Identification of the Discount Factor in DDC Models

Øystein Daljord

Chicago Booth

Dynamic Structural Econometrics Course

University of Chicago

July 12, 2019

Identification of the discount factor by an exclusion restriction

Standard DDC model known to be non-parametrically underidentified without further restrictions (Rust (1994), Magnac & Thesmar (2002))

→ Implies that counterfactual choice probabilities are not identified

In applications, the discount factor is often fixed at some plausible value, e.g. 0.95

→ Assumes the dynamics under study

Non-identification due to an incompleteness of the model

→ Identification requires further assumptions

Discuss intuitively appealing restrictions that identify the discount factor

Common intuition suggests variation that shifts future values, but not current payoffs is informative about time preferences

Yao et al (2012), *JMR*

- Identify the utility function from the terminal period. Once known, identify the discount factor in the next-to-terminal period, conditional on the utility function

Lee (2013), *AER*

- "[A] myopic model ($\beta = 0$) would be rejected if current hardware sales are influenced by variation in observed availability and estimated quality of future software."

deGroote & Verboven (2018), *AER*

- Households current responses to regulation that shifts future, but not current, electricity costs carry information about time preferences

Common intuition can be formalized as an exclusion restriction on the pay-off function

Covers most empirical strategies in the literature

- The exclusion restriction has economically transparent interpretation, similar to standard IV argument
- Leads to one moment condition that characterizes the identified set independently of the payoff function
  - → Can be used directly in estimation
  - → Fast to compute
- The identified set is finite
  - → Important classes of models are point identified
- The identified model has *empirical content*: the model under the exclusion restrictions can be rejected. Gives an intuitive specification test

1. Derive the main identification result
2. Examples of point identification, set identification, and falsification
3. Examples of applications
4. Briefly about set estimators
5. Compare to alternative results in the literature
   → Magnac & Thesmar (2002) and Fang & Wang (2015)
6. Extension to non-stationary models

Data, primitives, and assumptions: Rust (1994)

- Single agent problem
- Discrete time, infinite horizon
- Choice set $\mathcal{D} = \{1, 2, \ldots, K\}$
- Observable states: $\mathcal{X} = \{x_1, \ldots, x_J\}$
- Markov state transitions $\boldsymbol{f}_k(x) = [Pr(x_1|x, k), \ldots, Pr(x_J|x, k)]$
- Rational expectations
- Unobservable states $\varepsilon = \{\varepsilon_1, \ldots, \varepsilon_K\} \sim g$, independent of x (CI)
- Utilities $u_k(x, \varepsilon) = u_k(x) + \varepsilon_k$ (AS)
- Normalize $u_K(x) = 0$ for all $x \in \mathcal{X}$
- Observed choice probabilities $p_k(x)$, for all $k \in \mathcal{D}, x \in \mathcal{X}$

Primitives: $\{\boldsymbol{u}, \beta, \boldsymbol{f}, g\}$

Here, g is identified $EV1$ by assumption, and $\boldsymbol{f}$ by observed transitions

Parametric identification

The model with a linear-in-parameters utility function

$$u_d(z) = \theta' z,$$

where $z = [z_1, \ldots, z_r]$ are state variables, is point identified. See Komarova et al (2018)

- Includes the empirical specification in most applications
- Must believe the functional form

Two sets of key equations

Alternative specific value function

$$\mathbf{AVF}: v_k(x) = u_k(x) + \beta \int \mathbb{E}_\epsilon \left[\max_{k' \in D} \{ v_{k'}(x') + \varepsilon_{k'} \} \right] df_k(x'|x)$$
$$\text{for all } k \in \mathcal{D},\ x \in \mathcal{X}.$$

Reduced form (Hotz-Miller): the value contrasts are identified from the observed choices

$$\mathbf{HM}: \ln(p_k(x)) - \ln(p_K(x)) = v_k(x) - v_K(x)$$
$$\text{for all } k \in \mathcal{D}/\{K\},\ x \in \mathcal{X}.$$

Summarizes all the predictions the model makes about the data

The utilities and the discount factor are jointly underidentified

Need β to decompose the identified value contrasts

$$v_k(x) - v_K(x) = u_k(x) + \beta\left[\boldsymbol{f}_k(x) - \boldsymbol{f}_K(x)\right]\boldsymbol{v}$$

in a current utility and a continuation value component. Here, $\boldsymbol{v}_K = [v_K(x_1), \ldots, v_K(x_J)]'$

- No more predictions left in the structure
  - $\rightarrow$ $(K-1) \times J$ equations in **HM**
  - $\rightarrow$ $(K-1) \times J + 1$ unknown primitives $\boldsymbol{u}$ and β

One experimental approach

- Shift the continuation value component, holding current utilities fixed
- Current period choice response reflects time preferences

