Experiments in Legislative Bargaining (and a few research questions)

Duk Gyoo Kim

Yonsei University

March 2026

The purposes of this talk

- To introduce myself;
- To present one published paper that may have some low-hanging fruits following it;
- To provide some short research projects that you might be interested in (if time permits).

About myself

- Duk Gyoo KIM (kim.dukgyoo at yonsei.ac.kr)
- Office hours: Thursday 13:00–15:00 and 16:00–17:00

  To reserve: https://calendly.com/dukgyookim/officehours
  (Do so at least 12 hours before you plan to come.)
- Cornell Econ PhD, Caltech Postdoc (2yrs), AP at Mannheim, Germany (5yrs), and at Sungkyunkwan (3yrs).
- (Micro-based) public econ, political economy, behavioral econ. Experiments as a methodology

Experiments in Political Economy

From a perspective of an economist

- Some economists apply economic theories into the context of political economy.
- A subset of them combine theories with experiments. I am one of the few.
- I hope this talk to be a chance to introduce how economists experimentally study political economy.

Legislatures

- A policy is typically made by a legislature.
- Legislators are assumed to have different policy preferences because they are elected by different constituencies (e.g., different geographical districts).
- Consider a legislature consisting of n legislators indexed by i.
- Assume that n is odd and ≥ 3, and the legislature operates by majority rule.
- The legislators have to choose some policy p from some set of alternative policies P.
- Legislator i's utility if policy p is selected is $V_i(p)$.
- This utility function will reflect both the legislator's ideology and how the policy will impact him and his constituents.

Distributive Politics with no Condorcet Winners

- A policy $p^* \in P$ is said to be a *Condorcet Winner* if it would defeat or tie any other policy in a pairwise majority vote.
- If a unique Condorcet winner exists, it may be reasonable to think that the legislature would select it. The Condorcet winner exists
  - when a policy choice is binary,
  - when a preferred policy is on a single dimension,
- I am interested in one particular situation where a Condorcet winner doesn't exist: distributive politics

(Simplified) Distributive Politics

- Suppose that the legislators divide a fixed budget, normalized to 1, between projects located in the legislators' districts.
- A policy is a vector $p = (p_1, \cdots, p_n)$, where p_i is the spending in legislator i's district. $P = \{p \in [0,1]^n : \sum_i p_i = 1\}$.
- Assume that $V_i(p) = p_i$. Each legislator would prefer that all the projects are located in his own district.

The Legislative Bargaining Model

Baron and Ferejohn (1989)

- The legislative bargaining model builds on the agenda-setter model (Romer and Rosenthal, 1978).
- Consider a many-person divide-the-dollar game.
- One legislator is randomly selected to make a proposal about how to split a dollar, $p = (p_1, \ldots, p_n) \in [0,1]^n$, $\sum p_i = 1$.
- All legislators then vote for or against the proposal.
- If $q \in [1, n]$ or more legislators vote for the proposal, it is implemented. Otherwise, another legislator is randomly selected to make a policy proposal and repeat the proposal–voting procedure until the proposal is approved.

  ($q = 1 \Leftrightarrow$dictator. $q = n \Leftrightarrow$unanimity. $q = \frac{n+1}{2} \Leftrightarrow$majority.)
- Legislators' policy payoffs are discounted by δ^{t-1} if the policy is implemented in round t. $\delta \in (0,1]$.

The Legislative Bargaining Model

Baron and Ferejohn (1989)

Stationary Subgame-Perfect Equilibrium (SSPE)

- It is known that for sufficiently large δ, folk theorem applies: Virtually all allocations can be the equilibrium outcomes.
- Focus on stationary subgame-perfect strategies. SSPE is described by a player's proposal strategy and voting strategy.
- When selected as a proposer, the player offers $\frac{\delta}{n}$ to $q-1$ randomly-selected players and keeps $1-\delta\frac{q-1}{n}$ to herself.
- Nonproposers vote for a proposal if the offered payoff is $\geq \frac{\delta}{n}$.
- The first proposal is approved.

