Social norms and beliefs about collective action, constitutional compliance and intra-gender productivity

An experimental investigation

Universität Hamburg
Fakultät für Wirtschafts- und Sozialwissenschaften

Dissertation
zur Erlangung der Würde eines Doktors der
Wirtschafts- und Sozialwissenschaften „Dr. rer. pol.“
(gemäß der PromO vom 24. August 2010)
vorgelegt von

Konstantinos Pilpilidis
geboren in Thessaloniki, Griechenland in Hamburg am
07.09.2020

Hamburg, den 07.01.2022

urn:nbn:de:gbv:18-ediss-97958

Vorsitzender: Prof. Dr. Gerd Mühlheußer

Erstgutachter: Prof. Dr. Stefan Voigt

Zweitgutachter: Prof. Dr. Andreas Lange

Datum der Disputation: 27.01.2021

Erklärung

Hiermit erkläre ich, Konstantinos Pilpilidis, dass ich keine kommerzielle Promotionsberatung in Anspruch genommen habe. Die Arbeit wurde nicht schon einmal in einem früheren Promotionsverfahren angenommen oder als ungenügend beurteilt.

K. Pilpilidis

Hamburg, 07.09.2020 Unterschrift Doktorand

Eidesstattliche Versicherung

Ich, Konstantinos Pilpilidis, versichere an Eides statt, dass ich die Dissertation mit dem Titel:

Social norms and beliefs about collective action, constitutional compliance and intra-gender productivity: An experimental investigation

selbst und bei einer Zusammenarbeit mit anderen Wissenschaftlerinnen oder Wissenschaftlern gemäß den beigefügten Darlegungen nach § 6 Abs. 3 der Promotionsordnung der Fakultät Wirtschafts- und Sozialwissenschaften vom 24. August 2010 verfasst habe. Andere als die angegebenen Hilfsmittel habe ich nicht benutzt.

K. Pilpilidis

Hamburg, 07.09.2020 Unterschrift Doktorand

Unterschrift Verwaltung

Selbstdeklaration bei kumulativen Promotionen

Für alle der vorliegenden Artikel liegt die Eigenleistung fur das Konzept, die Planung, die Durchführung und der Manuskripterstellung bei 100 %.

K. Pilpilidis

Hamburg, 07.09.2020 Unterschrift Doktorand

Summary of the dissertation

This cumulative dissertation has two goals. First, it investigates the interaction between social norms and beliefs and individual behavior. Secondly, it introduces new tools for the study of social norms and beliefs. The first chapter acts as an introduction and a summary of the methods and the findings. The second chapter contains the first paper of the dissertation. The third chapter contains the second paper and the last chapter the third paper of the dissertation.

The second chapter discusses how monetary incentives influence whether citizens challenge an executive who abuses its authority for personal gain. During a lab experiment, I vary the costs and benefits of challenge and introduce a reward for loyalty to the executive. Participants who cannot monetarily profit from challenging challenge the executive who abuses its power. When the cost of challenge or the reward for loyalty increase, a participant who can monetarily profit from challenging is more likely to challenge. Contrary, the behavior of a participant who cannot monetarily profit from challenge does not change when the cost of challenge or the reward for loyalty increase. These results indicate that participants are partially driven by non-monetary motives, when they oppose an abuse of power.

The third chapter uses simulations to determine how constitutional enforcement influences the probability that unamendable constitutional provisions will become unpopular with the passage of time. The agents in the model organize their financial activities based on the society's laws, which are written by elected legislators. When the legislator is not stopped from violating the constitution, the probability that the next legislator will also violate the constitution increases. In contrast, the existence of mechanisms which stop violations (e.g. judicial review) significantly reduces the probability that a legislator will come to power who will try to violate the constitution. Constitutions which start with constraints to the legislator face a lower probability that they will be violated or that a legislator who wants to violate them will come to power, even when the constraints are removed later in time.

The literature divides gender discrimination into two types: taste and statistical discrimination. The third chapter disentangles the two using an online experiment. The participants are paid when they guess correctly the winner in a number of mixed-gender opponents' pairs. Before they submit their guesses, they learn the opponents' genders, ages and education. After submitting their guesses, they decide who in each pair should get a bonus knowing the opponents' genders, education and score in the competition. On the one hand, I find that most participants believe that men have higher scores on average compared to women. Yet, I find no further evidence that the participants statistically discriminate against women. On the other hand, women are less likely compared to men to get the bonus when they have a higher score than their male opponents. Further analysis shows that the women who are the most likely to win are the least likely to get the bonus. In other words, the participants taste discriminate.

Zusammenfassung der Dissertation

Diese kumulative Dissertation verfolgt zwei Ziele. Als Erstes untersucht sie die Interaktion zwischen sozialen Normen und beliefs und individuellem Verhalten. Als Zweites führt sie neue Instrumente für die Untersuchung von sozialen Normen und beliefs ein. Das erste Kapitel ist die Einleitung, in der die Methoden und die Befunde zusammengefasst werden. Das zweite Kapitel beinhaltet das erste Papier der Dissertation. Das dritte Kapitel beinhaltet das zweite Papier und das letzte Kapitel beinhaltet das dritte Papier dieser Dissertation.

Im zweiten Kapitel wird der Zusammenhang zwischen monetären Anreizen und ob zwei Bürger gegen die Exekutive vorgehen, die ihre Autorität zum persönlichen Vorteil missbraucht, erörtert. Während eines Labor Experiments variiere ich die Kosten und den Nutzen eines Widerstands und führe eine Belohnung für die Loyalität gegenüber der Regierung ein. Die Teilnehmer leisten Widerstand gegen die Regierung, auch wenn sie nicht finanziell von dem Widerstand profitieren können. Ein Teilnehmer, der finanziell von einem koordinierten Widerstand profitieren kann, leistet eher Widerstand, wenn die Kosten des Widerstands oder die Belohnung für die Loyalität steigen. Das Verhalten eines Teilnehmers, der von einem Widerstand monetär nicht profitieren kann, bleibt nach der Steigerung der Kosten des Widerstands oder der Belohnung für die Loyalität unverändert. Dies weist darauf hin, dass die Teilnehmer teilweise aus nicht-monetären Gründen gegen einen Missbrauch von Macht vorgehen.

Das dritte Kapitel verwendet Simulationen um festzustellen, wie die Durchsetzung der Verfassung die Wahrscheinlichkeit, dass mit der Zeit ein Verfassungsartikel, welcher mit einer Ewigkeitsklausel geschützt ist, unpopulär wird, beeinflusst. In Abhängigkeit der Gesetze, die der gewählte Gesetzgeber erlässt, organisieren die Agenten des Models ihre finanziellen Aktivitäten. Je mehr Gesetzgeber gegen die Verfassung verstoßen, desto wahrscheinlicher wird es, dass Gesetzgeber gewählt werden, die gegen die Verfassung ver-

stoßen wollen. Im Gegensatz dazu verringert die Existenz von Mechanismen, die Verfassungsverstöße verhindern, die Wahrscheinlichkeit erheblich, dass ein Gesetzgeber an die Macht kommt, der versucht, gegen die Verfassung zu verstoßen. Verfassungen, die von Anfang an mit Mechanismen versehen sind, die Verfassungsverstöße verhindern, haben eine höhere Wahrscheinlichkeit respektiert zu werden und führen zu einer gerringeren Wahrscheinlichkeit, dass Gesetzgeber gewählt werden, die gegen die Verfassung verstoßen wollen. Dies zeigt sich auch, wenn die Mechanismen später entfernt werden.

Die Literatur unterteilt Gender Diskriminierung in zwei Arten: Präferenzbasierte und statistische Diskriminierung. Das vierte Kapitel unterscheidet diese beiden mithilfe eines Online-Experiments. Die Teilnehmer werden bezahlt, wenn sie die Gewinner aus gemischte Paaren von Konkurrenten korrekt raten. Bevor sie ihre Einschätzung eingeben, werden sie über das Geschlecht, die Altersgruppe und den akademischen Grad der Konkurrenten informiert. Nachdem sie ihre Einschätzungen eingegeben haben, entscheiden sie sich welcher der Konkurrenten einen Bonus erhalten soll. Bevor sie über den Bonus entscheiden, erfahren sie das Geschlecht, den akademische Grad der Konkurrenten und die Punkte, die die Konkurrenten in deren Aufgabe erreicht haben. Einerseits stelle ich fest, dass die meisten Teilnehmer glauben, dass Männer im Vergleich zu Frauen im Durchschnitt höhere Punktzahlen erreicht haben. Ich finde jedoch keine weiteren Beweise dafür, dass die Teilnehmer einzelne Frauen statistisch diskriminieren. Anderseits erhalten Frauen im Vergleich zu Männern mit einer geringeren Wahrscheinlich den Bonus, wenn sie eine höhere Punktzahl als ihr männlicher Gegner haben. Weitere Analysen zeigen, dass Frauen, die am häufigsten gewinnen, am seltensten den Bonus erhalten. Dies bedeutet, dass die Teilnehmer anhand ihrer Präferenzen diskriminieren.

Contents

List of Figures

List of Tables

Acknowledgements

I would like to thank the Graduate School of the Faculty of Economics and the Institute of Law & Economics of the University of Hamburg for funding the two experiments reported in the dissertation.

I am deeply indebted to my first supervisor Prof. Dr. Voigt for giving me the opportunity to become a member of the academic community, for his guidance and critical comments during my dissertation, for asking the hard questions when they were necessary and for teaching me how to think as an academic. I would like to extend my sincerest thanks to my second supervisor Prof. Dr. Lange for helping me structure my research project and giving me valuable comments. I wish to thank the president of my dissertation committee Prof. Dr. Gerd Muehlheusser for generously offering his time, support, guidance and good will throughout the review of this document.

Furthermore, I am grateful to Dr. Stephan Michel and Prof. Dr. Jerg Gutmann, with whom not only we shared common research interests but also countless hours of debating, discussing, developing and presenting research ideas to each other. Research is a solitary work which one cannot do alone. I would also like to thank Marek Endrich and Tara Schau for helping me prepare and run the experimental sessions. The importance of their assistance cannot be overstated. The experiment were conducted in cooperation with the experimental lab of the University of Hamburg. Its staff took care of all the necessary red tape. I would be ever grateful for that.

Drafting this dissertation would have been a significantly more strenuous task without the support of my wife. She has been there with me through thick and thin. Her sacrifices make my efforts seem menial. She kept my back clear, took care of our newborn children while I was writing, sacrificed her sleep so that I can be fit for research the next day. Last, I want to thank the reader of this dissertation for his/her time. I hope that reading the dissertation will give you the same pleasure writing gave to me.

Chapter 1

Introduction

1.1 Motivation

During the political movement coined the Arab Spring, a series of new constitutions were drafted which were meant to modernize the Arab world. The time Egypt got its new constitution coincided with the start of this dissertation's drafting. The Egyptian government with the constitution committed to achieving *equality between women and men in all civil, political, economic, social, and cultural rights* (Article 11 of the Egyptian constitution of 2014) and prohibited the discrimination of citizens on the basis of sex (Article 53 of the Egyptian constitution of 2014). These principles of equality were considered so important for the development of Egypt that the President or the House of Representatives were allowed only to amend them to increase the guarantees given by the constitution.

Similar to the Egyptian constitution, the Tunisian constitution laid the foundations for gender equality by proclaiming that *male and female, have equal rights and duties, and are equal before the law without any discrimination* (Article 21 of the Tunisia constitution of 2014). Contrary to Egypt, Tunisians decided to protect the human rights and freedoms guaranteed by the constitution from any amendment.

Some of the goals that the masses hoped to achieve with the constitutions were to establish the rule of law and improve the position of women in the respective societies. This inspired me to delve into the question whether it is possible to achieve these goals using constitutions alone. During my search for an answer, I came upon social norms and beliefs.

Section 1.2 defines social norms and beliefs and how they are analyzed in this dissertation. Section 1.3 briefly discusses the results of the individual chapters of the dissertation. Then, section 1.4 synthesizes the results, discusses the lessons learned and concludes the introduction.

1.2 Social norms and beliefs

The term social norms has been given many definitions and rules which different definitions characterize as social norms have been given many names over the time (Ostrom 1986; North 1990; Woolcock 1998; Bicchieri 2006; Voigt 2019). I use the term to describe the rules about how an individual should behave which are known to the members of a society and are followed without third party enforcement. For example, in chapter two I show that participants of an experiment pay the cost of challenging the executive to support a victim of an executive's transgression, even when they cannot benefit monetarily from this action. Their behavior can be explained by an informal rule that dictates that they should be solidary to other participants. In chapter four, participants of another experiment act as if they believe that women should not be rewarded when they compete against men and win.

The term beliefs is used to describe an individual's expectations about how others will act (first-order beliefs). In chapter two, the participants of the experiment need to expect that the other participants will support them in stopping the transgression otherwise they will not oppose it. In chapter three, agents decide how to organize their financial activities based on their beliefs about what laws the legislator is more likely to introduce in the future. Last, in chapter four the participants of the experiment need to evaluate how much effort a set of competitors will put in a competition in order to choose a winner.

Social norms might not be compatible with formal institutions or monetary profit-maximization. Further, different individuals might not share the same set of beliefs.[1] Both can lead to coordination failure and sub-optimal decisions (Voigt 2019). In chapter two, the participants are not able to coordinate and challenge the executive optimally because they underestimate how frequently others will support them. In chapter three, the constitution grows out of sync with the beliefs about the optimal way to organize the financial life and is thus violated. In chapter four, women are not rewarded for their performance, when the participants expect men to be better in the task than women.

Despite the importance of social norms and beliefs, the available methods to detect them and quantify their impact are currently at their infancy (Voigt

1. Theory expects beliefs in equilibrium to correspond to the true likelihood of events and all individuals to have the same beliefs (Aumann 1976). Nevertheless, in real life information does not flow freely as theory assumes.

2018). As of today, the workhorse approach to detect social norms is to use experiments which expose individuals to conditions where the norm-compliant behavior differs from the profit-maximizing behavior. I follow that approach in the second and fourth chapters.

In the third chapter, I use agent-based modeling and simulations to predict how the legal order would look like if citizens organized their financial activities according to their belief about future laws and voted legislators according to their financial interests. Agent-based models have been the go-to approach for modeling emergent phenomena and the impact of social norms. When experiments reveal social norms and beliefs, agent-based models reveal how social norms and beliefs translate to social outcomes.

1.3 What we learn from each chapter

In chapter two, I take on the question how monetary incentives influence the probability of challenging an illegal act of the executive. Weingast (1995) proposes a model where two citizens have to coordinate in order to challenge and stop the executive from stealing their property. The model formalizes the decision to challenge as a collective action problem. The pay-off maximizing strategy, i.e. challenging only when the executive steals from both citizens, leads to sub-optimal social outcomes. Thus, Weingast argues that social norms for cooperation will develop. I bring Weingast's game to the lab to find out whether such norms can be detected in a sample of German students. Further, I ask how these norms perform when the incentives to challenge change.

I find that indeed participants challenge even if they cannot monetarily profit from it. Sadly, they do not challenge with the optimal frequency, i.e. they do not anticipate correctly how often their partner challenges. Interestingly, when the incentives not to challenge become stronger, those from whom the executive steals - henceforth the victims - are more likely to challenge her. This seemingly irrational response is rewarded with a reduction in coordination failure. Specifically, the victims challenge more and non-victims do not challenge less. These results can only be explained by social norms for cooperation as Weingast hypothesized.

In chapter three, I look into how the probability changes over time that an unamendable constitutional provision will be violated. I design an agent-based model where agents decide how to organize their financial activities based on the past laws and legal certainty. Further, how an agent organizes his financial activities influences which laws he perceives as ideal. Laws are introduced by an elected legislator. Each agent votes for a new legislator from a pool of candidates based on how similar the laws are which the agent and the candidate perceive as ideal. I implement two types of treatments. The legislator is exposed to different types of constraints for legislating and the constraints are either fixed over time or change.

Simulations show that imperfect enforcement of the constitution, i.e. allowing the legislator to introduce unconstitutional laws, leads to the agents' ideal laws drifting outside of the set of laws that the constitution allows. Further, the agents start voting in favor of candidates which declare that they will try to violate the constitution. Simulations also show that enforcing the constitution perfectly at the start of the game significantly decreases the probability that agents will vote for candidates whose ideal laws are not allowed by the constitution. Further, the probability of violations is significantly smaller even if the constraints are later removed.

In the last chapter, I turn my attention to the search for an explanation why constitutions and laws fail to promote gender equality. As a starting point, I look for evidence that women are considered less productive than men, i.e. are statistically discriminated, or whether there is a social norm that dictates that men should get a higher reward for the same amount of effort. In a pre-experiment, participants solve a real-effort task and collect points. Then, they are paired with another participant from the pre-experiment and whoever has the most points wins and is paid. During the online experiment, I ask a different set of participants to predict the winners of the pre-experiment and reward one member of each pair with a bonus. Before the participants make their guesses and choices about the boni, I reveal to them the gender of each of the participants of the pre-experiment.

The online-experiment detects weak evidence of statistical discrimination. Participants are not more likely to guess that women lost than they are to guess that they won against men. Nevertheless, they appoint different weights between men and women to how age and education influence performance. Contrariwise, I find strong evidence that the participants are more likely to reward women for losing against men rather than for winning. Indicatively, the women that are the most likely to perform well in the task are the least likely women to be rewarded for their performance.

1.4 Synthesis

The dissertation was inspired by topics prominent during the Arab Spring but produced insights which are valid for a wide number of societies and developed methods which can be applied for the analysis of the impact of social norms and beliefs globally. All three chapters employ experimental methods to answer their respective questions. The unit of analysis is always the individual, who acts under incomplete information and relies on beliefs and social norms to cover his information gaps.

The dissertation demonstrates the importance of understanding and finding ways to measure and predict the emergence of social norms and beliefs. Further, it provides a solution to the problem of disentangling norms and beliefs acting simultaneously. The experiments reported show that beliefs can be inconsistent with reality and uncertainty about the applying social norms can lead to sub-optimal decisions. To overcome the practical difficulties of running experiments with big groups of individual, I resort to simulations. Through an agent-based model, I show a vicious cycle, where constitutional violations of the past lead to future constitutional violations.

The trip in the world of social norms and beliefs is by far not close to its end. The following years social scientists will have to delve deeper into producing models that combine social norms and bounded rationality with classical economic theory. I hope that this dissertation will provide some assistance to their work.

References

Aumann, Robert J. 1976. Agreeing to Disagree. *Annals of Statistics* 4 (6): 1236–1239.

Bicchieri, Cristina. 2006. *The grammar of society: the nature and dynamics of social norms.* 278. Cambridge: Cambridge University Press.

North, Douglass C. 1990. *Institutions, Institutional Change and Economic Performance.* 164. Cambridge, MA: Cambridge University Press.

Ostrom, Elinor. 1986. An agenda for the study of institutions. *Public Choice* 48 (1): 3–25.

Voigt, Stefan. 2018. How to measure informal institutions. *Journal of Institutional Economics* 14 (1): 1–22.

———. 2019. *Institutional Economics An introduction.* 293. Cambridge: Cambridge University Press.