$\rightarrow$ Define an exclusion restriction which implies that experiment

Corrected value function a known mapping from data

Ex ante value function net of the expected value of the reference choice

$$\begin{aligned}\psi(x) &= \mathbb{E}_\epsilon\left[V(x) - v_k(x)\right]\\ &= \mathbb{E}_\epsilon\left[\max_{d\in\mathcal{D}}\{v_d(x) - v_K(x) + \epsilon_d\}\right]\end{aligned}$$

is a known mapping from the data derived from g only (Arcidiacono & Miller (2011)). For EV1, $\psi(x) = -\ln(p_K(x))$

Why useful?

$$\begin{aligned}\mathbb{E}_\epsilon\left[V(x)\right] &= \mathbb{E}_\epsilon\left[\max_{d\in\mathcal{D}}\{v_d(x) + \epsilon_d\}\right]\\ &= \mathbb{E}_\epsilon\left[\max_{d\in\mathcal{D}}\{v_d(x) - v_K(x) + \epsilon_d + v_K(x)\}\right]\\ &= \psi(x) + v_K(x)\end{aligned}$$

Use the recursive structure to write $v_K(x)$ in terms of data and primitives

Use that $\boldsymbol{v} = \boldsymbol{\Psi} + \boldsymbol{v}_K$ and $\boldsymbol{v}_K = \beta \boldsymbol{f}_K(x)\boldsymbol{v}$

$$
\begin{aligned}
v_K(x) &= \beta \boldsymbol{f}_K(x)[\boldsymbol{\Psi} + \boldsymbol{v}_K] \\
&= \beta \boldsymbol{f}_K(x)[\boldsymbol{\Psi} + \beta \boldsymbol{f}_K[\boldsymbol{\Psi} + \boldsymbol{v}_K]] \\
&= \beta \boldsymbol{f}_K(x)[\boldsymbol{\Psi} + \beta \boldsymbol{\Psi} + \beta^2 \boldsymbol{f}_K^2[\boldsymbol{\Psi} + \boldsymbol{v}_K]] \\
&= \beta \boldsymbol{f}_K(x)[\boldsymbol{I} - \beta \boldsymbol{f}_K]^{-1} \boldsymbol{\Psi}
\end{aligned}
$$

where $\boldsymbol{f}_K$ stacks $\boldsymbol{f}_K(x)$ etc.

Write the value contrasts in terms of primitives only

Combine with **HM** and **AVF** to get moment condition

$$\begin{aligned}
\ln(p_k(x)) - \ln(p_K(x)) =& v_k(x) - v_K(x) \\
=& \beta \left[\boldsymbol{f}_k(x) - \boldsymbol{f}_K(x)\right] \left[\boldsymbol{\Psi} + \boldsymbol{v}_K\right] + u_k(x) - u_K(x) \\
=& \beta \left[\boldsymbol{f}_k(x) - \boldsymbol{f}_K(x)\right] \left[\boldsymbol{I} - \beta \boldsymbol{f}_K\right]^{-1} \boldsymbol{\Psi} + u_k(x)
\end{aligned}$$

Exclusion restriction on utilities

Assume

$$u_k(x_1) - u_l(x_2) = 0$$

for some $k \in \mathcal{D}/\{K\}, l \in \mathcal{D}$ and $x_1, x_2 \in \mathcal{X}$, for either $x_1 \neq x_2$, or $k \neq l$, or both.

Difference the **HM** conditions corresponding to the indices of the exclusion restriction

$$\begin{aligned}
&\ln\left(p_k(x_1)/p_K(x_1)\right) - \ln\left(p_l(x_2)/p_K(x_2)\right) \\
&\qquad = \beta\left[\boldsymbol{f}_k(x_1) - \boldsymbol{f}_K(x_1) - \boldsymbol{f}_l(x_2) + \boldsymbol{f}_K(x_2)\right]\left[\boldsymbol{I} - \beta\boldsymbol{f}_K\right]^{-1}\boldsymbol{\Psi} + \\
&\qquad\qquad \underbrace{u_k(x_1) - u_l(x_2)}_{=0}
\end{aligned}$$

This moment condition contains all the information in the data on β

The identifying moment condition has a diff-in-diff structure

Suppose $l = k$ and $x_1 \neq x_2$

$$\underbrace{\ln\left(p_k(x_1)/p_K(x_1)\right) - \ln\left(p_k(x_2)/p_K(x_2)\right)}_{\text{diff-in-diff in log choice probabilities}}$$
$$= \beta \underbrace{\left[\boldsymbol{f}_k(x_1) - \boldsymbol{f}_K(x_1) - \boldsymbol{f}_k(x_2) + \boldsymbol{f}_K(x_2)\right] \boldsymbol{v}}_{\text{diff-in-diff in expected values}}$$

Relates shifts shifts in value contrasts to shifts in current period choice contrasts

1. 1st diff: Observed choices contrasted between k and K
2. 2nd diff: Contrasts differenced between x_1 and x_2

Similar to IV condition

- Shifts between states shift the continuation value without shifting the unobserved current period payoffs