If $\delta = 1$, $n = 3$, and $q = 2$, the proposer offers one member $\frac{1}{3}$ and keeps $\frac{2}{3}$. Here $\frac{1}{3}$ is the continuation value, the expected payoff of rejecting the offer and moving on to the next round.

The Legislative Bargaining Model: Experiments

Many interesting theoretical predictions. Changes in n, changes in q, changes in δ, heterogeneity in δ, proposer selection rule, length of potential bargaining rounds, etc. Refer to Baranski and Morton (2022) if interested. Main experimental findings are

- Minimum winning coalitions are most frequently observed.
- Delay becomes infrequent with experience.
- Proposer's partial rent extraction: There is a significant proposer power, but it is less than predicted by theory.

Roughly speaking, when $n = 3$, $q = 2$, and $\delta = 1$, typically observed allocation is about $(0.55, 0.45, 0)$ while theory predicts $(0.66, 0.33, 0)$.

The Legislative Bargaining Model: Experiments

Some of my work in this literature

- Kim (2019): If the proposer is randomly selected, but the previous proposers are not selected until everyone has the same number of proposal opportunities, then the 'partial proposer advantage' can be explained.
- Kim (2023): Among the possible reasons behind the 'partial proposer advantage,' fairness concern is not at all the driver.
- Kim and Kim (2022): Experimental examination of the effect of the competition to be selected as the proposer in a subsequent multilateral bargaining game.
- Kim and Lim (2024): The many-person divide-the-dollar game is fundamentally different from the many-person divide-the-penalty game.
- Baranski and Kim (2024): Experimental examination of the allocation of costs resulting from a negative externality, such as pollution-induced economic costs.

The Legislative Bargaining Model: Experiments

Some of my work in this literature

- Kim (2019): If the proposer is randomly selected, but the previous proposers are not selected until everyone has the same number of proposal opportunities, then the 'partial proposer advantage' can be explained.
- Kim (2023): Among the possible reasons behind the 'partial proposer advantage,' fairness concern is not at all the driver.
- Kim and Kim (2022): Experimental examination of the effect of the competition to be selected as the proposer in a subsequent multilateral bargaining game.
- Kim and Lim (2024): The many-person divide-the-dollar game is fundamentally different from the many-person divide-the-penalty game.
- Baranski and Kim (2024): Experimental examination of the allocation of costs resulting from a negative externality, such as pollution-induced economic costs.

Motivation

Many-person bargaining is everywhere, including:

- Budget appropriations among legislators
- Setting tax rates to different groups
- Free trade agreement among many countries
- Climate change summit to set the CO2 level of each country
- Search committee hiring a junior
- Location of noxious facilities
- Condominium board meeting for the use of common area

These may be modeled by a many-person divide-the-dollar game, but...

Motivation

Many-person bargaining is everywhere, including:

- Budget appropriations among legislators
- Setting tax rates to different groups
- Free trade agreement among many countries
- Climate change summit to set the CO2 level of each country
- Search committee hiring a junior
- Location of noxious facilities
- Condominium board meeting for the use of common area

the highlighted ones are about dividing a kind of losses, penalties, or burdens. Are these the same?

Key illustration

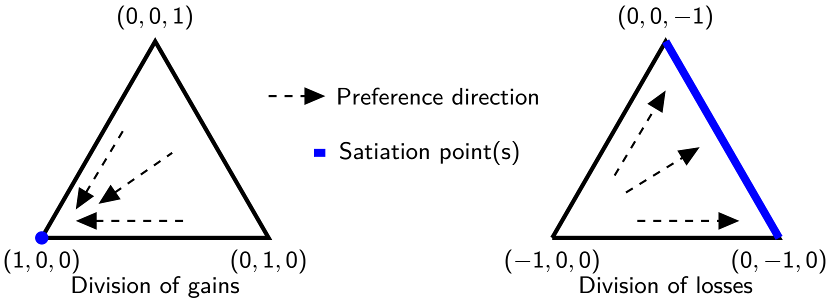


Figure: Different preference directions and satiation points

It is not about loss 'domain'

Adding endowment to shift the game to a gain domain doesn't affect the key difference.

