Identification of DDC Models and Counterfactuals

Paul T. Scott

Empirical IO
Winter 2016

Outline

Magnac and Thesmar (2002)

Kalouptsidi, Scott, and Souza-Rodrigues: single agent

Kalouptsidi, Scott, and Souza-Rodrigues: games

"Identifying Dynamic Discrete Decision Processes"
Magnac and Thesmar (2002)

Setup

- $x \in X$ - state variables
- $p_i(x)$ - choice probabilities (data)
- $u_i(x)$ - per-period utility from action i in state x
- $v_i(x)$ - conditional value function of action i in state x
- K - the reference action
- G - distribution of conditionally independent shocks
- q - the Hotz-Miller inversion function. i.e., $q_i(p(x)) = v_i(x) - v_K(x)$
- R - the surplus function, $R(v; G) = E_G(\max_i \{v_i + \varepsilon_i\})$

Lemma 1

Lemma 1 is basically a convenient restatement of the Hotz Miller inversion.

Lemma 1

For any action j and state x,

$$\begin{aligned} u_j(x) &= u_K(x) + q_j(p(x); G) \\ &\quad -\beta\left(E\left[v_K(x') \mid x, j\right] - E\left[v_K(x') \mid x, K\right]\right) \\ &\quad -\beta\left(E\left[R(q(p(x'); G)) \mid x, j\right] - E\left[R(q(p(x'); G)) \mid x, K\right]\right) \end{aligned}$$

Lemma 1, example

Let's derive Lemma 1 for the case of logit errors.

$$\begin{aligned}
\ln\left(\frac{p_j(x)}{p_K(x)}\right) &= v_j(x) - v_K(x) \\
&\Leftrightarrow \\
u_j(x) &= u_K(x) + \ln\left(\frac{p_j(x)}{p_K(x)}\right) \\
&\quad -\beta\left(E\left[\bar{V}(x')\,|x, j\right] - E\left[\bar{V}(x')\,|x, K\right]\right) \\
&= u_K(x) + \ln\left(\frac{p_j(x)}{p_K(x)}\right) \\
&\quad -\beta E\left[\ln\sum_i \frac{p_i(x')}{p_K(x')}\exp\left(v_K(x')\right)|x, j\right] \\
&\quad +\beta E\left[\ln\sum_i \frac{p_i(x')}{p_K(x')}\exp\left(v_K(x')\right)|x, K\right]
\end{aligned}$$

Lemma 1, example

Let's derive Lemma 1 for the case of logit errors.

$$\begin{aligned}
\ln\left(\frac{p_j(x)}{p_K(x)}\right) &= v_j(x) - v_K(x) \\
&\Leftrightarrow \\
u_j(x) &= u_K(x) + \ln\left(\frac{p_j(x)}{p_K(x)}\right) \\
&\quad -\beta\left(E\left[\bar{V}(x')\,|x,j\right] - E\left[\bar{V}(x')\,|x,K\right]\right) \\
&= u_K(x) + \ln\left(\frac{p_j(x)}{p_K(x)}\right) \\
&\quad -\beta\left(E\left[v_K(x')\,|x,j\right] - E\left[v_K(x')\,|x,K\right]\right) \\
&\quad -\beta\left(E\left[\ln\textstyle\sum_i \frac{p_i(x')}{p_K(x')}|x,j\right]\right. \\
&\quad \left. -E\left[\ln\textstyle\sum_i \frac{p_i(x')}{p_K(x')}|x,K\right]\right)
\end{aligned}$$

Lemma 1

Proposition 2

Let $C = \{c | c = (\beta, G, u_K(\cdot), v_K(\cdot))\}$.
(i) For a given $c \in C$, there exists one vector $(u_1(\cdot), \ldots, u_{K-1}(\cdot))$ compatible with $p(X)$.
(ii) Let $\mathbf{u}$ be the $(u_1(\cdot), \ldots, u_{K-1}(\cdot))$ vector associated with c and let $\mathbf{u}'$ be associated with $c' \neq c$. Then, $(\mathbf{u}, c)$ and $(\mathbf{u}', c')$ are observationally equivalent.