Weingast, Barry R. 1995. The economic role of political institutions: Market-preserving federalism and economic development. *Journal of Law, Economics, and Organization* 11 (1): 1–31.

Woolcock, Michael. 1998. *Theory and Society* 27 (2): 151–208.

Chapter 2

When the divided conquer: Challenging an executive abusing her power

2.1 Introduction

A century and a half after the Communist Manifesto called the proletarians of the world to unite and fight for their common interests, researchers have yet to reach a consensus what makes individuals follow such calls and fight for a common cause. This chapter focuses on the citizen's decision to challenge a government executive in order to stop her from using her power for personal gain.

A citizen surrenders his power to control the executive so that she can protect him from other citizens. As a side-effect, he cannot stop her from misusing her power without the support of other citizens. Since the seminal work of North and Weingast (1989), the mechanisms that stop misuse of the executive's power have been the subject of a sizeable literature. Empirical research confirms that coordinated challenges of the executive, such as civil disobedience, protest and revolution, lead not only to democratization but also to economic growth (North 1990; Acemoglu and Robinson 2006, 2012; Tarrow 1994; Aidt and Franck 2013; Aidt and Jensen 2014; Aidt and Franck 2015; Aidt and Leon 2016). A hurdle against coordinated challenge is that the executives can use divide-and-conquer strategies. In particular, they can transgress, reward loyalty or use repression against parts of the population (Wintrobe 2000; Posner, Spier and Vermeule 2010).

Existing work has not yet measured the impact of changes in the incentives for challenge on the probability an individual challenges an executive who uses divide-and-conquer strategies. This chapter fills the gap. The chapter's insights are based on a lab experiment. Although observational data can be externally valid, variation because of unobservable factors confounds the results. The experiment not only controls for confounding factors but also allows targeted manipulation of each participant's pay-off.

The game brought to the lab has been coined the Collective Resistance game -henceforth CR-game. The CR-game takes place between a dictator, i.e. a participant who divides the pay-offs, and two responders. The game begins with the dictator announcing whether she will claim part of the responders' initial endowments for herself. Then, both responders have to challenge this claim if they want to stop her. The dictator can take from the responders in two ways. Basic predation - henceforth BP- involves the dictator claiming part of both responders' endowments. Divide-and-conquer predation -henceforth DAC- involves claiming part of one responder's endowment -henceforth the victim- and (credibly) promising to leave intact the other responder's -henceforth the witness- endowment.

Cason and Mui (2007) develop the CR-game based on the sovereign-constituency transgression game of Weingast (1995, 1997). Their aim is to investigate whether allowing non-binding messaging (cheap talk) between responders leads to more mutual challenge and less taking from the dictator. They present evidence that both types of claims decrease with pre-play communication.[1] When Cason and Mui (2013) bring the CR-game back to the lab, they are interested in how repetition influences the type and frequency of the dictator's claims. Weingast (1995) argues that repetition allows responders to use history-dependent strategies, for example, tit-for-tat. Hence, it should lead to coordination on the pay-off superior equilibrium, i.e. that the dictator will take nothing, and the responders will not challenge. Nevertheless, Cason and Mui (2013, p. F248) point out that repetition is a *two-edged sword*. The dictator can also use history-dependent strategies and target DAC against the responder who challenged in earlier rounds. They show that indeed repetition alone cannot reduce predatory claims.

It becomes apparent that so far the CR-game has been used for the measurement of repeated interactions, where participants learn what the other

1. Later, they expand their findings by showing that richer forms of communication improve coordination and reduce claims further (Cason and Mui 2015b). They further show that groups act as the individuals in the CR-game (Cason and Mui 2015a). Rigdon and Smith (2010) use the CR-game to answer whether challenges result from fairness concerns or building a reputation for challenging. Last, Chong, Liu and Zhang (2016) bring the CR-game to the field to see whether the findings in the lab are consistent with findings on the field, which they are.

participants did in each period.[2] Because the CR-game is a coordination game with multiple equilibria, playing with each other repeatedly allows participants to anticipate their partner's actions and avoid coordination failure (Stahl and Wilson 1995). Further, challenging impacts the pay-offs of the other participants. Responders with different other-regarding preferences will value challenge differently. As the uncertainty about the other's valuation of challenge decreases, it becomes easier to play optimal strategies. Subsequently, it is not possible to disentangle the effect of incentives and the effect of the history of play.

The experiments conducted so far look into how different individuals respond on average to different treatments. They do not answer how a particular individual responds to different treatments. To expand the current findings and fill the gap, I design an one-shot CR-game which generates multiple independent observations per participant. Specifically, I use a conditional information lottery to ensure that participants cannot update their beliefs or that history of play can influence the participants' choices. Due to the innovation in the design, I can expose each participant to various pay-off structures and see how the probability that they challenge changes depending on the pay-offs.

Theory suggests that in a setting like the one-shot CR-game citizens can react in two ways to pay-off changes. As the net monetary benefit of challenge decreases, some citizens will be deterred from challenging and some citizens will challenge more. What will a particular citizen do depends on what value he appoints to stopping predatory claims independent of their monetary consequences.

The results of the experiment are the following. Participants on average underestimate the willingness of other participants to challenge. A participant as a victim challenges with a higher probability when the cost of challenge or the reward for loyalty, i.e. not challenging, increase. In contrast, as a witness the participant does not respond to pay-off changes. The fact that witnesses challenge at all shows that stopping predatory claims bears an intrinsic value for them.

I do not find robust evidence that participants challenge more when challenging is less destructive or when allowing a predatory claim would make the pie bigger. This implies that even though participants find it important to stop the dictator from taking part of their partner's endowment, they have a limited interest for how much the dictator will lose because they stopped

2. In (Rigdon and Smith 2010), the participants committed to their strategies in advance. However, they gain experience by playing with other participants for the first 10 periods of the experiment.

her or how much she would gain if she is not stopped.

The structure of the chapter is as follows. Section 2.2 presents the CR-game and the equilibrium behavior. Section 2.3 describes the experimental design and identification strategy. Section 2.4 discusses the experimental results. Section 2.5 concludes.

2.2 Theoretical predictions

2.2.1 The Collective Resistance Game

The CR-game spans over three phases and has four players: Nature, the dictator and two responders (A and B). First, Nature moves (t_0), then the dictator moves (t_1) and last the responders move simultaneously (t_2). The players begin the game with an initial endowment. The choices during the game determine how their initial endowments change. The structure of the section is as follows. This subsection defines the strategy space and the player's pay-offs. The next subsection derives the equilibria. When multiple equilibria exist, the subsection presents the conditions under which each Nash-equilibrium is risk dominant. The last subsection discusses the comparative statics and converts the insight of the model to testable hypotheses.

Nature draws stochastically a subjective cost for each responder and announces it to him at t_0. Henceforth, the subjective cost's monetary equivalent will be denoted as θ. θ can be positive or equal to zero and it is private information. Notwithstanding, the dictators and the responders know the distribution of θ.

At t_1, the dictator chooses between four actions. She can choose not to claim anything from the responders' initial endowments. She can claim part of the responders' endowments to boost her own endowment, i.e. choose BP. Last, she can claim a part of the victim's endowment for herself and credibly promise to leave the witness's endowment as it is, i.e. choose DAC. When the dictator chooses DAC, she decides who of the two responders will be the victim and who the witness, i.e. there are two ways to play DAC. The dictator cannot decide how much she will take. She only decides whether she will claim anything and from whom. At t_2, each responder chooses independently and simultaneously, whether to challenge the dictator (CH) or acquiesce (AC).

Depending on what the responders choose the dictator either takes what she claimed or pays a fine for making a claim. The dictator's pay-off function contains four elements. She begins the game with her initial endowment, E_D. She can claim L (for loot) from the responder(s) to gain αL from each responder. α, which is positive, represents how much of the loot ends in the pocket of the dictator. The intuition behind α is that if the dictator uses the loot to invest in productive activities, α is bigger than one. However, if the dictator has to use the loot to finance a corrupt system, for example bribe judges, α is smaller than one.

When both responders challenge the dictator, she pays a fine F and does not get the loot. The fine is paid although the dictator does not take anything. Last, if only one responder challenges after the dictator made a claim she has to reward the other responder for loyalty. I set the reward for loyalty to be smaller than what the dictator takes ($R < \alpha L$). The following equation translates the verbal description of the dictator's pay-offs into mathematical notation:

$$\begin{aligned}\Pi_D =& E_D - F\sum_i q_{iA}\sum_i q_{iB}\\ &+\alpha L(1-\sum_i q_{iA}\sum_i q_{iB})(2q_{BP}+q_{DAC_A}+q_{DAC_B})\\ &-R((1-\sum_i q_{iA})\sum_i q_{iB}+\sum_i q_{iA}(1-\sum_i q_{iB}))(q_{BP}+q_{DAC_A}+q_{DAC_B})\end{aligned}$$

where q_{iA} and q_{iB} is the probability that responder A and B, respectively, will challenge given that the dictator chose action $i \in [BP, NC, DAC_A, DAC_B]$ and q_{BP}, q_{DAC_A} and q_{DAC_B} are the probabilities with which the dictator plays each strategy for which:

$$q_{BP} + q_{DAC_A} + q_{DAC_B} + q_{NC} = 1$$

where q_{NC} is the probability with which the dictator will not make a claim. The strategy profile that maximizes the dictator's pay-off depends on the probability with which responders mutually challenge each action.

The responders' pay-off functions contain five elements. Each responder gets what remains of his initial endowment ($E - L$ or E) after the dictator moves. When the responder challenges, he has to pay the costs of challenge C. Whether he will protect his initial endowment by challenging depends on what the other responder does. If the other responder challenges, the dictator is not allowed to go through with her claim. If the other responder does not challenge, the dictator takes what she announced that she will take.

On top, the responder who did not challenge gets a reward for loyalty, R. If the loyal responder was the victim, the dictator takes less than claimed $L - R$. Last, witnessing the dictator taking from the other responder or losing part of the endowment from a predatory claim causes subjective costs θ. The following equation shows the n-th responders' expected pay-offs from playing the game with the m-th responder, i.e. before they see what the dictator chose:

$$\Pi_n = E - Cq_{in} + R(1-q_{in})q_{im} \\ -(L(q_{BP} + q_{DAC_n}) + \theta_n(1-q_{NC}))(1-q_{in}q_{im})$$

Responders maximize their pay-offs given the choice of the dictator. On the one hand, when the dictator does not claim parts of the responders' endowments as her own, the responders are strictly better-off to acquiesce independent of what the other responder does. On the other hand, when the dictator claims parts of the endowments, the optimal strategy depends on a number of factors.

The next subsection presents the responders' optimal strategies given that the dictator chose BP or DAC_A.[3] I discuss the pay-off for each strategy in terms of pay-off differences. Further, I assume that the dictator's pay-offs are irrelevant for the responders' pay-offs and I momentary blend them out of the discussion. Table 2.1 shows how a sub-game's pay-offs expressed in terms of pay-off differences look like.

		B (AC)	B (CH)
A	(AC)	$1-q_A$, $t(1-q_B)$	0, 0
	(CH)	0, 0	q_A, tq_B

Table 2.1: General form sub-game

where q_A and q_B are the responders pay-offs and t is a normalization factor for which:

$$q_A = \frac{L-C-R+\theta_A}{L-R+\theta_A} \quad |i \in [BP, DAC_A] \text{ and}$$

$$q_B = \begin{cases} \dfrac{-C-R+\theta_B}{-R+\theta_B} & |i = DAC_A \\ \dfrac{L-C-R+\theta_B}{L-R+\theta_B} & |i = BP \end{cases} \quad \text{and}$$

$$t = \begin{cases} \dfrac{-R+\theta_B}{L-R+\theta_A} > 0 \mid i = DAC_A \\ \dfrac{L-R+\theta_B}{L-R+\theta_A} > 0 \mid i = BP \end{cases}$$

3. DAC_A and DAC_B are symmetrical. Thus, the findings from the one sub-game can be applied to the other sub-game by re-labelling who is the victim and the witness. To avoid unnecessary redundancies, I discuss only one of the two sub-games.

2.2.2 Sub-game perfect Nash-equilibria

With complete information, the sub-game perfect equilibria (SPNE) are as follows. When $q_A \leq 0$ or $q_B \leq 0$, both responders best response is to acquiesce. This is true, when the dictator does not loot, when the witness of DAC has a low subjective cost ($\theta < C + R$) or the cost of challenge ($C + R$) is higher than the benefit of challenge ($L + \theta$).

When $q_A > 0$ and $q_B > 0$, multiple equilibria exist. One situation where this condition is fulfilled is the sub-game after the dictator chooses basic predation and $C + R < L$. In this case, both responders' best response is to match the pure strategy that the other responder plays, i.e. challenge when the other challenges and acquiesce when the other acquiesces. There is also one equilibrium where responder A challenges with probability equal to q_B and acquiesces with probability equal to $1 - q_B$ and responder B challenges with probability equal to q_A and acquiesces with probability equal to $1 - q_A$.

Due to the assumption that the distribution of θ is common knowledge, when there are multiple equilibria it is possible to calculate the equilibrium mixed strategy by replacing θ with its expected value. This exercise reveals that even when the expected value of θ is zero the equilibria in pure strategies do not change for the sub-game after BP. Contrary to the case of BP, for certain values of θ the witness's pay-offs difference is negative, i.e. the witness will acquiesce independent of what the victim does.

I will call a witness for whom $\theta \geq C + R$ social and a witness for which $\theta < C + R$ non-social. A non-social witness always acquiesces independent of what the other responder does. Responders extrapolate the probability with which they play a game with a social witness or a game with a non-social witness from the distribution of θ. Let a witness be social with a probability $0 \leq p \leq 1$ and let $0 \leq p' \leq 1$ be the probability that a victim challenges because he perceives p as sufficiently high, henceforth confident victim. A victim who is not confident will be henceforth referred to as insecure.

Table 2.2 represents the normal form of the sub-game after DAC_A with incomplete information. The first action in parentheses shows how a non-social witness (column-player) and an insecure victim (row-player) will play. The second action in parentheses represents how the social witness and the confident victim will play, respectively. For simplicity, the table does not contain the pay-offs of the non-social witness and the strictly dominated strategies are omitted.[4]

4. Because the non-social witness has $q_B < 0$ the witness's strategies (CH,CH) and (CH,AC) are strictly dominated by (AC,CH) and (AC,AC), respectively. The strategy where the insecure victim challenges and the confident victim acquiesces (CH,AC) is also omitted as dominated by (CH,CH).

		B (AC,AC)	B (AC,CH)
A	(AC,AC)	$1 - q_A$, $t(1 - q_B)$	$(1 - p)(1 - q_A)$, $t(1 - p)(1 - q_B)$
A	(AC,CH)	$(1 - p')(1 - q_A)$, $t(1 - p')(1 - q_B)$	$(1 - p)(1 - p')(1 - q_A) + pp'q_A$, $(1 - p)(1 - p')(1 - q_B) + pp'q_B$
A	(CH,CH)	0, 0	pq_A, tq_B

Table 2.2: Normal form DAC_A sub-game with incomplete information

Mutual acquiescence is an equilibrium independent of p and p'.

$$1 - q_A \geq (1 - p')(1 - q_A) > 0 \text{ and } 1 - q_B \geq (1 - p)(1 - q_B) \tag{2.1}$$

When the social witness challenges with probability q_B^* and acquiesces with a probability $1 - q_B^*$, the victim's expected pay-off is $[(1 - p) + p(1 - q_B^*)](1 - q_A)$ if he acquiesces and $pq_B^* q_A$, if he challenges. Hence, he is indifferent between challenge and acquiescence when the social witness challenges with a probability:

$$0 < q_B^* = \frac{1 - q_A}{p} \leq 1$$

The above implies that for $p < 1 - q_A$, a victim will acquiesce even if the social witness challenges with certainty, i.e. is insecure. For

$$p^* = \frac{C}{L - R}$$

and $p \geq p^*$, it is impossible for a victim to be insecure ($p' = 1$).

Let a victim challenge with probability q_i^* and acquiesce with probability $1 - q_i^*$, if he is insecure and challenge with probability q_c^* and acquiesce with probability $1 - q_c^*$, if he is confident. The social witness's expected pay-offs for each choice are:

$$\Pi_{CH} = tq_B[p'q_c^* + (1 - p')q_i^*]$$

$$\Pi_{AC} = t(1 - q_B)[p'(1 - q_c^*) + (1 - p')(1 - q_i^*)]$$

These expected pay-offs are equal when:

$$0 \leq q_c^* = \frac{1 - q_B + (1 - p')q_i^*}{p'} \leq 1$$

The victim's pay-offs increase with the probability of mutual challenge. Hence, the optimal strategy for the confident victim is to challenge with a higher probability than the insecure victim. Further, the probability with which an insecure victim challenges is a function of the probability with which a confident victim would challenge. Subsequently, the confident victim will challenge with certainty, if $q_i^* > 0$. If the confident victim would challenge with certainty, the optimal mix for an insecure victim becomes:

$$q_i^* = 1 - \frac{q_B}{1 - p'}$$

It is straightforward that for $p' \leq 1 - q_w$ the victim will only challenge, if he is confident. In this case, if the confident victim does not challenge with certainty, the witness will acquiesce. To sum up, the sub-game perfect Nash-equilibria (SPNE) are the following:[5]

1. After NC:
   (a) Unique equilibrium: Mutual acquiescence
2. After BP:
   (a) Pay-off inferior equilibrium: Mutual acquiescence.
   (b) Pay-off superior equilibrium: Mutual challenge.
   (c) Mixed strategy equilibrium: Responders challenge with probability:
   $$q = \frac{L - C - R + E(\theta)}{L - R + E(\theta)}$$
   and acquiesces with probability $1 - q$, where $E(\theta)$ is the expected value of θ.
3. After DAC_A:[6]
   (a) Pay-off inferior equilibrium: Mutual acquiescence.
   (b) Pay-off superior equilibrium: A non-social witness acquiesces.
      - If $p \geq p^*$ a social witness and the victim challenge.
      - If $p' \leq 1 - q_B$ and $p^* > p > 1 - q_A$ a social witness and a confident victim challenge but an insecure victim acquiesces.
   (c) Mixed strategy equilibrium: A non-social witness acquiesces.

5. The labeling of the equilibria as pay-off superior and inferior concerns only the sum $\Pi_A + \Pi_B$. Whether the pay-off superior equilibrium maximizes the total pay-offs, i.e. $\Pi_D + \Pi_A + \Pi_B$, depends on the size of α. If $\alpha > 1$, the size of the pie increases the more the dictator takes from the responders.

6. For the equilibria after DAC_B replace q_B, q_B^* and q_A with q_A, q_A^* and q_B accordingly.

- If $p \geq p^*$ a social witness and the victim challenge with probability q_B^* and q_c^*, and acquiesces with probability $1 - q_B^*$ and $1 - q_c^*$, respectively.
- If $p' > 1 - q_B$ but $p^* > p > 1 - q_A$ a social witness challenges with probability q_B^* and acquiesces with probability $1 - q_B^*$. A confident victim challenges with certainty. An insecure victim challenges with probability q_i^* and acquiesces with probability $1 - q_i^*$.