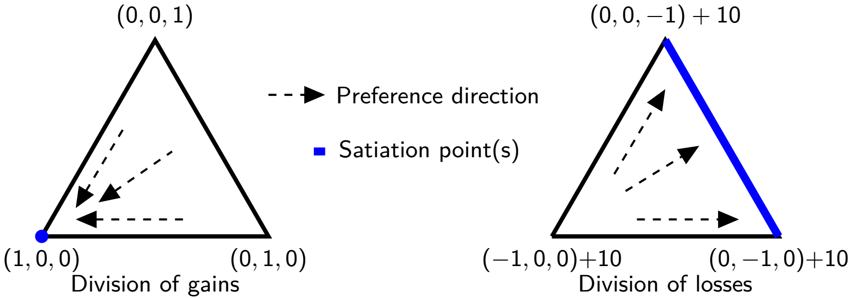


Figure: Different preference directions and satiation points

Preview of Results

- Unlike the divide-the-dollar game, the stationary equilibrium of the divide-the-penalty game is no longer unique in payoff.
- Among the continuum of the stationary equilibria, an equilibrium at one extreme is a mirror image of that in the divide-the-dollar game.
- Experimental evidence is consistent with the other extreme.
- Many interesting studies in multilateral bargaining on a gain domain are worth revisiting.

Plan of this talk

1. A quick recap of many-person divide-the-dollar (DD) game
2. A many-person divide-the-penalty (DP) game
3. Experiments
4. Results
5. Discussions

A quick recap of many-person divide-the-dollar (DD) game

Assume three members, a simple majority rule, and $\delta \leq 1$.

- In round t, a randomly selected member proposes how to split a surplus normalized to 1.
- The proposal is voted on. If two or more members vote for it, the proposal is implemented, and the game ends.
- If the proposal is rejected, they repeat the process, with a new proposer randomly selected. Payoff discounted by δ^{t-1}.
- When disagreed infinitely, everyone gets zero.

In the stationary subgame-perfect equilibrium, when $\delta = 1$, the randomly-selected proposer (say, member 1) in the first round offers (2/3, 1/3, 0).

A quick recap of many-person divide-the-dollar (DD) game

Theoretical predictions (Stationary Subgame-Perfect Equilibrium) and experimental findings

- **Minimum winning coalition**: A proposer forms a winning coalition whose size is the smallest for approval.
  $\Rightarrow$ Experimental evidence 👍
- **Full rent extraction of the proposer**: The proposer offers MWC members their continuation value, offers other members zero, and keeps all the remainder.
  $\Rightarrow$ Experimental evidence 👍👎
- **Full (utilitarian) efficiency**: The first-round proposal is approved without a delay.
  $\Rightarrow$ Experimental evidence 👍
- Uniqueness in payoff: The SSPE is essentially unique.

A many-person divide-the-penalty (DP) game

Assume 3 members, a simple majority, and $\delta \geq 1$.

1. In round t, the randomly recognized member proposes an allocation of -1 in terms of proportions: $p \in \{(p_1, p_2, p_3) | p_i \in [0, 1], \sum p_i = 1\}$
2. If 2 or more members vote for the proposal, the proposal is implemented, $-\delta^{t-1} p_i$ is accrued, and the game ends. If not, the game moves on to round $t + 1$.
3. In round $t + 1$, a new player is randomly recognized as the proposer. The game repeats at $t + 1$.
4. When disagreed infinitely, everyone gets -1.

