent of an instrument z_i, which we think of as some observed past exogenous shock that influences x_{i1}.

Policies

We assume conditions on the dynamic model such that σ is increasing in x and decreasing in u (u is a cost).

This gives policies that are cut-offs in u_{it}. For example, for outcome $a_{it} = 1$,

$$a_{it} = 1 \iff \bar{u}_2(x_{it}, w_{it}) < u_{it} \leq \bar{u}_1(x_{it}, w_{it})$$

with $\bar{u}_1$ and $\bar{u}_2$ increasing in x_{it}. This makes it easy to characterize the policies σ, which for discrete (x_i, w_i) are (without parametric restriction) just a discrete list of cutoffs. The sets $\mathcal{U}(a_i, x_i, w_i, \sigma)$ are also easy to describe.

See Chesher (2010) & Chesher and Rosen (2013) for the binary threshold crossing model. We have a multi-time period version, so we can impose *joint* independence.

Results Under Marginal Independence from the Instrument

Monte Carlos

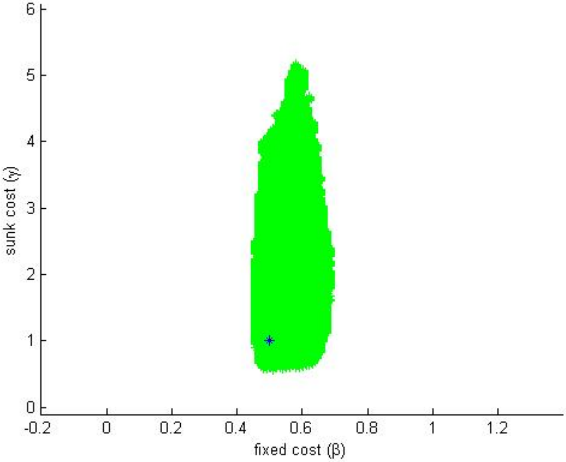


Figure: Projection of identified set for $\theta_\pi = (\beta, \gamma)$: marginal independence

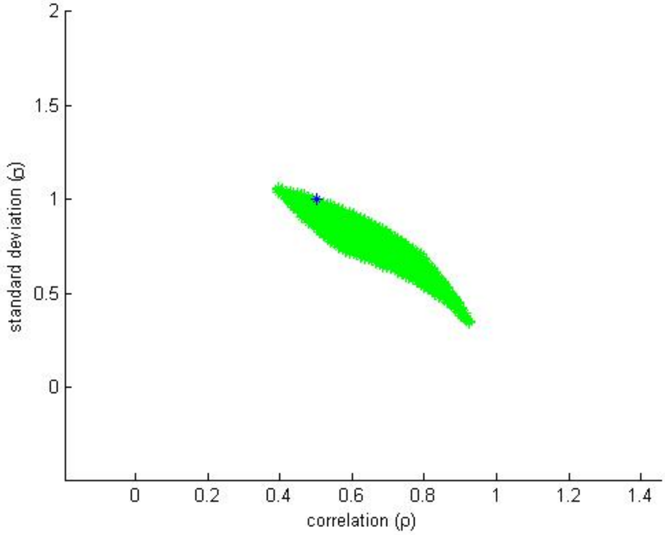


Figure: Projection of identified set for $\theta_u = (\rho, \sigma)$: marginal independence

Results Under Joint Independence

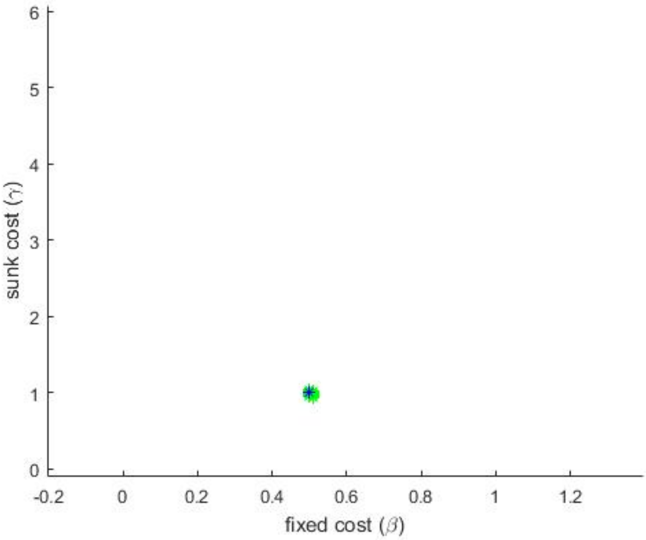


Figure: Projection of identified set for $\theta_\pi = (\beta, \gamma)$: joint independence

