UNIVERSITY OF CALIFORNIA,
IRVINE

The Economics of Cooperation and Conflict

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Economics

by

Patrick N. R. Julius

Dissertation Committee:
Professor John Duffy, Chair
Professor Michael McBride
Associate Professor Jean-Paul Carvalho

2021

DEDICATION

To my late father, Douglas Patrick Julius, who I wish had lived to see this day.
To my mother, Jo-Anne Marie Julius, and my husband-to-be, David Nicholas Ross,
without whose support I could never have made it this far.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

The third chapter was written in collaboration with Muhammad Bulatay, David Hales, and Weiwei Tasch.

Muhammed Bulutay is a graduate student at the Technische Unversitat Berlin.

David Hales is a graduate student at the University of California, Santa Barbara, and serves as Chairman of Global Innovations Bank, and Chairman and CEO of Global Fidelity Corp.

Weiwei Tasch is a post-doctoral researcher at the University of Pennsylvania.

This work was supported in part by grants from the University of California and the Joachim Herz Foundation.

Experimental sessions were run at the University of California, Irvine Experimental Social Sciences Laboratory and the University of California, Santa Barbara Experimental and Behavioral Economics Laboratory.

In addition to my dissertation committee, I would like to thank Professor Stergios Skaperdas and my friend Andrew Benson for comments which have improved this manuscript.

Portions of this dissertation have been submitted for publication at peer-reviewed journals, and one paper is currently under review at the Journal of Economic Behavior and Organization.

VITA

Patrick N. R. Julius

EDUCATION

Doctor of Philosophy in Economics **2021**
Recipient of the Fellowship in Honor of Christian Werner

University of California, Irvine *Irvine, CA*

Master of Arts in Economics **2014**
Distinguished Graduate

California State University, Long Beach *Long Beach, CA*

Bachelor of Science in Brain, Behavior, and Cognitive Science **2010**
Graduate with Distinction

University of Michigan *Ann Arbor, MI*

RESEARCH EXPERIENCE

Graduate Student Researcher **2018–2021**
University of California, Irvine *Irvine, California*

TEACHING EXPERIENCE

Teaching Assistant **2016–2018**
University of California, Irvine *Irvine, California*

ABSTRACT OF THE DISSERTATION

The Economics of Cooperation and Conflict

By

Patrick N. R. Julius

Doctor of Philosophy in Economics

University of California, Irvine, 2021

Professor John Duffy, Chair

Cooperation and conflict are central to economics and to human life in general. One of the most basic choices we must make in almost any situation is when to collaborate and when to compete, when to help and when to harm. This fundamental question of cooperation versus conflict recurs in a variety of domains, ranging from individuals to entire nations. Yet despite the many differences between these domains, certain common threads unite them, and one of my primary research goals has been to explore these common threads.

In the first chapter of this dissertation, I consider the question of cooperation and conflict at the smallest scale, between individuals in a group.

One of the simplest experimental representations of cooperation is the public goods game, which has been a staple of experimental work for decades. I devised a twist on this classic experimental design. Past public goods experiments have treated taxation as pure waste or assumed flat taxes. I study a more realistic setting in which taxes are progressive, tax revenue provides a public good that may be more or less efficient than contributions, and contributions are tax deductible. Though there is an interior dominant strategy equilibrium, mean contributions are consistently above this level, even when this is Pareto-harming. The motivation to contribute is resilient in the face of tax incentives—even when such "cooperation" is harmful to everyone.

In the second chapter of this dissertation, I consider the question of cooperation and conflict at the largest scale, between nation-states. While national conflict is the primary subject, it is worth nothing that the very existence of functional nation-states requires cooperation on an enormous scale. Thus, war is both a failure of cooperation and also an exemplar of it.

The central phenomenon I sought to better understand is the present rarity of international conflict: Wars are less frequent today than through most of history. Various explanations for this have been proposed, ranging from the spread of democracy, to globalized trade, to the invention of nuclear weapons. To investigate this question, I develop an economically microfounded model of international conflict as an indefinitely iterated game, showing that under typical parameters there exist many equilibria, some resulting in war, others in peace. I propose that international norms provide a plausible equilibrium selection mechanism for this iterated game. This offers a synthesis between the "realist" and "liberal" schools of international relations: The realists are correct that international norms are cheap talk, not written into the rules of the game; but the liberals are correct that norms can have real effects, when they select one equilibrium over another.

In the third chapter of this dissertation, I and my collaborators consider the question of cooperation and conflict on a more moderate scale: tacit collusion between firms in an oligopoly. This case is interesting because it is a form of cooperation on one scale that entails harm on another scale: Consumers would most likely be better off if firms chose *not* to cooperate with one another.

In particular, we hypothesized that tacit collusion may provide an explanation for the phenomenon of Asymmetric Price Transmission (APT), the tendency of prices to respond more rapidly to positive than to negative cost shocks. Using a laboratory experiment that isolates the effects of tacit collusion, we observe APT pricing behavior in markets with three or more sellers, but not in duopolies. Furthermore, we find that sellers accurately forecast others' prices, but nevertheless consistently set their own prices above the profit-maximizing response, particularly in the periods immediately following negative shocks. Overall, our findings support theories in which tacit collusion plays a central role in APT.

A few common insights can be gleaned from these three different lines of research:

First, cooperation is a very common outcome, even when circumstances are contrived to make it especially difficult. Cooperation only occurs in small space of possible equilibria in the international conflict model, yet in the real world cooperation is the most common outcome. In the experimental studies, matters are even more extreme: as these are finitely iterated games with a unique stage-game Nash equilibrium that is uncooperative, cooperation is theoretically not sustainable as an equilibrium at all. Yet on average we saw more cooperation than conflict in both experiments, and most participants in most sessions chose cooperatively at least some of the time.

Second, human beings are not perfect cooperators; the tension between conflict and cooperation is a persistent one. Obviously wars do occur in reality, even though they are rarer today than they once were. In experimental studies, cooperative behavior is very common, but so is uncooperative behavior, and most choices lie strictly between the optimal cooperative and competitive choices.

Third, the conditions under which cooperation can be sustained are subtle, complex, and often difficult to explain. This was particularly apparent in the public goods experiment, where neither self-interest, warm-glow altruism, nor inequity aversion could adequately explain the observed pattern of behavior. It was also apparent in the international conflict model, where surprisingly subtle changes in economic parameters could make a peaceful equilibrium collapse into war.

There remains a rich unexplored frontier of research in better understanding why and how human beings engage in cooperation and conflict.

Chapter 1

Experimental public goods games with progressive taxation

Abstract

Public goods can be provided through a mix of voluntary contributions and taxation. Existing experimental work has treated taxation as pure waste or assumed flat taxes. I study a more realistic setting in which taxes are progressive, tax revenue provides a public good that may be more or less efficient than contributions, and contributions are tax deductible. Though there is an interior dominant strategy equilibrium, mean contributions are consistently above this level, even when this is Pareto-harming. Using a ”stranger” re-matching design, I find no evidence of decay in contributions over time or end game effects.

This study was registered as AEARCTR-0003037.

Keywords: Experimental economics, behavioral economics, public goods, progressive taxation

JEL codes: D9, H2, H41

1.1 Introduction

Charitable contributions may be considered a form of public goods game, where the benefits of a donation are spread over a large population while the costs are borne primarily by the donor. Over 90% of all highly-developed countries offer a tax deduction for charitable contributions (Whitehead and Stiles, 2015), and worldwide, 77% of countries offer some form of donation incentive to corporate donors, while 66% of countries offer some form of donation incentive for individuals (Quick et al., 2014).

Public goods games have long been a staple of experimental economics; they were pioneered in the 1970s, and are still the subject of ongoing research(Kagel and Roth, 2016). Yet there are still vital aspects of public goods games which have not been studied experimentally. In particular, there has been very little work combining voluntary contributions with taxation that also provides a public good.

There have been a few experiments using public goods game with taxation, such as Sheremeta and Uler (2016); but in nearly all of these experiments, taxation was redistributive or simply wasteful; tax revenue did not provide a public good. Yet in the real world, many public goods are provided by government spending. Indeed, in some cases the public goods provided by government spending may be more efficient than those provided by private contributions. Contributions to charities that are less efficient than government spending may actually be Pareto-*harming*, in that both donor and recipient would have been better off if the money had been spent on government services that provide greater benefits for a lower cost.

The efficiency of public good provision by different charities and different government services varies widely: For instance, the most cost-effective charities, such as providing vaccines in very poor countries, can save a disability-adjusted life year for as little as $100 (Sinha et al., 2007), while many charitable organizations have little or no impact or are even outright fraudulent (Krasteva and Yildirim, 2016). Considering government services, public health programs to mitigate the spread of pandemics provide an important public good which saves many thousands of lives (as is particularly salient at the moment), and pollution regulation is a global public good that has been estimated to save hundreds of thousands of lives every year and provide trillions of dollars in net economic benefits (Landrigan et al., 2018), while from a global perspective national defense is almost all rent-seeking (and even from a national perspective is often highly inefficient).

Moreover, the real-world tax systems of the US and most highly-developed countries are progressive (Wagstaff et al., 1999); this changes the incentives for tax avoidance because the marginal benefit of a given contribution varies with the contributor's income. Despite this fact, previous public goods games experiments with taxation such as Eckel and Grossman (2003), Karlan and List (2007), and Sheremeta and Uler (2016) have used flat rather than progressive taxes.

This experiment modifies the standard public goods game in order to fill these gaps in the literature. Instead of the standard linear Voluntary Contribution Mechanism, I apply a progressive tax with deductible contributions, which is a realistic framing. The use of progressive taxation offers another advantage, aside from its realistic framing: it reliably generates an interior solution for the dominant strategy which can be varied by changing the parameters.

In one treatment, most similar to the standard public goods game, contributions to the group are beneficial, and the Pareto optimum is at the maximum contribution. I find results consistent with the rest of the literature: Participants systematically contribute above the dominant strategy level, resulting in a Pareto improvement over the dominant strategy equilibrium.

However, in an alternative treatment, the tax revenue is used to provide a different public good which is more efficient than the contributions, and the Pareto optimum is at the *minimum* contribution. Participants do contribute less under this treatment, but they still contribute more than the dominant strategy level, which is now Pareto *harming*. To my knowledge, this is the first experiment in the literature which tested a nonlinear public goods game that had an interior dominant strategy and a Pareto optimum at the maximum contribution.

In this experiment, the underlying tax function is smooth, allowing for a well-defined first-order condition in the model; but in practice the decision space is discrete. The progressive tax also has the benefit of inducing an interior dominant strategy and thus avoiding the *corner critique* (Plott and Smith, 2008) which most public goods experiments face: When the dominant strategy is a boundary solution at zero, any amount of error or noise in player behavior will result in the mean contribution being above the Nash equilibrium prediction, biasing the results. Moreover, many statistical tests are ineffective in analyzing such one-sided deviations, forcing reliance on strong parametric assumptions that may be poor approximations in the small samples typical of laboratory experiments. In this study I avoid such a critique by inducing an interior strictly dominant strategy.

This experiment allows for a clear test of other-regarding preferences: I am able to make sharp predictions as to how money-maximizing players would make allocations, which are at neither the boundaries nor the precise center of the decision space. I have chosen parameters so that both those with large and small endowments, if they are self-interested money-maximizers, will choose interior solutions for their donation choice. (I also have one treatment in each session where the dominant strategy is a corner solution, for comparison.) This interior solution allows for better statistical testing as well as opportunities for individuals to behave spitefully—paying a cost to harm other players that they deem worthy of punishment. Some of the results could be interpreted as indicating spite, but based on the results of other treatments I do not consider this the most parsimonious explanation.

1.2 Related Literature

Past research on public goods games has shown that a substantial number of players do not choose the strictly dominant strategy of zero contribution even in one-shot games, though they also do not generally achieve the Pareto efficient donation level either (Falkinger et al., 2000). As participants become more experienced, overcontribution tends to fall over time, particularly in finitely iterated games (Lugovskyy et al., 2017), but less so when players may choose to include or exclude players based on their past choices (Charness and Yang, 2014) or when they are allowed to punish free-riders (Fehr and Gachter, 2000; Coricelli et al., 2004; Fu et al., 2017). Yamakawa, Okano and Saijo (2016) found evidence that reciprocal motives better explain behavior in public goods games than what they call "one-shot motives" such as inequity aversion.

When the equilibrium contribution is zero, random errors can explain overcontribution behavior, since error results in over-contribution. Bayer et al. (2013) previously used a limited information environment to test whether error and confusion could explain overcontribution in public goods games, and found that they could not. Along similar lines, Corazzini and Tyszler (2015) found that quantal response equilibrium best fits the data under the assumption that a significant number of participants prefer better outcomes for the group, rather than under the assumption that all participants were self-interested but committing a high level of error.

Other studies using public goods games with interior Nash equilbria have found that overcontribution is a real effect and not simply the result of error (Plott and Smith, 2008). However, results for nonlinear public goods games have sometimes found mixed results which may be due to confusion among the participants (Cason and Gangadharan, 2015). Isaac and Walker (1998) found that the Nash equilibrium is not a good predictor of behavior in several nonlinear public goods games, though their design did not have a dominant strategy, and thus required participants to correctly anticipate the choices of others in order to achieve equilibrium.

This experiment also contributes to the policy evaluation literature on deductible contributions versus other methods of incentivizing donations. It remains controversial in the literature whether charitable tax deductions have a large effect on charitable contributions (Zampelli and Yen, 2017; Joulfaian, 2017). Eckel and Grossman (2003) found that tax deductions appear to be less efficient at motivating charitable contributions than donation matching, and Karlan and List (2007) found that the amount of matching is relatively unimportant, as long as there is a salient amount of matching. Adena and Huck (2017) found that matching contributions can create crowding-out, which is greatly reduced if the matched contributions go to a different, but still valued, charity. But the best policy for increasing donations may be even cheaper: Schulz et al. (2018) recently found that simply offering a list of default charities (a very low-cost nudge) was one of the most effective interventions at increasing donation rates.

Sheremeta and Uler (2016) conducted a similar experiment using a public goods game with taxation, and found that tax revenue was complementary to contributions, with more efficient government resulting in higher contributions. My results are different; I find that a higher MPCR on tax revenue results in lower contributions and more tax revenue. In their environment, tax revenue was purely redistributive or wasteful, while in mine, tax revenue provides a competing public good that may be more efficient than contributions. This difference may account for the different results. In addition, Sheremeta and Uhler's contributions were to actual charities, rather than other participants in the game, and their tax rate was flat rather than progressive; these factors may also have contributed to the different results.

1.3 Model

The game is structured as follows. Each player receives an endowment w_i, $0 \leq w_i \leq M$, and chooses an amount x_i to keep, contributing $w_i - x_i$ to the voluntary public good. Then, they pay taxes on what they chose to keep, according to a progressive tax function $\tau(x_i)$, $\tau'(x_i) > 0$, $\tau''(x_i) \geq 0$. In practice, I use $\tau(x_i) = \frac{x_i^2}{M}$, so that the average tax rate is simply the proportion of the maximum amount M chosen.

Voluntary contributions are multiplied by γ_V and returned to the entire group; thus the MPCR for voluntary contributions is $\frac{\gamma_V}{N}$. Tax revenue is multiplied by γ_T and returned to the group; thus the MPCR for taxation is $\frac{\gamma_T}{N}$. Because $\tau'(x_i) = \frac{2x_i}{M}$, it is possible to have marginal tax rates exceed 1 for endowments above $\frac{M}{2}$; thus I can test behavior under conditions of both $\gamma_V < \gamma_T$ and $\gamma_V > \gamma_T$. I will refer to the case where $\gamma_V < \gamma_T$ as "tax-efficient" and the case where $\gamma_V > \gamma_T$ as "tax-inefficient".

The monetary payoff y_i is as follows:

$$y_i = x_i - \tau(x_i) + \frac{\gamma_V}{N} \sum_j [w_j - x_j] + \frac{\gamma_T}{N} \sum_j [\tau(x_j)]$$

$$y_i = x_i - \frac{x_i^2}{M} + \frac{\gamma_V}{N} \sum_j [w_j - x_j] + \frac{\gamma_T}{N} \sum_j \left[\frac{x_i^2}{M}\right]$$

If players are self-interested and risk-neutral, their payoff is maximized by this first-order condition:

$$1 - \tau'(x_i) = \frac{\gamma_V}{N} - \frac{\gamma_T}{N}\tau'(x_i)$$

$$1 - \frac{2x_i}{M} = \frac{\gamma_V}{N} - \frac{\gamma_T}{N}\frac{2x_i}{M}$$

$$x_i = \frac{1 - \frac{\gamma_V}{N}}{1 - \frac{\gamma_T}{N}}\left[\frac{M}{2}\right] \tag{1.1}$$

This is a unique strictly-dominant strategy for each player, which is an interior solution for each initial endowment. If players are risk-neutral and self-interested, they should always play this strictly-dominant strategy, resulting in a unique dominant strategy equilibrium.

Proposition 1.1. *At the dominant strategy equilibrium, all players will choose to keep the amount* $x_i = \frac{1-\frac{\gamma_V}{N}}{1-\frac{\gamma_T}{N}}\left[\frac{M}{2}\right]$*, contributing the remainder of their endowment* $w_i - x_i$*.*

Proposition 1.1 follows directly from equation 1.1, and entails that voluntary contributions are higher when $\gamma_V > \gamma_T$ than when $\gamma_V < \gamma_T$. This is the most important treatment effect under consideration, so I have designated it proposition 1.2.

Proposition 1.2. *With all other parameters fixed, voluntary contributions will be higher when* $\gamma_V > \gamma_T$ *than when* $\gamma_V < \gamma_T$*.*

Proposition 1.1 also entails the following comparative static predictions. The marginal cost of donating money $\frac{\partial y_i}{\partial x_i}$ is decreasing in $\frac{\gamma_V}{N}$, indicating that a higher MPCR on contributions incentivizes more donation, and increasing in $\frac{\gamma_T}{N}$, indicating that a higher MPCR on tax revenue sector incentivizes less donation. Furthermore, since the amount *kept* x_i is constant across endowments, it follows that the amount *contributed* should increase one-to-one with endowment. Of these, the most important effect is the tax-efficiency effect.

g

1.4 Experimental Design

There are $N = 5$ participants in each game, participating in $T = 40$ rounds per session, with the round that is actually paid chosen randomly at the end; that is, a "pay-one" design.

Each participant is given an endowment w_i. Endowments are either 10 points ($5.00), 15 points ($7.50) or 20 points ($10.00). For rounds 1-10 and 20-30, all participants had an endowment of 15 points; for other rounds three had an endowment of 10 and the other two had an endowment of 20. These endowments were randomly re-assigned each round, so that each participant always had the same chance of getting each possible endowment. The use of heterogeneous endowments aids in testing for inequity aversion, discussed later in section 1.7.

I have opted for a "strangers" design where the participant groups are randomly re-matched each round, and a fixed, known number of rounds to allow for backward induction. This allows for many rounds of data collection while still preserving the same incentives as a one-shot static game.

The MPCR of contributions is $\frac{\gamma_V}{N} = 0.3$, and the MPCR of tax revenue is $\frac{\gamma_T}{N}$ which varies based on treatment.

As shown in table 1.1, the experiment has a 2×3 design, with tax-efficient and tax-inefficient treatments for each of the three endowments.

Income	10	15	20
Tax-inefficient ($\gamma_T/N = 0$)			
Tax-efficient ($\gamma_T/N = 0.5$)			

Table 1.1: Table illustrating 2×3 treatment design.

Contributions are chosen in increments of \$0.50, which ensures that the dominant strategy is always strict even in the discrete decision space. These parameters were chosen to make the changes in payoffs steep enough to be salient (each \$0.50 increment changes the participant's payoff by around \$0.05 to \$0.10 near the dominant strategy, and more further away from the peak), while still remaining reasonable, affordable total payoffs (including the \$7 participation fee, typically between \$20 and \$30 per participant for a 90-minute session). All endowments are public knowledge.

Each participant has the option to donate an amount $w_i - x_i$ to a public good, keeping x_i. Total contributions to the public good are multiplied by a factor $\gamma > 1$ then redistributed evenly to the entire group. Whatever is not donated to the public good x_i is then be considered "taxable income" and taxed according to a progressive tax, such that the total tax paid $\tau(x_i)$ is equal to $\frac{x_i^2}{20}$. The amount collected in taxes is then added to a separate public good, multiplied by γ_T.

Paying taxes in this experiment mandatory, as is generally the case in reality, unlike other experiments in which "taxation" was actually another voluntary contribution such as Li et al. (2011). The choice of MPCR reflects a government which in some treatments is more efficient at providing public goods than the private nonprofit sector, and in other treatments is less efficient.

For each treatment, these payoffs are presented to participants using an interactive slider, which participants can move to get immediate real-time feedback on the effects of each contribution choice on their payoff and that of the other participants in their group. Screenshots showing the full instructions and the interactive slider are provided in the appendix.

Once all choices have been made in each round, each player is informed of their own payoff as well as the contributions and payoffs of all other players.

All treatments were conducted within the same session using a within-participant design. The order was varied randomly between sessions.

Data was collected using the undergraduate subject pool at the Experimental Social Science Lab at the University of California, Irvine using oTree (Chen, Schonger, and Wickens 2016). I conducted the pilot studies and the first two sessions in person at the lab. Then, as a result of the COVID-19 pandemic, I conducted the remaining sessions online.

All of the participants were undergraduate students, representing a variety of majors.

I also collected demographic data on participants such as gender, age, and field of study, as there is some evidence that these characteristics can affect contributions in public goods games (Eckel and Grossman 2002; Frey and Meier 2005). However, regressions showed no statistically distinguishable effect of gender, and other demographic variables had too few cases in each cell to allow for sufficient statistical power.

Data analysis was conducted using R (R Core Team 2017) and specifically the `plm` package (Croissant and Millo 2008), and output using the `stargazer` package (Hlavac 2018).

1.5 Hypotheses

The model predicts that participants will choose the dominant strategy equilibrium. This leads to five hypotheses, all of which follow directly from equation 1.1.

1. *Dominant strategy equilibrium*: Participants will contribute the amount necessary to achieve the dominant strategy equilibrium.

2. *Pareto inefficiency*: Participants will not contribute the amount necessary to achieve the Pareto optimum.

3. *Tax efficiency response:* Participants will contribute more to the public good under the "tax-inefficient" condition than the "tax-efficient" condition.

4. *Endowment Response*: Contributions will be higher when endowments are higher.

Hypothesis 1 is what follows directly from the model: If players are rational and selfish, they will choose the dominant strategy.

The remaining hypothesis also follow from the model, and would be implied by hypothesis 1, but are more general: The dominant strategy equilibrium is Pareto-inefficient and results in the predicted responses to the tax efficient, MPCR, and endowment; but even if participants do not choose the dominant strategy level, they could still exhibit the same comparative static effects.

Table 1.2 shows the predicted results for each treatment, based on the model.

Table 1.2: Expected contribution under each treatment

	Taxes inefficient	Taxes efficient	Tax effect
Income = 10	3	0	-3
Income = 15	8	1	-7
Income = 20	13	6	-7

1.6 Results

1.6.1 Summary

The first session had 15 participants, the second session had 25 participants, and all subsequent sessions had 10 participants. Under the "stranger" matching protocol, each round all groups were randomly re-assigned, with each individual having an equal and independent probability of being matched into any group. All sessions were 40 rounds in length. This yields a total of 5200 observed choices. These observations are not completely independent, but with the random re-matching, reciprocal motives should be minimized, and in following sections I use panel data methods to control for any unobserved variation between individual participants.

Table 1.3 shows the mean level of contributions across each treatment cell, compared with the dominant strategy equilibrium prediction.

Table 1.3: Summary of results, comparing observed outcomes with DSE predictions. Asterices indicate statistically significant deviations from DSE predictions. *p<0.1; **p<0.05; ***p<0.01.

	Tax-inefficient		Tax-efficient		Difference	
Endowment	DSE	Observed	DSE	Observed	DSE	Observed
10	3	5.0***	0	2.3***	-3	-2.7
15	8	8.8	1	3.6***	-7	-5.2***
20	13	11.9**	6	5.8	-7	-6.1*

Table 1.4 shows the results of OLS regressions aggregating all the rounds in each session by treatment. This is a conservative approach that treats all participants in a session as a single unit. Each observation is a session mean of choices across all participants at a given endowment and tax treatment, for a total of $11 \times 3 \times 2 = 66$ observations.

The list below summarizes the findings for each hypothesis:

	Contribution		Deviation from DSE	
Constant	4.994***	(0.390)	1.994***	(0.390)
Income 15	3.681***	(0.551)	−1.319**	(0.551)
Income 20	6.876***	(0.551)	−3.124***	(0.551)
Tax-efficient	−2.672***	(0.551)	0.328	(0.551)
Tax-efficient × 15	−2.434***	(0.780)	1.566**	(0.780)
Tax-efficient × 20	−3.373***	(0.780)	0.627	(0.780)
Observations	66		66	
Adjusted R^2	0.861		0.514	
F Statistic (df = 5; 60)	81.409***		14.760***	
Note:			*p<0.1; **p<0.05; ***p<0.01	

Table 1.4: OLS regression, aggregated by treatment. Baseline is tax-inefficient with endowment 10. Standard errors in parentheses.

- Finding 1.1: Contributions deviate systematically from the dominant strategy equilibrium in nearly all treatments. Hypothesis 1 is not supported.
- Finding 1.2: Participants did not contribute the amount necessary to achieve the Pareto optimum. Hypothesis 2 is supported.
- Finding 1.3: Participants contributed more in the "tax-inefficient" treatment than in the "tax-efficient" treatment. Hypothesis 3 is supported.
- Finding 1.4: Participants contributed more when their endowment was higher. Hypothesis 4 is supported.

Figures 1.1 through 1.10 summarize the contributions within each session. On each graph, a translucent blue dot indicates an individual choice (resulting in darker dots where more participants made that choice), while the blue line represents the moving average of all choices in each round, and the green line represents the dominant strategy equilibrium contribution for that treatment. Consistent with the statistical tests, these graphs show that in almost every treatment, contributions were systematically above the dominant strategy.

1.6.2 Hypothesis 1: Dominant strategy equilibrium

I used non-parametric two-sided Wilcoxon signed-rank tests (with the standard continuity correction) to assess whether contributions deviated systematically from the dominant strategy equilibrium, except for the "tax-efficient" condition with endowment of 10, which has a dominant strategy of zero contribution for comparison with standard public goods games (this corner solution renders the Wilcoxon test invalid). I separated the sessions by ordering, to test for order effects in the treatments. There are some order effects: Deviations from the dominant strategy equilibrium appear to be larger when the "tax-efficient" condition was presented after the "tax-inefficient" condition, and when the unequal endowments were presented after the equal endowments.

The results of these tests are reported in table 1.5.

Table 1.5: Deviation from dominant strategy contribution, Wilcoxon signed-rank tests (p-values)

All sessions	Taxes inefficient	Taxes efficient
Income = 10	0.00000	N/A
Income = 15	0.00000	0.00000
Income = 20	0.00000	0.00354

With the Bonferroni correction for multiple comparisons, the corrected target p-value is 0.01. Even with this correction, the deviations from DSE are statistically significant in all 5 treatment cells.

Contributions are statistically distinguishable from the dominant strategy in all treatments. This indicates that participants are systematically deviating from the dominant strategy. Thus, hypothesis 1 is not supported: Participants are not choosing the dominant strategy equilibrium.

Finding 1.1. Dominant strategy equilibrium*: Contributions deviate systematically from the dominant strategy equilibrium in most treatments. Hypothesis 1 is not supported.*

1.6.3 Hypothesis 2: Pareto inefficiency

Average contributions are never at the Pareto optimum in any treatment, supporting hypothesis 2. Table 1.6 compares the mean contribution in each treatment to both the dominant strategy contribution and the Pareto-efficient contribution.

Except in the "tax-efficient" treatment with an endowment of 10, in which the two coincide, the mean contribution is typically closer to the dominant strategy contribution than the Pareto-efficient contribution. The only exception is in the tax-efficient treatment with an endowment of 15, where the mean contribution is statistically equidistant between the two.

As shown in table 1.6, the comparative static effects of the tax treatment are also much closer to the magnitude they would be under the dominant strategy, and are systematically very far from the magnitude they would be under Pareto-efficient allocations.

Finding 1.2. Pareto inefficiency*: Participants did not contribute the amount necessary to achieve the Pareto optimum. Hypothesis 2 is supported.*

Table 1.6: Comparison of contributions by treatment to dominant strategy and Pareto-efficient contribution levels.

	Tax-inefficient			Tax-efficient			Tax effect		
Endowment	DSE	Pareto	Observed	DSE	Pareto	Observed	DSE	Pareto	Observed
10	3	10	5.0	0	0	2.3	-3	-10	-2.7
15	8	15	8.8	1	0	3.6	-7	-15	-5.2
20	13	20	11.9	6	0	5.8	-7	-20	-6.1

1.6.4 Hypothesis 3: Tax efficiency response

The main effects of each treatment can also be assessed using Wilcoxon signed-rank tests. Results are reported in table 1.7 for the tax-efficiency treatment. I have reported results both for raw contribution choices and for deviations from the contribution level at the dominant strategy equilibrium.

The treatment effect on contributions is statistically distinguishable from zero, as predicted by theory. The treatment effect on deviations from the DSE is *also* statistically distinguishable from zero, which is not predicted by theory. Table 1.4 provides some insight into this effect: At the endowment of 15, deviations from the DSE are larger in the tax-efficient condition than in the tax-inefficient condition.

As table 1.3 shows, the tax treatment effect on contributions is always in the expected direction. This supports hypothesis 3.

Table 1.7: Main effect of tax treatment, Wilcoxon signed-rank tests (p-values)

	Contributions	Deviations from DSE
	0.00000	0.00000

Finding 1.3. *Participants contributed more in the "tax-inefficient" treatment than in the "tax-efficient" treatment. Hypothesis 3 is supported.*

1.6.5 Hypothesis 4: Endowment response

To estimate the effect of each treatment, I estimated random-effects models on treatment dummies with full interactions. The baseline case is the tax-inefficient treatment, with an endowment of 10. Results are reported in table 1.9. Since the dominant strategy is an interior solution, a censoring model is not necessary for evaluating the treatment effects.

Random-effects models are appropriate here because treatments are presented in a random order to all participants, so treatment is uncorrelated with participant characteristics and there is no reason to think that treatment order is correlated with participant characteristics. Moreover, a Hausman test comparing fixed effects and random effects on all the within-participant treatments was not statistically significant ($p > 0.50$).

Contributions were higher with larger endowments, and lower in the tax-efficient treatment; both of these comparative static predictions are consistent with theoretical predictions and the above findings for hypotheses 3 and 4. The magnitudes of the increase in contributions with endowment is somewhat less than one-to-one, resulting in contributions closer to the dominant strategy equilibrium for the endowment of 20 than the endowment of 10.

Contributions were lower in the tax-efficient treatment than in the tax-inefficient treatment, in the direction and approximate magnitude as the model predictions.

Table 1.10 shows the result of Wilcoxon signed-rank tests on the effects of changing endowments. All effects are statistically distinguishable from zero at $p < 0.01$, and as can be seen from table 1.3, all are in the expected direction where increasing endowment increases contributions. This result strongly supports hypothesis 4.

Table 1.8: Random Effects Models of Contributions

Constant	5.114***	(0.185)		
Tax-Efficient	−2.770***	(0.169)	−2.769***	(0.169)
Income 15	3.698***	(0.151)	3.698***	(0.151)
Income 20	6.733***	(0.190)	6.734***	(0.190)
Tax-Efficient × 15	−2.447***	(0.214)	−2.448***	(0.214)
Tax-Efficient × 20	−3.299***	(0.269)	−3.300***	(0.269)
Session Fixed Effects	No		Yes	
Observations	5,200		5,200	
Adjusted R^2	0.446		0.446	
Note:			*p<0.1; **p<0.05; ***p<0.01	

Table 1.9: Random effects models estimating the effects of each treatment on contributions.

Table 1.10: Effect of changing endowment, Wilcoxon signed-rank tests (p-values)

	10 to 15	15 to 20	10 to 20
Tax-efficient	0.00000	0.00000	0.00000
Tax-inefficient	0.00000	0.00000	0.00000

Finding 1.4. *Participants contributed more when their endowment was higher. Hypothesis 4 was supported.*

1.7 Other hypotheses: Altruism/spite effect, inequality effect

Public good games have been widely used to study other regarding preferences (Plott and Smith, 2008). In this section, I test whether such preferences can explain behavior in this experiment. The effects of altruism and inequity aversion can be readily predicted from the model.

Suppose that players assign some value β to the payoffs of others, which is received as a non-monetary payoff. $\beta > 0$ indicates altruism; $\beta < 0$ indicates spite. Still assuming that players are risk-neutral, this becomes their full payoff, including both monetary and non-monetary payoffs:

$$\pi_i = x_i - \frac{x_i^2}{M} + \gamma_V \left(\frac{1}{N} + \beta\right) \sum_j [w_j - x_j] + \gamma_T \left(\frac{1}{N} + \beta\right) \sum_j \left[\frac{x_i^2}{M}\right] + \beta \sum_j x_j - \beta \sum_j \left[\frac{x_i^2}{M}\right]$$

This is the new first-order condition:

$$0 = (1+\beta) - 2(1+\beta)\frac{x_i}{M} - \gamma_V \left(\frac{1}{N} + \beta\right) + \gamma_T \left(\frac{1}{N} + \beta\right) 2\frac{x_i}{M} + \beta$$

Again there is a dominant strategy, but it is not the same one as before.

$$x_i = \frac{M}{2} \frac{1 + \beta(1-\gamma_V) - \frac{\gamma_V}{N}}{1 + \beta(1-\gamma_T) - \frac{\gamma_T}{N}}$$

For $\beta > 0$, x_i will be larger than the selfish case if $1 - \gamma_V > 1 - \gamma_T$, i.e. if $\gamma_V < \gamma_T$; and if $\gamma_V > \gamma_T$, it will be smaller. For $\beta < 0$, this is reversed.

Since x_i is the amount kept, this means that altruistic participants would contribute more than the individual optimum if $\gamma_V > \gamma_T$, and less if $\gamma_V < \gamma_T$, while spiteful participants would do the opposite.

What about inequity aversion? Suppose participants have Fehr-Schmidt (1999) preferences, with envy α and guilt β, $\alpha \geq \beta \geq 0$:

$$\pi_i = y_i - \frac{\alpha}{N-1}\sum_j \max\{y_j - y_i, 0\} - \frac{\beta}{N-1}\sum_j \max\{y_i - y_j, 0\}$$

Since all participants get the same amount of the public goods, $y_i > y_j$ if and only if $x_i > x_j$. Thus we can write the individual's total payoff as:

$$\pi_i = x_i - \frac{x_i^2}{M} + \frac{\gamma_V}{N}\sum_j [w_j - x_j] + \frac{\gamma_T}{N}\sum_j \left[\frac{x_i^2}{M}\right] - \frac{\alpha}{N-1}\sum_j \max\{x_j - x_i, 0\} - \frac{\beta}{N-1}\sum_j \max\{x_i - x_j, 0\}$$

Let $N[x_j > x_i]$ be the number of participants with kept amounts x_j greater than x_i, and likewise $N[x_j < x_i]$ be the number of participants with kept amounts less than x_i. The first-order condition is then:

$$0 = 1 - 2\frac{x_i}{M} - \frac{\gamma_V}{N} + \frac{\gamma_T}{N}2\frac{x_i}{M} + \alpha\frac{N[x_j > x_i]}{N-1} - \beta\frac{N[x_j < x_i]}{N-1}$$

$$x_i = \frac{M}{2}\frac{1 - \frac{\gamma_V}{N} + \alpha\frac{N[x_j > x_i]}{N-1} - \beta\frac{N[x_j < x_i]}{N-1}}{1 - \frac{\gamma_T}{N}}$$

There is no longer a dominant strategy; an individual's optimal contribution depends upon the choices of others. However, we can still infer some comparative statics: If x_i is small, the α term will dominate and the individual will be incentivized to keep more (contribute less). If x_i is large, the β term will dominate and the individual will be incentivized to keep less (contribute more). Both of these effects will tend to equalize the final monetary payoffs across individuals, meaning that those who start with larger endowments will contribute more and those who start with smaller endowments will contribute less.

These results yield two additional hypotheses:

6. *Altruism/spite effect*: If participants contribute above the dominant strategy equilibrium when $\gamma_V > \gamma_T$, they will contribute below it when $\gamma_V < \gamma_T$, and vice-versa.

7. *Inequality effect*: Participants with larger endowments than the others in their group will contribute above the dominant strategy level by a larger amount.

As can be seen most clearly in table 1.6, contributions are on average above the dominant strategy level in all treatments, even when this is *further* from the Pareto-efficient level (i.e. in the tax-efficient case).

This refutes hypothesis 6. Results are not explained by either altruism or spite.

As can be seen in table 1.3, contributions are actually closer to the dominant strategy, or even below, when individual endowments are higher. Inequity aversion predicts the opposite: Those with smaller endowments should contribute less, and those with larger endowments should contribute more, in order to equalize the final payoffs.

This refutes hypothesis 7. Results are not explained by inequity aversion.

1.8 Additional Findings

1.8.1 Order effects

The order of the treatments was randomized across participants, in order to test for order effects. There do appear to be some order effects, as can be seen in table 1.11.

Table 1.11: Deviation from dominant strategy contribution, Wilcoxon signed-rank tests (p-values)

Tax-Efficient Last, Inequality Last 4 sessions	Taxes inefficient	Taxes efficient
Income = 10	0.00000	N/A
Income = 15	0.00000	0.00000
Income = 20	0.00000	0.13460
Tax-Efficient Last, Inequality First 2 sessions	**Taxes inefficient**	**Taxes efficient**
Income = 10	0.00000	N/A
Income = 15	0.01106	0.00007
Income = 20	0.05887	0.03893
Tax-Efficient First, Inequality Last 3 sessions	**Taxes inefficient**	**Taxes efficient**
Income = 10	0.00000	N/A
Income = 15	0.00000	0.00000
Income = 20	0.00000	0.00000
Tax-Efficient First, Inequality First 2 sessions	**Taxes inefficient**	**Taxes efficient**
Income = 10	0.00000	N/A
Income = 15	0.42180	0.00000
Income = 20	0.01528	0.65560

These p-values are not corrected for multiple comparisons. Applying the Bonferonni correction, the corrected target p-value is 0.0025. With this correction, the deviations from DSE are statistically significant in 13 out of 20 treatment cells.

1.8.2 Learning and end-game effects

Participants often change their behavior over the course of an experimental game; this has been attributed to learning about the game and about the behavior of other participants (Andreoni and Croson 2008). This particular environment is more complex than most other public goods games that have been used in experiments, which means that learning might be particularly important in understanding participant choices.

In order to detect changes in participant behavior that might reflect learning or reaction to the behavior of other participants, I performed Mann-Whitney tests on the first half and second half of each treatment sequence. Those results are reported in table 1.12. Contributions early in a treatment are no more likely to be above than below contributions late in a treatment, indicating that learning was not the cause of contributions above the DSE.

Table 1.12: Changes between each half of treatment (evidence of learning), Mann-Whitney U test (p-values)

Rounds	1-5/6-10	11-15/16-20	21-25/26-30	31-35/36-40
	0.2784	0.2934	0.06427	0.9521

1.8.3 Testing for decay over time

Another common result in public goods games experiments is an overall decay of contributions over time (Chaudhuri, Paichayontvijit, and Smith 2017). In order to test for such decay, I estimated random-effects models using a quadratic time trend. Results are reported in table 1.13.

Controlling for treatment effects, there is no evidence of decay in contributions over time. A quadratic time trend in contributions is not statistically distinguishable from zero with or without participant fixed effects.

In fact, there is a small but statistically significant *increasing* trend in contributions over time. This lack of decay over time could be a result of the "stranger" matching protocol which effectively resets the game at each round; however, contribution decay has been observed in some other experiments even with "stranger" matching protocols (Andreoni and Croson 2008).