Thus, there is an observationally equivalent model associated with each element of $C = \{c | c = (\beta, G, u_K(\cdot), v_K(\cdot))\}$

Restrictions for identification

- It's common to do something like assuming $u_K(x) = 0$ for all x
- This is not innocent. Note that restricting $u(x) = 0$ for a single x is an innocent normalization, but restricting payoffs to be flat across states is a substantive assumption.
    - Sometimes such restrictions are very natural. Think about what the restriction is in Rust (1987).

Identifying the discount factor

Idea behind Magnac and Thesmar's Proposition 4:

- Suppose you have different values of state variables which give the same current profits, but have different expectations for future values of the state variables.
  - e.g., perhaps we observe both current and futures prices
- In this case, one can identify the discount factor because we have something that shifts continuation values rather than shifting both current profits and continuation values.

"Identification of Counterfactuals in Dynamic Discrete Choice Models"
Kalouptsidi, Scott, and Souza-Rodrigues (2016)

The need for restrictions

Consider a utility function $\pi(a, s)$ where a is an action selected from some discrete set and s is a state variable.

- In static model, normalizing $\pi(J, s) = 0$ for some action J, and all states s is harmless.

The need for restrictions

Consider a utility function $\pi(a, s)$ where a is an action selected from some discrete set and s is a state variable.

- In static model, normalizing $\pi(J, s) = 0$ for some action J, and all states s is harmless.
- In a dynamic model, this is not harmless in general; profits in state s' affect incentives in state s.

The need for restrictions

Consider a utility function $\pi(a, s)$ where a is an action selected from some discrete set and s is a state variable.

- In static model, normalizing $\pi(J, s) = 0$ for some action J, and all states s is harmless.
- In a dynamic model, this is not harmless in general; profits in state s' affect incentives in state s.
- Unfortunately, identification of DDC models requires a normalization like this or other restrictions (Magnac and Thesmar, 2002).

The need for restrictions

Consider a utility function $\pi(a, s)$ where a is an action selected from some discrete set and s is a state variable.

- In static model, normalizing $\pi(J, s) = 0$ for some action J, and all states s is harmless.
- In a dynamic model, this is not harmless in general; profits in state s' affect incentives in state s.
- Unfortunately, identification of DDC models requires a normalization like this or other restrictions (Magnac and Thesmar, 2002).
- However, it's not clear that *counterfactuals* are under-identified in the same way. It could be the case that every model consistent with the data produces the same counterfactual behavior.

"the entire dynamic discrete choice project thus appears to be without empirical content, and the evidence from it at the whim of investigator choices about functional forms of estimating equations and application of ad hoc exclusion restrictions."
– Heckman and Navarro (2007)

Literature

- Recently, some papers have explored when counterfactuals are identified, offering cases that are and that aren't: Aguirregabiria (2010), Norets and Tang (2014), Arcidiacono and Miller (2015), Aguirregabiria and Suzuki (2014).
- Kalouptsidi, Scott, and Souza-Rodrigues offer a characterization (necessary and sufficient conditions), providing results that one can use to check whether virtually ever counterfactual in the literature is identified.
- In the subsequent slides, I offer proofs of relevant special cases which are simpler than the proofs in any of these papers.

Identification of Counterfactuals

- A counterfactual can involve a change in the utility function, $\pi \to \widetilde{\pi}$:

  $$\widetilde{\pi} = h(\pi),$$

  where h is some differentiable function, and a change in the transition matrix $F \to \widetilde{F}$

- We say that a counterfactual is identified if all utility functions consistent with observed data all generate the same counterfactual CCPs $\widetilde{p}$.

Result 1: Lump sum transfers

Result

In a single-agent setting, if $\widetilde{F} = F$ and $\widetilde{\pi} = \pi + g$, where g is a known vector, then counterfactual choice probabilities $\widetilde{p}$ are identified.

