Counterfactuals with Latent Information

Dirk Bergemann
Yale

Benjamin Brooks
UChicago

Stephen Morris
MIT

Applications Workshop

April 20, 2021

Overview

- Many (or most?) interesting problems in economics involve agents making decisions in the face of partial and differential information about the world
- Standard empirical approach: Assume a lot of parametric structure on information (affiliated private values, private entry costs, global games, etc) and estimate information and fundamentals from the data
- We can then use the identified parameters to generate counterfactual predictions

This paper

- Information is modeled non-parametrically as an **information structure**, following Harsanyi (1967)
- An information structures is considered admissible if there is a Bayes Nash equilibrium on that information structure that could have generated the data
- The problem of partially identifying the information structures that rationalize the data is intractable
- But if we fix a counterfactual of interest, we show that you can treat the information structure as a nuisance parameter and effectively solve it out
- We then obtain a tractable characterization of counterfactual predictions that are consistent with the data

Plan for the talk

1. Exposit main idea in a toy version of the Roy (1951) selection model
2. Generalize to games
3. Tightening counterfactuals

A simple labor-choice problem

- A unit mass of workers
- Each worker chooses $a \in A = \{0, 1\}$, where $a = 1$ means enter the labor market, and $a = 0$ means not enter
- The potential long-run wage of a worker is $\theta \in \Theta = \{-1, 1\}$
- Payoff from action a when the wage is θ is just $a\theta$
- Each worker has some partial information about the long-run wage before making the decision
- Expected utility preferences: Suppose a worker has a belief p that $\theta = 1$; optimal action is enter if and only if

$$p \cdot 1 + (1 - p) \cdot (-1) \geq 0 \iff p \geq 1/2$$

Information

- Let μ denote the prior distribution of θ
- We can represent workers' information about θ with an **information structure** (S, π)
  - Each worker gets a signal s in the finite set S
  - $\pi(s|\theta)$ is the conditional probability of s given θ
- The interim probability that the wage is θ given the signal s is just

$$\frac{\pi(s|\theta)\mu(\theta)}{\sum_{\theta'} \pi(s|\theta')\mu(\theta')}$$

- Example: $S = \mathbb{R}$ and $s = \theta + \epsilon$, where ϵ is a noise term that is conditionally iid across θ
- (Signals are meant to be an abstract representation of information, and probably not something we could measure)

Strategies and outcomes

- The probability of a worker choosing a when the signal is s is given by the **strategy** $\sigma(a|s)$
- σ is optimal if it maximizes

$$\sum_{\theta\in\{-1,1\}} \sum_{s\in S} \sum_{a\in\{0,1\}} a\theta\sigma(a|s)\pi(s|a)\mu(\theta)$$

Strategies and outcomes

- The probability of a worker choosing a when the signal is s is given by the **strategy** $\sigma(a|s)$
- σ is optimal if it maximizes

$$\sum_{\theta \in \{-1,1\}} \sum_{s \in S} \sum_{a \in \{0,1\}} a\theta\sigma(a|s)\pi(s|a)\mu(\theta)$$

- A strategy **induces** an **outcome distribution** ϕ over (a, θ):

$$\phi(a, \theta) = \sum_{s \in S} \sigma(a|s)\pi(s|\theta)\mu(\theta)$$

- Worker welfare: $\sum_{a,\theta} a\theta\phi(a, \theta)$

Strategies and outcomes

- The probability of a worker choosing a when the signal is s is given by the **strategy** $\sigma(a|s)$
- σ is optimal if it maximizes

$$\sum_{\theta\in\{-1,1\}} \sum_{s\in S} \sum_{a\in\{0,1\}} a\theta\sigma(a|s)\pi(s|a)\mu(\theta)$$

- A strategy **induces** an **outcome distribution** ϕ over (a,θ):

$$\phi(a,\theta) = \sum_{s\in S} \sigma(a|s)\pi(s|\theta)\mu(\theta)$$

- Worker welfare: $\sum_{a,\theta} a\theta\phi(a,\theta)$
- NB If workers knew θ exactly, we would have $a=1$ iff $\theta=1$, but because workers have partial information about θ, a may be imperfectly correlated with θ under ϕ

Equilibrium outcomes

- Which ϕ could be induced by an information structure and optimal strategy?
- Obvious necessary conditions:

$$\sum_{\theta} \theta\phi(0,\theta) \leq 0 \text{ and } \sum_{\theta} \theta\phi(1,\theta) \geq 0 \tag{$*$}$$

Equilibrium outcomes

- Which ϕ could be induced by an information structure and optimal strategy?
- Obvious necessary conditions:

$$\sum_{\theta} \theta\phi(0,\theta) \leq 0 \text{ and } \sum_{\theta} \theta\phi(1,\theta) \geq 0 \qquad (*)$$

- Why? Whatever information a worker has, they at least know what action they took under their optimal strategy, so
  - When the worker entered, entering the labor market must have yielded a non-negative payoff in expectation
  - When the worker did not enter, entering must have yielded a non-positive payoff in expectation
- We refer to $(*)$ as **obedience constraints**

Sufficiency

- In fact, these conditions are sufficient as well, in the sense that if ϕ satisfies $(*)$, then there exists a μ, (S, π), and optimal strategy σ that induce ϕ

Sufficiency

- In fact, these conditions are sufficient as well, in the sense that if ϕ satisfies $(*)$, then there exists a μ, (S, π), and optimal strategy σ that induce ϕ
- Just take $S = A = \{0, 1\}$, $\mu(\theta) = \phi(0, \theta) + \phi(1, \theta)$, $\sigma(a|a) = 1$, and

$$\pi(s|\theta) = \frac{\phi(s, \theta)}{\phi(0, \theta) + \phi(1, \theta)}$$

- Each worker just learns the action they "drew" from ϕ

Sufficiency

- In fact, these conditions are sufficient as well, in the sense that if ϕ satisfies $(*)$, then there exists a μ, (S, π), and optimal strategy σ that induce ϕ
- Just take $S = A = \{0, 1\}$, $\mu(\theta) = \phi(0, \theta) + \phi(1, \theta)$, $\sigma(a|a) = 1$, and

$$\pi(s|\theta) = \frac{\phi(s, \theta)}{\phi(0, \theta) + \phi(1, \theta)}$$

- Each worker just learns the action they "drew" from ϕ
- If ϕ satisfies $(*)$, we say it is an **equilibrium outcome**