Multiple Stationary Subgame-Perfect Equilibria

Focus on stationary strategies only. SSPE is unique in payoff in the DD game, but not in the DP game. For illustration, consider $n = 3$, $\delta = 1$, and a simple majority.

- One equilibrium allocation is $(0, -1/3, -2/3)$. We call it Most Egalitarian (ME).
- Another equilibrium allocation is $(0, 0, -1)$. We call it Utmost Inequality (UI).
- Anything between ME and UI is a SSPE. For example, $(0, -0.1, -0.9)$ is another equilibrium allocation.

There are a lot more to say in theory.

Why experiment?

- Equilibrium selection: We honestly do not know which equilibrium will be selected.

$\Rightarrow$ If the observed patterns are similar to the mirror image of the SSPE on the division of gains, "we are good to go."

$\Rightarrow$ Otherwise, it is worth revisiting all the important studies on a gain domain, if the main motivating situations are about dividing a loss.

(Part of) Experiments

Table: Experimental Treatments

		Voting Rule	
		Majority	Unanimity
Group Size	3	***M3***	U3
	5	***M5***	U5

- 12 to 15 Games ("Days") to divide $-50 * n$ tokens. $\delta = 1.2$, in the second round ("Meeting") divide $-50 * \delta * n$.
- 400 tokens endowed each Day. Worth discussing, but the results (both theoretical and experimental) are not affected.
- At HKUST. Random rematch. Between-subject. Random payment. 271 subjects. 16 $(= 4 \times 4)$ session

Predictions recap, Majority

ME equilibrium is the mirror image of the SSPE in the DD game.

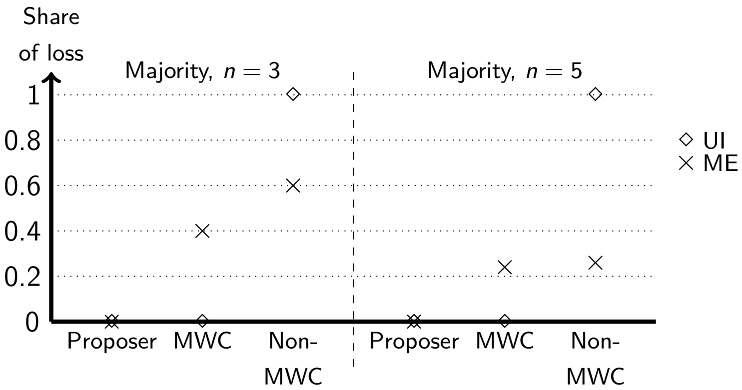


Figure: Hypotheses from theoretical predictions

Results

UI equilibrium is clearly dominant. MWC is clearly formed.

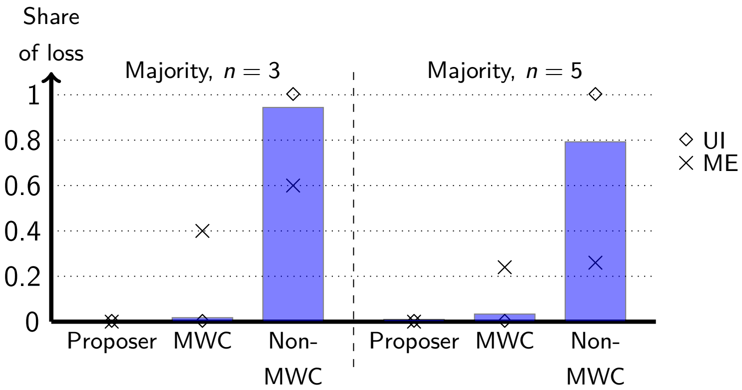
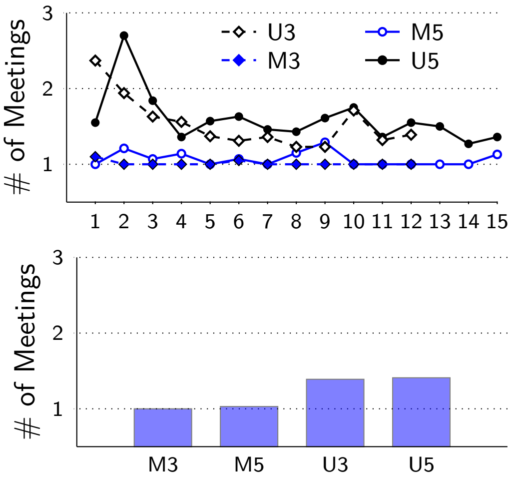


