# Lecture 6
# Behavioral conflict and violence

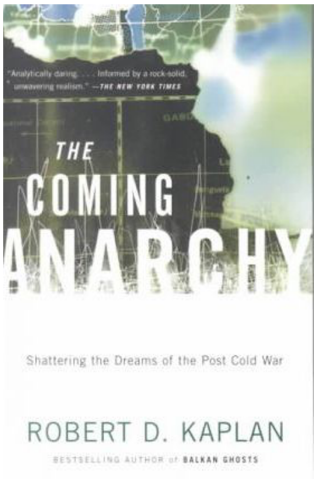
Christopher Blattman

19 May 2021

# Finishing up our typology of bargaining failures

1. Commitment problems
2. Incomplete information + incentives to misrepresent
3. Agency problems
   - 3.1 Absence of formal institutional checks
   - 3.2 Absence of informal checks (social norms & preferences)
   - 3.3 Absence of economic incentives/integration
4. Non-standard preferences
   - 4.1 Glory / relative status
   - 4.2 Utility from violence
   - 4.3 Violence as a means to an end
   - 4.4 Fairness & reciprocity
5. Misperception
   - 5.1 Errors in belief formation
   - 5.2 Decision-making under arousal

# The rationalist push to studying violence was partly a reaction to more lurid explanations




- Common journalistic explanations for war and terrorism include:
  - State control and communal norms weakening
  - An underlying bloodthirsty nature unleashed
  - Magnified by ancient hatreds between groups
  - Especially among uncivilized "others"
  - Compounded by younger, more urban, rootless populations
- Criticized as the "New Barbarism"
- Have we over-corrected?

# Today: What would behavioral theories of contentious politics look like?

- The kinds of biases that are important for bargaining and political behavior are arguably quite different from the biases in other kinds of decision-making
- I'll review hypotheses suggested by historical, ethnographic, and psychological accounts
  - Typically economists and political scientists have not formally tested these, yet
- Quantitative evidence or better theory on any one of these would be a major contribution
  - The same applies to voting, strikes, boycotts, riots, protests, etc.
- We'll also see that behavioral explanations lie beneath the surface of many theoretical and empirical papers (though seldom explicit or tested)

# A behavioral framework for thinking about political behavior

- Economic decision-making under uncertainty is founded on rational expected utility:
  - max $p(s)U(x|s)$

- Matt Rabin (2004) categorized all of behavioral economics into:
  1. Non-standard preferences
     - what does $U(x|s)$ really look like? What is in our utility function?
     - Nothing irrational  permits us to use standard tools and models
  2. How do people really form beliefs $p(s)$?
     - What heuristics, biases, and systematic mistakes do we make
     - Implies we may have the wrong $p(s)$, or fail to update, but we maximize
  3. Lack of stable utility maximization
     - Do people really max $p(s)U(x|s)$?
     - When do we fail to carefully calculate?

- In general, behavioral economics has done relatively little on strategic interactions

# Contents

# Passarelli & Tabellini (2017): An example of a model introducing fairness and emotions into decision making

- Some people have "expressive preferences" based in fairness
  - Participation has psychological rewards commensurate with the feeling of aggrievement, and these rewards are traded off against other considerations
  - These expressive preferences arise from a social norms — the government violating an expectation of fair behavior, such as failure to deliver a "policy entitlement", a reference point

- Expressive preferences are augmented by others' expression
  - There is a preference (not a strategic) complementarity: if expected participation is large, then more individuals are attracted to the protest for the same level of aggrievement

- But individuals behave rationally, weighing the pros and cons of participation, taking these non-standard preferences into account

# More formally

Individual $j$ in group $i$ chooses to riot if benefits are larger than costs:

$$p_i \lambda_i a_i - \mu - \epsilon_{ij} \geq 0$$

- $p_i$ is the proportion of your group participating
- $\lambda_i$ is the size of your group
- $a_i$ is the aggrievement caused by the policy to members of group $i$
- $\mu$ is the certain cost and risk of violent repression
- $\epsilon_{ij}$ is the idiosyncratic component of the cost or benefit of participation, uniformly distributed with mean 0 and density $1/2\sigma_{ij}$