The moment condition contains all the information in the data on β

$$
\begin{aligned}
&\ln\left(p_k(x_1)/p_K(x_1)\right) - \ln\left(p_k(x_2)/p_K(x_2)\right) \\
&\qquad = \beta\left[\boldsymbol{f}_k(x_1) - \boldsymbol{f}_K(x_1) - \boldsymbol{f}_k(x_2) + \boldsymbol{f}_K(x_2)\right]\left[\boldsymbol{I} - \beta\boldsymbol{f}_K\right]^{-1}\boldsymbol{\Psi} \\
&\qquad = \beta\left[\boldsymbol{f}_k(x_1) - \boldsymbol{f}_K(x_1) - \boldsymbol{f}_k(x_2) + \boldsymbol{f}_K(x_2)\right]\left[\boldsymbol{\Psi} + \beta\boldsymbol{f}_K\boldsymbol{\Psi} + \beta^2\boldsymbol{f}_K^2\boldsymbol{\Psi} + \ldots\right.
\end{aligned}
$$

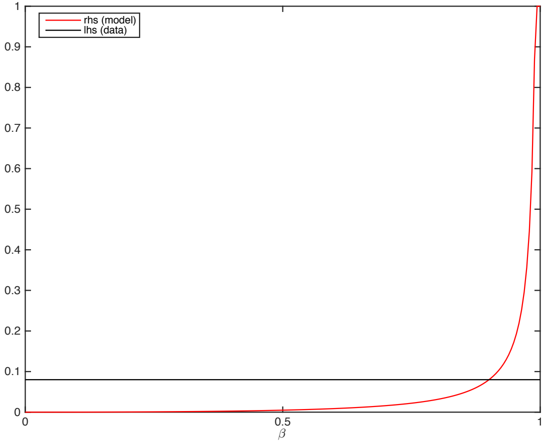
Moment condition is a real valued convergent power series in $\beta \in [0, 1)$, hence generally not a unique zero on its domain

Three cases

1. Unique solution (point identified)
2. No solution (falsified)
3. Multiple solutions (set identified)

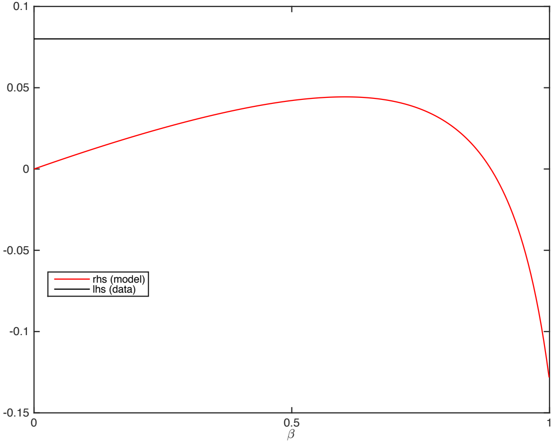
Case 1: Point identified

Plot of moment condition for some population data



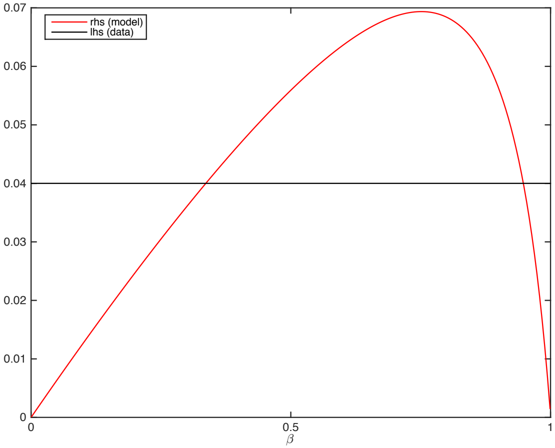
Case 2: Falsified

No fixed point for $\beta \in [0, 1)$



Case 3: Set identified

Locally identified: Point identified in a small neighbourhood of a fixed point



Abbring & Daljord (2019) main result

Exclusion restriction

$$u_k(x_1) = u_l(x_2) \tag{1}$$

for $k \in \mathcal{D}/\{K\}, l \in \mathcal{D}, x_1, x_2 \in \mathcal{X}$, and where either $k \neq l$, $x_1 \neq x_2$, or both

Theorem

Suppose that (1) *holds and that either* $p_k(x_1)/p_K(x_1) \neq p_l(x_2)/p_K(x_2)$, *or the rank condition*

$$[\boldsymbol{f}_k(x_1) - \boldsymbol{f}_K(x_1) - \boldsymbol{f}_l(x_2) + \boldsymbol{f}_K(x_2)]\,\boldsymbol{\Psi} \neq 0$$

holds. Then, the identified set $\mathcal{B}$ *is a closed discrete subset of* $[0, 1)$.