Figure: Proposed Shares, Majority

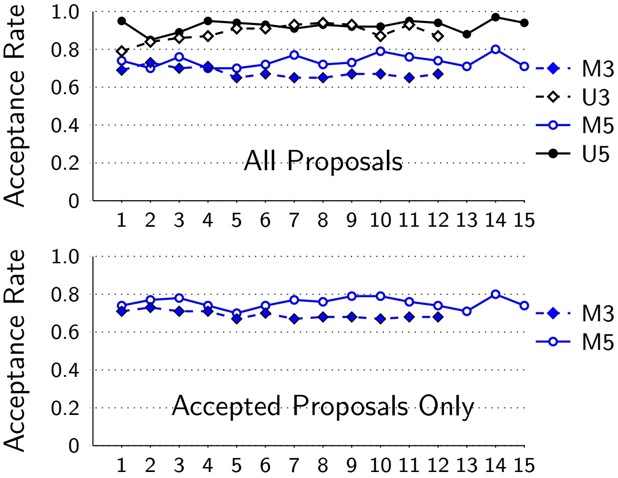
Results

Agreement is made without (or with little) delay.



Results

In M5, the size of the winning coalition is larger.



Take-away messages

- DD vs. DP is not like a half-full vs. a half-empty cup of water, but like a cup of clean water vs. a cup of filthy water.
- Theoretical predictions, contrary to our naïve conjectures, are quite different.
- The most unequal allocation is accepted under a majority. This has little to do with fairness concern.

If the main motivation for studying legislative bargaining model is to gain insights about distributive politics on the losses, think twice before applying the many-person divide-the-dollar game.

Low-hanging fruits

- I introduce this paper because there might be low-hanging fruits that graduate students can give a try.
- Note that many multilateral bargaining games considered in the literature have the distribution of penalties and losses in mind, but their theoretical predictions are clearly missing other equilibria.
- It is worth revisiting all the important studies on a gain domain, if the main motivating situations therein were about dividing a loss.

One of low-hanging fruits: Allocation of Pollution Costs

Baranski and Kim (2024)

How to allocate the costs of public bads? Public bads are byproducts of individuals' self-interested actions;

- The budget deficits due to the excessive private spending for locally-targeted projects $\Rightarrow$ insufficient public infrastructure.
- Firms' and countries' profit-maximizing decisions $\Rightarrow$ GHG emissions.

The costs induced by public bads need to be charged in a certain form of allocation of burdens (e.g., taxes), often resulting from political bargaining.

Research Question: How does **posterior bargaining of the costs** affect the voluntary production of public bads?

- How would a voting rule affect?
- How would a size of negative impact to the society affect?

In a nutshell, we studied

A two-stage game where:

- In stage 1, everyone pollutes to increase the private gain. The pollution cost is determined.
- In stage 2, they play the divide-the-penalty game, where the penalty is the pollution cost.

  (Notice that there are multiple equilibria in stage 2, which opens many possibilities of stage-1 behavior.)

In a nutshell, we found

- Pollution is very hard to deter.
- Both theoretically and experimentally, unanimity doesn't help to reduce pollution. ("If you mess up that much, why wouldn't I, who has the same veto power, do that much?")
- A "threat" to allocate the entire costs to one member whose environmental harm was the largest doesn't help to reduce pollution. ("If everyone fully pollutes, it all boils down to a lottery. Why would I reduce pollution?")
- An increase in α (severity parameter of pollution costs) reduces pollution to some degree in both majority and unanimity voting rules, but it decreases the agreement rate.