Sufficiency

- In fact, these conditions are sufficient as well, in the sense that if ϕ satisfies $(*)$, then there exists a μ, (S, π), and optimal strategy σ that induce ϕ
- Just take $S = A = \{0, 1\}$, $\mu(\theta) = \phi(0, \theta) + \phi(1, \theta)$, $\sigma(a|a) = 1$, and
$$\pi(s|\theta) = \frac{\phi(s, \theta)}{\phi(0, \theta) + \phi(1, \theta)}$$
- Each worker just learns the action they "drew" from ϕ
- If ϕ satisfies $(*)$, we say it is an **equilibrium outcome**
- Could use this characterization to provide an informationally robust prediction for behavior, consistent with a given prior μ
- This is a special case of **Bayes correlated equilibrium** (Bergemann and Morris, 2013, 2016)

Example

- A family of outcomes parametrized by $\alpha \in [0, 1/2]$:

$a \backslash \theta$	-1	1
0	α	$1/2 - \alpha$
1	$1/2 - \alpha$	α

- Both values of θ are equally likely under the prior

Example

- A family of outcomes parametrized by $\alpha \in [0, 1/2]$:

$$\begin{array}{ccc} a \backslash \theta & -1 & 1 \\ 0 & \alpha & 1/2-\alpha \\ 1 & 1/2-\alpha & \alpha \end{array}$$

- Both values of θ are equally likely under the prior
- This is an equilibrium outcome iff $\alpha \geq 1/4$

Example: Full information

- $\alpha = 1/2$:

$a \backslash \theta$	-1	1
0	$1/2$	0
1	0	$1/2$

- So $a = 1$ if and only if $\theta = 1$
- Only possible if the signal reveals **full information** about the wage
- In other words, for every $s \in S$, $\pi(s|\theta) > 0$ for at most one θ, so that the worker's belief is either 0 or 1

Example: No information

- $\alpha = 1/4$:

$a \backslash \theta$	-1	1
0	$1/4$	$1/4$
1	$1/4$	$1/4$

- Action is independent of the wage: Only possible if the signal reveals **no information** about the state, so for every s, $\pi(s|-1) = \pi(s|1)$
- An optimal strategy is $\sigma(0|s) = \sigma(1|s) = 1/2$ for each s

Example: Partial information

- $\alpha = 3/8$:

$a\backslash\theta$	$-1)$	1
0	3/8	1/8
1	1/8	3/8

- Action is imperfectly correlated with the wage

Example: Partial information

- $\alpha = 3/8$:

$a\backslash\theta$	$-1)$	1
0	3/8	1/8
1	1/8	3/8

- Action is imperfectly correlated with the wage
- Lots of information structure/optimal strategy pairs induce this outcome, e.g.,
  - Half of the workers observe full information and half observe no information (and randomize)

Example: Partial information

- $\alpha = 3/8$:

$a \backslash \theta$	$-1)$	1
0	$3/8$	$1/8$
1	$1/8$	$3/8$

- Action is imperfectly correlated with the wage
- Lots of information structure/optimal strategy pairs induce this outcome, e.g.,
  - Half of the workers observe full information and half observe no information (and randomize)
  - Or, every worker receives a signal $s \in \{-1, 1\}$ according to

$$\pi(s|\theta) = \begin{cases} 3/4 & \text{if } s = \theta; \\ 1/4 & \text{otherwise} \end{cases}$$

  - Optimal strategy: $\sigma(1|1) = 1$ and $\sigma(1|-1) = 0$

Counterfactuals

- Now suppose the preceding problem is (partially) **observed**, in the sense that we can measure ϕ but we don't know (S, π)
  (I'll generalize in a few slides to the case where ϕ is partially observed)

Counterfactuals

- Now suppose the preceding problem is (partially) **observed**, in the sense that we can measure ϕ but we don't know (S, π)
(I'll generalize in a few slides to the case where ϕ is partially observed)
- What might happen to worker welfare in a **counterfactual problem** where the labor market changes, and average wages either rise or fall?

Counterfactuals

- Now suppose the preceding problem is (partially) **observed**, in the sense that we can measure ϕ but we don't know (S, π)
  (I'll generalize in a few slides to the case where ϕ is partially observed)
- What might happen to worker welfare in a **counterfactual problem** where the labor market changes, and average wages either rise or fall?
- Formally, the wage becomes $\theta + z$, so that worker utility is $a(\theta + z)$
- We maintain that the workers' information about wages remains the same (i.e., information is "latent but fixed")

Counterfactual prediction

- We say that ϕ is **rationalized** by $(\mu, (S, \pi))$ if there is an optimal strategy σ for the observed problem such that $(\mu, (S, \pi), \sigma)$ induces ϕ
- Similarly, $\widehat{\phi}$ is rationalized by $(\mu, (S, \pi))$ if there is an optimal strategy $\widehat{\sigma}$ in the counterfactual problem such that $(\mu, (S, \pi), \widehat{\sigma})$ induces $\widehat{\phi}$

Counterfactual prediction

- We say that ϕ is **rationalized** by $(\mu, (S, \pi))$ if there is an optimal strategy σ for the observed problem such that $(\mu, (S, \pi), \sigma)$ induces ϕ
- Similarly, $\widehat{\phi}$ is rationalized by $(\mu, (S, \pi))$ if there is an optimal strategy $\widehat{\sigma}$ in the counterfactual problem such that $(\mu, (S, \pi), \widehat{\sigma})$ induces $\widehat{\phi}$
- $\widehat{\phi}$ is a **counterfactual prediction** if there exists $(\mu, (S, \pi))$ that rationalizes both $\widehat{\phi}$ and ϕ
- We will characterize the set of counterfactual predictions

Straw man: Partially identifying information

- Note that if we observe ϕ, then we observe the marginal on θ, which must be μ
- So the real latent parameter is (S, π)
- We could first (partially) identify the set of (S, π) that rationalize ϕ, and for each such information structure, compute optimal strategies when wages change to $\theta + z$
- But since S is an arbitrary set, this question isn't well posed
- We could in principle normalize S so that signals are identified with interim beliefs about θ
- But now we are talking about identifying possible distributions of interim beliefs, which is an infinite dimensional parameter...

