Counterfactual Analysis for Structural Dynamic Discrete Choice Models

Myrto Kalouptsidi, Yuichi Kitamura, Lucas Lima, Eduardo Souza-Rodrigues

Harvard, Yale, Harvard, Toronto

Georgetown University
September 14^{th}, 2021

Dynamic Models in Applied Work

- One of the main reasons to estimate structural models:
  - perform counterfactual analysis
- We focus on Dynamic Discrete Choice models (DDC)
- Fundamental identification problem:
  - Observationally equivalent models
    can generate different behavioral responses
    in a counterfactual environment

Dynamic Models in Applied Work

- Discrete choice models
  - From data on choices, we recover differences in payoffs

Dynamic Models in Applied Work

- Discrete choice models
  - From data on choices, we recover differences in payoffs
- Typical solution in applied work:
  - Model restrictions to select among observationally equivalent models
  - Are counterfactuals invariant to such restrictions?
    - More often than not, no
    - Often little guidance in choosing restrictions

Dynamic Models in Applied Work

- Discrete choice models
  - From data on choices, we recover differences in payoffs
- Typical solution in applied work:
  - Model restrictions to select among observationally equivalent models
  - Are counterfactuals invariant to such restrictions?
    - More often than not, no
    - Often little guidance in choosing restrictions
- How much can be learned about counterfactual outcomes of interest for a large (and empirically relevant) class of counterfactuals?

This Paper

- We develop a tractable framework to address this question
- Avoid assumptions point-identifying payoffs
  - $\Rightarrow$ model parameters and counterfactuals partially identified
- Characterize and compute:
  - Counterfactual behavior (CCPs)
  - Low-dimensional counterfactual outcomes of interest
- Asymptotically valid inference procedure
- Show how different restrictions can lead to informative bounds in applied examples

This Paper

- Offer solution to practitioners for wide range of applications
  - Flexible in terms of model restrictions, counterfactuals, outcomes
  - Tractable: many cases handled by well behaved constrained min/max
  - Paper: guide to practitioners, code

This Paper

- Offer solution to practitioners for wide range of applications
  - Flexible in terms of model restrictions, counterfactuals, outcomes
  - Tractable: many cases handled by well behaved constrained min/max
  - Paper: guide to practitioners, code
- We hope our results can help:
  - Understanding the impact of restrictions on counterfactual outcomes
  - Including confidence sets around counterfactual results

Related Lit

- *Identification of DDC models:*
  - Rust (1994), Magnac and Thesmar (2002), Heckman and Navarro (2007), Pesendorfer and Schmidt-Dengler (2008), Blevins (2014), Heckman, Humphries and Veramendi (2016), Bajari, Nekipelov, Chu, Park (2016), Abbring and Daljord (2020)
- *Estimation of DDC models:*
  - Miller (1984), Pakes (1986), Wolpin (1984), Rust (1987), Hotz and Miller (1993), Aguirregabiria and Mira (2002, 2007), Pakes, Ostrovsky and Berry (2007), Bajari, Benkard, and Levin (2007), Pesendorfer and Schmidt-Dengler (2008), Arcidiacono and Miller (2011), Srisuma and Linton (2012), Aguirregabiria and Magesan (2013, 2018), Kalouptsidi, Scott, and Souza-Rodrigues (2020)
- *Partial Identification of DDC models:*
  - Bajari, Benkard, and Levin (2007), Norets and Tang (2014), Morales, Sheu, and Zahler (2015), Dickstein and Morales (2018), Berry and Compiani (2019)
- *Identification of Counterfactuals in DDC models:*
  - Aguirregabiria (2010), Aguiregabiria and Suzuki (2014); Norets and Tang (2014), Arcidiacono and Miller (2017), Kalouptsidi, Scott, Souza-Rodrigues (2017, 2018)
- *Inference:*
  - Chernoff (1954), Romano and Shaikh (2008, 2012), Andrews and Soares (2010), Kitamura and Stoye (2018), Kaido, Molinari and Stoye (2019)
- *"Marschak's Maxim" principle:*
  - Marschak (1953), Ichimura and Taber (2000, 2002), Heckman (2000, 2010), Manski (2007), Blundell, Browning, and Crawford (2008), Blundell, Kristensen and Matzkin (2014, 2017), Adao, Costinot and Donaldson (2017), Bejara (2018), Mogstad, Santos and Torgovtsky (2018), Kitamura and Stoye (2019), Christensen and Connault (2020)

Outline

1. **Model**
   - Restrictions
   - Identification
2. Counterfactuals
   - Definitions and Parameters of Interest
   - Main Results
   - Example: Firm Entry/Exit Model
3. Inference
4. Empirical Application: Export subsidies

Model: Environment

- Time $t = 0, 1, ..., \infty$
- Agent i in period t chooses action $a_{it} \in \mathcal{A} = \{1, ..., A\}$
- Agent i's state at t is x_{it}, ϵ_{it}
- State transition $F\left(x_{it+1} | a_{it}, x_{it}\right)$
- Per period payoffs

$$\pi\left(a, x_{it}\right) + \varepsilon_{ait}$$

  where $\varepsilon_{it} = \left(\varepsilon_{0it}, ..., \varepsilon_{Ait}\right)$ iid from cdf G, e.g. logit
- Agent chooses sequence of actions to maximize

$$E \sum_{t=0}^{\infty} \beta^t \left(\pi\left(a_{it}, x_{it}\right) + \varepsilon_{ait}\right)$$

Model: Behavior

- Bellman equation:

$$V\left(x_{it},\varepsilon_{it}\right) = \max_{a\in\mathcal{A}}\left\{\pi\left(a,x_{it}\right)+\varepsilon_{ait}+\beta E\left[V\left(x_{it+1},\varepsilon_{it+1}\right)|a,x_{it}\right]\right\}$$

Model: Behavior

- Bellman equation:

$$V(x_{it}, \varepsilon_{it}) = \max_{a \in \mathcal{A}} \left\{ \pi(a, x_{it}) + \varepsilon_{ait} + \beta E\left[V(x_{it+1}, \varepsilon_{it+1}) \mid a, x_{it} \right] \right\}$$

- Conditional value function

$$v_a(x_{it}) = \pi(a, x_{it}) + \beta E\left[V(x_{it+1}, \varepsilon_{it+1}) \mid a, x_{it} \right]$$

Model: Behavior

- Bellman equation:

$$V\left(x_{it},\varepsilon_{it}\right) = \max_{a\in\mathcal{A}}\left\{\pi\left(a,x_{it}\right)+\varepsilon_{ait}+\beta E\left[V\left(x_{it+1},\varepsilon_{it+1}\right)|a,x_{it}\right]\right\}$$

- Conditional value function

$$v_a\left(x_{it}\right)=\pi\left(a,x_{it}\right)+\beta E\left[V\left(x_{it+1},\varepsilon_{it+1}\right)|a,x_{it}\right]$$

- Conditional choice probabilities (CCPs)

$$p_a\left(x_{it}\right)=\Pr\left(v_a\left(x_{it}\right)+\varepsilon_{ait}\geq v_j\left(x_{it}\right)+\varepsilon_{jit},\ \forall\ j\in\mathcal{A}\right)$$

Model: Behavior

- Bellman equation:

$$V\left(x_{it},\varepsilon_{it}\right) = \max_{a\in\mathcal{A}}\left\{\pi\left(a,x_{it}\right)+\varepsilon_{ait}+\beta E\left[V\left(x_{it+1},\varepsilon_{it+1}\right)|a,x_{it}\right]\right\}$$

- Conditional value function

$$v_a\left(x_{it}\right)=\pi\left(a,x_{it}\right)+\beta E\left[V\left(x_{it+1},\varepsilon_{it+1}\right)|a,x_{it}\right]$$