Result 2: Changes in transition process

Result

In a single-agent setting, if a counterfactual changes the transition process from F to $\widetilde{F}$, but the utility function is unchanged, then counterfactual choice probabilities $\widetilde{p}$ are not identified unless

$$(I - \beta F_a)(I - \beta F_J)^{-1} - \left(I - \beta \widetilde{F}_a\right)\left(I - \beta \widetilde{F}_J\right)^{-1} = 0 \qquad (1)$$

for all $a \neq J$, where $F_a \in \mathbb{R}^{X \times X}$ is the transition matrix conditional on action a, and I is an identity matrix of size X.

Value functions

Define the *ex ante value function*, which represents the expectation of the value function before idiosyncratic shocks are realized:

$$V\left(x_{it}\right) \equiv \int V\left(x_{it}, \varepsilon_{it}\right) dG\left(\varepsilon_{it}\right),$$

and the *conditional value function*, which represents the expected discounted payoffs conditional on particular action before the realization of idiosyncratic shocks:

$$v_a\left(x_{it}\right) \equiv u\left(a, x_{it}\right) + \beta E\left[V\left(x_{it+1}\right) | a, x_{it}\right].$$

CCPs

The agent's optimal policy is given by the conditional choice probabilities (CCPs):

$$p_a(x_{it}) = \int 1\{v_a(x_{it}) + \varepsilon_{it}(a) \geq v_j(x_{it}) + \varepsilon_{it}(j), \text{ for all } j \in \mathcal{A}\} \, dG(\varepsilon_{it})$$

where $1\{\cdot\}$ is the indicator function. Let p be the vector of CCPs, $p_a(x_{it})$ with $a \in \mathcal{A}$ and $x \in \mathcal{X}$.

Relationships

Some important relationships:

$$\pi_a = v_a - \beta F_a V, \quad \text{for } a = 1, ..., A \tag{2}$$

$$v_a - v_j = \sigma \phi_{aj}, \text{ for } a = 1, ..., A, \ a \neq j \tag{3}$$

$$V = v_a + \sigma \psi_a, \quad \text{for } a = 1, ..., A, \tag{4}$$

where $\pi_a, v_a, V, \phi_{aj}, \psi_a \in \mathbb{R}^X$, with $\pi_a(x) = \pi(a, x)$; F_a is the transition matrix with (m, n) element equal to $\Pr(x_{it+1} = x_n | a, x_{it} = x_m)$. Equation (2) defines the conditional value function; (3) restates the Hotz-Miller lemma; and (4), the Arcidiacono-Miller lemma.

One more relationship

Combining the above relationships, we can represent the under-identification problem in terms of a "strong normalization"

Lemma 1

Let $J \in \mathcal{A}$ be some reference action. For each $a \neq J$, the payoff function π_a can be represented as an affine transformation of π_J:

$$\pi_a = A_a \pi_J + b_a, \tag{5}$$

where $A_a = (I - \beta F_a)(I - \beta F_J)^{-1}$ and $b_a = \sigma(A_a \psi_J - \psi_a)$.

Proof of Lemma 1

Fix the vector $\pi_J \in \mathbb{R}^X$. Then,

$$\pi_a = v_a - \beta F_a V = V - \sigma\psi_a - \beta F_a V = (I - \beta F_a)\, V - \sigma\psi_a$$

where for $a = J$

$$V = (I - \beta F_J)^{-1} (\pi_J + \sigma\psi_J)\,.$$

After substituting for V, we have

$$\pi_a = (I - \beta F_a)\,(I - \beta F_J)^{-1} (\pi_J + \sigma\psi_J) - \sigma\psi_a.$$

$(I - \beta F_J)$ is invertible because F_J is a stochastic matrix and hence the largest eigenvalue is equal or smaller than one. The eigenvalues of $(I - \beta F_J)$ are given by $1 - \beta\lambda$, where λ are the eigenvalues of F_J. Because $\beta < 1$ and $\lambda \leq 1$, we have $1 - \beta\lambda > 0$.

