Elephants and Game Theory

University of Mannheim Department of Political Science

Research Design

Filippo Panfoli 2202955

filippo.panfoli@students.uni-mannheim.de | PanfoliF |

0009-0000-7861-8036 | https://github.com/PanfoliF

Replication materials / LaTeX sources: PanfoliF/Research-Proposal-RD

Version: v0.0 | **Last updated:** January 6, 2026

Keywords: game theory

JEL: C72; D82

Prepared for: Research Design, taught by Prof. Dr. Sean Carey

January 6, 2026

Abstract

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

1 Introduction

In this article, I aim to examine the demographic dynamics of Brazil, with a particular focus on gender-related policies. The Brazilian case offers a compelling context for analysis, as the evolution of gender and LGBTQ+ rights in the country has followed a trajectory distinct from that of Europe. Both Brazilian legislation and culture have been at the forefront of progressive reforms, partly due to historical and political circumstances. Notably, LGBTQ+ movements played a significant role in supporting the current Constitution, which has protected such rights since 1988. Knowledge and insights about the historical context and institutional settings are drawn from Encyclopedia Britannica and **southey__history__2012**.

The author aims to study the impact of the institutional settings on demographic and social variables. In particular, emphasis will be on the effects of marriage and education laws on variables such as fertility and emancipation of women. The author will use the introduction of different laws and use them to assess the impact on variables that describe the marital and educational status of women.

alves__context__2017 identified two main sources of change through a qualitative analysis. Following in his footsteps, the article will mainly focus on two laws: the "National Policy on Sexual and Reproductive Rights in 2005" and the "National Family Planning Policy in 2007". These norms will serve as instruments to predict a changes in emancipation of women. They are expected to positively affect the outcome variables.

A preliminary analysis of the literature on education in developing countries, highlighted a study from **turmena__reforma__2022**. The journal article constitutes the milestone of this article and serves as main reference for the literature on education in Brazil. This paper suggests that the "Law No. 5.692/1971 (Reforma do Ensino de 1ž e 2ž Graus)" had a major effects on education. I will exploit data related to this law. The analysis I intend to conduct will depend on the availability of data.

Potential data sources include IPUMS and the Instituto Brasileiro de Geografia e Estatística (IBGE). Additionally, aggregated data may be retrieved from the Brazilian Education Panel Databases, which covers the period from 1996 to 2015. Eventually, the article from **rubiane__daniele__cardoso__de__almeida__demographic__2023** offers panel data on some demographic aspects.

To establish causality between the dependent and independent variables, the analysis will employ causal inference techniques. As a possible solution to the identification problem the study will provide results from a Difference-in-Difference. This technique is able to isolate the effect produced by the introduction of the laws in the variables that approximate fertility or women emancipation. Nonetheless, without appropriate control variables no identification strategy is reliable. Fol-

lowing best practices of the political science field, only a deep analysis of the literature will provide for suitable control variables.

2 Causal Inference

The discussion so far motivates a comparison to frameworks that model how unobserved factors influence observed outcomes. Causal inference (Pearl, 2009) suggests that in order to answer the question of interest, we may also think of an unobserved confounder that that is mistakenly excluded from the framework. I argue that an omitted variable is present in this setting.

2.1 Potential Outcomes Framework

This section formalises the intuition using a DAG. To make the subsequent discussion accessible, it first outlines how the potential-outcomes framework applies to this set-up.

A potential outcome is a counterfactual. Thus, it is a scenario that we do not observe; we think of it as the closest plausible scenario that might have happened had the actual one not occurred.

In the case of the two interactions presented here, each of them may be considered the counterfactual scenario of the other. Since counterfactual outcomes do not exist, we use real and observed cases to mimic them. The underlieng belief is that no researcher is as informative as reality itself in producing examples of what might have happened. If we rely on observation, the role of the researcher is limited to selecting the most appropriate counterfactual from those offered by the natural world.

Hypothesis: absent any causal effect of emotions, the instance involving the elephant and the sheperd would have been identical to the situation offered by the other elephant and the tourist.