The linked problem

- Instead, we can just solve out (S, π) in the following way
- An equivalent way to conceptualize the counterfactual is that each worker actually made two choices:

The linked problem

- Instead, we can just solve out (S, π) in the following way
- An equivalent way to conceptualize the counterfactual is that each worker actually made two choices:
  - Enter/not enter at the observed wage θ (represented by $a \in \{0, 1\}$)

The linked problem

- Instead, we can just solve out (S, π) in the following way
- An equivalent way to conceptualize the counterfactual is that each worker actually made two choices:
  - Enter/not enter at the observed wage θ (represented by $a \in \{0, 1\}$)
  - Enter/not enter at the counterfactual wage $\theta + z$ ($\widehat{a} \in \{0, 1\}$)

The linked problem

- Instead, we can just solve out (S, π) in the following way
- An equivalent way to conceptualize the counterfactual is that each worker actually made two choices:
  - Enter/not enter at the observed wage θ (represented by $a \in \{0, 1\}$)
  - Enter/not enter at the counterfactual wage $\theta + z$ ($\widehat{a} \in \{0, 1\}$)
- Both problems have the same underlying state θ, and utility is

$$a\theta + \widehat{a}(\theta + z)$$

- We refer to this as the **linked problem**

The linked problem

- Instead, we can just solve out (S, π) in the following way
- An equivalent way to conceptualize the counterfactual is that each worker actually made two choices:
  - Enter/not enter at the observed wage θ (represented by $a \in \{0, 1\}$)
  - Enter/not enter at the counterfactual wage $\theta + z$ ($\hat{a} \in \{0, 1\}$)
- Both problems have the same underlying state θ, and utility is

$$a\theta + \hat{a}(\theta + z)$$

- We refer to this as the **linked problem**
- As analysists, we observe (a, θ) but not $\hat{a}$

Equilibrium outcomes in the linked problem

- The workers in the linked problem have information (S, π) and an optimal strategy $\overline{\sigma}$ that maps s to a distribution over $(a, \widehat{a})$
- Proceeding by analogy, we can characterize equilibrium outcomes for the linked problem, i.e., distributions $\overline{\phi}$ on $(a, \widehat{a}, \theta)$
- Choosing $(a, \widehat{a})$ must be optimal when $(a, \widehat{a})$ was chosen

Equilibrium outcomes in the linked problem

- The workers in the linked problem have information (S, π) and an optimal strategy $\overline{\sigma}$ that maps s to a distribution over $(a, \widehat{a})$
- Proceeding by analogy, we can characterize equilibrium outcomes for the linked problem, i.e., distributions $\overline{\phi}$ on $(a, \widehat{a}, \theta)$
- Choosing $(a, \widehat{a})$ must be optimal when $(a, \widehat{a})$ was chosen
- Using the additive separability of payoffs across the observed and counterfactual problems, this reduces to

$$\sum_{\theta} \theta \overline{\phi}(0, \widehat{a}, \theta) \leq 0 \text{ and } \sum_{\theta} \theta \phi(1, \widehat{a}, \theta) \geq 0 \ \forall \widehat{a}$$
$$\sum_{\theta} (\theta + z) \overline{\phi}(a, 0, \theta) \leq 0 \text{ and } \sum_{\theta} (\theta + z) \phi(a, 1, \theta) \geq 0 \ \forall a \qquad (**)$$

Characterization of counterfactual predictions

Theorem

$\widehat{\phi}$ is a counterfactual prediction if and only if there exists an equilibrium outcome $\overline{\phi}$ of the linked problem (i.e., an outcome satisfying $(**)$) such that

(i) *The marginal of $\overline{\phi}$ on (a, θ) is ϕ*

(ii) *The marginal of $\overline{\phi}$ on $(\widehat{a}, \theta)$ is $\widehat{\phi}$.*

Proof sketch: Only if

- If $\widehat{\phi}$ is a counterfactual prediction, then there is a $(\mu, (S, \pi))$ and optimal strategies σ and $\widehat{\sigma}$ that induce ϕ and $\widehat{\phi}$, respectively

Proof sketch: Only if

- If $\widehat{\phi}$ is a counterfactual prediction, then there is a $(\mu, (S, \pi))$ and optimal strategies σ and $\widehat{\sigma}$ that induce ϕ and $\widehat{\phi}$, respectively
- Now consider the linked problem with prior/information $(\mu, (S, \pi))$
- Straightforward to show that the following strategies are optimal:

$$\overline{\sigma}(a, \widehat{a}|s) = \sigma(a|s)\widehat{\sigma}(\widehat{a}|s)$$

(Follows from the additive structure of payoffs, so that the correlation between a and $\widehat{a}$ given θ does not affect the worker's payoff)

Proof sketch: Only if

- If $\widehat{\phi}$ is a counterfactual prediction, then there is a $(\mu, (S, \pi))$ and optimal strategies σ and $\widehat{\sigma}$ that induce ϕ and $\widehat{\phi}$, respectively
- Now consider the linked problem with prior/information $(\mu, (S, \pi))$
- Straightforward to show that the following strategies are optimal:

$$\overline{\sigma}(a, \widehat{a}|s) = \sigma(a|s)\widehat{\sigma}(\widehat{a}|s)$$

  (Follows from the additive structure of payoffs, so that the correlation between a and $\widehat{a}$ given θ does not affect the worker's payoff)
- This prior/information/strategy induce an equilibrium outcome $\overline{\phi}$ for the linked problem
- Easy to check that ϕ and $\widehat{\phi}$ are marginals of $\overline{\phi}$

Proof sketch: If

- If $\overline{\phi}$ is an equilibrium outcome of the linked problem, then there are $(\mu, (S, \pi))$ and an optimal strategy $\overline{\sigma}$ that induce $\overline{\phi}$