Because there are multiple equilibria, an equilibrium concept is necessary to predict the game's equilibrium. Applying the criterium proposed by Schelling (1960), responders will coordinate to challenge basic predation even without communication because they would have agreed and honor their agreement to challenge basic predation, if they could communicate (tacit bargaining equilibrium). Further, Harsanyi and Selten (1988) suggest that as the uncertainty about the pay-off increases, responders will converge to the equilibrium with the largest Nash-product (risk dominant equilibrium). After BP, the risk dominant SPNE is mutual challenge. Based on either criterium, the strategic profiles where the dictator plays BP with some probability are dominated by the profiles where the dictator plays BP with zero probability.

In the DAC sub-games, it is difficult to argue that there is a tacit bargaining equilibrium. The reward for loyalty makes promises of the witness to challenge non-credible. The Nash product for mutual acquiescence is $t(1-q_A)(1-q_B)$. The Nash-product for the equilibrium challenge is $tpp'q_Aq_B$. Beyond the point where the two Nash products are equal, the pay-off superior equilibrium becomes risk dominant. When $pp' < 1$, the Pareto superior equilibrium is risk dominant, if:

$$1 > pp' > \frac{C^2}{(-C - R + \theta_B)(L - C - R + \theta_A)} \quad \text{(critical probabilities)}$$

Assuming that the condition set above is fulfilled, the dictator is not expected to play DAC because she will anticipate mutual acquiescence with a high probability. Contrary, she is expected to play NC and the responders will not challenge. Otherwise, the dictator is expected to play DAC and both responders are expected not challenge because they will anticipate that the other responder will not challenge.

2.2.3 Comparative statics

Let $r = f(C, L, R)$ denote the riskiness of challenge, where:

$$\frac{\partial r}{\partial C} > 0, \frac{\partial r}{\partial L} < 0 \text{ and } \frac{\partial r}{\partial R} > 0$$

On the one hand, increasing riskiness can deter responders from challenging. First, the critical probabilities with which the other responder has to be confident or social are functions increasing in the riskiness. Second, as riskiness increases, the marginal witness will no longer be social, i.e. p will shrink. This automatically makes victims more insecure, i.e. p' also shrinks. Last, if riskiness is high, mutual challenge is less likely to be the risk dominant equilibrium. Hence, even social and confident responders will acquiesce. All these different factors deterring challenge will be referred to as the deterrence effect of riskiness.

On the other hand, the first derivatives of q_B^*, q_c^* and q_i^* for r are positive. In other words, as long as the responders are not deterred from challenging, they will challenge with a higher probability as riskiness increases. The intuition behind this result is simple. The set of θ's values for which $\theta \geq C+R$ shrinks with riskiness. As a consequence, the probabilities of playing with a social witness or a confident victim shrink.

By increasing the probability with which they challenge, the witness and the victim stop the further shrinkage of the set of the θ's values that ensure that a responder will challenge. Otherwise, there will be no responder left who would be willing to challenge. In layman terms, the responders have to act as zealots to avoid that only zealots would be willing to challenge the dictator. I will call the attempt of the social/confident responders to prevent a cascade due to the deterrence effect of riskiness, the compensatory response.

Provided that a marginally social/confident ($\theta = C + R$) responder anticipates a compensatory response, he has no reason to stop challenging. Which of the two effects prevails or whether they cancel each other out depends on the rate with which responders exit the pool of challengers and the distribution of θ. The following hypotheses will be tested.

Hypothesis 2.1a (Deterrence effect) *The likelihood that the responder challenges decreases as the riskiness of challenge increases.*

Hypothesis 2.1b (Compensatory response) *The likelihood that the responder challenges increases as the riskiness of challenge increases.*

The theoretical insights are based on a number of assumptions. The model treats subjective costs as aversion against predatory actions, which is independent of the game's pay-offs or who is the victim. In addition, what the dictator gets out of the game is not part of the responder's best response function. Responders are indifferent if the dictator leaves the game with nothing. They also do not care if predatory claims lead to welfare increases.

Nevertheless, experimental research has shown that people care about the impact of their choices on the pay-off of others. Inequality aversion, envy and inefficiency aversion are well documented other-regarding preferences (Fehr and Schmidt 1999; Bolton and Ockenfels 2000; Engelmann and Strobel 2004; Cooper and Kagel 2015). Thus, it is possible that the distribution of subjective costs depends on the player's actions and pay-offs. I test for two aspects that could influence subjective costs. Specifically, responders might be less prone to challenge if predatory claims constitute redistribution without losses, i.e. $\alpha = 1$, or financial growth, i.e. $\alpha > 1$.

Hypothesis 2.2 (Inefficiency aversion I) *The likelihood of challenge increases when taking reduces the sum of all the pay-offs.*

Because it is a one-shot game, there is no reason to assume that responders punish the dictator to stop future predatory claims (pro-social punishment). Thus, responders might be averted to challenge because of the destruction of resources following their action, i.e. the dictator's losses. Nevertheless, if responders try to punish pro-socially, i.e. try to teach the dictator a lesson for the future, the exact opposite would be true. Responders would be more likely to challenge the dictator, if they can destroy her endowment.

Hypothesis 2.3 (Inefficiency Aversion II) *The likelihood of challenge increases when the dictator's losses from mutual challenge decrease.*

2.3 Experimental design

2.3.1 Experiment

To test the predictions, I ran a lab experiment using z-Tree (Fischbacher 2007) in the experimental lab of the University of Hamburg.[7] The experiment consisted of four 90 minutes sessions, where the participants earned on average 14.50 Euros for their participation.

In each session, 28 participants played the role of the responders and one participant the role of the dictator. The computer decided the roles, which did not change during the session, and divided the responders randomly in 14 pairs when each session began. All sessions exposed the participants to nine variations of the CR-game.

In detail, the session's participants had to play the CR-game 38 times (9 variations x 4 sub-games + 2 random sub-games of random variations). Each

7. I recruited all participants through hroot (Bock, Nicklisch and Baetge 2012).

time -henceforth period, the dictator was shown which variation of the CR-game will be played in the period and she chose a sub-game. The computer informed the responders, which variation and sub-game were relevant for the period and each responder chose between options X (challenge) and Y (acquiesce) for this sub-game. The computer informed the responders what the other responder chose and their pay-offs from this period. Then, the next period began.

The instructions made clear that participants interact with other participants only in one period and only this period is paying. The computer chose the period randomly before each session and the period was not disclosed to the participants. During the paying period, the dictator chose the sub-game, and the computer communicated her choice to all responder-pairs in the session. Then, the choices of the responders determined the pay-offs.

For the rest of the periods, a pre-generated information feed determined which sub-game of which variation will be played in the period and what the computer will tell the responders that the other responder did. This design choice ensured that all participants saw all variations and all sub-games of the CR-game.

To make sure the instructions were understood, participants answered a set of questions about this aspect of the experiment. Each session began when all session's participants answered all the questions without help correctly. The instructions for the experiment can be found in appendix A.1.1.

Bardsley (2000) coined this design a conditional information lottery. A conditional information lottery generates observations for all outcomes without allowing for an updating of priors. Responders cannot distinguish when the information corresponds to actual choices or is part of the feed. Hence, the optimal strategy is to disregard the information of each previous period and treat each new period as if it is the paying period. Further, participants take one decision at a time, observe its potential implications and re-evaluate their strategy in the next period. An advantage of the conditional information lottery compared to the strategy method is that responders react to a potentially true choice of the dictator and not a hypothetical "how would you react if" question.

The information feed was generated before the experiment began and was common for all sessions. All sub-games of all variations of the CR-game could be played at least once. Two sub-games, which the computer randomly chose, could be played more than once. The responders did not know how many times they were to play each sub-game. The extra periods are used to conceal the paying period in which the dictators chose the sub-game. The information feed can be found in Appendix A.1.3.

In three out of the four sessions the dictator's choice for the paying period coincided with the random information feed. Hence, the responders played the sub-games in an identical order for these sessions. In the one session that the dictator's choice did not coincide with the feed, the responders did not play one sub-game at all and played three instead of two sub-games twice.

At the end of the session, the computer informed the participants what the dictator and their partner chose in the paying period and how much they earned from participating. Then, they filled out a questionnaire about their socio-demographic characteristics and their estimation of what the other participants did in the entire experiment. After the end of the session, the participants were paid individually and no responder could at any point identify the dictator or their partner.

2.3.2 Treatments

Responders come to the lab influenced by unobservable pay-off irrelevant cues. Such cues act as noise. To reduce the noise, I generate a common cue. Before each session began, I asked the participants to vote whether they were in favor of a 20c Euro reward for choosing X. This reward does not change the pure strategy equilibria. Participants knew that the computer would either allow the majority to decide or introduce the reward disregarding the votes. When the computer disregarded the votes, it chose randomly whether to disclose to the participants whether the majority was in favor of the reward.

This procedure allowed later in the analysis to control for a potential distortion by a democracy premium as Dal Bó, Foster and Putterman (2010) and Grossman and Baldassarri (2012) suggest. Before voting, the participants did not know the treatment that applied to their session. During the analysis, I find no evidence that the source of the common cue (vote or randomness) or information about the votes had any impact on behavior.

The computer disregarded the votes in two sessions. From these two sessions, the participants learned what the majority voted for in one session. Since voting in favor of the reward made responders better-off and the probability of being the dictator was small, only 27.7% of all participants voted against the reward. Thus, there was no session where the reward was not introduced.

All participants saw nine variations of the CR-game.[8] The pay-off for mutual challenge was 14 Euros in all variations. Table 2.3 presents the responders' initial endowments (E) and the dictators' initial endowments (E_D), the amount the dictator could take (L), the cost of challenge (C) and the

8. The pay-offs were given in the form of points and at the end of each session, for every point earned in the paying period the participant got 10c Euro.

reward for loyalty (R) in Euros. Further, it reports the portion of the dictator's initial endowment that was lost after fining (F/E_D) and what portion of the money taken reached the dictator's pocket in each variation (α).

Variation	E_D	E	L	C	R	F/E_D	α
Base	15	15	6	1	0	1	0.5
C_L	16	14.5	3	0.5	0	1	0.5
C_H	14	15.5	9	1.5	0	1	0.5
R_M	15	15	8	1	2	1	0.5
R_H	15	15	12	1	6	1	0.5
α_M	15	15	6	1	0	1	1
α_H	15	15	6	1	0	1	1.5
F_L	15	15	6	1	0	0	0.5
F_M	15	15	6	1	0	0.5	0.5

Table 2.3: Pay-off elements of each CR-game variation (in Euro)

The victim's optimal mixed strategy was always to challenge with a probability $\frac{83+E(\theta)}{100+E(\theta)}$. To ensure that the victim's best response function is the same between different variations of the game, the initial endowments and the loot were varied freely. Otherwise, all variations were different in one aspect from the *Base* variation.

Variations C_L and C_H differ in terms of the cost of challenge from the *Base* variation. Variations R_M and R_H differ in terms of the reward for loyalty. Variations α_M and α_H have different αs. Specifically, in variation α_M the sum of pay-offs remains the same after a taking (redistribution). Whereas, taking makes the pie bigger in variation α_H.

Variations F_L and F_M impose a different fine for claiming a part of the responders' endowment. In variation F_L, the dictator keeps her entire initial endowment, i.e. pays no fine after mutual challenge. As a result, the sum of pay-offs after mutual challenge is higher than the sum of pay-offs after mutual acquiescence. In variation F_M, the dictator pays as a fine half her initial endowment. Mutual challenges still lead to welfare losses. However, the victims gain 10 (6-1+6-1) Euro from mutual challenge when the dictator loses only 7.5 Euro from her initial endowment.

2.4 Experimental results

2.4.1 Descriptive statistics

There were 112 responders in total and two of them were non-students.[9] Their ages varied between 18 and 44 years old ($\mu = 25.64$, SD=4.85) with four responders being over the age of 35. 60 responders stated that they were women. Three dictators were women and one was a man.

The dictators chose basic predation 14 times in total. All dictators chose it in the first period. Each female dictator chose basic predation one more time, i.e. two times in total. The male dictator chose basic predation eight times in total. The male dictator not only chose BP more than the female dictators combined but he also chose *DAC* as many times as the female dictators chose *DAC* combined, i.e. 24 out of the 48 times *DAC* was chosen in the four sessions. Because the conditional information lottery disconnects the choices of the dictators from the choices of the responders, I do not discuss the dictators' choices further.

The frequency with which the responders challenged varied between four and 37 times during the 38 periods ($\mu = 16.65$, SD=5.78). The one responder who challenged 37 times submitted her choices faster than the rest of the responders. This behavior could have resulted from boredom, inattention, or disinterest. I drop her from the sample as an extreme outlier. Table 2.4 shows how many times the rest of the responders challenged per sub-game.

Sub-game	Acquiesced	Challenged	Total
BP	73	898	971
DAC	1,202	907	2,109
NC	1,115	23	1,138
Total	2,390	1,828	4,218

Table 2.4: Challenge between sub-games

Responders acted as expected. Only 50 participants never challenged as a witness. Pooling together the times each responder challenged in each sub-game reveals that no responder challenged more frequently as a witness than as a victim. Only one responder challenged more frequently as a victim of DAC rather than as a victim of basic predation ($\chi^2 = 2.81$, p-value=0.09). 76 responders never failed to challenge after BP. The rest challenged between

9. Non-students do not challenge less often than the students in the sample. As a robustness test, I also estimate the regressions without the non-students. The results do not change significantly. Also, in a random effects regression, being a student does not have a significant effect on the likelihood to challenge.

one and five out of the ten or eleven times -depending on the session- they played the BP sub-game.

Eleven responders challenged after NC (no claim from the dictator). The distribution of challenges after NC pro responder is: one responder did it six times, one responder did it five times, one did it three times, one did it twice and the rest once. From these eleven responders, six challenged NC in the seventh period. In this period, the responders played NC for the first time. The four preceding periods the computer informed them that the dictator chose basic predation. This combination of events could have led to slips of the hand. Four of the responders challenged no claims only in the 37th period. For these cases, mental fatigue could have played a role.

No systematic correlation of challenges after NC and other characteristics of the period in which the decision was made (treatment, number of period etc.) or the responder's characteristics can be found. As a robustness test, I always also estimate regressions where I exclude the two responders which challenged NC the most frequently from the sample. Their exclusion has no significant impact on the regressions' results.

2.4.2 Empirical strategy

As a next step, I estimate a number of fixed effect linear models to determine what makes responders challenge with a higher probability. The full model is:

$$
\begin{aligned}
Challenge_{it} = & \beta_1 \text{Sub-game}_{it} \\
& + \beta_2 \text{Treatment}_{it} \\
& + \beta_3 (\text{Sub-game}_{it} \times \text{Treatment}_{it}) \\
& + \gamma \text{Controls}_{it} + u_i + \epsilon_{it}
\end{aligned}
\tag{2.2}
$$

where Sub-game is a matrix containing three dummy variables. The variables capture whether the responder was a victim of basic predation (BP), the victim of DAC (DAC_v) or the witness (DAC_w) in the observed period. The baseline of comparison is being neither a victim nor a witness (NC sub-game). Treatment is a matrix containing one dummy for each variation of the CR-game, excluding the Base variation. u_i are the responders' fixed effects and ϵ is the random error term. In the model, β_3 captures whether and how much a responder adjusts his response to riskiness or different dictator's pay-offs depending on whether he is a victim of DAC/BP or a witness.

Last, matrix Controls contains additional controls. Particularly, I control for two sources of noise: experience and fatigue. It is possible that responders optimize their strategies and learn how to play as they play. The times the responder played any sub-game of a variation and any variation of a sub-game in previous periods serve as measures of previous exposure to the strategic situation in which he decides.

What the computer presents as the dictator's and the other responder's choices is the same between responders because of the common information feed. However, the responder's choices influence how they experience the information they get. Hence, I include dummies for the sub-game and the response of the other responder in the previous period. Although they are irrelevant for the following period, they can influence a participant's emotional state. For example, participants might get disappointed that they did not coordinate or excited if they experience coordination success.

Participants often try to avoid mental effort even if that means making mistakes (Kool et al. 2010). Moreover, as fatigue increases, concentration on the task becomes difficult and the cost of effort increases. According to Helton and Russell (2015), the time elapsed between the completion of a task and the beginning of a new task is a reliable predictor of the capacity to concentrate on the new task. In the experiment, a new period began when all participants chose an action. Depending on how fast a responder answered compared to the slowest responder, he had time to regroup and relax for the next period. The computer recorded automatically the seconds elapsed from the moment the server (experimenter's PC) send the signal to the clients (participants' PCs) and the moment the server received the reply from each client.[10] From this information, I calculate the rest time, which is included in the model. Responders rested on average approximately a minute between choice and new information.

Kahneman (2003, 2011) has shown that individuals use intuition (System-I) when they perceive a decision as trivial or time-sensitive. Further, research has shown that fast decisions are more likely to be in favor of redistribution or compliant to social norms (Chen and Fischbacher 2020; Caplin and Martin 2016; Alós-Ferrer, Garagnani and Hügelschäfer 2016; Rand 2016). To control for intuitive rather than deliberative responses the matrix contains a dummy for System-I thinking. Rand, Greene and Nowak (2012) suggest that participants use System-I thinking when they decide within a ten seconds window. One dummy is constructed, which takes the value of one when the server received the reply of the client within ten seconds from the dispatch of the signal, as the literature advises. The fastest decision was taken in approximately eight seconds and 155 decisions were faster than the ten seconds threshold (3.6% of all decisions).

10. The rest times are accurate. The lab in Hamburg connects all clients directly via a router to the z-Tree server. The resulting network delay is below 1ms.

For the ten seconds threshold to be meaningful, participants must be able to immediately execute their choice. However, during the experiment, responders had to use a mouse and find the button on a screen. Hence, I construct two additional dummies to control for the delay in the execution of a decision. The one captures whether the responder decided within twelve seconds (6.3 % of all decisions) and the other within fifteen seconds (10.3% of all decisions). I run regressions with each of the three dummies. The results are robust to different specifications of the Controls matrix. I report the results with the fifteen seconds dummy because it produces the models with the best fit.

I expect responder's characteristics which are time invariant to impact the propensity to challenge. Nevertheless, I am not interested to estimate their impact on the decision to challenge. Subsequently, I use a fixed effects specification. Based on a Sargan-difference test as proposed by Arellano (1993) and Wooldridge (2002, p.290) a random-effects estimator would be inconsistent ($\chi^2 = 137$, p-value< 0.000). Additionally, I cluster the standard errors on the responder's level.

Although the outcome is binary, I rely on linear models instead of logistic regressions for several reasons. The major advantage of the linear model is the easiness of interpreting the coefficients of the interaction terms. The focus of the chapter is the likelihood of challenge as a function of the interaction between the pay-off structure (variations) and being a victim or a witness. The coefficients of interaction terms in the linear models describe how much more likely is a challenge in a certain variation when the responder is a victim or a witness (additive relationship).