Equilibrium participation rate is an increasing function of group aggrievement and a decreasing function of costs and risk:

$$p*_i = \frac{\sigma_i - \mu}{2\sigma_i - \lambda_i a_i}$$

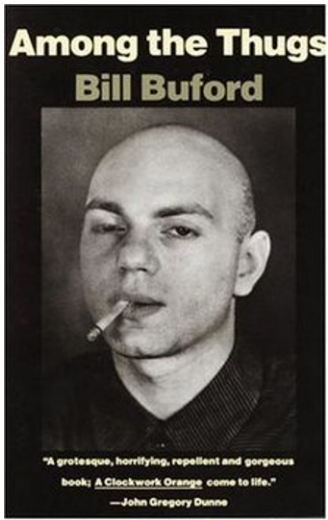
# Other thoughts

- Layers in a number of other elements, e.g. Reference points are endogenously determined, and are set by some sense of constraints facing the government
- Implications:
  - Means that rational, far-sighted governments may wish to restrain their future selves
  - Political power or influence here comes from a group's ease or technology of mobilization
  - Capacity for unrest causes an "excessive" amount of redistribution

- Feels a bit overfit to European protests
- Layers in many different "nonstandard" assumptions that interact
- An important step, but one might like to see a collection of models that consider a menu of these and similar "nonstandard" elements and illustrates how equilibrium changes with different combinations

# Contents

# Do (some) people get joy or pleasure from violence?




"There, on the streets of Fulham... I felt myself to be hovering above myself, capable of perceiving everything in slow motion and overwhelming detail.

I realized later that I was on a druggy high, in a state of adrenaline euphoria. And for the first time I am able to understand the words they use to describe it.

That crowd violence was their drug. What was it like for me? An experience of absolute completeness."

# Do (some) people get joy or pleasure from violence?

- As a manifestation of group participation and solidarity
  - Participant observers in British soccer hooliganism, the Vietnam War, and mobs demanding sacrifice all describe an overwhelming (though often momentary) joy in group violence (Broyles Jr 1984, Girard 1977, Buford 2001)
  - Others argue violence is a means of group bonding (Cohen 2017)

- As a manifestation of social preferences
  - "Schadenfreude": An extreme version of parochial altruism, where we not only have preferences for the well being of in-group, we take pleasure in seeing out-groups suffering (Chen and Li, 2009; Cikara et al., 2011; Kalin and Sambanis, 2018, Glowacki et al. 2017)

- Akin to accounts of "expressive preferences" for voting, or pleasure in the act of collective protest (Pearlman 2003)

- Little hard evidence of whether and how much this affects economic and political behavior

# Arguably more important: Disutility from competing beliefs

e.g. Identity-based competition and conflicts

- Are ethnic, religious, nationalist, or ideological conflicts motivated by a desire to eliminate or subjugate a rival group, or exterminate a heretic set of beliefs or ideologies?
  - Catholic elimination of heretical beliefs in early Europe
  - Hindu nationalism in modern day India
  - Nazi persecution of Jews
  - Polarization of left and right wing political groups

- This theory is distinct from the in-group preferences we have seen
  - Not a case of parochial altruism, but rather disutility from co-existence with another group, or from compromise with another group

# Disutility from compromise or competing beliefs

- Could be a preference possessed by group leaders, or a preference they inculcate and exploit among followers to achieve material ends
  - e.g. Models of polarization and conflict Esteban & Ray
- Someone committed to standard preferences could argue these are examples of commitment problems or agency problems, but preference-based explanations are also plausible
- To the best of my knowledge, very little quantitative work exists

# Note: Conflict is still a "puzzle" in these cases

- Means to achieving this end include:
  - Violence
  - Forced assimilation
  - Social engineering and other "nation-building"
- Thus question would be: When is war the less costly means of achieving these ends?
- Agency problems (leaders who do not internalize costs of war to the rest of the group) might interact with these preference-based explanations