Table 1.13: Random effects model with time trend, testing for decay of contributions over time.

Round	−0.017	−0.017
	(0.018)	(0.017)
Round^2	0.0003	0.0003
	(0.0004)	(0.0004)
Tax-efficient	-2.702^{***}	-2.773^{***}
	(0.188)	(0.170)
Income 15	3.760^{***}	3.693^{***}
	(0.169)	(0.152)
Income 20	6.886^{***}	6.717^{***}
	(0.209)	(0.191)
Tax-efficient × 15	-2.533^{***}	-2.462^{***}
	(0.238)	(0.215)
Tax-efficient × 20	-3.460^{***}	-3.281^{***}
	(0.296)	(0.269)
Constant	5.206^{***}	8.988^{***}
	(0.208)	(0.556)
Participant fixed effects	No	Yes
Observations	5,200	5,200
R^2	0.400	0.526
Adjusted R^2	0.399	0.514
Residual Std. Error	3.697 (df = 5192)	3.325 (df = 5063)
F Statistic	494.224^{***} (df = 7; 5192)	41.386^{***} (df = 136; 5063)
Note:		*p<0.1; **p<0.05; ***p<0.01

1.9 Discussion

Participant behavior is systematic and responds to treatment effects in a reliable way. However, the results are not entirely consistent with self-interested behavior, altruistic behavior or most theories of other-regarding preferences. Participants contribute above the dominant strategy equilibrium in both the tax-inefficient *and* tax-efficient conditions, which would indicate altruism in the former case but spite in the latter case. Participants with high endowments overcontribute, if anything, *less* than those with low endowments, which inequity aversion also cannot explain.

Warm glow altruism might be able to explain this effect, if individuals have a preference for the act of giving intentionally (as opposed to being forced to give via taxation). There is evidence of warm-glow preferences in past experiments(Crumpler and Grossman, 2008), which has important implications for the optimal design of tax policy regarding charitable contributions(Diamond, 2006).

Some of this effect could be explained through the use of a simple anchoring heuristic: Participants anchor to a "fair split" heuristic, giving half of their endowment, and then partially adjust toward the dominant strategy equilibrium. When asked to provide written explanations for their behavior at the end of the experiment, a few participants responded explicitly with similar heuristics, such as: "The more income I had, the more I chose to contribute, usually just slightly over half." This is also consistent with behavior in dictator games with taking(Cappelen et al., 2013), where adding a "take" option effectively moves the "middle" to a different position, reducing contributions.

A related possibility is that participants may be trying to conform to some social norm; if the norm is "contribute half", then a norm-based model similar to that of López-Pérez (2008) is a good fit for the data, as it would also result in behavior intermediate between contributing half and following the dominant strategy. Whether this is anchoring or a norm, in either case, the question still remains why it is to contribute half, and not something else, such as contributing all, contributing none, or contributing 3/4. Future research could adjudicate between these two models by attempting to induce a specific norm while also providing anchors that are distinct from the norm, to see which effect pulls participants more strongly.

Another possible explanation for overcontribution even under the tax-efficient condition is what might be called *tax aversion*, a willingness to pay additional cost simply to avoid having to pay taxes, even if the result is less after-tax income, or *preference for control*, a willingness to incur cost in order to be able to decide how much to contribute rather than have it taken without control. The former could be tested by running the same experiment again with more neutral framing, renaming "contribution" and "tax" as something like "account 1" and "account 2". If tax aversion is a cause of overcontribution, contributions should be reduced under such framing. The latter could be tested by making "account 2" also a voluntary contribution.

Other participants indicated intentions of reciprocity or mimicry—several mentioned trying to "match" or "copy" what other participants had done. This would have the effect of reducing between-participant variation, but does not by itself explain any particular level of contribution. (Indeed, it effectively amounts to the same equilibrium selection problem as occurs in the indefinitely iterated game, even though in the finitely iterated game such cooperation is not actually a subgame perfect equilibrium.) The lack of end game effects suggests that apparent indefiniteness was not a major factor; while it is unclear whether they were applying backward induction, participants seemed to understand that they were in a finite game and yet deviated from the dominant strategy regardless.

The most important effects came from the change in MPCR on taxes, which in real-world terms reflects the perceived efficacy or legitimacy of public good provision by the government. While Sheremeta and Uler (2016) found that government transfers are complementary to voluntary contributions, I found that participants government spending and contributions as substitutes: when tax revenue became more efficient, participants contributed less and paid more in taxes. This difference may arise because in my environment, tax revenue provides a public good, while in theirs, it is redistributive or wasteful. This difference could have important implications for the design of tax policy; the deduction on charitable contributions may have quite different effects depending on whether taxpayers perceive government spending to be efficiently providing public goods or wastefully redistributing income.

Note that in the lab, the MPCR is known and highly credible, while in reality the benefit of tax revenue and private contributions is uncertain and the perceived benefit depends upon the trust an individual has in government or other institutions. Since I found that increasing the MPCR of tax revenue decreased contributions, this suggests that citizens with high levels of trust in government are likely to contribute less to charity, preferring to have their income taxed and used for public goods provided by the government, while citizens with low levels of trust in government are likely to contribute more to charity, preferring to avoid taxation. This also suggests that improving trust in government could have the perverse effect of crowding-out charitable contributions (though if that trust is well-placed, overall public goods provision should still increase). This is a potential topic for future research, comparing countries by charitable contributions versus trust in government-provided public goods.

If tax aversion or preference for control turns out to be an important effect, then the charitable tax deduction remains a useful policy, even if the size of the deduction is relatively unimportant. The preference to avoid taxes may motivate some donors to contribute even though their after-tax income is not improved by the donation; thus, having *some* tax deduction may be beneficial, even if it is not particularly large.

The high prevalence of apparent anchoring or norm-following behavior suggests that behavioral nudges should be highly effective at increasing rates of contribution. In order to change behavior, it is not necessary to change the real monetary incentives, only the salient anchors and norms. This is consistent with empirical literature on nudges(Schulz et al., 2018) showing quite large effects of nudges on charitable contributions.

1.10 Conclusion

Using a progressive tax with deductible contributions, I induce an interior dominant strategy equilibrium in public goods games using a framing that is similar to real-world tax policy in most countries around the world. I find that contributions above the dominant strategy are a consistent, systematic behavior, not explicable by random error or tremble. Using a "stranger" re-matching design, I find no evidence of decay of contributions over time or end game effects.

However, I also find that altruism is not a good explanation for behavior, since when the MPCR of tax revenue is higher than that of contributions, the Pareto efficient outcome is for all participants to contribute zero, which is strictly less than the dominant strategy contribution. Yet even in such cases I still observe contributions that are systematically higher than the dominant strategy. Participants were not ignoring the change in MPCR, as it did have strong effects on their behavior—but they did not adjust in a way that would be consistent with altruism or seeking the Pareto-efficient outcome.

The result may be explained by heuristic behavior, likely some form of anchoring, where participants anchor to giving half their endowment (perhaps based on intuitive ease of calculation, or perhaps some perceived social norm), and then partially adjust toward the dominant strategy equilibrium. This model explains both the strong treatment effects of the change in MPCR and the fact that overcontribution persists even when undercontribution is Pareto-improving.

The effect of changing the MPCR on tax revenue suggest that taxes may be perceived as a substitute for charitable contributions, such that citizens are likely to contribute less if they trust government provision of public goods more (and vice-versa). The high prevalence of anchoring and norm-following is consistent with empirical evidence that behavioral nudges can have large effects on contribution behavior.

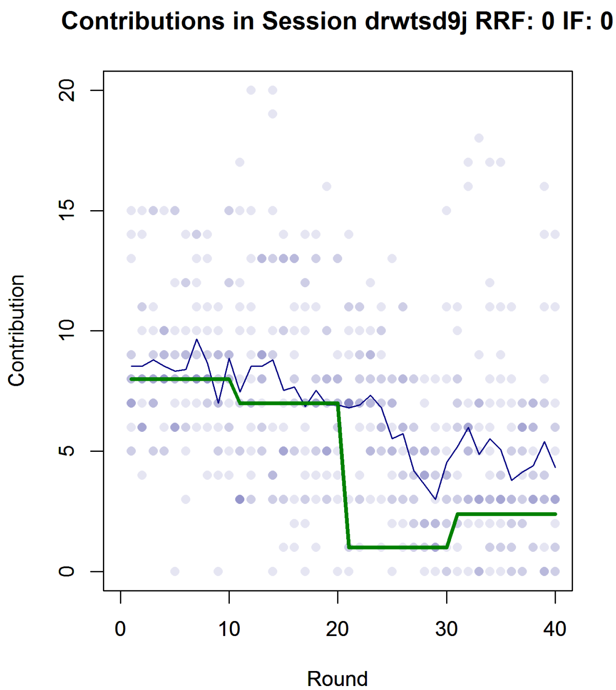


Figure 1.1: Contributions in session drwtsd9j with Tax-Efficient last, Unequal Endowments last

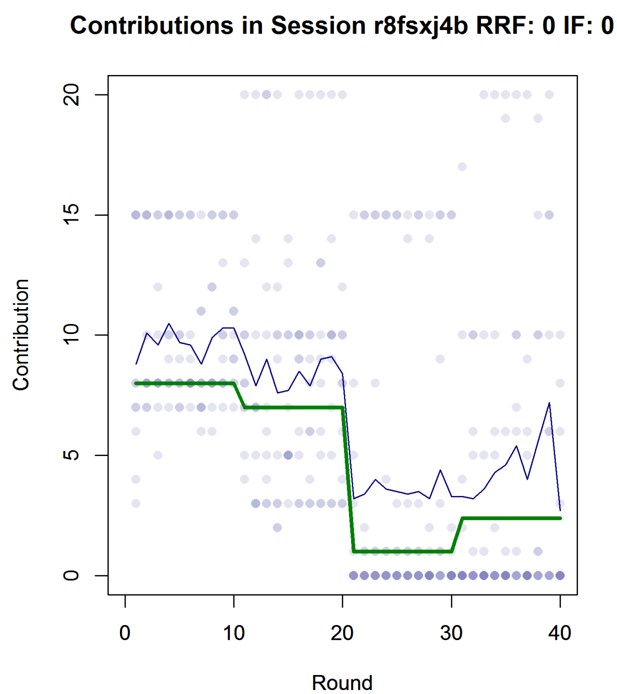


Figure 1.2: Contributions in session r8fsxj4b with Tax-Efficient last, Unequal Endowments last

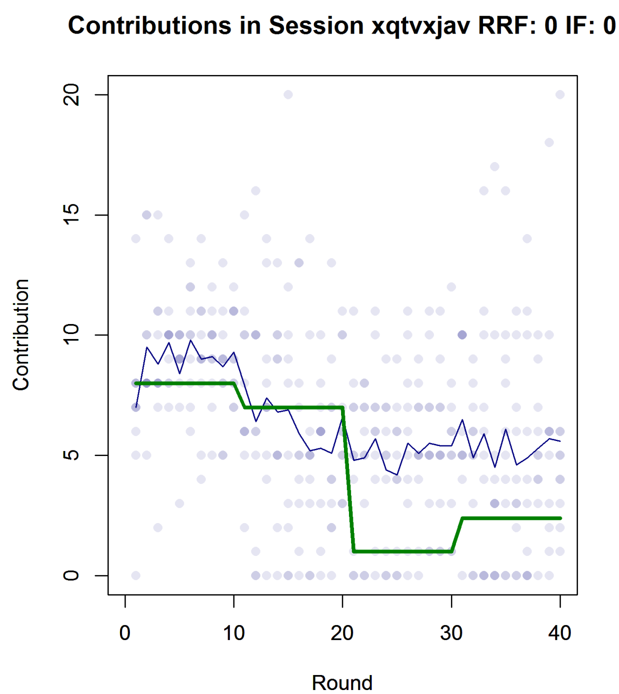


Figure 1.3: Contributions in session xqtvxjav with Tax-Efficient last, Unequal Endowments last

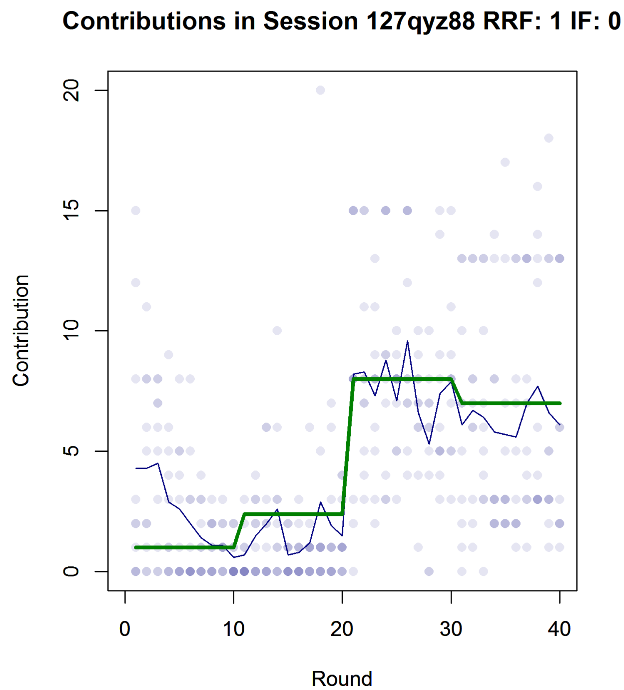


Figure 1.4: Contributions in session 127qyz88 with Tax-Efficient first, Unequal Endowments last

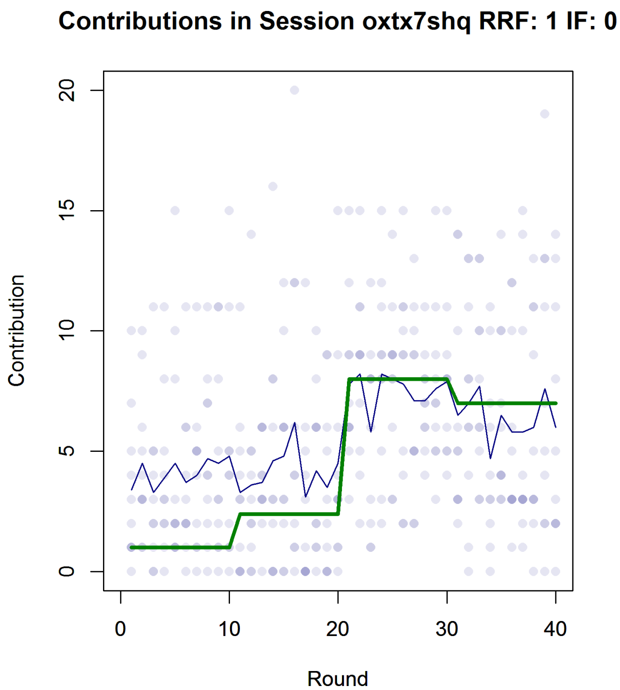


Figure 1.5: Contributions in session oxtx7shq with Tax-Efficient first, Unequal Endowments last

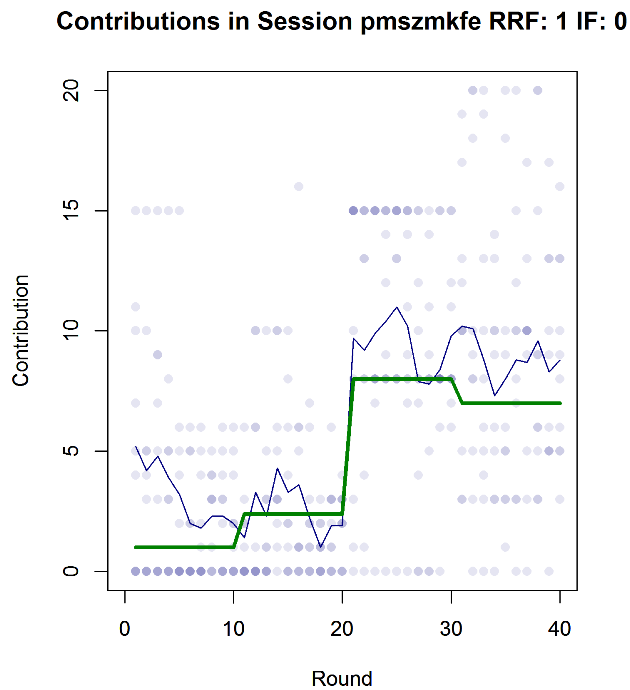


Figure 1.6: Contributions in session pmszmkfe with Tax-Efficient first, Unequal Endowments last

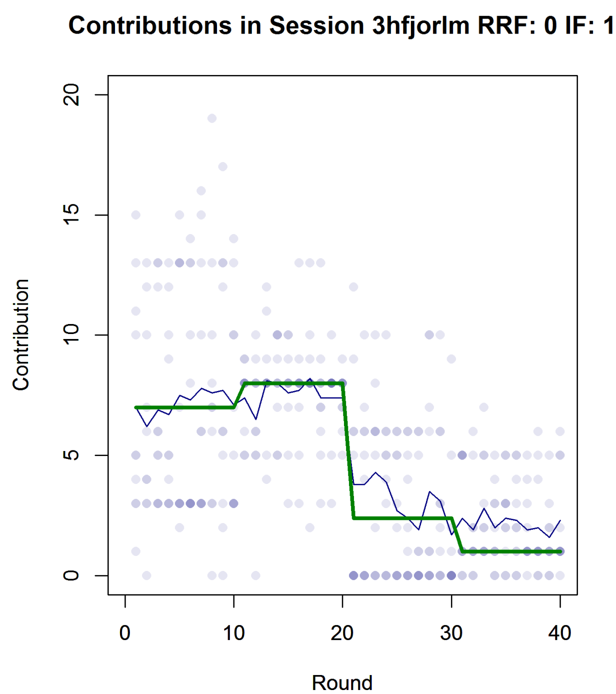


Figure 1.7: Contributions in session 3hfjorlm with Tax-Efficient last, Unequal Endowments first

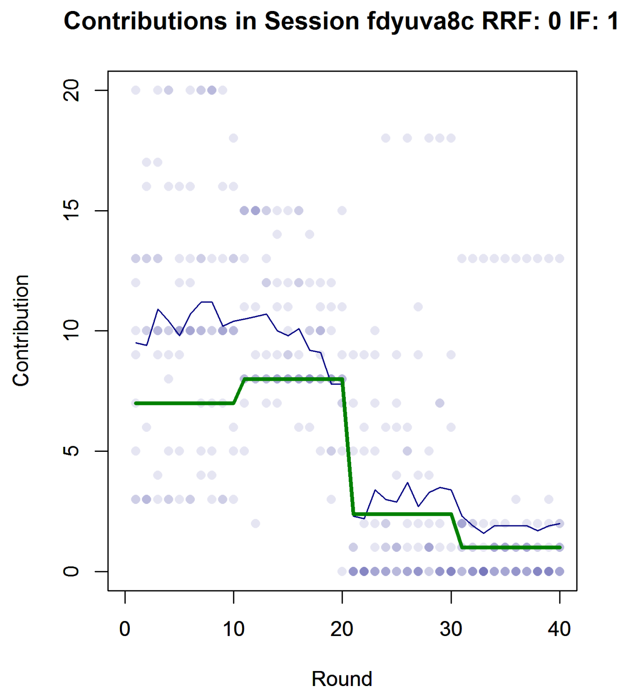


Figure 1.8: Contributions in session fdyuva8c with Tax-Efficient last, Unequal Endowments first

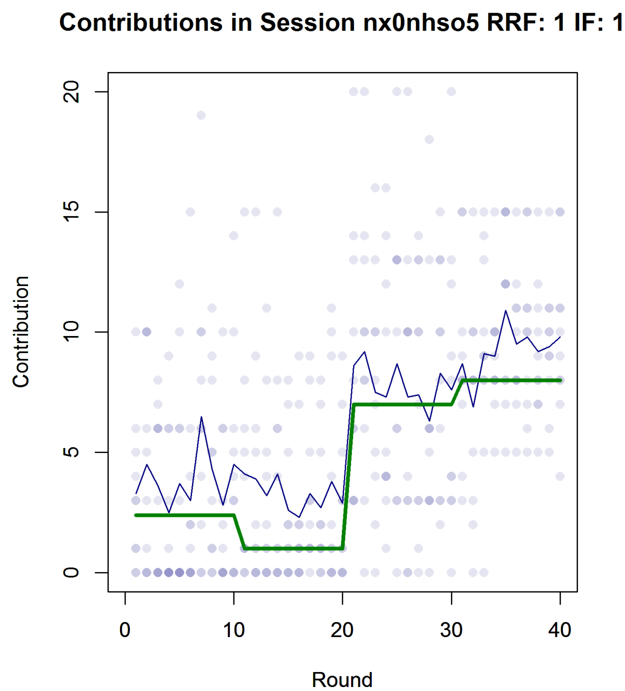


Figure 1.9: Contributions in session nx0nhso5 with Tax-Efficient first, Unequal Endowments first

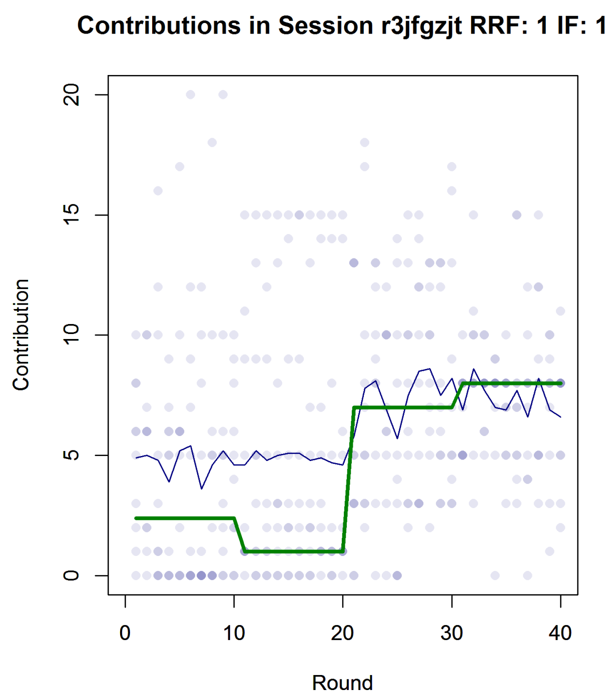


Figure 1.10: Contributions in session r3jfgzjt with Tax-Efficient first, Unequal Endowments first

Chapter 2

International Norms of Peace, War, and Sanction: A Microfounded Iterated Game Approach

Abstract

Wars are less frequent today than through most of history. Various explanations for this have been proposed, ranging from the spread of democracy, to globalized trade, to the invention of nuclear weapons.

In this paper I develop an economically microfounded model of international conflict as an indefinitely iterated game, showing that under typical parameters there exist many equilibria, some resulting in war, others in peace. I propose that international norms provide a plausible equilibrium selection mechanism for this iterated game. This offers a synthesis between the "realist" and "liberal" schools of international relations: The realists are correct that international norms are cheap talk, not written into the rules of the game; but the liberals are correct that norms can have real effects, when they select one equilibrium over another.

2.1 Introduction

Why is war now so rare? The idea that war is rare may come as a shock to some, given how frequently wars are found in the news. But if we compare to the number of wars that *could* be happening and would seem to be strategically advantageous, or to the rate at which wars have occurred throughout most of human history, wars today are indeed extremely rare.

From 1500 to 1800, there was never a period of more than 25 years in which no major powers were at war with one another (Roser, 2016). Well into the 20th century, wars between major powers were a frequent occurrence. But something changed after the Second World War (Gaddis, 1992). Wars are less frequent now than they were even in the 1980s. This trend toward greater peace has been referred to as the "Long Peace"(Kriesberg, 2007).

One leading theory of this shift is that it is the result of nuclear weapons, which make war too costly to contemplate. Yet this cannot explain why nuclear powers rarely invade non-nuclear powers, particularly when those non-nuclear powers are not closely allied with any other nuclear powers. Yes, the US has invaded Iraq and Afghanistan, and participated in more limited military interventions in a variety of places around the world. Yet even the fact that these interventions were small and often covert demands explanation—as this was certainly not the case for unilateral invasions by major powers in the 16th century—and there are literally dozens of countries that the US *could* have invaded without expecting nuclear retaliation but didn't. Moreover, no country has used nuclear weapons in war since the two atomic bombs that the US dropped on Japan in 1945, even in cases where the defending country had no nuclear weapons and all major nuclear powers were either neutral or actively on the same side (as in the Iran-Iraq War)—yet in these cases theory predicts that nuclear weapons should make war *more* likely rather than less (Rauchhaus, 2009).

The primary purpose of this paper is to develop economically microfounded models of international conflict that can explain the Long Peace and derive conditions under which it might be expected to continue into the future. The key insight is *multiple equilibria*: the iterated game has many possible equilibrium outcomes.

I propose that equilibria are selected and sustained by mutual adherence to emergent *norms*. Once a norm has emerged, it can be sustained against shocks up to a certain size; and equilibrium norms can exist resulting in either peace or war, in a path-dependent way. This approach which uses norms as an equilibrium selection mechanism was pioneered by Binmore (Binmore, 2010), though it is not the only approach to understanding norms in terms of game theory(Paternotte and Grose, 2013).

Understanding how these norms form and break could have profound implications: Due to the extreme long-tail risks that nuclear war poses, the future of human civilization may rest upon whether we can continue to avoid major wars in a world with nuclear arsenals ((Sidel, 1989; Bostrom, 2002; Kroenig, 2015). Even a few small insights into how to make peace more likely have a very large expected payoff for humanity.

2.2 Related Literature

The recent period of peace between major powers has been referred to by many scholars as the "Long Peace". The term was first used in print by Gaddis (1986), but there is now an extensive literature in political science and international relations on the Long Peace, ranging from descriptive historical evidence (Gaddis, 1992) to analyses suggesting it is not as unusual as previously thought (Siverson and Ward, 2002), and positing explanations for its occurrence including international trade (Bearce, 2003), the spread of democracy (Mesquita et al., 1999), US military hegemony, (Lebow, 1994), a reduction in the supply of willing soldiers (Inglehart et al., 2015), and the advent of nuclear weapons (Rauchhaus, 2009). These explanations are not necessarily mutually exclusive; the "Liberal Peace" account unites democracy, trade, and international institutions as the mechanisms for sustaining peace (Doyle, 2005).

Game theory also has a long tradition in political science, but has largely been restricted to simple static games, often static 2×2 games and in fact most often specifically the Prisoner's Dilemma. Game theory is much richer than this, and by excluding multiplayer games, games with broader strategy spaces, and, above all, indefinitely iterated games, political scientists may have been missing out on important insights that game theory could have provided.

Iterated game theory is extremely rich; in a sense, *too* rich, as the folk theorems (Fudenberg and Tirole, 1991) allow for a vast multiplicity of equilibria even in quite simple games such as the Prisoner's Dilemma. Since a wide variety of real-world phenomena have uncertain repetition and are thus indefinitely iterated games, this poses a problem: It is very difficult to predict behavior when there are a vast number of possible equilibria. Even if we assume that *some* equilibrium is reached, we cannot say *which* equilibrium will be reached: This is the "equilibrium selection problem" (Harsanyi and Selten, 1988).

Yet it may be that "solving" the equilibrium selection problem is the wrong approach altogether. Perhaps we should take the game theory seriously and allow that the world may be full of genuine strategic uncertainty which human beings are forced to navigate. Perhaps the vast multiplicity of outcomes is not a bug but a feature: We in fact *do* observe distinct equilibrium outcomes as we look across times and places. There have been periods of war and periods of peace, both within nations, between neighboring nations, and across the world as a whole. The question then becomes how real-world individuals and institutions can navigate this strategic uncertainty and achieve an equilibrium outcome.

One answer is to use *norms* as an equilibrium selection mechanism (Binmore, 2010). We know that a great deal of human behavior is motivated by social norms—indeed, I would argue that, to a first approximation, all human behavior is social norms. Norms decide where we live, what we wear, what we eat, how we speak, whom we interact with, even what we believe. Yet the formal modeling of norms in economic theory remains very limited, and different researchers have used greatly different approaches: Some treat norms as a form of social preferences (Bicchieri, 2005), others use them to expand the available strategy space via signaling mechanisms (Gintis, 2010). Norms have also been used extensively in the international relations literature, particularly within the "constructivist" school(Checkel, 1997; Finnemore and Sikkink, 2001). In most of this literature, norms are *contrasted* with rationality—norms are given as a reason to do something which is distinct or even contrary to rationality. But in my approach, norms are a *subset* of rationality—they are the form of rationality that selects equilibria in coordination games.

In a game with a vast space of possible equilibria, some mechanism is needed to coordinate behavior on one particular equilibrium. Norms, which may be explicit or implicit, simply tell players which strategies others intend to play, incentivizing them to coordinate on one particular equilibrium outcome.

This has relevance for one of the central conflicts in international relations theory, namely the "realist" versus "liberal" schools (Powell, 1994; Reus-Smit and Reus-Smit, 2008). Realists argue that international treaties and international law are just cheap talk, worth no more than the paper they were printed on. Liberals argue that well-designed international institutions can maintain peace and promote economic development.

I argue that *both* sides are correct: International treaties *are* indeed cheap talk, and they *can* indeed maintain peace and promote development. That is to say, cheap talk can have real effects, because we are in an indefinitely iterated game with a vast number of equilibria. Simply *saying* that you will choose one equilibrium over another, without any direct enforcement compelling you to follow through, can create incentives for both you and others to genuinely conform to that equilibrium.

2.3 Static Model

Consider a normal-form game with 2 players. Each player i represents a country with capital stock K_i and population L_i.

Each country produces GDP using a Cobb-Douglas production function $Y_i = K_i^\alpha L_i^{1-\alpha}$.

The utility of a country in each period (e.g. one year) is given by a representative agent with constant relative risk aversion of 1. The use of risk-averse agents is realistic, and an important contribution of this paper: The literature on contests with risk aversion remains relatively sparse (Konrad and Schlesinger, 1997; Cornes and Hartley, 2012; Treich, 2010; Cornes and Hartley, 2003).

$$\pi_i = \ln\left[\frac{Y_i}{L_i}\right] = \ln\left[\frac{K_i^\alpha L_i^{1-\alpha}}{L_i}\right] = \alpha \ln\left[\frac{K_i}{L_i}\right]$$

For technical reasons, I will assume that each country's GDP is at least 1, i.e. $\alpha \ln\left[\frac{K_i}{L_i}\right] \geq 1$.

Each country chooses a unilateral strategy toward the other country separately and privately. A country may choose "Trade" or "War". If at least one country chooses "War", war occurs between them. If both countries choose "Trade", trade occurs between them.

Countries can choose to apply a sanction $s \in [0,1]$, which reduces the quantity of trade goods received by both countries, if trade occurs. If the two countries choose different levels of sanctions, the highest is applied. Countries may also choose to pay tribute τ as a direct transfer of income from one country to the other. Country i pays $\tau_{ij} \geq 0$ to country j and receives $\tau_{ji} \geq 0$ from country j. The two are separate because τ_{ij} is under the control of country i while τ_{ji} is under the control of country j.

If trade occurs, the quantity traded follows a gravity model; gravity models have been used extensively in the trade literature, have well-developed microfoundations, and are empirically well-supported(Isard, 1954; Bergstrand, 1985; Krugman, 1991; Oguledo and Macphee, 1994; Kepaptsoglou et al., 2010; Anderson, 2011).

The total value of trade between countries with GDP Y_i and Y_j is, where γ is a constant and D_{ij} reflects the distance (both geographical and cultural) between the two countries:

$$\frac{\gamma(1-s)}{D_{ij}} Y_i Y_j$$

This results in the following payoff for trade:

$$\pi_i(\text{trade}) = \ln\left[\frac{Y_i}{L_i} + \frac{\gamma(1-s)}{D_{ij}}\frac{Y_i Y_j}{L_i} + \frac{\tau_{ji}-\tau_{ij}}{L_i}\right] = \ln\left(\frac{Y_i}{L_i}\right) + \ln\left[1 + \frac{\gamma(1-s)}{D_{ij}} Y_j + \frac{\tau_{ji}-\tau_{ij}}{Y_i}\right]$$

$$\pi_i(\text{trade}) = \alpha \ln\left[\frac{K_i}{L_i}\right] + \ln\left[1 + \frac{\gamma(1-s)}{D_{ij}} Y_j + \frac{\tau_{ji}-\tau_{ij}}{Y_i}\right] \tag{2.1}$$

First, note that this payoff is strictly decreasing in s; thus, in static Nash equilibrium, no country would ever choose positive sanctions. This is intuitive; since sanctions also hurt the country that imposes them, there is no incentive to impose them. This payoff is also strictly increasing in τ_{ji} and strictly decreasing in τ_{ij}; thus each country would prefer to receive tribute, but in equilibrium will give none.

Second, note that this payoff is strictly increasing in both K_j and L_j: Trade with a larger country is more beneficial due to the greater quantity of imports and exports. Nothing constrains total trade from exceeding GDP for small countries; this is entirely realistic, as some small countries such as Taiwan and Singapore do indeed have trade in excess of 100% of GDP.

If war occurs, each country's payoff is decided by a contest function. Contest functions are now widely used in the analysis of conflict (Skaperdas, 1996). They provide a simple but powerful tool for analyzing conflicts where effort invested is directly opposed to the efforts of others—of which warfare is a prime example. I will use the standard Tullock ratio-form contest function as is commonly found in the literature (Cubel and Sanchez-Pages, 2016; Jia et al., 2013) and has been empirically validated as a reasonable approximation to important real-world conflicts (Hwang, 2009).

If war occurs, each country chooses its level of military spending g_i. These levels are chosen secretly and simultaneously, but only if war occurs. If trade occurs, both countries spend zero on military spending. This may be unrealistic, as peacetime military spending, while small, is generally not zero. Peacetime military spending can have other effects, such as increasing readiness or providing a commitment mechanism, but these are outside the scope of this paper.

Military spending comes with a cost. With this utility function, the most microfounded cost function would be as follows:

$$c(g_i) = \ln\left[\frac{Y_i}{L_i}\right] - \ln\left[\frac{Y_i - g_i}{L_i}\right] = -\ln\left[1 - \frac{g_i}{Y_i}\right]$$

Unfortunately, this turns out to be intractable with a ratio-form contest function. I will therefore take a first-order Taylor approximation to this function, as follows:

$$-\ln\left[1-\frac{g_i}{Y_i}\right] \approx -\ln\left[1-\frac{g_i}{Y_i}\right]\Bigg|_{g_i=0} + \frac{-\frac{1}{Y_i}}{1-\frac{g_i}{Y_i}}\Bigg|_{g_i=0}(g_i)+\dots$$

This results in the following cost function:

$$c(g_i) = \frac{g_i}{Y_i}$$

This first-order approximation obeys the following intuitively appealing properties:

1. Cost is non-negative: $c_i(g_i) \geq 0$
2. Cost is monotonically increasing and (weakly) convex: $\frac{\partial c_i}{\partial g_i} > 0$, $\frac{\partial c_i}{\partial g_i}2 = 0$
3. The cost of zero military spending is zero: $c_i(0) = 0$
4. The real cost of the same amount of military spending is higher for countries with lower GDP: $\frac{\partial c_i}{\partial Y_i} < 0$

This approximation is valid only if $\frac{g_i}{Y_i}$ is small, but this is reasonable, as except in extreme circumstances (such as the World Wars) most countries maintain military spending below 10% of GDP. At $\frac{g_i}{Y_i} = 0.10$, the approximation error is only approximately 0.005, or 5%.

If a country wins in war, they get to claim some capital from the other country. The amount of capital currently held by country i which is contestable is κ_i (and likewise κ_j for country

j). War also destroys some proportion of each country's capital ω; I will assume this is constant across countries.[1]

These are the payoffs for winning and losing:

$$\pi_i(\text{win}) = \ln\left[\frac{(K_i+\kappa_j)^\alpha (1-\omega)^\alpha}{L_i^\alpha}\right] = \alpha\ln\left[\frac{K_i+\kappa_j}{L_i}\right] + \alpha\ln(1-\omega)$$

$$\pi_i(\text{lose}) = \ln\left[\frac{(K_i-\kappa_i)^\alpha (1-\omega)^\alpha}{L_i^\alpha}\right] = \alpha\ln\left[\frac{K_i-\kappa_i}{L_i}\right] + \alpha\ln(1-\omega)$$

Combining all these yields the overall expected payoff for war:

[1]This assumption is largely innocuous, since any country j that suffers more destruction can instead be thought of as transferring more capital κ_j which is ultimately destroyed.

For known values of K_i before war and K_{iw} after winning and K_{il} after losing, this is a general solution of this system:

$$\omega = 1 - \frac{(K_{iw}+K_{jl})}{(K_i+K_j)} = 1 - \frac{(K_{il}+K_{jw})}{(K_i+K_j)}$$

$$\kappa_j = \frac{K_{iw}}{1-\omega} - K_i$$

$$\kappa_i = \frac{K_{jw}}{1-\omega} - K_j$$

For instance, let us consider an extreme case where country i has nothing to lose. Let ω_i be the true rate of capital destruction for country i, ω_j be the true rate of capital destruction for country j, and κ_{j0} be the true amount of capital transferred upon victory. Then $K_{il} = (1-\omega_i)K_i$, $K_{jw} = (1-\omega_j)K_j$, $K_{iw} = (1-\omega_i)K_i + (1-\omega_j)\kappa_{j0}$, $K_{jl} = (1-\omega_j)(K_j-\kappa_{j0})$.

$$\omega = 1 - \frac{(1-\omega_i)K_i + (1-\omega_j)\kappa_{j0} + (1-\omega_j)(K_j-\kappa_{j0})}{(K_i+K_j)} = \frac{\omega_i K_i + \omega_j K_j}{K_i+K_j}$$

$$\kappa_j = \frac{(1-\omega_i)K_i + (1-\omega_j)\kappa_{j0}}{1-\omega} - K_i = \frac{\omega-\omega_i}{1-\omega}K_i + \frac{1-\omega_j}{1-\omega}\kappa_{j0}$$

$$\kappa_i = \frac{(1-\omega_j)K_j}{1-\omega} - K_j = \frac{\omega-\omega_j}{1-\omega}K_j$$

This does result in $\kappa_i < 0$, which may make it seem as though country i may actually prefer losing to not fighting at all. However, this is not actually the case, because of the effect of ω_i. Since $K_{il} \le K_i$ as long as $\omega_i \ge 0$, it follows that $\pi_i(\text{lose}) \le \alpha\ln\left[\frac{K_i}{L_i}\right]$, even when $\kappa_i < 0$. The only assumption required is that country i does not end up with more capital after losing the war than they would have if they had not fought at all.

$$\pi_i(\text{war}) = \frac{g_i}{g_i + g_j}\alpha \ln\left[\frac{K_i + \kappa_j}{L_i}\right] + \frac{g_j}{g_i + g_j}\alpha \ln\left[\frac{K_i - \kappa_i}{L_i}\right] + \alpha \ln(1-\omega) - \frac{g_i}{Y_i}$$

$$\pi_i(\text{war}) = \frac{g_i}{g_i + g_j}\alpha \ln\left[\frac{K_i + \kappa_j}{K_i - \kappa_i}\right] + \alpha \ln\left[\frac{K_i - \kappa_i}{L_i}\right] + \alpha \ln(1-\omega) - \frac{g_i}{Y_i}$$

Define R_i as follows:

$$R_i = \ln\left[\frac{K_i + \kappa_j}{K_i - \kappa_i}\right] = \ln\left[\frac{1 + \frac{\kappa_j}{K_i}}{1 - \frac{\kappa_i}{K_i}}\right] \tag{2.2}$$

We may think of this as the "return on victory" for country i. R_i is increasing in the capital stock κ_j that can be claimed from country j, relative to country i's current capital stock K_i, and also increasing in the capital κ_i that country i stands to lose if defeated. Note that $R_i > 0$.