Proof sketch: If

- If $\overline{\phi}$ is an equilibrium outcome of the linked problem, then there are $(\mu, (S, \pi))$ and an optimal strategy $\overline{\sigma}$ that induce $\overline{\phi}$
- Straightforward to check that

$$\sigma(a|s) = \sum_{\widehat{a}} \overline{\sigma}(a, \widehat{a}|s);$$

$$\widehat{\sigma}(\widehat{a}|s) = \sum_{a} \overline{\sigma}(a, \widehat{a}|s)$$

are optimal strategies for the observed and counterfactual problems

Proof sketch: If

- If $\overline{\phi}$ is an equilibrium outcome of the linked problem, then there are $(\mu, (S, \pi))$ and an optimal strategy $\overline{\sigma}$ that induce $\overline{\phi}$
- Straightforward to check that

$$\sigma(a|s) = \sum_{\widehat{a}} \overline{\sigma}(a, \widehat{a}|s);$$
$$\widehat{\sigma}(\widehat{a}|s) = \sum_{a} \overline{\sigma}(a, \widehat{a}|s)$$

are optimal strategies for the observed and counterfactual problems

- Moreover, these strategies induce equilibrium outcomes, which are precisely the marginals of $\overline{\phi}$ $\square$

Solving out the nuisance parameter

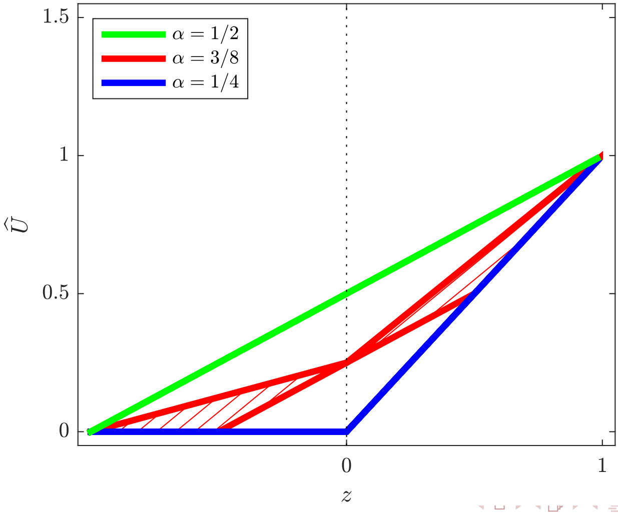
- Essentially, we solved out (S, π) by observing that it is without loss to consider information structures in which $s = (a, \widehat{a})$, and "obeying their signals" is an optimal strategy
- This is an instance of the revelation principle for information (e.g., Myerson 1986, Bergemann and Morris 2013, 2016)
- This normalization works as long as we fix a particular counterfactual of interest

Bounding counterfactual welfare

- How to use this result? Well, for any counterfactual quantity that is linear in $\widehat{\phi}$, we can compute the set of possible counterfactual values by solving a linear program
- For example, bounds for counterfactual worker welfare solve:

$$\begin{aligned} &\max / \min_{\overline{\phi},\widehat{\phi}\geq 0} \sum_{\widehat{a},\theta} \widehat{a}(\theta+z)\widehat{\phi}(\widehat{a},\theta) \\ \text{s.t. } &\sum_{\widehat{a}} \overline{\phi}(a,\widehat{a},\theta) = \phi(a,\theta)\ \forall(a,\theta) \\ &\sum_{a} \overline{\phi}(a,\widehat{a},\theta) = \widehat{\phi}(\widehat{a},\theta)\ \forall(\widehat{a},\theta) \\ &\text{and } (**) \end{aligned}$$

Numerical example: Counterfactual worker welfare



Comments on the figure

- For $\alpha = 1/4$ (no information) and $\alpha = 1/2$ (full information), there are point predictions for welfare for all z
- This is because information is "point identified" by the data, up to redundancy/labeling of signals
- In fact, Blackwell's theorem implies that these information structures attain minimum/maximum counterfactual welfare across all possible information structures

Comments on the figure

- For $\alpha = 1/4$ (no information) and $\alpha = 1/2$ (full information), there are point predictions for welfare for all z
- This is because information is "point identified" by the data, up to redundancy/labeling of signals
- In fact, Blackwell's theorem implies that these information structures attain minimum/maximum counterfactual welfare across all possible information structures
- For $\alpha = 3/8$, information is **not** point identified from the data
- In the local counterfactual ($z = 0$) there is still a point prediction because of a simple replication argument
- But for $z \neq 0$, there is a fat set of counterfactual welfare levels, with the upper and lower bounds corresponding to the two rationalizing information structures that I previously described

Censored data

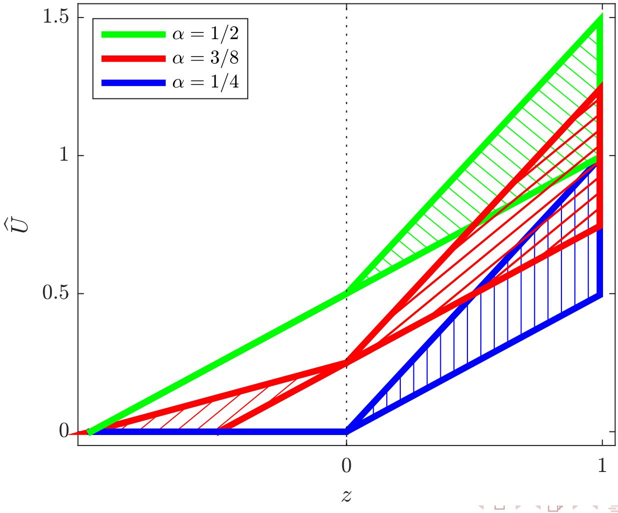
- If we view θ as a potential wage, then it is natural to think that we would not observe θ for workers who did not enter the labor market
- We can repeat the earlier exercise, except that now, the data restriction is that

$$\sum_{\widehat{a}} \overline{\phi}(1, \widehat{a}, \theta) = \phi(1, \theta) \ \forall \theta;$$

$$\sum_{(\widehat{a}, \theta)} \overline{\phi}(0, \widehat{a}, \theta) = 1 - \phi(1, -1) - \phi(1, 1)$$

- This replaces the constraint that the marginal of $\overline{\phi}$ on (a, θ) is ϕ
- Otherwise, the linear program remains the same