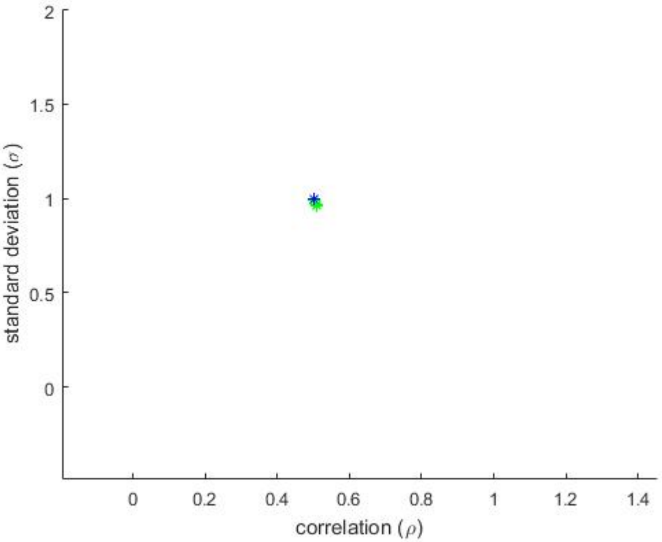


Figure: Projection of identified set for $\theta_u = (\rho, \sigma)$: joint independence

Results under Incorrect Assumption of Exogeneity

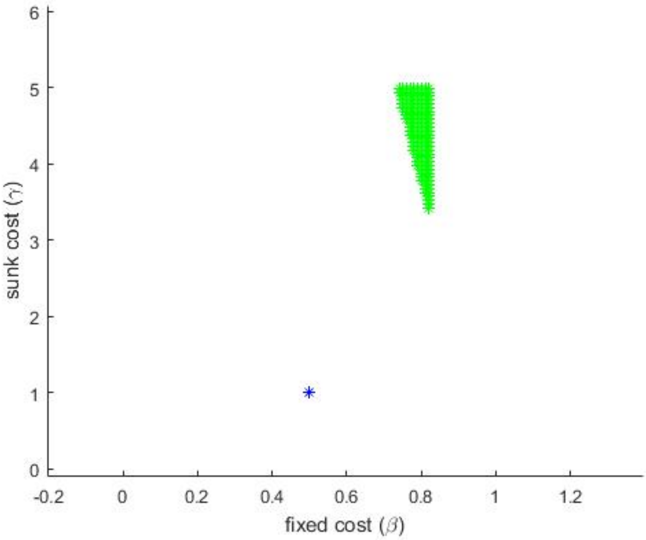


Figure: Projection of identified set for $\theta_\pi = (\beta, \gamma)$: exogeneity

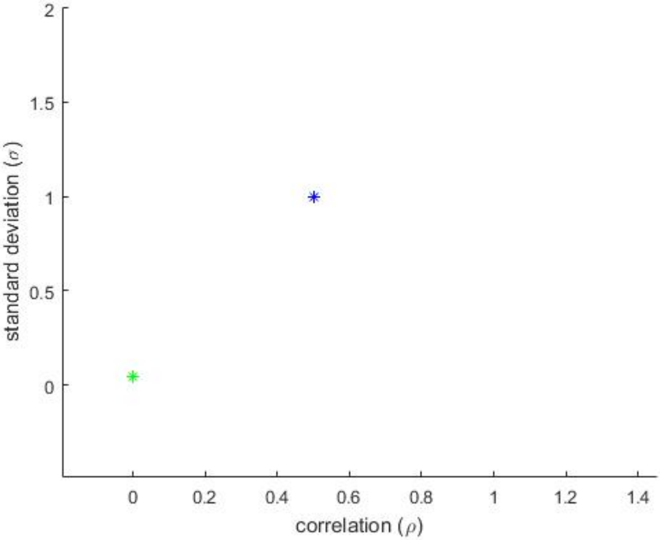


Figure: Projection of identified set for $\theta_u = (\rho, \sigma)$: exogeneity

Inference

We can apply inference procedures from the moment inequalities literature.