This is a useful result, proposition 2.1:

Proposition 2.1. *The return on victory R_i is always positive, and depends only the* ratios *of contestable capital $\frac{\kappa_i}{K_i}$ and $\frac{\kappa_j}{K_i}$, and not the actual quantity of capital. It is increasing in $\frac{\kappa_j}{K_i}$ and decreasing in $\frac{\kappa_i}{K_i}$. It is independent of K_j.*

Thus we have:

$$\pi_i(\text{war}) = \frac{g_i}{g_i + g_j}\alpha R_i + \alpha \ln\left[\frac{K_i - \kappa_i}{L_i}\right] + \alpha \ln(1-\omega) - \frac{g_i}{Y_i}$$

Maximizing this payoff with respect to g_i yields the following first-order condition:

$$\frac{g_j}{(g_i + g_j)^2} R_i = \frac{1}{Y_i}$$

A similar first-order condition is found for g_j:

$$\frac{g_i}{(g_i + g_j)^2} R_j = \frac{1}{Y_j}$$

These can be used to find closed-form solutions for the the best-response functions as well as the Nash equilibrium level of military spending:

$$0 = g_j^2 + (2g_i - R_i Y_i) g_j + g_i^2$$

$$g_j = \frac{R_i Y_i}{2} - g_i + \sqrt{R_i Y_i} \sqrt{\frac{R_i Y_i}{4} - g_i} \qquad (2.3)$$

Whether the positive or negative root is the true maximum depends upon whether $R_i Y_i > R_j Y_j$ or $R_i Y_i < R_j Y_j$. We can think of this product $R_i Y_i$ as the "will to fight", as it is the product of GDP and return on victory. The country with the larger will to fight will have the positive root and the higher military spending, while the other will have the negative root and lower military spending. If $R_iY_i = R_jY_j$, in equilibrium the two roots will converge to a single repeated root.

Solving for the equilibrium levels of military spending yields the following:

$$(g_i + g_j)^2 = R_iY_ig_j = R_jY_jg_i$$

$$\frac{g_i}{g_j} = \frac{R_iY_i}{R_jY_j}$$

$$g_i^* = R_iY_i\frac{R_iY_iR_jY_j}{(R_iY_i + R_jY_j)^2} \tag{2.4}$$

This is the equilibrium probability of victory for country i:

$$P_i[\text{win}] = \frac{g_i^*}{g_i^* + g_j^*} = \frac{R_iY_i}{R_iY_i + R_jY_j}$$

Note that military spending is higher (and thus probability of victory is higher) if and only if will to fight is higher. This yields proposition 2.2:

Proposition 2.2. *Military spending, and thus the probability of victory, for country i is increasing in own GDP Y_i and decreasing in enemy GDP Y_j. It is increasing in own return on victory R_i and decreasing in enemy return on victory R_j.*

This also allows us to compute the equilibrium payoff for war:

$$\pi_i(\text{war}) = \frac{R_iY_i}{R_iY_i + R_jY_j}\alpha R_i + \alpha \ln\left[\frac{K_i - \kappa_i}{L_i}\right] + \alpha\ln(1-\omega) - \frac{R_iY_i}{Y_i}\frac{R_iY_iR_jY_j}{(R_iY_i + R_jY_j)^2}$$

$$\pi_i(\text{war}) = R_i\left(\alpha - \frac{R_jY_j}{R_iY_i + R_jY_j}\right)\frac{R_iY_i}{R_iY_i + R_jY_j} + \alpha \ln\left[\frac{K_i - \kappa_i}{L_i}\right] + \alpha\ln(1-\omega)$$

$$\pi_i(\text{war}) = R_i\left(\frac{R_iY_i}{R_iY_i + R_jY_j} - (1-\alpha)\right)\frac{R_iY_i}{R_iY_i + R_jY_j} + \alpha \ln\left[\frac{K_i - \kappa_i}{L_i}\right] + \alpha\ln(1-\omega) \quad (2.5)$$

By inspection, $\pi_i(\text{war})$ is increasing in return on victory R_i and total GDP Y_i, and decreasing in population L_i and destructiveness ω.

It is not apparent whether π_i is increasing or decreasing in secure capital $K_i - \kappa_i$, because R_i is decreasing in secure capital.

When is war preferable for country i?

War produces a higher expected payoff for country i precisely when $\Delta_i > 0$, where $\Delta_i = \pi_i(\text{war}) - \pi_i(\text{trade})$:

$$\begin{aligned}\Delta_i = {} & R_i\left(\frac{R_iY_i}{R_iY_i + R_jY_j} - (1-\alpha)\right)\frac{R_iY_i}{R_iY_i + R_jY_j} \\ & + \alpha \ln\left[\frac{K_i - \kappa_i}{L_i}\right] + \alpha\ln(1-\omega) \\ & - \alpha\ln\left[\frac{K_i}{L_i}\right] - \ln\left[1 + \frac{\gamma(1-s)}{D_{ij}}Y_j + \frac{\tau_{ji} - \tau_{ij}}{Y_i}\right]\end{aligned}$$

$$\Delta_i = R_i \left(\frac{R_i Y_i}{R_i Y_i + R_j Y_j} - (1-\alpha) \right) \frac{R_i Y_i}{R_i Y_i + R_j Y_j} \tag{2.6}$$

$$+ \alpha \ln \left[1 - \frac{\kappa_i}{K_i} \right] + \alpha \ln(1-\omega) \tag{2.7}$$

$$- \ln \left[1 + \frac{\gamma(1-s)}{D_{ij}} Y_j + \frac{\tau_{ji} - \tau_{ij}}{Y_i} \right] \tag{2.8}$$

I will refer to Δ_i as country i's *propensity for war.* A positive propensity for war indicates that war has a higher expected payoff than trade, while a negative propensity for war indicates the opposite. Note that propensity for war is independent of population L_i.

Proposition 2.3. *Δ_i is increasing in α if and only if $R_i P_i[\text{win}] > -\ln\left[1 - \frac{\kappa_i}{K_i}\right] - \ln(1-\omega)$. For any given value of the other parameters, there exists a threshold $\overline{R}$ such that for $R_i < \overline{R}$, Δ_i is increasing in α, and for any $R_i > \overline{R}$, Δ_i is decreasing in α.*

Proposition 2.3 means that, depending on the return on victory, the proportion of capital that is contestable, and the destructiveness of war, increased capital intensity of production can either increase or decrease the propensity for war. Capital intensity matters because capital can be claimed in war but labor cannot.[2] The combination of higher destructiveness and higher capital intensity together tends to decrease propensity for war.

Note that since war redistributes some capital and destroys other capital, but does not increase total output, trade always results in a higher surplus than war; if one country's

[2] I am thus implicitly excluding the ancient practice of using prisoners of war for forced labor. This is relatively rare in modern warfare, and certainly not common enough to provide a substantial proportion of GDP for most countries. Perhaps more significantly, I am also ignoring the loss of life in war, as with representative agent utility loss of life has the perverse effect of increasing utility by mechanically increasing per-capita GDP. This would imply that countries *prefer* to lose soldiers in warfare, which is obviously untrue. More realistically there would most likely be some psychic cost due to loss of life that offsets this effect, but to keep the analysis simpler I have simply excluded it from the model. This is in a sense conservative: I am assuming that the psychic cost is no larger than necessary to offset the gain in per-capita GDP, so that a country is indifferent about losing population to war. If we were to add a more realistic (larger) psychic cost of loss of life, this would reduce propensity for war.

propensity for war is positive, it follows that the other's propensity for war must be negative. Thus all wars in this model are unilateral: One country attacks another which would have preferred not to fight.

In principle, this means that there is some possible transfer that a country which expects to lose could make to its opponent in exchange for not fighting, which would be Pareto-improving. In practice, however, such a contract seems impossible to enforce; upon receiving the transfer, the aggressive country could simply invade anyway, and would in fact be in an even more advantageous position when doing so. In an indefinitely iterated game, such tribute might be more feasible, but it is not a panacea, as will be shown in section 2.12.

Once a war has started, the invaded country does have an incentive to continue fighting, rather than immediately surrender, because fighting at least offers some *probability* of retaining their capital stock (or even claiming some of the attacking country's). They would have preferred to trade rather than fight, but they prefer to fight rather than surrender. In the model I haven't even given countries the option of surrendering, but under fairly general conditions they would never want to take such an option anyway.

Proposition 2.4 is apparent by inspection.

Proposition 2.4. Δ_i *is decreasing in* γ *and increasing in* s *and* D_{ij}*; that is, smaller gains from trade, greater distance, or stricter sanctions will increase the incentives for war.* Δ_i *is decreasing in* τ_{ij} *and increasing in* τ_{ji}*; that is, giving tribute will increase a country's propensity for war, while receiving tribute will decrease it.*

We may wonder how increased capital translates into increased military spending. Proposition 2.5 provides the somewhat surprising answer. There are two countervailing effects: While more capital increases GDP and thus reduces the real cost of military spending, it also reduces the return on capital and makes the prize of winning less appealing.

Proposition 2.5. *Optimal military spending g_i^* is increasing in own capital K_i if and only if the following condition holds:*

$$\frac{\alpha}{K_i}R_i > \frac{\kappa_i + \kappa_j}{(K_i + \kappa_j)(K_i - \kappa_i)}$$

Optimal military spending is increasing in enemy capital K_j if and only if the following condition holds:

$$\frac{R_iY_i - R_jY_j}{R_iY_i + R_jY_j} \cdot \frac{\alpha}{K_j}R_j \geq \frac{\kappa_i + \kappa_j}{(K_j + \kappa_i)(K_j - \kappa_j)}$$

In particular, optimal military spending is decreasing in enemy capital if $R_iY_i < R_jY_j$, that is, if own will to fight is less than enemy will to fight and the war was already a probable loss.

As war becomes more destructive, we would expect it to become less appealing. Proposition 2.6 supports this intuition, and is apparent by inspection.

Proposition 2.6. *As the proportion of capital destroyed by war ω increases, the propensity for war Δ_i decreases.*

Another useful comparison is between the effects of κ_i and κ_j. If one country's territory is more contestable than the other's, which country is more likely to win? If $\kappa_i > \kappa_j$ and $K_i = K_j$, which is larger? R_i or R_j?

Proposition 2.7 provides a form of the "Life-Dinner Principle": "The rabbit runs for his life, but the fox only runs for his dinner."(Dawkins et al., 1979) The country that has more to lose in the war has a stronger incentive to fight.

Proposition 2.7. *If countries have equal capital $K_i = K_j$ and labor $L_i = L_j$, then the country with more contestable territory $\kappa_i > \kappa_j$ also has a higher return on victory $R_i > R_j$, and will invest more in military spending $g_i^* > g_j^*$ and therefore have a higher probability of victory.*

While the country with more contestable territory is more likely to win in war, they will also inherently suffer more costs, and as a result be less willing to go to war. (Similarly, the hare is more likely to prefer not to encounter the fox than vice-versa.)

Proposition 2.8. *If countries have equal capital $K_i = K_j = K$ and labor $L_i = L_j$: If $\frac{R_i}{R_i+R_j} < 1-\alpha$, then the country with more contestable territory $\kappa_i > \kappa_j$ has a lower propensity for war $\Delta_i < \Delta_j$.*

Moreover, it is always the case that:

$$\Delta_i - \Delta_j \leq \alpha \ln \left[\frac{K + \kappa_j}{K - \kappa_j} \right] < \alpha R_i$$

Proposition 2.9 shows that for realistic levels of capital intensity, two countries of equal capability will never want to go to war with one another. War requires asymmetry in strength, to justify the risk of loss.

Proposition 2.9. *War is only desirable if the two countries are not equal. If $\alpha < \frac{1}{2}$, two countries with identical characteristics will both have $\Delta_i < 0$.*

There is also a more general result:

Proposition 2.10. *If $\alpha < \frac{1}{2}$, then if $\Delta_i > 0$, then $\Delta_j < 0$. All wars are unilateral in the sense that one side would prefer not to fight.*

2.4 Dynamic Model

In real life, countries do not face one another in static one-shot games; they interact with each other over time, with some small chance of another government collapsing or being replaced by a new government before the next interaction. This is best represented as an indefinitely iterated game. I will assume that the game terminates each round with a probability $\delta \in (0, 1)$, which does not change between rounds. This actually seems quite realistic for countries, as the fall of nations is generally not predictable far in advance, and unlike individuals, countries do not have an apparent typical "lifespan"; some countries (such as Tanganyika) have formed and dissolved within a few years, while others (such as the United Kingdom) have spanned multiple centuries. The probability of a country collapsing or radically changing its policy stance may not be constant over time (and is definitely not constant across countries), but it is surely always strictly between 0 and 1, and allowing the probability to vary would complicate the model without materially affecting the conclusions.[3]

Proposition 2.9 shows that war will not be desirable if the two countries are equal. Therefore, without loss of generality, suppose that country i is the stronger country, the one that wants to fight.

We might suppose that country j can adopt a strategy of offering tribute: $\tau_{ji} > 0$. They continue to offer this tribute each round as long as country i does not attack them.

In particular, country j chooses its level of tribute τ_{ji} such that country i's expected payoff is the same whether or not a war starts, i.e. $\Delta_i = 0$. Since country j's payoff will then be higher than under war, this is a Pareto improvement.

[3] There is one possible case where time-varying termination probability could affect the results: If the probability mass function for the termination time of a country is sufficiently fat-tailed that it has no well-defined expectation value (e.g. Pareto distribution with $\alpha \leq 1$), a situation can arise where the "shadow of the future" has an infinite expected utility. Expected utility theory can behave quite strangely in such circumstances, and it is an open question how we should interpret such results (Bostrom, 2011).

The necessary level of tribute τ_{ji}^* can be found by setting Δ_i in equation 2.6 equal to zero and solving for τ_{ji}:

$$\tau_{ji}^* = Y_i \exp\left[R_i\left(\frac{R_iY_i}{R_iY_i + R_jY_j} - (1-\alpha)\right)\frac{R_iY_i}{R_iY_i + R_jY_j}\right. \tag{2.9}$$

$$\left. + \alpha \ln\left[1 - \frac{\kappa_i}{K_i}\right] + \alpha \ln(1-\omega)\right] \tag{2.10}$$

$$- Y_i\left[1 + \frac{\gamma}{D_{ij}}Y_j\right] \tag{2.11}$$

By assumption, $\Delta_i > 0$; thus the positive terms in equation 2.5 outweigh the negative terms. Hence, the exponent will be positive, and the term $\exp[\ldots]$ will be larger than 1. Note however that the term $\left[1 + \frac{\gamma}{D_{ij}}Y_j\right]$ is also larger than 1, due to gains from trade.

Note that nothing ensures $\tau_{ji} < Y_j$, particularly when $Y_i \gg Y_j$. Because countries can transfer capital by war but only GDP by tribute, there may not be an amount of tribute that country j can afford to pay which will satisfy country i.

Proposition 2.11. *In the iterated Trade-War game, if $\Delta_i > 0$, then for any given value of the other parameters, there exists a threshold $\underline{Y} > 0$ such that for any $Y_j < \underline{Y}$, there is no possible tribute $\tau_{ji} < Y_j$ that would give country i a high enough payoff to convince them to not choose "War". In such circumstances, there is no subgame-perfect equilibrium in which tribute is sufficient to prevent conflict.*

The intuition is as follows: Because country i is larger and richer than country j, their marginal utility of income is much lower than that of country j, but their marginal productivity of capital is much higher. Hence there is no amount of tribute that country j can afford to pay which will placate them; they would rather claim the capital directly rather than accept the meager tribute that country j can provide by using it themselves.

Does this mean that war in such circumstances is inevitable? Not necessarily.

Earlier I showed above that positive sanctions would not be chosen in the static game; in the iterated game, this is no longer the case. For δ sufficiently small, there is then an equilibrium generated by a "Grim Trigger" retaliation strategy in which sanctions are deployed as a threat to stop unilateral invasions; on the equilibrium path, this results in a Pareto improvement over equilibria in which war occurs.

Suppose that each country adopts the following strategy:

1. Always choose "Trade". Always set tribute to zero.
2. In the first round, set the sanction s to zero with all countries.
3. If war does not occur, set the sanction to zero with all countries.
4. If war occurs, set the sanction s to 1 with any countries that ever chose "War", and zero with all other countries.

If all countries maintain this strategy, trade will occur in every round, without sanction or tribute. The payoff from this for each country is:

$$\pi_i(\mathrm{GT}) = \pi_i(\mathrm{trade})$$

$$\pi_i(\mathrm{GT}) = \alpha \ln\left[\frac{K_i}{L_i}\right] + \ln\left[1 + \frac{\gamma}{D_{ij}} Y_j\right]$$

The most tempting deviation from this equilibrium is for country i to unilaterally attack another country $j \neq i$. If this occurs, this will be country i's expected payoff in the first period:

$$\pi_i(\text{war}, t=0) = R_i \left(\frac{R_i Y_i}{R_i Y_i + R_j Y_j} - (1-\alpha) \right) \frac{R_i Y_i}{R_i Y_i + R_j Y_j} + \alpha \ln \left[\frac{K_i - \kappa_i}{L_i} \right] + \alpha \ln(1-\omega)$$

After choosing war in the first period, country i will face sanctions in later periods. But country i may have more capital than before (if they won the war) or less (if they lost it). If they win the war, they will get a payoff of $\pi_i(\text{GT won})$; if they lose, they will get a payoff of $\pi_i(\text{GT lost})$:

$$\pi_i(\text{GT won}) = \alpha \ln \left[\frac{K_i + \kappa_j}{L_i} \right] + \alpha \ln(1-\omega)$$

$$\pi_i(\text{GT lost}) = \alpha \ln \left[\frac{K_i - \kappa_i}{L_i} \right] + \alpha \ln(1-\omega)$$

This is their expected long-run payoff:

$$\pi_i(\text{war vs. GT}) = \pi_i(\text{war}, t=0) + \frac{(1-\delta)}{\delta} \left[P(i\,\text{wins}) \pi_i(\text{GT won}) + P(j\,\text{wins}) \pi_i(\text{GT lost}) \right]$$

$$
\begin{aligned}
\pi_i(\text{war vs. GT}) = {} & R_i \left(\frac{R_i Y_i}{R_i Y_i + R_j Y_j} - (1-\alpha) \right) \frac{R_i Y_i}{R_i Y_i + R_j Y_j} \\
& + \alpha \ln \left[\frac{K_i - \kappa_i}{L_i} \right] + \frac{1}{\delta} \alpha \ln(1-\omega) \\
& + \frac{(1-\delta)}{\delta} \frac{R_i Y_i}{R_i Y_i + R_j Y_j} (\alpha R_i) \\
& + \frac{(1-\delta)}{\delta} \left(\alpha \ln \left[\frac{K_i - \kappa_i}{L_i} \right] \right)
\end{aligned}
$$

$$
\begin{aligned}
\pi_i(\text{war vs. GT}) = {} & R_i \left(\frac{R_i Y_i}{R_i Y_i + R_j Y_j} - (1-\alpha) \right) \frac{R_i Y_i}{R_i Y_i + R_j Y_j} \\
& + \frac{1}{\delta} \alpha \ln \left[\frac{K_i - \kappa_i}{L_i} \right] + \frac{1}{\delta} \alpha \ln(1-\omega) \\
& + \frac{(1-\delta)}{\delta} \frac{R_i Y_i}{R_i Y_i + R_j Y_j} (\alpha R_i)
\end{aligned}
$$

We can define a new the propensity for war Δ'_i in the dynamic game as follows:

$$
\Delta'_i = \pi_i(\text{war vs. GT}) - \pi_i(\text{GT})
$$

$$
\begin{aligned}
\Delta_i' = {} & R_i \left(\frac{R_i Y_i}{R_i Y_i + R_j Y_j} - (1-\alpha) \right) \frac{R_i Y_i}{R_i Y_i + R_j Y_j} \\
& + \frac{1}{\delta} \alpha \ln \left[\frac{K_i - \kappa_i}{L_i} \right] + \frac{1}{\delta} \alpha \ln(1-\omega) \\
& + \frac{(1-\delta)}{\delta} \frac{R_i Y_i}{R_i Y_i + R_j Y_j} (\alpha R_i) \\
& - \alpha \ln \left[\frac{K_i}{L_i} \right] - \frac{1}{\delta} \ln \left[1 + \frac{\gamma}{D_{ij}} Y_j \right]
\end{aligned}
$$

Note first that as $\delta \to 1$, $\Delta_i' \to \Delta_i$.

The $\frac{1-\delta}{\delta}$ and $\frac{1}{\delta}$ terms become larger as δ decreases; but is their net contribution positive or negative? This depends on the parameters.

This term reflects the fact that winning is better than losing:

$$
\frac{R_i Y_i}{R_i Y_i + R_j Y_j} (\alpha R_i) > 0
$$

This term reflects the fact that losing capital makes your own GDP lower:

$$
\alpha \ln \left[\frac{K_i - \kappa_i}{L_i} \right] < \alpha \ln \left[\frac{K_i}{L_i} \right]
$$

This term reflects the losses from destruction in war:

$$
\alpha \ln(1-\omega) < 0
$$

And this term reflects the loss of trade revenue from being sanctioned:

$$-\ln\left[1+\frac{\gamma}{D_{ij}}Y_j\right]<0$$

The question is then: Can the positive term dominate the others? Can the return on victory be so high that war is desirable even under the shadow of the future?

It turns out that this is possible. Under such circumstances, $\Delta_i' > \Delta_i$, and the shadow of the future actually *increases* the incentives for war. In such circumstances, sanctions may be insufficient to deter conflict.

However, if γ is sufficiently large, then we will have $\Delta_i' < \Delta_i$. In this case, Δ_i' decreases without bound as $\delta \to 0$. Thus, there is always some threshold $\underline{\delta}$ under which the Grim Trigger equilibrium can be sustained, and conflict can be prevented.

By the Folk Theorems (Fudenberg and Tirole, 1991), this equilibrium attains the highest possible total welfare under the most permissive conditions on δ, and is in that sense optimal.

By similar reasoning, there can also be a "Tit-For-Tat" equilibrium in which sanctions are only imposed for one round (or some finite number of rounds) rather than indefinitely. The permissible range for δ to sustain this latter equilibrium is smaller than the "Grim Trigger" equilibrium; however, the "Tit-For-Tat" equilibrium has the advantage that should a government appear that engages in war irrationally, sanctions will only be imposed until shortly after that government is removed, rather than on into the indefinite future. Hence, if the "Tit-For-Tat" equilibrium is feasible, it would be preferable in real life to the "Grim Trigger" equilibrium.

In many cases, a full sanction of $s = 1$ would not be necessary, and instead sanctioning countries could choose some partial sanction $s \in (0, 1)$ that is still a sufficient threat. (If we treat Δ_i' as a function of s, choose the s that makes $\Delta_i'(s) = 0$.) Like using "Tit-For-Tat" instead of "Grim Trigger", this narrows the circumstances in which peace can be sustained, but makes the outcome better if war should end up occurring.

This can be generalized to proposition 2.12.

Proposition 2.12. *In the indefinitely iterated Trade/War game, $\Delta_i' < \Delta_i$ if and only if the following holds:*

$$\ln\left[1 - \frac{\kappa_i}{K_i}\right] + \ln(1-\omega) + \frac{R_i Y_i}{R_i Y_i + R_j Y_j} R_i - \ln\left[1 + \frac{\gamma}{D_{ij}} Y_j\right] < 0$$

If $\Delta_i' < \Delta_i$, then there is a threshold $\underline{\delta}$ such that for any termination probability $\delta < \underline{\delta}$, there is at least one subgame-perfect equilibrium in which war never occurs because all countries adopt a strategy in which they would impose sanctions against countries that have gone to war. On the equilibrium path, sanctions will never actually be imposed.

There are many other possible equilibria of the indefinitely iterated game.

In particular, each country playing its dominant strategy in each round is always a Nash equilibrium of the iterated game. This means that even under conditions where a "Tit-For-Tat" equilibrium is feasible, an equilibrium where war occurs is also feasible.

The fact that there are many equilibria of this game implies that some mechanism is needed to select which equilibrium actually occurs. I propose that the mechanism is *international norms*: Major powers meet and discuss what their strategies will be, agreeing to adopt a particular strategy in the iterated game.

Such agreements can be cheap talk with no direct enforcement mechanism for lying—a common complaint by the "realist" school of international relations against international treaties (Goldsmith and Posner, 2000). Yet in fact as long as the proposed strategy sustains an equilibrium in the iterated game, and sufficiently many countries *believe* that it will be followed, that cheap talk agreement can become a self-sustaining norm.

In this sense, *both* the "liberal" *and* "realist" schools may be correct: A treaty is nothing more than cheap talk between agents that act in their own self-interest, but the correct choice of cheap talk can actually shift the entire world into a different equilibrium. It *does* matter how treaties are written and which countries sign them, because that is the equilibrium selection mechanism in the indefinitely iterated game.

In the absence of a treaty, the obvious norm is for each country to do whatever serves it; this results in frequent war. But in the presence of well-designed treaties, countries can be held back from engaging in war by the threat of sanctions from other countries.

2.5 Multilateral Sanctions

There are of course conditions under which the "Grim Trigger" unilateral sanction regime above can fail to sustain peace.

This raises the question of whether multilateral sanctions could succeed where unilateral sanctions fail.

Note first that multilateral *tribute* can be difficult to sustain. Let the set of countries that are neutral to the conflict between i and j be N.

If countries $n \in N$ other than j are strong enough that country i doesn't want to attack them, then they have no incentive to provide tribute. Their payoff would be strictly worse under a tribute regime than if country i is simply allowed to attack country j. It is not a subgame-perfect equilibrium to pay tribute to protect other countries if you do not need the protection yourself.

Therefore, to model multilateral sanctions, suppose that other countries are strong enough that country i does not want to attack them. All of these countries are trading partners for country i, so in the absence of sanctions this is country i's trade payoff:

$$\pi_i(\text{trade}) = \alpha \ln \left[\frac{K_i}{L_i}\right] + \ln \left[1 + \sum_{n \in \{j\} \cup N} \frac{\gamma(1-s)}{D_{in}} Y_n\right]$$

$$\begin{aligned}
\Delta_i' = {} & R_i \left(\frac{R_i Y_i}{R_i Y_i + R_j Y_j} - (1-\alpha)\right) \frac{R_i Y_i}{R_i Y_i + R_j Y_j} \\
& + \frac{1}{\delta} \alpha \ln \left[\frac{K_i - \kappa_i}{L_i}\right] + \frac{1}{\delta} \alpha \ln(1-\omega) \\
& + \frac{(1-\delta)}{\delta} \frac{R_i Y_i}{R_i Y_i + R_j Y_j} (\alpha R_i) \\
& - \alpha \ln \left[\frac{K_i}{L_i}\right] - \frac{1}{\delta} \ln \left[1 + \sum_{n \in \{j\} \cup N} \frac{\gamma}{D_{ij}} Y_n\right]
\end{aligned}$$

The multilateral "Grim Trigger" is the same as before: Sanctions are imposed on any countries that ever chose "War".

The condition for a "Grim Trigger" equilibrium becomes:

$$\ln \left[1 - \frac{\kappa_i}{K_i}\right] + \ln(1-\omega) + \frac{R_i Y_i}{R_i Y_i + R_j Y_j} R_i - \ln \left[1 + \sum_{n \in \{j\} \cup N} \frac{\gamma}{D_{in}} Y_n\right] < 0 \qquad (2.12)$$

Compare this to the condition for unilateral sanctions to be effective, from proposition 2.12:

$$\ln\left[1 - \frac{\kappa_i}{K_i}\right] + \ln(1-\omega) + \frac{R_i Y_i}{R_i Y_i + R_j Y_j} R_i - \ln\left[1 + \frac{\gamma}{D_{ij}} Y_j\right] < 0 \tag{2.13}$$

By construction, this holds:

$$\sum_{n \in \{j\} \cup N} \frac{\gamma}{D_{in}} Y_n \geq \frac{\gamma}{D_{ij}} Y_j$$

Thus, it is possible to have inequality 2.12 be true while inequality 2.13 is false. In such circumstances, a multilateral sanction regime can sustain peace in subgame-perfect equilibrium where a unilateral sanction regime could not.

Since on the equilibrium path sanctions are never actually imposed, there is no downside for a country to participating in the multilateral sanction regime, even if the country is sure that they have sufficient military strength to resist any attack. Thus, unlike a multilateral tribute equilibrium, a multilateral sanction equilibrium can be sustained even with many large, powerful countries.

2.6 Discussion

Many recent changes in the world may have contributed to the Long Peace, such as trade globalization, economic growth, the spread of democracy, and the invention of nuclear weapons. But this model suggests that certain aspects are crucial:

- While tribute was common in antiquity, it is quite rare today. This model suggests that greater inequality between countries (due to differences in economic growth) may explain this: Larger countries do not benefit enough from tribute from small countries to be sufficiently deterred.

- While in the past it was common for countries to engage in warfare without punishment, now it is typical for the international community to impose sanctions against countries countries that engage in aggressive war.

- Whereas territory claimed through conquest was once treated as legitimate, now the international community typically does not recognize territory claimed through aggression.

- Due to more integrated global trade networks, more advanced civilian and military technology, and greater capital accumulation in general, the costs of military conflict in terms of lost trade opportunities and damage to capital stocks are higher than ever.

Most of these changes are not changes in technology or material circumstances, but changes in *norms*. While the invention of new military technologies (and particularly nuclear weapons) is relevant, it is not the sole cause of the Long Peace.

The reason norms are so important is that nations in conflict are not engaged in a one-shot static game; they are engaged in an indefinitely iterated game. This is important because indefinitely iterated games have a vast number of possible equilibria, which means that some sort of equilibrium selection mechanism is necessary. International norms can provide this selection mechanism.

This raises another issue however: How are norms themselves chosen? For every equilibrium in the vast space of possibilities, there is a corresponding norm that would generate that equilibrium. Much as in a bargaining game, there is reason to think that a process of negotiation would likely lead toward norms that are Pareto-efficient, but this is by no means guaranteed. There are almost always Pareto-inefficient equilibria and norms that can sustain such equilibria. In some cases there can be multiple Pareto-efficient equilibria (for example, in this model, there are cases when either a sanction equilibrium or a tribute equilibrium is sustainable and both are Pareto-efficient), and then the outcome of negotiation becomes even less clear. Further research, perhaps using evolutionary game theory, may be able to better explain why peaceful norms took so long to emerge, why they emerged when they did, and whether they will continue to be stable in the future. Understanding this process could have vital policy implications for future leaders and diplomats hoping to sustain the Long Peace into the twenty-first century.

In this model sanctions are used as a threat to sustain peace, but sanctions are never actually imposed on the equilibrium path. The real world is not so simple: While sanctions are rarer than free trade, they are nonetheless imposed with some regularity. Moreover, when sanctions are actually imposed, they are often ineffective (Lektzian and Biglaiser, 2013; Bapat et al., 2013; Pape, 1997; Morgan and Schwebach, 1997; Vries et al., 2014; Zhou and Cuyvers, 2011). Extensions of the model to imperfect information may be able to explain this: Upon facing a threat of sanctions, nation might try to "call the bluff" if they think that other nations will back down—only to find that sanctions are in fact imposed. Such a model may also be able to explain the apparent ineffectiveness of sanctions based on a self-selection effect: Sanctions are only ever imposed on the countries that are most stubborn and least likely to respond to sanctions. However, the details of this more sophisticated model are outside the scope of this paper.

2.7 Conclusion

By analyzing interstate conflict as an indefinitely iterated game, this economically micro-founded model provides some insights into the possible causes of the Long Peace. By the folk theorems, indefinitely iterated games have a vast number of possible equilibria. International norms provide a mechanism for selecting particular equilibria, and good norms can result in peaceful, Pareto-efficient outcomes. The threat of sanctions can be used in a "Grim Trigger" or "Tit for Tat" fashion to deter warfare.

This provides a partial synthesis between the "realist" and "liberal" schools of international relations: The realists are correct that international treaties and norms are nothing more than cheap talk, while the liberals are correct that these treaties and norms can have real effects on outcomes if they are mutually believed and adhered to. Norms cannot create an equilibrium where none existed, but they can shift behavior from one equilibrium to another and sustain that equilibrium once it is established.

Future research may be able to gain insight into how such international norms are formed and sustained, and why they have changed so radically over historical time. The implications of such knowledge are profound: In a world of nuclear proliferation, understanding how to sustain peace could decide whether human civilization survives into the next century.

Chapter 3

Imperfect Tacit Collusion and Asymmetric Price Transmission

Abstract

We investigate the role tacit collusion plays in Asymmetric Price Transmission (APT), the tendency of prices to respond more rapidly to positive than to negative cost shocks. Using a laboratory experiment that isolates the effects of tacit collusion, we observe APT pricing behavior in markets with 3, 4, 6, and 10 sellers, but not in duopolies. Furthermore, we find that sellers accurately forecast others' prices, but nevertheless consistently set their own prices above the profit-maximizing response, particularly in the periods immediately following negative shocks. Overall, our findings support theories in which tacit collusion plays a central role in APT.

The phenomenon of Asymmetric Price Transmission (APT), that is, that supplier prices rise quickly after positive input cost shocks, but fall relatively more slowly after similarly-sized negative cost shocks, has been repeatedly documented in the literature such that we can rightly describe it as a stylized fact.[1] However, while empirical evidence for the APT phenomenon is ample, identification of its causal forces is not settled. Many theoretical explanations have been proposed, but the empirical literature has yet to conclusively determine which of these are valid or are most influential.

Empirical studies of APT predominantly examine aggregate-level variables (e.g. inflation, concentration) proposed to be relevant in the theory literature. The focus on such variables occurs because firm-level determinants are either not directly observable, or are not adequately measurable in panel data form. This approach yields helpful correlations between such variables, but the effort to identify causal relationships has met with only limited success, most notably in the context of firm-level underpinnings of the phenomenon. While the search for accurate firm-level data should certainly be continued, and where discovered used to further inform our understanding of pricing behavior, experimental methods offer a comparative advantage: testing theories that involve variables which are unobservable in the field (e.g. agents' information sets) lie outside the reach of empirical methods;[2] if however these same variables can be controlled through experimental design, we can overcome this obstacle to testing theory.

A question of primary interest is whether tacit collusion drives APT-like pricing behavior.[3] The field data does not convincingly exclude the possibility that market competitors secretly communicate, given the strong legal and even criminal incentives for firms to conceal – or

[1] See Section 3.1.1 for an overview of the evidence.

[2] Meyer and von Cramon-Taubadel (2004) and Frey and Manera (2007) provide extensive discussions of methodological issues in econometric tests of APT.

[3] In this paper we will use the term "tacit collusion" to mean the phenomenon in which suppliers coordinate on prices above the competitive equilibrium level, through the channel of publicly visible pricing alone. Tacit collusion can also take the form of coordination on quantities below competitive equilibrium levels, but in this paper we will focus strictly on the role of coordination on prices.

avoid engaging in – such activities. This provides an obvious challenge for identification and motivates turning to the controlled setting of the laboratory, where we can directly observe competitor behavior and credibly prevent communication between sellers.[4]

An argument put forth by Borenstein et al. (1997) is that a variation of the "trigger price" model of oligopolistic coordination, originally introduced by Green and Porter (1984), may explain the emergence of APT-type dynamics through tacit collusion. In their model, when positive shocks occur firms immediately raise prices in order to preserve profit margins; however, when negative shocks occur firms react adaptively, holding prices at pre-shock levels until they see convincing evidence that a rival has cut their prices. Rapidly lowering prices in response to a downward cost shock could be perceived as defection from a mutually beneficial regime of tacit collusion, thus inviting retaliation from other firms. In contrast, rapidly raising prices in response to an upward cost shock poses no such threat to one's competitors, and therefore incurs no corresponding risk of retaliation. Although their arguments are sound, and consistent with a deep empirical literature finding correlations suggestive of tacit collusion, Borenstein et al. (1997) conclude that they are unable to conclusively draw support for this hypothesis from their data. As no other empirical study of which we are aware has accomplished this either, we thus find motivation to turn to the laboratory to examine the role of tacit collusion in driving APT dynamics.[5]

A second question of interest is whether the number of competing sellers in a given market plays a significant role in the realization of the APT phenomenon. Notably, in his broad study of U.S. wholesale and retail markets, Peltzman (2000) finds a negative relation between the number of competitors in a market and the magnitude of APT observed. As with

[4]Furthermore, the laboratory may be the only environment in which we can reliably detect collusion, since the non-collusive prices or profits are unavailable without imposition of strong structural assumptions.

[5]There are some studies that regress the estimated asymmetry with measures of market concentration as Loy et al. (2016). Counter-intuitively, the authors find that asymmetry decreases with higher concentration in German milk market. However, it is difficult to associate this estimate with the causal impact of collusion on APT as higher concentration index may stem from higher efficiency or product differentiation rather than conduct.

any empirical study, however, this study does not exclude the possibility that explicit (but unobserved) communication between firms lies behind this result. Several (non-APT focused) studies of experimental oligopoly markets find that there is an inverse relation between the number of sellers in a market and the size of deviations from the Nash equilibrium (NE) outcomes (for example, see Huck et al. (2004), Dufwenberg and Gneezy (2000), and Fonseca and Normann (2012)). However, we are unaware of any experimental study that specifically studies the role of the number of sellers in driving the APT phenomenon. We therefore incorporate the number of sellers in our markets as a treatment variable in our experimental design.

To our knowledge, Bayer and Ke (2018) is the only experimental study that directly targets the topic of APT. The authors' study employs a Bertrand duopoly setting in which sellers' costs either increase, decrease or stay constant at the halfway point of the experiment. With two extensions of this baseline condition, they further test the impacts of search costs and asymmetric information on APT. They find APT across all treatments, even in the absence of search and information frictions. They argue that the asymmetry can be explained with a backward-looking learning model: If a seller fails (manages) to sell the good in the period prior to the shock, it is more (less) likely that she will adjust her price downwards (upwards) in the following period. The authors' results support this regularity when the shock is negative, but not when it is positive. Hence, although this learning model may account for the downward rigidity, it falls short of explaining the asymmetry.[6]

While Bayer and Ke (2018)'s study provides a useful benchmark to our own, our design choices differ substantially from theirs, as we pursue different research questions. Whereas we aim to assess the roles of cooperative behavior and tacit collusion on pricing asymmetries,

[6] Bayer and Ke (2018) also reason that following positive cost shocks sellers will reason that other sellers will all immediately raise their prices, and so they do the same, while following a negative cost shock sellers do not see any reason to cut their prices unless and until they subsequently lose sales. They cite factors such as bounded rationality as explanations for this behavior, but do not offer a more precise explanation of the channels through which the observed behavior emerges.

they deliberately try to attenuate their impacts to isolate the role of learning.[7] In particular, in their experiment sellers whose stores are not visited by a buyer receive only limited information on the market price, due to the feedback structure. In our experiment, we inform sellers of the average market price of the other sellers, as we want to create the conditions in which price signalling can be studied more explicitly.

In our experimental setting, subjects play the role of sellers and a computer plays the role of buyers. Each seller faces demand that linearly decreases with one's own price and linearly increases with the average price of others. We vary the size of groups across sessions as 2, 3, 4, 6, and 10, while calibrating the demand function to hold the best-response functions of each seller identical, across all group sizes. This approach allows us to isolate and study the impact of group size on the realization of APT through the coordination channel. Throughout our experiment, sellers experience a series of input price shocks – either large or small – that shift the NE price either up or down. Through this design, we are able (i) to test whether APT emerges despite the absence of market frictions and information asymmetries that are often theorized to be the causal forces behind pricing asymmetries; and, (ii) if APT does occur, to assess the impact of number of sellers on the magnitude of the resulting asymmetries. To our knowledge, ours is the first experiment that study the role of number of sellers in shaping APT.