Counterfactual worker welfare with censored data



Comments on censored data

- Prediction is the same for $z < 0$ because in this range, whenever the worker didn't enter in the observed problem, they would still not enter in the counterfactual
- For $z > 0$, there are workers who didn't enter in the observed problem but may choose to enter in the counterfactual
- But whether they find it optimal to enter in the counterfactual depends on the censored wages

Comments on censored data

- Prediction is the same for $z < 0$ because in this range, whenever the worker didn't enter in the observed problem, they would still not enter in the counterfactual
- For $z > 0$, there are workers who didn't enter in the observed problem but may choose to enter in the counterfactual
- But whether they find it optimal to enter in the counterfactual depends on the censored wages
- For example, when the data came from $\alpha = 1/2$, the actual ϕ showed $\theta = -1$ when $a = 0$
- But with censored data, it is possible that $\mathbb{E}[\theta|a = 0] = 0.5$, which is still consistent with the worker not entering, but then they strictly prefer to enter in the counterfactual with $z > 0$

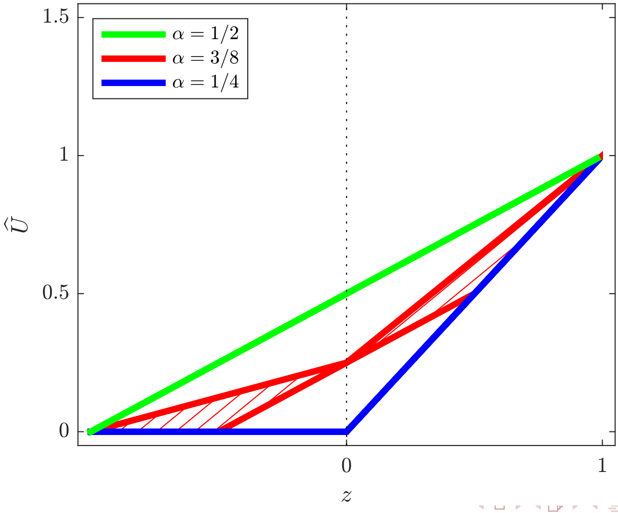
Fixed versus variable information

- Fixing information is the natural thing to do when we want to vary one parameter at a time
- But of course, when one part of the economy changes, others may change as well, and information is no exception

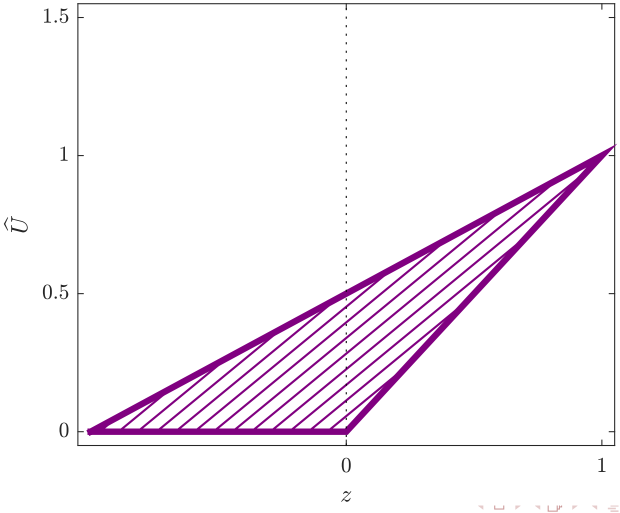
Fixed versus variable information

- Fixing information is the natural thing to do when we want to vary one parameter at a time
- But of course, when one part of the economy changes, others may change as well, and information is no exception
- In a **variable information** counterfactual, we only hold the prior μ fixed, but we allow information to vary
- This is the approach taken in two recent papers: Tamer, Syrgkanis, and Ziani (2018) and Magnolfi and Roncoroni (2021)
  (Later versions of Magnolfi and Roncoroni also do fixed information counterfactuals, following this paper)
- Variable information necessarily leads to a larger counterfactual prediction...

Counterfactual worker welfare with fixed information



Counterfactual worker welfare with variable information



Generalization with many players

- We now extend the analysis to general games
- Finite set of players N
- Finite set of states Θ
- A **game form** $\mathcal{G}$ consists of:
  - Finite actions A_i for each $i \in N$, $A = \times_{i \in N} A_i$
  - Utility functions $u_i : A \times \Theta \to \mathbb{R}$
- The **prior** $\mu \in \Delta(\Theta)$

Information

- An **information structure** $\mathcal{I}$ consists of
    - Finite set of signals S_i for each $i \in N$
    - $S = \times_{i \in N} S_i$
    - $\pi : \Theta \to \Delta(S)$

Information

- An **information structure** $\mathcal{I}$ consists of
  - Finite set of signals S_i for each $i \in N$
  - $S = \times_{i \in N} S_i$
  - $\pi : \Theta \to \Delta(S)$
- NB as in the single-player case, the signal encodes beliefs about the state, but with many players, it also encodes beliefs about others' signals
- Mertens and Zamir (1985): signals can be "canonically" represented as infinite hierarchies of beliefs, although the hierarchies do not capture pure coordination devices that the players may have access to
- In short, there is no simple way to represent information so that it can be directly partially identified from the data

Strategies, equilibrium, outcomes

- A triple $(\mu, \mathcal{G}, \mathcal{I})$ is a Bayesian game
- Strategies: $\sigma_i : S_i \to \Delta(A_i)$
- Expected payoff from a strategy profile $\sigma = (\sigma_i)_{i \in N}$:

$$U_i(\sigma) = \sum_{\theta, s, a} u_i(a, \theta)\sigma(a|s)\pi(s|\theta)\mu(\theta)$$

- σ is a **(Bayes Nash) equilibrium** if $U_i(\sigma) \geq U_i(\sigma_i', \sigma_{-i})$ for all i, σ_i'

Outcomes

- An outcome of $\mathcal{G}$ is a distribution $\phi \in \Delta(A \times \Theta)$
- ϕ is induced by $(\mu, \mathcal{I}, \sigma)$ if for all (a, θ),

$$\phi(a, \theta) = \sum_{s \in S} \sigma(a|s)\pi(s|\theta)\mu(\theta)$$

- ϕ is a **Bayes correlated equilibrium** (BCE) if for all i, a_i, a_i',

$$\sum_{a_{-i},\theta} \phi(a_i, a_{-i}, \theta) u_i(a_i, a_{-i}, \theta) \geq \sum_{a_{-i},\theta} \phi(a_i, a_{-i}, \theta) u_i(a_i', a_{-i}, \theta)$$