Sketch of proof

Under the stated assumptions, the moment condition

$$\begin{aligned}&\ln\left(p_k(x_1)/p_K(x_1)\right) - \ln\left(p_l(x_2)/p_K(x_2)\right)\\ &\qquad = \beta\left[\boldsymbol{f}_k(x_1) - \boldsymbol{f}_K(x_1) - \boldsymbol{f}_l(x_2) + \boldsymbol{f}_K(x_2)\right]\left[\boldsymbol{\Psi} + \beta\boldsymbol{f}_K\boldsymbol{\Psi} + \beta^2\boldsymbol{f}_K^2\boldsymbol{\Psi} + \ldots\right.\end{aligned}$$

is a non-constant real analytic function of β on $[0, 1]$ with the only limit point at 1

So at most finitely many zeros on $[0, 1)$

Magnac & Thesmar's identification result

Important paper on identification of DDC models

Showed that the discount factor can be point identified by a different exclusion

- Hard to understand what the restriction means

Recall the reduced form: **HM** and **AVF** function give moment condition

$$\ln(p_k(x)) - \ln(p_K(x)) = \beta \left[\boldsymbol{f}_k(x) - \boldsymbol{f}_K(x)\right] \boldsymbol{\Psi} + U_k(x)$$

where

$$U_k(x) = u_k(x) - u_K(x) + \beta \left[\boldsymbol{f}_k(x) - \boldsymbol{f}_K(x)\right] \boldsymbol{v}_K$$

is the *current value function*

M&T's exclusion restriction is defined on the current value function

Assume

$$U_k(x_1) - U_l(x_2) = 0$$

for some $k \in D/\{K\}, l \in D$, for either $k \neq l$, or $x_1 \neq x_2$, or both.

Difference the corresponding **HM** moments

$$\begin{aligned}\ln\left(p_k(x_1)/p_K(x_1)\right) - \ln\left(p_l(x_2)/p_K(x_2)\right) = \\ \beta\left[\boldsymbol{f}_k(x_1) - \boldsymbol{f}_K(x_1) - \boldsymbol{f}_l(x_2) + \boldsymbol{f}_K(x_2)\right]\boldsymbol{\Psi} + \\ \underbrace{U_k(x_1) - U_l(x_2)}_{=\,0,\text{ by restriction}}\end{aligned}$$

Moment condition now linear in β

$\rightarrow$ has either no solution (falsified) or one solution (point identified)

The current value restriction has unintuitive economic interpretation

Current value function represents a value contrast between two choice sequences

$$U_k(x) = \underbrace{u_k(x) + \beta \boldsymbol{f}_k(x)\boldsymbol{v}_K}_{\text{choose } k \text{ now, } K \text{ next period, optimally ever after}} - \underbrace{u_K(x) + \beta \boldsymbol{f}_K(x)\boldsymbol{v}_K}_{\text{choose } K \text{ now, } K \text{ next period, optimally ever after}}$$

→ Does not correspond to common economic concepts or choice sequences

Hard to assess how plausible the identifying assumption is if we don't understand what it means

The literature sometimes mixes up $U(x)$ and $u(x)$

Bayer et al (2016), ECTA

> "*Magnac and Thesmar (2002) [...] showed that dynamic models are identified with an appropriate exclusion restriction, in particular, a variable that shifts expectations, but not current utility.*"

Magnac & Thesmar's exclusion restriction generally identifies a different parameter than ours

Rust (1987): Zurcher's choice

Renewal actions: Suppose choice K resets the state transition process

$$\boldsymbol{f}_K = \begin{bmatrix} 1 & \cdots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 1 & \cdots & 0 & 0 \end{bmatrix},$$

Then M&T's and our moment condition coincide

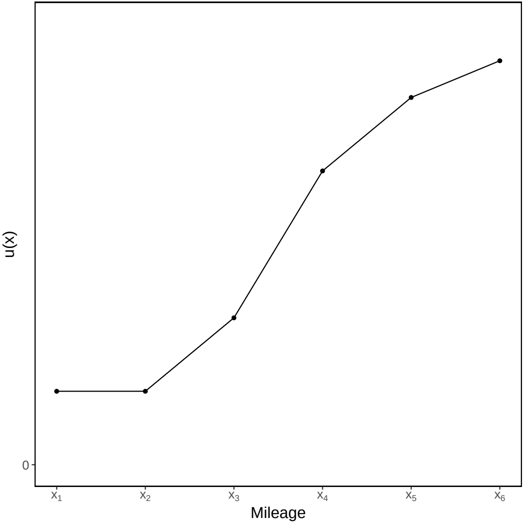
- Since $v_K(x) = \beta\left(m(x_1) + v_K(x_1)\right)$ for all $x \in \mathcal{X}$ (state independent), then $U_k(x) = u_k(x) + \beta[\boldsymbol{f}_k(x) - \boldsymbol{f}_K(x)]\boldsymbol{v}_K = u_k(x)$,

so point identified

Example of single action, one-period *finite dependence* (Arcidiacono & Miller (2011))