Proof of Results 1-2

Write out the definition of the conditional value function in vector notation,

$$v_a = \pi_a + \beta F_a V,$$

and then use Arcidiacono and Miller's Lemma to substitute for the ex ante value function V:

$$v_a = \pi_a + \beta F_a \left(v_a + \psi_a (p) \right).$$

This allows us to express the conditional value in terms of the payoff function and transition matrix:

$$v_a = \left(I - \beta F_a \right)^{-1} \left(\pi_a + \beta F_a \psi_a (p) \right). \tag{6}$$

Proof of Results 1-2

- Write out that equation for two different actions, a and J, and apply Lemma 1:

$$\begin{aligned} v_a &= (I-\beta F_a)^{-1}\left(A_a \pi_J + b_a(p) + \beta F_a \psi_a(p)\right), \\ v_J &= (I-\beta F_J)^{-1}\left(\pi_J + \beta F_J \psi_J(p)\right). \end{aligned} \tag{7}$$

- Difference and apply the HM inversion:

$$\begin{aligned} \psi_J(p) - \psi_a(p) &= (I-\beta F_a)^{-1}\left(A_a u_J + b_a(p) + \beta F_a \psi_a(p)\right) \\ &\quad - (I-\beta F_J)^{-1}\left(u_J + \beta F_J \psi_J(p)\right). \end{aligned} \tag{8}$$

Proof of Results 1-2

- This equation should also hold for the counterfactual utility function $h(\pi)$ and transation matrix $\tilde{F}$:

$$\begin{aligned}\psi_J(\widetilde{p}) - \psi_a(\widetilde{p}) &= \left(I - \beta \widetilde{F}_a\right)^{-1} \left(h_a(\pi) + \widetilde{b}_a(\widetilde{p}) + \beta \widetilde{F}_a \psi_a(\widetilde{p})\right) \\ &\quad - \left(I - \beta \widetilde{F}_J\right)^{-1} \left(h_J(\pi) + \beta \widetilde{F}_J \psi_J(\widetilde{p})\right)\end{aligned}$$

- This allows us to see both Results 1 and 2 . . .

Proof of Result 1

- To see result 1, plug in
  $h_a(\pi) = \pi_a + g_a$
  $h_J(\pi) = \pi_J + g_J$
  $\pi_a = A_a \pi_J + b_a$

$$
\begin{aligned}
\psi_J(\widetilde{p}) - \psi_a(\widetilde{p}) \quad = \quad & (I - \beta F_a)^{-1} (A_a \pi_J + \\
& b_a(p) + g_a + \widetilde{b}_a(\widetilde{p}) + \beta F_a \psi_a(\widetilde{p})\Big) \\
& -(I - \beta F_J)^{-1} (\pi_J + g_J + \beta F_J \psi_J(\widetilde{p}))
\end{aligned}
$$

  noting that $\widetilde{F} = F$ for the counterfactuals of Result 1.
- Recall that $A_a = (I - \beta F_a)(I - \beta F_J)^{-1}$

Proof of Result 1

$$
\begin{aligned}
\psi_J(\widetilde{p}) - \psi_a(\widetilde{p}) &= (I - \beta F_a)^{-1} \big(A_a \pi_J + \\
&\quad b_a(p) + g_a + \widetilde{b}_a(\widetilde{p}) + \beta F_a \psi_a(\widetilde{p})\big) \\
&\quad - (I - \beta F_J)^{-1} (\pi_J + g_J + \beta F_J \psi_J(\widetilde{p}))
\end{aligned}
$$

- Recall that $A_a = (I - \beta F_a)(I - \beta F_J)^{-1}$, implying

$$
\begin{aligned}
\psi_J(\widetilde{p}) - \psi_a(\widetilde{p}) &= (I - \beta F_J)^{-1} A_a \pi_J + \\
&\quad (I - \beta F_J)^{-1} \left(b_a(p) + g_a + \widetilde{b}_a(\widetilde{p}) + \beta F_a \psi_a(\widetilde{p})\right) \\
&\quad - (I - \beta F_J)^{-1} (\pi_J + g_J + \beta F_J \psi_J(\widetilde{p}))
\end{aligned}
$$