- Conditional choice probabilities (CCPs)

$$p_a\left(x_{it}\right)=\Pr\left(v_a\left(x_{it}\right)+\varepsilon_{ait}\geq v_j\left(x_{it}\right)+\varepsilon_{jit},\ \forall\ j\in\mathcal{A}\right)$$

- E.g. if logit shocks

$$p_a\left(x_{it}\right)=\frac{\exp\left(v_a\left(x_{it}\right)\right)}{\sum_{j\in\mathcal{A}}\exp\left(v_j\left(x_{it}\right)\right)}$$

Model (Non)Identification

- Assume for now: β and G are known
- Model Identification:
  - the main object of interest is π
  - data: choices at different states (p, F)

Model (Non)Identification

- Assume for now: β and G are known
- Model Identification:
  - the main object of interest is π
  - data: choices at different states (p, F)
- Intuition for non-identification:
  - π has AX parameters
  - but there are only $(A-1)\,X$ observed CCPs p
  - thus, X free parameters, or X restrictions required to obtain π

Model (Non)Identification

- Assume for now: β and G are known
- Model Identification:
  - the main object of interest is π
  - data: choices at different states (p, F)
- Intuition for non-identification:
  - π has AX parameters
  - but there are only $(A-1)\,X$ observed CCPs p
  - thus, X free parameters, or X restrictions required to obtain π
- Example: binary choice $\{0, 1\}$
  - want to obtain π_0, π_1
  - observe p_1 (p_0 is uninformative since $p_0 = 1 - p_1$)

Model (Non)Identification

- Underidentification problem: π_1 is an affine transformation of π_0, ▸ argument

$$\pi_1 = M_1 \pi_0 + b_1 (p)$$

where b_1 known function of p and

Model (Non)Identification

- Underidentification problem: π_1 is an affine transformation of π_0, ▸ argument

$$\pi_1 = M_1 \pi_0 + b_1 \left(p \right)$$

where b_1 known function of p and

$$M_1 = \left(I - \beta F_1 \right) \left(I - \beta F_0 \right)^{-1}$$

Model (Non)Identification

- Underidentification problem: π_1 is an affine transformation of π_0, ▸ argument

$$\pi_1 = M_1 \pi_0 + b_1 (p)$$

where b_1 known function of p and

$$M_1 = (I - \beta F_1)(I - \beta F_0)^{-1}$$

- To *point*-identify π, set $\pi_0 = 0$ or some other X restrictions

Model (Non)Identification

- Underidentification problem: π_1 is an affine transformation of π_0, ▸ argument

$$\pi_1 = M_1 \pi_0 + b_1(p)$$

  where b_1 known function of p and

$$M_1 = (I - \beta F_1)(I - \beta F_0)^{-1}$$

- To *point*-identify π, set $\pi_0 = 0$ or some other X restrictions
- Example of logit:
  - Static model ($M_1 = I$),

$$\pi_1 = \pi_0 + \underbrace{\log p_1 - \log p_0}_{b_1(p)}$$

Model (Non)Identification

- Underidentification problem: π_1 is an affine transformation of π_0, ▸ argument

$$\pi_1 = M_1 \pi_0 + b_1(p)$$

where b_1 known function of p and

$$M_1 = (I - \beta F_1)(I - \beta F_0)^{-1}$$

- To *point*-identify π, set $\pi_0 = 0$ or some other X restrictions
- Example of logit:
  - Static model ($M_1 = I$),

$$\pi_1 = \pi_0 + \underbrace{\log p_1 - \log p_0}_{b_1(p)}$$

  - Dynamic model,

$$\pi_1 = M_1 \pi_0 + \underbrace{\log p_1 - M_1 \log p_0}_{b_1(p)}$$

Non-Identification of Payoffs

- Generally, take a reference action J
- Easy to show π_a, all $a \neq J$, is an affine transformation of π_J

$$\pi_a = M_a \pi_J + b_a(p)$$

where

$$M_a = (I - \beta F_a)(I - \beta F_J)^{-1}$$

and $b_a(p)$ is a known function of p ▸ proof

- Stack it for compact notation to

$$\mathbf{M}\pi = b_{-J}(p)$$

▸ stacking

Model Restrictions

- Instead impose additional $d \leq X$ restrictions:

$$R^{eq}\pi = r^{eq}$$

- Examples:
  - Exclusion Restrictions
  - Parametric Assumptions
  - "Normalizations"

Model Restrictions

- Instead impose additional $d \leq X$ restrictions:

$$R^{eq}\pi = r^{eq}$$

- Examples:
  - Exclusion Restrictions
  - Parametric Assumptions
  - "Normalizations"
- Impose m inequality restrictions:

$$R^{iq}\pi \leq r^{iq}$$

- Examples: monotonicity, concavity, supermodularity

Model Identification

- The sharp identified set for the payoffs π is

$$\Pi^I = \left\{ \pi \in \mathbb{R}^{AX} : \mathbf{M}\pi = b_{-J}(p), R^{eq}\pi = r^{eq}, R^{iq}\pi \leq r^{iq} \right\}$$

- Π^I is a convex polyhedron of dimension $X - d$
- π is point identified when $[\mathbf{M}', R^{eq\prime}]'$ is full rank

Outline

1. Model
    - Restrictions
    - Identification
2. **Counterfactuals**
    - **Definitions and Parameters of Interest**
    - Main Results
    - Example: Firm Entry/Exit Model
3. Inference
4. Empirical Application: Export subsidies

Counterfactuals: Definitions

- A counterfactual is a transformation that creates the new primitives $\left\{\widetilde{\mathcal{A}}, \widetilde{\mathcal{X}}, \widetilde{F}, \widetilde{\pi}, \widetilde{\beta}, \widetilde{G}\right\}$: ▸ examples from the applied lit

  - $\widetilde{\mathcal{A}}$ and $\widetilde{\mathcal{X}}$ are new sets of actions and states
  - $\widetilde{F}$ is the new transition matrix
  - $\widetilde{\pi}$ is the new payoff, $\widetilde{\pi} = \mathcal{H}\pi + g$
  - $\widetilde{\beta}$ is the new discount factor
  - $\widetilde{G}$ is the new distribution of ε

- These new primitives imply a new optimal behavior $\widetilde{p}$

Target Parameters of Interest

- In many cases researcher not interested in $\widetilde{p}$ per se
  - $\widetilde{p}$ can be a huge object, hard to interpret
  - rather, we might care about some outcome of $\widetilde{p}$
- Examples:
  - Long-run dynamics
  - Welfare measures (CS and/or profits, gov expenditures)
  - Average treatment effects
- Let $\theta \in \Theta \subset \mathbb{R}^n$ be a (low-dimensional) object of interest:

$$\theta = f(\widetilde{p}, \pi; p, F)$$

Outline

1. Model
   - Restrictions
   - Identification
2. **Counterfactuals**
   - Definitions and Parameters of Interest
   - **Main Results**
   - Example: Firm Entry/Exit Model
3. Inference
4. Empirical Application: Export subsidies

Main Results

- Baseline world (our data):

$$\mathbf{M}\pi = b_{-J}(p)$$

- Analogous relation holds in counterfactual world:

$$\tilde{\mathbf{M}}\tilde{\pi} = \tilde{b}_{-J}(\tilde{p})$$

Main Results

- Baseline world (our data):

$$\mathbf{M}\pi = b_{-J}(p)$$

- Analogous relation holds in counterfactual world:

$$\tilde{\mathbf{M}}\tilde{\pi} = \tilde{b}_{-J}(\tilde{p})$$

and substituting $\tilde{\pi} = \mathcal{H}\pi + g$,

$$\tilde{\mathbf{M}}\left(\mathcal{H}\pi + g\right) = \tilde{b}_{-J}(\tilde{p})$$

or

$$\left(\tilde{\mathbf{M}}\mathcal{H}\right)\pi = \tilde{b}_{-J}(\tilde{p}) - \tilde{\mathbf{M}}g$$