Zurcher's choice

If states have a natural ordering, shape restrictions can be cast as exclusion restrictions on pay-offs



Zurcher's choice

If states have a natural ordering, shape restrictions can be cast as exclusion restrictions on pay-offs

- If $u_k(x_1) = a u_k(x_2)$ for known a, then point identification
- If $u_k(x)$ is known for some $x \in X$ or $k \in \mathcal{D} \backslash \{K\}$, then point identification
  - $\rightarrow$ $u_k(x_1) = 0$, used by Rust (1987) to identify the replacement cost, can alternatively identify the discount factor (but does not work in those data)

Particularly useful for cardinal pay-offs, e.g. costs or profits

Finite dependence

If choosing a given single action ρ times resets the state distribution, then ρ period single action finite dependence

- By Theorem 2 in Abbring & Daljord (2019), ρ upper bounds the number of discount factors in the identified set

Models with one period, single action finite dependence popular in the literature

- Includes optimal stopping problems, some adoption problems, ...
- Scott (2013), Traiberman (2018), de Groote and Verboven (2018), Diamond et al. (2018)

Identification in non-stationary models

Unlike in stationary models, an assumption of stationary utilities has identifying power in non-stationary models

- Standard intuition: utilities are identified in the static terminal period choice problem, e.g. Yao et al (2012)
- Can be represented as exclusion restrictions on utilities. Same as in Bajari et al (2017)
- Potentially plausible assumption
- For some special cases, e.g. finite dependence, don't need long panels

Non-stationary models

Time index $\boldsymbol{u}_{k,t}$, hence

$$\boldsymbol{v}_{k,t} = \boldsymbol{u}_{k,t} + \beta \boldsymbol{f}_k \left[\boldsymbol{\Psi}_{t+1} + \boldsymbol{v}_{K,t+1} \right]$$

where $\boldsymbol{v}_{K,t} = \sum_{\tau=t+1}^{T} (\beta \boldsymbol{f}_K)^{\tau-t} \boldsymbol{\Psi}_\tau$.

Exclusion restriction $u_{k,t}(x_1) = u_{l,t'}(x_2)$ gives moment condition

$$\begin{aligned} \ln\left(p_{k,t}(x_1)/p_{K,t}(x_1)\right) - \ln\left(p_{l,t'}(x_2)/p_{K,t'}(x_2)\right) = \\ \beta \Bigg(\left[\boldsymbol{f}_k(x_1) - \boldsymbol{f}_K(x_1)\right] \left[\sum_{\tau=t+1}^{T} (\beta \boldsymbol{f}_K)^{\tau-t-1} \boldsymbol{\Psi}_\tau \right] - \\ \left[\boldsymbol{f}_l(x_2) - \boldsymbol{f}_K(x_2)\right] \left[\sum_{\tau=t'+1}^{T} (\beta \boldsymbol{f}_K)^{\tau-t'-1} \boldsymbol{\Psi}_\tau \right] \Bigg) \end{aligned}$$

Non-stationary models

Theorem

Suppose that

$$u_{k,t}(x_1) = u_{l,t'}(x_2)$$

for $k \in \mathcal{D}/\{K\}$, $l \in \mathcal{D}$, $x_1 \in \mathcal{X}$, $x_2 \in \mathcal{X}$, $1 \leq t' < T$, *and* $t' \leq t \leq T$; *with either* $k \neq l$, *or* $x_1 \neq x_2$, *or* $t' < t$, *or a combination of the three. If either* $p_{k,t}(x_1)/p_{K,t}(x_1) \neq p_{l,t'}(x_2)/p_{K,t'}(x_2)$ *or*

$$[\boldsymbol{f}_k(x_1) - \boldsymbol{f}_K(x_1)]\boldsymbol{\Psi}_{t+1} - [\boldsymbol{f}_l(x_2) - \boldsymbol{f}_K(x_2)]\boldsymbol{\Psi}_{t'+1} \neq 0,$$

then there are no more than $T - t'$ *points in the identified set.*

Proof: Same intuition, different theorem (fundamental theorem of algebra)

Corollary: If $t' = T - 1$, then point identified

Results extend with minor modifications if the reference utility is known

Important recent results show that the common normalization $\boldsymbol{u}_K = 0$ implies that a large class of counterfactuals are not identified, e.g. Kalouptsidi et al (2018)

Suppose that both $u_k(x_1) - u_l(x_2)$ and $\bar{u}_K(x_1) - \bar{u}_K(x_2)$ are known, for some choices $k \in \mathcal{D}/\{K\}$ and $l \in \mathcal{D}$, and known states $x_1 \in \mathcal{X}$ and $x_2 \in \mathcal{X}$; with either $k \neq l$, $x_1 \neq x_2$, or both. Then

$$\begin{aligned} &\ln\left(p_k(x_1)/p_K(x_1)\right) - \ln\left(p_l(x_2)/p_K(x_2)\right) - \Delta^2 u \\ &\qquad = \beta\left[\boldsymbol{f}_k(x_1) - \boldsymbol{f}_K(x_1) - \boldsymbol{f}_l(x_2) + \boldsymbol{f}_K(x_2)\right]\left[\boldsymbol{I} - \beta \boldsymbol{f}_K\right]^{-1} \bar{\boldsymbol{\Psi}}, \end{aligned}$$

with $\Delta^2 u \equiv u_k(x_1) - u_l(x_2) - \bar{u}_K(x_1) + \bar{u}_K(x_2)$ and $\bar{\boldsymbol{\Psi}} \equiv \boldsymbol{\Psi} + \bar{\boldsymbol{u}}_K$ known.