Our contributions to the literature are two-fold: First, we document the prevalence of the APT phenomenon through experiments in which we possess strict control over the environment. In particular, our results indicate that the APT may emerge even in the absence of market frictions and information asymmetries that are often theorized to be the causes of

[7] Although Bayer and Ke (2018) exert effort to minimize the role played by tacit collusion with their study, their typed-stranger matching protocol significantly reduces but does not completely eliminate the possibility that subjects might repeatedly interact, and thus have the opportunity to establish reputation over time. By contrast, the perfect-stranger matching protocol, in which a subject is assured they will be matched with another only once in a session, does eliminate this possibility. Moreover, the duopoly setting of their study makes collusion presumably more reachable, since coordination is easier when there is only one other market participant. As a result, it is hard to assess the extent of the role to which cooperative behavior played in their study.

pricing asymmetries. This suggests that in markets with three or more sellers, the presence of agents who attempt to coordinate on prices via price signaling may suffice for APT pricing dynamics to emerge. In our duopoly markets, however, our results suggest that coordination on prices can be so successful that rather than the APT phenomenon, persistent pricing at near-monopoly levels may instead emerge. Second, by calibrating demand based on the number of sellers in a *ceteris paribus* manner, we are able to isolate and perform hypothesis tests on the effects of increased group size on APT. For markets with three or more sellers, we find no significant difference in either the magnitude of observed APT, or the rate at which post-shock price behavior converges to NE-implied prices. Together, the results of our study support theories that highlight the role of tacit collusion on APT. We conclude that APT may be the product environments in which collusion is significant, but imperfect.

3.1 Related Literature

3.1.1 Field Evidence

Bacon (1991) provides an early empirical study suggesting that retail gasoline prices in the United Kingdom experience faster and more concentrated responses to crude oil price increases, than they do to similar crude oil price decreases. Bacon termed this phenomenon "Rockets and Feathers," and since this paper was published dozens of other researchers have detected the presence of this sort of asymmetry in a variety of consumer and intermediate goods markets.

Peltzman (2000) provides one of the most comprehensive empirical examinations of APT. He conducts a broad study of pricing behavior of 77 consumer and 165 producer goods markets in the U.S., and he concludes that in more than two-thirds of these markets prices rise faster than they fall, in response to input cost changes. Peltzman also seeks correlations between various features of markets and industries, and the degree to which evidence of APT is present. Most notably, he finds that markets with fewer competitors tend to exhibit more pricing asymmetry, while on the other hand markets with higher levels of concentration tend to be less likely to exhibit pricing asymmetry, as in Loy et al. (2016). Peltzman's study, however, does not provide an explanation for these correlations.

In an early survey of field evidence, Meyer and von Cramon-Taubadel (2004) find that (excluding Peltzman (2000)'s study), symmetry in price response is rejected in almost one-half of all cases in the literature. Their survey also shows that different test methods yield highly varying rejection rates (between 6% and 80%). Frey and Manera (2007) and Perdiguero-García (2013) provide meta-regression analyses with more comprehensive and recent data sets. Both studies confirm that APT is very likely to occur but also emphasize the variation of reported outcomes. Their results show that this heterogeneity can be explained with several factors as characteristics of data (e.g., data frequency) and of the employed econometric model. Most notably, Perdiguero-García (2013) reports that the asymmetry tends to decrease in more competitive segments of the industry.

3.1.2 Theoretical Explanations

There is a growing body of literature on the theoretical accounts of APT, an unsurprising fact given that pricing asymmetries are not predicted by standard price competition models.[8] These studies propose explanations of APT mainly by introducing market frictions,

[8] A notable exception is the case of Markov-perfect equilibria, and in particular the case of the Edgeworth cycle. In this phenomenon, firms undercut each others' prices successively until prices approach marginal cost; at this point, one of the firms decides with some positive probability to spike its price, and once this occurs the cycle is repeated, yielding each firm positive economic profits. Maskin and Tirole (1988) further show that these cycles provide a case where asymmetric pricing can be sustained in equilibrium. However, the Edgeworth cycle model requires that firms make price decisions alternately; the model does not support an equilibrium when price decisions are made simultaneously or continuously. Moreover, the emergence of the phenomenon seems in practice to be limited to environments in which competitors rapidly and publicly change prices (see for example Byrne and De Roos (2019) for an interesting case in Perth, Australia petrol markets, in which a government mandate for retail suppliers to publish their prices daily seems to have

information asymmetries or boundedly rational agents into the underlying models. One reason there is such a variety in the way different studies explain the APT is because these studies typically focus on specific market structures (e.g., wholesale petroleum markets) and their idiosyncrasies. In this subsection, we review some of these studies in an attempt to categorize as well as to highlight discrepancies.[9]

Borenstein et al. (1997) consider the role of search costs in facilitating APT. They hypothesize that negative cost shocks in the presence of costly search provide firms temporary pricing power, which they then use to delay reductions in prices, yielding temporarily superior profits. Benabou and Gertner (1993) and Yang and Ye (2008) also develop explanations based on consumer search costs, but also on the volatility of input costs. They reason that volatility should reduce search incentives for consumers; producers, realizing that short-term demand elasticity is increased as a result, thereby yielding them temporarily increased pricing power, respond by reducing prices more slowly. Reagan and Weitzman (1982) and Borenstein and Shepard (1996) propose explanations based on inventory costs, reasoning that it is relatively more costly for manufacturers and suppliers facing capacity constraints or sharply rising short-term production costs to deal with unanticipated increases in demand resulting from price drops, than it is to respond to corresponding drops in demand due to price increases. Ball and Mankiw (1994) consider a menu-cost model in conjunction with positive trend inflation as an explanation of APT. In another study, Ahrens et al. (2017) show that the presence of consumers with loss aversion may explain why prices are more sluggish to adjust downwards than upwards in response to permanent demand shocks.

The various explanations and models described above provide differing implications for government policy: if APT occurs due to collusion, there may be room for regulation to improve economic efficiency; if however APT is primarily caused by the presence of inventory

facilitated the emergence of a weekly cycle of Edgeworth-like pricing dynamics that persisted for many years.). The Edgeworth cycle model therefore applies to a relatively narrow range of market contexts.

[9]For more exhaustive surveys of theoretical explanations, see Meyer and von Cramon-Taubadel (2004) and Brown and Yucel (2000).

costs, asymmetric menu costs, or search costs, then regulation that controls pricing behavior may actually induce inefficiency rather than attenuate it. Given the robust evidence of the widespread existence of APT and its non-trivial magnitude and impact on consumer outcomes, identifying which theories describe the asymmetric pricing behavior is key to informing effective public policy.

3.1.3 APT and Experiments

Despite the many possible explanations that have been proposed, the empirical literature yields only mixed evidence that is often inconclusive due to identification issues. This suggests there is room for further research to shed light on the phenomenon. We consider the advantages of experimental methods in isolating and studying causal determinants of APT.[10] In this subsection, we summarize the most relevant literature to our study.

There are two studies of which we are aware – in addition to Bayer and Ke (2018) – that conduct market experiments with APT-related results. Deck and Wilson (2008) investigate gasoline markets and find that retail prices adjust asymmetrically to changes in station costs in zones with clustered stations, but not in zones with stations that are relatively isolated from competitors. Cason and Friedman (2002) find weak evidence of APT in posted offer markets where customers incur switching costs. While these studies examine their findings on APT, their experimental designs are optimized to investigate questions regarding the structure of gasoline markets (e.g., zone pricing, divorcement) and of consumer markets (e.g., switching costs), not to identify causes of APT. In particular, sellers' costs in both experiments follow random-walk shocks, which may not be salient enough to detect APT.

[10] The usage of experimental methods in macroeconomic research is becoming more and more prevalent. See Duffy (2016) and Cornand and Heinemann (2019) for recent surveys.

Our study distinguishes itself from this string of literature by examining APT with larger, persistent shocks.[11]

Apart from studies that directly target APT, price competition experiments that study the impact of group size on tacit collusion are also relevant to the current paper. Dufwenberg and Gneezy (2000) provide an early evidence for such a relation through an oligopoly game that corresponds to a discrete version of the Bertrand model. They find that winning prices tend to converge to NE levels in groups of three or four competitors, but stay consistently high in duopolies. Morgan et al. (2006) find that increasing the number of sellers from 2 to 4 decreases the prices paid by some consumers (the ones informed about the entire distribution of prices) but not for others (the ones who buy with motives other than prices). Abbink and Brandts (2008) also find that there is a negative relationship between the number of competing firms and price levels.[12] Nevertheless, as in Dufwenberg and Gneezy (2000), they find that collusive pricing is the modal outcome in duopolies. Fonseca and Normann (2012), Orzen (2008), Davis (2009) and Horstmann et al. (2018) provide further evidence that collusive prices are very likely to be observed in duopolies. Average prices approach considerably close to the NE in the baseline condition of these studies (fixed matching, no communication, symmetric sellers etc.) when the number of sellers is 3 or greater.

The main conclusion of these studies is that persistent coordination over collusive prices is unlikely in markets other than duopolies. This, however, does not preclude the possibility that players might manage to coordinate temporarily on high prices, following negative shocks. Experiments also indicate that increasing the number of sellers often leads to more competitive outcomes (in terms of price and output), which in turn should make APT less

[11] Fehr and Tyran (2001) also employ large positive and negative shocks and report APT-like behavior in a price-setting game. However, the authors do not analyze the phenomenon, nor do they probe its implications. In another related experimental study, Duersch and Eife (2019) consider Bertrand duopolies with zero marginal cost in either inflationary, deflationary or constant price environments. They find that real prices are significantly lower in the inflationary environment compared to non-inflationary environments.

[12] Their results are particularly interesting since in their price competition setting, there exist multiple equilibria.

likely. Although, the meta-analyses of Fiala and Suetens (2017) and Horstmann et al. (2018) on oligopoly experiments indicate that there may not be a linear relationship between the number of competing firms and the degree of tacit collusion. Horstmann et al. (2018) argue that this result may stem from the relatively small number of studies that provide pairwise comparisons and the lack of statistical power in these studies. Our study contributes to the literature through improvements of these axis.

3.2 Method

3.2.1 Pricing Game

We develop a variant of classical price competition related to the "Linear city" model of Hotelling (1929) and the "Circular city" model of Salop (1979), and employ this in our experimental environment. In this setting, the demand facing seller $i \in N$ in period $t \in T$ is equal to

$$q_{i,t}(p_{i,t}, p_{-i,t}; \delta, \gamma) = \begin{cases} \delta - \gamma(p_{i,t} - p_{-i,t}), & p_{i,t} \in [p^{min}, p^{max}] \\ 0, & \text{otherwise} \end{cases} \tag{3.1}$$

where δ and γ are parameters of demand, $p_{i,t}$ is the price set by seller i and $p_{-i,t}$ is the average price chosen by the rival sellers in the same market (i.e. $p_{-i,t} \equiv \frac{1}{N-1}\sum_{j \neq i}^{N-1} p_{j,t}$) at period t. p^{min} represents the price floor and p^{max} is the representative consumer's valuation of the good.[13,14] Given the own-demand specification in (3.1), seller profits are calculated as

$$\pi_{i.t} = (p_{i,t} - mc_t)q_{i,t} - f, \tag{3.2}$$

where $q_{i,t}$ is quantity demanded from seller i as defined in (3.1), mc_t is marginal cost that shifts every $\overline{T}$ periods that comprise a round (denoted $r \in R$) and f is fixed cost. Sellers set their prices in each period simultaneously from a discrete set that is bounded as $p_{i,t} \in [mc_t, p^{max}]$, such that the price floor is equal to the marginal cost of that round.

In the described game, there is a unique symmetric stage-game NE which can be retrieved from the first-order condition of the profit-maximization problem. This NE corresponds to the unique subgame-perfect Nash equilibrium (SPNE) of the finitely-repeated game by backward induction. In this NE, all sellers set their prices equal to

$$p_t^{NE} = mc_t + \frac{\delta}{\gamma}, \tag{3.3}$$

with each seller achieving (current-period) profits of $\pi^{NE} = \frac{\delta^2}{\gamma} - f$. Sellers may achieve the joint profit maximum (JPM) if they each set their prices to the maximum price p^{max}. In addition to the NE price, the maximum price constitutes a second likely focal point upon which sellers may attempt to coordinate.

[13]This reduced form demand function can be represented by either Hotelling's or Salop's address models for $N = 2$, and by Salop's model for $N = 3$, by setting $\delta = \frac{L}{N}$ and $\gamma = \frac{L}{c}\frac{(N-1)}{N}$. Here, L corresponds to the number consumers that are equidistantly located along a linear (circular) city, and c is their travel cost. For $N \geq 4$ we depart from Salop's model in two ways. First, we do not assume that sellers directly compete only with the two adjacent sellers along the circle. We instead assume symmetric cross-price demand elasticities for all market participants, regardless of their "address". Second, we calibrate the values of parameters L and c by group size N in such a way that δ and γ remain constant for all markets. The former makes the strategic space easier for subjects to understand and the latter allows us to make *ceteris paribus* comparison across treatment groups.

[14]Note that a linear specification of demand is the direct result of an assumption of quadratic utility (see Amir et al. (2017) for a proof).

The best-response function derived from the first-order condition can be expressed as:

$$p_{i,t}^{BR} = \frac{1}{2}\left(mc_t + \frac{\delta}{\gamma} + \underset{i,t-1}{\mathbb{E}}[p_{-i,t}]\right) \tag{3.4}$$

where $p_{i,t}^{BR}$ and $\mathbb{E}_{i,t-1}[p_{-i,t}]$ represent the best-response action and the conditional expectation of seller i for the average price of others, respectively.

In this pricing game, neither own-demand nor own-profit depend on the number of sellers. These only depend on own-price and the average price of rival sellers. The best-response action is also independent of N for a wide range of expectation models, including rational expectations. This feature assures that the incentives given to the sellers of different group sizes are matched and the market power of each seller is *ex-ante* equal. We consider this as necessary for ensuring a *ceteris paribus* comparison between the treatment conditions.

3.2.2 Experimental Design

Sellers interact repeatedly in the described pricing game for R rounds, which are each composed of $\overline{T}$ periods. Marginal cost mc_t fluctuates at the beginning of each round, modeling large exogenous cost shocks, but remains invariant throughout the remainder of each round. Our experimental manipulations consist of varying the size of markets across sessions in a between-subjects design, and of varying the size and direction of shocks across rounds in a within-subjects design. We implement a fixed-matching protocol during a session.

The calibration of the experimental game is summarized in Table 3.1. The experiment consists of 5 rounds of 15 periods each, with a new marginal cost announced at the beginning of each round. The sequence of shocks is identical across all treatments: Marginal cost starts at \$0.90 in Round 1, drops to \$0.50 in Round 2, rises to \$1.30 in Round 3, falls again to \$0.50 in Round 4, then rises to \$0.90 for Round 5.

Table 3.1: Experimental design parameters

General parameters	
Number of periods per round	$\overline{T} = 15$
Number of rounds per session	$R = 5$
Demand parameters	$\delta = 8.50$, $\gamma = 7.275$
Fixed cost	$f = 1$
Maximal/reservation price	$p^{max} = 3$
Varying parameters	
Group size across treatments	$N \in \{2, 3, 4, 6, 10\}$
Marginal cost across rounds	$mc : (0.90, 0.50, 1.30, 0.50, 0.90)$
Cost shock sequence	$\Delta mc \equiv \eta : (-0.40, +0.80, -0.80, +0.40)$
NE price across rounds	$p^{NE} : (2.07, 1.67, 2.47, 1.67, 2.07)$

3.2.3 Procedures

Experimental sessions were conducted at the University of California, Santa Barbara's Experimental and Behavioral Economics Laboratory (EBEL) using the z-Tree platform (Fischbacher, 2007), between September and December of 2018. A total of 245 subjects were recruited from the experimental economics subject pool of the same univerity, using the ORSEE tool (Greiner, 2015). Subjects were allocated to markets of size 2, 3, 4, 6 and 10, with a total of 36, 39, 52, 48 and 70 subjects assigned to each group size condition, respectively. This setup yields 59 independent markets for the analysis.[15]

At the beginning of each experiment, subjects are provided written instructions which are also read to them aloud by an experimenter. Subjects then proceed to take a short comprehension quiz.[16] In the main part of the experiment, each subject plays the role of sellers and makes a series of 75 pricing decisions, whereas consumer behavior is simulated by computer. We also elicit subjects' one-period-ahead expectations about the average price chosen by rival sellers (i.e., $\mathbb{E}_{t-1}[p_{-i,t}]$). These expectations are not rewarded separately, to avoid creating hedging issues. Subjects are able to set a price between the marginal cost and the maximum

[15]In one session (20 subjects), the data from the final period (of 60 periods) is lost due to technical reasons. All the analysis in the results section is performed based on all the available data.

[16]We reviewed answers for each subject and provided explanations where needed. See Online Appendix I for all the experimental material.

price (of \$3.00), in increments of \$0.01. Once all subjects set their prices and expectations, they are individually notified by the computer of the average price established by the others in their market, reminded of their own price, and shown their own resulting payoff for that period. Subjects are able to track the previous values of these outcomes through a history box that is available in their screen (see Online Appendix I.D).

We notify subjects that a new cost shock will occur at the beginning of each new round, either an increase or decrease, of either \$0.40 or \$0.80. We reveal the magnitude and direction of each shock immediately prior to the first period of each respective round. At that time, we also hand out copies of a printed payoff table corresponding to the new marginal cost. These tables assist subjects in estimating the profits they will receive, conditional on the hypothetical prices they and others may set in each period of that round (see Online Appendix I.C).

Sessions lasted a total of 90 to 125 minutes. Subjects were paid \$18.66 on average (a minimum of \$10.89 and a maximum of \$28.50), which includes the \$5.00 show-up fee and \$3.00 for the completion of the optional survey (no subject declined this offer). The remaining payoff is determined as the average payoff of a randomly chosen round of the game.

3.3 Hypothesis

This experimental setup allows us to test the following hypotheses:

Hypothesis 1: *Prices respond symmetrically to (equally sized) positive and negative shocks.*

As the experimental design specifically avoids any of the features outlined in Section 3.1.2 (e.g., frictions, information asymmetries), theory suggests prices should react symmetrically. This can further be verified by reviewing the linearity of best-response function with respect to its arguments. In a directed counter-hypothesis, we predict prices to react asymmetrically to shocks. In particular, we expect downward rigidity in line with the arguments of Borenstein et al. (1997). We can test this hypothesis by exploiting the exogenous within-subjects treatment variations in marginal cost.

Our second hypothesis concerns tacit collusion and coordination:

Hypothesis 2: *Sellers' market power is invariant to the number of sellers in their market, and is unaffected by the existence of periodic shocks.*

Market power, that is the ability of sellers to raise prices over marginal cost, should be invariant to the number of rival sellers, given that we have calibrated demand in such a way that both the profit and best-response functions are independent of group size. Moreover, in the absence of frictions and the ability of competitors to communicate, the theory predicts a constant markup for all levels of marginal cost. However, if tacit collusion occurs, we expect to observe higher market power (i) in smaller markets, and (ii) in the periods occurring soon after shocks. For (i), we expect to observe effective coordination more often in smaller markets, where there are fewer sellers to dampen the strength of price signals. For (ii), we expect that shocks may boost the market power of sellers (at least temporarily), as such shocks may play the role of a coordination device. We can test this hypothesis by using the between-subjects treatment variations in group size, and within-subject treatment variations in marginal cost.

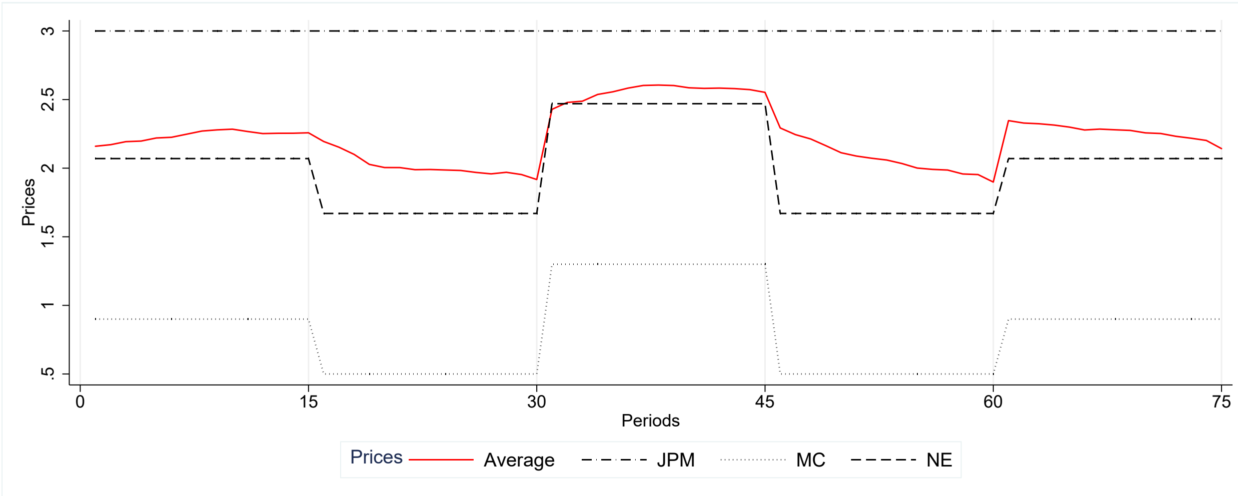
Finally, our third hypothesis concerns individual pricing strategies:

Hypothesis 3: *Conditional on expectations, pricing behavior follows the best-response function.*

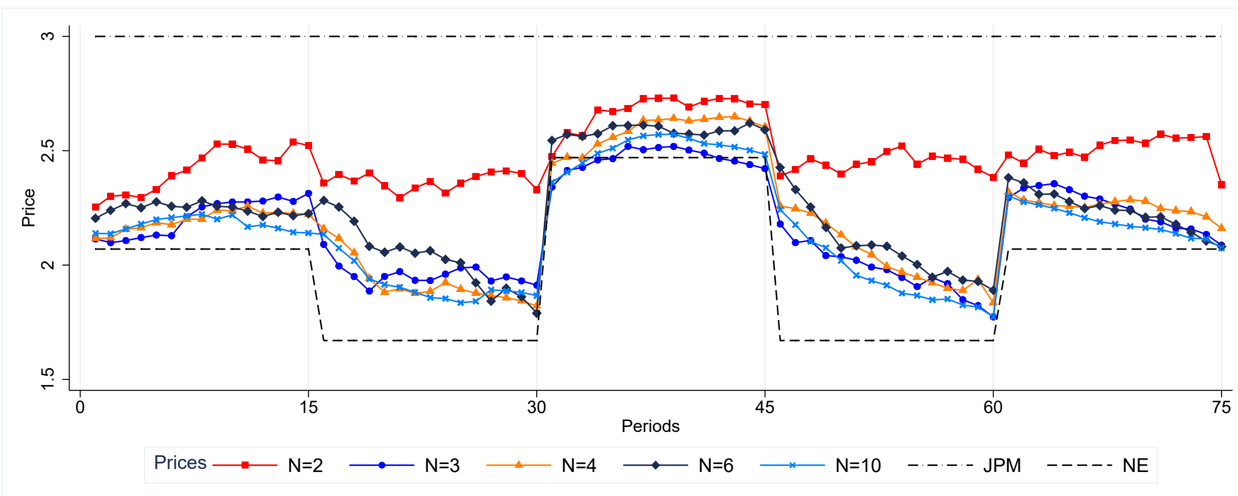
This hypothesis is built on the following rationale: The Rational Expectations Hypothesis (REH) of Muth (1961) admits the possibility of expectation errors at the individual level, but which should tend to cancel out in aggregate. Also, after observing $t-1$ periods of price history, a seller may learn that the others do not behave consistently with the predictions of REH. Nevertheless, conditional on expectations, sellers should select the best-response action as this maximizes their profit. As we elicit subjects' guesses on the average price set by others, we can test this hypothesis without assuming a specific expectation model.

3.4 Results

Figure 3.1 provides a depiction of the average price per period, as the average of all market prices and as broken out by group size. Here, market price refers to the average of all prices in market m (i.e., $p_{m,t} = \frac{1}{N}\sum_{i=1}^{N} p_{i,t}$). The reader can readily discern that for groups of size 3 and greater, average prices rise rapidly after positive cost shocks, while they fall more slowly after negative cost shocks. By contrast, for groups of size 2, it is not immediately obvious whether average pricing behavior is affected by cost shocks. A second observation that is immediately clear is that average prices are generally above the NE price, with deviations being higher following negative shocks compared to the positive ones. Overall, the visual inspection of the data suggests the presence of market behavior consistent with APT.



(a) Average of all market prices



(b) Average market prices by group size

Figure 3.1: Average pricing behavior across periods and group sizes.

3.4.1 Estimation of Asymmetry

We follow Peltzman (2000) and estimate the coefficients of the distributed lag model (DLM) to measure the magnitude of APT. This model can be expressed as:

$$\Delta p_{i,t} = \sum_{k=0}^{K} b_{t-k} \cdot \Delta mc_{t-k} + \sum_{k=0}^{K} c_{t-k} \cdot (\mathbb{1}[\Delta mc_{t-k} > 0] \cdot \Delta mc_{t-k}) + \epsilon_{i,t} \tag{3.5}$$

where the change in output price (i.e., $\Delta p_{i,t} = p_{i,t} - p_{i,t-1}$) is modelled as a function of the lagged changes in marginal cost (i.e., Δmc_{t-k}). The indicator variable $\mathbb{1}[\Delta mc_{t-k} > 0]$ takes the value 1 if the change in marginal cost in period $t-k$ is positive and equal to 0 otherwise. The sum of interaction coefficients $\sum_{k=0}^{K} c_{t-k}$ reflects the magnitude of asymmetry and its persistence over K periods.

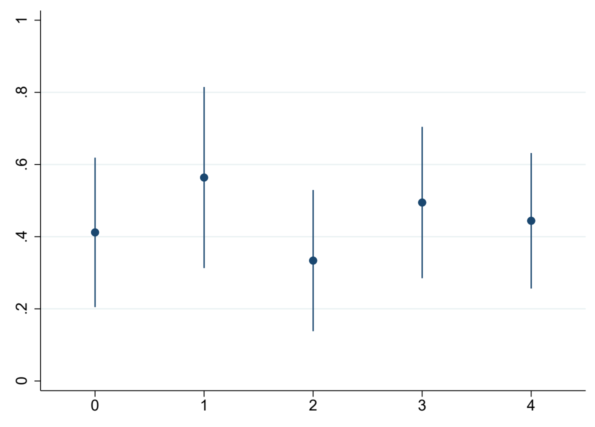


Figure 3.2: Cumulative response after K periods. Dots refer to $\sum_{k=0}^{K} c_{t-k}$. Lines represent 95% confidence intervals.

We estimate model (3.5) with Ordinary Least Squares (OLS) regressions in a step-wise manner. Figure 3.2 reports the estimated asymmetry for $K = 4$.[17] Estimates indicate that the APT is both strong and persistent. Immediate price reactions are 32.9 cents greater in magnitude for positive than for negative 80-cent shocks.

We now assess the reaction of prices to equally sized shocks between our treatment groups with non-parametric tests. We compare immediate pass-through rates of shocks that are β_0^+

[17] We report the full set of results in Online Appendix II.A. All estimations employ robust standard errors that are clustered at market level. We also include a set of indicator variables that are specific to each group size (i.e., $\mathbb{1}[N = s]$), the lagged change in the average price of rival sellers (i.e., $\Delta p_{-i,t-1}$), a three-way interaction term (between $\mathbb{1}[\Delta mc_{t-k} > 0]$, Δmc_{t-k} and group size specific indicator variables) and autoregressive terms amongst the set of regressors to check the robustness of estimates. The significance of asymmetry coefficients as well as their magnitude are robust to the inclusion of these variables.

and β_0^- calculated as:

$$\begin{aligned} p_{i,t+\tau}^+ &= p_{i,t-1}^+ + \beta_\tau^+ \eta^+ \\ p_{i,t+\tau}^- &= p_{i,t-1}^- + \beta_\tau^- \eta^- \end{aligned} \tag{3.6}$$

where η^+ (η^-) reflects either the large or small positive (negative) shock and $t-1$ corresponds to the period just before the shock. Note that the demand function we describe in (3.1) is perfectly inelastic, so that shifts in the NE price following cost shocks are exactly equal to the magnitude of the cost shock itself. In accordance, we can denote the cases $\beta_0^+ = 1$ and/or $\beta_0^- = 1$ as incidences of "full pass-through" of cost shocks. The ratio of β_0^+ and β_0^- thus conveys information on the degree of APT in immediate cost-shock responses. A ratio of 1 would indicate the absence of APT.

Table 3.2: Asymmetry in the immediate pass-through rates

	All	$N>2$	$N=2$	$N=3$	$N=4$	$N=6$	$N=10$
Small shocks							
β_0^-	0.159	0.115	0.411	0.558	0.158	-0.144	0.0154
	(0.0676)	(0.0739)	(0.162)	(0.247)	(0.138)	(0.130)	(0.0985)
β_0^+	1.119	1.270	0.244	1.305	1.209	1.233	1.322
	(0.0672)	(0.0659)	(0.196)	(0.172)	(0.121)	(0.121)	(0.123)
p-value	0.000	0.000	0.411	0.028	0.000	0.000	0.000
Large shocks							
β_0^-	0.324	0.313	0.391	0.303	0.432	0.206	0.304
	(0.0362)	(0.0405)	(0.0734)	(0.107)	(0.0811)	(0.0706)	(0.0712)
β_0^+	0.639	0.718	0.181	0.537	0.779	0.945	0.619
	(0.0396)	(0.0400)	(0.111)	(0.116)	(0.0636)	(0.0796)	(0.0643)
p-value	0.000	0.000	0.035	0.202	0.013	0.000	0.006
Observations	245	209	36	39	52	48	70

Table 3.3: The averages of pass-through rates by differing group sizes are reported. Below averages, standard errors are reported in parentheses. p-values correspond to the result of the Wilcoxon signed-rank test on the equality of pass-through rates for small or large shocks (i.e. $H_0 : \beta_0^+ = \beta_0^-$).

Table 3.2 provides average value of pass-through rates for different aggregation levels. First, we note that the hypothesis of full pass-through can generally be rejected.[18] Second, we test APT in the immediate post-shock responses by testing the equality of immediate pass-through rates for equally sized shocks as $H_0 : \beta_0^+ = \beta_0^-$ via Wilcoxon signed-rank tests. The pooled data and the data for groups of size greater than 2 suggest rejecting the null. For groups of size 3, we reject symmetry for the smaller but not for the larger shock. For duopolies, we see that the asymmetry is reversed; the average price response following the larger cost shock is significantly greater for the negative than for the positive cost shock. We do not attempt to reconcile this puzzling result, but simply note that the data for duopolies do not suggest pricing behavior in line with the "Rockets and Feathers" phenomenon. Taken together with the estimates of the DLM, we reach the first two results of our paper:

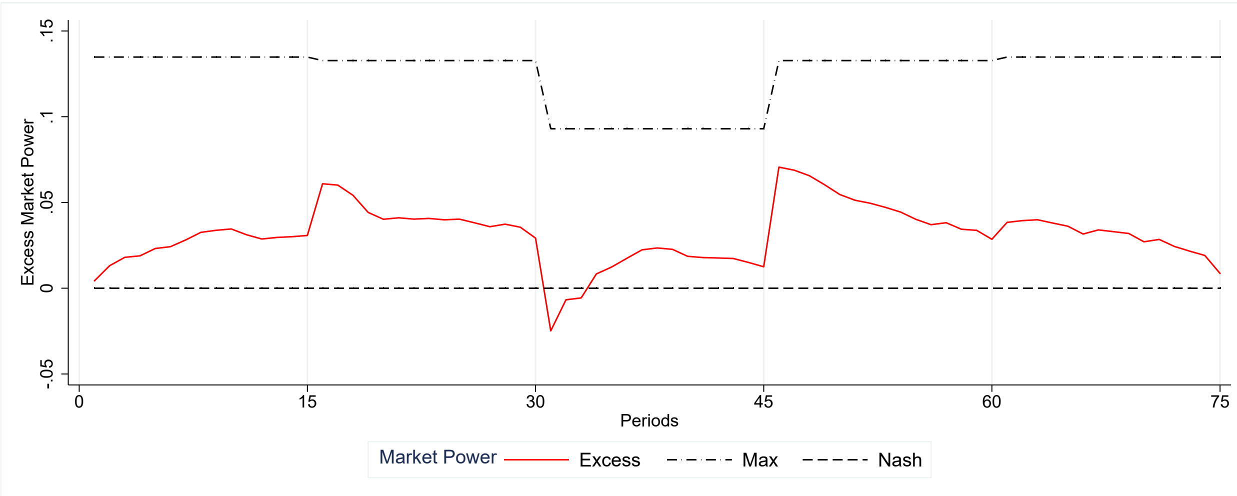
Result 1.1: *Prices do not react symmetrically to equally sized positive and negative shocks.*

Result 1.2: *Price reactions in duopoly markets are not consistent with APT.*

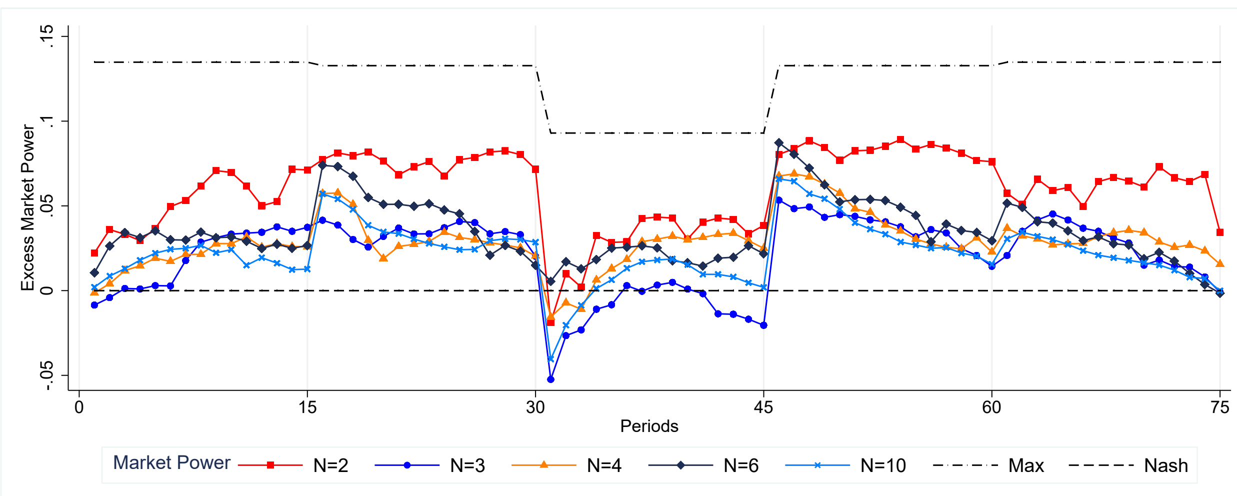
3.4.2 Market Power

We now turn to our second hypothesis. We follow the literature in applying the Lerner index as the relevant measure of market power: $L_{i,t} = \frac{p_{i,t} - mc_t}{p_{i,t}}$ (Lerner, 1934). We propose that the difference between the observed Lerner index (i.e., $L_{i,t}$) and the "theoretical" Lerner index, that is the index that would be relevant if behavior was consistent with NE predictions (i.e., $L_t^{NE} = \frac{p_t^{NE} - mc_t}{p_t^{NE}}$), provides a measure of "excess" market power due to collusion. We further propose this as an appropriate measure of tacit collusion, as our price competition structure incorporates homogeneous goods, and we control marginal costs. Thus, we do not suffer the

[18]Exceptions consist of the small positive shock (i.e., $\eta^+ = 0.40$) and $N = 6$ for the large positive shock.



(a) Average of excess market power



(b) Excess market power by group size

Figure 3.3: Excess market power across periods and group sizes. In both subfigures, "Max" refers to the maximum excess market power that can be observed (i.e., when $L_{i,t} = L_t^{max} = \frac{p^{max} - MC_r}{p^{max}}$) and "Nash" refers to the case $L_{i,t} = L_t^{NE}$.

identification problem of observational studies. Our measure of excess market power can then be expressed as:

$$L_{i,t}^x = L_{i,t} - L_t^{NE} = mc_t \left(\frac{1}{p_t^{NE}} - \frac{1}{p_{i,t}} \right). \tag{3.7}$$

Figure 3.3 depicts the average of our measure of excess market power, by period and treatment. Upon visual examination, one can immediately see that excess market power generally lies above the theoretical "Nash" level, consistent with an environment in which tacit collusion exists. Also, this average measure reaches its highest levels during the second and fourth rounds, the two rounds that immediately follow negative shocks. Following the large positive shock at the beginning of the third round, excess market power falls so much that it turns negative for several periods. Following the small positive shock at the beginning of the fifth round, excess market power does not react notably.

We test veracity of these observations by performing OLS regressions.[19] We consider the following specification:

$$L^x_{i,t} = \alpha + \sum_{s \neq 2} \delta_s \cdot \mathbb{1}[N = s] + \sum_{e \neq 1} \gamma_e \cdot \mathbb{1}[r = e] + \epsilon_{i,t}, \tag{3.8}$$

where the excess market power of seller i in period t is modeled as a function of group size- and round-specific indicator variables. Our main hypothesis consists of the joint nullity of all coefficients.

Table 3.4 reports the estimates in a step-wise manner. In model (5), we truncate the data to the periods where shocks shift the marginal cost (i.e., periods 16, 31, 46 and 61) and replace the dependent variable with the change in excess market power as $\Delta L^x_{i,t}$. This allows us to interpret the estimates of round specific indicator variables as the immediate effect of cost shocks on the tacit collusion in model (5).