2.2 Application of the Potential Outcomes Framework

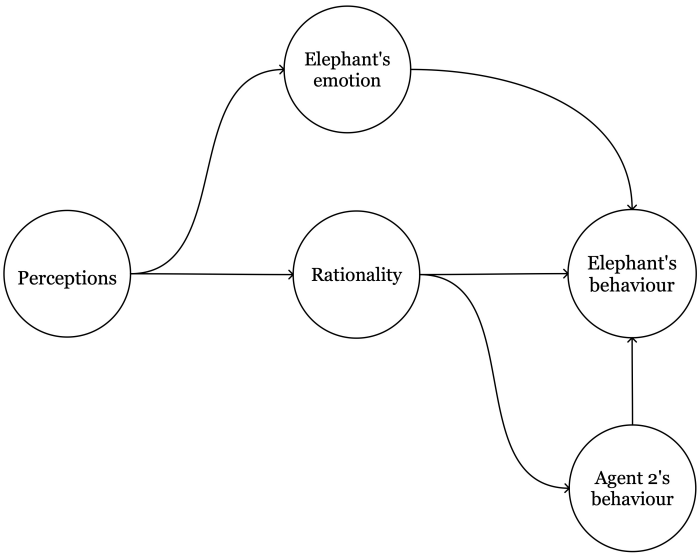

Figura 1: *Causal DAG.*

The graph shows five circles that represent our variables. We temporarily set aside the emotion node to describe the baseline case.

In the baseline set-up agents are represented by perceptions, rationality and behaviour. The former two shape the latter, which is the outcome variable of interest. A change in perceptions corresponds to a change in actions. Rationality is modelled as the decision rule or the algorithm agents have in their minds. It is the function that represents utility. Two players with the same utility function act in the same way when facing the same game. This means that a change in rationality corresponds to a chenge in behaviour.

Our outcome variable is the only one we can observe and measure. In the two occurrences described by Safina (2015), we *observe* the conduct of elephants and humans. However, we can only *infer* rationality and perceptions.

When we represent two agents using three variables, we make use of three circles for each player. Nonetheless, the figure presents only one circle for perceptions and only one circle for rationality.

This choice reflects a simplification. In order to ease the analysis, we assume that both agents perceive reality in the same way. Thus, both receive the same stimuli from the environment. We also assume that they process these stimuli through the same algorithm. They weigh the same benefits and costs in the same way. These assumptions symplify the baseline setting because restrict the set of potential outcomes: we consider only two possible outcomes: rationality in state

A or B and perceptions in state A or B.

How do we distinguish between the two agents in the Causal DAG? They differ in the realised behavioural outcome. While human conduct is constant, elephant is different in the two instances (**??**), this is why we have one circle for each player.

We have 4 circles, three of which are constant in the two occurrences. The fifth one represents elephant's emotions. This variable provides the key source of variation across cases. Thus, it can explain the change in elephant behaviour.

2.3 The Rationale for using Causal Inference
From the General Case to a Single Case–and Back Again

von Neumann and Morgenstern (1990, chap. 2, p. 14) write: "An almost exact theory of a gas, containing about 10^{25} freely moving particles, is incomparably easier than that of the solar system, made up of 9 major bodies; and still more than that of a multiple star of three or four objects of about the same size. This is, of course, due to the excellent possibility of applying the laws of statistics and probabilities in the first case. [...] The problem must be formulated, solved and understood for small numbers of participants before anything can be proved about the changes of its character in any limiting case of large numbers".

To pursue this suggestion, we move the abstract general case of rational-agent model to analyse specific interactions in which the representation fails to adequately map reality.

We then turn to causal inference in order to reconstruct a more general account: namely, to move from the idiosyncrasies of a single interaction to a *new* general framework. This requires abstracting away from irrelevant features and identifying a shared causal mechanism. The result is a revised pattern: a new map that can guide inference, organise the evidence, capture the data-generating process.

3 Literature Review and Limitations

3.1 A Review of Models of Conflicts

Further exploration of the conflict-games literature may either support or qualify the analysis presented so far. In any case, already-established models provide a useful point of comparison. Some of the papers that stand out include Schelling (1960, 1966). On the other hand, it may be of interest to focus on McKelvey and Palfrey (1995) to further develop the design.

3.1.1 A Review of Models of Conflicts with a Focus on Bargaining

One prominent strand of the literature examines on bargaining models of conflict. Key contributions to this field number Fearon (1995, 1996, 1998) and Schelling (1960, 1966). Reiter (2003) offers an in-depth review of the field. He provides

an historical perspective and discusses more recent works. Additionally, he summarises the most important ideas.