Example: Reward program design and targeting (Daljord, Mela, Yao, Sprigg (2019))

80% of Americans part of some loyalty program (Reuters/Ipsos (2018))

- Reward structure: buy H within period T, get R

Reward programs pose a dynamic choice problem

Mixed evidence on performance of reward programs

- Who should be offered which kind of reward structures?
- Depends on customers reward sensitivity and cost of rewards
- Typically no variation in design parameters within customer in data, so hard to learn reward sensitivity from the data

Three-step approach of Rust (2019)

Data from International Hotels Group where customers were randomized to different $\mathcal{S} = \{H, R, T\}$

Has dynamic programming improved decision making?

1. *Estimate* preferences for stays, rewards, and time using a DDC model of hotel stays
   - → allow for heterogeneous preferences
2. *Optimize*. Choose
   - → *a set of* reward structures $\{\boldsymbol{H}, \boldsymbol{R}, \boldsymbol{T}\}$
   - → *an assignment* of existing customers to those reward structures

   and calculate profits
3. *Validate* optimized reward structure through RCTs

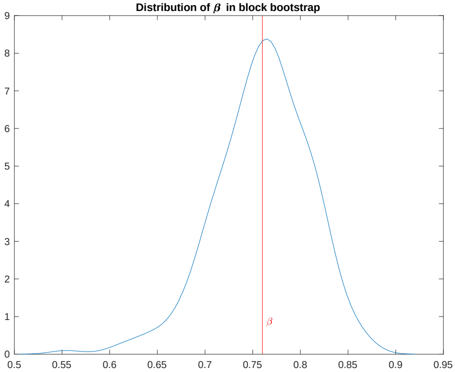
Task: Estimate preferences

A customer chooses $d \in \{0, 1, 2\}$ stays in each of $t \in \{1, \dots, T\}$ weeks

- Utility of a stay $u_{d,t}(x) + \epsilon_{d,t}$, with $\epsilon_{d,t} \sim EV1$
  - → x is number of stays left to the reward, $x \in \{-1, 0, 1, ..., H\}$
  - → $u_{d,t}(0)$ is the utility of reward. Lasts one period
  - → $u_{d,t}(-1)$ is the utility after the reward is collected
- No purchase utility $u_{0,t}(x) = 0$ for all t and x
- Exclusion restriction: $u_{d,t}(x) = u_{d,t}(x')$ for all $x, x' \neq 0$, all d, all t. Substantive
- Observable state transitions $\boldsymbol{f}_d(x)$ fully controlled by choice

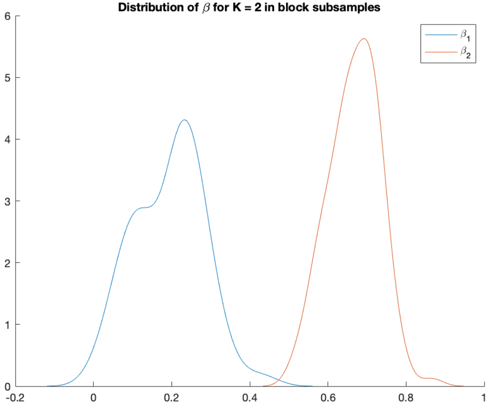
Discount factor estimates

Homogenous model, weekly level, one group



Some parameter estimates

Group fixed effects of Bonhomme et al (2018), two groups



Ex: does the incremental lift cover the costs?

Assign a customer to a particular reward structure or not?

Yes, if it creates incremental profits

- When targeted, then $Pr(d_t = d | \boldsymbol{u}_g, \beta_g, \mathcal{S})$ is a dynamic problem
- When not targeted, then $Pr(d_t = d | \boldsymbol{u}_g, ., \varnothing)$ is a static problem

Then

- Incremental revenue: draw from choice probs, have p and $c(R)$
- Expected cost is simulated share of customers that reach the hurdle
- Computationally straightforward

Repeat across candidate reward structures. Evaluate in RCT

Example: Stockpiling

Understanding the empirically large sales responses to temporary price cuts is a major theme of modern empirical quant marketing literature, e.g. Blattberg & Neslin (1989), Blattberg et al (1995)

- Earliest micro-econometric work in marketing measured promotion effects on demand, e.g., Guadagni & Little (1983)
- Decomposed the total price effect into incidence, brand choice, and quantity choice, e.g., Chiang (1991), Chintagunta (1993)