Contrary, the coefficients of interaction terms in non-linear models represent a multiplicative relationship. Because the size of the interaction effect depends on the value of the covariates, the coefficients do not capture the full interaction effect on the likelihood of challenge. Further, insignificant coefficients do not necessarily mean that the propensity of the responders to challenge is not significantly different in two treatments or that the effect of one treatment is not modulated by another treatment (Norton, Wang and Ai 2004; Cornelißen and Sonderhof 2009; Greene 2010). Last, there is currently no unbiased estimation process for the marginal effects of interaction terms in panel non-linear models (Greene and Zhang 2019). As a result, with non-linear models it is impossible at the moment to gain an understand about the size of the effect of interrelated variables.

Another advantage is that fixed effect linear models are consistent with cluster-adjusted standard errors. Responders in the experiment play consistently over time and are influenced in their future choices by their past experiences in the game. As a result, different individuals can respond differently to treatments over time. All the existing non-linear equivalents to the linear fixed effect specification require that there is no serial correlation in the idiosyncratic error (Wooldridge 2019).

2.4.3 Regression analysis

Table 2.5 reports the results of regression analysis. The model reported in column one regresses the dependent variable only on the Sub-game matrix. The model reported in column two also contains the Treatment matrix. In column three, the controls are added. In column four, the Treatment matrix is interacted with the Sub-game matrix for the first time.

The model reported in column five contains all the variables of the full model plus a variable for the times a responder challenged in previous periods. Previous moral behavior can create a pressure for consistency or license individuals to act immorally in future rounds (Mullen and Monin 2016). This variation is not captured by the fixed effects because the number of challenges in the past evolves over time. For comparability with the other coefficients reported in the table, columns four and five do not report the coefficients of the model but the marginal effect of the variables calculated based on the results of the model.

The reported coefficients show by how much the likelihood to challenge increases in each sub-game in comparison to the likelihood to challenge in the NC sub-game. As already presented, 100 responders never challenged NC, seven challenged NC once and only four responders challenged NC several times. In other words, the likelihood to challenge NC is not significantly different than zero. As a result, the coefficients can be interpreted as the likelihood to challenge in each sub-game.

The size and the sign of the coefficients give no reason to question the conjecture that the responder's utility function contains an element $\theta > 0$. Specifically, the coefficients for BP are all greater than the predicted mixed strategy when the subjective cost is expected to zero by the participants.[11] One could however argue that the result can be explained by the fact that mutual challenge is the risk and pay-off dominant equilibrium, when the dictator chooses BP. More definitive evidence supporting the subjective costs conjecture is that the coefficient for DAC_w which captures the likelihood that a witness will challenge is robustly significantly different than zero.

The likelihood of challenge was smaller than optimal, when the dictator used divide-and-conquer predation. If the average victim had consistent beliefs, i.e. expected: 1) 62 of the witnesses to be social ($p = 0.55$), and 2) that

11. Wald-test $H_0 : \partial BP = 0.833$, χ^2=14, p-value< 0.000.

	DV: i-th responder's decision in t-th period				
	(1)	(2)	(3)	(4) (ME)	(5) (ME)
DAC_v	0.627***	0.625***	0.618***	0.614***	0.640***
	(0.031)	(0.031)	(0.031)	(0.036)	(0.037)
DAC_w	0.193***	0.192***	0.183***	0.171***	0.194***
	(0.028)	(0.029)	(0.028)	(0.028)	(0.035)
BP	0.905***	0.903***	0.894***	0.889***	0.919***
	(0.017)	(0.017)	(0.019)	(0.019)	(0.024)
Past CH					0.014***
					(0.004)
Treatment	No	Yes	Yes	Yes	Yes
Controls	No	No	Yes	Yes	Yes
Adj-R^2	0.583	0.584	0.586	0.588	0.591

Observations: 4,218 from 111 participants. Clustered on the participant's level robust SE are in parentheses. Controls: rest time (seconds), times the variation and the sub-game were played in $T < t$, other responder's choice in $t-1$, sub-game in $t-1$ and a System-I thinking (<15s) dummy. Participant's dummies are used as fixed effects. (*) p<0.1; * p<0.05; ** p<0.01; *** p<0.001.

Table 2.5: Likelihood of challenge

a social witness challenges with a likelihood equal to 0.17 (q_v^*), they would challenge with a likelihood equal to 0.83. Excluding the 29 victims, who always challenged and the four victims who always acquiesced after DAC, only 20% of the rest of the victims challenged DAC with a likelihood equal to 0.83 or higher.

The victim of DAC is estimated to challenge with a likelihood a little higher than 0.61. Such a low likelihood of challenge can be explained by expectations that the witnesses will not be social $p < 0.16$. Sadly, the exact size of coordination failure is difficult to estimate without a precise measure of the victim's θ and beliefs about the distribution of θ. To see how the likelihood of mutually challenging BP or DAC changed depending on the CR-game's variation, I replace the Sub-game matrix with a dummy taking the value of one when the dictator did not choose NC. Table 2.6 reports the marginal effect based on this model.

The likelihood of challenge decreases when the cost of challenge decreases. This finding is in line with the deterrence hypothesis. The likelihood of challenge also decreases when the reward for loyalty is introduced but it is

Marginal effects: CR-game variations				
	ME	Robust SE	95% Conf. Interval	p-value
C_L	-0.301	0.025	-0.079 : 0.019	0.227
C_H	-0.038	0.022	-0.082 : 0.006	0.093
R_M	0.045	0.024	-0.002 : 0.093	0.061
R_H	-0.031	0.023	-0.078 : 0.015	0.179
α_M	0.069	0.032	0.006 : 0.131	0.032
α_H	-0.040	0.018	-0.077 : -0.003	0.033
F_L	0.056	0.019	0.018 : 0.094	0.004
F_M	0.103	0.029	0.045 : 0.160	0.001

The coefficients are based on the full model with the variable for previous challenges. Observations: 4,218 from 111 participants. Adj. R^2: 0.425.

Table 2.6: Marginal response to pay-off manipulation

not at its highest level. Contrary, when the reward is at its highest level the responders are as likely to challenge the dictator as they are when there is no reward. This finding is consistent with the compensatory response hypothesis and the theoretical finding that the deterrence effect kicks-in only when riskiness exceeds a certain level.

Surprisingly, the responders are more likely to challenge a dictator who is efficient and pockets everything that she steals rather than a dictator with a leaky bucket, i.e. one who gains less from stealing than what she steals. Less surprising, when the dictator makes the pie bigger, even though it is in her favor they are less likely to challenge.

Last, the likelihood of challenge increases, when the size of the fine decreases. The responders are 5.6 percentage points more likely to challenge, when they play the F_L variation, where there is no fine, compared to the Base variation. Further, they are 10.4 percentage points more likely to challenge compared to the Base variation in the F_M variation, where the fine is half the dictator's initial endowment.

The difference of the two marginal effects is significant.[12] Nevertheless, the likelihood of challenge does not linearly decrease in the size of the fine, as would be expected if the responders cared purely for efficiency. It seems that when the fine decreases the responders indeed find it easier to challenge due to less inefficiency introduced by their actions. Notwithstanding, when there is no fine they do not want to invest the resources to challenge the dictator. One explanation for such a behavior could be that they perceive

12. Wald-test $H_0 : \frac{\partial Challenge_i}{\partial F_L} = \frac{\partial Challenge_i}{\partial F_M}$, χ^2=3, p-value= 0.086.

a challenge as more appealing when "it teaches the dictator a lesson" (prosocial punishment) but the lesson has to be proportional to the wrong it wants to discourage.

Based on the aforementioned results, the null cannot be rejected for both hypotheses 2.2 and 2.3 (inefficiency aversion I and II). Responders seem to care about efficiency to some degree. Nevertheless, there is a trade-off between efficiency and setting incentives for the dictator. The non-results are interesting from a methodological point of view. They show the importance of controlling for the pay-off structure of the CR-game and how parameterization can bias the findings.

To investigate the validity of the deterrence and compensation hypotheses, I will now return to the main model and look how the responders adjust the likelihood with which they challenge depending on the variation of the CR-game and whether they were the victim. The model whose coefficients are reported in table 2.6 assumes that there is no heterogeneity between the witnesses and the victims and pools the observations from the BP, DAC_A and DAC_B sub-game. Subsequently, the model's results tell how the likelihood that a predatory claim will be mutually challenged changes between the variations averaging the response of victims and witnesses.

Not going further than the effect on the average responder leads to misguided conclusions. The equilibrium strategy is different between sub-games and victims and witnesses are expected to respond differently to the treatments. It is possible that the insignificant coefficients are the result of canceling out. Table 2.7 reports the coefficients for the variables in matrices Sub-game and Treatment and their interaction from the full model with a control for past challenges.

			DAC_v		DAC_w		BP	
			+0.440***	(0.052)	+0.159***	(0.038)	+0.852***	(0.042)
C_L	−0.033	(0.034)	+0.166*	(0.075)	+0.037	(0.063)	−0.000	(0.060)
C_H	−0.065(*)	(0.034)	+0.151*	(0.065)	−0.016	(0.057)	+0.117(*)	(0.065)
R_M	−0.118(*)	(0.065)	+0.372***	(0.089)	+0.117	(0.093)	+0.124(*)	(0.065)
R_H	−0.040(*)	(0.022)	+0.249***	(0.069)	+0.011	(0.057)	+0.137**	(0.050)
α_M	−0.056	(0.087)	+0.248**	(0.085)	+0.055	(0.099)	+0.102	(0.091)
α_H	−0.061	(0.047)	+0.155*	(0.076)	+0.037	(0.070)	+0.078	(0.069)
F_L	−0.040(*)	(0.023)	+0.266***	(0.069)	+0.074	(0.048)	+0.046	(0.061)
F_M	+0.031	(0.021)	+0.254**	(0.089)	+0.007	(0.073)	+0.033	(0.057)

Observations: 4,128 from 111 participants. The coefficients are based on a dynamic linear fixed effects model. Clustered on the participant's level robust standard errors are in parentheses. Adj R^2: 0.591. *** p<0.001, ** p<0.01, * p<0.05, (*) p<0.1.

Table 2.7: Likelihood of challenge: Interaction effects

The gray cells show the difference to the likelihood of challenge between the Base variation and other variations (vertically) and NC and other sub-games (horizontally). The other coefficients represent the additional change to the likelihood of challenge in the variation for the specific sub-game or equivalently the additional change to the likelihood of challenge for a specific sub-game, when each variation of the CR-game is played. The coefficients in the first row are smaller than the marginal effects shown in table 2.5. However, this is no reason for concern because the results are qualitative the same. To avoid repetition, I will directly present the rest of the results.

Participants were less likely to challenge in the C_H variation than in the Base variation. The likelihood of challenge also decreased when a reward for loyalty was provided -even though a higher reward did not lead to significantly less challenges.[13] These results confirm the deterrence hypothesis (Hypothesis 2.1a). Whenever the deterrence effect kicked-in (variations C_H, R_M and R_H) the victims' likelihood of challenge increased, independent of whether the dictator wanted to take from the endowments of one or both responders. The interaction terms of C_H, R_M and R_H and the DAC_v and BP dummies are positive and significantly different than zero. Despite the deterrence effect, the average victim was more likely to challenge when riskiness got higher.[14] In other words, the compensatory response is stronger than the deterrence effect for the victims.

Victims of DAC also challenge with a higher likelihood in the C_L variation. C_L was one of the two variations which were played more than once by all participants. The positive and significant coefficient is driven by the fact that the second time responders face DAC in this variation, the victims are 56.5% more likely to challenge.

The non-victims do not respond to the riskiness of challenge differently when they are witnesses and when the dictator makes no claim. This is evidence that they do not try to compensate the victims to convince them to challenge. The result can further be interpreted as an indication that social witnesses do not expect the compensatory response. If they did, they would have challenged with a higher likelihood, because as the victim challenges with a higher likelihood, the social witness profits more from challenging.

The results indicate that increasing the riskiness of challenge reduces coordination failure. In particular, the witnesses do not adjust their response but the victims who challenged less than optimal challenge more frequently. This result is not surprising. As the stakes increase, participants try to play

13. Wald-test $H_0 : \beta_{R_M} = \beta_{R_H}$, χ^2=1.4, p-value= 0.243.

14. Wald-tests H_0: $\beta_{BP}+\beta_{BP\times K_H}$=0, $\chi^2 = 275$, p-value< 0.000; H_0: $\beta_{BP}+\beta_{BP\times R_M}$=0, $\chi^2 = 421$, p-value< 0.000; H_0: $\beta_{BP}+\beta_{BP\times R_H}$=0, $\chi^2 = 484$, p-value< 0.000; H_0: $\beta_{DAC_v}+\beta_{DAC_v\times K_H}$=0, $\chi^2 = 83$, p-value< 0.000; H_0: $\beta_{DAC_v} + \beta_{DAC_v\times R_M}$=0, $\chi^2 = 127$, p-value< 0.000; H_0: $\beta_{DAC_v} + \beta_{DAC_v\times R_H}$=0, $\chi^2 = 165$, p-value< 0.000.

the game more rationally.

The victim of *DAC* challenges with a higher likelihood as the fine increases. Although this response is not rational in the one-shot game, it is intuitive. A low fine does not deter future violations. Although the game is one-shot, it is not uncommon that participants in experiments try to teach other participants a lesson for their life after the experiment.

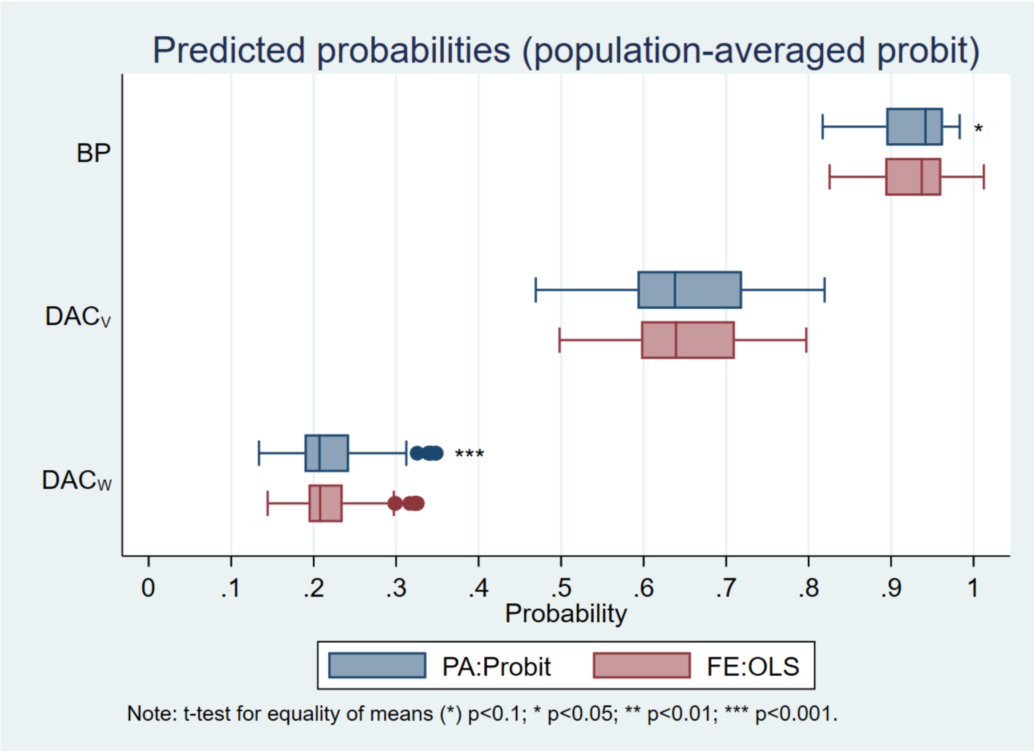


Figure 2.1: Predicted probabilities pro sub-game of the CR-game

Despite the advantages of the linear model, the predictions for high probabilities, like the probability to challenge BP, and low probabilities like the probability to challenge as a witness, can be inconsistent. As a robustness test, I estimate a population-averaged probit model with robust standard errors. The coefficients of this model are reported in Appendix A.1.4. Because statistical comparison of the coefficients between the linear and the non-linear model is not informative, I use the predicted probabilities to compare the results. Figure 2.1 visualizes with boxplots the distribution of the predicted probabilities from the two models. As can be seen, the predicted probabilities are not significantly different from the predictions of the linear model.[15]

15. A two-sided paired t-test ($H_0: \mu_{Probit} \neq \mu_{OLS}$) confirms that the results are not

To sum up, the experiment provides clear evidence that participants expect to experience a non-monetary benefit for stopping the dictator from taking. Nevertheless, their beliefs about how likely is their partner to challenge are inconsistent. As a result, they fail to challenge with the optimal likelihood. The compensatory response is stronger than the deterrence effect for the victims. Contrary, the deterrence effect is stronger for the witnesses. Based on the results of the experiment, I can reject the null for hypothesis 2.1a for the witnesses but not for the victims and I can reject the null for hypothesis 2.1b for the victims but not the witnesses of predation.

2.5 Conclusion

The focus of this chapter is whether changes to the incentives deter or encourage citizens to challenge the executive who misuses her power for personal gain. I investigate three factors which could influence the probability of a challenge: the riskiness of challenge (the cost of challenge and whether there is a reward for loyalty), the distributional effects of predatory actions and the destructiveness of a successful challenge.

Theoretical analysis reveals two forces influencing behavior. When the monetary reward for loyalty is higher than the monetary benefit of challenge responders will be deterred from challenging. However, the optimal strategy is also a function of how much responders suffer psychologically when they lose part of their initial endowment or their partner loses part of his endowment. Given that this psychological cost is sufficiently high, responders will challenge with a higher probability when riskiness increases to counteract the aforementioned deterrence effect.

Victims of takings underestimate the willingness of the other responders to challenge when the dictator uses divide-and-conquer tactics. This result of the experiment reveals the existence of a dangerous backdoor to reduce citizens' control over the executive. Discriminatory laws or unequal application of the law (biased enforcement or use of discretionary measures) allows the dictators/executives to manipulate the riskiness of challenge asymmetrically for different parts of the society. If such strategies are available to the executive, the executive could maximize deterrence and minimize compensatory responses.

The victims are more likely to challenge when the riskiness goes up, as the compensation hypothesis predicts. However, the witnesses of divide-and-conquer predation underestimate the willingness of the victims to challenge. As a result, the witnesses do not adjust the probability with which they

significantly different: p-value< 0.000

challenge accordingly. Interestingly, this discrepancy means that as the riskiness of challenge increases, coordination failure decreases. In other words, individuals are more likely to play equilibrium strategies.

Last, the experiment does not show clear evidence that the responders care about the final pay-offs as a whole. Specifically, there is no robust evidence that a dictator who increases the size of the pie and a dictator who wastes resources by stealing are not treated the same way. Moreover, there is evidence that responders want to punish the dictator who committed a predatory act. As a result, the likelihood of challenge robustly decreases when the dictator pays no fine for trying to steal from the responders. At the same time, when the fine is too high responders are less likely to challenge. Such a behavior reveals the existence of a dilemma between setting incentives for the dictator and minimizing waste.