Main Results

Proposition

The sharp identified set for the counterfactual CCP $\widetilde{p}$ is

$$\widetilde{\mathbf{P}}^{I} = \left\{ \begin{array}{c} \widetilde{p} \in \widetilde{\mathbf{P}} : \exists \pi \in \mathbb{R}^{AX} \text{ such that} \\ \mathbf{M}\pi = b_{-J}(p), \\ R^{eq}\pi = r^{eq}, \;\; R^{iq}\pi \leq r^{iq}, \\ (\widetilde{\mathbf{M}}\mathcal{H})\pi = \widetilde{b}_{-J}(\widetilde{p}) - \widetilde{\mathbf{M}}g \end{array} \right\}$$

$\widetilde{\mathbf{P}}^{I}$ *is a connected manifold*

The dimension of $\widetilde{\mathbf{P}}^{I}$ *is given by* rank(CR)*, where* C *depends on the counterfactual and* R *depends on the restrictions*

When Π^{I} *is bounded,* $\widetilde{\mathbf{P}}^{I}$ *is compact*

Main Results

Proposition

The sharp identified set for θ is

$$\Theta^I = \left\{ \begin{array}{c} \theta \in \Theta : \exists \left(\widetilde{p}, \pi\right) \in \widetilde{\mathbf{P}} \times \mathbb{R}^{AX} \textit{ such that} \\ \theta = f\left(\widetilde{p}, \pi, p, F\right), \ \mathbf{M}\pi = b_{-J}(p), \\ R^{eq}\pi = r^{eq}, \ R^{iq}\pi \leq r^{iq}, \\ (\widetilde{\mathbf{M}}\mathcal{H})\pi = \widetilde{b}_{-J}\left(\widetilde{p}\right) - \widetilde{\mathbf{M}}g \end{array} \right\}$$

When f is continuous, Θ^I is a connected set

When Π^I is bounded, Θ^I is compact

When θ is a scalar, Θ^I is a finite interval $\left[\theta^L, \theta^U\right]$

Main Results

- When θ is a scalar, the bounds of the interval Θ^I can be computed by the following optimization problems:

$$\theta^U \equiv \max_{(\widetilde{p},\pi)} f\left(\widetilde{p},\pi;p,F\right) \quad \text{and} \quad \theta^L \equiv \min_{(\widetilde{p},\pi)} f\left(\widetilde{p},\pi;p,F\right)$$

subject to

$$\begin{aligned} \mathbf{M}\,\pi &= b_{-J}(p), \\ R^{eq}\,\pi &= r^{eq}, \\ R^{iq}\,\pi &\leq r^{iq}, \\ (\widetilde{\mathbf{M}}\mathcal{H})\,\pi &= \widetilde{b}_{-J}\left(\widetilde{p}\right) - \widetilde{\mathbf{M}}g \end{aligned} \tag{1}$$

Main Results

- When θ is a scalar, the bounds of the interval Θ^I can be computed by the following optimization problems:

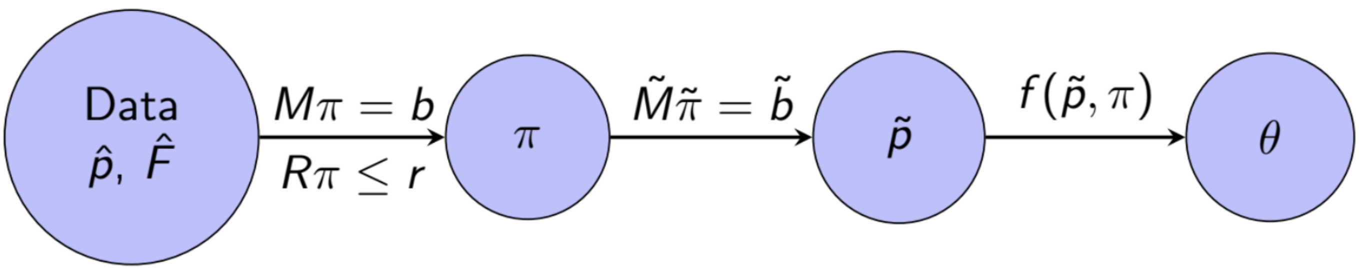
$$\theta^U \equiv \max_{(\widetilde{p},\pi)} f\left(\widetilde{p}, \pi; p, F\right) \quad \text{and} \quad \theta^L \equiv \min_{(\widetilde{p},\pi)} f\left(\widetilde{p}, \pi; p, F\right)$$

subject to

$$\begin{array}{rcl} \mathbf{M}\,\pi & = & b_{-J}(p), \\ R^{eq}\,\pi & = & r^{eq}, \\ R^{iq}\,\pi & \leq & r^{iq}, \\ (\widetilde{\mathbf{M}}\mathcal{H})\,\pi & = & \widetilde{b}_{-J}\left(\widetilde{p}\right) - \widetilde{\mathbf{M}}g \end{array} \tag{1}$$

- Can solve using standard software (eg, knitro)
- We develop a stochastic search algorithm when gradient of f costly to compute and/or high-dimensional state space ‣ Implementation Details
- Can extend to unknown β and/or unknown G up to parametric form ‣ Details

Main Results



Outline

1. Model
   - Restrictions
   - Identification
2. **Counterfactuals**
   - Definitions and Parameters of Interest
   - Main Results
   - **Example: Firm Entry/Exit Model**
3. Inference
4. Empirical Application: Export subsidies

Example: Firm Entry/Exit Model

- Actions: $\mathcal{A} = \{\text{out}, \text{in}\} = \{0, 1\}$

- States:
  - current status, $a_{-1} \equiv a_{t-1} \in \{0, 1\}$
  - market demand, $w_t \in \left\{w^l, w^h\right\}$

Example: Firm Entry/Exit Model

- Actions: $\mathcal{A} = \{\text{out}, \text{in}\} = \{0, 1\}$
- States:
  - current status, $a_{-1} \equiv a_{t-1} \in \{0, 1\}$
  - market demand, $w_t \in \left\{w^l, w^h\right\}$
- Payoff:

$$\pi(a, a_{-1}, w) = \begin{cases} a_{-1} \underbrace{(vp - fc)}_{\text{variable profit-FC}} + (1 - a_{-1}) \left(vp - fc \underbrace{-ec}_{\text{entry cost}}\right) & \text{if } a = 1 \text{ (in)} \\ a_{-1} \underbrace{s}_{\text{scrap value}} + (1 - a_{-1}) \underbrace{0}_{\text{if already out, 0}} & \text{if } a = 0 \text{ (out)} \end{cases}$$

Example: Firm Entry/Exit Model

- Need $X = AW = 4$ restrictions to point-identify π
  - $\pi_0\,(0, w) = 0$ (and vp known)
  - **Usually either $fc = 0$ or $s = 0$**

Example: Firm Entry/Exit Model

- Need $X = AW = 4$ restrictions to point-identify π
  - $\pi_0(0, w) = 0$ (and vp known)
  - **Usually either $fc = 0$ or $s = 0$**
- What can go wrong?
  - mistakenly setting $s = 0 \Rightarrow$
    - firm value lower $\Rightarrow$ entry cost downward biased (even sign reversals)
  - mistakenly setting $fc = 0 \Rightarrow$
    - per period profits upward biased $\Rightarrow$ entry cost/scrap values blown up