Reiter (2003) says: "James Fearon prominently contributed to this scholarship with his 1995 paper "Rationalist Explanations for War." He usefully highlighted the point that if states agreed on the outcome of a possible war, they could probably avoid war. Central to Fearon's paper is the importance of focusing on states' disagreement over their capabilities and/or resolve. [...] Fearon uses a bargaining model to argue that if two states in dispute know the outcome of a possible war, they should in general prefer to reach a deal that would reflect the hypothetical postwar political settlement, rather than fight, reach that same settlement, and also suffer the costs of war. This view assumes that fighting itself is costly–that the belligerent always suffers some negative utility, no matter how the issue at stake is settled. In fact, this assumption becomes necessary when one asserts that two states would always prefer to reach a bargain without fighting rather than fight and then reach the same bargain. Within his bargaining model, Fearon develops three conditions under which war is possible."

To broaden the overview, some other salient articles comprise Filson and Werner (2004), Kirshner (2000), A. Smith and Stam (2004), Wagner (2000), and Wittman (1979).

3.2 Limits of Theoretical Models

This discussion is necessarily selective and cannot provide a comprehensive assessment of the formal-modelling literature. In the previous pages I have presented a personal re-elaboration along with a brief summary of the relevant literature in this area. This is certainly not enough to assess the state of the models of this essay. However, the exposition I provide should not be taken as comprehensive, since there is no such thing as *a single correct or definitive modelling framework.*

In this respect, it is possible to further expand the discussion on limitations. The literature offers some very valuable contributions, among others, I discuss: Ashworth et al. (2021), Bräuninger and Swalve (2020), and Pareto (1916). These essays presents different ways to combine empirical and theoretical researc together with some insightful interpretations of the meaning and scope of models.

3.3 Connection between Game Theory and Biology

There are many important applications of game theory in biology. Particularly influential examples include: Axelrod and Hamilton (1981) and J. M. Smith (1973). Future work may shed light on more recent and equally notable applications. A vast literature covers the topic of evolutionary games as well: Hofbauer and Sigmund (2003), Kandori et al. (1993), Nowak (2006), and Taylor and Jonker (1978). In particular, Hofbauer and Sigmund (2003, chap 12) offers a clear treatise of powerful analytical tools together with applications of these game-theoretic tools in biology.

Unfortunately, a more detailed discussion of these works must be postponed.

Conclusion

The puzzle motivating this paper was deliberately simple: two elephant-human encounters occur in broadly similar environments, yet they generate sharply different behavioural trajectories. In one occasion, the elephant's threat is forceful but calibrated; in the other, violence is followed by a striking form of care.

Standard signalling logic is largely successful in explaining why an elephant might issue a credible threat and why a human might respond cautiously. What it does not readily explain is the combination of escalation and subsequent protection–an outcome that looks puzzling from the standpoint of canonical rational-agent models. The central claim advanced here is that the apparent inconsistency largely reflects omitted elements of the model: the relevant "shock" can be internal, and the relevant constraint can be affective.

The causal-inference lens is useful. Treating the two episodes as potential outcomes of a shared underlying interaction makes the research design explicit: the cases serve as mutual counterfactuals, and the goal is to isolate which latent factor, omitted in standard game-theoretic frameworks, generates the divergence. Within this framing, surprise functions as an informational discontinuity that plausibly shifts the elephant's perception of threat and thus the equilibrium-relevant mapping from signals to actions. Crucially, however, surprise alone cannot account for the post-attack behaviour. That second step motivates the introduction of empathy as a regulator: a force that can attenuate aggression, reshape the "intensity" of signals, and generate assistance once the situation is reinterpreted.

This conceptual move has a representational payoff. Instead of treating emotions as noise that perturbs otherwise coherent strategic choice, the paper sketches a way to represent emotions as structured, competing influences over cognition. The adaptation of the vote-buying framework provides a particularly transparent metaphor: thoughts play the role of passive "voters", while emotions act as active players allocating scarce intensity to secure agenda-setting power–here, the determination of behaviour.

In equilibrium, the elephant's observable trajectory corresponds to a stable emotional configuration: neither emotion can improve its outcome given the other's allocation. Read in this way, "attack, then care" is not a contradiction so much as the observable consequence of a reallocation of cognitive control after a misinterpretation is recognised.

The broader implication is that many phenomena labeled as "shocks" in political science and international relations may be mischaracterised when treated as purely exogenous disturbances. A framework that makes these mechanisms explicit can therefore complement familiar strategic models, not by abandoning rationality, but by widening the set of state variables that rational choice is al-

lowed to condition on.