The findings are applicable to many fields of social life. For instance, Posner, Spier and Vermeule (2010) give a very elaborate account of situations where individuals face the problem of controlling an administrator, director or executive. For example, minority shareholders cannot stop single-handedly a CEO who acts against their interests. To protect their interests, they can either sell out or get the support of other shareholders. Research shows that companies where the latter response prevails have a higher company's share value (Denes, Karpoff and McWilliams 2017). The insights of the experiment expand our understanding of what facilitates shareholder activism.

So far, I stressed the advantages of experiments. Notwithstanding, there are also disadvantages as by any method. In the lab, the stakes are low. The experiment simulates the small acts of everyday resistance, for example demonstrating, filing lawsuits and signing petitions to stop an act of the executive which is illegal. However, it bears little resemblance to revolutions, where citizens risk their well-being. Last, revolts do not lead to an instantaneous success or such a clear result as in the lab. It is often the accumulation of small victories and continuous struggle that lead to the end result. Although, this argumentation is valid, it exaggerates the goals and instances of challenge. Not many of us will be asked to lead the opposition against a repressive military regime. Maybe if the mechanism that enable everyday resistance are in place we will not have to.

References

Acemoglu, Daron, and James A Robinson. 2006. De Facto Political Power and Institutional Persistence. *American Economic Review* 96 (2): 325–330.

Acemoglu, Daron, and James A Robinson. 2012. *Why Nations Fail: The Origins of Power, Prosperity, and Poverty.* 2nd. London: Profile Books LTD.

Aidt, Toke S, and Raphaël Franck. 2013. How to get the snowball rolling and extend the franchise: Voting on the Great Reform Act of 1832. *Public Choice* 155 (3-4): 229–250.

———. 2015. Democratization Under the Threat of Revolution: Evidence From the Great Reform Act of 1832. *Econometrica* 83 (2): 505–547.

Aidt, Toke S, and Peter S Jensen. 2014. Workers of the world, unite! Franchise extensions and the threat of revolution in Europe, 1820–1938. *European Economic Review* 72:52–75.

Aidt, Toke S, and Gabriel Leon. 2016. The Democratic Window of Opportunity: Evidence from Riots in Sub-Saharan Africa. *Journal of Conflict Resolution* 60 (4): 694–717.

Alós-Ferrer, Carlos, Michele Garagnani and Sabine Hügelschäfer. 2016. Cognitive reflection, decision biases, and response times. *Frontiers in Psychology* 7 (1402): 1–21.

Arellano, Manuel. 1993. On the testing of correlated effects with panel data. *Journal of Econometrics* 59 (1-2): 87–97.

Bardsley, Nicholas. 2000. Control without deception: Individual behaviour in free-riding experiments revisited. *Experimental Economics* 3 (2000): 215–240.

Bock, Olaf, Andreas Nicklisch and Ingmar Baetge. 2012. *hroot: Hamburg registration and organization online tool.* WiSo-HH.

Bolton, Gary E, and Axel Ockenfels. 2000. ERC : A Theory of Equity, Reciprocity, and Competition. *American Economic Review* 90 (1): 166–193.

Caplin, Andrew, and Daniel Martin. 2016. The dual-process drift diffusion model: Evidence from response times. *Economic Inquiry* 54 (2): 1274–1282.

Cason, Timothy N, and Vai Lam Mui. 2007. Communication and coordination in the laboratory collective resistance game. *Experimental Economics* 10 (3): 251–267.

———. 2013. Coordinating resistance through communication and repeated interaction. *Economic Journal* 124 (574): 226–256.

———. 2015a. Individual versus group play in the repeated coordinated resistance game. *Journal of Experimental Political Science* 2 (1): 94–106.

———. 2015b. Rich communication, social motivations, and coordinated resistance against divide-and-conquer: A laboratory investigation. *European Journal of Political Economy* 37:146–159.

Chen, Fadong, and Urs Fischbacher. 2020. Cognitive processes underlying distributional preferences: A response time study. *Experimental Economics* 23:421–446.

Chong, Dennis, Mingxing Liu and Qi Zhang. 2016. Collective resistance under authoritarianism: Elite–mass strategies in an experimental Game. *Political Behavior* 38 (4): 951–976.

Cooper, David J, and John H Kagel. 2015. Other-regarding preferences: A selective survey of experimental results. Chap. 4 in *The handbook of experimental economics vol.2,* 217–274. Princeton, NJ: Princeton University Press.

Cornelißen, Thomas, and Katja Sonderhof. 2009. Partial effects in probit and logit models with a triple dummy-variable interaction term. *Stata Journal* 9 (4): 571–583.

Dal Bó, Pedro, Andrew Foster and Louis Putterman. 2010. Institutions and behavior: Experimental evidence on the effects of democracy. *American Economic Review* 100 (5): 2205–2229.

Denes, Matthew R., Jonathan M. Karpoff and Victoria B. McWilliams. 2017. Thirty years of shareholder activism: A survey of empirical research. *Journal of Corporate Finance* 44:405–424.

Engelmann, Dirk, and Martin Strobel. 2004. Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments. *American Economic Review* 94 (4): 857–869.

Fehr, Ernst, and Klaus M Schmidt. 1999. A theory of fairness, competition and cooperation. *Quarterly Journal of Economics* 114 (3): 817–868.

Fischbacher, Urs. 2007. *Z-Tree: Zurich toolbox for ready-made economic experiments.*

Greene, William. 2010. Testing hypotheses about interaction terms in nonlinear models. *Economics Letters* 107 (2): 291–296.

Greene, William, and Qiushi Zhang. 2019. Nonlinear and Related Panel Data Models. Chap. 3 in *Panel data econometrics,* edited by Mike Tsionas, 45–96. London: Academic Press.

Grossman, Guy, and Delia Baldassarri. 2012. The Impact of Elections on Cooperation: Evidence from a Lab-in-the-Field Experiment in Uganda. *American Journal of Political Science* 56 (4): 964–985.

Harsanyi, John C, and Reinhard Selten. 1988. *A general theory of equilibrium selection in games.* Cambridge, MA: MIT Press.

Helton, William S, and Paul N Russell. 2015. Rest is best: The role of rest and task interruptions on vigilance. *Cognition* 134:165–173.

Kahneman, Daniel. 2003. Maps of bounded rationality: Psychology for behavioral economics. *American Economic Review* 93 (5): 1449–1475.

———. 2011. *Thinking Fast and Slow.* 499. New York, NY: Farrar, Straus / Giroux.

Kool, Wouter, Joseph T McGuire, Zev B Rosen and Matthew M Botvinick. 2010. Decision making and the avoidance of cogntive demand. *Journal of Experimental Psychology: General* 139 (4): 665–682.

Mullen, Elizabeth, and Benoît Monin. 2016. Consistency versus licensing effects of past moral behavior. *Annual Review of Psychology* 67 (1): 363–385.

North, Douglass C. 1990. *Institutions, Institutional Change and Economic Performance.* 164. Cambridge, MA: Cambridge University Press.

North, Douglass C, and Barry R Weingast. 1989. Constitutions and commitment: The evolution of institutions governing public choice in seventeenth-century England. *Journal of Economic History* 49 (4): 803–832.

Norton, Edward C, Hua Wang and Chunrong Ai. 2004. Computing interaction effects and standard errors in logit and probit models. *The Stata Journal: Promoting communications on statistics and Stata* 4 (2): 154–167.

Posner, Eric A, Kathryn E Spier and Adrian Vermeule. 2010. Divide and Conquer. *Journal of Legal Analysis* 2 (2): 417–471.

Rand, David G. 2016. Cooperation, fast and slow: Meta-analytic evidence for a theory of social heuristics and self-interested deliberation. *Psychological Science* 27 (9): 1192–1206.

Rand, David G, Joshua D Greene and Martin A Nowak. 2012. Spontaneous giving and calculated greed. *Nature* 489 (7416): 427–430.

Rigdon, Mary L, and Doug Smith. 2010. The role of solidarity and reputation building in coordinating collective resistance.

Schelling, Thomas C. 1960. *The Strategy of Conflict.* 322. Cambridge MA: Harvard University Press.

Stahl, Dale O, and Paul W Wilson. 1995. On player's models of other players: Theory and experimental evidence. *Games and Economic Behavior* 10 (1): 218–254.

Tarrow, Sidney G. 1994. *Power in Movement.* 328. New York, NY: Cambridge University Press.

Weingast, Barry R. 1995. The economic role of political institutions: Market-preserving federalism and economic development. *Journal of Law, Economics, and Organization* 11 (1): 1–31.

———. 1997. The Political Foundations of Democracy and the Rule of Law. *American Political Science Review* 91 (2): 245–263.

Wintrobe, Ronald. 2000. *The Political Economy of Dictatorship.* Cambridge MA: Cambridge University Press.

Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data.* 752. Cambridge, MA: MIT press.

———. 2019. Correlated random effects models with unbalanced panels. *Journal of Econometrics* 211 (1): 137–150.

Chapter 3

An eternal constitution: A silver bullet or a poisonous pill

3.1 Introduction

One common characteristic of big-C constitutions is that they are what legal scholars call *generally entrenched*. This means that their provisions are harder to revise compared to laws. In an often-used metaphor, the constitution is compared to the rope which tied Ulysses to the mast allowing him to resist the charm of the Sirens and drive his ship to safety. Extending the metaphor, entrenchment is the wax in the ears of Ulysses' sailors, which prevents them from hearing his pleas to be unbound.

Unamendability, the highest degree of entrenchment, can be divided in three broad types. Absolute unamendability (or eternity clauses) prohibits all amendments of a constitutional provision. Temporal unamendability (or constitutional locks) prohibits the amendment of constitutional provisions during certain time intervals. For example, some countries in transition prohibited the amendment of the entire constitution from anything between a month and multiple legislative period after its introduction (Gilbert, Guim and Weisbuch 2019). Last, situational unamendability prohibits the amendment of the constitution when certain conditions are fulfilled. For instance, some constitutions prohibit amendments during a state of emergency (Bjørnskov and Voigt 2018).

If the spread of a constitutional rule is a metric for its success, absolute unamendability carries the day compared to other forms of unamendability. Figure 3.1 shows the percentage of constitutions which also contain different types of unamendability from 1901 to 2015.[1] It can be seen that since World

1. The values of the figure are calculated based on the following two datasets: number

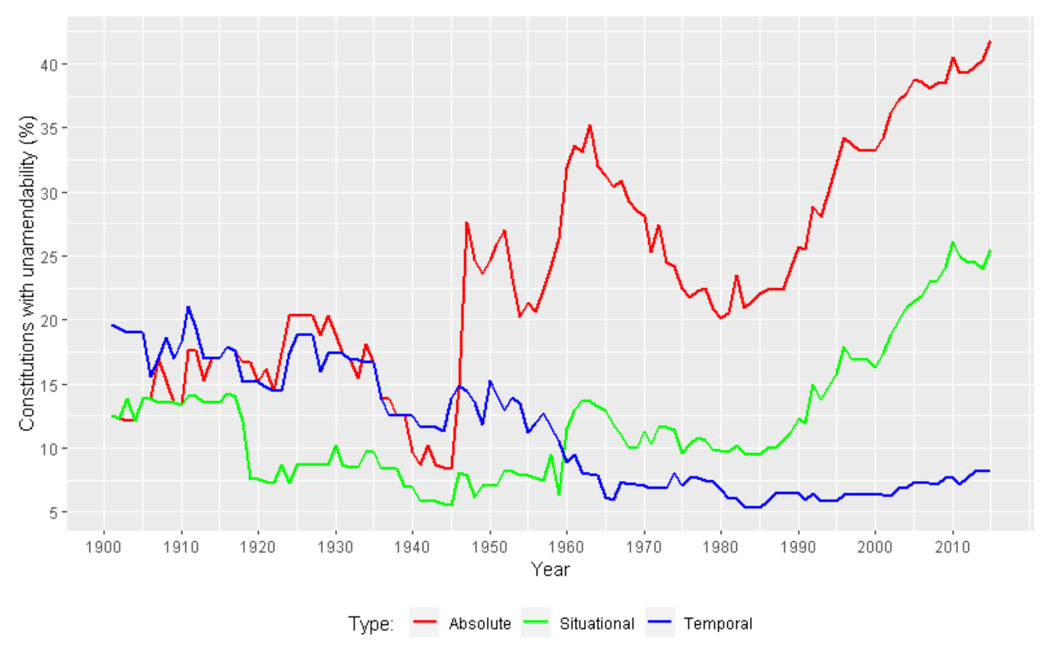


Figure 3.1: Time series plot: Percentage of written constitutions with unamendable provisions

War II absolute unamendability has been the predominant form of unamendability. Today, every second constitution in force prohibits the amendment of at least one of its provisions.

Eternity clauses follow constitutions from their birth. Indicatively, less than ten constitutions globally the last 150 years introduced eternity clauses with an amendment (Pilpilidis 2018, Fn.14). By taking certain issues out of the hands of the legislator, unamendability is meant to protect parts of the constitution from frequent change. However, it entails that a generation which is no longer in power decides the limits of the law. As the population changes over time, not being able to adjust the constitution can lead to constitutional replacement or a decline in constitutional compliance.

At the time of writing this chapter, we know little to nothing about whether not being able to adjust a provision to its time has an impact on constitutional compliance or on the evolution of unconstitutional preferences. This chapters narrows this gap. In particular, I use agent-based modeling to model how the probability changes over time that a legislator will come to power who will try to violate a constitutional provision given that the

of constitutions with unamendability (Hein 2018), total number of constitutions (Zackary Elkins, Ginsburg and Melton 2020).

provision cannot be amended. To the best of my knowledge, this chapter constitutes one of the first applications of agent-based modeling to answer a question from the field of constitutional economics.

The agents represent citizens in a society who try to maximize their income over a number of legislative sessions. The constitution is exogenously given in the model, i.e. the agents cannot re-draft or amend it. Every legislative session, agents elect a legislator who introduces laws. Laws represent legal constraints to financial activities, which are perfectly enforced. These laws can belong to the set of laws allowed by the constitution or not. If an agent manages to comply with the legislator's laws, he earns income. Otherwise, he pays fines.

How much an agent earns depends on how close the laws in force are to an agent's ideal laws. An agent's ideal laws represent the optimal legal framework for an agent to generate income. Because the agent is allowed to change the way he generates income based on his beliefs about which will be the laws in the future, the agent's ideal laws change from one legislative session to another.

The constitution is meant to limit which laws the legislator can propose. Nevertheless, whether it is enforced depends on the constraints put to the legislator. I implement four types of constraints to the legislator. First, the legislator is only allowed to introduce constitutional laws (strict constraints). In case the legislator introduces an unconstitutional law, the law is reviewed and annulled. Constitutional review happens automatically and without delay or uncertainty about the result, i.e. there is perfect constitutional enforcement.

Second, the legislator is allowed to introduce any law (no constraints). Unconstitutional laws are not reviewed, annulled or opposed by the agents. Third, the laws are reviewed but the legislator has the right to call for a referendum to ask for the permission of the agents to introduce an unconstitutional law (weak constraints). Last, the legislator is only an agenda-setter and all laws have to be ratified by a referendum to come to force (direct democracy).

I find that the probability that a legislator will get elected who will want to introduce unconstitutional laws drops significantly with strict constraints. The legislator being submitted to weak constraints increases both the probability of violations and the probability that a legislator who will try to violate the constitution comes to power. Indicatively, a legislator who can introduce any law she wants and a legislator who has to ask for the majority's consent to introduce unconstitutional laws are almost as likely to violate the constitution.

I further find that when a constitution is perfectly enforced at the start of the game, it becomes self-enforcing, i.e. it is less likely that legislators will come to power that will try to introduce unconstitutional laws. This remains true even if after a few legislative sessions the legislator is allowed to introduce any laws she wants without having to face constitutional review, i.e. if the constraints are removed. Respectively, if the laws are not reviewed in the first legislative sessions, even if the constitution is enforced perfectly later, it is very likely that the agents that are elected as legislators will try to violate the constitution.

The results suggest that combining unamendability with perfect enforcement in the beginning of a constitution's life is a silver bullet against future violations and the emergence of majorities or politicians in power that might jeopardize the rule of law. Nevertheless, combining unamendability with imperfect enforcement is a recipe for failure. Unamendability in this case will become a poisonous pill that will lead to constitutional replacement.

Section 3.2 presents the findings of previous research about the determinants and effects of amendment constraints in general and unamendability in particular. It further discusses the benefits and weaknesses of agent-based models (henceforth ABM). Section 3.3 presents the model. The section describes the actors, their action space and the timeline of a game played over a finite number of legislative sessions. I run two simulation experiments. In one of them, the legislator's constraints are constant over time. In the second simulation experiment, the legislator's constraints change during the game. Section 3.4 discusses the results of the simulations. Last, section 3.5 concludes.

3.2 Literature review

3.2.1 What do we know about unamendability

Legal theory equates amending unamendable provisions with constitution-making, i.e. having to re-bargain all aspects of the constitution (Roznai 2017b; Albert 2019). If unamendability works as legal theorists assume, then the cost to amend an unamendable provision becomes larger than the cost to adjust to the provision. Unamendability has two faces. On the one hand, it could protect the constitution from transient majorities which could act against democracy and the rule of law. On the other hand, it allows a "dead" generation to decide about the future of the current generation.[2]

2. For a review of the literature introducing and discussing what has been coined the "dead hand problem" and the trade-off between democracy and constitutionalism see: (Roznai 2015, 2018)

Despite the abundance of normative theories to justify unamendability, there are currently no empirical findings about the determinants of unamendability. Albert (2013) presents a number of cases where unamendability was used to express the importance of different constitutional provision. Although the paper does not define what determines whether a society considers a provision important, it provides at least one explanation for the use of eternity clauses.

Concerning the determinants of amendment constraints in general, Hein (2019) shows that the introduction of amendment constraints is a function of a country's colonial and constitutional history. In the former case, amendment rules act as a signal that the legal certainty which the colonial power provided will also be provided by the new government. In the latter case, drafters attempt to signal that a constitutional change is not a constitutional break. In both cases the introduction of amendment constraints only loosely corresponds to a concern about future majorities, which could jeopardize the rule of law. In that sense, Hein (2019) and Albert (2013) agree.

Versteeg and Zackin (2016) demonstrate empirically that the drafters of constitutions in US states consistently make constitutional amendment easy, when they make the constitutional provisions more specific. In other words, they present a trade-off between how much power the legislator gets and to what extent she can increase her power. When a provision is formulated in a general way, the legislator is given leeway in how she can regulate a policy area. Moreover, judges are implicitly enabled to adjust the interpretation of the provision to adapt to changes in the society. However, the drafters use amendment constraints to avoid that the legislators and judges uncontrollably expand their power. Contrary, when the legislator and judges have small leeway in how they can implement the constitution, the drafters allow them to adjust the constitution to emerging needs.

Beyond the determinants of eternity clauses, one could think of at least two questions which need an answer to understand the effects of unamendability. Firstly, one can ask whether unamendability actually prevents constitutional change. Secondly, assuming that constitutions with unamendable provisions stay in force for a number of years and the unamendable provisions are not amended, one wonders whether unamendable provisions are violated more or less often than other provisions as time passes.