Illustration

Example: Firm Entry/Exit Model

- For partial identification, we instead impose some basic restrictions:

  1. $fc \geq 0$, $ec \geq 0$, (always assume $\pi_0(0, w) = 0$ and vp known)
  2. $vp - fc \leq ec \leq \frac{\mathbb{E}[vp - fc]}{1-\beta}$, and $\pi_1(1, w^h) \geq \pi_1(1, w^l)$
  3. s does not depend on w

Example: Firm Entry/Exit Model

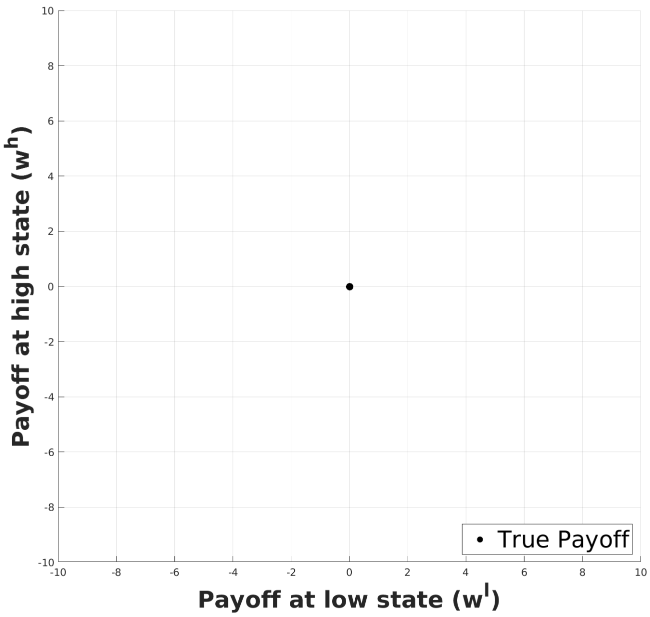


Figure: Payoff of choosing out when already out

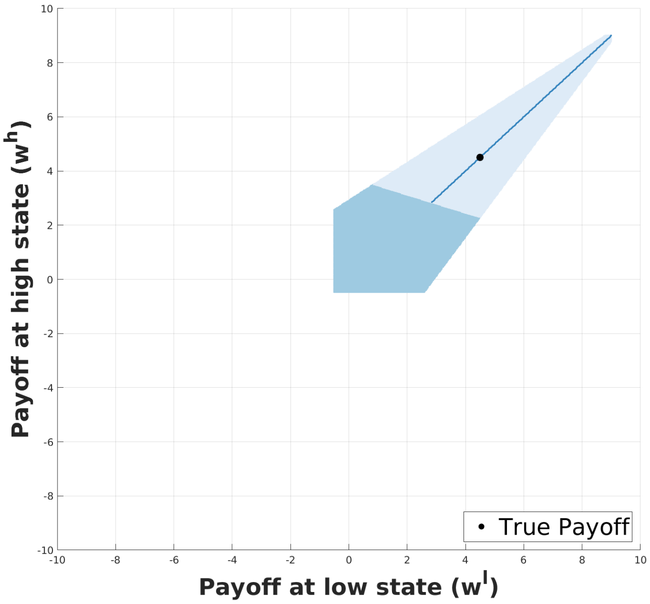
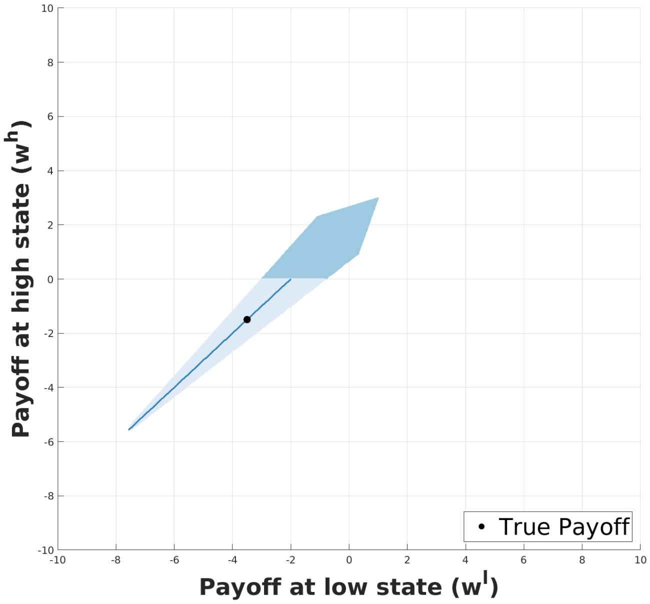


Figure: Payoff of exiting (s)

Example: Firm Entry/Exit Model



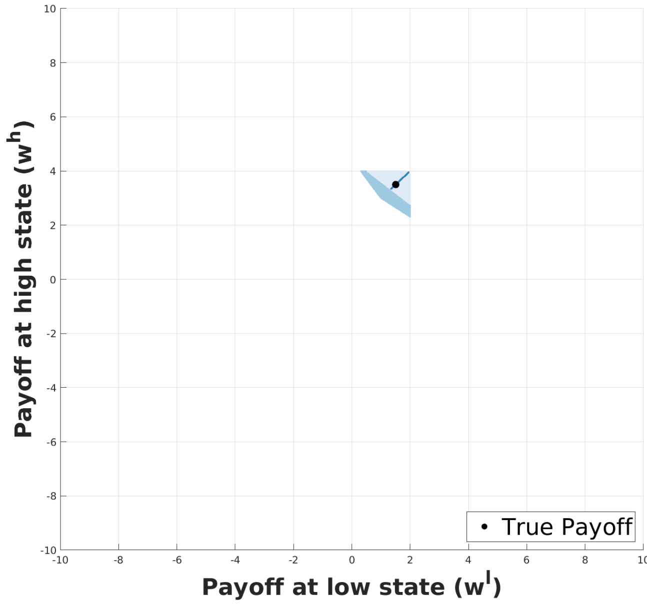


Figure: payoff of entering $(vp - fc - ec)$ Figure: payoff of staying in $(vp - fc)$

Example: Firm Entry/Exit Model

- Counterfactual: entry subsidy
  - Reduce entry cost *ec* by 20%

- Formally, $\widetilde{\pi} = \mathcal{H}\pi$, with

$$\widetilde{\pi}_0 = \pi_0,$$

$$\widetilde{\pi}_1 = \begin{bmatrix} vp - fc - \tau \times ec \\ vp - fc \end{bmatrix}$$

with $\tau = 0.8$ Details

- $\widetilde{\mathbf{P}}^I$ is two-dimensional, even in the absence of Restrictions 1–3