  → purchase acceleration accounts for 15-50% of promotion effect
- Literature largely based on static models that assume away potential for stock-piling and inter-temporal substitution

  → all purchase acceleration classified as incremental

DDC models of stock-piling

Small extant literature using differentiated products DDC models has assumed discount factors based on interest rates, e.g., $\beta = 0.9995$ at weekly level

→ e.g., Erdem et al (2003) finds that sales responses are mostly due to purchase acceleration and category expansion, and not brand switching

→ e.g., Hendel and Nevo (2005) finds that price elasticities are 30% larger in static models

These findings depend on the assumed discount factor

- Lab studies find consumers are much more impatient, e.g., $\beta = 0.7$ yearly
- Ching & Osborne (2019) uses exclusion restrictions on unobservable inventory, finds $\beta = 0.7$
- The discount factor is identified by the functional form

All stock-piling DDC models impose identifying restrictions

From Daljord & Dubé (2019)

- Two detergents a and b
- Suppose a state is $p = [p_a, p_b]$, with $p_a, p_b \in \{p^h, p^l\}$
- Four states

$$p \in \{p_1, p_2, p_3, p_4\} = \left\{ \begin{array}{c} [p^h, p^h] \\ [p^h, p^l] \\ [p^l, p^h] \\ [p^l, p^l] \end{array} \right\}$$

- Indirect utility depends only on own prices e.g. $u_a(p) = u_a(p_a)$
- Then $u_a(p_1) = u_a(p_2)$, $u_a(p_3) = u_a(p_4)$ etc

Intuition

- Independent variation in other products prices may shift expectations of future prices, but not the current utility of this product

Use the moment conditions directly in estimation

Use two-step approach

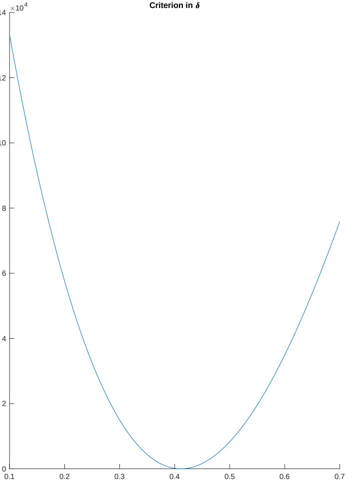
1. Estimate $p(x)$ and $\boldsymbol{f}(x)$ in a first step
2. Construct the moment condition.

$$\ln\left(p_k(x_1)/p_K(x_1)\right) - \ln\left(p_l(x_2)/p_K(x_2)\right) \\ = \beta\left[\boldsymbol{f}_k(x_1) - \boldsymbol{f}_K(x_1) - \boldsymbol{f}_l(x_2) + \boldsymbol{f}_K(x_2)\right]\left[\boldsymbol{I} - \beta\boldsymbol{f}_K\right]^{-1}\boldsymbol{\Psi}$$

3. Minimum distance $\hat{\beta} = \arg\min\, m_T(\beta)' m_T(\beta)$

Easy to program (optional data exercise available)

If point identified, its a standard m-estimator and usual asymptotics apply



Set identification and multiple exclusion restrictions

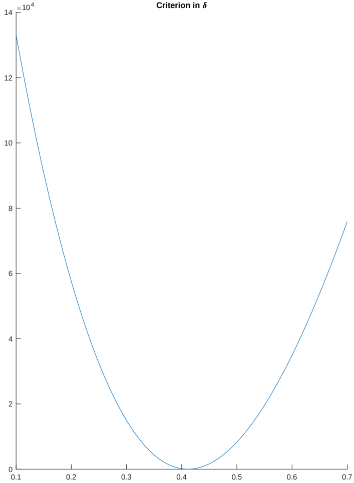
Multiple exclusion restrictions are often available

- with exclusion restrictions on variables
- stationary utilities in finite horizon models

Moment conditions may have multiple zeros individually, but share only one

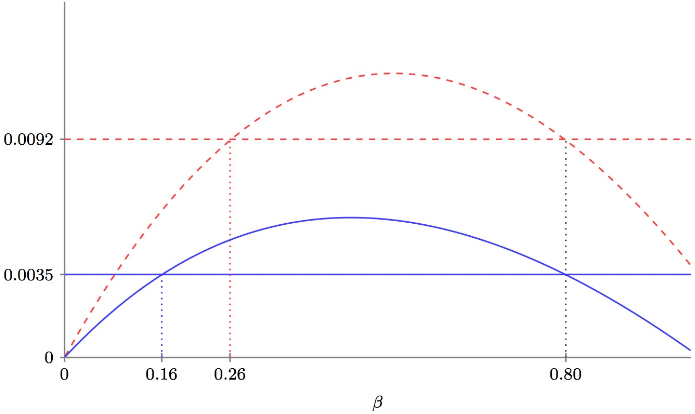
If point identified, its a standard m-estimator and usual asymptotics apply