Concerning the first question, Hein (2020) demonstrates that courts in Europe have consistently for the last fifty years used amendment rules and particularly eternity clauses to stop constitutional amendments. This provides evidence that eternity clauses at least in democratic countries with relatively independent judiciaries are able to stop some unconstitutional formal changes. Nonetheless, Hein does not look at all amendments done in the sample's countries for the sampled years. Thus, it is not clear how many of the unconstitutional amendments were detected and stopped by the judiciary.

Unamendability is a barrier to formal amendments. However, constitutional change can take different forms and can be accomplished with different means (Contiades and Fotiadou 2013). A provision can also be changed through replacement. In this case, the courts are not able to stop the change of an unamendable provision (Landau and Dixon 2015). Theory suggests that constitutions which are hard to change might be in a particular danger of replacement (Gavison 2002; Zachary Elkins, Ginsburg and Melton 2009). For example, the constitutional provision defining the polity was unamendable in the Greek constitution of 1973. One year later, a new constitution was drafted which changed the polity.

As mentioned, unconstitutional amendments are equated to constitution-making. Thus, the same constitutional change can be classified as constitutional replacement or amendment depending on whether eternity clauses exist. For that reason, it is not possible to derive a meaningful understanding of the relationship between eternity clauses and constitutional replacement using empirical methods without developing a new definition for the latter.

Last, to measure the effectiveness of eternity clauses in stopping formal amendments one needs to make eternity clauses comparable. Constitutional provisions can be double-entrenched, i.e. protected with unamendable eternity clauses or single entrenched, i.e. protected with amendable eternity clauses. For example, Roznai (2017a) discusses how eternity clauses in Portugal were not protected with eternity clauses. Thus, the provisions that they protected were amended after another amendment removed the eternity clauses.

Further, a simple reading of unamendable provisions reveals that some of them are general, e.g. eternity clauses protect "human dignity" in Germany (Art. 1 of the current German constitution), whereas others are specific, e.g. "titles of nobility [...] are neither awarded nor recognized to Greek citizens" (Art. 4 of the current Greek constitution).[3] When the protected provisions are general, it is possible to informally change them e.g. by changing their interpretation, without having to formally amend or replace them (Voigt 1999; Albert 2015). Tsebelis (2017) provides evidence that shorter and more general constitutions tend to be less frequently formally amended.

There is a number of indicators which try to quantify how the design of amendment rules influences the difficulty of amendment. However, the indicators are not correlated with each other or predict the rate of amendment

3. There is currently no dataset quantifying the specificity of the provisions protected with eternity clauses.

(Ginsburg and Melton 2015). Ginsburg and Melton (2015) claim that formal amendment rules are not a good predictor for the frequency of amendments. They stipulate that the best predictor for the difficulty of amending the constitution is how often it was amended in the past, which they refer to as the *amendment culture*.

Concerning the second question, i.e. whether unamendable provisions are more likely to be violated as time passes compared to provisions which are amended over time, there is currently no research (Voigt 2020b, p. 45ff.). Zachary Elkins, Ginsburg and Melton (2016) present three possible paths for how the efficacy of constitutional rights changes over time.

Constitutional efficacy can be stable over time. The authors call this state stasis. Time "could lead people to impute wisdom to [the constitutional provision's] content, which might mean that the cost of violating the provision increases" (Zachary Elkins, Ginsburg and Melton 2016, p. 242). The authors call this trend maturation. Contrary, as time passes a constitutionally protected right can become outdated, i.e. no longer reflect the preferences in the population and thus be violated. The authors call this trend decay.

The authors show empirically that the protection of rights increases (matures) faster when the rights are entrenched in the constitution rather than when they are not. Further, constitutional rights are more likely to be in a state of stasis rather than decay compared to rights which are not constitutionally entrenched. In the analysis, the authors show that which of the three paths is taken and how fast is the rate of maturation or decay depends on a number of factors such as a country's regime and the degree of judicial independence.

Based on the findings and assuming that unamendability stops amendments, unamendability would be expected to be positively correlated with a state of stasis or with the emergence of maturation. However, one should be careful with making such a conclusion. Eternity clauses often protect provisions establishing procedural rules. The authors only look at constitutional rights and not violations of procedural rules. Further, the politicians who are entrusted with protecting constitutional rights are often the ones who decide whether they will constitutionally entrench them, i.e. the results suffer from endogeneity.

So far, the literature relied on a very simple model in combination with empirical analysis to quantify the effect of amendment rules. The laws were treated as a commodity and the constitution as a price catalog. If the price is right, politicians buy the commodity, i.e. legislate. Otherwise, they either try to change the price (amend or replace) or they steal the commodity, i.e. violate the constitution. Combining this framework with empirical analysis researchers tried to estimate how much the price of legislating changes with different constitutional rules. Models following this approach assume that what politicians do has no effect on the value of the law and the constitution or the difficulty of violating the latter.

Zachary Elkins, Ginsburg and Melton (2016) provide good reasons to reconsider this assumption. In order to understand how the ability of the government to violate an unamendable provision and the willingness of the citizens to accept a violation changes over time, we need a dynamic framework. This paper proposes such a framework using agent-based modelling. The next subsection discusses the innovations introduced with ABM and how agent-based modeling can be used to provide the missing dynamic models.

3.2.2 What can we learn from agent-based models

Previous work in constitutional economics analyzed constituted states as systems in equilibrium. The underlying assumption of this approach is that the conditions under which a state is created and develops determine which one of many equilibria the state will reach. Thus, by analyzing the environmental, cultural and institutional factors which effect a state's development one can deterministically predict a state's formal institutions.

Vice-versa, by observing the state's development, e.g. level of democracy, economic growth, control over the government etc., one can extrapolate the effect of formal institutions (Voigt 2020a). The problem with this approach is that it cannot explain institutional change and the volatility in the effect of institutions. If states deterministically converged to an equilibrium, neither should be present.

ABM provide a tool to analyze the impact of institutions stochastically. The goal of agent-based modeling is to simulate the conditions and interactions in a complex system in order to see how reality would look like, if individuals acted in a certain way (Chen and Venkatachalam 2017). Agent-based modeling is particularly useful to discover and analyze behavioral patterns which emerge when the agents act in groups. Moreover, it is useful to estimate the probability distribution of different events when there is path-dependency. This sub-section gives a short introduction in ABM and some of agent-based modeling's basic concepts.

Schelling (1969) was the first to use agent-based modeling to study a social phenomenon.[4] He was interested in the question whether residential

4. Agents developed out of Von Neumann's cellular automata (1966). In contrast to cellular automata which are autonomous but do not interact with other automata, agents interact with the environment and other agents. Pool and Kessler (1965) were the first to run a computer simulation with automata which process information according to biases. These automata did not respond to other automata.

segregation would come about between two groups with recognizably different characteristics (e.g. skin color or gender) when the members of the groups have mildly discriminatory preferences.

Schelling's agents had two states. They could be content or discontent. An agent's state was a function of two factors: his location relative to other agents and his type ('+' or 'o'). An agent was content when the number of neighbors of its own type did not drop under a certain threshold and discontent agents moved to an area where they could be content. In this stylized setting, Schelling shows that residential segregation happens even when agents are happy to live in neighborhoods where members of their group are the minority.

This very first ABM was calculated by Schelling using coins on a checkers' board. The agents could not evolve over time and their action space was limited. Since then, the development of modern computers automated "moving pawns on a checker's board" and allowed for the emergence of the field of Agent-based Computational Economics (ACE). Modern ABM allow for the modeling of heterogeneous agents with multiple evolving strategies solving complex problems (Chen 2016).

ABM usually assume that the agents are boundedly rational. A fully rational agent acts under full information and thus is able to calculate an optimal strategy based on the probability with which other rational agents will play different strategies. As a result, he will not re-evaluate his strategy based on experience. Contrary, a boundedly rational agent starts with a set of prior beliefs about the probability that other agents will play different strategies. Then, during play the agent collects information about the likelihood of different strategies and re-evaluates his priors. As a result, the strategies that a boundedly rational agent plays are a function of previous play.

One major contribution of ABM is that they provide a tool to estimate the probability that different equilibria will emerge when agents learn, explore the strategy space and evolve their strategies while interacting with other agents, i.e. while playing. Axelrod (1984, 1986) was one of the first to use agents playing evolving strategies in a computer simulation. He allowed agents to learn from each other by making less successful agents imitate more successful ones. These first implementations of agents that learn and evolve use unintelligent learning. Agents learn only what they can see, i.e. what the modeler includes as a potential strategy profile in the model.

Holland and Miller (1991) present how using different algorithms for exploration and learning introduces intelligent learning, i.e. agents which learn by trying out new strategies which are not hard-coded by the modeler. Intelligent agents can choose their strategy stochastically and improve the mix of strategies they play through experience and experimentation. The authors conclude that by making the agent's decision-processes closer to human cognitive processes ABM provide a tool to conduct two important tasks: experimentation of system dynamics and checking whether the results of mathematical modeling hold when the system becomes complex.

Another advantage of ABM is that they do not require oversimplifying a system. For instance, Kollman, Miller and Page (1992, 1997) simulate competition for votes between parties (in multiple jurisdictions) which discover the voters' ideological preferences over a number of polls and elections. Parties have platforms, which cover different policy areas. Moreover, they can be either office-seeking, i.e. pure vote-maximizers, or ideological, i.e. want to win election but with a platform that is as close to their ideal platform as possible. Voters have diverse political preferences and the strength of their preferences differs between policy areas.

To calculate the probability that platforms will converge by describing the system as a Markov chain one needs to be able to enumerate all possible states. The problem is that the policy space is n-dimensional and the combinations of policies that different parties can propose is countless. Moreover, the transition matrix is updated dynamically as parties gather information. Although it is practically impossible to calculate whether political platforms will converge in the described setting, through agent-based modeling in combination with simulations, it is possible to estimate the likelihood and degree of convergence.

Methodological individualism goes hand in hand with agent-based modeling. Hence, the lack of ABM in constitutional economics is striking. One explanation could be that computational results are not generalizable (Lehtinen and Kuorikoski 2007). Mathematical models describe social behavior as a natural phenomenon. Any researcher using the same set of assumptions should be able to come to the same set of conclusions. ABM are less predictable. Slight changes to the parameter values of a model can lead to quantitatively different results. As Laver and Sergenti (2012, p. 7) put it "computational results are good only for those parameter settings that have actually been investigated".

The other major weakness of ABM is that they are costly in terms of computational capacity and memory. Despite the advancement of information technology, simulating complex systems with many agents acting under different institutions is still difficult with conventional machines and requires knowledge of coding and algorithmic thinking (Bonabeau 2002). Both skills are to this day not taught in most faculties of economics or political sciences.

This section began by presenting the sparse literature about the determinants and effects of eternity clauses. Due to the lack of explicit theories about the effect of eternity clauses, the general theories developed to answer how amendment rules constrain amendments and whether amendment rules influence constitutional compliance were discussed. Concerning constitutional compliance three potential paths were presented. Maturation involves the occurrence of less constitutional violations over time. Decay involves the occurrence of more violations over time and stasis is the phenomenon where the number of violations stay constant over time.

Testing with observational data which path unamendable provisions tend to take is difficult. The provisions protected with eternity clauses vary in their specificity. Thus, the provisions they protect might be changed without being amended. Further, we do not know how the preferences of the people between different generations of the population differ and whether the stability of the constitution influences the preferences of the people in a generation.

The second part of the section discussed how ABM were used in the past and how they can be used to recreate complex systems for experimentation. The concept of an evolving agent was introduced and some examples of ABM were presented. The next section presents the model which is used to see how the probability that an unamendable provision will be violated changes over time.

3.3 The model

3.3.1 The society, the policy space and the constitution

The game is played between N agents and nature over S legislative sessions. For brevity, the agents as a whole will be referred to as the society. An agent's goal is to maximize his income from D financial activities. The society has a constitution which cannot be re-drafted or amended.

One part of the constitution defines that agents elect a legislator who introduces laws every legislative session and the legislator's constraints. All agents are given the right to run as candidates for the position of the legislator. All agents have one vote and agents have the right to abstain from voting. The legislator can only introduce universally applied laws. The constitution also establishes mechanisms that ensure that the legislator's laws which are introduced in accordance with the legislator's constraints are enforced to the letter. For the purposes of this paper, this part of the constitution is never violated.

Another part of the constitution -henceforth the constitutional core- defines the set of laws which can be introduced by the legislator to regulate each of the D activities. The legislator might try to violate this part of the constitution. Whether she will succeed depends on the constraints put on her. The legislator's constraints will be described in detail shortly.

Each activity constitutes a policy area and each policy area can be regulated with a single law. The law constrains how the agents are allowed to extract income in a policy area. The constitutional core defines the set of laws allowed to regulate an activity independently of the laws allowed to regulate other activities. This means that all agents and nature know whether a law is unconstitutional without having to consider the other laws of the society.

All possible ways to regulate an activity, i.e. all possible laws, are ranked and presented with values on an interval $[d_{min}, d_{max}]$, where:

$$|d_{min}| = |d_{max}| \text{ and } d_{min} < 0 < d_{max}$$

Laws are given higher absolute values when they introduce more constraints or incentivize financial activities with stronger incentives. Further, laws with opposing goals are on a different side of d_0. For example, an emission tax of 10% would be represented with a value greater than a tax of 2%. If emission taxes are given values greater than zero, subsidies for industrial production get values smaller than zero.

The midpoint of the interval ($d_0 = 0$) represents the free market, i.e. that agents are allowed to decide freely how to extract income from an activity without any constraints to their freedom of contracts, having to pay taxes or expect to receive subsidies. Henceforth, the legal order, i.e. the set of laws in force, after a legislative session s will be denoted as a point $i^s_* = (d^s_1, d^s_2, ..., d^s_D)$ in a D-dimensional policy space I, where d_d is the law in each policy area.

The constitutional core C is a body within the policy space I and will be written as:

$$C = ([d^l_1, d^r_1], [d^l_2, d^r_2], ..., [d^l_D, d^r_D])$$

The point d^l of an interval $[d^l, d^r]$ represent the most leftist law (smallest d) and the point d^r represents the most rightist law (largest d) that the constitution allows and both are points within the $[d_{min}, d_{max}]$ interval. Henceforth, coverage will denote the portion of all laws which are allowed by a constitutional provision:

$$\text{Coverage}_d = \frac{d^r_d - d^l_d}{d_{max} - d_{min}}$$

3.3.2 The timeline and the agents

Figure 3.2 shows the timeline of a legislative session. A test condition represents a node of the game where the agents' or nature's action space changes

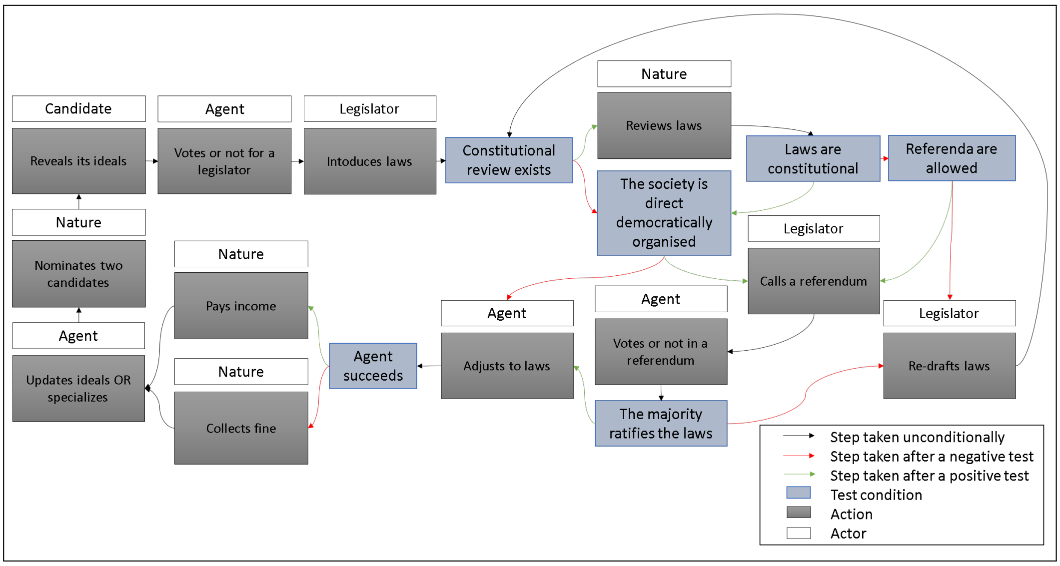


Figure 3.2: Flowchart: The legislative session

based on different conditions. There are two types of test conditions. In one type, the action space of the legislator changes depending on the constraints imposed on her by the constitution, e.g. whether there is constitutional review or all laws need to be ratified by a referendum. In another type, the action space changes because of agents' actions, e.g. whether the majority of the agents ratify an unconstitutional law or an agent fails to comply to laws. Both test conditions are represented with blue boxes in the figure.

The black arrows link two actions or actions and test conditions which always take place one after the other. The green (red) arrows link the test conditions with the actions that follow when a condition is (not) fulfilled. The white boxes represent the actors which can act at each node and the gray boxes the type of action the actors can take.

Briefly without going into detail, the game's timing is as follows. Nature nominates candidates for the position of the legislator. The candidates reveal their ideal laws and agents vote. The candidate who receives the majority of the votes becomes the legislator and introduces laws. Depending on the constitution, these laws are either submitted to constitutional review or not. Moreover, depending on whether the constitution allows for (mandatory) referenda the legislator can/has to call for a referendum to ratify the (reviewed) laws.

Laws which are reviewed as being unconstitutional and/or are rejected by a referendum are re-drafted or the legislator decides not to change the existing law. The legislative phase ends when all laws are either passed without review, are reviewed as being constitutional and/or are ratified by a referendum, the legislator accepts the laws in force or any combination of the above.

Then, each agent tries to comply with the laws. Nature collects fines from the agents who fail to comply and pays income to the agents who manage to comply with the laws. After an agent is paid or pays, he decides whether to change the way he earns income. He can either change the means with which he earns income or specialize more in using the means he already uses.

Each agent begins having three characteristics randomly selected by nature: his ideal states (one for each policy area), which map to his ideal laws (d^s_{nd}), his current states (p^s_{nd}) and his ability to adjust his current states to new laws (a^s_{nd}). In the paper, the term state encompasses the methods, equipment, partnerships or other legal arrangements that an agent uses to generate income in one policy area. All of the above characteristics are private information only known to the agent (and nature).

An ideal law reflects the legal constraints under which an agent can earn at full capacity through an activity ($d^s_{nd} \in [d_{min}, d_{max}]$). At $s = 1$, the agent's ideal laws mirror the agent's current states:

$$d^1_{nd} = p^1_{nd} \; \forall \; n \in N \; \& \; d \in D$$

Because the laws are perfectly enforced, agents can only earn income through an activity d if their current state corresponds to the law, i.e. $p^s_{nd} = d^s_d$. To make the narrative more intuitive, imagine that the agents produce D goods and the law defines the production and quality standards for the goods to be marketable. How much an agent can earn by producing in accordance with a law (P^s_{nd}) increases in the agent's ability to generate income given the law in force (E^s_{nd}) and decreases in the time the agents needs to comply to the law (t^s_{nd}).