Example: Firm Entry/Exit Model

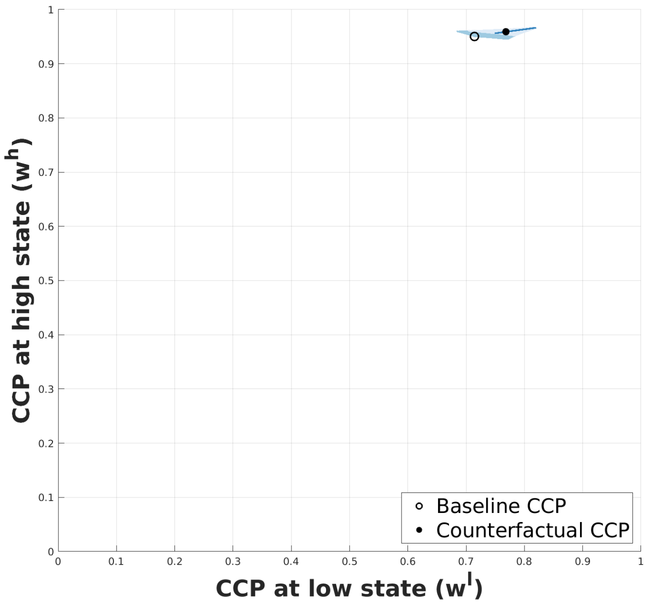


Figure: Prob entry

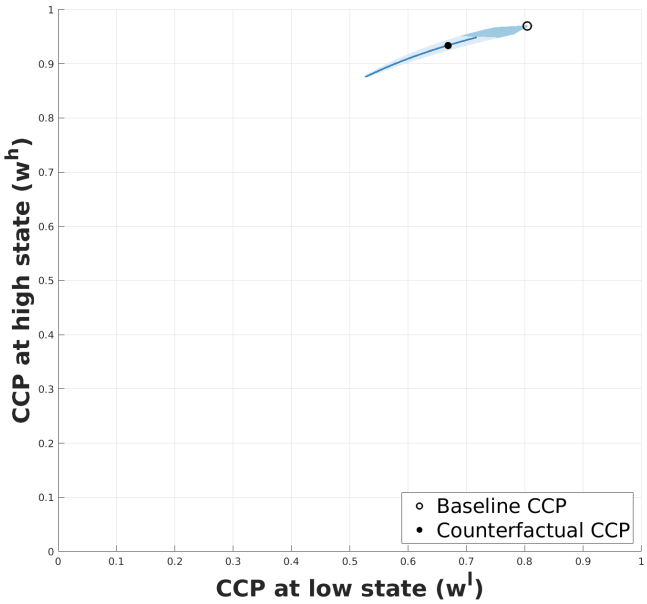


Figure: Prob stay

Example: Firm Entry/Exit Model

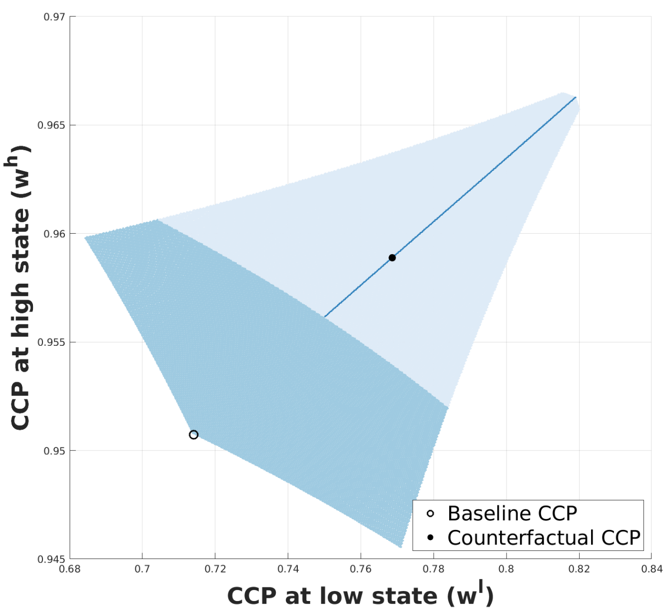


Figure: Prob Entry

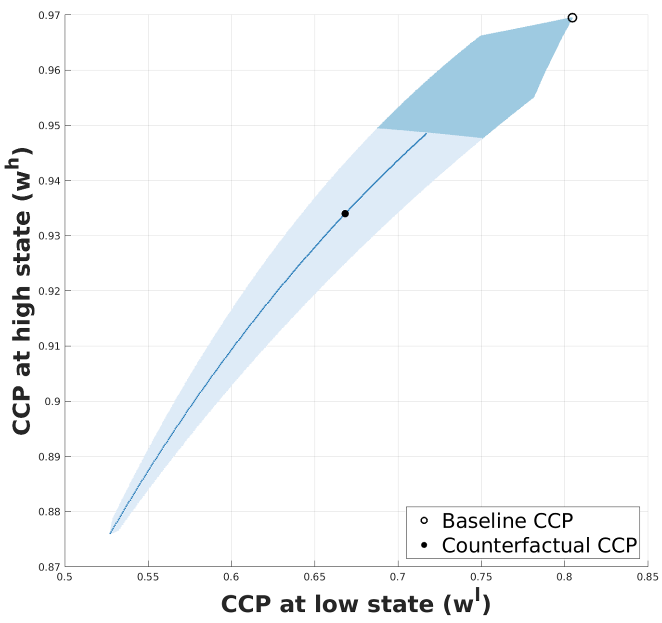


Figure: Prob Stay

Example: Firm Entry/Exit Model

Outcome of Interest	Target parameter	Sharp Identified Sets		
	True	Restriction 1	Restrictions 1–2	Restrictions 1–3
Change in Prob. of Being Active	-0.0638	[-0.1235, 0.0000]	[-0.1235, -0.0341]	[-0.1235, -0.0421]
	True	Restriction 1	Restrictions 1–2	Restrictions 1–3
Change in Consumer Surplus	-0.0875	[-0.1735, 0.0000]	[-0.1735, -0.0474]	[-0.1735, -0.0573]
	True	Restriction 1	Restrictions 1–2	Restrictions 1–3
Change in the Value of the Firm	0.9513	[0.0000, 1.8229]	[0.4489, 1.8229]	[0.6388, 1.8229]

graph

Outline

1. Model
   - Restrictions
   - Identification
2. Counterfactuals
   - Definitions and Parameters of Interest
   - Main Results
   - Example: Firm Entry/Exit Model
3. **Inference**
4. Empirical Application: Export subsidies

Inference: Partially-Identified Counterfactual

- Use data on actions and states $\{a_{it}, x_{it}\}$, $i = 1, ..., N$ and $t = 1, ..., T$
  - Estimate $\hat{p}_N$ and $\hat{F}_N$ (e.g. frequencies)
  - Asymptotics in N
- We build on Kitamura and Stoye (2018, *Ecta*)
  - Construct confidence sets by inverting a test

Inference: Partially-Identified Counterfactual

- Fix counterfactual outcome $\theta = \theta_0$
- Consider the optimization problem

$$J(\theta_0) := \min_{\substack{(\widetilde{p},\pi)\,:\, R^{iq}\pi \leq r^{iq}, \\ \mathcal{R}^{eq}(\theta_0,\pi,\widetilde{p};p,F)=0}} [b_{-J}(p) - \mathbf{M}\pi]' \, \Omega \, [b_{-J}(p) - \mathbf{M}\pi]$$

- If constraints satisfied at θ_0, then $J(\theta_0) = 0$, otherwise $J(\theta_0) > 0$
- Thus the hypothesis $H_0 : \theta = \theta_0$ is equivalent to $H_0 : J(\theta_0) = 0$
- To test it, use empirical counterpart (plug in estimators $\hat{p}_N, \hat{F}_N$)
- Use subsampling to obtain critical values theorem implementation monte carlo results

Outline

1. Model
   - Restrictions
   - Identification
2. Counterfactuals
   - Definitions and Parameters of Interest
   - Main Results
   - Example: Firm Entry/Exit Model
3. Inference
4. **Empirical Application: Export subsidies**

Empirical Application: Export supply and subsidies

- Dynamic model of export behavior by Das, Roberts, and Tybout (2007, *Ecta*)
  - Policy question: horserace of different export promoting subsidies

Empirical Application: Export supply and subsidies

- Dynamic model of export behavior by Das, Roberts, and Tybout (2007, *Ecta*)
  - Policy question: horserace of different export promoting subsidies
- Same model as firm entry/exit example:
  - choice: enter/exit in exporting
  - states: lagged choice, market variables (exchange rates, demand/cost)