# One variant on these identity-based accounts: "Value-rational violence"

Max Weber 1978, Varshney 2003

- In *Economy and Society*, Max Weber contrasts two forms of motivation:
  1. Instrumental rationality: Strict cost-benefit calculus with respect to goals
  2. Value rationality: Actions are determined by a conscious belief in the value for its own sake of some ethical, aesthetic, religious, or other form of behavior, independently of its prospects of success

- Disutility from competing beliefs/identities can be instrumental or value-rational

- Value-rationality implies a certain *increasing returns to preferences* — situations where people are willing to make extreme sacrifices or take great personal risks
  - This could help explain some individual behavior (e.g. Suicide attacks)
  - If these values are collective, could be a major source of bargaining failure

# Are "indivisibilities" value-rational preferences not commitment problems?

## e.g. Payoff asymmetries in Caselli et al 2015

| | | Player $B$ | |
|---|---|---|---|
| | | Action 0 | Action 1 |
| Player $A$ | Action 0 | 0,0 | $x + c_A, -x + c_B$ |
| | Action 1 | $x + c_A, -x + c_B$ | $x + c_A, -x + c_B$ |

- If A and B can negotiate, they should be able to find peaceful settlement within $(x, -x)$
- Indivisible resources lead to a commitment problem. But why should oil be indivisible?
- One option: An *unwillingness to compromise*?

# An unwillingness to compromise is a story underlying many historical accounts of conflict

- Refusal to share sacred religious sites (Hassner 2003, 2006, 2009)
- Another instance is one where the idea of compromise on some ideological value or principle is itself abhorrent
  - Liberty and self-determination in separatist anti-colonial movements
  - A common explanation of the American Revolution (Bailyn 1992, Thompson 2019)

# Contents

# Jha & Shayo 2020: Valuing peace

- People have different personal exposure to risks and returns from conflict and peace, and may not internalize the gains from peace
  - Hypothesis: Jewish Israelis not internalizing costs of conflict

- Design an experiment
  - Can exposure to financial markets help individuals internalize the economic costs of conflict?
  - Can this exposure change individual attitudes towards war and peace, or even their votes?

- What is this a case of? Is this a behavioral paper?

# Jha & Shayo 2020: Valuing peace

- People have different personal exposure to risks and returns from conflict and peace, and may not internalize the gains from peace
  - Hypothesis: Jewish Israelis not internalizing costs of conflict
- Design an experiment
  - Can exposure to financial markets help individuals internalize the economic costs of conflict?
  - Can this exposure change individual attitudes towards war and peace, or even their votes?
- What is this a case of? Is this a behavioral paper?
  - A Coasian view of the world
  - Is the failure one of learning, agency problems, or limited attention?

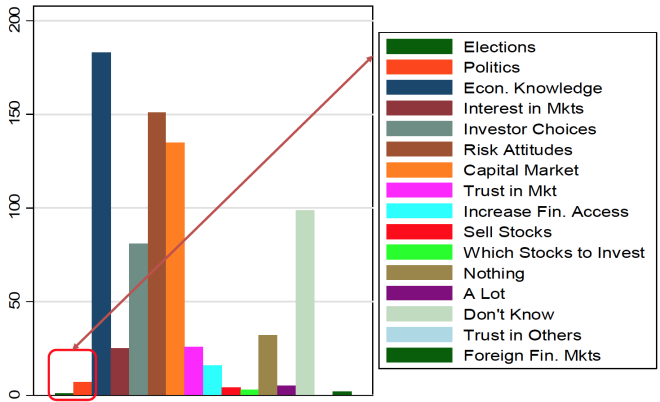
# The experiment

- Internet marketing panel of 60,000 Jewish Israelis invited to a study of investor behavior
  - 1345 likely voters "win" lottery to get a stock portfolio, oversampling centrists
  - Randomly assign to low stakes financial asset conditions worth $50–100:
    - Israeli stocks
    - Palestinian stocks
    - Voucher (tradable for stocks)
    - Control
  - Encourage them to trade on specific online platform over 4-7 weeks
- Sample useful because **they will not know that later social surveys are linked to the experiment**
- Outcomes: attitudes and votes
  - Main measure: vote in the March 2015 Israeli general elections
  - Other measures: selfreported attitudes towards peace deal

