ECON 289, Lecture 3
Information design
and informationally robust predictions

Instructor: Ben Brooks

Spring 2023

Bayesian games and their limitations

- Harsanyi type spaces and Bayesian games are powerful tools for analyzing strategic behavior in the presence of (possibly higher order) uncertainty
- But there are two issues, one technical and one conceptual:

Bayesian games and their limitations

- Harsanyi type spaces and Bayesian games are powerful tools for analyzing strategic behavior in the presence of (possibly higher order) uncertainty
- But there are two issues, one technical and one conceptual:
  1. Even though we know how to model rich higher order information, it can be quite challenging to solve for equilibrium, either numerically or analytically, without strong assumptions

Bayesian games and their limitations

- Harsanyi type spaces and Bayesian games are powerful tools for analyzing strategic behavior in the presence of (possibly higher order) uncertainty
- But there are two issues, one technical and one conceptual:
  1. Even though we know how to model rich higher order information, it can be quite challenging to solve for equilibrium, either numerically or analytically, without strong assumptions
  2. Moreover, unlike the actions and the payoffs, which often have physically observable counterparts, the type space is more abstract... Which type space is empirically relevant?

Bayesian games and their limitations

- Harsanyi type spaces and Bayesian games are powerful tools for analyzing strategic behavior in the presence of (possibly higher order) uncertainty
- But there are two issues, one technical and one conceptual:
  1. Even though we know how to model rich higher order information, it can be quite challenging to solve for equilibrium, either numerically or analytically, without strong assumptions
  2. Moreover, unlike the actions and the payoffs, which often have physically observable counterparts, the type space is more abstract... Which type space is empirically relevant?
- One response to the second issue is to be agnostic about which is the correct type space... and as we'll see, this can also help with the first issue

Incomplete information correlated equilibrium

- There is an analogy here with correlated equilibrium, which captures the analyst's uncertainty about which correlation devices are available to the players
  - Aumann's theorem says that any behavior that is consistent with a common prior and common knowledge of rationality is a correlated equilibrium
  - Moreover, correlated equilibrium often ends up being **more** tractable than Nash under an arbitrary correlation device, due to the linearity of the obedience constraints in the joint distribution of actions
- A natural idea is to adapt correlated equilibrium to games of incomplete information
- We will operationalize this idea using **Bayes correlated equilibrium** (BCE) (Bergemann and Morris, 2013, 2016)
- NB lots of alternative definitions of incomplete info correlated equilibrium, notably Cotter (1991) and Forges (1993)

Basic environment

- Payoff relevant state $\theta \in \Theta$
- Players $i = 1, \ldots, n$, actions A_i
- Payoff functions $u_i : A \times \Theta \to \mathbb{R}$
- Fix a common prior Harsanyi type space $\mathcal{T} = (\Theta, T, \pi)$
- Assume all sets are finite

Bayes Nash equilibrium

- Together, $(\Theta, \{A_i\}, \{u_i\})$ and $\mathcal{T}$ constitute a Bayesian game
- Player i's **strategies** are mappings $\sigma_i : T_i \to \Delta(A_i)$
- Under the strategy profile $\sigma = (\sigma_1, \ldots, \sigma_n)$, player i's expected payoff is

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

- A **Bayes Nash equilibrium** is a strategy profile σ such that for all i and σ_i',

$$U_i(\sigma) \geq U_i(\sigma_i', \sigma_{-i})$$

Implications of rationality in Bayesian games

- Suppose that the agents know their types $t_i \in T_i$ for sure
- What are the outcomes that are consistent with common knowledge of rationality and consistent with a common prior such that the distribution of (θ, t) is $\pi(\theta, t)$?
- This is equivalent to asking, what are the outcomes that could arise in a Bayes Nash equilibrium when the players know "at least as much" as they know under $\mathcal{T}$?
- For example, in the auction context, maybe $\theta = (\theta_1, \ldots, \theta_N)$ encodes the player's values, and we want to assume that each player knows their own value, $t_i = \theta_i$, but we are uncertain about what else the players might know beyond their own value

Expansions and outcomes

- An **expansion** of $\mathcal{T} = (T, \pi)$ is a type space $\mathcal{T}' = (T \times T', \pi')$ such that for all t

$$\sum_{t'} \pi'(\theta, t, t') = \pi(t)$$

- In other words, $\mathcal{T}'$ is equivalent to observing $\mathcal{T}$ plus the an extra dimension t'
- An **outcome** is just a joint distribution μ over $\Theta \times T \times A$
- Any type space $\mathcal{T}'$ that expands $\mathcal{T}$ and strategy profile σ on $\mathcal{T}'$ **induce** the outcome given by

$$\mu(\theta, t, a) = \sum_{\{t' | (t,t') \in T'\}} \pi'(\theta, t, t')\sigma(a|t, t')$$

Bayes correlated equilibrium

- A **Bayes correlated equilibrium** of $(\Theta, \{A_i\}, \{u_i\}, \mathcal{T})$ is an outcome such that