- Fact (BM '16): ϕ is a BCE of $\mathcal{G}$ if and only if there exists $(\mu, \mathcal{I})$ and an equilibrium σ of $(\mu, \mathcal{G}, \mathcal{I})$ that induce ϕ

Linked games

- Given game forms $\mathcal{G} = (A_i, u_i)_{i \in N}$ and $\widehat{\mathcal{G}} = (\widehat{A}_i, \widehat{u}_i)_{i \in N}$, the **linked game** $\overline{G} = (\overline{A}_i, \overline{u}_i)_{i \in N}$ is defined by
    - $\overline{A}_i = A_i \times \widehat{A}_i$
    - Thus $\overline{A} = A \times \widehat{A}$
    - $\overline{u}_i(a, \widehat{a}, \theta) = u_i(a, \theta) + \widehat{u}_i(\widehat{a}, \theta)$
- We refer to $\mathcal{G}$ and $\widehat{\mathcal{G}}$ as **component games** of $\overline{\mathcal{G}}$

Counterfactual predictions

- Fix an **observed game** $\mathcal{G}$ and an **unobserved game** $\widehat{\mathcal{G}}$
- Let $M \subseteq \Delta(A \times \Theta)$
- $\widehat{\phi} \in \Delta(\widehat{A} \times \Theta)$ is a **counterfactual prediction consistent with** M if there exists
  - $(\mu, \mathcal{I})$
  - an equilibrium σ of $(\mu, \mathcal{G}, \mathcal{I})$
  - an equilibrium $\widehat{\sigma}$ of $(\mu, \widehat{\mathcal{G}}, \mathcal{I})$

  such that $(\mu, \mathcal{I}, \sigma)$ induce an outcome in M and $(\mu, \mathcal{I}, \widehat{\sigma})$ induce $\widehat{\phi}$

Counterfactual predictions

- Fix an **observed game** $\mathcal{G}$ and an **unobserved game** $\widehat{\mathcal{G}}$
- Let $M \subseteq \Delta(A \times \Theta)$
- $\widehat{\phi} \in \Delta(\widehat{A} \times \Theta)$ is a **counterfactual prediction consistent with** M if there exists
  - $(\mu, \mathcal{I})$
  - an equilibrium σ of $(\mu, \mathcal{G}, \mathcal{I})$
  - an equilibrium $\widehat{\sigma}$ of $(\mu, \widehat{\mathcal{G}}, \mathcal{I})$

  such that $(\mu, \mathcal{I}, \sigma)$ induce an outcome in M and $(\mu, \mathcal{I}, \widehat{\sigma})$ induce $\widehat{\phi}$
- In words, $\widehat{\phi}$ is an equilibrium outcome of the unobserved game for some prior and information structure, and there is also an equilibrium of the observed game under that prior and information structure that is consistent with M

Main result

Theorem

The distribution $\widehat{\phi}$ is a counterfactual prediction consistent with M if and only if there exists a BCE $\overline{\phi}$ of the linked game $\overline{\mathcal{G}}$ such that

(i) *the marginal of $\overline{\phi}$ on $A \times \Theta$ is in M, and*

(ii) *the marginal of $\overline{\phi}$ on $\widehat{A} \times \Theta$ is $\widehat{\phi}$.*

Remarks

- We have generalized form the selection example in three ways:
    - Many players
    - Arbitrary sets of actions/states
    - General data restrictions
- Presumably, the most useful restrictions are linear in ϕ

Remarks

- We have generalized form the selection example in three ways:
  - Many players
  - Arbitrary sets of actions/states
  - General data restrictions
- Presumably, the most useful restrictions are linear in ϕ
- Some notable examples:
  - $M = \{\phi\}$ for some $\phi \in \Delta(A \times \Theta)$
    (the entire outcome is observed)
  - $M = \{\phi \in \Delta(A \times \Theta) | \mathrm{marg}_A \phi = \psi\}$ for some $\psi \in \Delta(A)$
    (Only actions are observed and the state is unobserved)
  - $M = \{\phi \in \Delta(A \times \Theta) | \mathrm{marg}_A \phi \in \Psi\}$ for some $\Psi \subseteq \Delta(A)$
    (Only some aspects of the action distribution are observed, e.g., the high bid in an auction)

Assumptions that are without loss

- All players receive signals from the same information structure
  (Variation in information can be represented by just adding more signals)

Assumptions that are without loss

- All players receive signals from the same information structure
  (Variation in information can be represented by just adding more signals)
- All players play the same equilibria of $\mathcal{G}$ and $\widehat{\mathcal{G}}$
  (Can always add public correlation devices to $\mathcal{I}$ that are used to coordinate on different equilibria)

Assumptions that are without loss

- All players receive signals from the same information structure
  (Variation in information can be represented by just adding more signals)
- All players play the same equilibria of $\mathcal{G}$ and $\widehat{\mathcal{G}}$
  (Can always add public correlation devices to $\mathcal{I}$ that are used to coordinate on different equilibria)
- Preferences are known
  (Uncertainty about preferences can be embedded in θ)

Assumptions that are without loss

- All players receive signals from the same information structure
  (Variation in information can be represented by just adding more signals)
- All players play the same equilibria of $\mathcal{G}$ and $\widehat{\mathcal{G}}$
  (Can always add public correlation devices to $\mathcal{I}$ that are used to coordinate on different equilibria)
- Preferences are known
  (Uncertainty about preferences can be embedded in θ)
- μ is held fixed in between observed and counterfactual outcomes
  (Indeterminacy between μ and u)

Assumptions that are without loss

- All players receive signals from the same information structure
(Variation in information can be represented by just adding more signals)
- All players play the same equilibria of $\mathcal{G}$ and $\widehat{\mathcal{G}}$
(Can always add public correlation devices to $\mathcal{I}$ that are used to coordinate on different equilibria)
- Preferences are known
(Uncertainty about preferences can be embedded in θ)
- μ is held fixed in between observed and counterfactual outcomes
(Indeterminacy between μ and u)
- One observed game and one counterfactual game
(In the paper, we show how to combine data from multiple games, and also simultaneously make joint predictions in more than one counterfactual)