(If time permits) potential research project (1/3)

Ambiguous Effort, Public Goods, and Redistribution

The stylized observation

Standard theory suggests monetary incentives should align strictly with individual performance. Yet, in many firms, managers prefer "flattened" wage structures.

Why?

- Individual performance metrics are imperfect.
- Low individual output is **ambiguous**:
  1. **Shirking:** The employee is lazy.
  2. **Pro-social Sacrifice:** The employee spent time helping others (mentoring, team culture, securing external sources) at the their own expenses. Public goods increase probabilistically.

Research Gap: How do third parties (spectators) redistribute when they cannot distinguish between a "Free Rider" and an "Unlucky Altruist"?

Literature Review

- **Meritocratic Fairness** (Konow, 2000; Cappelen et al., 2007, 2013)
  - Distinguishes *Effort* vs. *Luck*.
  - **Limitation:** Effort is usually observable in these designs.
- **Social Risk & Betrayal** (Bortolotti et al., 2025)
  - Finds that people redistribute more when inequality is caused by others (social risk) than by nature.
  - **Limitation:** Focuses on *victims* of social risk, not *contributors* to it.
- **Potential Contribution**
  - We introduce an **Identification Problem**.
  - We test if spectators give the "benefit of the doubt" to low performers who might be hidden public-good providers.

Experimental Design: The Production Phase

The Setup:

- Groups of 3 players.
- Fixed Time Budget: 10 slots per player.

Three Tasks:

- **Task A: High Effort (H)**
  - 12 tokens per each time budget.
  - Requires a high level of effort. Private benefit, no risk.
- **Task B: Low Effort (L)**
  - 4 tokens per each time budget.
  - Requires a zero or low level of effort. Private benefit, no risk.
- **Task B: Public Good (P)**
  - No guaranteed tokens.
  - Requires a high level of effort (same as H)
  - 24 tokens for *everyone* if successful with a 20% success rate.

Everyone becomes a spectator for other groups.

After the tasks are done, each player works as a Spectator for other K different groups.

- Based on the observations (varying by treatment), a spectator decides how to reallocate the payoffs of the *other* groups.
- The computer randomly selects one of their K decisions to implement.
- No portfolio effects, no incentives to misreport (although unincentivized)

Why would the spectator role be uneasy? Suppose the spectator observes someone got a low payoff. Is this person lazy? Or an unlucky public-good contributor?

Main Hypothesis: Ambiguity increases redistribution compared to transparency.

Treatments

We manipulate the information available to the Spectator.

Treatment	Information Set
T1: Max Ambiguity	Spectator sees only final payoffs.
T2: Partial Ambiguity	Spectator sees final payoffs AND total group P tasks.
T3: No Ambiguity	Spectator sees final payoffs AND task choices.

All that I wrote above are tentative: We can think of many variations (e.g., payoff function of each task, competition aspect, interaction of public goods (resources) and rewards of individual tasks), and different approaches (e.g., instead of being a spectator, allocation rule can be collectively determined before the tasks.)

(If time permits) potential research project (2/3)

The Illusion of Randomness in Economic Choice

Humans are bad at generating randomness.

- When asked to "randomly" pick a number between 1 and 10, we systematically avoid "easy"(5) or "salient"(1 and 10) options, creating predictable patterns.
- Many (23–33%) pick 7. Another 'random-looking' num is 3.

Why does this matter for economics?

- Economic theory relies heavily on the assumption that agents can randomize effectively, especially when asked to play a **mixed strategy**. Consider games in sports (a kicker and a goalie), tax compliance (IRS and tax evaders), and security patrols (attackers and defenders).
- **Hypothesis:** We can **exploit** those who fail to randomize.