Bayes correlated equilibrium

- A **Bayes correlated equilibrium** of $(\Theta, \{A_i\}, \{u_i\}, \mathcal{T})$ is an outcome such that

$$\sum_{a \in A} \mu(\theta, t, a) = \pi(\theta, t)$$

Bayes correlated equilibrium

- A **Bayes correlated equilibrium** of $(\Theta, \{A_i\}, \{u_i\}, \mathcal{T})$ is an outcome such that

$$\sum_{a \in A} \mu(\theta, t, a) = \pi(\theta, t)$$

and for all i, t_i, a_i, and a_i',

$$\sum_{\theta, t_{-i}, a_{-i}} \mu(\theta, (t_i, t_{-i}), (a_i, a_{-i}))(u_i(a_i, a_{-i}, \theta) - u_i(a_i', a_{-i}, \theta)) \geq 0$$

Bayes correlated equilibrium

- A **Bayes correlated equilibrium** of $(\Theta, \{A_i\}, \{u_i\}, \mathcal{T})$ is an outcome such that

$$\sum_{a \in A} \mu(\theta, t, a) = \pi(\theta, t)$$

and for all i, t_i, a_i, and a_i',

$$\sum_{\theta, t_{-i}, a_{-i}} \mu(\theta, (t_i, t_{-i}), (a_i, a_{-i}))(u_i(a_i, a_{-i}, \theta) - u_i(a_i', a_{-i}, \theta)) \geq 0$$

- We refer to the first of these as the **marginal constraints** and the second as the **obedience constraints**
- Like with CE, we have an interpretation that a disinterested mediator sees (θ, t) and can make secret "recommendations" to the players, and the resulting outcome is a BCE iff it is an equilibrium for the agents to "obey" the mediator

Epistemic characterization

Theorem

An outcome μ is a BCE of $(\Theta, \{A_i\}, \{u_i\}, \mathcal{T})$ iff there exists some expansion $\mathcal{T}'$ of $\mathcal{T}$ and a BNE σ of the game $(\Theta, \{A_i\}, \{u_i\}, \mathcal{T}')$ such that μ is induced by $\mathcal{T}'$ and σ.

- So, BCE are the outcomes that are consistent with rationality and players knowing $\mathcal{T}$
- That play follows a BCE is a "safe" prediction, that entails weak assumptions about types, aside from the common prior
- Also, like correlated equilibrium, it is described as the intersection of a family of linear incentive constraints, which makes it analytically tractable

Proof: Only if

- First, if μ is a BCE, then we can define the expansion to have type spaces $T_i' = T_i \times A_i$ and prior $\pi'(\theta, t, a) = \mu(\theta, t, a)$
- The strategies are the identity mapping, i.e.,

$$\sigma_i(a_i | t_i, a_i) = 1$$

- IC follows from the obedience constraint; if player i deviates to σ_i', then the expected payoff is

Proof: Only if

- First, if μ is a BCE, then we can define the expansion to have type spaces $T_i' = T_i \times A_i$ and prior $\pi'(\theta, t, a) = \mu(\theta, t, a)$
- The strategies are the identity mapping, i.e.,

$$\sigma_i(a_i | t_i, a_i) = 1$$

- IC follows from the obedience constraint; if player i deviates to σ_i', then the expected payoff is

$$\sum_{\theta, t', a} \pi'(\theta, t, a) \sigma_i'(a_i' | t_i, a_i) u_i(a_i', a_{-i}, \theta)$$

Proof: Only if

- First, if μ is a BCE, then we can define the expansion to have type spaces $T_i' = T_i \times A_i$ and prior $\pi'(\theta, t, a) = \mu(\theta, t, a)$
- The strategies are the identity mapping, i.e.,

$$\sigma_i(a_i | t_i, a_i) = 1$$

- IC follows from the obedience constraint; if player i deviates to σ_i', then the expected payoff is

$$\sum_{\theta, t', a} \pi'(\theta, t, a) \sigma_i'(a_i' | t_i, a_i) u_i(a_i', a_{-i}, \theta)$$

$$= \sum_{t_i, a_i, a_i'} \sigma_i'(a_i' | t_i, a_i) \sum_{\theta, t_{-i}, a_{-i}} \mu(\theta, (t_i, t_{-i}), (a_i, a_{-i})) u_i(a_i', a_{-i}, \theta)$$

Proof: Only if

- First, if μ is a BCE, then we can define the expansion to have type spaces $T_i' = T_i \times A_i$ and prior $\pi'(\theta, t, a) = \mu(\theta, t, a)$
- The strategies are the identity mapping, i.e.,

$$\sigma_i(a_i | t_i, a_i) = 1$$

- IC follows from the obedience constraint; if player i deviates to σ_i', then the expected payoff is

$$\begin{aligned}
&\sum_{\theta, t', a} \pi'(\theta, t, a)\sigma_i'(a_i' | t_i, a_i) u_i(a_i', a_{-i}, \theta) \\
&= \sum_{t_i, a_i, a_i'} \sigma_i'(a_i' | t_i, a_i) \sum_{\theta, t_{-i}, a_{-i}} \mu(\theta, (t_i, t_{-i}), (a_i, a_{-i})) u_i(a_i', a_{-i}, \theta) \\
&\leq \sum_{t_i, a_i, a_i'} \sigma_i'(a_i' | t_i, a_i) \sum_{\theta, t_{-i}, a_{-i}} \mu(\theta, (t_i, t_{-i}), (a_i, a_{-i})) u_i(a_i, a_{-i}, \theta) \\
&= U_i(\sigma)
\end{aligned}$$