First, we reject the joint nullity of all coefficients in all models except model (2) at a confidence level of 0.01. The fact that the constant α is positive and significant in model (1)

[19] The non-parametric counterpart of this test is available in Online Appendix II.B

Table 3.4: Excess market power

	(1)	(2)	(3)	(4)	(5)
Constant	0.032	0.060	0.025	0.054	0.004
	(0.004)	(0.011)	(0.004)	(0.011)	(0.007)
$N = 3$		-0.039		-0.039	0.018
		(0.015)		(0.015)	(0.020)
$N = 4$		-0.031		-0.031	0.028
		(0.014)		(0.014)	(0.010)
$N = 6$		-0.026		-0.026	0.047
		(0.014)		(0.014)	(0.010)
$N = 10$		-0.038		-0.038	0.029
		(0.012)		(0.012)	(0.011)
$r = 2$			0.017	0.017	
			(0.003)	(0.003)	
$r = 3$			-0.014	-0.014	-0.084
			(0.004)	(0.004)	(0.009)
$r = 4$			0.023	0.023	0.028
			(0.004)	(0.004)	(0.006)
$r = 5$			0.005	0.005	-0.020
			(0.004)	(0.004)	(0.007)
Observations	18355	18355	18355	18355	980
Adjusted R^2	-	0.052	0.054	0.106	0.225
F-statistic	-	2.607	31.289	19.288	27.003

Table 3.5: Results of OLS regressions on model 3.8 are reported. In model 5, the dependent variable is the change in excess market power, $\Delta L^x_{i,t}$. Below estimates, robust standard errors that are clustered at the market level are reported in parentheses.

indicate the overall presence of tacit collusion. Second, the coefficients of round-specific indicator variables in model (3) indicate that tacit collusion is higher during the second and fourth rounds, and lower during the third round relative to the first round. In model (5) where we truncate the data, the coefficient of rounds 3 and 5 are negative and that of round 4 is positive. Furthermore, we reject the hypothesis $H_0 : \alpha + \delta_s + \gamma_e = 0$ at a confidence level of 0.05 (i.e., $\alpha + \delta_{\{N=3,4,10\}} + \gamma_5 = 0$). We can thus say that immediately after a negative (the large positive) shock, the excess market power increases (decreases). Third, coefficients of group size specific indicator variables are negative and significant, although at marginal level for $N = 6$ (p-value$= 0.071$) in model (2). Here, we also reject the hypothesis $H_0 : \alpha + \delta_s = 0$

for all s (p-value< 0.01). This suggests that tacit collusion is present in all markets but its magnitude is smaller when $N > 2$. The sign of these coefficients in model (5) suggests that markets larger than size 3 increase their market power in response to the first negative shock. Lastly, we generally reject the hypothesis $H_0 : \alpha + \delta_s + \gamma_e = 0$ in the most unrestricted model (4) (11 times out of 15 tests at p-value< 0.05). The overall interpretation of these tests provide the basis of our second result:

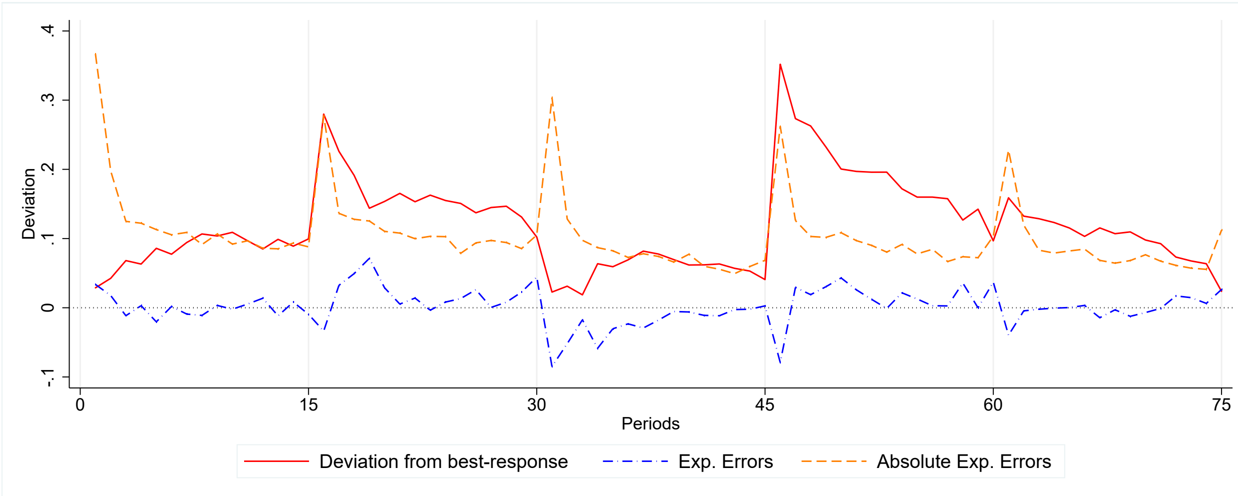
Result 2: *Excess market power (i.e., tacit collusion) is not invariant to shock direction and group size. It is persistently higher in duopolies, and in larger-sized markets it rises following negative cost shocks.*

3.4.3 Deviations from Best-Response

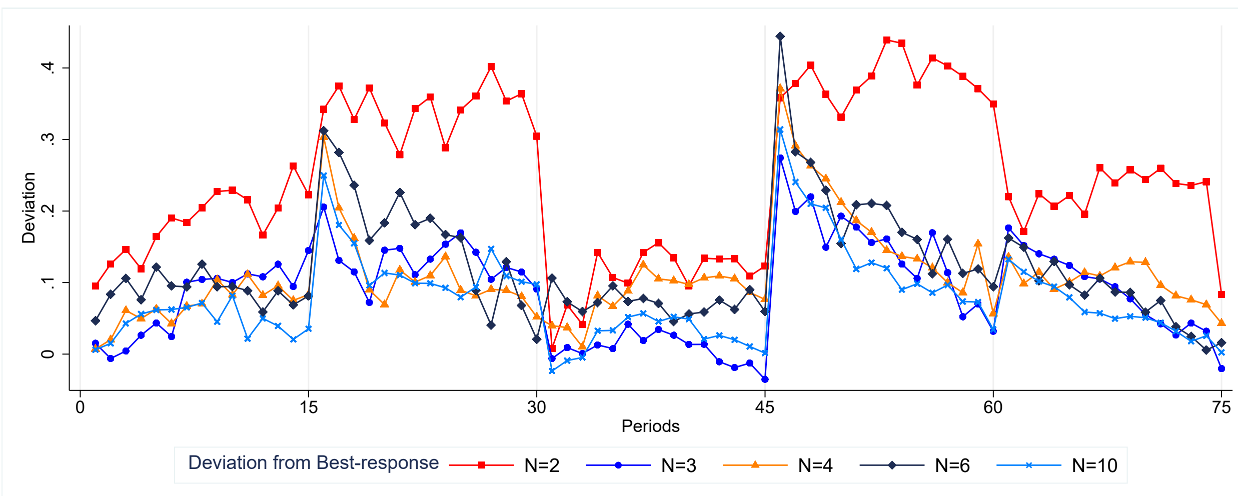
Finally, we assess deviations from subjects' best-response actions. Deviations from the best-response action can be attributed either to error, or alternatively to strategic motives. To argue that the deviations we observe in our experiments are not entirely emerging from erroneous behavior, we compare the magnitudes of such deviations to the average magnitude of expectation errors (i.e., $\mathbb{E}_{i,t-1}[p_{-i,t}] - p_{-i,t}$) and the average of absolute expectation errors.[20]

Figure 3.4(a) depicts the average value of these deviations over time. The average expectation errors are remarkably close to zero, with no obvious trend across periods. Although this suggests that beliefs are on average correct, it does not imply the complete absence of errors: The average measure of absolute expectation errors lies well above zero throughout the experiment. The latter peaks following cost shocks but subsequently trends downward. These results on expectation errors are consistent with those of prior experiments in which prices are strategic complements (e.g., Hommes et al. 2005). However, deviations from the

[20] We label these latter two as "errors" rather than as "deviations" as there is no strategic benefit to knowingly submitting inaccurate guesses/expectations in our experiment.



(a) Average of different deviations across all markets



(b) The deviations from best-response action by group size

Figure 3.4: Deviations from best-response action and errors in expectations.

best-response action reveal a different and rather interesting pattern: They peak sharply following negative shocks and remain high during these rounds, but do not peak similarly following positive shocks. The second graph in Figure 3.4 depicts the average of deviations from best-response action by group size. The same pattern can be traced across our treatment groups.

We perform OLS regressions to study deviations from best-response. Consider the following specification:

$$p_{i,t} - p_{i,t}^{BR|\mathbb{E}} = \alpha + \sum_{s \neq 2} \delta_s \cdot \mathbb{1}[N = s] + \sum_{e \neq 1} \gamma_e \cdot \mathbb{1}[r = e] + \epsilon_{i,t} \tag{3.9}$$

where the deviation of subject i's price from the best-response action conditional on the submitted guess (i.e., $p_{i,t}^{BR|\mathbb{E}}$) is modeled as a function of group size- and round-specific indicator variables. The theory postulates the joint nullity of all coefficients.

Table 3.6 reports the estimates in a step-wise manner. In model (5), we truncate the data to the periods where shocks shift marginal cost, the same way in previous section and replace the dependent variable with the change in deviation from best-response action.

The joint nullity of all coefficients can be rejected in all specifications. The hypotheses $H_0 : \alpha + \gamma_s = 0$ in model (2) and $H_0 : \alpha + \delta_s = 0$ in model (3) can be rejected at significance of 0.01. This points out to the following two results: (i) Sellers deviate more (less) from the associated best-response action following negative (positive) shocks and (ii) deviations are lower when $N > 2$. We see that the sign of group size indicator coefficients in models (4) and (5) are flipped. In model (4), they reflect the fact that groups of size 3 and larger deviate less, on average, relative to duopolies. In model (5), they correspond to the immediate reaction of these groups to the first negative shock. These deviations rise further when a large negative shock shifts the marginal cost down ($\hat{\gamma}_4 = 0.131$) while they drop significantly in response to the large positive shock ($\hat{\gamma}_3 = -0.260$). In consequence, we reach to the following results:

Result 3.1: *Sellers deviate on average from their best-response action.*

Table 3.6: Deviations from best-response

	(1)	(2)	(3)	(4)	(5)
Constant	0.120	0.083	0.248	0.212	0.044
	(0.013)	(0.012)	(0.041)	(0.040)	(0.028)
$N = 3$			-0.159	-0.159	0.122
			(0.049)	(0.049)	(0.055)
$N = 4$			-0.138	-0.138	0.163
			(0.050)	(0.050)	(0.051)
$N = 6$			-0.128	-0.128	0.210
			(0.051)	(0.051)	(0.041)
$N = 10$			-0.170	-0.170	0.144
			(0.044)	(0.044)	(0.046)
$r = 2$		0.080		0.080	
		(0.010)		(0.010)	
$r = 3$		-0.028		-0.028	-0.260
		(0.010)		(0.010)	(0.031)
$r = 4$		0.112		0.112	0.131
		(0.015)		(0.015)	(0.025)
$r = 5$		0.018		0.018	-0.118
		(0.013)		(0.013)	(0.024)
Observations	18355	18355	18355	18355	980
Adjusted R^2	-	0.044	0.051	0.095	0.170
F-statistic	-	37.840	4.013	19.546	40.442

Table 3.7: Results of OLS regressions on model 3.9 are reported. In model 5, the dependent variable is the change in deviation from the best-response action, $\Delta(p_{i,t} - p_{i,t}^{BR|\mathbb{E}})$. Robust standard errors are clustered at the market level and are reported in parentheses.

Result 3.2: *Deviations from the best-response action grow (shrink) following negative (positive) shocks.*

3.5 Discussion

Our results point to the co-appearance of asymmetric price transmission and tacit collusion. The latter seems to be the result of strategic behavior, as our analysis of deviations from best-response action reveals. These findings are consistent with theories that cast tacit collusion as having a significant role in the emergence of APT, such as the trigger price model in Borenstein et al. (1997). Most of the other theoretical explanations of APT in the literature cannot account for the pricing behavior observed in our results. We can reasonably exclude, for example, the influence of explicit collusion (i.e., involving direct communication), capacity constraints, inventory limitations, (a)symmetric menu costs, consumer loss aversion, (a)symmetric search costs, contexts of alternating price moves and price lockup periods, and so forth, as being necessary conditions for APT, since these features are excluded by our design.

We cannot, however, claim a monotonic relation between the magnitude of APT and tacit collusion: the pricing behavior of duopolies in our experiment is revealed to be fairly symmetric. We explain the exceptionality of the result for duopoly markets as follows: In duopolies, collusion is so strong that sellers are, by and large, able to maintain cooperative (tacitly collusive) pricing over a sustained period of time, with pricing showing no reversion to Nash. We therefore argue that APT requires significant, but imperfect, collusion.[21]

If tacit collusion is indeed a significant causal force behind APT, then our work has important implications for antitrust enforcement policy against collusion and price-fixing. In particular, regulators may consider APT in a market as a signal for the presence of collusion between firms in that market. Since many real-world interactions between competitive firms are repeated indefinitely, such collusion may even be sustainable as a NE. Further research is needed to determine whether collusion is an important cause of APT behavior in field settings, and if so, whether suitable forms of regulatory intervention might exist to reduce such collusion without increases in inefficiency.

[21]The fact that our duopolies reached almost stable collusion, while larger markets did not, is consistent with the literature we review in Section 3.1.3. This can be attributable to a combination of two factors: first, coordination between market participants becomes increasingly difficult with each new seller, and three may well be the number from which the difficulties and costs involved in maintaining coordination start to exceed the marginal benefits; second, our duopolies are unique in that each participant can deduce the choices made by the other participant by observing aggregate market outcomes. In a triopoly or larger market, by contrast, it is not possible for sellers to detect whether an aggregate market outcome is due to the defection of a single competitor, or from a broader but shallower defection by multiple competitors.

We propose that follow-on research may yield further insights into the mechanisms through which tacit collusion leads to APT, as well as potential policy responses that might diminish its frequency and magnitude. In particular, future experiments should address the impact of different levels of information transparency. Most notably, testing the effects of providing feedback on individual prices and/or payoffs of rivals on APT may provide particularly helpful insights. The latter is shown to lead to more rivalistic outcomes in experimental oligopoly studies, as it initiates imitation dynamics (Fiala and Suetens, 2017), while the former can lead to more collusive levels. Nevertheless, both may reduce the degree of asymmetry in price transmission. Another area of needed research is to explore the roles of market power and market concentration in shaping APT pricing behavior. In our experimental design we explicitly kept the theoretical market power of sellers the same across markets of varying sizes to study the impact of group size on coordination channel, alone. Finally, future studies may benefit from testing the robustness of our findings to alternative demand specifications, or to the introduction of human subjects acting as buyers.

Bibliography

Abbink, Klaus and Jordi Brandts (2008) "Pricing in Bertrand competition with increasing marginal costs," *Games and Economic Behavior*, Vol. 63, No. 1, pp. 1–31.

Adena, M. and S. Huck (2017) "Matching donations without crowding out? Some theoretical considerations, a field, and a lab experiment," *Journal of Public Economics*, Vol. 148, pp. 32–42, DOI: 10.1016/j.jpubeco.2017.02.002.

Ahrens, Steffen, Inske Pirschel, and Dennis J Snower (2017) "A theory of price adjustment under loss aversion," *Journal of Economic Behavior & Organization*, Vol. 134, pp. 78–95.

Amir, Rabah, Philip Erickson, and Jim Jin (2017) "On the microeconomic foundations for linear demand for differentiated products," *Journal of Economic Theory*, Vol. 169, pp. 641–665.

Anderson, James E. (2011) "The Gravity Model," *Annual Review of Economics*, Vol. 3, No. 1, pp. 133–160, URL: `https://doi.org/10.1146/annurev-economics-111809-125114`, DOI: 10.1146/annurev-economics-111809-125114.

Bacon, Robert W (1991) "Rockets and feathers: the asymmetric speed of adjustment of UK retail gasoline prices to cost changes," *Energy Economics*, Vol. 13, No. 3, pp. 211–218.

Ball, Laurence and N Gregory Mankiw (1994) "Asymmetric price adjustment and economic fluctuations," *Economic Journal*, Vol. 104, No. 423, pp. 247–261.

Bapat, Navin A., Tobias Heinrich, Yoshiharu Kobayashi, and T. Clifton Morgan (2013) "Determinants of Sanctions Effectiveness: Sensitivity Analysis Using New Data," *International Interactions*, Vol. 39, No. 1, pp. 79–98, URL: `https://doi.org/10.1080/03050629.2013.751298`, DOI: 10.1080/03050629.2013.751298.

Bayer, Ralph-C and Changxia Ke (2018) "What causes rockets and feathers? An experimental investigation," *Journal of Economic Behavior & Organization*, Vol. 153, pp. 223–237.

Bayer, Ralph-C., Elke Renner, and Rupert Sausgruber (2013) "Confusion and learning in the voluntary contributions game," *Experimental Economics*, Vol. 16, No. 4, pp. 478–496, URL: `https://link.springer.com/article/10.1007/s10683-012-9348-2`, DOI: 10.1007/s10683-012-9348-2.

Bearce, David H. (2003) "Grasping the Commercial Institutional Peace," *International Studies Quarterly*, Vol. 47, No. 3, pp. 347–370, URL: `https://academic.oup.com/isq/article/47/3/347/1922632`, DOI: 10.1111/1468-2478.4703003.

Benabou, Roland and Robert Gertner (1993) "Search with learning from prices: does increased inflationary uncertainty lead to higher markups?" *Review of Economic Studies*, Vol. 60, No. 1, pp. 69–93.

Bergstrand, Jeffrey H. (1985) "The Gravity Equation in International Trade: Some Microeconomic Foundations and Empirical Evidence," *The Review of Economics and Statistics*, Vol. 67, No. 3, pp. 474–481, URL: `https://www.jstor.org/stable/1925976`, DOI: 10.2307/1925976.

Bicchieri, Cristina (2005) *The Grammar of Society: The Nature and Dynamics of Social Norms*: Cambridge University Press, Google-Books-ID: 4N1FDIZvcI8C.

Binmore, Ken (2010) "Social norms or social preferences?" *Mind & Society*, Vol. 9, No. 2, pp. 139–157, URL: `https://doi.org/10.1007/s11299-010-0073-2`, DOI: 10.1007/s11299-010-0073-2.

Borenstein, Severin, A Colin Cameron, and Richard Gilbert (1997) "Do gasoline prices respond asymmetrically to crude oil price changes?" *Quarterly Journal of Economics*, Vol. 112, No. 1, pp. 305–339.

Borenstein, Severin and Andrea Shepard (1996) "Sticky prices, inventories, and market power in wholesale gasoline markets,"Technical report, National Bureau of Economic Research.

Bostrom, Nick (2002) "Existential risks: analyzing human extinction scenarios and related hazards," *Journal of Evolution and Technology*, Vol. 9, URL: `https://ora.ox.ac.uk/objects/uuid:827452c3-fcba-41b8-86b0-407293e6617c`.

——— (2011) "Infinite Ethics," *Analysis and Metaphysics*, No. 10, pp. 9–59, URL: `https://www.ceeol.com/search/article-detail?id=77729`.

Brown, Stephen PA and Mine K Yucel (2000) "Gasoline and crude oil prices: why the asymmetry?" *Economic & Financial Review*, No. Third Quarter, pp. 23–29.

Byrne, David P and Nicolas De Roos (2019) "Learning to coordinate: A study in retail gasoline," *American Economic Review*, Vol. 109, No. 2, pp. 591–619.

Cappelen, Alexander W., Ulrik H. Nielsen, Erik Ø. Sørensen, Bertil Tungodden, and Jean-Robert Tyran (2013) "Give and take in dictator games," *Economics Letters*, Vol. 118, No. 2, pp. 280–283, URL: `http://www.sciencedirect.com/science/article/pii/S0165176512005733`, DOI: 10.1016/j.econlet.2012.10.030.

Cason, Timothy N and Daniel Friedman (2002) "A laboratory study of customer markets," *BE Journal of Economic Analysis & Policy*, Vol. 2, No. 1, pp. 1–43.

Cason, Timothy N. and Lata Gangadharan (2015) "Promoting cooperation in nonlinear social dilemmas through peer punishment," *Experimental Economics*, Vol. 18, No. 1, pp. 66–88, URL: `https://doi.org/10.1007/s10683-014-9393-0`, DOI: 10.1007/s10683-014-9393-0.

Charness, Gary and Chun-Lei Yang (2014) "Starting small toward voluntary formation of efficient large groups in public goods provision," *Journal of Economic Behavior & Organization*, Vol. 102, pp. 119–132, URL: `http://www.sciencedirect.com/science/article/pii/S0167268114000754`, DOI: 10.1016/j.jebo.2014.03.005.

Checkel, Jeffrey T. (1997) "International Norms and Domestic Politics:: Bridging the Rationalist?Constructivist Divide," *European Journal of International Relations*, Vol. 3, No. 4, pp. 473–495, URL: `https://doi.org/10.1177/1354066197003004003`, DOI: 10.1177/1354066197003004003.

Corazzini, Luca and Marcelo Tyszler (2015) "Preference for Efficiency or Confusion? A Note on a Boundedly Rational Equilibrium Approach to Individual Contributions in a Public Good Game," URL: `https://www.hindawi.com/journals/jam/2015/961930/`, DOI: 10.1155/2015/961930.

Coricelli, Giorgio, Dietmar Fehr, and Gerlinde Fellner (2004) "Partner Selection in Public Goods Experiments," *Journal of Conflict Resolution*, Vol. 48, No. 3, pp. 356–378, URL: `https://doi.org/10.1177/0022002704264143`, DOI: 10.1177/0022002704264143.

Cornand, Camille and Frank Heinemann (2019) "Experiments in macroeconomics: methods and applications," in *Handbook of Research Methods and Applications in Experimental Economics*: Edward Elgar Publishing.

Cornes, Richard and Roger Hartley (2003) "Risk Aversion, Heterogeneity and Contests," *Public Choice*, Vol. 117, No. 1, pp. 1–25, URL: `https://doi.org/10.1023/A:1026136008608`, DOI: 10.1023/A:1026136008608.

——— (2012) "Risk aversion in symmetric and asymmetric contests," *Economic Theory*, Vol. 51, No. 2, pp. 247–275, URL: `https://doi.org/10.1007/s00199-009-0490-9`, DOI: 10.1007/s00199-009-0490-9.

Crumpler, Heidi and Philip J. Grossman (2008) "An experimental test of warm glow giving," *Journal of Public Economics*, Vol. 92, No. 5, pp. 1011–1021, URL: `http://www.sciencedirect.com/science/article/pii/S0047272708000029`, DOI: 10.1016/j.jpubeco.2007.12.014.

Cubel, María and Santiago Sanchez-Pages (2016) "An axiomatization of difference-form contest success functions," *Journal of Economic Behavior & Organization*, Vol. 131, pp. 92–105, URL: `http://www.sciencedirect.com/science/article/pii/S0167268116301780`, DOI: 10.1016/j.jebo.2016.08.011.

Davis, Douglas (2009) "Pure numbers effects, market power, and tacit collusion in posted offer markets," *Journal of Economic Behavior & Organization*, Vol. 72, No. 1, pp. 475–488.

Dawkins, Richard, John Richard Krebs, J. Maynard Smith, and Robin Holliday (1979) "Arms races between and within species," *Proceedings of the Royal Society of London. Series B. Biological Sciences*, Vol. 205, No. 1161, pp. 489–511, URL: https://royalsocietypublishing.org/doi/abs/10.1098/rspb.1979.0081, DOI: 10.1098/rspb.1979.0081.

Deck, Cary A and Bart J Wilson (2008) "Experimental gasoline markets," *Journal of Economic Behavior & Organization*, Vol. 67, No. 1, pp. 134–149.

Diamond, Peter (2006) "Optimal tax treatment of private contributions for public goods with and without warm glow preferences," *Journal of Public Economics*, Vol. 90, No. 4, pp. 897–919, URL: http://www.sciencedirect.com/science/article/pii/S0047272705000897, DOI: 10.1016/j.jpubeco.2005.06.001.

Doyle, Michael W. (2005) "Three Pillars of the Liberal Peace," *American Political Science Review*, Vol. 99, No. 3, pp. 463–466, URL: https://www.cambridge.org/core/journals/american-political-science-review/article/three-pillars-of-the-liberal-peace/B704A7319ADFE0208E0B3D0FFB30A860, DOI: 10.1017/S0003055405051798.

Duersch, Peter and Thomas A Eife (2019) "Price competition in an inflationary environment," *Journal of Monetary Economics*, Vol. 104, pp. 48–66.

Duffy, John (2016) "Macroeconomics: a survey of laboratory research," *Handbook of Experimental Economics*, Vol. 2, pp. 1–90.

Dufwenberg, Martin and Uri Gneezy (2000) "Price competition and market concentration: an experimental study," *International Journal of Industrial Organization*, Vol. 18, No. 1, pp. 7–22.

Eckel, Catherine and Philip Grossman (2003) "Rebate versus matching: does how we subsidize charitable contributions matter?" *Journal of Public Economics*, Vol. 87, No. 3-4, pp. 681–701, URL: https://econpapers.repec.org/article/eeepubeco/v_3a87_3ay_3a2003_3ai_3a3-4_3ap_3a681-701.htm.

Falkinger, Josef, Ernst Fehr, Simon Gächter, and Rudolf Winter-Ember (2000) "A Simple Mechanism for the Efficient Provision of Public Goods: Experimental Evidence," *American Economic Review*, Vol. 90, No. 1, pp. 247–264, URL: https://www.aeaweb.org/articles?id=10.1257/aer.90.1.247, DOI: 10.1257/aer.90.1.247.

Fehr, Ernst and Simon Gachter (2000) "Cooperation and Punishment in Public Goods Experiments," *American Economic Review*, Vol. 90, No. 4, pp. 980–994, URL: https://www.aeaweb.org/articles?id=10.1257/aer.90.4.980, DOI: 10.1257/aer.90.4.980.

Fehr, Ernst and Jean-Robert Tyran (2001) "Does money illusion matter?" *American Economic Review*, Vol. 91, No. 5, pp. 1239–1262.

Fiala, Lenka and Sigrid Suetens (2017) "Transparency and cooperation in repeated dilemma games: a meta study," *Experimental Economics*, Vol. 20, No. 4, pp. 755–771.

Finnemore, Martha and Kathryn Sikkink (2001) "TAKING STOCK: The Constructivist Research Program in International Relations and Comparative Politics," *Annual Review of Political Science*, Vol. 4, No. 1, pp. 391–416, URL: `https://doi.org/10.1146/annurev.polisci.4.1.391`, DOI: 10.1146/annurev.polisci.4.1.391.

Fischbacher, Urs (2007) "z-Tree: Zurich toolbox for ready-made economic experiments," *Experimental economics*, Vol. 10, No. 2, pp. 171–178.

Fonseca, Miguel A and Hans-Theo Normann (2012) "Explicit vs. tacit collusion—The impact of communication in oligopoly experiments," *European Economic Review*, Vol. 56, No. 8, pp. 1759–1772.

Frey, Giliola and Matteo Manera (2007) "Econometric models of asymmetric price transmission," *Journal of Economic Surveys*, Vol. 21, No. 2, pp. 349–415.

Fu, Tingting, Yunan Ji, Kenju Kamei, and Louis Putterman (2017) "Punishment can support cooperation even when punishable," *Economics Letters*, Vol. 154, pp. 84–87, URL: `http://www.sciencedirect.com/science/article/pii/S0165176517300265`, DOI: 10.1016/j.econlet.2017.01.016.

Fudenberg, Drew and Jean Tirole (1991) *Game Theory*, Cambridge, Mass: The MIT Press, 1st edition.

Gaddis, John Lewis (1986) "The Long Peace: Elements of Stability in the Postwar International System," *International Security*, Vol. 10, No. 4, pp. 99–142, URL: `https://muse.jhu.edu/article/446174/summary`.

——— (1992) "The Cold War, the Long Peace, and the Future," *Diplomatic History*, Vol. 16, No. 2, pp. 234–246, URL: `https://academic.oup.com/dh/article/16/2/234/338721`, DOI: 10.1111/j.1467-7709.1992.tb00499.x.

Gintis, Herbert (2010) "Social norms as choreography," *Politics, Philosophy & Economics*, Vol. 9, No. 3, pp. 251–264, URL: `https://doi.org/10.1177/1470594X09345474`, DOI: 10.1177/1470594X09345474.

Goldsmith, Jack Landman and Eric A. Posner (2000) "Moral and Legal Rhetoric in International Relations: A Rational Choice Perspective," SSRN Scholarly Paper ID 250042, Social Science Research Network, Rochester, NY.

Green, Edward J and Robert H Porter (1984) "Noncooperative collusion under imperfect price information," *Econometrica*, Vol. 52, No. 1, pp. 87–100.

Greiner, Ben (2015) "Subject pool recruitment procedures: organizing experiments with ORSEE," *Journal of the Economic Science Association*, Vol. 1, No. 1, pp. 114–125.

Harsanyi, John C. and Reinhard Selten (1988) *A General Theory of Equilibrium Selection in Games*, Vol. 1: The MIT Press, URL: `https://ideas.repec.org/b/mtp/titles/0262582384.html`.

Hommes, Cars, Joep Sonnemans, Jan Tuinstra, and Henk Van de Velden (2005) "Coordination of expectations in asset pricing experiments," *Review of Financial Studies*, Vol. 18, No. 3, pp. 955–980.

Horstmann, Niklas, Jan Krämer, and Daniel Schnurr (2018) "Number effects and tacit collusion in experimental oligopolies," *Journal of Industrial Economics*, Vol. 66, No. 3, pp. 650–700.

Hotelling, Harold (1929) "Stability in competition," *Economic Journal*, Vol. 39, No. 153, pp. 41–57.

Huck, Steffen, Hans-Theo Normann, and Jörg Oechssler (2004) "Two are few and four are many: number effects in experimental oligopolies," *Journal of Economic Behavior & Organization*, Vol. 53, No. 4, pp. 435–446.

Hwang, Sung-Ha (2009) "Contest success functions: Theory and evidence," Working Paper 2009-04, Working Paper.

Inglehart, Ronald F, Bi Puranen, and Christian Welzel (2015) "Declining willingness to fight for one?s country: The individual-level basis of the long peace," *Journal of Peace Research*, Vol. 52, No. 4, pp. 418–434, URL: `https://doi.org/10.1177/0022343314565756`, DOI: 10.1177/0022343314565756.

Isaac, R. Mark and James M. Walker (1998) "Nash as an organizing principle in the voluntary provision of public goods: Experimental evidence," *Experimental Economics*, Vol. 1, No. 3, pp. 191–206, URL: `https://doi.org/10.1007/BF01669342`, DOI: 10.1007/BF01669342.

Isard, Walter (1954) "Location Theory and Trade Theory: Short-Run Analysis," *The Quarterly Journal of Economics*, Vol. 68, No. 2, pp. 305–320, URL: `https://academic.oup.com/qje/article/68/2/305/1896234`, DOI: 10.2307/1884452.

Jia, Hao, Stergios Skaperdas, and Samarth Vaidya (2013) "Contest functions: Theoretical foundations and issues in estimation," *International Journal of Industrial Organization*, Vol. 31, No. 3, pp. 211–222, URL: `http://www.sciencedirect.com/science/article/pii/S0167718712000811`, DOI: 10.1016/j.ijindorg.2012.06.007.

Joulfaian, David (2017) "Is Charitable Giving by the Rich Really Responsive to the Income Tax?" SSRN Scholarly Paper ID 1952889, Social Science Research Network, Rochester, NY.

Kagel, John H. and Alvin E. Roth (2016) *The Handbook of Experimental Economics*: Princeton University Press, Google-Books-ID: E2aYDwAAQBAJ.

Karlan, Dean and John A. List (2007) "Does Price Matter in Charitable Giving? Evidence from a Large-Scale Natural Field Experiment," *American Economic Review*, Vol. 97, No. 5, pp. 1774–1793, URL: `https://www.aeaweb.org/articles?id=10.1257/aer.97.5.1774`, DOI: 10.1257/aer.97.5.1774.

Kepaptsoglou, Konstantinos, Matthew G. Karlaftis, and Dimitrios Tsamboulas (2010) "The Gravity Model Specification for Modeling International Trade Flows and Free Trade Agreement Effects: A 10-Year Review of Empirical Studies," *The Open Economics Journal*, Vol. 3, No. 1, URL: https://benthamopen.com/ABSTRACT/TOECONSJ-3-1.

Konrad, Kai A. and Harris Schlesinger (1997) "Risk Aversion in Rent?Seeking and Rent?Augmenting Games," *The Economic Journal*, Vol. 107, No. 445, pp. 1671–1683, URL: https://academic.oup.com/ej/article/107/445/1671/5064007, DOI: 10.1111/j.1468-0297.1997.tb00074.x.

Krasteva, Silvana and Huseyin Yildirim (2016) "Information, competition, and the quality of charities," *Journal of Public Economics*, Vol. 144, pp. 64–77, URL: http://www.sciencedirect.com/science/article/pii/S0047272716301554, DOI: 10.1016/j.jpubeco.2016.10.008.

Kriesberg, Louis (2007) "Long Peace or Long War: A Conflict Resolution Perspective," *Negotiation Journal*, Vol. 23, No. 2, pp. 97–116, URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1571-9979.2007.00132.x, DOI: 10.1111/j.1571-9979.2007.00132.x.

Kroenig, Matthew (2015) "The History of Proliferation Optimism: Does It Have a Future?" *Journal of Strategic Studies*, Vol. 38, No. 1-2, pp. 98–125, URL: https://doi.org/10.1080/01402390.2014.893508, DOI: 10.1080/01402390.2014.893508.

Krugman, Paul R (1991) *Geography and trade*, Leuven, Belgium; Cambridge, Mass.; London: Leuven University Press ; The MIT Press, OCLC: 318372085.

Landrigan, Philip J., Richard Fuller, and Ma et. al. Zhong (2018) "The Lancet Commission on pollution and health," *The Lancet*, Vol. 391, No. 10119, pp. 462–512, URL: https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(17)32345-0/abstract, DOI: 10.1016/S0140-6736(17)32345-0, Publisher: Elsevier.

Lebow, Richard Ned (1994) "The long peace, the end of the cold war, and the failure of realism," *International Organization*, Vol. 48, No. 2, pp. 249–277, URL: https://www.cambridge.org/core/journals/international-organization/article/long-peace-the-end-of-the-cold-war-and-the-failure-of-realism/398DDDB473ABC35F54E5B9CE8CF26EF9, DOI: 10.1017/S0020818300028186.

Lektzian, David and Glen Biglaiser (2013) "Investment, Opportunity, and Risk: Do US Sanctions Deter or Encourage Global Investment?" *International Studies Quarterly*, Vol. 57, No. 1, pp. 65–78, URL: https://academic.oup.com/isq/article/57/1/65/1801585, DOI: 10.1111/j.1468-2478.2012.00761.x.

Lerner, Abba P (1934) "The concept of monopoly and the measurement of monopoly power," *Review of Economic Studies*, Vol. 1, No. 3, pp. 157–175.

Loy, Jens-Peter, Christoph R Weiss, and Thomas Glauben (2016) "Asymmetric cost pass-through? Empirical evidence on the role of market power, search and menu costs," *Journal of Economic Behavior & Organization*, Vol. 123, pp. 184–192.

López-Pérez, Raúl (2008) "Aversion to norm-breaking: A model," *Games and Economic Behavior*, Vol. 64, No. 1, pp. 237–267, URL: `http://www.sciencedirect.com/science/article/pii/S0899825608000031`, DOI: 10.1016/j.geb.2007.10.009.

Lugovskyy, Volodymyr, Daniela Puzzello, Andrea Sorensen, James Walker, and Arlington Williams (2017) "An experimental study of finitely and infinitely repeated linear public goods games," *Games and Economic Behavior*, Vol. 102, pp. 286–302, URL: `http://www.sciencedirect.com/science/article/pii/S0899825617300052`, DOI: 10.1016/j.geb.2017.01.004.

Maskin, Eric and Jean Tirole (1988) "A theory of dynamic oligopoly, II: Price competition, kinked demand curves, and Edgeworth cycles," *Econometrica*, Vol. 56, No. 3, pp. 571–599.

Bueno de Mesquita, Bruce Bueno de, James D. Morrow, Randolph M. Siverson, and Alastair Smith (1999) "An Institutional Explanation of the Democratic Peace," *American Political Science Review*, Vol. 93, No. 4, pp. 791–807, URL: `https://www.cambridge.org/core/journals/american-political-science-review/article/an-institutional-explanation-of-the-democratic-peace/E9BF44FBC2B3E931100FAFCD0AAD85F7`, DOI: 10.2307/2586113.

Meyer, Jochen and Stephan von Cramon-Taubadel (2004) "Asymmetric price transmission: a survey," *Journal of Agricultural Economics*, Vol. 55, No. 3, pp. 581–611.

Morgan, John, Henrik Orzen, and Martin Sefton (2006) "An experimental study of price dispersion," *Games and Economic Behavior*, Vol. 54, No. 1, pp. 134–158.

Morgan, T. Clifton and Valerie L. Schwebach (1997) "Fools Suffer Gladly: The Use of Economic Sanctions in International Crises," *International Studies Quarterly*, Vol. 41, No. 1, pp. 27–50, URL: `https://www.jstor.org/stable/2600906`.

Muth, John F (1961) "Rational expectations and the theory of price movements," *Econometrica*, pp. 315–335.

Oguledo, Victor and Craig R. Macphee (1994) "Gravity models: a reformulation and an application to discriminatory trade arrangements," *Applied Economics*, Vol. 26, No. 2, pp. 107–120, URL: `https://doi.org/10.1080/00036849400000066`, DOI: 10.1080/00036849400000066.

Orzen, Henrik (2008) "Counterintuitive number effects in experimental oligopolies," *Experimental Economics*, Vol. 11, No. 4, pp. 390–401.

Pape, Robert A. (1997) "Why Economic Sanctions Do Not Work," *International Security*, Vol. 22, No. 2, pp. 90–136, URL: `https://www.jstor.org/stable/2539368`, DOI: 10.2307/2539368.

Paternotte, Cedric and Jonathan Grose (2013) "Social Norms and Game Theory: Harmony or Discord?" *The British Journal for the Philosophy of Science*, Vol. 64, No. 3, pp. 551–587, URL: `https://academic.oup.com/bjps/article/64/3/551/1470919`, DOI: 10.1093/bjps/axs024.

Peltzman, Sam (2000) “Prices rise faster than they fall,” *Journal of Political Economy*, Vol. 108, No. 3, pp. 466–502.

Perdiguero-García, Jordi (2013) “Symmetric or asymmetric oil prices? A meta-analysis approach,” *Energy policy*, Vol. 57, pp. 389–397.

Plott, Charles R. and Vernon L. Smith (2008) *Handbook of Experimental Economics Results*: Elsevier, Google-Books-ID: wAK_aKs3EPoC.

Powell, Robert (1994) “Anarchy in international relations theory: the neorealist-neoliberal debate,” *International Organization*, Vol. 48, No. 2, pp. 313–344, URL: https://www.cambridge.org/core/journals/international-organization/article/anarchy-in-international-relations-theory-the-neorealistneoliberal-debate/77011ECF30FE5DF0F0962CC300E4173C, DOI: 10.1017/S0020818300028204.

Quick, Elaine, Toni Ann Kruse, and Adam Pickering (2014) “Rules to Give By: A Global Philanthropy Legal Environment Index,” p. 345, URL: https://idis.org.br/wp-content/uploads/2014/12/RULES-TO-GIVE-BY-FINAL-with-Country-Reports.pdf.

Rauchhaus, Robert (2009) “Evaluating the Nuclear Peace Hypothesis: A Quantitative Approach,” *Journal of Conflict Resolution*, Vol. 53, No. 2, pp. 258–277, URL: http://journals.sagepub.com/doi/10.1177/0022002708330387, DOI: 10.1177/0022002708330387.

Reagan, Patricia B and Martin L Weitzman (1982) “Asymmetries in price and quantity adjustments by the competitive firm,” *Journal of Economic Theory*, Vol. 27, No. 2, pp. 410–420.

Reus-Smit, Christian and Duncan Reus-Smit (2008) *The Oxford Handbook of International Relations*: Oxford University Press, URL: https://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199219322.001.0001/oxfordhb-9780199219322, DOI: 10.1093/oxfordhb/9780199219322.001.0001.

Roser, Max (2016) “War and Peace,” *Our World in Data*, URL: https://ourworldindata.org/war-and-peace.

Salop, Steven C (1979) “Monopolistic competition with outside goods,” *The Bell Journal of Economics*, Vol. 10, No. 1, pp. 141–156.

Schulz, Jonathan F., Petra Thiemann, and Christian Thöni (2018) “Nudging generosity: Choice architecture and cognitive factors in charitable giving,” *Journal of Behavioral and Experimental Economics*, Vol. 74, pp. 139–145, URL: http://www.sciencedirect.com/science/article/pii/S2214804318301575, DOI: 10.1016/j.socec.2018.04.001.

Sheremeta, Roman M. and Neslihan Uler (2016) “The Impact of Taxes and Wasteful Government Spending on Giving,” SSRN Scholarly Paper ID 2702098, Social Science Research Network, Rochester, NY.

Sidel, Victor W. (1989) "Weapons of Mass Destruction: The Greatest Threat to Public Health," *JAMA*, Vol. 262, No. 5, pp. 680–682, URL: https://jamanetwork.com/journals/jama/fullarticle/378214, DOI: 10.1001/jama.1989.03430050096035.

Sinha, Anushua, Orin Levine, Maria D Knoll, Farzana Muhib, and Tracy A Lieu (2007) "Cost-effectiveness of pneumococcal conjugate vaccination in the prevention of child mortality: an international economic analysis," *The Lancet*, Vol. 369, No. 9559, pp. 389–396, URL: http://www.sciencedirect.com/science/article/pii/S0140673607601950, DOI: 10.1016/S0140-6736(07)60195-0.

Siverson, Randolph M. and Michael D. Ward (2002) "The Long Peace: A Reconsideration," *International Organization*, Vol. 56, No. 3, pp. 679–691, URL: https://www.cambridge.org/core/journals/international-organization/article/long-peace-a-reconsideration/3B2C1187408A151CEC777F1574BBBF33, DOI: 10.1162/002081802760199926.

Skaperdas, Stergios (1996) "Contest success functions," *Economic Theory*, Vol. 7, No. 2, pp. 283–290, URL: https://doi.org/10.1007/BF01213906, DOI: 10.1007/BF01213906.