Looks quadratic around a unique zero, typically asymptotically normal



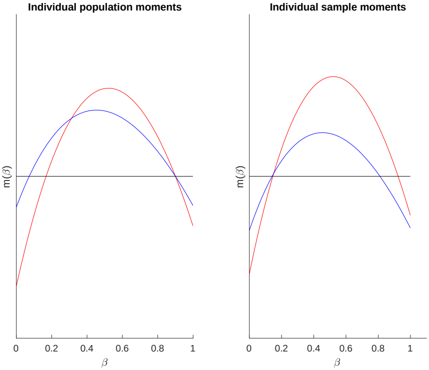
Example of multiple exclusion restrictions

Moment conditions may give set identification individually, but point identification collectively



Finding zero in finite samples

Population moments (left) perturbed by sampling variation (right)



Can't tell from the data if the shared zero is in the lower or upper region

Set identification suggests set estimators

Without point identification, can not find consistent point estimators

But β is still set identified, so consistent set estimators on the table

Point identification a special case of set identification

- Consistent set estimators converge in probability to a set if set identified and a point if point identified

Class of set estimators based on inverting the criterion function

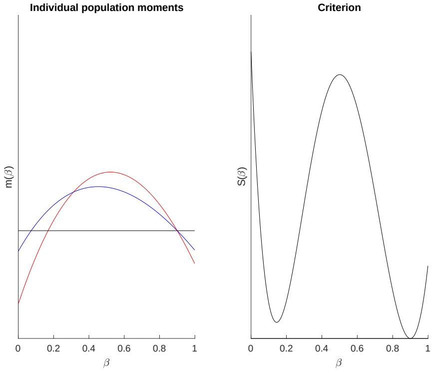
e.g. Romano & Shaikh (2010) and Chernozhukov et al (2007)

- Criterion function S, e.g. Minimum Distance
- Identified set $\mathcal{B} = \arg\min_{\beta \in [0,1)} S_\infty(\beta)$
- Estimator

$$C_n(c) = \{\beta \in [0,1) : a_n S_n(\beta) \leq c\}$$
$$\text{s.t.} \ \lim_{n\to\infty} Pr\{\mathcal{B} \subseteq C_n(c)\} = 1 - \alpha.$$

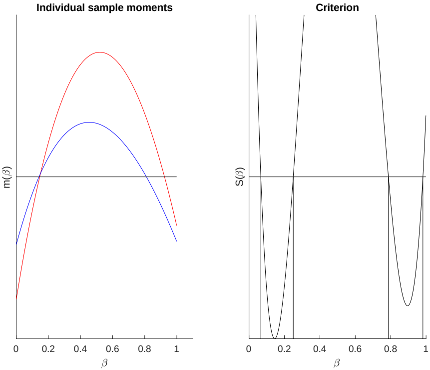
Example: set estimator

Population moments have one shared zero → point identified



Identified set is the zero of the population criterion function

Example: set estimator



Estimated set contains the identified set with probability $1 - \alpha$

- As the set estimator converges to the identified set, here the true discount factor, the subset to the left vanishes

Fang & Wang (2015) *generic identification* result for hyperbolic discount functions is incorrect

In a model of partially naive agents with exclusion restrictions, F&W's Proposition 2 states that $(\delta, \tilde{\delta}, \beta)$ are *generically identified*

- Geometric discounting a special case with present bias parameters $(\delta, \tilde{\delta}) = (1, 1)$

Seemingly a more general result

Fang & Wang (2015) *generic identification* argument

Specifies a DDC model with more exclusion restrictions than free parameters

- F&W Proposition 2 states that the discount function parameters are *generically identified* in the range of the model
  - $\rightarrow$ means that it is identified for *almost all* data that the model can generate
- Appeals to the transversality theorem
  - $\rightarrow$ with more moment conditions than free parameters, there is generically no set of primitives that can rationalize the data
- But by assumption, some set of primitives generated the data
- Therefore, except for a very small subset, there is a unique set of primitives that rationalizes the data

So, generically point identified

Fang & Wang's proof is incorrect and incomplete

Abbring & Daljord (2019) gives two counterexamples

1. The model is nowhere identified in the set of data that can be generated by the model
2. The model is everywhere identified in the set of data that can be generated by the model

Neither example is consistent with the Fang & Wang's claim that the model is identified for almost all data that can be generated by the model

- The proof of the result is incorrect and incomplete and has nothing to say about identification

Identification of the partially naive DDC model an open question

Summary

1. Common intuition on identifying variation can be implemented through exclusion restrictions
   - → Gives a simple moment condition that characterizes the identified set
   - → Important classes of models are point identified. More examples in the paper
   - → Can be used directly in estimation
2. The identified model is falsifiable
   - → Allows you to test the model in an intuitive way
3. Argument extends to present-biased discount functions (joint with Abbring and Iskhakov)