Proof: If

- We simply verify the induced outcome is a BCE:

$$\mu(\theta, t, a) = \sum_{t'} \pi'(\theta, t, t')\sigma(a|t, t')$$

- Suppose there exists i, t_i, a_i, a_i', such that

$$\sum_{\theta, t_{-i}, a_{-i}} \mu(\theta, (t_i, t_{-i}), (a_i, a_{-i}))(u_i(a_i, a_{-i}, \theta) - u_i(a_i', a_{-i}, \theta)) < 0$$

Proof: If

- We simply verify the induced outcome is a BCE:

$$\mu(\theta, t, a) = \sum_{t'} \pi'(\theta, t, t')\sigma(a|t, t')$$

- Suppose there exists i, t_i, a_i, a_i', such that

$$\sum_{\theta, t_{-i}, a_{-i}} \mu(\theta, (t_i, t_{-i}), (a_i, a_{-i}))(u_i(a_i, a_{-i}, \theta) - u_i(a_i', a_{-i}, \theta)) < 0$$

- Consider the σ_i' that plays a_i' whenever t_i would play a_i under σ_i
- Resulting payoff is

$$\begin{aligned} U_i(\sigma) + \sum_{\theta, t_{-i}, t', a_{-i}} & \pi'(\theta, t_i, t_{-i}, t')\sigma(a_i, a_{-t}|t_i, t_{-i}, t')(u_i(a_i', a_{-i}, \theta) - u_i(a, \theta)) \\ &= U_i(\sigma) + \sum_{\theta, t_{-i}, a_{-i}} \mu(\theta, t_i, t_{-i}, a_i, a_{-i})(u_i(a_i', a_{-i}, \theta) - u_i(a, \theta)) \\ &> U_i(\sigma) \end{aligned}$$

which contradicts σ being a BNE $\square$

A "revelation principle"

- Our original motivation was to make "safe" predictions by considering the range of outcomes across lots of type spaces and equilibria
- In particular, suppose we have some welfare function $\phi(a, t, \theta)$, and we are interested in the range of possible values
- This is essentially the optimization problem of maximizing

$$\sum_{a,t,t',\theta} \pi'(\theta, t, t')\sigma(a|t, t')\phi(a, t, \theta)$$

  over all equilibria on expansions $(T \times T', \pi')$ of (T, π)
- This problem looks intractable! The space of expansions is vast, and characterizing BNE on a given expansion is often quite demanding
- But, the BCE characterization shows is that it is actually a finite dimensional LP
- Why? WLOG to optimize the expectation of ϕ over all BCE...

The BCE LP

- Maximizing a Bayesian objective $\phi(a, t, \theta)$ over BCE:

$$\max_{\mu(\theta,t,a)\geq 0} \phi(a,t,\theta)\mu(\theta,t,a)$$

$$\text{s.t.} \sum_{\theta,t_{-i},a_{-i}} \mu(\theta,t_i,t_{-i},a_i,a_{-i})\left(u_i(a_i,a_{-i},\theta)-u_i(a_i',a_{-i},\theta)\right) \geq 0 \ \forall i,a_i,t_i,a_i'$$

$$\sum_a \mu(\theta,t,a) = \pi(\theta,t) \ \forall \theta,t$$

- Dual:

$$\min_{\alpha_i(t_i,a_i,a_i')\geq 0,\lambda(\theta,t)} \sum_{\theta,t} \lambda(\theta,t)\pi(\theta,t)$$

$$\text{s.t. } \lambda(\theta,t) \geq \phi(a,t,\theta) + \sum_{i,a_i'} \alpha_i(t_i,a_i,a_i')(u_i(a_i,a_{-i},\theta)-u_i(a_i',a_{-i},\theta))$$

Interpreting α_i

- The dual variables $\alpha_i(t_i, a_i, a_i')$ are Lagrange multipliers on the obedience constraints
- As pointed out previously, if we aggregate constraints with weights proportional to the optimal multipliers, the value of the objective cannot change
- In this context, such a sum across (t_i, a_i, a_i') would be

$$\sum_{t_i, a_i, a_i'} \alpha_i(t_i, a_i, a_i')\mu(\theta, t_i, t_{-i}, a_i, a_{-i})\left(u_i(a_i, a_{-i}, \theta) - u_i(a_i', a_{-i}, \theta)\right) \geq 0$$

- What does this constraint mean?

Interpreting α_i

- The dual variables $\alpha_i(t_i, a_i, a_i')$ are Lagrange multipliers on the obedience constraints
- As pointed out previously, if we aggregate constraints with weights proportional to the optimal multipliers, the value of the objective cannot change
- In this context, such a sum across (t_i, a_i, a_i') would be

$$\sum_{t_i, a_i, a_i'} \alpha_i(t_i, a_i, a_i')\mu(\theta, t_i, t_{-i}, a_i, a_{-i})\left(u_i(a_i, a_{-i}, \theta) - u_i(a_i', a_{-i}, \theta)\right) \geq 0$$

- What does this constraint mean? It corresponds to a **probabilistic deviation** in the normal form, where player i type t_i deviates from a_i to a_i' with prob proportional to α_i
- So, in other words, the optimal α_i telling us which are the relevant deviations in the normal form
- We could drop all others without changing the optimal BCE