With discrete IVs, we work with unconditional moment inequalities (e.g. Andrews and Soares (2010) and Bugni, Canay, and Shi (2017)).

With continuous IVs, we have conditional moment inequalities (e.g. Andrews and Shi (2013)).

When object of interest is of lower dimension than the structural parameters, one can avoid grid search over the entire parameter space. The price to pay is that one needs to solve highly nonlinear optimization problems, but this has become common in the literature.

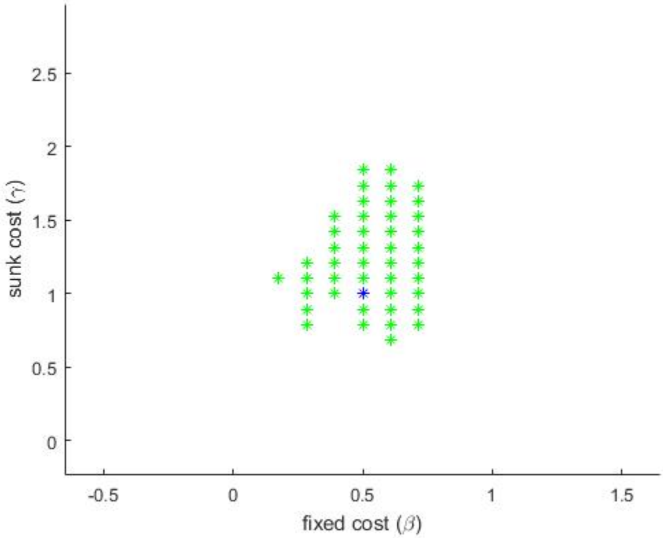


Figure: Projection of 95% confidence set for $\theta_\pi = (\beta, \gamma)$

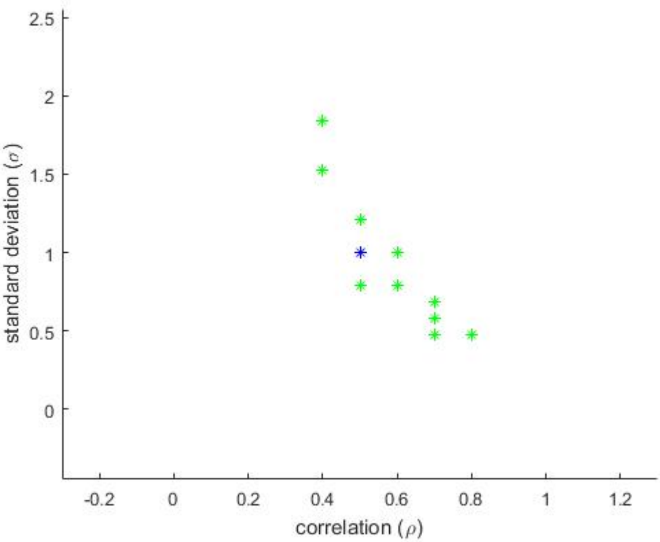


Figure: Projection of 95% confidence set for $\theta_u = (\rho, \sigma)$

Illustration: dynamic duopoly

Consider the problem faced by two firms choosing how many stores to open in each of many markets over time. Given the number of stores in a market, the two firms engage in Bertrand competition and each of them charges the same price across its own (within market) stores. Let j index stores and $k\left(j\right)$ be the firm owning store j.

Consumer i shopping at store j gets utility

$$u_{ij} = \overline{\delta} - \alpha p_{k(j)} + \epsilon_{ij},$$

where ϵ_{ij} is EV-distributed.

Market Shares and Prices in the Illustration

The market share of firm k at time t is given by

$$s_{kt}\left(p_t, x_t\right) = \frac{x_{kt} \exp^{\overline{\delta} - \alpha p_{kt}}}{1 + \sum_{r=1}^{2} x_{rt} \exp^{\overline{\delta} - \alpha p_{rt}}},$$

where $p_t = (p_{1t}, p_{2t})$, $x_t = (x_{1t}, x_{2t})$, and x_{kt} denotes the number of stores that firm k has open at time t.

The first order condition for firm k's static prices at time t then determines the equilibrium prices in each market at any point in time and this in turn determines each firm's variable profits.