At the same time, the limits of the present contribution should be stated plainly. A fuller treatment would require specifying payoffs, information structures, and equilibrium selection more rigorously. Future work could therefore proceed by tightening the formal model, including the relationship between surprise and empathy as distinct but interacting state variables.

The main message is modest: when game theory explains puzzling conduct by invoking "surprise", some times it may be pointing towards a missing internal mechanism. Making that mechanism explicit–rather than treating it as residual noise–can turn an ad hoc rationalisation into a tractable extension of formal modelling, one that is better aligned with the texture of real strategic behaviour.

Acknowledgements

Artificial intelligence-based tools were employed solely to improve linguistic clarity and grammar. No AI system contributed to the development of the research questions, theoretical framework or conclusions presented in this paper.

Bibliography

Books

Ashworth, S., Berry, C. R., & Bueno de Mesquita, E. (2021). *Theory and credibility: Integrating theoretical and empirical social science*. Princeton University press.

Hofbauer, J., & Sigmund, K. (2003). *Evolutionary games and population dynamics* (Transferred to digital printing). Cambridge University Press.

Pearl, J. (2009). *Causality: Models, Reasoning, and Inference* (2nd ed.) [Chapters 1 and 2 define causal diagrams (DAGs) and the structural model framework.]. Cambridge University Press.

Safina, C. (2015). *Beyond Words: What Animals Think and Feel* [Explores animal cognition and emotions, linking behavioral evidence with questions of consciousness.]. Henry Holt and Company.

Schelling, T. C. (1960). *The Strategy of Conflict* [Classic work on strategic behavior and conflict theory.]. Harvard University Press.

Schelling, T. C. (1966). *Arms and influence* [Classic work on deterrence, coercion, and the strategic use of uncertainty in conflict.]. Yale University Press.

von Neumann, J., & Morgenstern, O. (1990). *Theory of games and economic behavior* (6., paperback printing). Princeton University Press.

Articles

Axelrod, R., & Hamilton, W. D. (1981). The Evolution of Cooperation.

Fearon, J. D. (1995). Rationalist explanations for war. *International Organization*, *49*(3), 379–414. https://doi.org/10.1017/S0020818300033324

Fearon, J. D. (1998). Bargaining, enforcement, and international cooperation. *International Organization*, *52*(2), 269–305. https://doi.org/10.1162/002081898753162820

Filson, D., & Werner, S. (2004). Bargaining and fighting: The impact of regime type on war onset, duration, and outcomes. *American Journal of Political Science*, *48*(2), 296–313. https://doi.org/10.1111/j.0092-5853.2004.00071.x

Kandori, M., Mailath, G. J., & Rob, R. (1993). Learning, Mutation, and Long Run Equilibria in Games. *Econometrica*, *61*(1), 29. https://doi.org/10.2307/2951777

Kirshner, J. (2000). Rationalist explanations for war? *Security Studies*, *10*(1), 143–150. https://doi.org/10.1080/09636410008429423

McKelvey, R. D., & Palfrey, T. R. (1995). Quantal Response Equilibria for Normal Form Games. *Games and Economic Behavior*, *10*(1), 6–38. https://doi.org/10.1006/game.1995.1023

Nowak, M. A. (2006). Five rules for the evolution of cooperation. *Science*, *314*(5805), 1560–1563. https://doi.org/10.1126/science.1133755

Reiter, D. (2003). Exploring the Bargaining Model of War. *Perspective on Politics*, *1*(1), 27–43. https://doi.org/10.1017/S1537592703000033

Smith, A., & Stam, A. C. (2004). Bargaining and the nature of war. *Journal of Conflict Resolution*, *48*(6), 783–813. https://doi.org/10.1177/0022002704268026

Smith, J. M. (1973). The logic of Animal Conflict. *246*.

Taylor, P. D., & Jonker, L. B. (1978). Evolutionary stable strategies and game dynamics. *Mathematical Biosciences*, *40*(1–2), 145–156. https://doi.org/10.1016/0025-5564(78)90077-9

Wagner, R. H. (2000). Bargaining and War. *American Journal of Political Science*, *44*(3), 469–484. https://doi.org/10.2307/2669259

Wittman, D. (1979). How a war ends: A rational model approach. *Journal of Conflict Resolution*, *23*(4), 743–763. https://doi.org/10.1177/002200277902300408

Files

File.txt