Informationally robust predictions and information design

- When we compute extremal BCE, we are "designing" information to maximize a linear objective
- This is a metaphor for the analyst trying to understand the range of possible behavior in a given environment
- A more literal interpretation is that there is an actual agent who knows (θ, t) and can commit to a rule that sends players private signals
- This is the perspective often taken in the literature on strategic communication with commitment, e.g., Bayesian persuasion
- Both interpretations are valid, but for the latter to be plausible, the information designer has to both have a lot of commitment power and the ability to implement any (recommendation) information structure

Application: First-price auctions

- Next we will discuss a rich application of BCE to the study of first-price auctions/Bertrand competition, based on Bergemann, Brooks, and Morris (2017)
- n bidders
- Values $v_i \in [\underline{v}, \overline{v}]$

Application: First-price auctions

- Next we will discuss a rich application of BCE to the study of first-price auctions/Bertrand competition, based on Bergemann, Brooks, and Morris (2017)
- n bidders
- Values $v_i \in [\underline{v}, \overline{v}]$
- Bidders do not necessarily know anything about the value
- In other words, the base type space $\mathcal{T}$ is no info, i.e., $|T_i| = 1$ for each i

Application: First-price auctions

- Next we will discuss a rich application of BCE to the study of first-price auctions/Bertrand competition, based on Bergemann, Brooks, and Morris (2017)
- n bidders
- Values $v_i \in [\underline{v}, \overline{v}]$
- Bidders do not necessarily know anything about the value
- In other words, the base type space $\mathcal{T}$ is no info, i.e., $|T_i| = 1$ for each i
- The bidders compete in a first-price auction:
  - Bids $b_i \in \mathbb{R}_+$
  - High bidder wins
  - Winner pays their bid

Connection to Bertrand

- We can equivalently interpret the model as one of Bertrand competition (i.e., a procurement auction)
- Seller $\iff$ consumer with unit demand, willingness to pay w
- Bidder $\iff$ firm
- Value $v_i \iff$ cost c_i
- Bid $b_i \iff$ take-it-or-leave-it price p_i
- Firm wins if $w - p_i \geq \max\{0\} \cup \{w - p_j | j \neq i\}$, ties broken uniformly
- Winner's payoff $v_i - b_i \iff p_i - c_i$

Historical analysis of the FPA

- Strategic analysis of FPA started with Vickrey (1961) who studied the symmetric **independent-private value model**, where v_i are iid draws from a CDF F with pdf f
- Symmetric monotonic pure strategy equilibrium, in which $b_i = \beta(v_i)$, where $\beta(v_i) = \frac{1}{F^{n-1}(v_i)} \int_{w=\underline{v}}^{v_i} w dF^{n-1}(w)$
- Same (interim) welfare outcome as in second-price auction, where the high value bidder wins and pays the expected second-highest value

Historical analysis of the FPA

- Strategic analysis of FPA started with Vickrey (1961) who studied the symmetric **independent-private value model**, where v_i are iid draws from a CDF F with pdf f
- Symmetric monotonic pure strategy equilibrium, in which $b_i = \beta(v_i)$, where $\beta(v_i) = \frac{1}{F^{n-1}(v_i)} \int_{w=\underline{v}}^{v_i} w dF^{n-1}(w)$
- Same (interim) welfare outcome as in second-price auction, where the high value bidder wins and pays the expected second-highest value
- To see that this is an equilibrium, note that the bidding function solves the ODE

$$\beta'(w) = \frac{dF^{n-1}(w)}{F^{n-1}(w)}(w - \beta(w))$$

- If a bidder with value v bids $\beta(w)$, the resulting surplus is $(v - \beta(w))F^{n-1}(w)$, which has derivative

$$(v - \beta(w))dF^{n-1}(w) - \beta'(w)F^{n-1}(w) = (v - w)dF^{n-1}(w),$$

so that surplus is single-peaked at $w = v$

Known solutions of the FPA beyond IPV

- Vickrey also solved some asymmetric examples, revenue equivalence breaks
- Wilson (1977) who constructed a monotonic pure strategy equilibrium of the **mineral rights model**, where there are pure common values, and bidders' observe the value plus iid noise
- When values are **interdependent** (as opposed to private), bidders must take into account the **winner's curse**, i.e., adverse selection from winning the auction:
  If you win when others' signals are low, then you will tend to update downwards about the value from the event that you win, and must bid less accordingly
- When value/noise distributions are fixed and $N \to \infty$, the bidders compete away their rents
- Both models later generalized by Milgrom and Weber's (1982) **symmetric affiliated values** model
- Equilibria have also been found for a handful of other cases...

Towards a theory with richer information

- The classical literature focused on settings with monotonic pure strategy equilibria
- This essentially limits us to environments with linearly ordered types
- In particular, for bidders with higher types to bid more, we in general need higher types to be associated with
  - higher expected values
  - higher conditional distributions of others' types (so that higher types $\implies$ others bidding more aggressively)
- This are strong assumptions... In practice, we would not think there is necessarily such a tight link between information about one's own value and about others information. What happens if types are multidimensional? Separate information about own and others' values? What kinds of behavior could we be missing?
- To analyze this question, we will study the BCE of the FPA

Common values

- Focus mostly on **common values**: $v_i = v \sim F$
- Assume that F has a strictly positive density $f(v)$

Common values

- Focus mostly on **common values**: $v_i = v \sim F$
- Assume that F has a strictly positive density $f(v)$
- No reserve price $\implies$ Total surplus is

$$\hat{v} = \int_{v=0}^{1} v f(v) dv$$

- The split between seller and buyers depends on information
- What is the range of things that might happen?