Empirical Application: Export supply and subsidies

- Dynamic model of export behavior by Das, Roberts, and Tybout (2007, *Ecta*)
  - Policy question: horserace of different export promoting subsidies
- Same model as firm entry/exit example:
  - choice: enter/exit in exporting
  - states: lagged choice, market variables (exchange rates, demand/cost)
- Estimation:
  - Use Colombian plant-level panel data
  - Impose:
    - $\pi_0(0) = 0$ and vp known (estimated from trade model)
    - $s = 0$, fc, ec invariant over states
    - $\Rightarrow$ over-identified
- Here: Impose Restrictions 1–3

Empirical Application: Export supply and subsidies

- Three counterfactual policies:
  1. export revenues subsidies (2%)
  2. fixed costs subsidies (28%)
  3. entry costs subsidies (25%)

Empirical Application: Export supply and subsidies

- Three counterfactual policies:

  1. export revenues subsidies (2%)
  2. fixed costs subsidies (28%)
  3. entry costs subsidies (25%)

- Outcome of interest θ: is a benefit–cost ratio

  - long-run returns to subsidies:

$$\frac{\text{increase in exports}}{\text{government spending}}$$

  - *highly non-linear* in $\tilde{p}, \pi$ expressions

Empirical Application: Export supply and subsidies

	Restriction 1	Restrictions 1–2	Restrictions 1–3
		Revenue Subsidies	
Estimated Identified Set	15.13	15.13	15.13
90% Confidence Interval	(11.15, 18.90)	(11.15, 18.90)	(11.15, 18.90)
		Fixed Costs Subsidies	
Estimated Identified Set	[8.41, 30.82]	[11.11, 13.35]	[11.92, 12.60]
90% Confidence Interval	(7.17, 35.28)	(8.75, 15.13)	(9.65, 14.46)
		Entry Costs Subsidies	
Estimated Identified Set	[5.55, 24.04]	[7.85, 17.28]	[8.88, 9.36]
90% Confidence Interval	(3.42, 34.96)	(6.03, 24.15)	(6.94, 16.04)

Findings

- DRT find that

$$\text{revenue} > \text{FC} > \text{entry subsidy}$$

- We find that
  - Restrictions 1-3 DRT's ranking
  - Restrictions 1-2: confirm only

$$\text{revenue} > \text{FC}$$

  - Ranking hinges on assumption that s is invariant over states
    - intuition: scrap values that change over states reduce the inefficiency of entry cost subsidies

Conclusion

- This paper
  - Develop unified framework to investigate how much one can learn about counterfactual outcomes in DDC models for large class of counterfactuals
  - Computationally tractable procedure for estimation and inference
- Aid practitioners in terms of
  - assess which empirical results survive under basic restrictions
  - understand impact of common restrictions imposed on counterfactuals
  - include confidence sets around their counterfactual results
- Extensions to dynamic games

Appendix

AM Lemma

- Proof of AM Lemma:

$$V(x) = \int \max_{a \in A} \{v_a(x) + \varepsilon(a)\} \, dG(\varepsilon)$$

$$\begin{aligned}
&= \int \max_{a \in A} \{v_{\mathbf{1}}(x) - v_j(x) + \varepsilon(1), ..., v_J(x) - v_j(x) + \varepsilon(J)\} \, dG(\varepsilon) + v_j(x) \\
&= \int \max_{a \in A} \{\phi_{\mathbf{1}j}(p(x)) + \varepsilon(1), ..., \phi_{Jj}(p(x)) + \varepsilon(J)\} \, dG(\varepsilon) + v_j(x) \\
&= \psi_j(p(x)) + v_j(x)
\end{aligned}$$

Model Non-Identification: Basic Equation

Base equations people use to recover unknowns:

- conditional value functions

$$v_1 = \pi_1 + \beta F_1 V$$

$$v_2 = \pi_2 + \beta F_2 V$$

- Hotz-Miller inversion (ϕ known function of p):

$$v_1 - v_2 = \phi(p)$$

- Arcidiacono-Miller lemma (ψ known function of p):

$$V = v_1 + \psi_1(p)$$

$$V = v_2 + \psi_2(p)$$

Model Non-Identification: Basic Equation

- System:

$$\begin{bmatrix} I & 0 & \beta F_1 & -I & 0 \\ 0 & I & \beta F_2 & 0 & -I \\ 0 & 0 & 0 & I & -I \\ 0 & 0 & I & -I & 0 \\ 0 & 0 & I & 0 & -I \end{bmatrix} \begin{bmatrix} \pi_1 \\ \pi_2 \\ V \\ v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \phi(p) \\ \psi_1(p) \\ \psi_2(p) \end{bmatrix}$$

Model Non-Identification: Basic Equation

- System:

$$\begin{bmatrix} I & 0 & \beta F_1 & -I & 0 \\ 0 & I & \beta F_2 & 0 & -I \\ 0 & 0 & 0 & I & -I \\ 0 & 0 & I & -I & 0 \\ 0 & 0 & I & 0 & -I \end{bmatrix} \begin{bmatrix} \pi_1 \\ \pi_2 \\ V \\ v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \phi(p) \\ \psi_1(p) \\ \psi_2(p) \end{bmatrix}$$

- Can obtain π_1 as a function of π_2:

$$\pi_1 = A_1 \pi_2 + b_1(p)$$

where A_1, b_1 functions of F and p Go Back

Model Non-Identification: Basic Equation

- We have:

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

  - Pick $a = J$
$$V = (I - \beta F_J)^{-1} (\pi_J + \psi_J)$$
  - Substitute for V:
$$\pi_a = (I - \beta F_a)(I - \beta F_J)^{-1} (\pi_J + \psi_J) - \psi_a$$
  - So that
$$M_a = (I - \beta F_a)(I - \beta F_J)^{-1}$$
$$b_a = M_a \psi_J - \psi_a$$

Go Back

Model Identification

- It is convenient to stack the relations

$$\pi_a = M_a \pi_J + b_a(p)$$

as follows:

$$\pi_{-J} = \mathbf{M}_{-J} \pi_J + b_{-J}(p)$$

or

$$\pi_{-J} - \mathbf{M}_{-J} \pi_J = b_{-J}(p)$$

or

$$[I, -\mathbf{M}_{-J}] \begin{bmatrix} \pi_{-J} \\ \pi_J \end{bmatrix} = b_{-J}(p)$$

- In the end we obtain

$$\mathbf{M}\pi = b_{-J}(p)$$

where $\mathbf{M} = [I, -M_{-J}]$, and $\pi = [\pi'_{-J}, \pi'_J]'$ Go Back

Examples of Counterfactuals

- "*Change Set of Actions*": $\widetilde{\mathcal{A}}$
  - Rust and Phelan (1997): eliminate social security
  - Gilleskie (1998): restrict access to medical care in first days of illness
  - Crawford and Shum (2005): switching drug not possible
  - Keane and Wolpin (2010): eliminate welfare program
  - Keane and Merlo (2010): eliminate private jobs option for politicians after leaving congress
  - Rosenzweig and Wolpin (1993): add farmer insurance option
- "*Change Transitions*": $\widetilde{F}$
  - Hendel and Nevo (2007): change consumer price long-run mean
  - Collard-Wexler (2013): change volatility of demand shocks
  - Kalouptsidi (2014): impact of time to build
  - Chan, Hamilton, and Papageorge (2015): change individual health status
- "*Change Discount Factor*": $\widetilde{\beta}$
  - Conlon (2012): myopic consumers in the LCD TV industry