Store Entry

Each firm incurs a fixed cost for each open store at time t and a sunk cost if it decides to open a new store in the next period. Again, $a_{kt} = x_{kt+1}$ for all k. Flow profits are

$$\pi_{kt} = \tilde{\pi}_k \left(x_{it}, w_{it}\right) - \beta x_{it} u_{kt} - \gamma \left(a_{kt} - x_{kt}\right) \mathbb{I}\left\{a_{kt} \geq x_{kt}\right\} \cdot \eta_{kt},$$

where β is the fixed cost, γ is the sunk cost and η_{kt} and u_{kt} are shocks that are unobserved to the econometrician. We assume that u_{kt} is common knowledge and possibly correlated over time and across players, and therefore correlated with x_{it}. For simplicity in the illustration, η_{kt} is assumed to be private information, iid over time and across players. This shock ensures existence of a dynamic equilibrium given the discrete u_{kt}.

For simplicity in a preliminary illustrative, we let each u_{kt} take on only two values, so that u_{it} takes on four values. We simulate a model with a high degree of serial correlation, $u_{it+1} = u_{it}$ with probability $\rho = 0.7$ and switches to one of the other 3 values with 0.1 probability each.

As in the single agent case, we use past values of w as instruments. The instrument for firm 1 has a correlation of 0.81 with x_{1t} and of 0.56 with x_{2t}.

Identified Policies

The structural parameters are the fixed cost β, the sunk cost γ and the correlation parameter ρ . First, we find all the policies that satisfy the GIV conditions. In the example, there are only 64 possible states and therefore a policy is a vector of length 64 with elements equal to the number of stores (1 or 2). We enumerate all possible (monotonic) policies and check the GIV conditions for each of them. Only 5335 policies survive.

Identified Parameters

In the second step, we go from the identified set for the policy to the identified set for the structural parameters. Specifically, for each candidate value θ in a grid, we check whether there exists a policy in the GIV identified set such that when the opponent plays according to that policy, it is optimal for a firm to play the same policy in a symmetric equilibrium. As in the existing literature with only i.i.d. shocks, we do not rely on uniqueness of the (theoretical) equilibria.

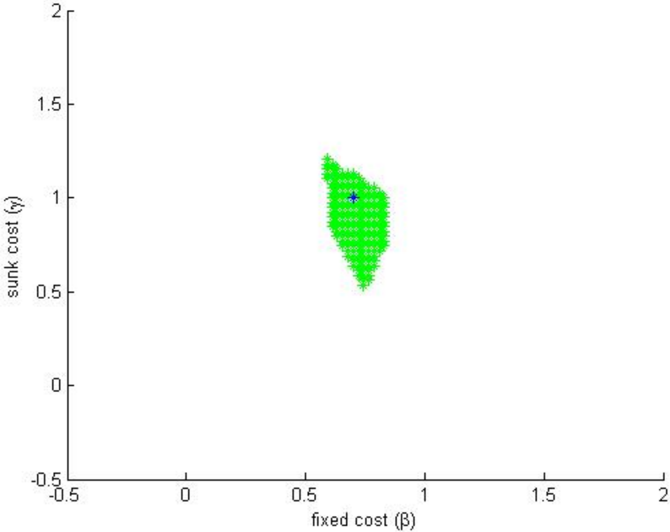


Figure: Projection of identified set for (β, γ)

A Hint at An Application

- Begin with the small-town ready-mix concrete example of Collard-Wexler (2014), which is based on public data. The number of cement firms is shifted by "construction demand," proxied by local construction employment.
- In the spirit of Pakes, Ostrovsky, and Berry (2007), we estimate "variable profit" as a function of the number of firms from (limited) public data, outside of the dynamic model.
- This leaves us with the four "dynamic parameters" of our single-agent example (sunk cost & mean fixed cost, plus the standard deviation & serial correlation of the unobservable).
- We use the Abbring and Campbell (2010) model of the number of firms in dynamic oligopoly, with their "refinement" to unique equilibrium. This makes the problem more like the single-agent case.

Collard-Wexler on Concrete

Collard-Wexler data are for the years 1994 to 2006. He assumes serially correlated errors and effectively estimates the policy-function of the dynamic model, which involve thresholds in the unobservables. He solves the incompleteness problem due to initial conditions by directly modeling the initial unobservables and there is no multiple equilibria problem under the Abbring-Campbell "last in-first out" conditions. He doesn't estimate structural parameters, focusing just on the implied evolution of states over time.