Example 1: public information

- Same information: $T_1 = T_2 = \cdots = T_n$
- $\eta(t|v) = 0$ if $t_i \neq t_j$ for any i and j

Example 1: public information

- Same information: $T_1 = T_2 = \cdots = T_n$
- $\eta(t|v) = 0$ if $t_i \neq t_j$ for any i and j
- The bidders compete away their rents and simply bid the interim expectation of their value given the public signal
- Revenue is $R = \hat{v}$ and $U_i = 0$

Example 2: Insider trading

- Engelbrecht-Wiggans, Milgrom, and Weber (1983)
- One bidder is informed:
  - $T_1 = [\underline{v}, \overline{v}]$
  - $|T_2| = \cdots = |T_n| = 1$
  - $t_1 = v$ with prob 1

Example 2: Insider trading

- Engelbrecht-Wiggans, Milgrom, and Weber (1983)
- One bidder is informed:
  - $T_1 = [\underline{v}, \overline{v}]$
  - $|T_2| = \cdots = |T_n| = 1$
  - $t_1 = v$ with prob 1
- There is an equilibrium in which the informed bidder bids

$$\beta(v) = \frac{1}{F(v)} \int_{w=\underline{v}}^{v} w\, f(w) dw$$

- Uninformed bidders "simulate" independent draws s_i from $(F(v))^{1/(n-1)}$, and bid $\beta(s_i)$

Example 2: Insider trading

- Engelbrecht-Wiggans, Milgrom, and Weber (1983)
- One bidder is informed:
  - $T_1 = [\underline{v}, \overline{v}]$
  - $|T_2| = \cdots = |T_n| = 1$
  - $t_1 = v$ with prob 1
- There is an equilibrium in which the informed bidder bids

$$\beta(v) = \frac{1}{F(v)} \int_{w=\underline{v}}^{v} w\, f(w) dw$$

- Uninformed bidders “simulate” independent draws s_i from $(F(v))^{1/(n-1)}$, and bid $\beta(s_i)$
- Interpretation: the high bidder bids as if they are in the two bidder IPV model, with values drawn from F
- In the aggregate, the uninformed bidders “simulate” a second value drawn from F, and bid as if it were their true value

Proof that this is an equilibrium

- That the informed bidder is playing a best response is immediate

Proof that this is an equilibrium

- That the informed bidder is playing a best response is immediate
- If an uninformed player bids $\beta(w)$, they win if they outbid the other uninformed bidders (which is independent of the value) and if the true value is less than w
- But when they win, they pay the expected value conditional on it being less than w, so the net surplus is zero! $\square$

Proof that this is an equilibrium

- That the informed bidder is playing a best response is immediate
- If an uninformed player bids $\beta(w)$, they win if they outbid the other uninformed bidders (which is independent of the value) and if the true value is less than w
- But when they win, they pay the expected value conditional on it being less than w, so the net surplus is zero! $\square$
- NB as in IPV, expected revenue is the expected second-highest of two draws from F, which is strictly below $\hat{v}$

Minimum revenue

- These examples show that there is a non-degenerate range of revenues in the common value FPA
- How low can revenue go? Can $R = 0$ if there is a common prior and a Nash equilibrium?

Minimum revenue

- These examples show that there is a non-degenerate range of revenues in the common value FPA
- How low can revenue go? Can $R = 0$ if there is a common prior and a Nash equilibrium? No!
- That could only happen if everyone bids zero and ties, so bidder surplus is $\hat{v}/N$
- But then any bidder could deviate to a bid of ϵ, and earn $\hat{v} - \epsilon$

Minimum revenue

- These examples show that there is a non-degenerate range of revenues in the common value FPA
- How low can revenue go? Can $R = 0$ if there is a common prior and a Nash equilibrium? No!
- That could only happen if everyone bids zero and ties, so bidder surplus is $\hat{v}/N$
- But then any bidder could deviate to a bid of ϵ, and earn $\hat{v} - \epsilon$
- This suggests that revenue should be bounded away from zero, because if the bid distribution is too low, then the temptation to bid higher will be too great:
  The probability of winning would increase too fast relative to the additional cost
- But what is the exact minimum?