Enriching the model

- The theorem is "bare bones" in the sense that there is minimal structure on actions/information
- One can add more structure in order to tighten the counterfactual
- Two leading examples:
  - Bounds on information
  - Payoff shifters/instruments

Bounds on information

- In a sense, the outcome of the observed game imposes both a lower bound and an upper bound on information
  - Lower bound: any information structure that rationalizes the data must be at least as informative as the action in the observed game (This can be made precise using the individual sufficiency order on information structures)
  - Upper bound: players cannot have so much information that they would not be willing to take their observed actions
- One can generalize both of these ideas

General lower bounds on information

- In particular, we can introduce an exogenously given information structure $\underline{\mathcal{I}} = (\underline{S}, \underline{\pi})$, and suppose that players observe at least their signals in $\underline{\mathcal{I}}$
- It is straightforward to generalize the theorem, where an outcome is now a joint distribution on $(a, \underline{s}, \theta)$, and the obedience constraints are that a_i is optimal conditional on $(a_i, \underline{s}_i)$
- Leading example is **private values**: $\Theta = \times_{i \in N} \Theta_i$, $\underline{S}_i = \Theta_i$, and $\underline{\pi}(s = \theta | \theta) = 1$

General upper bounds on information

- There is no simple way to impose an upper bound on information that is a direct counterpart to the lower bound and preserves the linear structure
- But one could impose upper bounds using "hypothetical" games
- In particular, there is nothing special about which represent limits on how precise players information can be about the state or about others information
- For example, one could conjecture a hypothetical game in which players just guess the state, and impose an upper bound on their payoffs

Payoff shifters/instruments

- A common device in applied work: There is an exogenous "payoff shifter" ω_i that affects player i's utility from an action, and the econometrician can observe the mapping from payoff shifters to actions
- We can incorporate such shifters into our framework:
  - Ω_i: Set of values for player i's payoff shifter
  - $\Omega = \times_{i \in N} \Omega_i$
  - $\eta(\omega|\theta)$: The exogenous distribution of $\omega \in \Omega$ given $\theta \in \Theta$
  - Preferences are of the form $u_i(a, \omega, \theta)$

Payoff shifters/instruments

- A common device in applied work: There is an exogenous "payoff shifter" ω_i that affects player i's utility from an action, and the econometrician can observe the mapping from payoff shifters to actions
- We can incorporate such shifters into our framework:
  - Ω_i: Set of values for player i's payoff shifter
  - $\Omega = \times_{i \in N} \Omega_i$
  - $\eta(\omega|\theta)$: The exogenous distribution of $\omega \in \Omega$ given $\theta \in \Theta$
  - Preferences are of the form $u_i(a, \omega, \theta)$
- Assumptions:
  - Each player observes their own ω_i
  - Signals in $\mathcal{I}$ are conditionally independent of ω given θ

Reduction to the original model

- We can reduce this problem to our original model, by creating a new game form $\tilde{\mathcal{G}}$ in which player i's actions $\tilde{A}_i$ is the set of pure-strategies that map Ω_i to A_i
- Given a pure-strategy profile $\tilde{a} \in \tilde{A}$, the expected utility is

$$\tilde{u}_i(\tilde{a}, \theta) = \sum_{\omega \in \Omega} u_i(\tilde{a}(\omega), \omega, \theta)\eta(\omega|\theta)$$

- One can then apply our theorem to the reduced game $\tilde{\mathcal{G}}$
- This crucially relies on us knowing η (or being able to estimate it)

Example with payoff shifters

- As an extreme example, in the selection game, suppose that the payoff from entering the labor market is $\theta + 1 - 2\omega$, where ω is a payoff shifter that is uniform on $[-1, 1]$
- The analyst observes the joint distribution of (a, ω)
- Let p denote the interim belief of the worker that $\theta = 1$

Example with payoff shifters

- As an extreme example, in the selection game, suppose that the payoff from entering the labor market is $\theta + 1 - 2\omega$, where ω is a payoff shifter that is uniform on $[-1, 1]$
- The analyst observes the joint distribution of (a, ω)
- Let p denote the interim belief of the worker that $\theta = 1$
- Not entering $(a = 0)$ is optimal if and only if

$$p \cdot (1 + 1 - 2\omega) + (1 - p) \cdot (-1 + 1 - 2\omega) \leq 0 \iff p \leq \omega$$

Example with payoff shifters

- As an extreme example, in the selection game, suppose that the payoff from entering the labor market is $\theta + 1 - 2\omega$, where ω is a payoff shifter that is uniform on $[-1, 1]$
- The analyst observes the joint distribution of (a, ω)
- Let p denote the interim belief of the worker that $\theta = 1$
- Not entering ($a = 0$) is optimal if and only if

$$p \cdot (1 + 1 - 2\omega) + (1 - p) \cdot (-1 + 1 - 2\omega) \leq 0 \iff p \leq \omega$$

- Thus, the probability of choosing $a = 0$ given ω is precisely the probability that $p \leq \omega$, and the joint distribution of (a, ω) point identifies the distribution of workers' interim beliefs

Multi-player example: Entry

- Classic problem in applied IO: Modeling firm entry into a market (e.g., Bresnahan and Reiss, 1990; Ciliberto and Tamer, 2009)
- Firms $i = 1, 2$ decide whether or not to enter
- Private entry cost $c_i \in \{0, C\}$
- Monopoly profit is X, duopoly profit is $X - \Delta$
- Payoffs:

$a_1 \backslash a_2$	N	E
N	$(0, 0)$	$(0, X - c_2)$
E	$(X - c_1, 0)$	$(X - \Delta - c_1, X - \Delta - c_2)$

Observed outcome and counterfactual

- Parameters for observed game: $X = 3$, $\Delta = C = 2$
- Observed outcome: all cost profiles are equally likely, firms enter if and only if they have low cost