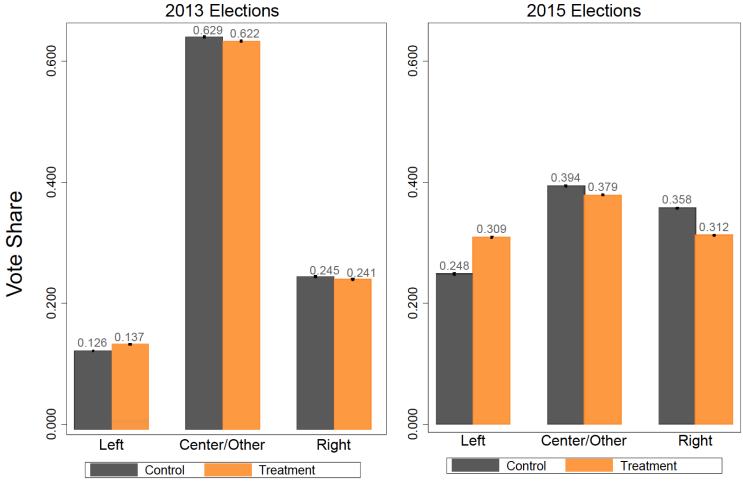
# Good lesson in reducing risk of experimenter demand

## Investor framing + "independent" post-experiment measurement

Post experiment survey question: "What do you think the researchers will learn from study?"



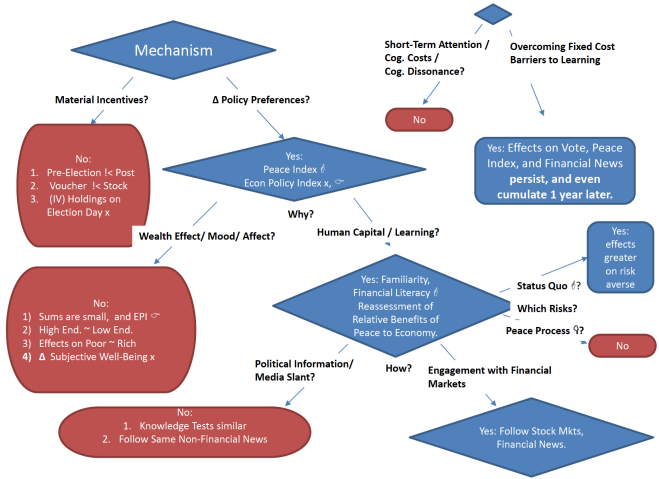
# Shift leftward: Treatment increases likelihood of voting for left parties (propeace initiatives) by 4-6 pp




N=1311. The center bars include 71 and 20 individuals who voted for for 'other' parties in 2013 and 2015, respectively, as well as 1 and 27 individuals who did not vote in 2013 and 2015, respectively.

# Pre-publication versions framed as overcoming fixed cost of learning about financial markets, not a behavioral bias

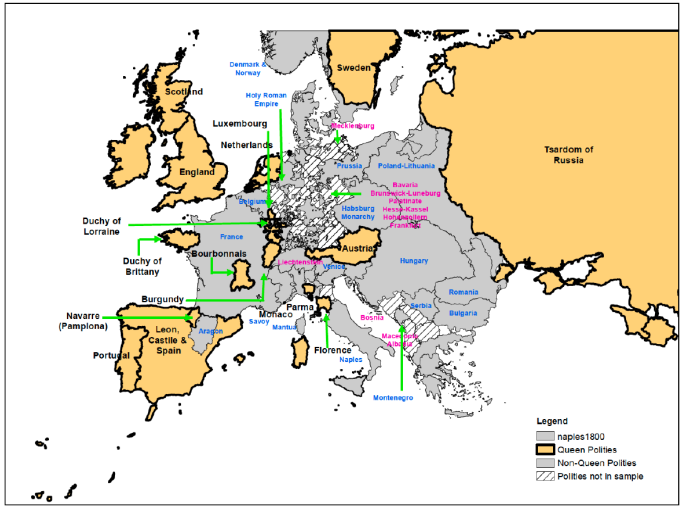
## How would you test these if you could design the experiment over again?