Examples of Counterfactuals

- "*Change Payoffs*": $\widetilde{\pi} = \mathcal{H}\pi + g$
  - "Additive Change":
    - Keane and Wolpin (1997): college tuition subsidies
    - Schiraldi (2011), Li and Wei (2014): car scrappage subsidies
    - Duflo, Hanna and Ryan (2012): teacher bonus incentives
  - "Proportional Change":
    - Das, Roberts and Tybout (2007), Varela (2013), Lin (2015), and Igami (2015): percentage changes (subsidies) in entry/sunk costs
  - "Change in Types":
    - Keane and Wolpin (2010): replace parameters of minorities by those of white women (investigate racial-gap in labor market)
    - Eckstein and Lifshitz (2011): use preference/costs parameters of 1955's cohort for other cohorts
    - Ryan (2012): replace the after-change (CAAA) entry cost by the before policy entry cost (cement industry)

Go Back

Unknown G

- Can extend to case where β is unknown and/or G is known only up to parametric form
- If $\beta \in \mathcal{B} \subset [0,1)$, $\mathcal{B}$ compact and convex
  - If $G(\varepsilon;\lambda)$, $\lambda \in \Lambda \subset \mathbb{R}^L$, Λ compact and convex
  - Then upper bounds is

$$\theta^U \equiv \max_{(\widetilde{p},\pi,\beta,\lambda)\in\widetilde{\mathbf{P}}\times\mathbb{R}^{AX}\times\mathcal{B}\times\Lambda} f\left(\widetilde{p},\pi,\beta,\lambda;p,F\right) \tag{2}$$

subject to

$$\begin{array}{rcl} \mathbf{M}(\beta)\,\pi & = & b_{-J}(p;\beta,\lambda), \\ R^{eq}\,\pi & = & r^{eq}, \\ R^{iq}\,\pi & \leq & r^{iq}, \\ (\widetilde{\mathbf{M}}(\beta)\mathcal{H})\,\pi & = & \widetilde{b}_{-J}\left(\widetilde{p};\beta,\lambda\right) - \widetilde{\mathbf{M}}(\beta)g \end{array} \tag{3}$$

  - More costly, but computationally feasible ▸ Go Back

Example: Firm Entry/Exit Model

- Usually either $fc = 0$ or $s = 0$
- What can go wrong? (From KSS)

States: (a_{-1}, w)	**True Profit**	**Profit 1** scrap value $= 0$	**Profit 2** fixed cost $= 0$
$a =$ out			
$\pi(a, a_{-1} = 0, w_H) = 0$	0	0	0
$\pi(a, a_{-1} = 0, w_L) = 0$	0	0	0
$\pi(a, a_{-1} = 1, w_H) = s$	10	0	120
$\pi(a, a_{-1} = 1, w_L) = s$	10	0	120
$a =$ in			
$\pi(a, a_{-1} = 0, w_H) = -ec$	-9	0.5	-113.5
$\pi(a, a_{-1} = 0, w_L) = -ec$	-9	0.5	-113.5
$\pi(a, a_{-1} = 1, w_H) = vp(w_H) - fc$	8	7.5	13.5
$\pi(a, a_{-1} = 1, w_L) = vp(w_L) - fc$	-5.33	-5.83	0.167

Example: Firm Entry/Exit Model

Counterfactual: additive entry subsidy $\tilde{\pi} = \pi + g$ (From KSS)

States: (a_{-1}, w)	**Baseline**	**True CF**	**Estimated CF** scrap value $= 0$	**Estimated CF** fixed cost $= 0$
CCP				
$(a_{-1} = 0, w_H)$	93.61%	94.95%	94.95%	94.95%
$(a_{-1} = 0, w_L)$	72.99%	80.33%	80.33%	80.33%
$(a_{-1} = 1, w_H)$	99.99%	99.99%	99.99%	99.99%
$(a_{-1} = 1, w_L)$	0.48%	0.29%	0.29%	0.29%

Example: Firm Entry/Exit Model

Counterfactual: Proportional entry subsidy $\tilde{\pi} = \mathcal{H}\pi$ (From KSS)

States: (a_{-1}, w)	**Baseline**	**True CF**	**Estimated CF** scrap value $= 0$	**Estimated CF** fixed cost $= 0$
CCP				
$(a_{-1} = 0, w_H)$	93.61%	94.95%	93.53%	99.87%
$(a_{-1} = 0, w_L)$	72.99%	80.33%	72.53%	99.81%
$(a_{-1} = 1, w_H)$	99.99%	99.99%	99.99%	90.59%
$(a_{-1} = 1, w_L)$	0.48%	0.29%	0.49%	0.00%

Go Back

Example: Firm Entry/Exit Model

- Counterfactual: reduce entry cost ec by 20%
- Formally, $\widetilde{\pi} = \mathcal{H}\pi$, with

$$\mathcal{H}_{00} = I, \text{ and } \mathcal{H}_{11} = \begin{bmatrix} \tau I & (1-\tau)I \\ 0 & I \end{bmatrix},$$

  with $\tau = 0.8$
- So that

$$\widetilde{\pi}_0 = \mathcal{H}_{00}\pi_0 = \pi_0,$$

$$\widetilde{\pi}_1 = \mathcal{H}_{11}\pi_1 = \begin{bmatrix} vp - fc - \tau ec \\ vp - fc \end{bmatrix}$$

- $\widetilde{\mathbf{P}}^I$ is two-dimensional, even in the absence of Restrictions 1–3 Go Back

Example: Firm Entry/Exit Model

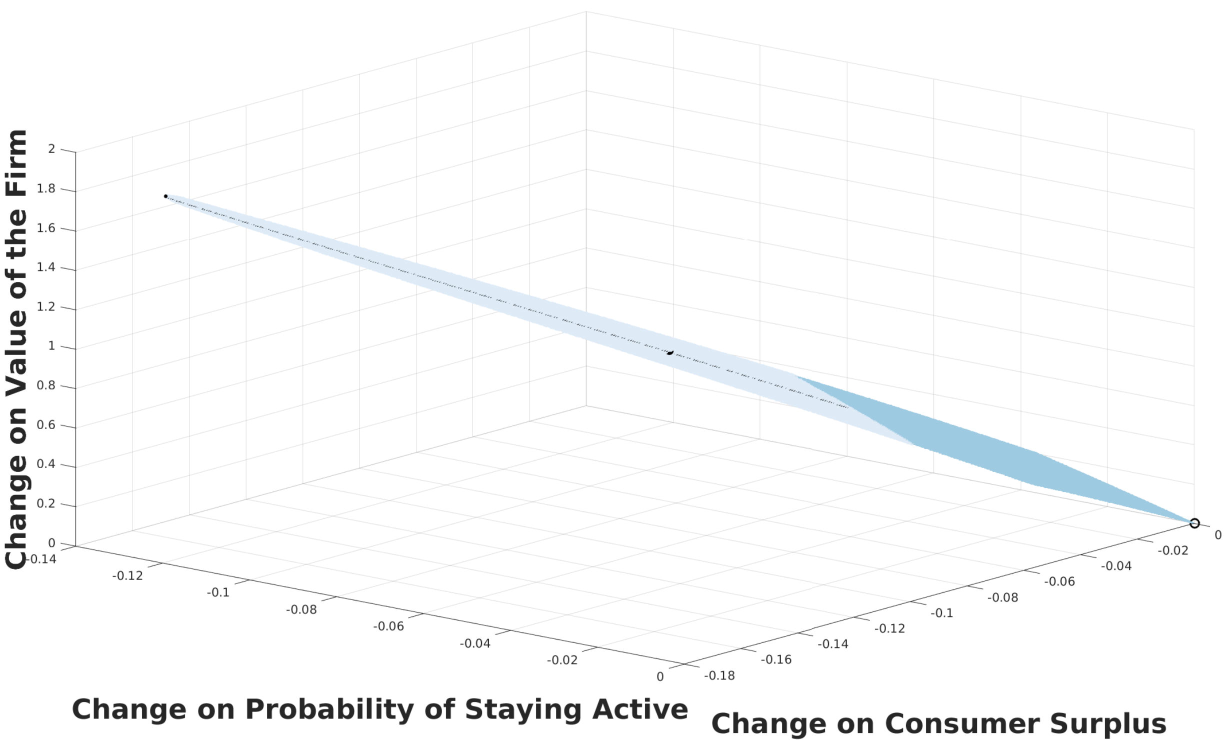


Figure: Outcomes of Interest, θ

Go Back

Inference: Partially-Identified Counterfactual

Theorem

Choose the subsample size h_N such that $h_N \to \infty$ and $h_N/N \to 0$ as $N \to \infty$. Then, under some regularity conditions,