Our VERY Preliminary Work

We extend this to an incomplete model (with an initial conditions problem) and attempt to set-estimate the parameters. For now, we use a very (!) coarse grid and we are using a somewhat ad-hoc set of test sets. There are a number of econometric issues to consider, including whether we should adjust for a large number of moments.

For instruments, we are using long-past interactions of income & population.

Test Sets

There are too many test sets, so we use some simple sets involving, *e.g.*, two-period behavior. We also use some sets that intuitively help to identify (say) sunk cost. For example, the set associated with the event "there is at least one entry in the market over the sample period."

Test Sets
simulation

For some test sets, the probabilities are hard to compute, so use a simulation procedure to obtain test sets. For a given observable event and a given candidate policy function, we

- take many draws for the unobservables
- obtain the equilibrium number of firms over time implied by the candidate policy function
- compute the fraction of markets where the event $\mathcal{E}$ occurs

This gives us the probability on the r.h.s. of the Chesher-Rosen inequality. On the l.h.s. we simply have the probability of the event $\mathcal{E}$, conditional on values of the instruments.

Very Preliminary Results

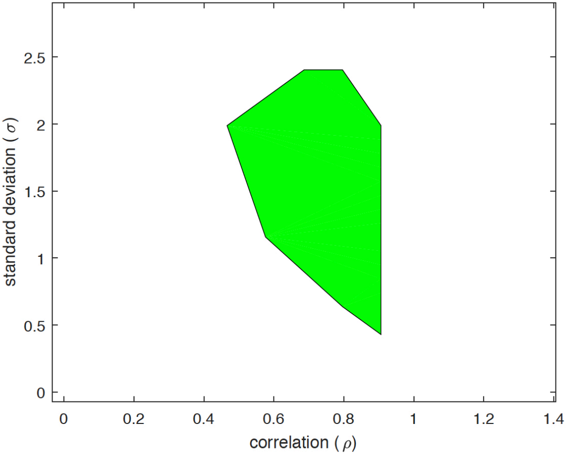


Figure: Concrete Data: Prelim Projection of the Estimated Set for (ρ, σ)

Very Preliminary Results
continued

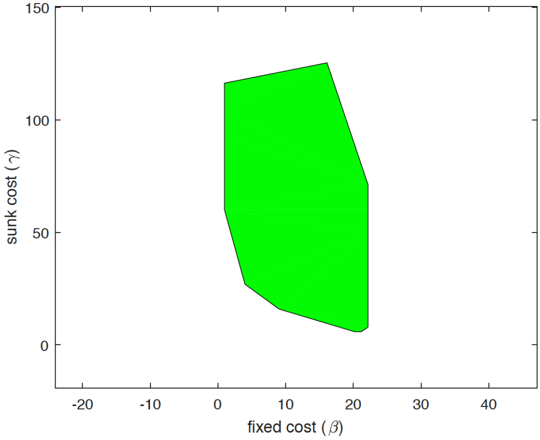


Figure: Concrete Data: Prelim Projection of the Estimated Set for (β, γ)

Initial Conditions in the Single-Agent Case

In our single-agent case, as in Collard-Wexler, we could complete model with an *ad-hoc* equation (or process) that determines the initial x_{i1}, for example,

$$x_{i1} = f(z_i, u_{i0})$$

This would give us equalities in condition (1) and then in simple discrete cases the GIV method is equivalent to MLE. This gives an "IV" interpretation to MLE applied to the complete model. Note: MLE might or might not be point identified, past exogenous shifters can still be important.

An Alternative: Dynamic Panel Restrictions

Note that our model of serial correlation

$$u_{it} = \rho u_{t-1} + \sigma \nu_{it}$$

suggests an additional "dynamic panel data" style restriction, that the "innovation" ν_{it} is independent (or mean independent) of past *endogenous* states $(x_{it-1}, x_{it-2}, \ldots)$.

This "dynamic panel" restriction is not really in the spirit of the current paper, but it suggests an approach that may be very helpful when instruments for current market structure are not available or not very powerful.

A Last Practical Concern

By a classic curse of dimensionality in dynamic games, the number of "natural" parameters in the policy function grows rapidly in the number of states and players. However, this problem also occurs in the "exogenous states" i.i.d. literature. In both cases, we need some lower dimensional approximation to the policy functions.