Treich, Nicolas (2010) "Risk-aversion and prudence in rent-seeking games," *Public Choice*, Vol. 145, No. 3, pp. 339–349, URL: https://doi.org/10.1007/s11127-009-9569-x, DOI: 10.1007/s11127-009-9569-x.

de Vries, Anthonius, Clara Portela, and Borja Guijarro-Usobiaga (2014) "Improving the Effectiveness of Sanctions: A Checklist for the EU," SSRN Scholarly Paper ID 2520336, Social Science Research Network, Rochester, NY.

Wagstaff, Adam, Eddy van Doorslaer, and Hattem et. al. van der Burg (1999) "Redistributive effect, progressivity and differential tax treatment: Personal income taxes in twelve OECD countries," *Journal of Public Economics*, Vol. 72, No. 1, pp. 73–98, URL: http://www.sciencedirect.com/science/article/pii/S0047272798000851, DOI: 10.1016/S0047-2727(98)00085-1.

Whitehead, Anita and Timothy A.A. Stiles (2015) "Donations in Other Countries: Are they tax-deductible?"Technical report, KPMG.

Yamakawa, Takafumi, Yoshitaka Okano, and Tatsuyoshi Saijo (2016) "Detecting motives for cooperation in public goods experiments," *Experimental Economics*, Vol. 19, No. 2, pp. 500–512, URL: https://link.springer.com/article/10.1007/s10683-015-9451-2, DOI: 10.1007/s10683-015-9451-2.

Yang, Huanxing and Lixin Ye (2008) "Search with learning: understanding asymmetric price adjustments," *RAND Journal of Economics*, Vol. 39, No. 2, pp. 547–564.

Zampelli, Ernest M. and Steven T. Yen (2017) "The Impact of Tax Price Changes on Charitable Contributions to the Needy," *Contemporary Economic Policy*, Vol. 35, No. 1, pp. 113–124, URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/coep.12177, DOI: 10.1111/coep.12177.

Zhou, Weifeng and Ludo Cuyvers (2011) "Linking International Trade and Labour Standards: The Effectiveness of Sanctions under the European Union’s GSP," *Journal of World Trade*, Vol. 45, No. 1, pp. 63–85, URL: `http://kluwerlawonline.com/abstract.php?area=Journals&id=TRAD2011003`.

Appendix A

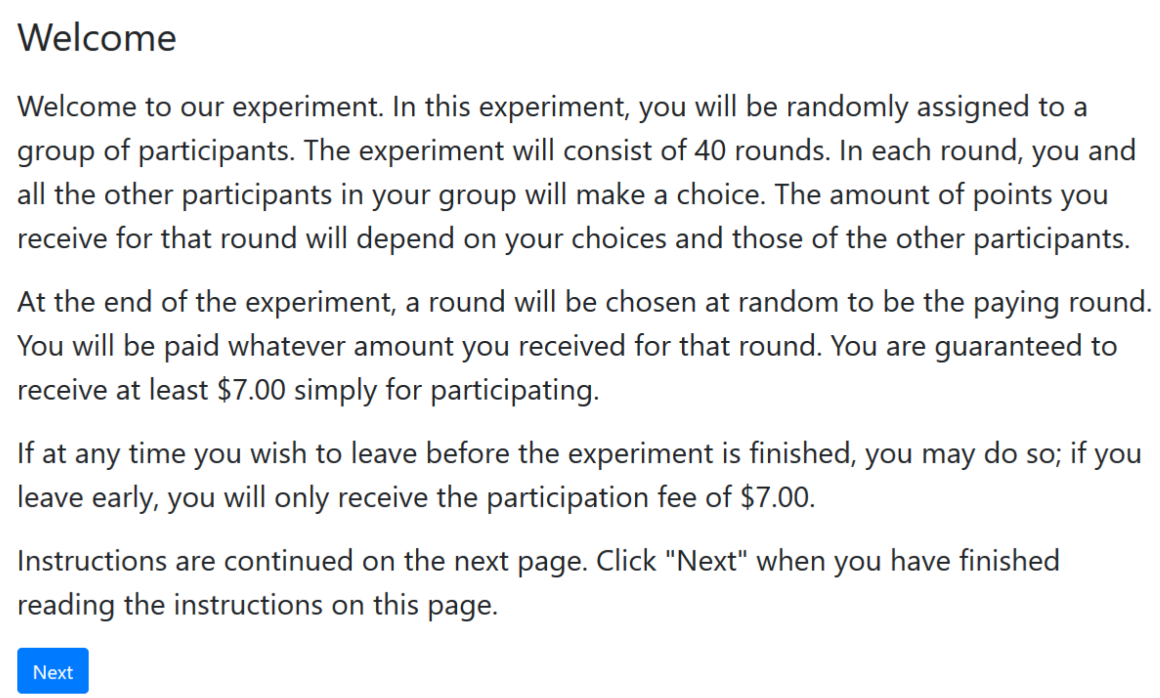
Supplemental Appendix: Experimental Public Goods Games with Progressive Taxation

A.1 Screenshots and Instructions

The following screenshots show the instructions presented to the participants, as well as the interface that they used to make their choices. Other than instructions common to all experiments at the lab (e.g. "Please sign in, then sit down at a computer and silence your cell phone"), these were the only instructions provided to participants. Participants were allowed to ask individual questions in private, but none did so.

Of particular note is the slider interface in figure A.9. As participants moved the slider to make their choices, the figures in the table below it would update in real time to tell them what their contribution, tax payment, and payoff would be as a result of their choice. The initial position was randomized each round in case participants decided to anchor to its initial position; however, regressions showed no evidence that such anchoring occurred.

Figures A.5 and A.7 show the screens participants saw just before the tax-inefficient treatment, while figures A.6 and A.8 show the screens they saw just before the tax-efficient condition. The parameter values shown are from the “low-MPCR” session; they were adjusted accordingly for the other sessions.



Welcome

Welcome to our experiment. In this experiment, you will be randomly assigned to a group of participants. The experiment will consist of 40 rounds. In each round, you and all the other participants in your group will make a choice. The amount of points you receive for that round will depend on your choices and those of the other participants.

At the end of the experiment, a round will be chosen at random to be the paying round. You will be paid whatever amount you received for that round. You are guaranteed to receive at least $7.00 simply for participating.

If at any time you wish to leave before the experiment is finished, you may do so; if you leave early, you will only receive the participation fee of $7.00.

Instructions are continued on the next page. Click "Next" when you have finished reading the instructions on this page.

Next

Figure A.1: Screenshot of instructions, page 1.

Task

You will be assigned to a different group of participants each round, chosen randomly. There will always be a total of 5 participants in your group, including yourself.

In each round, you will be given a certain **starting amount** . This amount may change between rounds. You will have the option to give some or all of your **starting amount** to your group. You will choose how much to give by moving a slider. The initial position of the slider will be set randomly each round.

You will have 1 minute, 15 seconds to make a decision in each round. If time runs out, whatever the slider currently shows will be submitted as your choice.

Any points you choose to give will be multiplied by 1.0, and then shared equally with everyone in the group, including yourself. This means that for each 1.0 point you choose to give, you will get back 0.2 points.

Instructions are continued on the next page. Click "Next" when you have finished reading the instructions on this page.

Next

Figure A.2: Screenshot of instructions, page 2.

Payoffs

The total amount of points you get in each round will be the sum of five components:

- Your **after-tax income**, which is your **starting amount** minus your contribution and any tax you paid
- Your share of the **contributions** that **you** gave to your group.
- Your share of the **tax revenue** collected from **you**.
- Your share of the **contributions** that **others** gave to your group.
- Your share of the **tax revenue** collected from **others** in your group.

The part of your payoff that is the result of your own choices will be referred to as your **net payoff**. Your **total payoff** is your **net payoff** plus what you receive as a result of choices made by others in your group.

Instructions are continued on the next page. Click "Next" when you have finished reading the instructions on this page.

Next

Figure A.3: Screenshot of instructions, page 3.

Taxes

Your **starting amount** minus your **contribution** to your group is your **taxable income**. The following table shows, in increments of 5.0 points, how much tax you would pay, and what you would be left with as after-tax income:

Taxable income	Tax rate	Tax paid	After-tax income
0	0.0 %	0.0	0.0
5	25.0 %	1.25	3.75
10	50.0 %	5.0	5.0
15	75.0 %	11.25	3.75
20	100.0 %	20.0	0.0

Instructions are continued on the next page. Click "Next" when you have finished reading the instructions on this page.

Next

Figure A.4: Screenshot of instructions, page 4.

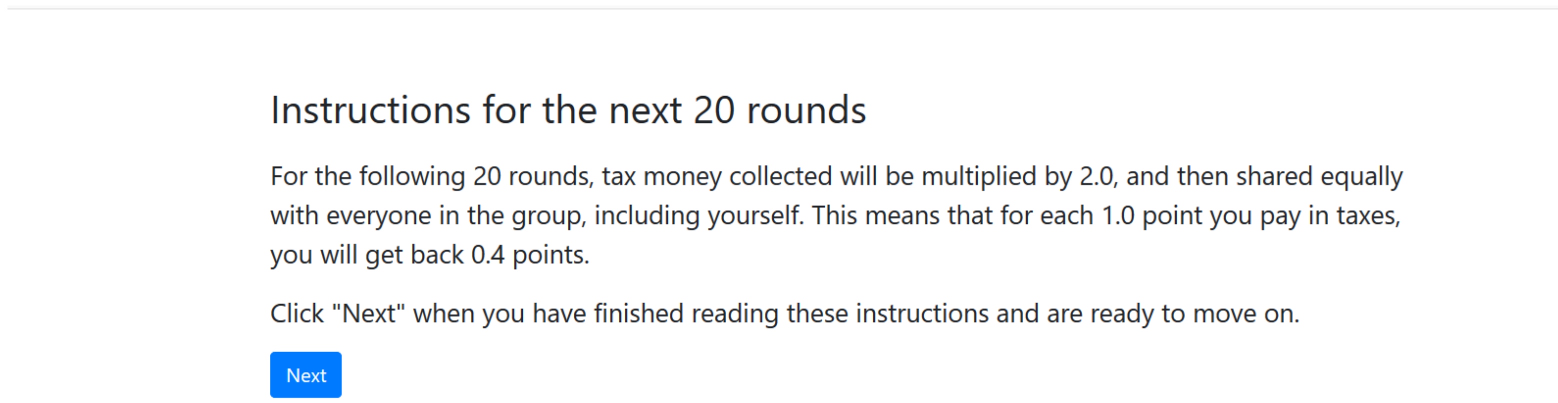
Instructions for the next 20 rounds

For the following 20 rounds, tax money collected is simply lost. Neither you nor anyone in your group will receive any part of it.

Click "Next" when you have finished reading these instructions and are ready to move on.

Next

Figure A.5: Screenshot of instructions, page 5, tax-inefficient condition.



Instructions for the next 20 rounds

For the following 20 rounds, tax money collected will be multiplied by 2.0, and then shared equally with everyone in the group, including yourself. This means that for each 1.0 point you pay in taxes, you will get back 0.4 points.

Click "Next" when you have finished reading these instructions and are ready to move on.

Next

Figure A.6: Screenshot of instructions, page 5, tax-efficient condition.

Summary

The following equations summarize the payoffs:

$$A - C = B$$

$$B - D = E$$

$$E + 0.2\ C = F$$

- (A) Your starting amount
- (C) Your contribution
- (B) Your taxable income
- (D) The tax you pay
- (E) Your after-tax income
- (F) Your net payoff

Your total payoff will be equal to your net payoff, plus whatever you receive in contributions from the group.

Figure A.7: Screenshot of instructions, page 6, tax-inefficient condition.

Summary

The following equations summarize the payoffs:

$$A - C = B$$

$$B - D = E$$

$$E + 0.2\ C + 0.4\ D = F$$

- (A) Your starting amount
- (C) Your contribution
- (B) Your taxable income
- (D) The tax you pay
- (E) Your after-tax income
- (F) Your net payoff

Your total payoff will be equal to your net payoff, plus whatever you receive in contributions and tax revenue from the group.

Figure A.8: Screenshot of instructions, page 6, tax-efficient condition.

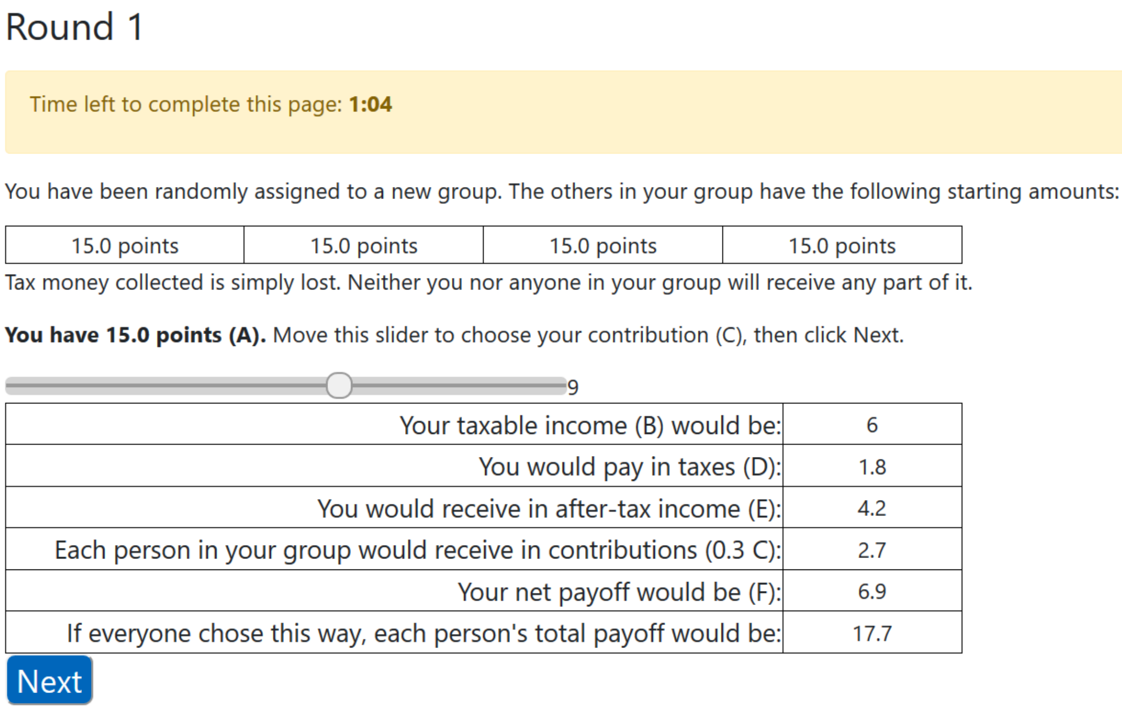


Figure A.9: Screenshot of slider interface

A.2 Robustness checks

As a robustness check, I also ran OLS regressions on the treatment effects and all interactions, as well as participant fixed-effects models for all the within-participant effects. These results are reported in table A.1 for contributions and table A.2 for deviations from the dominant strategy equilibrium. Results are substantially similar to the random-effects models.

Similarly, I ran a participant fixed-effect regression testing for decay of contributions over time, shown in table A.3. Once again the results are substantially similar to the random-effects model.

Table A.1: OLS and Fixed-Effect Models of Contributions

Baseline: High MPCR, tax-inefficient, income = 10	5.185***	5.578***	5.300***
	(0.164)	(0.503)	(0.334)
Income = 15	3.866***	3.809***	3.233***
	(0.207)	(0.195)	(0.422)
Income = 20	6.638***	6.495***	5.950***
	(0.259)	(0.246)	(0.528)
Tax-efficient	-2.782***	-2.880***	-1.722***
	(0.232)	(0.219)	(0.472)
Participant fixed-effects	No	Yes	No
Low MPCR			-0.400
			(0.422)
Reversed order			0.100
			(0.422)
Tax-efficient × income = 15	-2.276***	-2.179***	-1.364**
	(0.293)	(0.276)	(0.597)
Tax-efficient × income = 20	-3.199***	-2.955***	-2.711***
	(0.366)	(0.350)	(0.747)
Low MPCR × income = 15			0.375
			(0.534)
Low MPCR × income = 20			0.470
			(0.668)
Reversed order × income = 15			1.271**
			(0.534)
Reversed order × income = 20			1.320**
			(0.668)
Low MPCR × tax-efficient			-1.204**
			(0.597)
Reversed order × tax-efficient			-1.551***

Table A.2: OLS and Fixed Effects Models of Deviation from DSE

Baseline: High MPCR, tax-inefficient, income = 10	2.569***	3.070***	2.300***
	(0.164)	(0.504)	(0.334)
Income = 15	-1.134***	-1.191***	-1.767***
	(0.207)	(0.196)	(0.422)
Income = 20	-3.362***	-3.505***	-4.050***
	(0.259)	(0.247)	(0.528)
Tax-efficient	-0.167	-0.265	1.278***
	(0.232)	(0.220)	(0.472)
Participant fixed-effects	No	Yes	No
Low MPCR			0.600
			(0.422)
Reversed order			0.100
			(0.422)
Tax-efficient × income = 15	1.339***	1.437***	2.636***
	(0.293)	(0.277)	(0.597)
Tax-efficient × income = 20	0.417	0.661*	1.289*
	(0.366)	(0.350)	(0.747)
Low MPCR × income = 15			0.375
			(0.534)
Low MPCR × income = 20			0.470
			(0.668)
Reversed order × income = 15			1.271**
			(0.534)
Reversed order × income = 20			1.320**
			(0.668)
Low MPCR × tax-efficient			-2.204***
			(0.597)
Reversed order × tax-efficient			-1.551***

Table A.3: Fixed Effects Models Testing Decay of Contributions over Time

	Dependent variable:
	player.contribution
Taxes efficient	-3.025^{***}
	(0.219)
Income = 15	4.306^{***}
	(0.203)
Income = 20	6.495^{***}
	(0.244)
Taxes efficient × income = 15	-2.294^{***}
	(0.277)
Taxes efficient × income = 20	-2.955^{***}
	(0.345)
round	0.095^{***}
	(0.021)
round^2	-0.001^{**}
	(0.001)
Observations	2,600
R^2	0.501
Adjusted R^2	0.487
F Statistic	363.014^{***} (df = 7; 2528)
Note:	*p<0.1; **p<0.05; ***p<0.01

A.3 Finite mixture model regression results

These tables show detailed regression results for each of the types in the finite mixture model.

Type 1: Even splitters (7-19)	Estimate	Std. Error	z-value	$Pr(>\lvert z \rvert)$
(Intercept)	2.82057	0.59089	4.7734	0.00000***
Income = 15	-2.74838	0.73125	-3.7585	0.00017***
Income = 20	-5.17880	0.87422	-5.9239	0.00000***
Taxes efficient	-0.03900	0.84778	-0.0460	0.96330
Taxes efficient × income = 15	3.95512	1.06072	3.7287	0.00019***
Taxes efficient × income = 20	5.87084	1.25335	4.6841	0.00000***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

Table A.4: Coefficient estimates for participants of type 1, "even splitters", in the 8-type finite mixture model.

Type 2: Spiteful (6-10)	Estimate	Std. Error	z-value	$Pr(>\lvert z \rvert)$
(Intercept)	-0.89931	0.37725	-2.3838	0.01713*
Income = 15	-0.20825	0.47617	-0.4373	0.66187
Income = 20	-3.38252	0.61003	-5.5448	0.00000***
Taxes efficient	2.94893	0.52284	5.6402	0.00000***
Taxes efficient × income = 15	2.55190	0.66762	3.8224	0.00013***
Taxes efficient × income = 20	1.99974	0.86939	2.3002	0.02144*

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

Table A.5: Coefficient estimates for participants of type 2, "spiteful", in the 8-type finite mixture model.

Type 3: Selfish (7)	Estimate	Std. Error	z-value	$Pr(>\|z\|)$
(Intercept)	0.11765	0.146546	0.8028	0.42209
Income = 15	-0.01765	0.192671	-0.0916	0.92702
Income = 20	0.19814	0.281284	0.7044	0.48117
Taxes efficient	0.02941	0.231709	0.1269	0.89899
Taxes efficient × income = 15	-0.25798	0.291517	-0.8850	0.37617
Taxes efficient × income = 20	-0.76187	0.376507	-2.0235	0.04302*

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

Table A.6: Coefficient estimates for participants of type 3, "selfish", in the 8-type finite mixture model.

Type 4: Naive (6)	Estimate	Std. Error	z-value	$Pr(>\|z\|)$
(Intercept)	2.63198	0.25180	10.4526	0.00000***
Income = 15	-3.00205	0.32615	-9.2045	0.00000***
Income = 20	-6.37208	0.44678	-14.2623	0.00000***
Taxes efficient	3.33635	0.36528	9.1337	0.00000***
Taxes efficient × income = 15	2.83655	0.47317	5.9948	0.00000***
Taxes efficient × income = 20	3.46128	0.63679	5.4355	0.00000***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

Table A.7: Coefficient estimates for participants of type 4, "naive", in the 8-type finite mixture model.

Type 5: Confused? (4)	Estimate	Std. Error	z-value	$Pr(>\|z\|)$
(Intercept)	4.89773	0.54361	9.0097	0.00000***
Income = 15	-1.44453	0.65943	-2.1906	0.02848*
Income = 20	-4.13682	0.75054	-5.5118	0.00000***
Taxes efficient	-2.00468	0.71641	-2.7982	0.00514**
Taxes efficient × income = 15	4.92945	0.89092	5.5330	0.00000***
Taxes efficient × income = 20	3.67232	1.08574	3.3823	0.00072***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

Table A.8: Coefficient estimates for participants of type 5, "confused", in the 8-type finite mixture model.

Type 6: Giving target 1 (19-25)	Estimate	Std. Error	z-value	$Pr(>\|z\|)$
(Intercept)	2.95004	0.25643	11.5044	0.00000***
Income = 15	-1.64300	0.31705	-5.1822	0.00000***
Income = 20	-3.55250	0.38718	-9.1754	0.00000***
Taxes efficient	0.13389	0.34525	0.3878	0.69817
Taxes efficient × income = 15	1.68141	0.43580	3.8582	0.00011***
Taxes efficient × income = 20	-1.52130	0.55073	-2.7624	0.00574**

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

Table A.9: Coefficient estimates for participants of type 6, "giving target 1", in the 8-type finite mixture model.

Type 7: Generous (12-15)	Estimate	Std. Error	z-value	$Pr(>\|z\|)$
(Intercept)	4.36761	0.35560	12.2823	0.00000***
Income = 15	0.50656	0.44040	1.1502	0.25005
Income = 20	-2.46663	0.54181	-4.5526	0.00001***
Taxes efficient	-3.82121	0.50223	-7.6085	0.00000***
Taxes efficient × income = 15	-1.28148	0.62382	-2.0543	0.03995*
Taxes efficient × income = 20	-1.05955	0.76271	-1.3892	0.16477

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

Table A.10: Coefficient estimates for participants of type 7, "generous", in the 8-type finite mixture model.

Type 8: Giving target 2 (4)	Estimate	Std. Error	z-value	$Pr(>\|z\|)$
(Intercept)	3.63333	0.35030	10.3722	0.00000***
Income = 15	-2.70830	0.46341	-5.8443	0.00000***
Income = 20	-6.53326	0.70063	-9.3249	0.00000***
Taxes efficient	-2.34170	0.52543	-4.4567	0.00001***
Taxes efficient × income = 15	1.26665	0.67835	1.8673	0.06187.
Taxes efficient × income = 20	0.24161	0.93506	0.2584	0.79611

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

Table A.11: Coefficient estimates for participants of type 8, "giving target 2", in the 8-type finite mixture model.

A.4 Empirical distribution functions of contributions

Empirical distribution functions for each treatment are shown in figures A.10 through A.27. The vertical blue line indicates the dominant strategy contribution.

There is significant probability mass at the dominant strategy contribution in most treatments, but also significant mass elsewhere, and the distribution around the dominant strategy is not symmetric. Many of the "tax-inefficient" treatments have significant mass at the maximum contribution, and many of the "tax-efficient" treatments have significant mass at the minimum contribution; in each case, this is the Pareto-efficient contribution level.

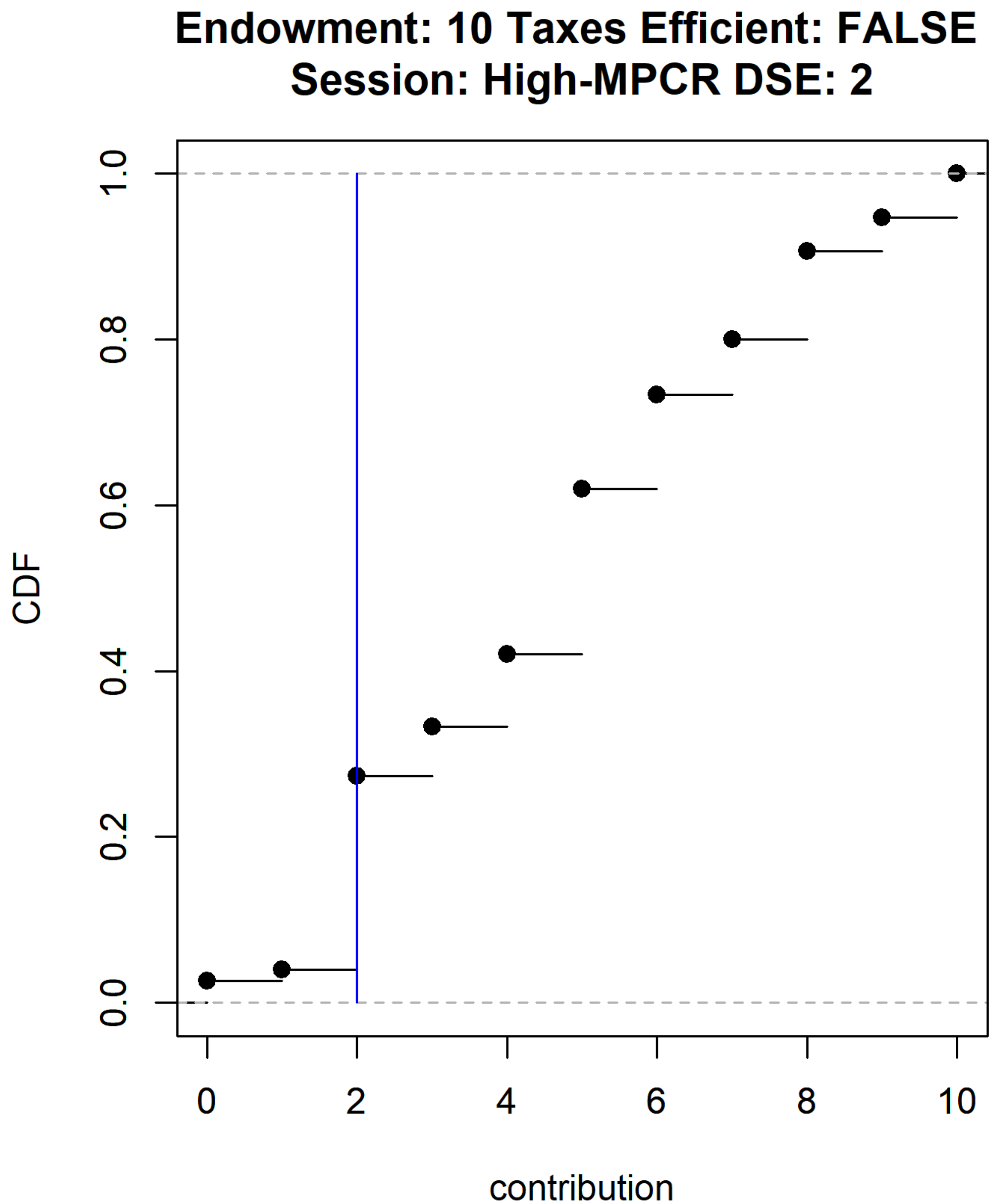


Figure A.10: Empirical distribution function of contributions with income = 10, tax-inefficient, low-MPCR session

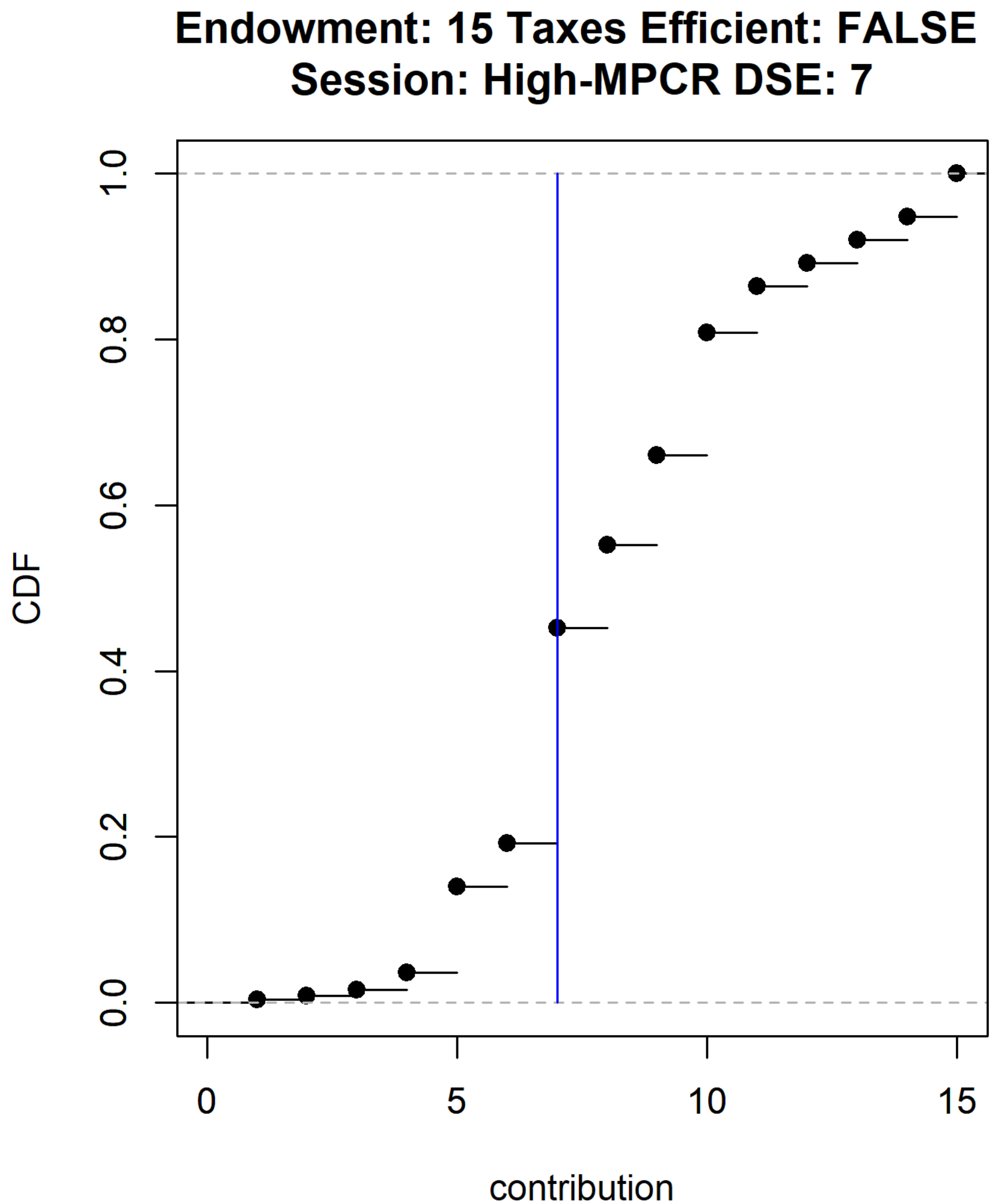


Figure A.11: Empirical distribution function of contributions with income = 15, tax-inefficient, low-MPCR session

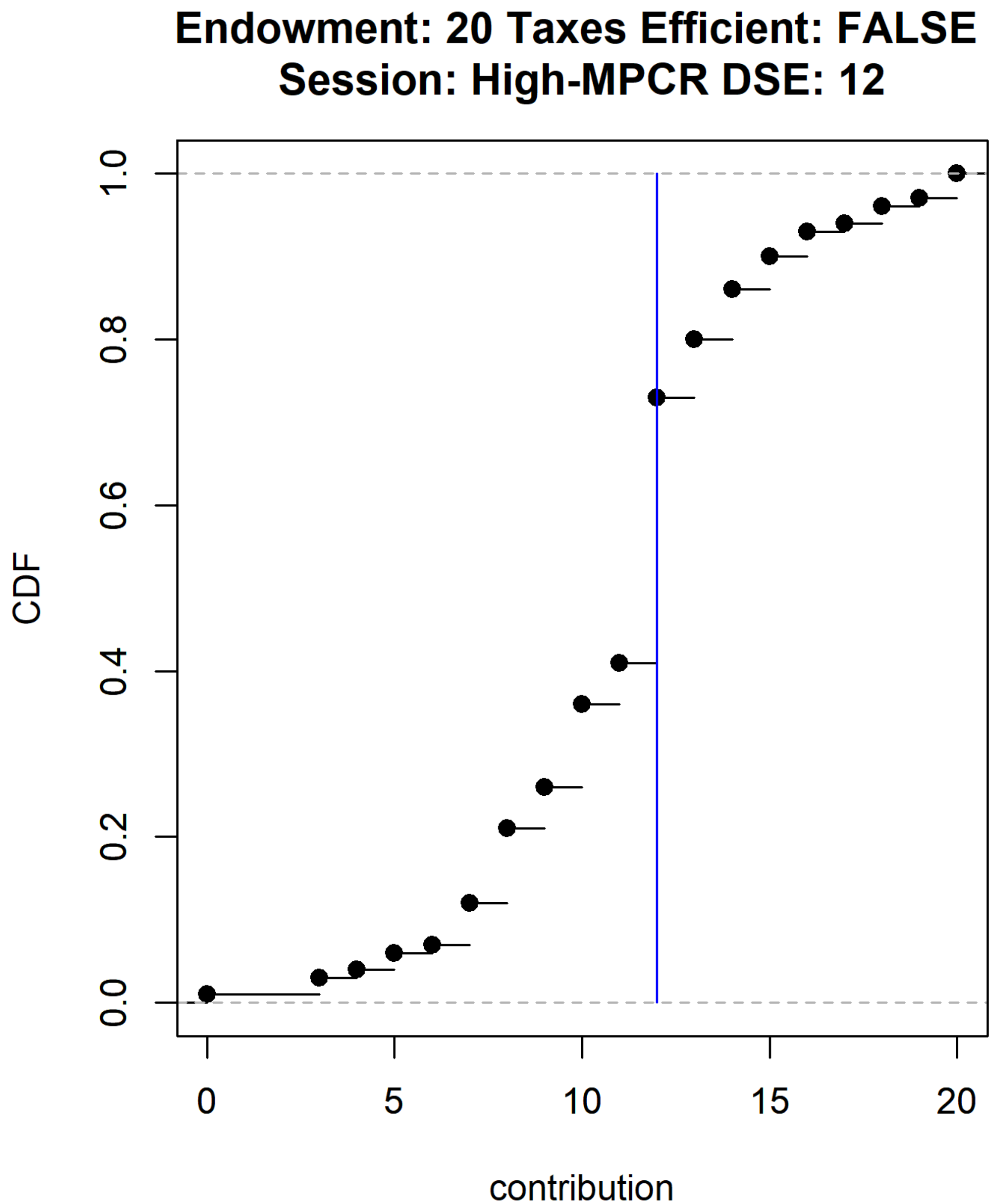


Figure A.12: Empirical distribution function of contributions with income = 20, tax-inefficient, low-MPCR session

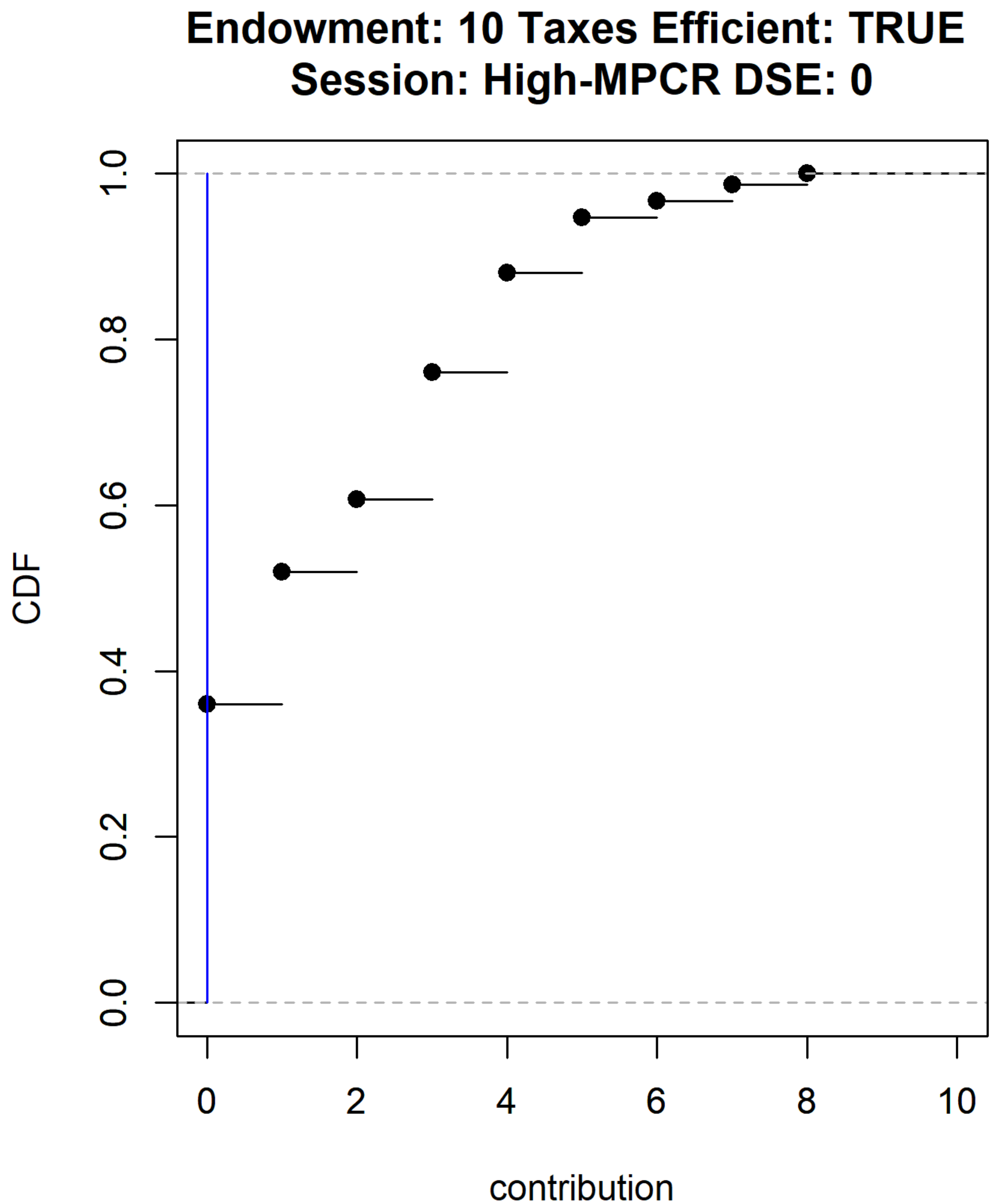


Figure A.13: Empirical distribution function of contributions with income = 10, tax-efficient, low-MPCR session