$$\liminf_{N\to\infty} \inf_{\mathfrak{p}\in\mathcal{P}_{\theta_0}} \Pr\{N\widehat{J}_N(\theta_0) \leq \widehat{c}_{1-\alpha}\} = 1-\alpha, \qquad \text{for every } \theta_0 \in \Theta^I,$$

where $\widehat{c}_{1-\alpha}$ is the $1-\alpha$ quantile of $h_N \widehat{J}^*_{h_N}(\theta_0)$, with $0 \leq \alpha \leq \frac{1}{2}$.

The asymptotically uniformly valid $1-\alpha$ confidence set for θ is the collection of θ_0's such that the test does not reject the null $H_0' : J(\theta_0) = 0$.

Go Back

Implementation Details

- Computation of bounds θ^L, θ^U
  - constrained maximization algorithms
  - do provide gradient if feasible
  - show how to calculate gradient when counterfactuals involve the ergodic cdf of states
- If gradient cannot be provided and/or state space is very large
  - propose a stochastic search algorithm
  - directions in the space of $\tilde{p}$ avoid solving the nonlinear equation directly
  - directions in the space of π keep track of model restrictions ▸ Go Back

Implementation Details: Inference procedure

- Exploit the relationship between the optimization for θ^L, θ^U and the optimization in the inference procedure
- Estimate θ^L, θ^U with the full sample
- No point in the interval $\left[\theta^L, \theta^U\right]$ is rejected
- Move to the right (left) by testing a neighboring point $\theta^U + \eta$ $(\theta^L - \eta)$ using sub-sampling
- Continue until the first time a point rejects the null
- Use continuity of the functions involved to pass the solution of the latest as starting value

Inference Details

- At significance level α the test rejects $\theta = \theta_0$ if $J(\theta_0) > c_{1-\alpha}$
  - $c_{1-\alpha}$ is determined by the data as the $(1-\alpha)$ quantile of $h_N J_{h_N}(\theta_0)$, where h_N the subsample size
- Subsampling preferred to bootstrap
  - Under weak conditions it enables us to establish asymptotically uniformly valid critical values
  - So that for all θ the probability of not rejecting θ is asymptotically $1-\alpha$
  - The proof takes into account the randomness imposed by samples in both the point to be projected and the set that contains the projection and avoids convexity assumptions inapplicable in our case

Go Back

Monte Carlo Results

Target Parameter: θ = Long-run Average Probability of Being Active			
Small State Space: $X = 16$			
$T = 5, N = 100$	Restrictions 1	Restrictions 1–2	Restrictions 1–3
True Identified Set	[0.7500, 0.9036]	[0.7500, 0.8763]	[0.7500, 0.8662]
Average Estimated Bounds	[0.7583, 0.9036]	[0.7579, 0.8727]	[0.7580, 0.8651]
Average Bias	[0.0083, 0.0000]	[0.0079, -0.0036]	[0.0080, -0.0011]
Confidence Sets: Average Endpoints	[0.6729, 0.9214]	[0.6734, 0.8951]	[0.6757, 0.8870]
Confidence Sets: Average Length	0.2485	0.2217	0.2113
Coverage Probability (90% nominal)	0.9060	0.9010	0.9050
Time Estimation (sec)	0.04	0.05	0.01
Time Inference (min)	2	2	1
$T = 15, N = 1000$	Restrictions 1	Restrictions 1–2	Restrictions 1–3
True Identified Set	[0.7500, 0.9036]	[0.7500, 0.8763]	[0.7500, 0.8662]
Average Estimated Bounds	[0.7507, 0.9036]	[0.7507, 0.8761]	[0.7507, 0.8661]
Average Bias	[0.0007, -0.0000]	[0.0007, -0.0002]	[0.0007, -0.0001]
Confidence Sets: Average Endpoints	[0.7296, 0.9079]	[0.7296, 0.8806]	[0.7297, 0.8713]
Confidence Sets: Average Length	0.1782	0.1510	0.1417
Coverage Probability (90% nominal)	0.9090	0.9010	0.9040
Time Estimation (sec)	0.04	0.04	0.01
Time Inference (min)	2	2	0.7

Monte Carlo Results

Large State Space: $X = 250$			
$T = 5, N = 100$	Restriction 1	Restrictions 1–2	Restrictions 1–3
True Identified Set	[0.7504, 0.9060]	[0.7504, 0.8787]	[0.7503, 0.8685]
Average Estimated Bounds	[0.7612, 0.9027]	[0.7605, 0.8701]	[0.7593, 0.8638]
Average Bias	[0.0108, -0.0033]	[0.0102, -0.0086]	[0.0090, -0.0047]
Confidence Sets: Average Endpoints	[0.6678, 0.9253]	[0.6602, 0.9096]	[0.6621, 0.8979]
Confidence Sets: Average Length	0.2575	0.2494	0.2358
Coverage Probability (90% nominal)	0.8960	0.9090	0.9080
Time Estimation (sec)	7	8	0.7
Time Inference (min)	578	477	252
$T = 15, N = 1000$	Restrictions 1	Restrictions 1–2	Restrictions 1–3
True Identified Set	[0.7504, 0.9060]	[0.7504, 0.8787]	[0.7503, 0.8685]
Average Estimated Bounds	[0.7532, 0.9064]	[0.7532, 0.8790]	[0.7510, 0.8685]
Average Bias	[0.0028, 0.0004]	[0.0028, 0.0003]	[0.0007, 0.0000]
onfidence Sets: Average Endpoints	[0.7321, 0.9106]	[0.7287, 0.8845]	[0.7288, 0.8757]
Confidence Sets: Average Length	0.1786	0.1558	0.1469
Coverage Probability (90% nominal)	0.9070	0.9000	0.9020
Time Estimation (sec)	6	6	0.6
Time Inference (min)	505	457	150

Note: T = number of periods, N = number of markets, X = number of states.

Go Back

Empirical Application: Export supply and subsidies

- Export revenues subsidies (2%)

$$\theta_R = \frac{(\widetilde{f}^* - f^*)' \times \mathbf{R}^f}{\widetilde{f}^{*\prime} \times 0.02 \times \mathbf{R}^f}$$

- Fixed costs subsidies (28%)

$$\theta_F = \frac{(\widetilde{f}^* - f^*)' \times \mathbf{R}^f}{\widetilde{f}^{*\prime} \times 0.28 \times \begin{bmatrix} 0 \\ fc \end{bmatrix}}$$

- Entry costs subsidies (25%)

$$\theta_E = \frac{(\widetilde{f}^* - f^*)' \times \mathbf{R}^f}{\widetilde{f}^{*\prime} \times 0.25 \times \begin{bmatrix} ec \circ \widetilde{p}_1 \\ 0 \end{bmatrix}}$$

where $\mathbf{R}^f$ the export revenue and f^* the ergodic pdf over states

Go Back