The existing empirical literature often uses an (internally inconsistent) "double parametric" specification, with (for example) a "nonparametric" entry equation estimated in the first stage by a highly parametric probit equation. Our approach suggests, at least, estimating that probit via GIV methods rather than MLE treating states as econometrically exogenous.

Conclusion

- Market structure in dynamic IO models should not be assumed to be exogenous.
- General IV models are one natural way to handle the econometric endogeneity of states, while preserving much of the intuition of existing two-step methods.
- Intuition of IV: past exogenous shifters (and regulation, etc) are correlated with today's state.
- May have set-identified policies and/or structural parameters. The data, the IV restriction and Bellman's equation together restrict the identified set.

Needed: more applications!

References I

Abbring, J. H., and J. R. Campbell (2010): "Last-In First-Out Oligopoly Dynamics," *Econometrica*, 78(5), 1491–1527.

Andrews, D. W., and X. Shi (2013): "Inference based on conditional moment inequalities," *Econometrica*, 81(2), 609–666.

Andrews, D. W., and G. Soares (2010): "Inference for Parameters Defined by Moment Inequalities Using Generalized Moment Selection," *Econometrica*, 78(1), 119–157.

Arcidiacono, P., and R. Miller (2011): "Conditional Choice Probability Estimation of Dynamic Discrete Choice Models with Unobserved Heterogeneity," *Econometrica*, 7(6), 1823–1868.

References II

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

Beresteanu, A., F. Molinari, and I. Molchanov (2011): "Sharp Identification Regions in Models with Convex Moment Predictions," *Econometrica*, 79(6), 1785–1821.

Berry, S., and E. Tamer (2007): "Identification in Models of Oligopoly Entry," in *Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress*, ed. by W. N. R. Blundell, and T. Persson, vol. 2. Cambridge University Press.

Bugni, F. A., I. A. Canay, and X. Shi (2017): "Inference for subvectors and other functions of partially identified parameters in moment inequality models," *Quantitative Economics*, 8(1), 1–38.

References III

Chernozhukov, V., and C. Hansen (2005): "An IV Model of Quantile Treatment Effects," *Econometrica*, 73(1), 245–261.

Chernozhukov, V., H. Hong, and E. Tamer (2007): "Estimation and Confidence Regions for Parameter Sets in Econometric Models1," *Econometrica*, 75(5), 1243–1284.

Chesher, A., and A. Rosen (2015): "Characterizations of identified sets delivered by structural econometric models," working paper CWP 63/15, Cemmap, Institute for Fiscal Studies UCL.

Ciliberto, F., and E. Tamer (2009): "Market Structure and Multiple Equilibria in Airline Markets," *Econometrica*, 77(6), 17911828.

Collard-Wexler, A. (2014): "Mergers and Sunk Costs: An Application to the Ready-Mix Concrete Industry," *American Economic Journal: Microeconomics*, 6(4), 407–447.

References IV

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

Kasahara, H., and K. Shimotsu (2009): "Nonparametric Identification of Finite Mixture Models of Dynamic Discrete Choices," *Econometrica*, 77(1), pp. 135–175.

Keane, M. P., and K. I. Wolpin (1997): "The Career Decisions of Young Men," *Journal of Political Economy*, 105(3), 473–522.

Magnac, T., and D. Thesmar (2002): "Identifying Dynamic Discrete Decision Processes," *Econometrica*, 70(2), 801–816.

Manski, C. F. (2003): *Partial Identification of Probability Distributions*. Springer, New York.

References V

Manski, C. F., and E. Tamer (2002): "Inference on Regressions with Interval Data on a Regressor or Outcome," *Econometrica*, 70(2), 519–546.

Pakes, A., M. Ostrovsky, and S. Berry (2007): "Simple Estimators for the Parameters of Dynamic Games, with Entry/Exit Examples," *RAND Journal of Economics*, 38(2), 373–399.

Pesendorfer, M., and P. Schmidt-Dengler (2008): "Asymptotic Least Squares Estimators for Dynamic Games," *Review of Economic Studies*, 75(3), 901–928.

Rust, J. (1987): "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55(5), 999–1033.

References VI

TAMER, E. (2003): "Incomplete Simultaneous Discrete Response Model with Multiple Equilibria," *The Review of Economic Studies*, 70(1), pp. 147–165.