Evidence from simulations

- On your problem set, I asked you to solve for the revenue minimizing BCE when the common value is standard uniform and there are two bidders
- In that exercise, you took a large discrete grid of values and bids
- The solution has a number of distinct features:

Evidence from simulations

- On your problem set, I asked you to solve for the revenue minimizing BCE when the common value is standard uniform and there are two bidders
- In that exercise, you took a large discrete grid of values and bids
- The solution has a number of distinct features:
  1. Minimum expected revenue is ≈ 0.161 when there are 51 values

Evidence from simulations

- On your problem set, I asked you to solve for the revenue minimizing BCE when the common value is standard uniform and there are two bidders
- In that exercise, you took a large discrete grid of values and bids
- The solution has a number of distinct features:
  1. Minimum expected revenue is ≈ 0.161 when there are 51 values
  2. Bids are independent

Evidence from simulations

- On your problem set, I asked you to solve for the revenue minimizing BCE when the common value is standard uniform and there are two bidders
- In that exercise, you took a large discrete grid of values and bids
- The solution has a number of distinct features:
  1. Minimum expected revenue is ≈ 0.161 when there are 51 values
  2. Bids are independent
  3. The high bid is perfectly comonotonic with the value

Evidence from simulations

- On your problem set, I asked you to solve for the revenue minimizing BCE when the common value is standard uniform and there are two bidders
- In that exercise, you took a large discrete grid of values and bids
- The solution has a number of distinct features:
  1. Minimum expected revenue is ≈ 0.161 when there are 51 values
  2. Bids are independent
  3. The high bid is perfectly comonotonic with the value
  4. The binding obedience constraints are precisely those associated with upward deviations

Evidence from simulations

- On your problem set, I asked you to solve for the revenue minimizing BCE when the common value is standard uniform and there are two bidders
- In that exercise, you took a large discrete grid of values and bids
- The solution has a number of distinct features:
  1. Minimum expected revenue is ≈ 0.161 when there are 51 values
  2. Bids are independent
  3. The high bid is perfectly comonotonic with the value
  4. The binding obedience constraints are precisely those associated with upward deviations
  5. Moreover, the Lagrange multiplier on the obedience constraint only depends on the deviation, not the recommendation!
- We will use these key pieces of evidence to reverse engineer the BCE and the proof of optimality

Constructing the revenue-minimizing BCE

- From the BCE logic, without loss to take signals equal to bids, b_i iid draws from G, the distribution in the simulation
- But we don't have to! Indeed, we could draw real signals s_i from another distribution $\tilde{G}$ for signals, and then there is a monotonic bidding function $\beta(s_i)$, and $G(\beta(s_i)) = \tilde{G}(s_i)$, e.g., we could take $\tilde{G}$ to be uniform if we want, and then $\beta = G^{-1}$
- Is there a more convenient choice of units?

Constructing the revenue-minimizing BCE

- From the BCE logic, without loss to take signals equal to bids, b_i iid draws from G, the distribution in the simulation
- But we don't have to! Indeed, we could draw real signals s_i from another distribution $\tilde{G}$ for signals, and then there is a monotonic bidding function $\beta(s_i)$, and $G(\beta(s_i)) = \tilde{G}(s_i)$, e.g., we could take $\tilde{G}$ to be uniform if we want, and then $\beta = G^{-1}$
- Is there a more convenient choice of units?
- The true value is comonotonic with the highest bid, so it's monotonic with the highest signal as well, i.e., $v = \lambda(\max_i s_i)$ for some non-decreasing λ, and $F(\lambda(\max_i s_i)) = (\tilde{G}(\max_i s_i))^n$

Constructing the revenue-minimizing BCE

- From the BCE logic, without loss to take signals equal to bids, b_i iid draws from G, the distribution in the simulation
- But we don't have to! Indeed, we could draw real signals s_i from another distribution $\tilde{G}$ for signals, and then there is a monotonic bidding function $\beta(s_i)$, and $G(\beta(s_i)) = \tilde{G}(s_i)$, e.g., we could take $\tilde{G}$ to be uniform if we want, and then $\beta = G^{-1}$
- Is there a more convenient choice of units?
- The true value is comonotonic with the highest bid, so it's monotonic with the highest signal as well, i.e., $v = \lambda(\max_i s_i)$ for some non-decreasing λ, and $F(\lambda(\max_i s_i)) = (\tilde{G}(\max_i s_i))^n$
- Seems natural to pick units for the signals so that SOMETHING is simplified... Since we don't even know G yet, might as well make λ the identity, in which case $\tilde{G} = F^{1/n}$, and the highest signal **equals** the highest value
- Then the winning bid is $\beta(\max_i s_i) = \beta(v)$!

The winning bid function

- But what is the correct choice of β?
- Well, it must be that bidders are indifferent between following their equilibrium strategy and **all** upward deviations
- Consider a bidder of type t_i who bids $\beta(s_i)$ for $s_i \geq t_i$
- The resulting payoff is

$$t_i F^{(N-1)/N}(s_i) + \int_{x=t_i}^{s_i} x dF^{(N-1)/N}(x) - \beta(s_i) F^{(N-1)/N}(s_i)$$

The winning bid function

- But what is the correct choice of β?
- Well, it must be that bidders are indifferent between following their equilibrium strategy and **all** upward deviations
- Consider a bidder of type t_i who bids $\beta(s_i)$ for $s_i \geq t_i$
- The resulting payoff is

$$t_i F^{(N-1)/N}(s_i) + \int_{x=t_i}^{s_i} x dF^{(N-1)/N}(x) - \beta(s_i) F^{(N-1)/N}(s_i)$$

- Why? Clearly they will win and pay $\beta(s_i)$ if and only if every other bidder's type is less than s_i
- And the value conditional on winning is the maximum of t_i and the highest of the others types