Proof of Result 1

- The π_J terms cancel, leaving us with

$$\begin{aligned}\psi_J(\widetilde{p}) - \psi_a(\widetilde{p}) &= (I - \beta F_J)^{-1}\left(b_a(p) + g_a + \widetilde{b}_a(\widetilde{p}) + \beta F_a \psi_a(\widetilde{p})\right) \\ &\quad - (I - \beta F_J)^{-1}\left(g_J + \beta F_J \psi_J(\widetilde{p})\right)\end{aligned}$$

- Thus, the equation that counterfactuals CCPs $\widetilde{p}$ must satisfy does not depend on π_J, meaning any "normalization" will result in the same counterfactual behavior.

- For Result 2, the fact that $\widetilde{F} \neq F$ means that the π_J term does not cancel out unless the transition matrices satisfy the special condition described in the theorem. When the counterfactual changes F, the condition $\widetilde{p}$ must satisfy depends on the value of π_J, and so the counterfactual behavior is sensitive to "normalizations."

"On the Non-Identification of Counterfactuals in Dynamic Games"
Kalouptsidi, Scott, and Souza-Rodrigues (2016)

An entry game

- Let's consider a dynamic entry game introduced by Pesendorfer and Schmidt-Dengler.
- Two players indexed by $i = A, B$
- In each period t, the players simultaneously choose whether to be active in the market ($a_{it} = 1$) or not ($a_{it} = 0$)
- state variable which equals the player's action in the previous period: $s_{it} = a_{i,t-1}$
- The state of the game is simply the pair of states, $x_t = (s_{At}, s_{Bt})$, which is common knowledge

Payoffs

- As a baseline case, we consider the following payoff function:

$$u_i\left(a_i, a_{-i}, s_i, s_{-i}\right) = \begin{cases} 0 & \text{if } a_i = 0, s_i = 0 \\ \phi & \text{if } a_i = 0, s_i = 1 \\ \pi_1 - c & \text{if } a_i = 1, s_i = 0, a_{-i} = 0 \\ \pi_1 & \text{if } a_i = 1, s_i = 1, a_{-i} = 0 \\ \pi_2 - c & \text{if } a_i = 1, s_i = 0, a_{-i} = 1 \\ \pi_2 & \text{if } a_i = 1, s_i = 1, a_{-i} = 1 \end{cases}$$

- A symmetric equilibrium for this model: each player enters with probability .576 when both players were not active in the previous period. Each player remains active with probability .595 when both players competed previously. When only one player was active previously, the incumbent remains active with probability .842 and the other firm enters with probability .305.
- We also consider another payoff function which is consistent with the same baseline equilibrium.

Table: Entry Game: Payoff Functions

$\pi(a_i, a_i, s_i, s_{-i})$		Model 1				Model 2			
		$(a_i, a_{-i}) =$				$(a_i, a_{-i}) =$			
s_i	s_{-i}	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$
0	0	0	0	1	-1.4	0	0	0	-0.506
0	1	0	0	1	-1.4	0	0	0	-1.11
1	0	0.1	0.1	1.2	-1.2	0	0	0	1.51
1	1	0.1	0.1	1.2	-1.2	0	0	0	-0.399

Table: Entry Game: Payoff Functions

$P\left(\text{active}\middle\|s_i, s_i\right)$		Baseline	CF – fixed opponent		CF – equilibrium	
			Model 1	Model 2	Model 1	Model 2
s_i	s_{-i}					
0	0	0.576	0.527	0.527	0.634	0.516
0	1	0.305	0.257	0.257	0.207	0.243
1	0	0.842	0.875	0.875	0.983	0.836
1	1	0.595	0.643	0.643	0.692	0.612