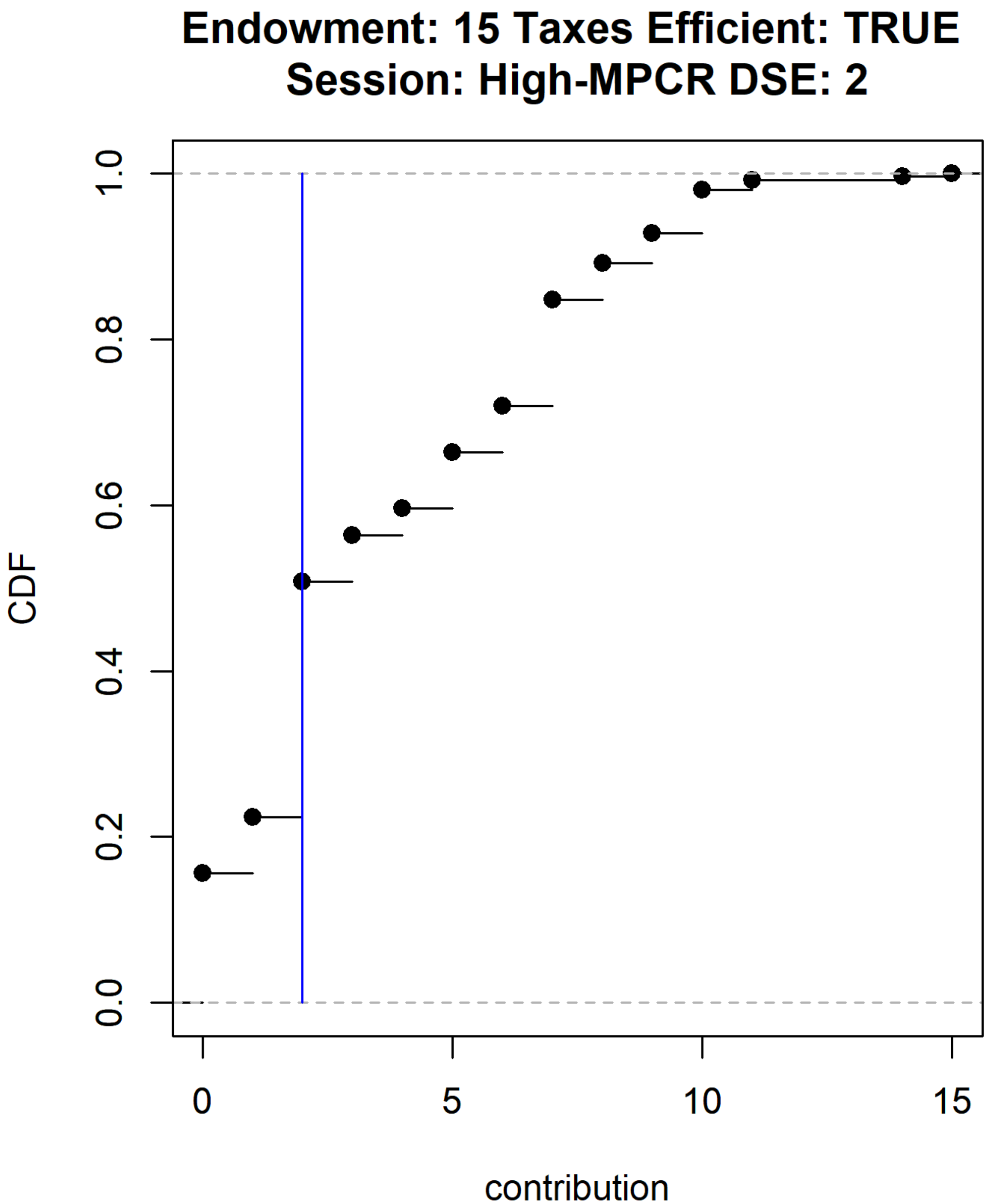


Figure A.14: Empirical distribution function of contributions with income = 15, tax-efficient, low-MPCR session

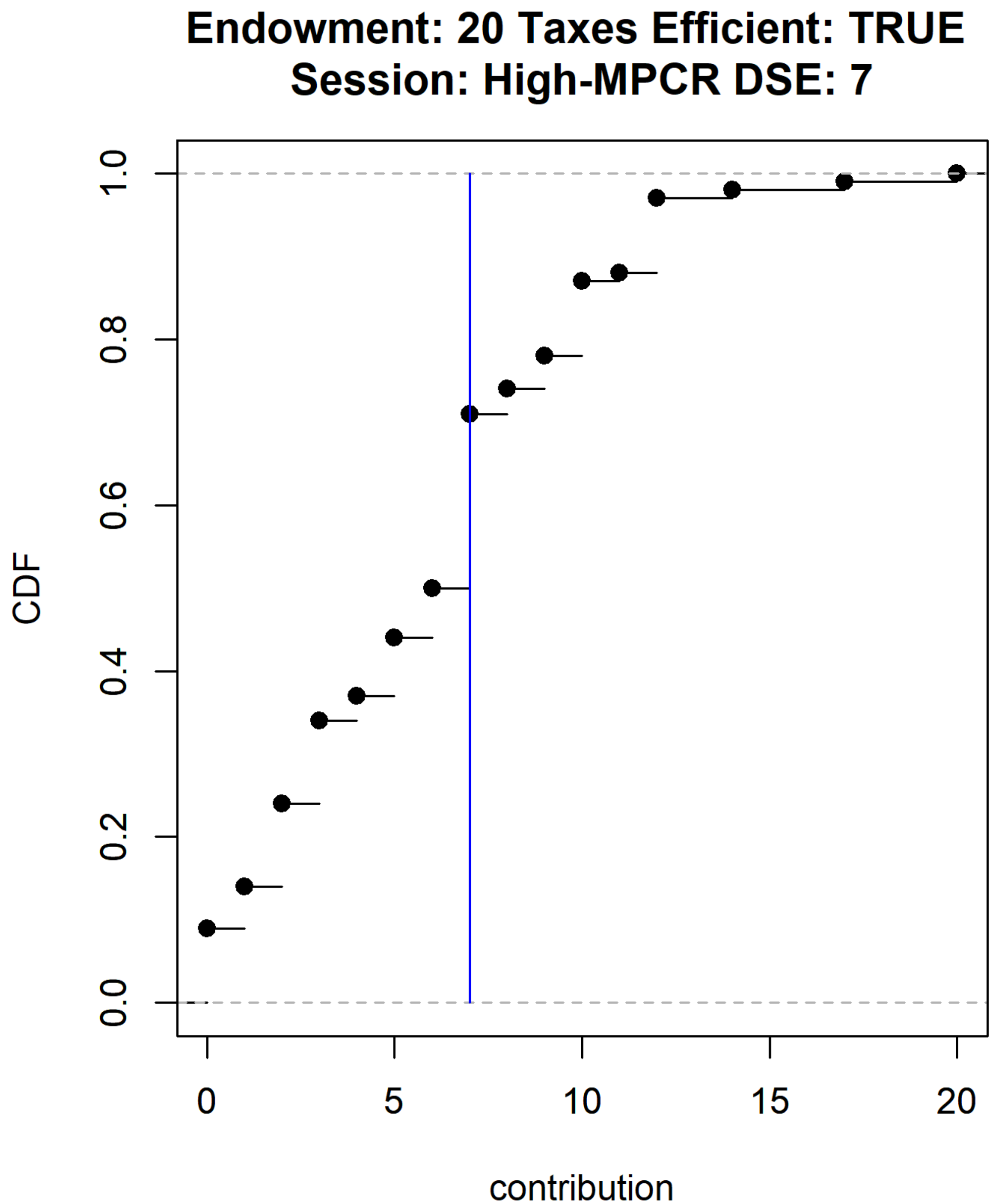


Figure A.15: Empirical distribution function of contributions with income = 20, tax-efficient, low-MPCR session

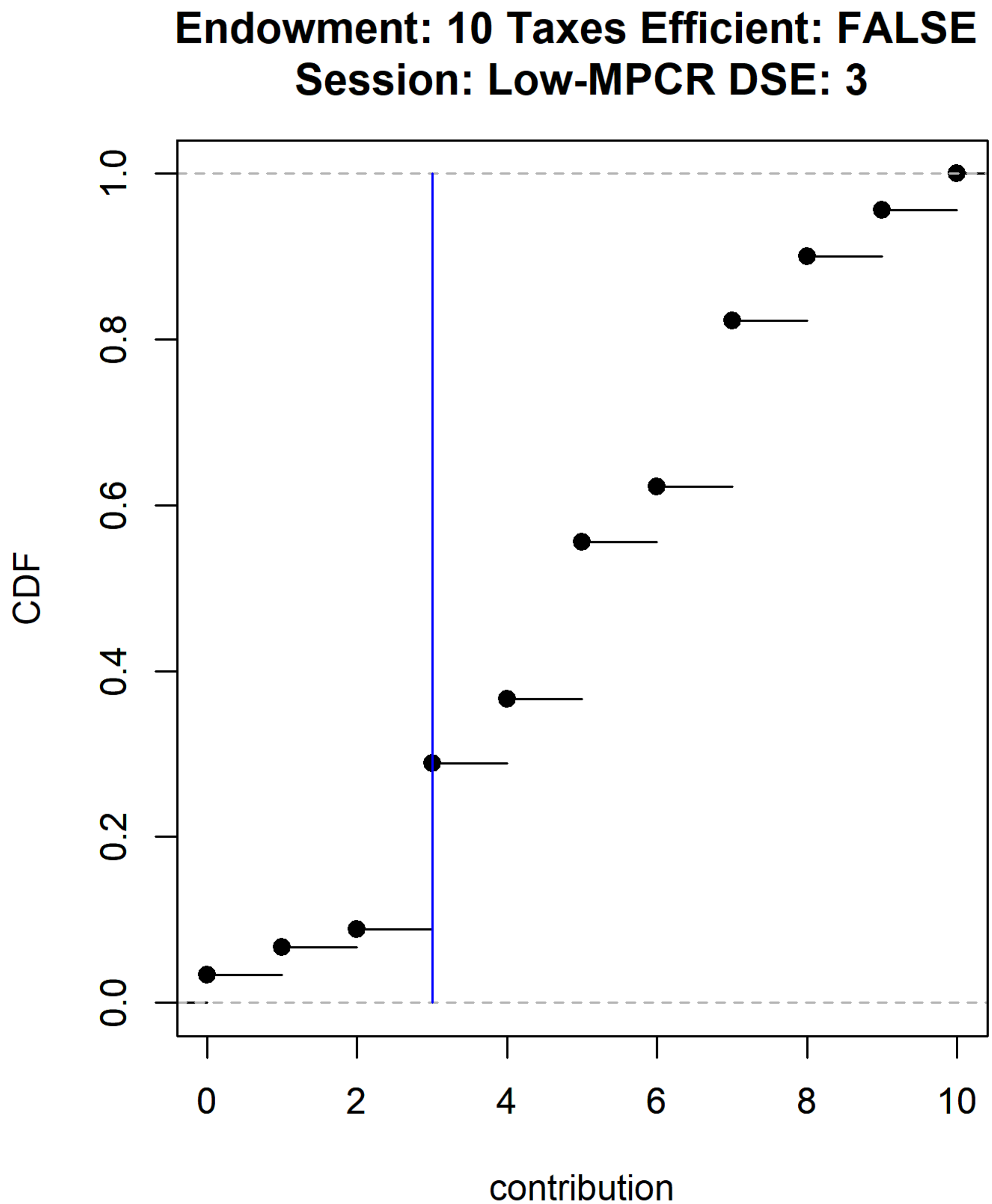


Figure A.16: Empirical distribution function of contributions with income = 10, tax-inefficient, high-MPCR session

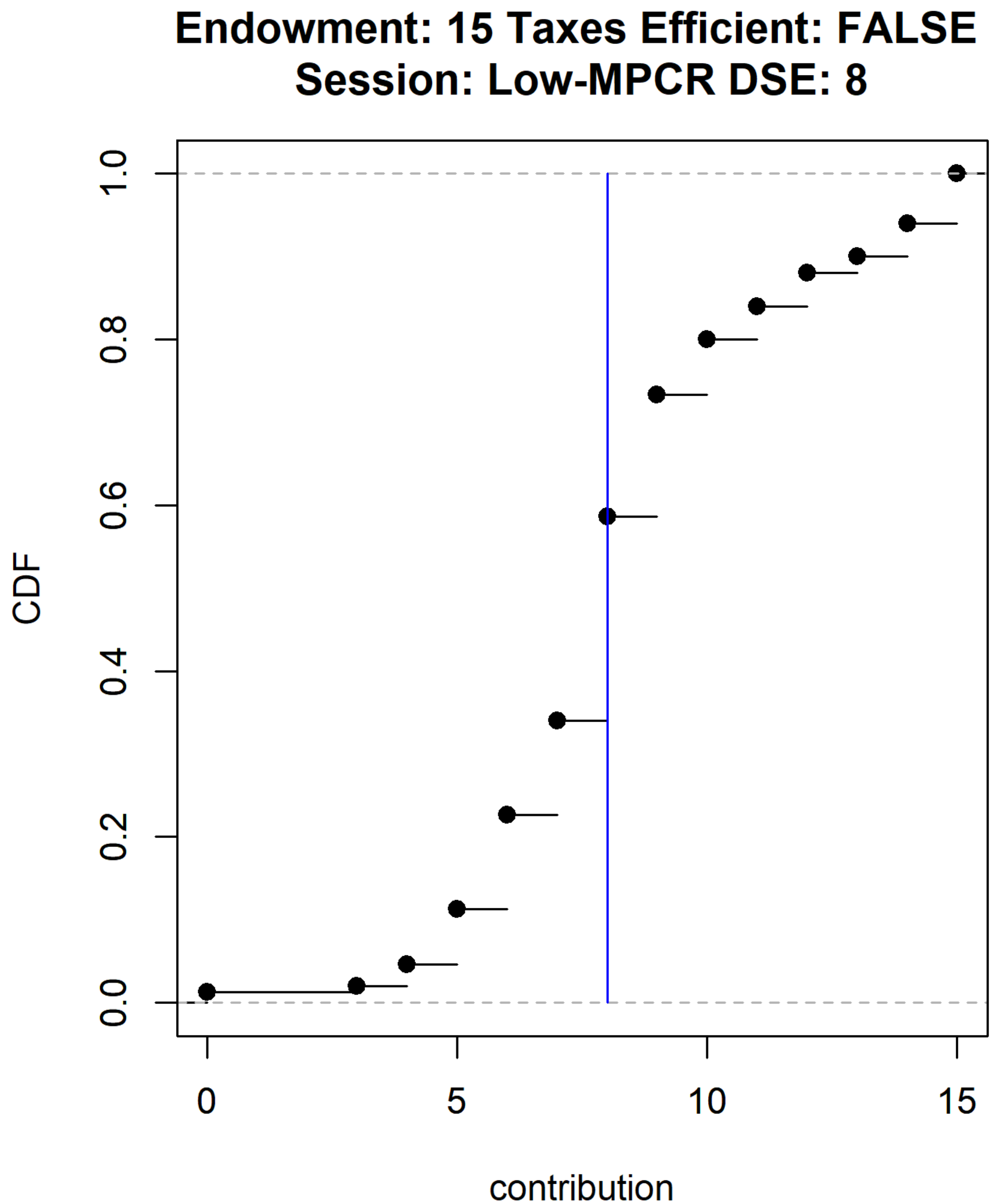


Figure A.17: Empirical distribution function of contributions with income = 15, tax-inefficient, high-MPCR session

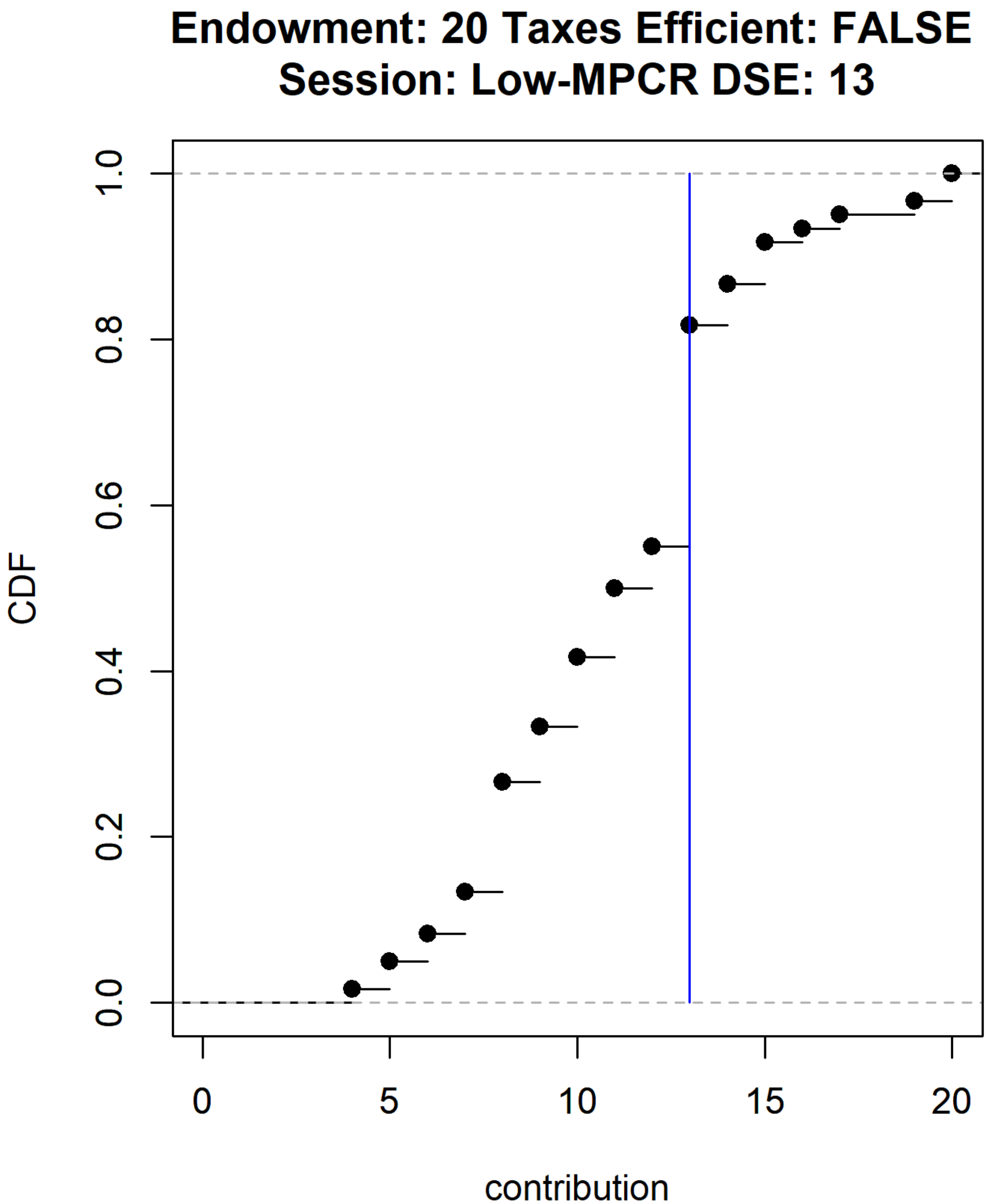


Figure A.18: Empirical distribution function of contributions with income = 20, tax-inefficient, high-MPCR session

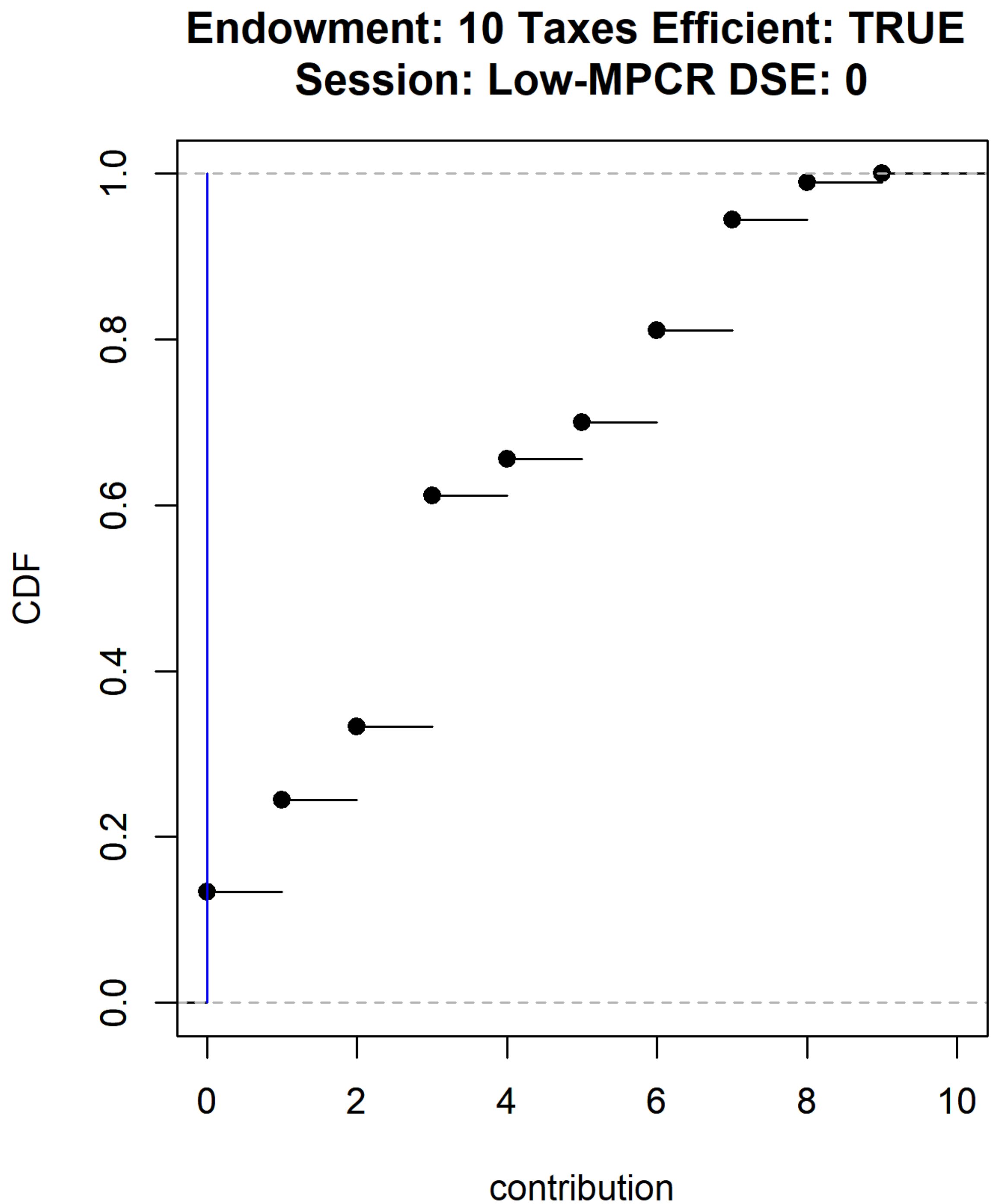


Figure A.19: Empirical distribution function of contributions with income = 10, tax-efficient, high-MPCR session

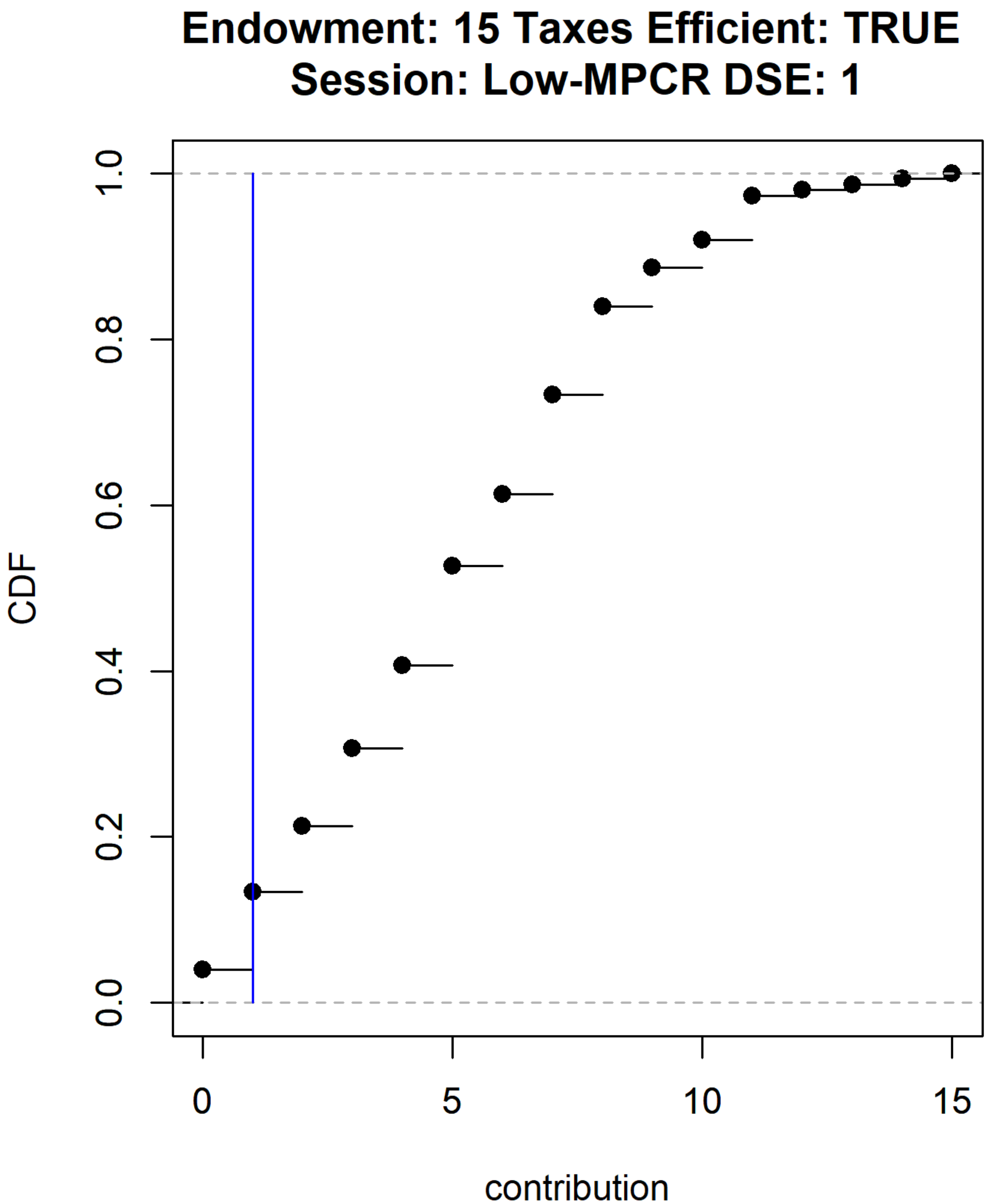


Figure A.20: Empirical distribution function of contributions with income = 15, tax-efficient, high-MPCR session

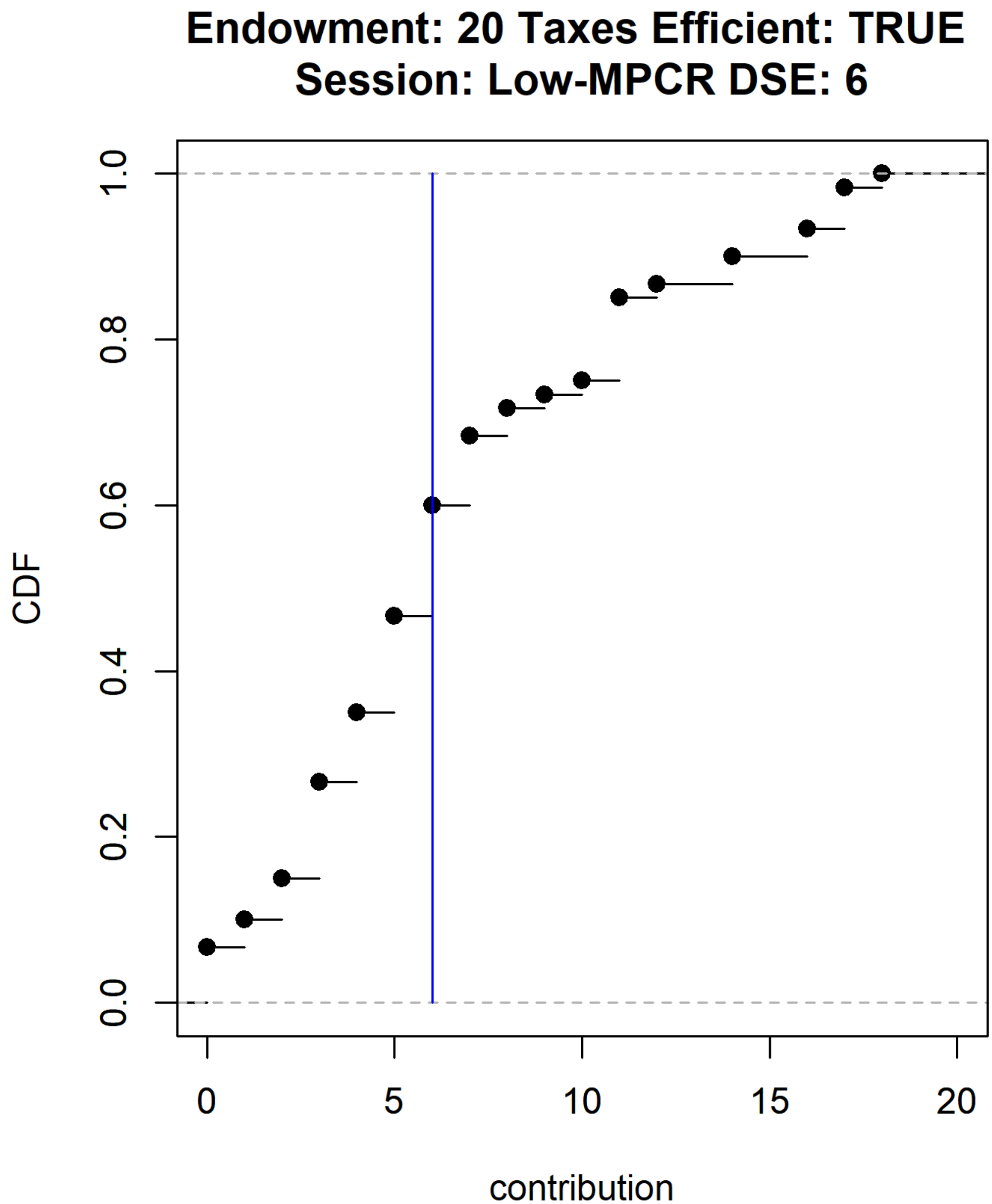


Figure A.21: Empirical distribution function of contributions with income = 20, tax-efficient, high-MPCR session

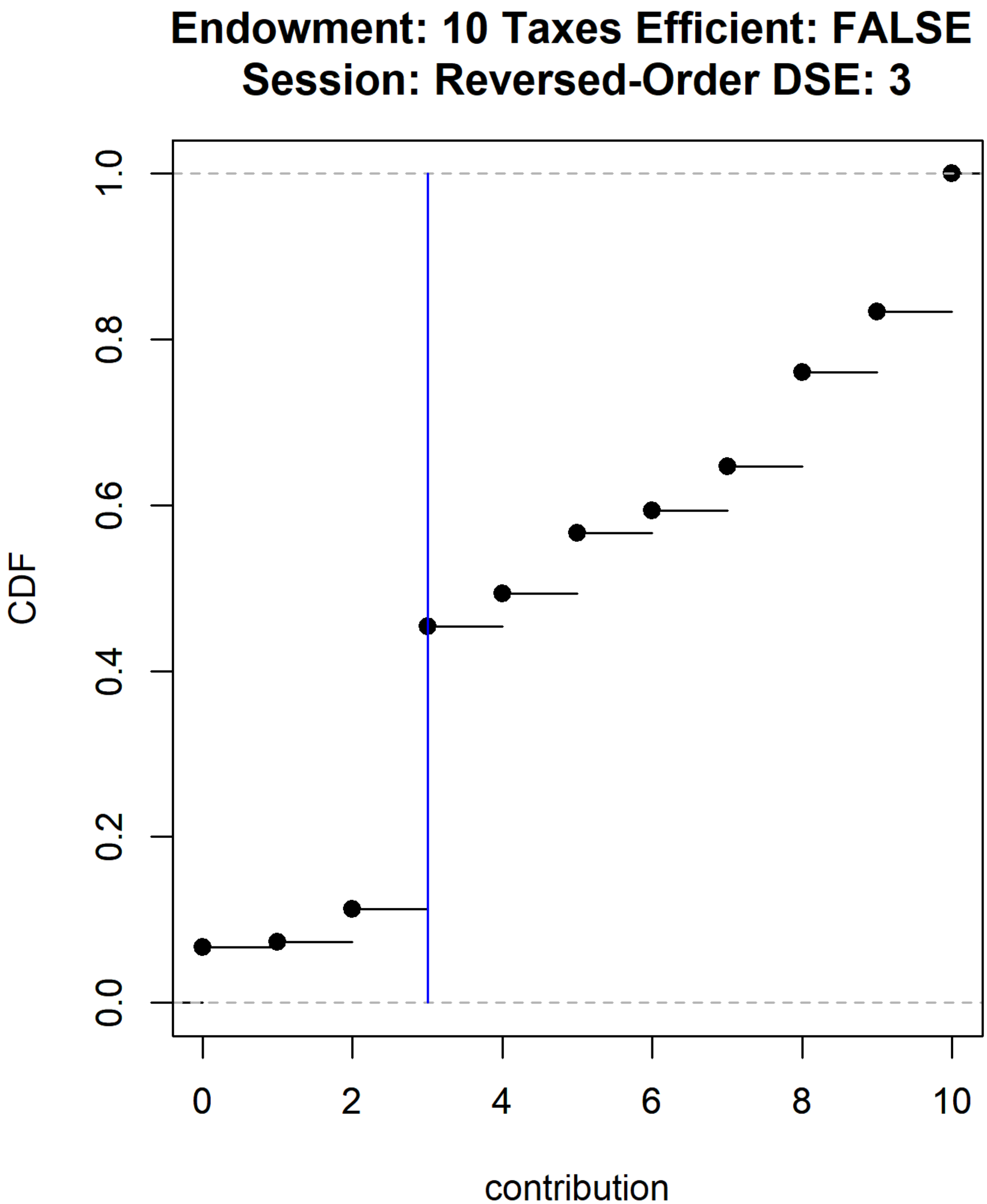


Figure A.22: Empirical distribution function of contributions with income = 10, tax-inefficient, reversed-order session

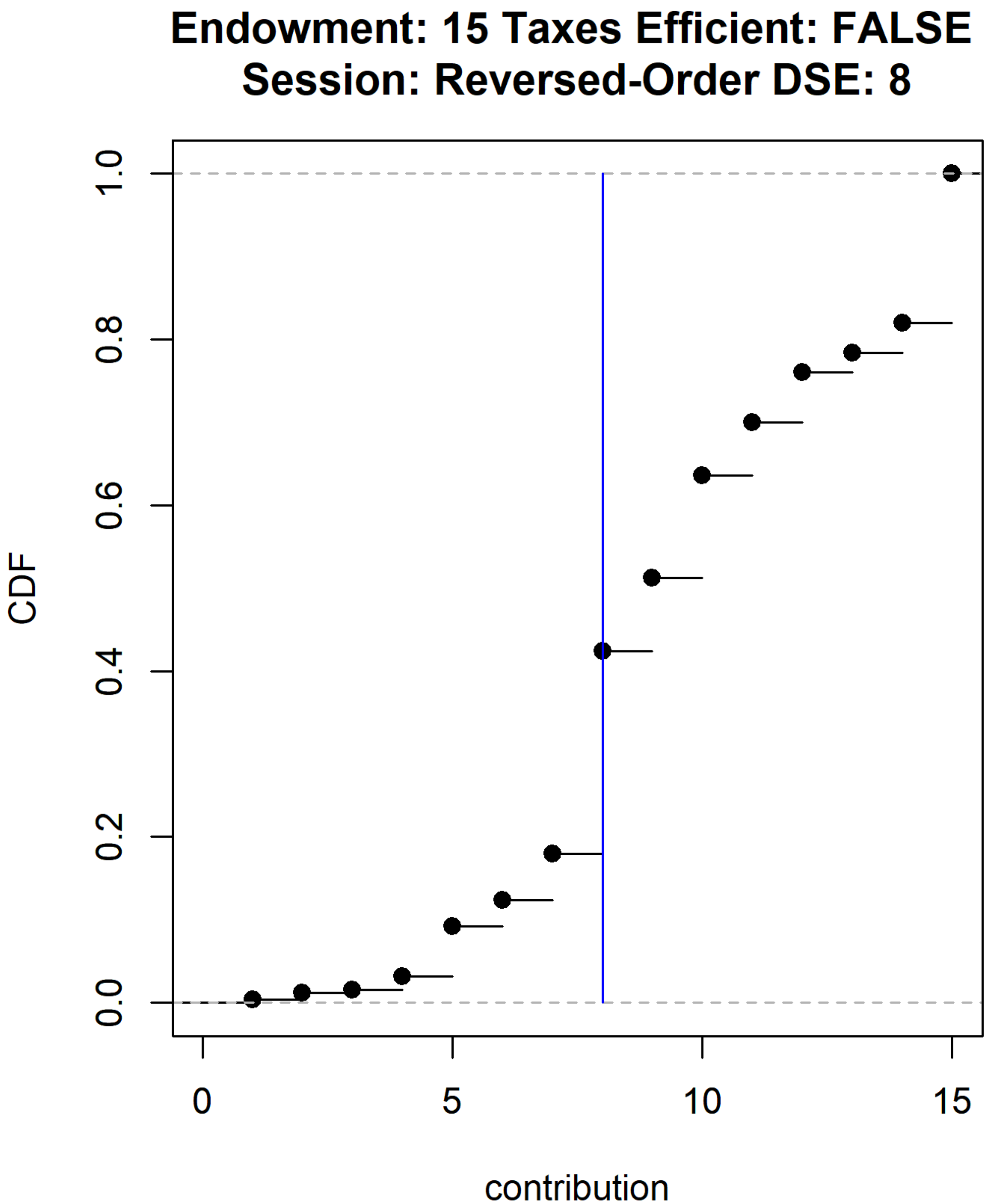


Figure A.23: Empirical distribution function of contributions with income = 15, tax-inefficient, reversed-order session

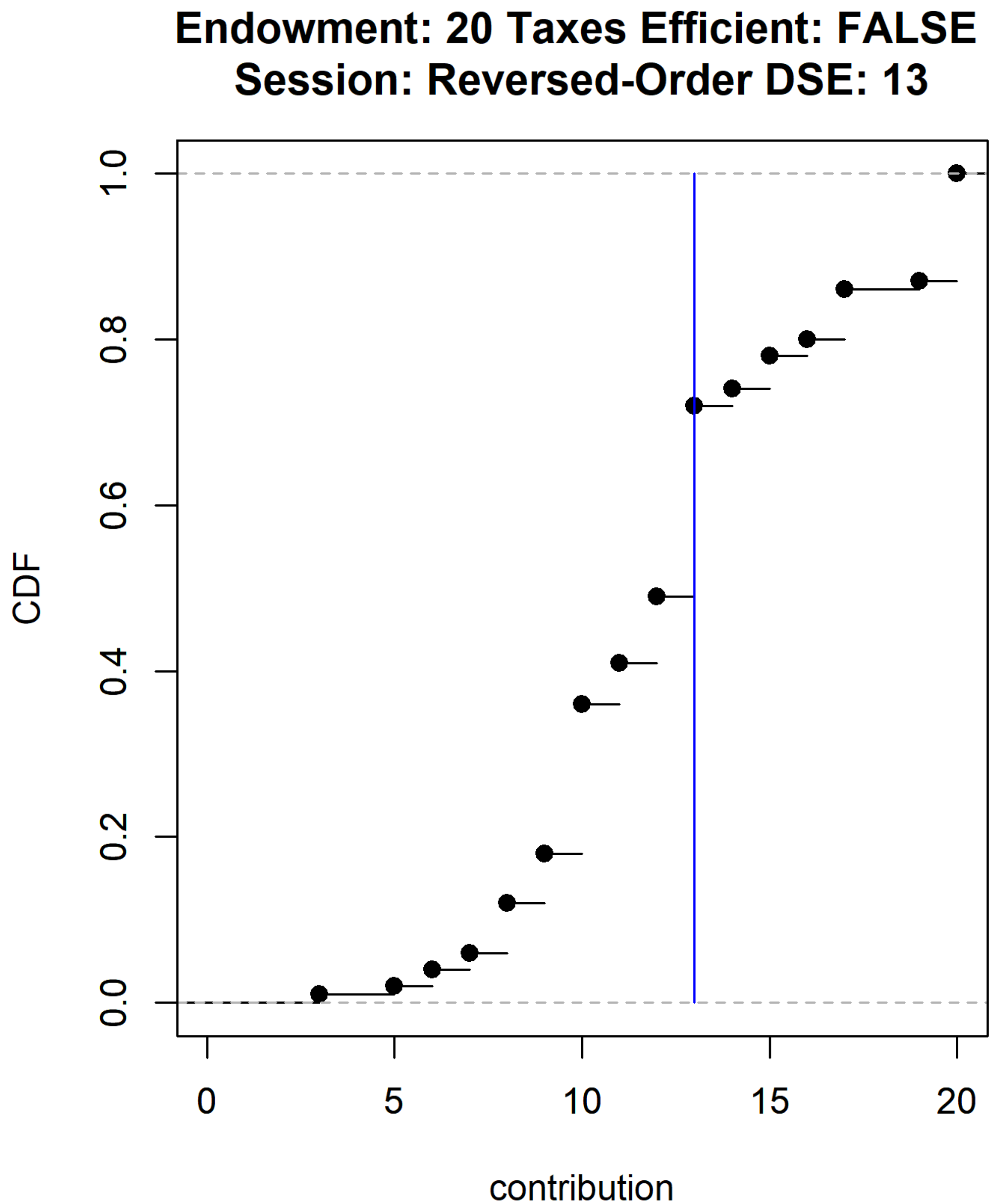


Figure A.24: Empirical distribution function of contributions with income = 20, tax-inefficient, reversed-order session