From Psychology: Perceived Randomness

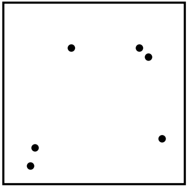
True Randomness (Poisson Process):

- Clusters happen naturally.
- Gaps are irregular.

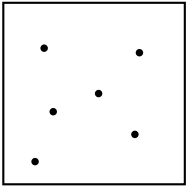
Human “Randomness”:

- We believe randomness means “evenness.”
- We actively avoid clusters (the “Gambler’s Fallacy”).
- We create “repulsion zones” around past events.

True Random



Human “Random”



The Identification Problem

When we observe a person avoiding a "cluster" in a Hide-and-Seek game, why do they do it?

- **Possible Channel 1: Cognitive Bias**: "I feel that a random distribution *must* be spread out. I am avoiding the cluster because it looks 'wrong' to me." → **The Gambler's Fallacy**
- **Possible Channel 2: Strategic Belief**: "I believe *my opponent* will look in the cluster. I am avoiding it to outsmart them." → **Level-k Reasoning**

To separate these, we need a 2×2 factorial design.

Experimental Design

We vary two dimensions: **Visual History** (Information) and **Opponent Type**.

	Opponent: Algorithm (Plays Uniform Nash)	**Opponent: Human** (Plays Strategically)
No History	Cell 1: Baseline	Cell 2: Blind Game
History	Cell 3: Minefield	Cell 4: The Duel

- Between "Baseline" and "Minefield": If subjects avoid craters here, it increases the predictability of their actions, potentially reducing their payoffs due to the cognitive bias.
- Between "Minefield" and "Duel": Do subjects avoid craters *more* when playing with a human?

Dimension 1: The Opponent Type

Treatment A: The Computer Opponent

- Subjects are explicitly told: "The seeker is a computer program that picks locations **completely at random**."
- **Optimal Strategy:** Pick any point $x \in [0, 1]^2$ with equal probability.
- **Prediction:** Behavioral agents will still exhibit "repulsion effect," reducing their effective hiding space.

Treatment B: The Human Opponent

- Subjects play against another subject in the lab.
- **Optimal Strategy:** Depends on beliefs about the opponent's sophistication.
- **Prediction:** Agents will engage in "Level-k" thinking (e.g., "He thinks I won't hide near the old bomb, so I should!").

Theoretical Framework: Level-k in Space

For the Duel, we model behavior using Level-k (still in progress):

- **Level-0 (Random):** The Level-0 seeker picks a uniform random point.
- **Level-1 (The "Biased" Hider):** best responds to Level-0, but has the "Repulsion Bias." $\Rightarrow$ Doughnut-shaped density around craters.
- **Level-2 (The "Smart" Seeker):** believes Hider is Level-1, so try to look more in the spaces with no old craters.
- **Level-3 (The "Contrarian" Hider):** believes Seeker is Level-2 (looking in gaps), so try to hide near to the old craters.

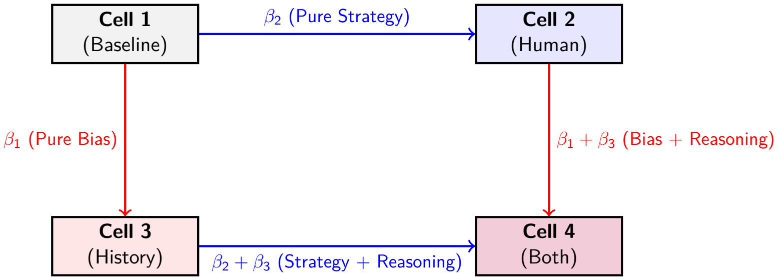
Hypotheses

- **H1: The illusion of randomness (Cell 1 and Cell 2)**: Hiders will avoid corners, edges, and centers.
- **H2: The further illusion of randomness (Cell 1 vs. Cell 3)**: Even against a computer player, Hiders will avoid "old craters."
- **H3: Strategic Sophistication (Cell 4 vs. Cell 3)**: Hiders with a human Seeker will exhibit **greater** repulsion from craters.
- **H4: Marginal effects (Cell4-Cell3-Cell1 or Cell4-Cell2-Cell1)**: See next slide.