The winning bid function

- But what is the correct choice of β?
- Well, it must be that bidders are indifferent between following their equilibrium strategy and **all** upward deviations
- Consider a bidder of type t_i who bids $\beta(s_i)$ for $s_i \geq t_i$
- The resulting payoff is

$$t_i F^{(N-1)/N}(s_i) + \int_{x=t_i}^{s_i} x dF^{(N-1)/N}(x) - \beta(s_i) F^{(N-1)/N}(s_i)$$

- Why? Clearly they will win and pay $\beta(s_i)$ if and only if every other bidder's type is less than s_i
- And the value conditional on winning is the maximum of t_i and the highest of the others types
- Bidders are indifferent to upward deviations

$$\iff \beta(s_i) = \underline{\beta}_i(s_i) \equiv \frac{1}{F^{(N-1)/N}(x)} \int_{x=\underline{v}}^{s_i} x dF^{(N-1)/N}(x)$$

i.e., the Vickrey equilibrium as if types were private values!

Minimum expected revenue

- In the uniform case, we would have

$$\underline{\beta}(s_i) = \frac{1}{s_i^{\frac{N-1}{N}}} \int_{x=0}^{s_i} x d(x^{(N-1)/N}) = \frac{N-1}{2N-1} s_i$$

- Since the high signal is uniform (and is equal to the value!) we get that expected revenue is

$$\underline{R} = \frac{1}{2} \frac{N-1}{2N-1}$$

- When $N = 2$, $\underline{R} = 1/6 \approx 0.167$, not far from the numerical simulation

Interpretation

- Why would this structure minimize expected revenue?
- Well, it exhibits an extreme form of the winner's curse
- Conditional on one's own signal being s_i, all you learn is that the value is in $[s_i, \overline{v}]$
- But winning when the signal is s_i means that the value is exactly s_i!
- So, winning is very bad news about the value, so much so that bidder's behave as if their signal is the true value, even though it is only a lower bound
- Still, we have yet to establish that this is in fact minimum revenue...

A formal result

Theorem

Minimum revenue in the FPA across all BCE is $\int_v \underline{\beta}(v)F(dv)$.

- To proof the theorem, we need to use what we learned about the optimal Lagrange multipliers, which are positive only for upward deviations, and depend only on the deviation, not the recommendation!
- In our unit on linear programming, I observed that if we aggregate constraints with weights proportional to the multipliers, the optimal value will not change
- In this case, the simulation is telling us that we can add together all the constraints that deviate to b_i from recommendations $b_i' < b_i$ without changing the value

A formal result

Theorem

Minimum revenue in the FPA across all BCE is $\int_v \underline{\beta}(v)F(dv)$.

- To proof the theorem, we need to use what we learned about the optimal Lagrange multipliers, which are positive only for upward deviations, and depend only on the deviation, not the recommendation!
- In our unit on linear programming, I observed that if we aggregate constraints with weights proportional to the multipliers, the optimal value will not change
- In this case, the simulation is telling us that we can add together all the constraints that deviate to b_i from recommendations $b_i' < b_i$ without changing the value
- This corresponds to a **uniform upward deviation** to b_i: Bid the maximum of your recommendation and b_i
- The simulation is telling us to ignore all other deviations!

Symmetry

- It will be easier to apply these deviations if we first observe that when minimizing revenue, it is without loss to look at symmetric BCE
- Why? The obedience constraints are linear, so if μ is the revenue minimizing BCE, then so is the distribution μ' obtained by "permuting" the bidders identities, i.e., giving bidder i's recommendation to bidder $\xi(i)$ for some permutation ξ
- Obviously, the permutation does not change expected revenue, and by randomizing over all permutations, we obtain a symmetric revenue-minimizing BCE

Winning bid distributions

- Now, write $H(b|v)$ for the distribution of the winning bid conditional on the common value v
- We wish to show that the H that minimizes expected revenue subject to obedience must place probability one on $b = \underline{\beta}(v)$
- Notice that by symmetry, each bidder's equilibrium surplus must be

$$\frac{1}{N}\int_v \int_{b'} (v - b')H(db'|v)F(dv)$$

- Now, what would a bidder's surplus be from a uniform deviation up to b? Assuming that there is zero probability of a tie, it must be

$$\int_v \left((v-b)H(b|v) + \frac{1}{N}\int_{b' \geq b} (v - b')H(db'|v) \right) F(dv)$$

- So, the uniform upward constraint is just that

$$\frac{1}{N}\int_v \int_{b' \leq b} (v - b')H(db'|v)F(dv) \geq \int_v (v-b)H(b|v)F(dv)$$

Comonotonicity

$$\frac{1}{N}\int_v \int_{b' \leq b} (v - b') H(db'|v) F(dv) \geq \int_v (v - b) H(b|v) F(dv)$$

- We can IBP to get

Comonotonicity

$$\frac{1}{N}\int_v \int_{b'\leq b}(v-b')H(db'|v)F(dv) \geq \int_v (v-b)H(b|v)F(dv)$$

- We can IBP to get

$$\frac{1}{N}\int_v \int_{b'\leq b} H(b'|v)db'F(dv) \geq \frac{N-1}{N}\int_v (v-b)H(b|v)F(dv)$$

Comonotonicity

$$\frac{1}{N}\int_v \int_{b' \leq b} (v - b')H(db'|v)F(dv) \geq \int_v (v - b)H(b|v)F(dv)$$