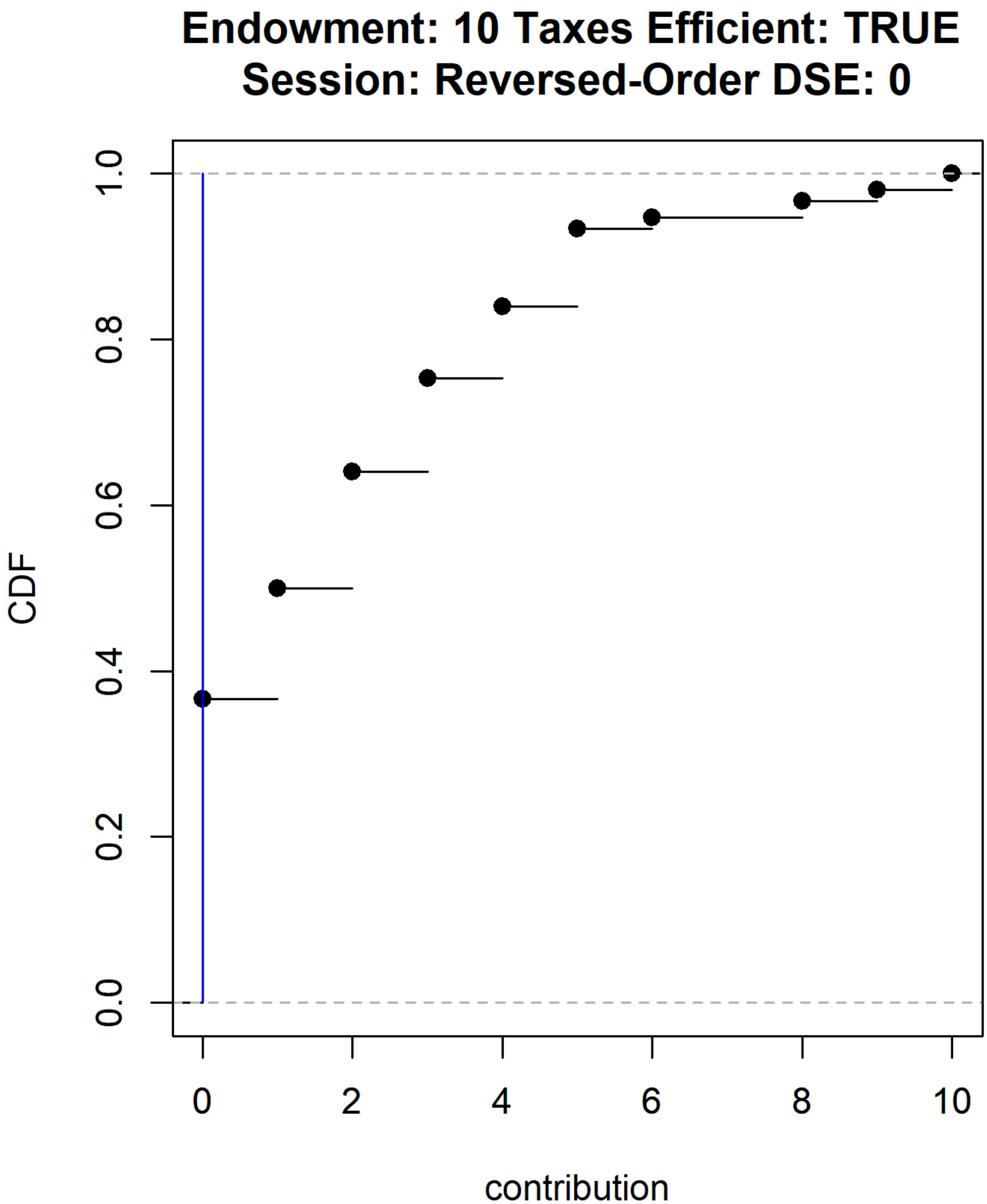


Figure A.25: Empirical distribution function of contributions with income = 10, tax-efficient, reversed-order session

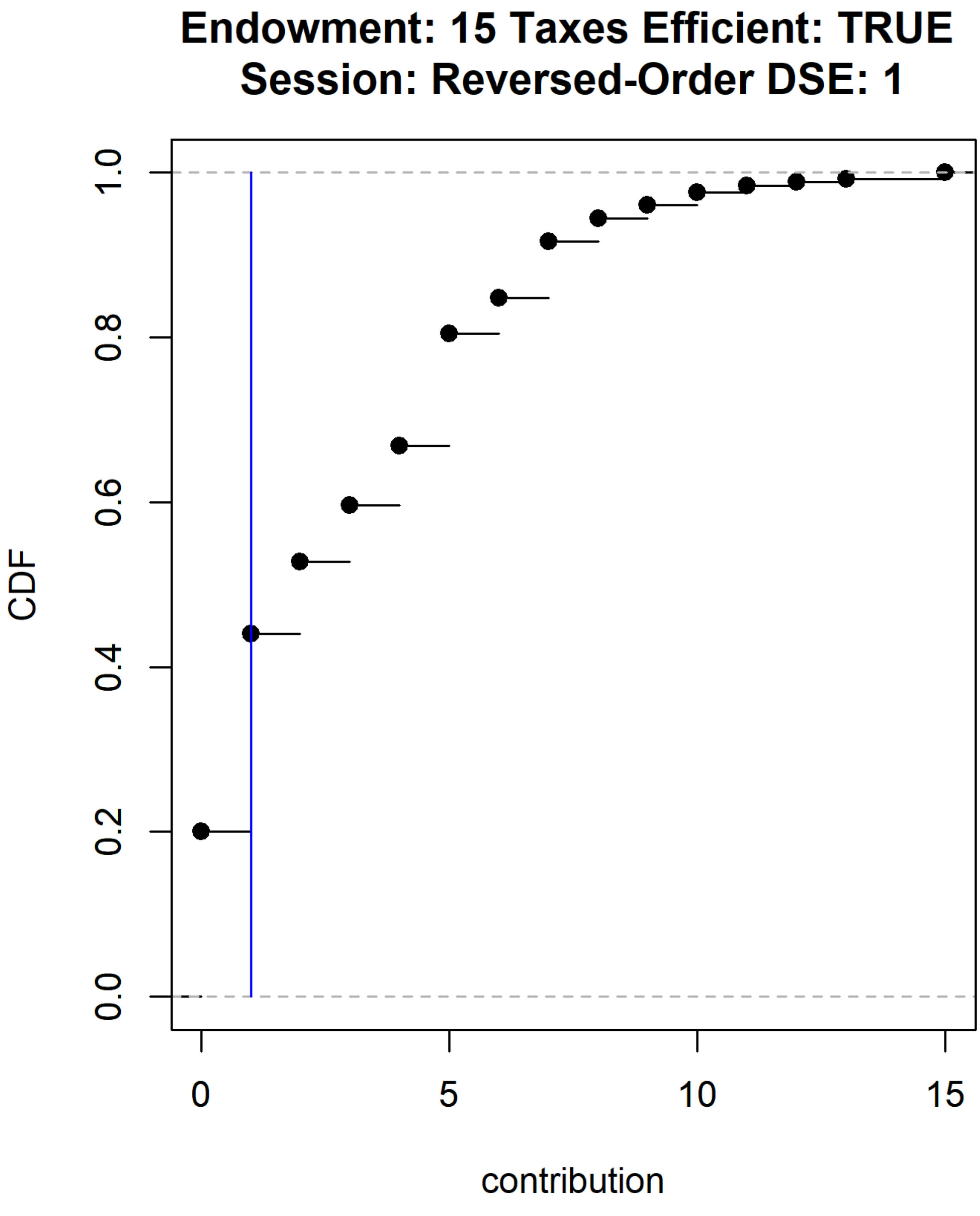


Figure A.26: Empirical distribution function of contributions with income = 15, tax-efficient, reversed-order session

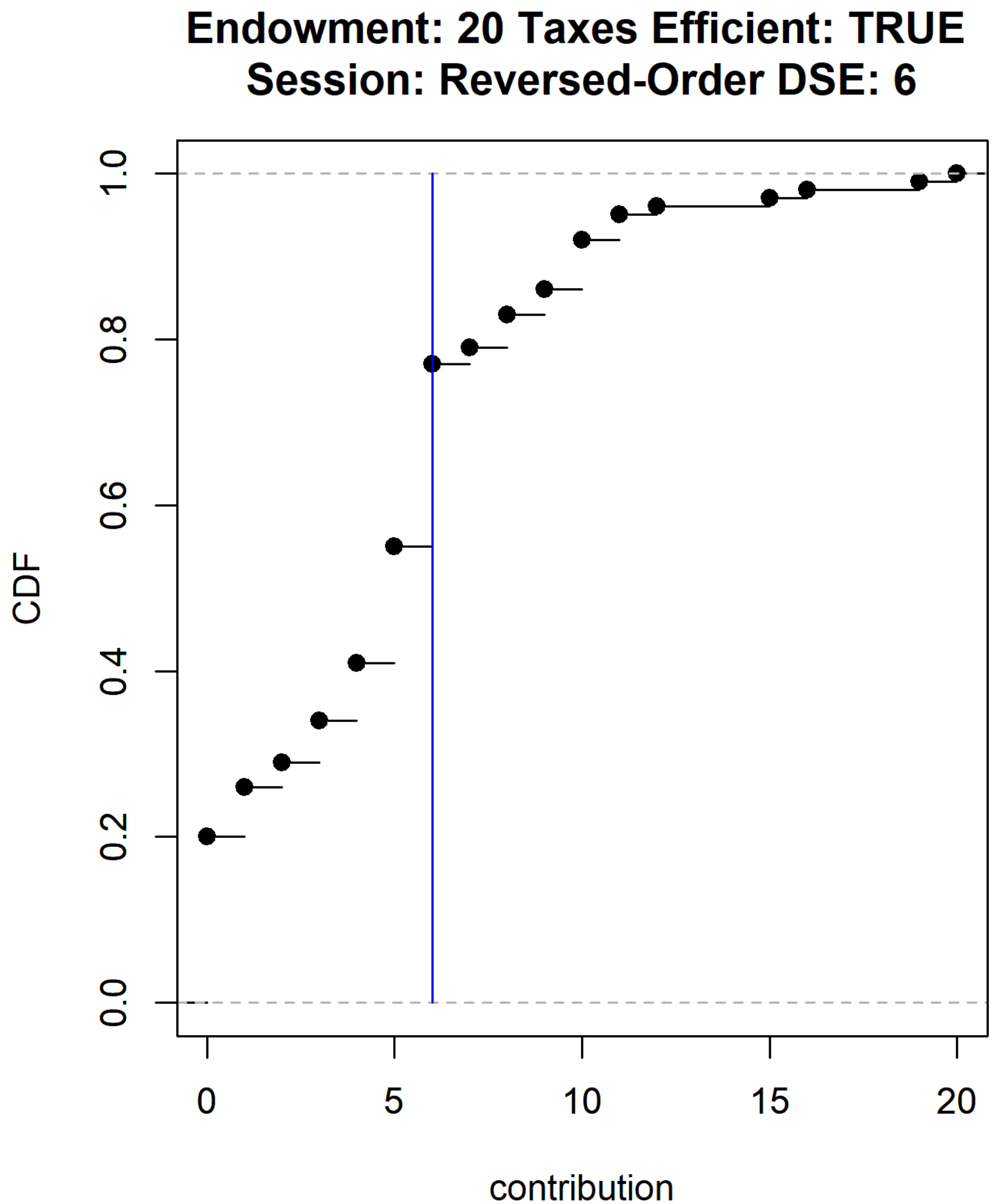


Figure A.27: Empirical distribution function of contributions with income = 20, tax-efficient, reversed-order session

Appendix B

Supplemental Appendix: International Norms of Peace, War, and Sanction

B.1 Proofs

B.1.1 Proof of proposition 2.3

Proposition 2.3. Differentiate Δ_i with respect to α:

$$\frac{\partial \Delta_i}{\partial \alpha} = \frac{R_i^2 Y_i}{R_i Y_i + R_j Y_j} + \ln\left[1 - \frac{\kappa_i}{K_i}\right] + \ln(1 - \omega)$$

The first term is strictly positive, and strictly less than R_i. The other two terms are strictly negative.

The first term dominates if and only if:

$$R_i P_i[\text{win}] > -\ln\left[1 - \frac{\kappa_i}{K_i}\right] - \ln(1-\omega)$$

The right-hand side is constant in R_i. The left-hand side is strictly increasing in R_i, bounded at zero below, and unbounded above.

Therefore for any given value of the other parameters, there exists some threshold $\overline{R}$ such that for any $R_i < \overline{R}$, Δ_i is decreasing in α, while for any $R_i > \overline{R}$, Δ_i is decreasing in α.

□

B.1.2 Proof of proposition 2.5

Proposition 2.5. Note that R_i depends on K_i, so we must consider $\frac{\partial R_i}{\partial K_i}$; but R_j does not depend upon K_i.

$$R_i = \ln\left[\frac{K_i + \kappa_j}{K_i - \kappa_i}\right]$$

$$\frac{\partial R_i}{\partial K_i} = \frac{1}{K_i + \kappa_j} - \frac{1}{K_i - \kappa_i}$$

$$\frac{\partial R_i}{\partial K_i} = -\frac{\kappa_i + \kappa_j}{(K_i + \kappa_j)(K_i - \kappa_i)} < 0$$

$$g_i^* = R_i Y_i \frac{R_i Y_i R_j Y_j}{(R_i Y_i + R_j Y_j)^2}$$

$$g_i^* = R_j Y_j \frac{1}{(1 + \frac{R_j Y_j}{R_i Y_i})^2}$$

$$\frac{\partial g_i^*}{\partial K_i} = R_j Y_j \left(-\frac{2}{\left(1 + \frac{R_j Y_j}{R_i Y_i}\right)^3} \right) \left(-\frac{R_j Y_j}{(R_i Y_i)^2} \right) \left[R_i \frac{\partial Y_i}{\partial K_i} + \frac{\partial R_i}{\partial K_i} Y_i \right]$$

$$\frac{\partial Y_i}{\partial K_i} = \frac{\partial}{\partial K_i} \left[K_i^\alpha L_i^{1-\alpha} \right] = \alpha K_i^{\alpha-1} L_i^{1-\alpha} = \frac{\alpha}{K_i} Y_i$$

$$\frac{\partial g_i^*}{\partial K_i} = \frac{2}{R_i Y_i + R_j Y_j} Y_i \left[\frac{\alpha}{K_i} + \frac{1}{R_i} \frac{\partial R_i}{\partial K_i} \right]$$

Since $\frac{\alpha Y_i}{K_i} > 0$ but $\frac{\partial R_i}{\partial K_i} < 0$, the overall sign of this derivative can vary. There are two countervailing effects: As capital increases, GDP increases, and thus it becomes less costly (in utility) to spend more. But a larger capital stock means a lower return on capital, and so the prize for winning also becomes less valuable.

Ignoring strictly positive terms:

$$\frac{\partial g_i^*}{\partial K_i} \propto Y_i \left[\frac{\alpha}{K_i} - \frac{1}{R_i} \frac{\kappa_i + \kappa_j}{(K_i + \kappa_j)(K_i - \kappa_i)} \right]$$

Thus optimal military spending is increasing in own capital if and only if:

$$\frac{\alpha}{K_i} R_i > \frac{\kappa_i + \kappa_j}{(K_i + \kappa_j)(K_i - \kappa_i)}$$

Now consider the effect of the other country's capital stock.

$$\frac{\partial R_j}{\partial K_j} = -\frac{\kappa_i + \kappa_j}{(K_j + \kappa_i)(K_j - \kappa_j)} < 0$$

$$\frac{\partial}{\partial K_j} [R_j Y_j] = Y_i \left[\frac{\alpha}{K_j} + \frac{1}{R_j} \frac{\partial R_j}{\partial K_j} \right]$$

$$\frac{\partial g_i^*}{\partial K_j} = \frac{\partial}{\partial K_j} \left[\frac{R_j Y_j}{\left(1 + \frac{R_j Y_j}{R_i Y_i}\right)^2} \right]$$

$$\frac{\partial g_i^*}{\partial K_j} = \frac{\left(1+\frac{R_jY_j}{R_iY_i}\right)^2 \frac{\partial}{\partial K_j}\left[R_jY_j\right] - 2R_jY_j\left(1+\frac{R_jY_j}{R_iY_i}\right)\frac{1}{R_iY_i}}{\left(1+\frac{R_jY_j}{R_iY_i}\right)^4}$$

$$\frac{\partial g_i^*}{\partial K_j} = \frac{\left(1+\frac{R_jY_j}{R_iY_i}\right)^2 Y_j\left[\frac{\alpha}{K_j}+\frac{1}{R_j}\frac{\partial R_j}{\partial K_j}\right] - 2\frac{R_jY_j}{R_iY_i}\frac{\partial Y_j}{\partial K_j}\left(1+\frac{R_jY_j}{R_iY_i}\right)}{\left(1+\frac{R_jY_j}{R_iY_i}\right)^4}$$

The denominator is strictly positive, so consider only the numerator:

$$\frac{\partial g_i^*}{\partial K_j} \propto \left(1+\frac{R_jY_j}{R_iY_i}\right)\left[\left(1+\frac{R_jY_j}{R_iY_i}\right)Y_j\left[\frac{\alpha}{K_j}-\frac{1}{R_j}\frac{\kappa_i+\kappa_j}{(K_j+\kappa_i)(K_j-\kappa_j)}\right]-2\frac{R_jY_j}{R_iY_i}\frac{\alpha}{K_j}Y_j\right]$$

Multiply through by $(R_iY_i)^2$ (which is strictly positive):

$$\frac{\partial g_i^*}{\partial K_j} \propto (R_iY_i+R_jY_j)\left[(R_iY_i+R_jY_j)\,Y_j\left[\frac{\alpha}{K_j}-\frac{\kappa_i+\kappa_j}{R_j(K_j+\kappa_i)(K_j-\kappa_j)}\right]-2R_jY_j^2\frac{\alpha}{K_j}\right]$$

$$\frac{\partial g_i^*}{\partial K_j} \propto (R_iY_i+R_jY_j)\left[(R_iY_i+R_jY_j)\,Y_j\left[\frac{\alpha}{K_j}-\frac{\kappa_i+\kappa_j}{R_j(K_j+\kappa_i)(K_j-\kappa_j)}\right]-2R_jY_j^2\frac{\alpha}{K_j}\right]$$

$$\frac{\partial g_i^*}{\partial K_j} \propto (R_iY_i+R_jY_j)^2\left[\alpha\frac{R_iY_i-R_jY_j}{R_iY_i+R_jY_j}\cdot\frac{Y_j}{K_j}-\frac{\kappa_i+\kappa_j}{R_j(K_j+\kappa_i)(K_j-\kappa_j)}\right]$$

Divide through by $(R_iY_i + R_jY_j)^2 Y_j$ and multiply through by R_j:

$$\frac{\partial g_i^*}{\partial K_j} \propto \left[R_j \frac{R_iY_i - R_jY_j}{R_iY_i + R_jY_j} \cdot \frac{\alpha}{K_j} - \frac{\kappa_i + \kappa_j}{(K_j + \kappa_i)(K_j - \kappa_j)} \right]$$

If $R_jY_j > R_iY_i$, then $R_iY_i - R_jY_j < 0$, and therefore this derivative is negative.

If $R_jY_j < R_iY_i$, the derivative can be either negative or positive.

$\square$

B.1.3 Proof of proposition 2.2

We already obtained these expressions for the probability of victory:

$$P_i(\text{win}) = \frac{\sqrt{R_iY_i \ln\left[\frac{Y_j}{L_j}\right]}}{\sqrt{R_iY_i \ln\left[\frac{K_j}{L_j}\right]} + \sqrt{R_jY_j \ln\left[\frac{Y_i}{L_i}\right]}}$$

$$P_i(\text{win}) = \frac{\sqrt{R_iK_i^\alpha L_i^{1-\alpha} \ln\left[\frac{K_j}{L_j}\right]}}{\sqrt{R_iK_i^\alpha L_i^{1-\alpha} \ln\left[\frac{K_j}{L_j}\right]} + \sqrt{R_jK_j^\alpha L_j^{1-\alpha} \ln\left[\frac{K_i}{L_i}\right]}}$$

Thus, P_i(win is increasing in Y_i if and only if $\frac{Y_i}{\ln\left[\frac{Y_i}{L_i}\right]}$ is increasing in Y_i, and likewise P_i(win is increasing in K_i if and only if $\frac{K_i^\alpha}{\ln\left[\frac{K_i}{L_i}\right]}$ is increasing in K_i.

Differentiating these yields:

$$\frac{\partial}{\partial Y_i}\left[\frac{Y_i}{\ln\left[\frac{Y_i}{L_i}\right]}\right] = \frac{\ln\left[\frac{Y_i}{L_i}\right]-1}{\left(\ln\left[\frac{Y_i}{L_i}\right]\right)^2}$$

$$\frac{\partial}{\partial K_i}\left[\frac{K_i^\alpha}{\ln\left[\frac{K_i}{L_i}\right]}\right] = \frac{\alpha K_i^{\alpha-1}\ln\left[\frac{K_i}{L_i}\right]-K_i^\alpha\frac{1}{K_i}}{\left(\ln\left[\frac{K_i}{L_i}\right]\right)^2} = K_i^{\alpha-1}\frac{\alpha\ln\left[\frac{K_i}{L_i}\right]-1}{\left(\ln\left[\frac{K_i}{L_i}\right]\right)^2}$$

By assumption, $\ln\left[\frac{Y_i}{L_i}\right] = \alpha\ln\left[\frac{K_i}{L_i}\right] > 1$. Therefore both derivatives are positive. QED.

B.1.4 Proof of proposition 2.6

The effect of other capital destruction ω_j is apparent by inspection. The effect of own capital destruction ω_i is not as immediate as ω_j, because ω_i occurs in two terms. The direct effect of ω_i is to decrease the payoffs for war by destroying one's own capital; but there is also an indirect effect via increasing the return on victory. To prove proposition 2.6, we must account for this tradeoff and show that one side always dominates.

$$\frac{\partial R_i}{\partial \omega_i} = \frac{1}{(1-\omega_i)\left(1+\frac{K_i(1-\omega_i)}{\beta_j K_j(1-\omega_j)}\right)}$$

$$\frac{\partial \Delta_i}{\partial \omega_i} = -\frac{1}{1-\omega_i} + \frac{\partial \Delta_i}{\partial R_i}\frac{\partial R_i}{\partial \omega_i}$$

$$\frac{\partial \Delta_i}{\partial \omega_i} = \frac{1}{1-\omega_i}\left[\frac{\frac{g_i^*+g_j^*}{2\eta(L_i+L_j)}+\frac{1}{2}}{1+\frac{K_i(1-\omega_i)}{\beta_j K_j(1-\omega_j)}} - 1\right]$$

Is this positive or negative? Note that $\frac{g_i^*+g_j^*}{2\eta(L_i+L_j)}+\frac{1}{2}$ is a probability, so it cannot be larger than 1.

Moreover, $1+\frac{K_i(1-\omega_i)}{\beta_j K_j(1-\omega_j)} > 1$.

Therefore the -1 term dominates, and $\frac{\partial \Delta_i}{\partial \omega_i} < 0$. Increasing the amount of capital destroyed reduces the propensity to war for both sides.

QED.

B.1.5 Proof of proposition 2.7

Proposition 2.7. More generally, we have:

$$R_i - R_j = \ln\left[\frac{K_i+\kappa_j}{K_i-\kappa_i}\right] - \ln\left[\frac{K_j+\kappa_i}{K_j-\kappa_j}\right]$$

$$R_i - R_j = \ln\left[\frac{(K_i+\kappa_j)(K_j-\kappa_j)}{(K_i-\kappa_i)(K_j+\kappa_j)}\right]$$

$$R_i - R_j = \ln\left[K_iK_j + (K_j - K_i)\kappa_j - \kappa_j^2\right] - \ln\left[K_iK_j + (K_i - K_j)\kappa_i - \kappa_i^2\right]$$

If $K_i = K_j = K$, this reduces to:

$$R_i - R_j = \ln\left[\frac{K^2 - \kappa_j^2}{K^2 - \kappa_i^2}\right]$$

If $\kappa_i > \kappa_j$, the denominator is smaller, therefore the fraction is larger than one and $R_i > R_j$.

Since by assumption $K_i = K_j$ and $L_i = L_j$, $Y_i = Y_j$ and thus $R_iY_i > R_jY_j$. By proposition 2.2, government spending and probability of victory is higher for country i.

□

B.1.6 Proof of proposition 2.8

Proposition 2.8. By the symmetry of the situation, since $Y_i = Y_j = Y$ and $K_i = K_j = K$, the only difference between Δ_i and Δ_j is as follows:

$$\begin{aligned}\Delta_i - \Delta_j = R_i &\left(\frac{R_i}{R_i + R_j} - (1-\alpha)\right)\frac{R_i}{R_i + R_j} \\ &+ \alpha\ln\left[1 - \frac{\kappa_i}{K}\right] - \alpha\ln\left[1 - \frac{\kappa_j}{K}\right]\end{aligned}$$

This term represents the increased cost of going to war due to the risk of losing more capital:

$$\alpha \ln \left[1 - \frac{\kappa_i}{K}\right] - \alpha \ln \left[1 - \frac{\kappa_j}{K}\right] = \alpha \ln \left[\frac{K - \kappa_i}{K - \kappa_j}\right] < 0$$

The remaining terms represent the increased benefit of going to war due to a higher chance of victory.

By proposition 2.7, $R_i > R_j$. Therefore $\frac{R_i}{R_i+R_j} > \frac{1}{2}$.

If $\frac{R_i}{R_i+R_j} < 1 - \alpha$, the proposition is proved. Therefore suppose that $\frac{R_i}{R_i+R_j} > 1 - \alpha$.

The largest possible value of $\frac{R_i}{R_i+R_j}$ is 1; therefore $\Delta_i - \Delta_j$ is bounded above by:

$$\Delta_i - \Delta_j \leq \alpha R_i + \alpha \ln \left[\frac{K - \kappa_i}{K - \kappa_j}\right]$$

Recall that $R_i = \ln \left[\frac{K_i+\kappa_j}{K_i-\kappa_i}\right]$.

Therefore we have:

$$\Delta_i - \Delta_j \leq \alpha \ln \left[\frac{K + \kappa_j}{K - \kappa_i}\right] + \alpha \ln \left[\frac{K - \kappa_i}{K - \kappa_j}\right]$$

$$\Delta_i - \Delta_j \leq \alpha \ln \left[K_i + \kappa_j\right] - \alpha \ln \left[K_i - \kappa_i\right] + \ln \left[K - \kappa_i\right] - \ln \left[K - \kappa_j\right]$$

$$\Delta_i - \Delta_j \leq \alpha \ln\left[K_i + \kappa_j\right] - \ln\left[K - \kappa_j\right]$$

$$\Delta_i - \Delta_j \leq \alpha \left[\frac{K + \kappa_j}{K - \kappa_j}\right] < \alpha R_i$$

[[This bound is too weak. It is bounded above by something positive, so it could still be positive.]]

$\square$

B.1.7 Proof of proposition 2.9

Proposition 2.9. If we let $K_i = K_j = K$, $Y_i = Y_j = Y$, etc., equation 2.6 becomes:

$$\begin{aligned}\Delta_i = {} & R\left(\frac{RY}{RY + RY} - (1-\alpha)\right)\frac{RY}{RY + RY} \\ & + \alpha \ln\left[1 - \frac{\kappa}{K}\right] + \alpha \ln(1-\omega) \\ & - \ln\left[1 + \frac{\gamma(1-s)}{D_{ij}} Y + \frac{\tau - \tau}{Y}\right]\end{aligned}$$

$$
\begin{aligned}
\Delta_i &= R\left(\frac{1}{2} - (1-\alpha)\right)\frac{1}{2} \\
&+ \alpha \ln\left[1 - \frac{\kappa}{K}\right] + \alpha \ln(1-\omega) \\
&- \ln\left[1 + \frac{\gamma(1-s)}{D_{ij}} Y\right]
\end{aligned}
$$

For any $\alpha < \frac{1}{2}$, all four terms in this expression are negative, and therefore $\Delta_i < 0$.

$\square$

B.1.8 Proof of proposition 2.10

Proposition 2.10. Suppose that $\Delta_i > 0$. Recall equation 2.6:

$$
\begin{aligned}
\Delta_i &= R_i\left(\frac{R_i Y_i}{R_i Y_i + R_j Y_j} - (1-\alpha)\right)\frac{R_i Y_i}{R_i Y_i + R_j Y_j} \\
&+ \alpha \ln\left[1 - \frac{\kappa_i}{K_i}\right] + \alpha \ln(1-\omega) \\
&- \ln\left[1 + \frac{\gamma(1-s)}{D_{ij}} Y_j + \frac{\tau_{ji} - \tau_{ij}}{Y_i}\right] > 0
\end{aligned}
$$

Compare this against Δ_j:

$$\Delta_j = R_j \left(\frac{R_j Y_j}{R_i Y_i + R_j Y_j} - (1-\alpha) \right) \frac{R_j Y_j}{R_i Y_i + R_j Y_j}$$
$$+ \alpha \ln \left[1 - \frac{\kappa_j}{K_j} \right] + \alpha \ln(1-\omega)$$
$$- \ln \left[1 + \frac{\gamma(1-s)}{D_{ij}} Y_j + \frac{\tau_{ij} - \tau_{ji}}{Y_j} \right]$$

The sum $\alpha \ln \left[1 - \frac{\kappa_j}{K_j} \right] + \alpha \ln(1-\omega)$ is negative. (Note as in footnote 1 above that this is the case even if $\kappa_j < 0$, because of the effect of ω.)

Since country i is belligerent, country j will not be receiving tribute from them; there is no incentive for country i to try to stop country j from fighting. Thus, $\tau_{ji} = 0$.

Suppose that country j did not want to fight, and thus was willing to pay tribute $\tau_{ij} > 0$. Then it follows that $\Delta_j < 0$ and the proposition is proved.

Thus, the only case in which we could have $\Delta_j > 0$ would require that $\tau_{ji} = 0$. Therefore the trade term $- \ln \left[1 + \frac{\gamma(1-s)}{D_{ij}} Y_j + \frac{\tau_{ij} - \tau_{ji}}{Y_j} \right] \leq 0$.

In order for Δ_j to be positive, it is therefore necessary to have:

$$R_j \left(\frac{R_j Y_j}{R_i Y_i + R_j Y_j} - (1-\alpha) \right) \frac{R_j Y_j}{R_i Y_i + R_j Y_j} > 0$$

But for the same reason, in order for Δ_i to be positive, we must have:

$$R_i \left(\frac{R_i Y_i}{R_i Y_i + R_j Y_j} - (1-\alpha) \right) \frac{R_i Y_i}{R_i Y_i + R_j Y_j} > 0$$

By assumption, Δ_i is positive. But this implies that $\frac{R_i Y_i}{R_i Y_i + R_j Y_j} > 1 - \alpha$.

For realistic values of $\alpha < \frac{1}{2}$, this necessitates $R_i Y_i > R_j Y_j$. Which means that we must have

$$R_j \left(\frac{R_j Y_j}{R_i Y_i + R_j Y_j} - (1 - \alpha) \right) \frac{R_j Y_j}{R_i Y_i + R_j Y_j} < 0$$

Therefore $\Delta_j < 0$.

$\square$

B.1.9 Proof of proposition 2.11

Proposition 2.11. Equation 2.9 tells us what the minimum necessary tribute must be to ensure peace in equilibrium:

$$\begin{aligned} \tau_{ji}^* = Y_i \exp \Bigg[& R_i \left(\frac{R_i Y_i}{R_i Y_i + R_j Y_j} - (1 - \alpha) \right) \frac{R_i Y_i}{R_i Y_i + R_j Y_j} \\ & + \alpha \ln \left[1 - \frac{\kappa_i}{K_i} \right] + \alpha \ln(1 - \omega) \Bigg] \\ & - Y_i \left[1 + \frac{\gamma}{D_{ij}} Y_j \right] \end{aligned}$$

We want to see under what conditions we can ensure that $\tau_{ji}^* < Y_j$.

$$
\begin{aligned}
Y_j > Y_i \exp\Bigg[& R_i \left(\frac{R_i Y_i}{R_i Y_i + R_j Y_j} - (1-\alpha) \right) \frac{R_i Y_i}{R_i Y_i + R_j Y_j} \\
& + \alpha \ln \left[1 - \frac{\kappa_i}{K_i} \right] + \alpha \ln(1-\omega) \Bigg] \\
& - Y_i \left[1 + \frac{\gamma}{D_{ij}} Y_j \right]
\end{aligned}
$$

$$
\begin{aligned}
\frac{Y_j}{Y_i} + \left[1 + \frac{\gamma}{D_{ij}} Y_j \right] > \exp\Bigg[& R_i \left(\frac{R_i Y_i}{R_i Y_i + R_j Y_j} - (1-\alpha) \right) \frac{R_i Y_i}{R_i Y_i + R_j Y_j} \\
& + \alpha \ln \left[1 - \frac{\kappa_i}{K_i} \right] + \alpha \ln(1-\omega) \Bigg]
\end{aligned}
$$

$$
Y_j \left[\frac{1}{Y_i} + \frac{1}{Y_j} + \frac{\gamma}{D_{ij}} \right] > \left[1 - \frac{\kappa_i}{K_i} \right]^{\alpha} (1-\omega)^{\alpha} \exp \left[R_i \left(\frac{R_i Y_i}{R_i Y_i + R_j Y_j} - (1-\alpha) \right) \frac{R_i Y_i}{R_i Y_i + R_j Y_j} \right]
$$

For any given value of the other parameters, we can always find a threshold $\underline{Y} > 0$ such that for all $Y_j < \underline{Y}$, this inequality fails. In fact there is a closed-form expression for $\underline{Y}$ which is well-defined and positive for any values of the parameters in the domain:

$$
\underline{Y} = \frac{\left[1 - \frac{\kappa_i}{K_i} \right]^{\alpha} (1-\omega)^{\alpha}}{\left[\frac{1}{Y_i} + \frac{1}{Y_j} + \frac{\gamma}{D_{ij}} \right]} \exp \left[R_i \left(\frac{R_i Y_i}{R_i Y_i + R_j Y_j} - (1-\alpha) \right) \frac{R_i Y_i}{R_i Y_i + R_j Y_j} \right]
$$

$\square$

B.1.10 Proof of proposition 2.12

Proposition 2.12. We want to find the conditions under which $\Delta_i' < \Delta_i$.

$$\begin{aligned}
\Delta_i' = {} & R_i \left(\frac{R_i Y_i}{R_i Y_i + R_j Y_j} - (1-\alpha) \right) \frac{R_i Y_i}{R_i Y_i + R_j Y_j} \\
& + \frac{1}{\delta} \alpha \ln \left[\frac{K_i - \kappa_i}{L_i} \right] + \frac{1}{\delta} \alpha \ln(1-\omega) \\
& + \frac{(1-\delta)}{\delta} \frac{R_i Y_i}{R_i Y_i + R_j Y_j} (\alpha R_i) \\
& - \alpha \ln \left[\frac{K_i}{L_i} \right] - \frac{1}{\delta} \ln \left[1 + \frac{\gamma}{D_{ij}} Y_j \right]
\end{aligned}$$

$$\begin{aligned}
\Delta_i = {} & R_i \left(\frac{R_i Y_i}{R_i Y_i + R_j Y_j} - (1-\alpha) \right) \frac{R_i Y_i}{R_i Y_i + R_j Y_j} \\
& + \alpha \ln \left[1 - \frac{\kappa_i}{K_i} \right] + \alpha \ln(1-\omega) \\
& - \ln \left[1 + \frac{\gamma(1-s)}{D_{ij}} Y_j + \frac{\tau_{ji} - \tau_{ij}}{Y_i} \right]
\end{aligned}$$

We are interested in the difference between these two:

$$
\begin{aligned}
\Delta_i' - \Delta_i = {} & R_i \left(\frac{R_i Y_i}{R_i Y_i + R_j Y_j} - (1-\alpha) \right) \frac{R_i Y_i}{R_i Y_i + R_j Y_j} \\
& + \frac{1}{\delta} \alpha \ln \left[\frac{K_i - \kappa_i}{L_i} \right] + \frac{1}{\delta} \alpha \ln(1-\omega) \\
& + \frac{(1-\delta)}{\delta} \frac{R_i Y_i}{R_i Y_i + R_j Y_j} (\alpha R_i) \\
& - \alpha \ln \left[\frac{K_i}{L_i} \right] - \ln \left[1 + \frac{\gamma}{D_{ij}} Y_j \right] \\
& - R_i \left(\frac{R_i Y_i}{R_i Y_i + R_j Y_j} - (1-\alpha) \right) \frac{R_i Y_i}{R_i Y_i + R_j Y_j} \\
& - \alpha \ln \left[1 - \frac{\kappa_i}{K_i} \right] - \alpha \ln(1-\omega) \\
& - \frac{1}{\delta} \ln \left[1 + \frac{\gamma}{D_{ij}} Y_j \right]
\end{aligned}
$$

$$
\begin{aligned}
\Delta_i' - \Delta_i = {} & \frac{1-\delta}{\delta} \alpha \ln \left[1 - \frac{\kappa_i}{K_i} \right] + \frac{1-\delta}{\delta} \alpha \ln(1-\omega) \\
& + \frac{1-\delta}{\delta} \frac{R_i Y_i}{R_i Y_i + R_j Y_j} (\alpha R_i) - \frac{1-\delta}{\delta} \ln \left[1 + \frac{\gamma}{D_{ij}} Y_j \right]
\end{aligned}
$$

$\frac{1-\delta}{\delta} > 0$, so $\Delta_i' < \Delta_i$ if and only if:

$$
\ln \left[1 - \frac{\kappa_i}{K_i} \right] + \ln(1-\omega) + \frac{R_i Y_i}{R_i Y_i + R_j Y_j} R_i - \ln \left[1 + \frac{\gamma}{D_{ij}} Y_j \right] < 0
$$

The Grim Trigger strategy will yield a subgame-perfect equilibrium, provided that $\Delta_i' < 0$.

If $\Delta_i' < \Delta_i$, for any fixed value of Δ_i, decreasing δ will decrease Δ_i', and since $\frac{1-\delta}{\delta}$ is unbounded above, there is always some threshold $\underline{\delta}$ such that $\Delta_i' < 0$.

$\square$