If $d^s_d = p^s_{nd}$, then $t^s_{nd} = 0$. Otherwise, each agent changes its current state to comply to the law at a different rate. The time an agent needs to comply to the law in force is a function of the agent's current state, the law in force and the agent's adaptability:

$$\frac{\partial t^s_{nd}}{\partial |d^s_d - p^s_{nd}|} > 0 \text{ and } \frac{\partial t^s_{nd}}{\partial a^s_{nd}} < 0$$

Adaptability ($a^s_{nd} \in [0, d_{max} - d_{min}]$) represents how fast an agent can fulfill the legal standards or comply to legal requirements.

The n-th agent's ability to generate income from an activity d in a legislative session s (E^s_{nd}) depends on an agent's degree of specialization (S^s_{nd})

and the compatibility of an agent's ideal state and the law (C^s_{nd}):[5]

$$E^s_{nd} = f(S^s_{nd}, C^s_{nd}) \text{ where } \frac{\partial E^s_{nd}}{\partial S^s_{nd}} > 0 \text{ and } \frac{\partial E^s_{nd}}{\partial C^s_{nd}} > 0$$

Adaptability and specialization are negatively correlated in the model. In other words, a generalist is expected to work with general purpose techniques and equipment that can be easily repurposed or be cheaply replaced. This makes it easy to adapt. However, a generalist cannot sell at the price/produce at the quality that a specialist can, i.e. an agent with special purpose equipment which cannot be modified or repurposed.

An agent's ideal laws represent the laws that would allow the agent to reach its full earning potential. As the laws differ from the agent's ideal laws, the agent earns less than his full potential. Compatibility is a measure for the distance between an agent's ideal law and the society's law.

The agents whose current state is different than the laws after t act illegally. Illegality is punished with fines, which are proportional to the degree the agent's state differed from the law:

$$F^s_{nd} = d^s_d - p^{s+1}_{nd}$$

where p^{s+1}_{nd} is an agent's current state after t. Contrary to income, the fines do not differ between agents. However, the foregone income depends indirectly on S^s_{nd}. Thus, more specialized agents have more to lose from not compying to laws. An agent's wealth after a legislative session s is equal to:

$$w^s_n = \sum_{s=1}^{s} \sum_{d=1}^{D} P^s_{nd} - F^s_{nd}$$

Having presented the environment, the timing of the model in a general way and the agents, I will not discuss each part of the model in detail.

3.3.3 The action space and how agents and nature act

Nature nominates two candidates for the position of the legislators at the beginning of each legislative session. There is no limit to the times an agent can be nominated as a candidate. Nature chooses the two candidates among all agents stochastically. The probability that an agent is chosen as a candidate is a function increasing in an agent's wealth.

Becoming wealthier than other agents requires that the agent specializes. Highly specialized agents in the model have the most to lose from major legal

5. For simplicity, E^s_{nd}, S^s_{nd} and C^s_{nd} are normalized to vary between zero and one.

changes. Thus, it is reasonable that they are the ones most likely to try to set the laws. Since the agents' goal is to maximize their income, wealthier agents are also by definition more successful at playing the game. One could say that financial success is an indication of an agent's ability to make optimal decisions.

The assumption that citizens which are wealthier are more likely to run for office is not far from reality. Recent research has shown that income is correlated with intelligence in the eyes of voters (Dal Bó et al. 2017; Griffin, Newman and Buhr 2020). Thus, wealthy citizens have an advantage as political candidates. Carnes (2018) presents evidence of what he calls a cash-ceiling. Specifically, he shows that no senator or president in the USA ever worked a single day in his or her life in a low-income job. Giglioni (2020) presents examples of rich candidates who financed themselves into power also in Europe.

Making nature nominate the candidates stochasitcally reduces the model's complexity. Without any cost for being a candidate, all agents would run for the position of the legislator and will vote for themselves. Thus, a tie-breaker rule would have to decide the legislator every legislative session. Introducing a cost for running for office solves the problem of every agent running for office. However, to decide whether to pay the costs of running for office, the agents need to be given a measure for their electability, which law they will be able to introduce once in office and which laws will be introduced if they lose the elections. Because the agents do not know which laws other agents will try to introduce, they cannot estimate their electability or the expected benefits from running from office or getting elected.

Nature acts in all other cases deterministically. In particular, it always reviews laws when there is constitutional review. If a law does not pass the review, i.e. $d_d < d_d^l$ or $d_d > d_d^r$, nature always annuls the law. It collects a fine from all agents not complying to the laws, even the legislator, and pays an income to all agents complying with the law. Last, both fines collected and income paid are based on a formulas which is constant between agents and legislative sessions and known to all agents.

After being nominated, the candidates reveal their ideal laws truthfully. All agents believe the information the candidates give and vote for the candidate whose ideal laws are closer to theirs. If the candidates' ideal laws are equally apart from an agent's ideal laws, this agent does not bother to vote. The distance between the n-th agent and the c-th candidate is defined as:

$$dis_{nc}^s = \sum_{d=1}^{D} \frac{|d_{nd}^s - d_{cd}^s|}{a_{nd}^s}$$

The candidates' ideal laws reveal to the agents the first law that the candidate will try to introduce in each policy area. Whether the legislator will be able to introduce these laws lies not only in the hands of the legislator. It also depends on the legislator's constraints. Thus, it is a noisy signal for which laws the legislator will actually be able to introduce. An agent votes based on the distance, even if the constraints imposed by the constitution are expected to force both candidates after getting elected to introduce the same laws. In other words, it is assumed that an agent politically supports agents which have common interest as himself.

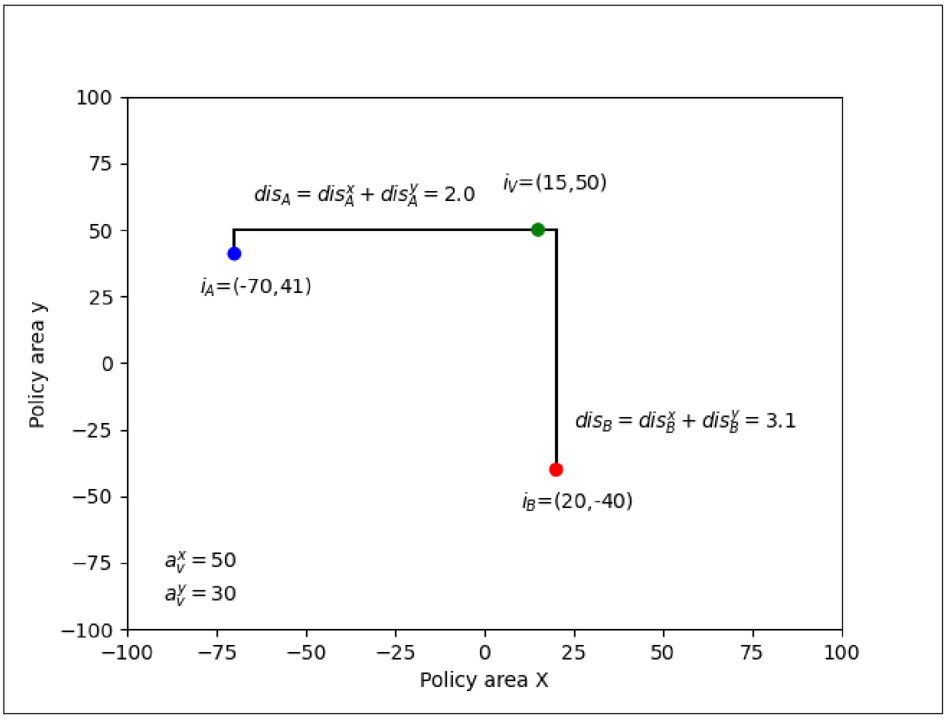


Figure 3.3: Example: Voting in elections

Figure 3.3 shows an example of how a voter V would decide between a candidate A and a candidate B. V earns income from two activities. Policy area X represents how one of the activities can be regulated -henceforth activity X - and policy area Y how the other activity can be regulated - henceforth activity Y. V's ideal legal order is $i_v = (15, 50)$. A's ideal legal order is in absolute terms further away compared to B's ideal legal order from V's ideal legal order. However, V is less adaptable in activity Y than in activity X ($a_v^x > a_v^y$). As a result, he weighs distance in policy area Y as more important than distance in policy area X. Because the total weighted distance to A is smaller than the total weighted distance to B ($dis_A < dis_B$),

V will vote in favor of A.

The candidate receiving the most votes wins. In case of a tie, nature appoints one candidate as the legislator randomly.[6] Winning the elections allows an agent to introduce laws. A legislator is not forced to change the law. If she does not, agents accept the status quo and do not try to force law-making.

The model allows for four types of constraints to the legislator. Table 3.1 divides the types of constraints to the legislator according to whether they entail constitutional review or the existence of the possibility to call a certain type of referendum. As a reminder, unconstitutional laws are reviewed and annulled if there is constitutional review. However, it is assumed that laws introduced with a referendum are not reviewed, i.e. the will of the People trumps the rule of law.

		Referenda are:		
		allowed	prohibited	obligatory
Constitutional review exists:	Yes	Weak constraints	Strict constraints	Direct democracy
	No	No constraints		

Table 3.1: Types of legislator's constraints

A constitution without constitutional review of the laws and no obligatory referenda does not constrain the legislator at all. The legislator can introduce any law she wants (no constraints). A constitution with constitutional review which prohibits referenda stops the legislator from introducing unconstitutional laws perfectly (strict constraints). Constitutions with strict constraints also include constitutions where referenda are possible but they are not allowed to introduce laws that violate the constitution.

In a direct democracy, the constitution requires that all laws are ratified by the majority of the agents. Thus, the legislator is only given power to make proposals, i.e. is an agenda-setter. Whether the legal order will contain unconstitutional laws depends on the results of the referenda following the proposals of the legislator.

Last, it is possible that the constitution allows for referenda although the constitution does not establish a direct democracy. The last type of constraints (weak constraints) provide a way for the legislator to re-introduce an annulled law. Specifically, the power to decide whether a constitutional violation will be stopped is taken from the hands of nature -in real life that

6. The probability of a candidate being appointed as the legislator after a tie is equal to $c_A/(c_A + c_B)$, where c_A is the agent's probability to be selected as a candidate and c_B his opponent's probability to be selected as a candidate.

would be the impartial and independent judges- and is put in the hands of the agents. Because the legislator has no reason to call for a referendum to introduce constitutional laws, she will only call referenda when constitutional review is expected to annul a law.[7]

When the legislator submits a law for ratification, agents vote in favor of the new law if they expect to earn more or pay a smaller fine with the new law than with the law in force. Agents which are indifferent between the law in force and the new law abstain from voting. The law is ratified, when the majority of the voters, i.e. those agents that turned up, is in favor of the new law.

A legislator begins by introducing her first best law, i.e. the law which could be introduced with the given constraints and maximizes her income from an activity. The location of the first best depends on the legislator's adaptability, current state, the location of her ideal law and her constraints. If a law is annulled and/or not ratified, the legislator introduces her next best law, i.e. the law that allows her to earn the next highest income. When a legislator ranks the law in force as the first/next best law, the legislator declares that she will not change the law.

When there are strict constraints, a legislator's first best is always a constitutional law. Contrary, the same legislator with the same ideals when there are weak or no constraints might rank an unconstitutional law as the first best. A highly specialized legislator ranks different laws primarily by focusing on how much time she would need to comply to the laws and then based on their compatibility with her ideal laws. Alternatively, a less specialized legislator ranks laws primarily according to their compatibility with her ideal laws and then according to the time she needs to comply to them.

The legislator does not care if the law will deviate to the left or to the right of her first best law. Thus, might alternate between introducing laws to the left and to the right of an annulled law in the policy space. After all laws are introduced (not annulled and/or ratified) and/or the legislator declares which laws will not be changed, the legislative phase is concluded.

Agents start to adjust their current states to the legal order simultaneously in all policy areas immediately after the legislative phase is concluded. Agents try to comply with the laws, even when it would be rational to sacrifice current income to maximize total wealth.[8] Whether agents manage to

7. The literature describes several reasons to use referenda as a mean to introduce laws. For example, politically controversial topics can be submitted to referenda. The legislator might want to try to create legitimacy etc. See: (Blume, Müller and Voigt 2009). Such considerations are not part of the agents' utility function in the model.

8. The reason for this design choice is that it makes the introduction of additional arbitrary parameter such as discount rates and risk preferences for the agents unnecessary.

comply to the laws in force depends on their adaptability and the distance between their current state and the law in force.

At the end of each legislative stage, agents decide whether they want to make long term changes to the way they earn income. These changes can take two forms. Agents either start using different methods, equipment, legal arrangements etc. or specialize more in how effective they use the methods, equipment etc. they already use. The former type of change results in the agents updating their ideal laws. For brevity, changing the methods, equipment, etc. used to earn income will be henceforth referred to as updating the ideal laws.

The degree of specialization and the agent's new ideal laws depend on the legal history of the society. The legal history of a policy area d at legislative session s will be defined as:

$$h_d^s = (d_d^1, d_d^2, ..., d_d^s)$$

The legal history describes the set of laws which were introduced to regulate the policy area from the first legislative session untill the session s. The term legal certainty will be used to describe how volatile the laws were until a session s and is defined as:

$$\sigma_d^s = \sqrt{\frac{1}{s-1}\sum_{i=1}^{s}(d_d^i - \overline{h_d^s})^2}$$

An agent decides whether to change how he earns his income after he sees how lucrative the session was. The probability that he updates his ideal laws or specializes depends on the degree to which the way he produces is profitable. The degree of profitability is set to be:

$$G_n^d = \frac{\sum_{s=s_0}^{s} P_{nd}^s - F_{nd}^s}{t(s-s_0)} < 1$$

where s_0 is the round at which an agent changed the way he earns income for the last time and $t(s-s_0)$ is the highest possible wealth that can be earned in $s-s_0$ sessions. The probability with which an agent updates his ideal laws is set to be equal to $1-G_n^d$.[9] When an agent does not adjust his ideal laws, he might specialize. The probability with which he specializes is set to be equal to $1-S_{nd}^s$. The intuition is that specialization becomes more difficult the more specialized an agent is. The probabilities of different changes to

9. If $G_n^d \leq 0$, then the probability is set to be equal to one.

how an agent earns income are:

$$Pr(\text{Updating}) = 1 - G_n^d$$
$$Pr(\text{Specialisation}) = G_n^d(1 - S_{nd}^s)$$
$$Pr(\text{No change}) = G_n^d S_{nd}^s$$

Updating involves the agents shifting their ideal laws towards a point in the policy space which they believe would in expectation allow them to earn more. Because the agents do not know the true distribution of future laws, they rely on their experiences to decide how to update. Nevertheless, how agents interpret their experiences differs. Hence, each agent stochastically chooses a point in the policy space towards which he shifts his ideal laws. The randomly drawn points follow a distribution with a mean $\overline{h_d^s}$, i.e. a mean equal to the average law, and a standard deviation equal to legal uncertainty σ_d^s.

By how much an agent shifts his ideal laws, i.e. how close he comes to the point in the policy space he has selected depends on the agent's adaptability and luck. Highly specialized agents find it hard to make major adjustments whereas lowly specialized agents find it easy. Both types of agents might not succeed in their plans to learn new methods or partially succeed.

Nature decides stochastically how close agents come to their randomly selected point in the policy space. The minimum distance an agent's ideal laws shift from their initial position is equal to one (unit of distance). The maximum distance is equal to an agent's adaptability ($1 \leq |d_{nd}^s - d_{nd}^{s+1}| \leq a_{nd}^s$).

The degree to which an agent specializes when he does not update his ideal laws is a function of legal uncertainty. If there is legal certainty, i.e. the laws are not volatile, the agents' income maximizing strategy is to specialize and capture more income rather than stay responsive to major legal changes. When the laws are volatile, agents are better-off to remain flexible in order to be able to extract income under any law.

Agents choose whether they will specialize high or low stochastically. A highly specialized agent can adjust to changes in an area equal to half the area's legal uncertainty and a lowly specialized agent can adjust to changes equal to the legal uncertainty. Both types of agents can always adjust by one unit of distance, even if there is zero uncertainty.

The strategy an agent uses to decide whether and how to change the way he earns income is a variation of a win stay-lose shift strategy with memory $s - s_0$ deep. When an ideal state is sufficiently profitable, the agent does not explore new ideal states. Contrary, he specializes. Ideal states which do not directly perform well are not immediately changed. However, long-term sub-optimal profitability leads to updating of the ideal state. Additionally, agents might return back to the way they produced in the past, if after exploration the profitability of the new ideals is low.

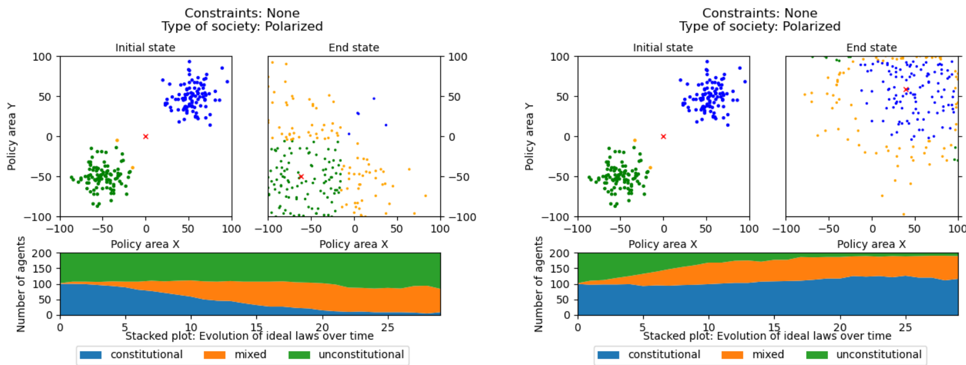
Because specialization increases profitability, it reduces the probability of updating indirectly. If an agent specializes too soon, he might remain in a sub-optimal ideal state for a number of sessions. Nevertheless, as the number of the sessions approaches infinity, the probability that the agent's ideal laws are equal to the average law and the probability that the agent reaches the maximum level of specialization approach one.

The location of the average law, e.g. whether it is within or outside of the constitutional core, as well as the number of rounds necessary for the majority of the agents to reach it is unknown and depends on the path that the state takes. Figure 3.4 gives two examples for how changing the way agents produce in order to adjust to their experiences can influence the probability of constitutional violations.

Each example consists of two scatterplots and a stacked area plot. The policy space is two-dimensional and the scatterplots show the positions of the agents' ideal legal orders in the policy space. The scatterplots on the top left corner of each example show the distribution of ideal laws at $s = 1$. The scatterplots on the top right corner of each example show the position of the agents' ideal laws after 30 legislative sessions. Below the scatterplots, the stacked area plots report the number of agents whose ideal law in one or both policy areas were constitutional during a session.

The green dots (and areas in the area plots) represent the (number of) agents whose ideal laws are both not allowed by the constitution. Yellow indicates agents whose ideal law in one policy area is allowed by the constitution but whose ideal law in the other policy area is not (mixed). Blue dots indicate agents whose ideal laws are both allowed by the constitution. Last, the red x marks show the free market in the left scatterplots ("law" in force when the game starts) and the average law when the game ends in the right scatterplots.