# Contents

# 18 polities in Europe had at least one queen, 1480-1913

## Dube & Harish 2020 JPE



# Are queenly regimes more or less likely to go to war than kingly ones?

- Of 194 monarchical reigns, 29 queens and 34 reigns
  - 24 reigns – Solo queens
  - 10 reigns – Co-ruling queens
- Empirical strategy: # of female contenders as instruments for whether queen in power
  - Queen less likely if previous monarchs had a first-born male
  - Queen more likely if previous monarchs had a sister
- 111 international wars and 28 civil wars

# Queenly regimes significantly more likely to find themselves at war

- Queens 27% more likely to be at external war than kings  a doubling over mean
  - IV coefficient  3 times larger than OLS, which *could* arise from more women coming to power in times of peace, or could arise from violations of IV assumptions
  - No effect on civil wars
- Generally speaking, queens did not initiate these wars
  - Especially if they were unmarried
- A (partly) behavioral interpretation?

# Queenly regimes significantly more likely to find themselves at war

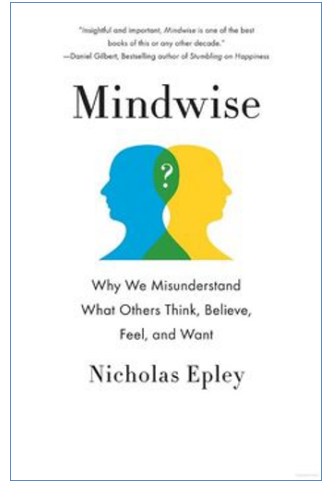
- Queens 27% more likely to be at external war than kings  a doubling over mean
  - IV coefficient  3 times larger than OLS, which *could* arise from more women coming to power in times of peace, or could arise from violations of IV assumptions
  - No effect on civil wars

- Generally speaking, queens did not initiate these wars
  - Especially if they were unmarried

- A (partly) behavioral interpretation?
  - Perceived weakness leads others to attack them
  - Implicitly a story of incomplete information (on type) plus systematic underestimation of women leaders
  - To the extent queens attack, it may be to signal strength (in face of prejudice)
  - Paper highlights incomplete information story, but not the error in beliefs

# Contents

# Overconfidence: The (not so) Newlywed game



# Are humans predictably overconfident?

- There are some trivial examples
  - Most drivers think they are above average (Svenson 1981)
  - Marathoners underestimate their time to completion (Krawczyk & Wilamowski 1984)
  - Almost all US high school students rated themselves as at least average at getting along with others" and a quarter put themselves in the top 1% (Camerer 2003)
- Economic forecasters are often far too confident in their precision (Alpert & Raiffa 1982)
- Most entrepreneurs think that their startup is more likely to succeed than their peers startups (Cooper et al 1988)
- Overconfident CEOs believe their company is undervalued and are less willing to raise capital by issuing new shares and more likely to attempt mergers (Malmendier & Tate 2005, 2008)
- I'm not aware of any evidence related to political behavior, bargaining, let alone violence

# Contents

# A (rare) limited attention model of conflict

## Acemoglu & Wolitsky 2014

- Tries to understand how incomplete information leads to prolonged conflicts
  - Decades-long civil wars, ethnic & religious conflicts, etc

- Model
  - Repeated coordination game where peace is the efficient outcome
  - In full-information repeated game settings, cooperation is enforced through tit-for-tat and grim trigger strategies
  - Authors propose that prolonged conflicts could arise from:
  - Incomplete information about whether defection was a mistake or act of a bad type
  - Limited memory – do not remember why conflict stated

- Show that if probability of bad types is small, errors can lead to long conflicts that resolve only after each side forgets who started the conflict
  - They conclude (probabilistically) that war began because of error, & begin to cooperate again

# Contents

# An alternative I find more compelling: Projection bias + Fairness

Recall the earlier "fairness" and incomplete information hypothesis

1. There is a distribution of "fair" and "selfish" types in society
2. The distribution of these types in the other group is uncertain, and the group may have incentives to misrepresent
3. Could help explain why incomplete information could lead to long conflicts
   - ▷ Wars break out to resolve information asymmetry
   - ▷ Fighting persists because initial attacks and violence lead to intrinsic preferences for violence, narrowing the bargaining range
4. BUT: Each side expects vengeant reactions, so they should strive to find bargains rather than risk war

What if we are bad at predicting what opponents perceive as unjust?