Potential Implications We can learn more about what the system can exploit humans with randomness illusion.

Decomposing the Drivers: Bias vs. Strategy

We can isolate the "Primary Driver" by comparing the marginal effects along two pathways.



- **Comparing β_1 vs. β_2:** Is the visual cue (red arrow) stronger than the strategic context (blue arrow)?
- **Testing β_3:** Does facing a Human make the Repulsion bias worse?
  - *Hypothesis 4:* $\beta_3 > 0$. Human opponent's (possibly limited) reasoning amplifies the perceived salience of the craters.

(If time permits) potential research project (3/3)

The Beauty Premium in Academia: Does Research Field Matter?

The stylized observation

A well-documented "beauty premium" exists in the labor market. The economics job market offers a unique natural experiment: highly standardized timing, with CVs and profile pictures universally available online.

Why does appearance matter in academia?

- **Halo Effect:** Attractiveness is subconsciously linked to competence.
- **Interpersonal Skills:** Attractive individuals often develop better communication and networking skills over time.

Research Gap: Existing studies treat academic economics as a single, homogeneous profession. But does the beauty premium vary by the *nature* of the research? Specifically, what happens when a field requires more "sales" and persuasion?

Literature Review

- **Looks and Academic Career** (Hale et al., JEBO 2023)
  - Finds that attractive PhD graduates place better and receive more citations.
  - **Limitation:** Focuses on Top-10 graduates overall and does not distinguish the required skill sets across sub-fields.
- **Task-based Beauty Premium** (Stinebrickner et al., 2019)
  - Shows the beauty premium exists mainly in jobs requiring interpersonal interaction, not in pure information/data-processing tasks.
- **Potential Contribution**
  - We introduce **Task Heterogeneity** into the academic job market.
  - We test if spectators (hiring committees, conference attendees) give a higher premium to candidates in *Applied Economics* (requiring "sales" and policy persuasion) compared to *Theoretical Economics* (pure logic).

Empirical Design: Data Construction

Data Sources:

- Scrape profiles of Job Market Candidates (JMCs) across Top 100 institutions.
- Collect Job Market Papers (JMPs), CVs, and placement outcomes.

Key Variables Construction:

- **Attractiveness ($Beauty_i$)**
  - Instead of human RAs, utilize Machine Learning (e.g., Face API) to generate objective attractiveness and perceived competence scores from profile pictures.
  - Other characteristics (e.g., trustworthiness, prudence, confidence) can be measured as well.
- **Research Type ($Applied_i$)**
  - Extract **JEL Classification Codes** from JMPs.
  - Classify as *Theory/Method* (e.g., C: Mathematical methods, D8: Information) vs. *Applied/Policy* (e.g., J: Labor, I: Health/Education).

Empirical Strategy and Hypotheses

We rely on an interaction model to test the heterogeneous effect of appearance.

$$\begin{aligned} Placement_i = &\beta_0 + \beta_1 Beauty_i + \beta_2 Applied_i \\ &+ \beta_3 (Beauty_i \times Applied_i) + \gamma X_i + \epsilon_i \end{aligned}$$

(* Different dependent variables can be considered.)

Hypotheses:

- **H1: The Baseline Premium:** $\beta_1 > 0$. Even in theory fields, a baseline beauty premium might exist due to unconscious bias in interviews.
- **H2: The "Sales" Channel (Main Hypothesis):** $\beta_3 > 0$. The premium is significantly larger for applied researchers, where presenting, persuading, and interacting with the public are crucial components of the job.

Note: We can easily extend this to long-term outcomes: Number of seminar flyouts, co-authorship networks, or citation counts, etc.