Observed outcome and counterfactual

- Parameters for observed game: $X = 3$, $\Delta = C = 2$
- Observed outcome: all cost profiles are equally likely, firms enter if and only if they have low cost
- This is an equilibrium at these parameters if firms know their costs (and only know their costs)
  - In fact, entering is a strictly dominant strategy for the low cost firm
  - If only the low cost firm enters, the high cost firm is indifferent:

$$\frac{1}{2}(X - C) + \frac{1}{2}(X - \Delta - C) = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot (-1) = 0$$

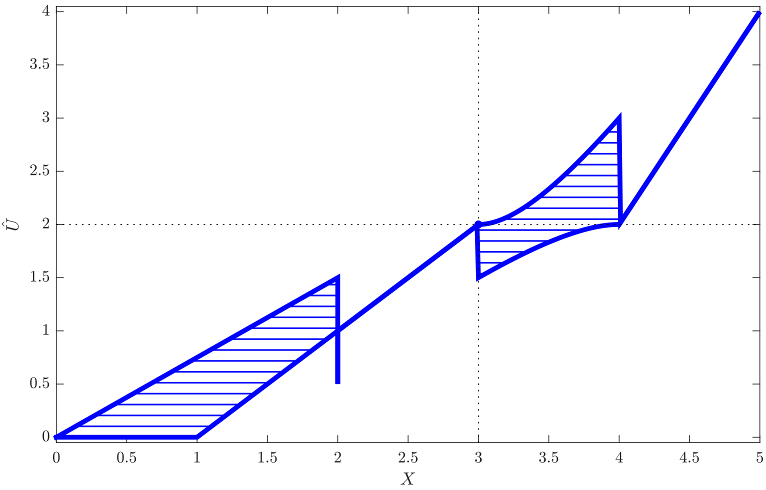
Observed outcome and counterfactual

- Parameters for observed game: $X = 3$, $\Delta = C = 2$
- Observed outcome: all cost profiles are equally likely, firms enter if and only if they have low cost
- This is an equilibrium at these parameters if firms know their costs (and only know their costs)
  - In fact, entering is a strictly dominant strategy for the low cost firm
  - If only the low cost firm enters, the high cost firm is indifferent:

$$\frac{1}{2}(X - C) + \frac{1}{2}(X - \Delta - C) = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot (-1) = 0$$

- We will look at counterfactual producer surplus as we vary X (so uniformly shifting profits from entry)

Counterfactual producer surplus in the entry game



Describing the equilibria

- In the paper, we give a full description of the information/equilibrium pairs that generate the bounds
- Key facts from the "data":

Describing the equilibria

- In the paper, we give a full description of the information/equilibrium pairs that generate the bounds
- Key facts from the "data":
  - Firms must at least know their own costs, since they are perfectly correlated with the observed entry decision

Describing the equilibria

- In the paper, we give a full description of the information/equilibrium pairs that generate the bounds
- Key facts from the "data":
    - Firms must at least know their own costs, since they are perfectly correlated with the observed entry decision
    - Entering is a strictly dominant strategy for low types, so we don't learn anything about low types' higher order beliefs

Describing the equilibria

- In the paper, we give a full description of the information/equilibrium pairs that generate the bounds
- Key facts from the "data":
  - Firms must at least know their own costs, since they are perfectly correlated with the observed entry decision
  - Entering is a strictly dominant strategy for low types, so we don't learn anything about low types' higher order beliefs
  - But high types can only be indifferent between entering and not entering if they have no information about the other firm's type (Recall discussion of upper bound on information)

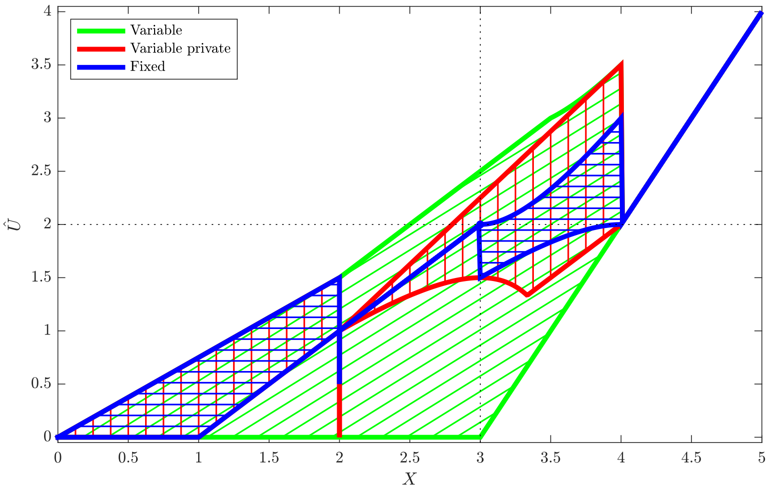
Describing the equilibria

- In the paper, we give a full description of the information/equilibrium pairs that generate the bounds
- Key facts from the "data":
  - Firms must at least know their own costs, since they are perfectly correlated with the observed entry decision
  - Entering is a strictly dominant strategy for low types, so we don't learn anything about low types' higher order beliefs
  - But high types can only be indifferent between entering and not entering if they have no information about the other firm's type (Recall discussion of upper bound on information)
- For $X > 2$, bounds are obtained using pure correlation devices used to maximize/minimize the probability that both high cost firms enter

Describing the equilibria

- In the paper, we give a full description of the information/equilibrium pairs that generate the bounds
- Key facts from the "data":
  - Firms must at least know their own costs, since they are perfectly correlated with the observed entry decision
  - Entering is a strictly dominant strategy for low types, so we don't learn anything about low types' higher order beliefs
  - But high types can only be indifferent between entering and not entering if they have no information about the other firm's type (Recall discussion of upper bound on information)
- For $X > 2$, bounds are obtained using pure correlation devices used to maximize/minimize the probability that both high cost firms enter
- Last figure... counterfactual with variable information

Counterfactual profit with variable information



Conclusion

- We have a new non-parametric approach to counterfactual predictions in games of incomplete information
- The main advantage: It circumvents partial identification of the information structure, an infinite dimensional nuisance parameter
- Already being used in practice by Magnolfi and Roncoroni (2021)
- We have nothing to say about how to do inference, and there are significant computational challenges that need to be overcome in operationalizing this methodology
- These are important directions for future work!