# What do we know about projection bias? Not much

- Survey after survey finds that people tend to exaggerate the extent to which others think, feel, and act as they do
  - Conservatives tend to think other people are more conservative than they are
  - Voters think that non voters were more likely to vote like themselves
- People even mispredict their future selves
  - Underappreciate taste changes (Loewenstein, ODonoghue, and Rabin 2003)
- I'm not aware of evidence from strategic interactions, violent or not

# What do we know about projection bias?

- Might we systematically mispredict:
  - What is perceived as an unjust act?
  - How a competitor will respond to aggression?
  - Costs of conflict to our future selves and others?

- Mispredicting what will be seen as a just respnse is the story of many identity based conflicts
  - e.g. Northern Ireland (English, 2003)

- One way to test: Some evidence that perspective-taking and empathy exercises reduce projection bias and errors
  - e.g. Bruneau & Saxe show that Israelis adjust attitudes in response to (1) being asked to write a story from the perspective of a Palestinian, or (2) a Palestinian giving them their perspective

# Contents

# Decision-making under arousal

## e.g. See Pearlman 2013 reading on Arab Uprisings

- Generally understudied: the role of affect or emotion on decision-making
- Some theory and evidence suggest that high levels of arousal can reduce the quality of decisions and provoke more rash and punitive responses
  - Specific biases may be tied to particular emotional states
  - Fearful emotional states increases risk aversion
  - Anger increases confidence, feelings of power, and reduces risk aversion (Lerner & Tiedens, 2006)
- Open question: do these emotional states also affect high-stakes and deliberated decisions by groups and leaders?

# Blattman et al 2017: Why does CBT help to reduce aggression and violence and criminality?

- One answer is rational: Changing identity changes relevant social prescriptions
- One is not: emotional regulation "slows down" thinking in emotionally charged situations and reduces the automatic and harmful use of violence