- We can IBP to get

$$\frac{1}{N}\int_{b' \leq b} H(b')db' \geq \frac{N-1}{N}\left(\int_v vH(b|v)F(dv) - bH(b)\right)$$

- So the only piece that depends on the correlation between v and b is $\int_v vH(b|v)F(dv)$

Comonotonicity

$$\frac{1}{N}\int_v \int_{b' \le b} (v - b')H(db'|v)F(dv) \ge \int_v (v - b)H(b|v)F(dv)$$

- We can IBP to get

$$\frac{1}{N}\int_{b' \le b} H(b')db' \ge \frac{N-1}{N}\left(\int_v vH(b|v)F(dv) - bH(b)\right)$$

- So the only piece that depends on the correlation between v and b is $\int_v vH(b|v)F(dv)$
- Holding $H(b)$ fixed, this term is minimized by making $H(b|v)$ as large as possible for small v and small for large v, i.e.,

$$H(b|v) = \begin{cases} 1 & \text{if } F(v) \le H(b) \\ 0 & \text{if } F(v) > H(b) \end{cases}$$

- In other words, v and b should be comonotonic! And put probability one on $\beta(v) = \min\{b|H(b) \ge F(v)\}$

The optimal winning bid distribution

- We can rewrite the constraint for a uniform deviation up to $\beta(v)$:

$$\int_{w \leq v} (w - \beta(v)) F(dw) \leq \frac{1}{N} \int_{w \leq v} (w - \beta(w)) F(dw)$$

  which rearranges to

$$\beta(v) \geq \Lambda(\beta)(v) \equiv \frac{1}{F(v)} \int_{w \leq v} \left(\frac{N-1}{N} w + \frac{1}{N} \beta(w) \right) F(dw)$$

- Facts:
  1. Λ is monotonic
  2. Λ is a contraction (in the sup norm) of modulus $1/N$
  3. β satisfies the uniform upward constraints iff $\beta \geq \Lambda(\beta)$

Final step

- So, if $\beta \geq \Lambda(\beta)$, then by induction, the sequence $\Lambda^k(\beta)$ is decreasing in k, and therefore also satisfies obedience
- Moreover, by the Banach fixed point theorem, $\Lambda^k(\beta)$ is converging to the unique fixed point of Λ
- It is straightforward to verify that this fixed point is $\underline{\beta}$
- Thus, $\underline{\beta}$ is lower than any other winning bid distribution that satisfies uniform upward constraints, and hence must be the revenue minimizing winning bid function $\square$

Robust surplus extraction in the large

- Recall that minimum revenue in the uniform example is $(N-1)/(4N-4)$
- NB This asymptotes to $1/4$, which is less than the total surplus of $1/2$, so bidders still get rents in the limit
- This is also true in the EMW model of proprietary info
- Very different from older positive results of Wilson (1977) and Milgrom (1979) that show asymptotic full surplus extraction of the FPA in the mineral rights model
- Begs the question, are there other mechanisms that asymptotically extract more revenue?
- Du (2018) constructs a sequence of mechanisms that asymptotically have revenue equal to $\hat{v}$, regardless of F
- He similarly uses BCE to construct lower bounds on revenue for these mechanisms that converges to $\hat{v}$
- We will return to this topic in our units on robust auction design

Generalizations

- BBM '17 generalize to interdependent values, where the bidders have possible different values $(v_1, \ldots, v_n)$
- Much of the analysis goes through with the role of the common value being played by the average of the $n-1$ lowest value, as long as values are **exchangeable**
- BBM '17 also look at other objectives:
  - Max revenue: Just the expected highest value
  - Min welfare: When $N = 2$ and values are iid, there is a BCE in which the bidder with the **lower** value always wins!

Known values

- Thus far, we have consider the case of **unknown values**, where each bidder may not know their own value for the good
- What happens with **known values**, i.e., each bidder knows (at least) their own value?
- Binary values: $v_i \in \{\underline{v}, \overline{v}\}$,
  - BBM 2021 show that the same techniques can be used to pin down minimum revenue/maximum bidder surplus
  - Dual interpretation: Bertrand competition, unknown number of firms, as in consumer search a la Varian (1980) and Burdett and Judd (1983)
- Maximum revenue is now non-trivial (subject to bidders using weakly undominated strategies)
  - Can use techniques of BBM '15 "The limits of price discrimination" to characterize maximum revenue

Open questions

- The literature has progressed slowly, mostly under the Bertrand interpretation
- Minimum revenue/max bidder surplus for asymmetric value distributions
- Minimum total surplus in general
- Asymmetric objectives (maximizing one bidder's objective)
- Minimum revenue/max bidder surplus beyond binary values, and for asymmetric known values

Broader applications

- In your final project for the course, I will ask you to apply some of the techniques we have explored to a novel problem
- One way to go would be to analyze BCE and robust predictions in a new game that has not been previously studied
- Game theory is littered with settings involving incomplete information which may benefit from the informationally robust perspective
- Beyond FPA/Bertrand competition, there is a host of other natural problems, including other auction formats, other forms of competition, investment games, electoral competition, bargaining, public goods, etc, almost all of which remain unexplored!
- Part of the challenge is identifying the right set of assumptions so that the extremal BCE are tractable
- A good rule of thumb is: If the solution is too complicated, make the model simpler!