TABLE 5—PROGRAM IMPACTS ON TIME PREFERENCES, SELF CONTROL SKILLS, AND IDENTITY/VALUES

| | | ITT regression: ($N = 947$) | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Therapy only | | | | Cash only | | | | Both | | | |
| | | | | *p*-value | | | | *p*-value | | | | *p*-value | |
| Outcome (*z*-score) | Control mean (1) | ITT (2) | SE (3) | Unadj. (4) | Adj. (5) | ITT (6) | SE (7) | Unadj. (8) | Adj. (9) | ITT (10) | SE (11) | Unadj. (12) | Adj. (13) |
| *Panel A. 2–5 weeks* | | | | | | | | | | | | | |
| Forward-looking time preferences | −0.202 | 0.179 | [0.098] | 0.068 | 0.419 | 0.071 | [0.099] | 0.476 | 0.833 | 0.318 | [0.099] | **0.001** | **0.019** |
| Patience | −0.093 | 0.187 | [0.073] | **0.010** | **0.046** | 0.116 | [0.073] | 0.114 | 0.280 | 0.267 | [0.074] | **<0.001** | **0.002** |
| Time inconsistency | 0.008 | −0.063 | [0.074] | 0.393 | 0.596 | −0.009 | [0.076] | 0.903 | 0.902 | −0.138 | [0.075] | 0.067 | 0.209 |
| Self-control skills | −0.037 | 0.085 | [0.098] | 0.383 | 0.833 | −0.147 | [0.104] | 0.159 | 0.595 | 0.037 | [0.096] | 0.696 | 0.833 |
| Impulsiveness | −0.010 | −0.011 | [0.101] | 0.915 | 0.995 | 0.180 | [0.108] | 0.094 | 0.619 | 0.104 | [0.095] | 0.275 | 0.919 |
| Conscientiousness | −0.077 | 0.109 | [0.105] | 0.301 | 0.919 | 0.046 | [0.106] | 0.664 | 0.995 | 0.163 | [0.105] | 0.120 | 0.684 |
| Perseverance/GRIT | −0.035 | 0.027 | [0.099] | 0.785 | 0.995 | −0.130 | [0.105] | 0.217 | 0.873 | 0.042 | [0.104] | 0.686 | 0.995 |
| Reward responsiveness | −0.010 | −0.071 | [0.106] | 0.503 | 0.980 | 0.107 | [0.107] | 0.321 | 0.919 | 0.013 | [0.105] | 0.901 | 0.995 |
| Identity and values | −0.169 | 0.192 | [0.095] | **0.043** | 0.319 | 0.199 | [0.092] | **0.031** | 0.259 | 0.268 | [0.097] | **0.006** | 0.067 |
| Attitudes toward use of violence | 0.100 | −0.206 | [0.094] | **0.028** | 0.123 | −0.187 | [0.096] | **0.051** | 0.176 | −0.180 | [0.097] | 0.065 | 0.176 |
| Index of appearance | −0.118 | 0.085 | [0.081] | 0.295 | 0.298 | 0.131 | [0.081] | 0.105 | 0.189 | 0.203 | [0.080] | **0.011** | 0.062 |
| *Panel B. 12–13 months* | | | | | | | | | | | | | |
| Forward-looking time preferences | −0.149 | 0.149 | [0.102] | 0.144 | 0.826 | 0.105 | [0.102] | 0.303 | 0.952 | 0.209 | [0.105] | **0.047** | 0.464 |
| Patience | −0.240 | 0.170 | [0.103] | 0.097 | 0.350 | 0.145 | [0.096] | 0.132 | 0.386 | 0.258 | [0.099] | **0.009** | **0.048** |
| Time inconsistency | 0.129 | −0.072 | [0.083] | 0.386 | 0.712 | 0.018 | [0.087] | 0.833 | 0.836 | −0.059 | [0.084] | 0.480 | 0.712 |
| Self-control skills | −0.070 | 0.159 | [0.090] | 0.080 | 0.631 | −0.025 | [0.095] | 0.794 | 0.992 | 0.244 | [0.095] | **0.011** | 0.154 |
| Impulsiveness | 0.082 | −0.178 | [0.096] | 0.064 | 0.431 | 0.006 | [0.098] | 0.951 | 0.961 | −0.212 | [0.099] | **0.032** | 0.265 |
| Conscientiousness | 0.018 | −0.065 | [0.097] | 0.506 | 0.961 | −0.028 | [0.100] | 0.779 | 0.961 | 0.044 | [0.097] | 0.648 | 0.961 |
| Perseverance/GRIT | −0.037 | 0.116 | [0.099] | 0.241 | 0.851 | 0.057 | [0.099] | 0.565 | 0.961 | 0.105 | [0.103] | 0.311 | 0.903 |
| Reward responsiveness | 0.072 | −0.165 | [0.102] | 0.106 | 0.580 | 0.084 | [0.100] | 0.397 | 0.938 | −0.242 | [0.102] | **0.018** | 0.177 |
| Identity and values | −0.021 | 0.013 | [0.089] | 0.882 | 0.992 | −0.101 | [0.089] | 0.255 | 0.962 | 0.034 | [0.090] | 0.704 | 0.904 |
| Anticriminal/ antiviolent values | 0.070 | −0.076 | [0.088] | 0.386 | 0.880 | 0.026 | [0.088] | 0.768 | 0.948 | −0.177 | [0.086] | **0.040** | 0.279 |

# Contents

# Research frontiers

- Several areas remain unexplored, not only in political economy
  - The effect of emotional states on decision-making
  - Projection bias in any strategic interaction
  - How fairness concerns affect strategic interactions
- A promising route is designing interventions to solve specific behavioral biases
  - This is essentially route followed by Jha & Shayo and Blattman et al