In both examples, the societies are identical at $s = 1$. The constitititutional core ($C = ([-15, 95], [-5, 93])$) and the legislator's constraints (none) are also identical. Despite starting as identical societies exposed to the same institutions, the societies take completely different paths. In example (a), the first legislators were elected from within the agents with unconstitutional ideal laws. A series of unconstitutional laws incentivized the agents with ideal laws which were allowed by the constitution to update and switch sides, i.e. move from the upper right quadrant of the policy space where ideal laws are constitutional to the bottom left quadrant. This response meant that it was easier for legislators with unconstitutional ideal laws to get elected and introduce more laws which were located on the lower left quadrant of the



(a) Example for decay (b) Example for maturation

Figure 3.4: Maturation and decay after 30 legislative sessions

policy space. As a result, more agents switched sides.

Contrary, in the society of example (b) the agents with constitutional ideal laws managed to directly get one of them elected as the legislator. This cascaded into agents with unconstitutional ideal laws changing sides. After a number of legislative sessions, only a handful of agents remained who had purely unconstitutional ideal laws. Example (b) can be seen as a case, where those that the constitution is meant to benefit are also in power.

When a country takes the path that the society took in example (a), the constitution will become constantly more unpopular. The agents will vote for candidates with unconstitutional ideal laws. After some point, it will become impossible to find a candidate who will introduce constitutional laws. When a country takes the path that the society in example (b) takes, the constitution becomes self-enforcing as time passes. Candidates with unconstitutional ideal laws do not get voted in office. As more time passes, the probability of finding a candidate with unconstitutional ideal laws approaches zero. The next section reports the results of simulations which were run to estimate the probability that a society takes the path of maturation or takes the path of decay.

3.4 Analysis

3.4.1 Keeping the legislator's constraints constant over time

In real life, it is impossible to expose the same society at the same point in time to different institutions in order to extrapolate the institutions' effects.

Luckily, this is not the case for simulation experiments. I run two sets of simulations. In the first set, the constraints are constant throughout the game. In the second set, the constraints change at some point in the game.

For the first set of simulations, I generate 60 societies consisting of 200 agents.[10] For each society, I generate 50 unique constitutional cores (3,000 constitutional cores in total). Each constitutional core has ten provisions. Keeping the constitutional core constant, I run one simulation pro constitutional core for each type of constraint (12.000 simulations). In total, I generate 3,600,000 observations (30 sessions $\times$ 10 provisions $\times$ 4 types of legislator's constraints $\times$ 50 constitutional cores $\times$ 60 societies)

I record whether and how a provision was violated every session. Further, I record whether the legislator would have introduced an unconstitutional law had she not been the subject to constraints and the legislator's ranking of the law that was introduced (e.g. first-best, second best, third best, etc.). Moreover, I count how many agents updated their ideal laws, specialized and the number of agents whose ideal laws were outside of the constitutional core in each policy area.

The majority of legislators served once (78.5%). 20 agents served nine times as the legislator, which is the maximum number of terms in office observed. In 83.9% of all sessions the legislator managed to introduce her most preferred law and in 9.2% her second best. The lowest ranked law that a legislator introduced was ranked as the 121rst-best law from 201 possible laws. In 54,3% of the sessions, the legislator did not attempt to introduce an unconstitutional law. The legislator's first-best law was unconstitutional but she was forced to introduce a constitutional law in 7.1% of all sessions. Subsequently, the likelihood that a legislator will manage to introduce an unconstitutional law is 5.4 times higher than the likelihood that she will be stopped.

In 38.7% of all sessions a constitutional violation took place. A quarter of these violations were passive violations, i.e. the legislator accepted an unconstitutional law which was in force. Another quarter were direct violations, i.e. the legislator introduced an unconstitutional law without a referendum or constitutional review. That means that approximately half of all unconstitutional laws were ratified by a referendum. The number of violations decreased over time only when the constitution was perfectly enforced, i.e. the legislator was subject to strict constraints. Otherwise, the number of violations in societies without strict constraints on average increased.

10. The societies can be divided in fifteen different society's types. Appendix A.2.1 presents in detail the simulations' parameters, the different types of societies and how the constitution is generated.

Constraints	Direct violations			Passive violations			Referenda			Sum of violations			Total
	Mean	SD	Max	Mean	SD	Max	Mean	SD	Max	Mean	SD	Max	
Strict				0.73	2.69	30				0.73	2.69	30 (1)	21,825
Weak				3.61	4.11	20	11.52	6.83	29	15.13	10.04	30 (616)	453,827
None	11.63	6.91	29	3.58	4.07	21				15.21	10.09	30 (696)	456,198
Dir. dem.				3.68	4.09	21	11.63	6.86	29	15.31	10.06	30 (659)	459,395
Total	348,790			347,906			694,549			1,391,245			

Table 3.2: Sum of violations pro provision

Table 3.2 groups the provisions according to the constraints' type and reports the mean, the standard deviation and the maximum number of times a provision was violated with any of the available means in the 30 legislative sessions. Next to the maximum number of violations pro provision in each category, I give the number of provisions which were violated the maximum amount of times. Only one constitutional provision was violated every single session with strict constraints. This number presents a great contrast to the over 600 provisions which were violated every single period with every other type of constraints.

To get a better understanding of which factors influence the probability of a violation, I use conditional logistic regression. Table 3.3 reports the results. The dependent variable takes the value of one when a provision d was violated in session s and zero otherwise.

In the first model, I stratify the observations by constitution. With the term constitution, I refer to the unique combination of a society, a constitutional core and a type of legislator's constraints. The second model stratifies the observations according to the constitutional core. These two models assume that violations within the same constitution follow a similar pattern. However, different patterns emerge with different constitutions. I stratify by society in the third model and I do not stratify in the fourth model.

For the explanatory variables, I use the number of the session, a continuous variable for the number of agents whose ideal laws are within the constitutional core ($\text{Inclusiveness}_{sd} \in [0, 200]$), the number of times a provision was violated in the past and the constaints' type. Due to the number of observations, statistical significance is a bad measure for whether the coefficients represent systematic correlations or randomness. For this reason, I will discuss statistical significance in correlation with economic significance, i.e. how big the effect is.

Violations decrease over time. However, the decrease is not economically significant. Inclusiveness has a robust and negative effect on violations. This effect is economically significant. If inclusiveness increases by five percentage points, the odds of a violation will be 1.16 times higher than without the increase.

DV: Dummy whether provision d was violated in session s				
	(1)	(2)	(3)	(4)
Session	-0.070^{***}	-0.072^{***}	-0.084^{***}	-0.084^{***}
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Inclusiveness	-0.010^{***}	-0.009^{***}	-0.020^{***}	-0.020^{***}
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Past violations	0.097^{***}	0.102^{***}	0.162^{***}	0.162^{***}
	(0.0002)	(0.0002)	(0.0004)	(0.0004)
Constant				1.866^{***}
				(0.009)
		Legislator's constraints		
Strict		-2.693^{***}	-0.713^{***}	-2.883^{***}
		(0.007)	(0.016)	(0.007)
Weak (Ref.)			2.171^{***}	
			(0.016)	
None		0.004^{**}	2.180^{***}	0.010^{***}
		(0.002)	(0.016)	(0.004)
Direct democracy		0.008^{***}	2.184^{***}	0.013^{***}
		(0.002)	(0.016)	(0.004)
AIC	14,292,365	17,492,452	3,076,798	3,076,889
Fixed-Effects	Constitution	Core	Society	-

Observations: 3,600,000. Coefficients represent logits. Robust SE are reported in the parentheses. *p<0.1; **p<0.05; ***p<0.01

Table 3.3: Likelihood of a violation without a switch

Past violations have an opposite effect. Each additional violation means that the odds of a future violation increase 1.17 times. The differences in the probability of a violation when the constitution does not constrain the legislator and a direct democracy, as well as a direct democracy and a society where the constitution imposes weak constraints, are economically insignificant. Only a constitution with strict constraints leads to a significantly lower probability of a violation.

To run a simulation to find out that less violations happen when there are less ways to violate is meaningless. Such a conclusion could be reached also using observational data. What currently observational data cannot fully capture is how often the legislator tries to violate the constitution or how

DV: Dummy whether the legislator tried to introduce an unconstitutional law for policy area d in session s		
	Main term	Interaction with session
Session	-0.076^{***}(0.0003)	
Strict constraints	-1.532^{***} (0.007)	$+0.111^{***}$ (0.0004)
Direct democracy	$+0.037^{***}$ (0.007)	-0.002^{***} (0.0004)
No constraints	-0.004 (0.007)	$+0.0003$ (0.0004)
Inclusiveness	-0.023^{***} (0.0001)	
Past violations	$+0.149^{***}$ (0.0003)	
Constant	$+2.076^{***}$ (0.009)	

Observations: 3,600,000. Coefficients represent logits. Robust SE are reported in the parentheses. Model's LogLikelihood =−1,963,517. *p<0.1; **p<0.05; ***p<0.01

Table 3.4: Likelihood of an attempt to violate without a switch

often the legislator would try to violate the constitution, if it knew it could not be stopped.

In the simulations, I record every attempt of the legislator to introduce a law. Further, I program the agents to attempt to introduce the most lucrative law for them even if they know that this law cannot be introduced. For that reason, I can also see how the probability that a legislator would want to violate a provision changes with different types of constitutions. Table 3.4 reports the results of an ordinary logistic regression where the dependent variable takes the value one when the legislator tried to introduce an unconstitutional law and zero otherwise.

Regression analysis shows that with a perfectly enforced constitution the probability of a violation does not only decrease because there are fewer means to violate the constitution but also because the legislator who gets elected is less likely to want to introduce unconstitutional laws. In other words, less agents with ideals outside of the constitutional core are elected to power.

A direct democracy is as likely to experience violations as a society which has a constitution that imposes no constraints on the legislator. Nevertheless, the legislator in a direct democracy is significantly more likely to try to violate the constitution. This shows that constitutional review is indeed a mechanism that constrains the legislator, even when the legislator is given ways to circumvent constitutional review. Further, direct democracy can force legislators which do not want to violate the constitution to violate it.

As time passes the probability that a legislator will come to power decreases. However, the change is not economically significant. As a control for heterogeneity, I also introduce a set of interaction terms between the type of the constraints and the variable session. The statistically and economically significant positive coefficient of the interaction term between strict constraints and the session reflects the fact that the probability cannot decrease as much over time because it is already very low.

So far, the constitution's type was not changed during the game. When the legislator introduced the first constitutional law, strict constraints ensured that the agents' ideal laws would start converging to a point within the constitutional core. When the first agents adjusted to unconstitutional laws, the lack of a mechanism to stop constitutional violations ensured that the future legislators would violate constitutions without constraints. It is thus yet unclear what would happen, if a society starts with a constitution which is perfectly enforced and suddenly switches to a setting where constitutional enforcement is put in the hands of the people or worse the legislator. The next subsection answers this question.

3.4.2 Changing the legislator's constraints during the game

In the previous subsection, the means the legislators possessed to violate the constitution were kept constant. However, this must not be the case in real life. Constitutional amendments can introduce new mechanisms to legislate. The judiciary might face a crisis which would not allow it to safeguard the constitution.

To account for such changes, I implement four possible switches. The first two switches simulate improvement and deterioration of constitutional enforcement. A constitution either starts with strict constraints, i.e. as the least likely constitution to be violated, and switches to a constitution without constraints or starts as a constitution without constraints, i.e. the most likely constitution to be violated, and switches to a constitution with strict constraints.

The second set of switches simulates democratization and democratic backslides. A society can starts as a direct democracy and switch to a representative democracy with binding non-obligatory referenda. Alternatively, a society starts as a representative democracy with referenda and at some point becomes a direct democracy.

I generate fifteen societies and a unique core for each society. I further program forty types of switches. Each constitution (= society + core) is exposed multiple times to each type of switch. The type of the switch is defined by which of the possible switches took place and in which round. In total, I generate 1,980,000 observations.

	DV: Dummy whether provision d was violated in session s			
	Weak constraints		Strict constraints	
Switched after	started	ended	started	ended
Never	(Reference group)		-0.430^{***} (0.002)	
1rst session	-0.004 (0.013)	$+0.002$ (0.003)	$+0.042^{*}$ (0.015)	$+0.086^{***}$ (0.003)
2nd session	$+0.025^{*}$ (0.009)	$-0.006^{(*)}$ (0.003)	$+0.022^{(*)}$ (0.010)	$+0.091^{***}$ (0.003)
3rd session	$+0.021^{*}$ (0.007)	$+0.008^{**}$ (0.003)	$+0.032^{***}$ (0.009)	$+0.063^{***}$ (0.003)
4th session	$+0.008$ (0.007)	$+0.004$ (0.003)	$+0.051^{***}$ (0.007)	$+0.059^{***}$ (0.003)
5th session	$+0.001$ (0.006)	-0.014^{***} (0.003)	$+0.062^{***}$ (0.007)	$+0.057^{***}$ (0.003)
6th session	$+0.012^{(*)}$ (0.005)	$-0.007^{(*)}$ (0.003)	$+0.022^{***}$ (0.006)	$+0.048^{***}$ (0.003)
7th session	$+0.040^{***}$ (0.005)	$+0.003$ (0.003)	$+0.008$ (0.006)	$+0.028^{***}$ (0.003)
8th session	$+0.024^{***}$ (0.005)	-0.009^{**} (0.003)	$+0.019^{***}$ (0.005)	$+0.031^{***}$ (0.003)
9th session	$+0.019^{***}$ (0.004)	-0.009^{**} (0.003)	$+0.015^{**}$ (0.005)	$+0.036^{***}$ (0.003)
	Direct democracy		No constraints	
Switched after	started	ended	started	ended
Never	-0.006^{**} (0.002)		0.0001 (0.002)	
1rst session	$+0.033$ (0.019)	-0.005 (0.004)	$+0.019$ (0.019)	-0.100^{***} (0.004)
2nd session	-0.011 (0.013)	$+0.015^{***}$ (0.004)	-0.011 (0.013)	-0.157^{***} (0.004)
3rd session	$+0.035^{***}$ (0.011)	-0.003 (0.004)	$+0.009$ (0.011)	-0.143^{***} (0.004)
4th session	$+0.031^{***}$ (0.009)	-0.004 (0.004)	$+0.018^{(*)}$ (0.009)	-0.150^{***} (0.004)
5th session	$+0.012$ (0.008)	$+0.009^{(*)}$ (0.004)	$+0.016$ (0.008)	-0.142^{***} (0.004)
6th session	$+0.026^{***}$ (0.008)	-0.001 (0.004)	$+0.001$ (0.008)	-0.114^{***} (0.004)
7th session	$+0.011$ (0.007)	$+0.0002$ (0.004)	$-0.016^{(*)}$ (0.007)	-0.124^{***} (0.004)
8th session	-0.002 (0.007)	$+0.008^{(*)}$ (0.004)	-0.009 (0.007)	-0.125^{***} (0.004)
9th session	$+0.018^{**}$ (0.006)	$+0.004$ (0.004)	$+0.002$ (0.006)	-0.107^{***} (0.004)
Inclusiveness	-1.466^{***} (0.002)			

Observations: 1,980,000. Adj. $R^2 = 0.526$. The variable *Inclusiveness* is rescaled to be between zero and one. The coefficients are based on a linear model with society's fixed-effects. Clustered on the society's level robust SE are reported in parentheses. *** $p < 0.001$, ** $p < 0.005$**, * $p < 0.01$, (*) $p < 0.05$.

Table 3.5: Likelihood of a violation with a switch

Table 3.5 reports the results of a linear regression with fixed effects on the constitutional core's level. I report the results from a linear model because of the ease of interpreting the coefficients of the interaction terms.[11] I also run a conditional logistic regression. The results of the linear and the non-linear model are not qualitatively different.

I differentiate between the probability of a provision being violated before a switch and after a switch given the type of the switch. The columns *started* report the differences in the probability of a violation before a switch given the type of constaints applicable. The columns *ended* report the differences after the constraints are changed. The table's rows *Never* show how much the probability of a violation with constraints of a given type differed to the

11. For an extensive discussion of the benefits of linear models compared to non-linear models see: subsection 2.4.2.

probability of a violation with weak constraints. Last, the row *Inclusiveness* reports how the probability of a violation changes when the portion of agents with ideals within the constitutional core changes.

As also shown in the previous subsection, constitutions which allow for referenda are more likely to be violated compared to constitutions that don't. The probability of a violation is marginally smaller in societies which give very limited power to the legislator, i.e. direct democracies, compared to societies that impose no constraints on the legislator. However, the difference is negligible. This regression also confirms the result of the previous section that whether the legislator must use the majority to introduce unconstitutional laws or can do it on her own does not matter. Any means to violate the constitution will lead to the constitution being violated.

The model shows further that as inclusiveness increases, the probability of a violation decreases. Indicatively, a 10 percentage points increase in inclusiveness reduces the probability of a violation by 14 percentage points. This result is independent of the constraints' type.

The coefficients in the *started* columns show how much a constitution that did not switch differed to a constitution with the same constraints before the switch. Any differences between constitutions with the same constraints before the switch and without a switch are merely the result of randomness. Hence, it is not surprising that most of the coefficients in the *started* columns are either statistically insignificant or economically insignificant.

The coefficients in the columns *ended* show how much more or less likely it was that the constitution would get violated if it switched at any point from one type of constraints to another. Each row of the columns *ended* shows how the effect of the switch differs depending on the session at which the switch was made.

Switching from a direct democracy to weak constraints does not have a robust effect on the likelihood of a violation. If a switch has an effect on violations, the effect should be either constant over time or be stronger the earlier the switch is made. The inconsistent pattern in the signs of the coefficients is an additional indication that this type of switch does not significantly change constitutional compliance. Switching from weak constraints to a direct democracy demonstrates the same pattern.

To sum up, with the simulated sample and the model used in the paper I cannot find an effect of increasing popular participation on constitutional compliance. This result is in line with the existing literature on the effects of direct democracy. For example, Blume, Müller and Voigt (2009) find that the use of direct democracy has an effect on a state's outcome, however it is not a panacea. Whether direct democracy improves a state's performance in terms of helping reduce total spending, budget deficit, productivity etc, or worsens it depends on a variety of factors.

Contrary, both improvement and deterioration of constitutional enforcement effects the probability of violations after the switch. If a constitution is perfectly enforced during the first session, the probability that its provisions will not be violated by an unconstrained legislator decreases by 10 percentage points compared to when the same society is not exposed for a single legislative session to a perfectly enforced constitution. Enforcing the constitution perfect for an additional session reduces the probability of a violation by another 6 percentage points. Respectively, by allowing the legislator to introduce unconstitutional laws only during the first legislative session, the probability that the constitution will be violated at any session after the switch becomes approximately nine percentage points higher. However, the effect does not get stronger the longer a society is exposed to a constitution without constraints.

The size of the effect and the number of sessions an effect needs to kick-in, as presented in the literature review, is a function of the model's parameters, i.e. how fast agents adapt to the legal order. If agents adapt more frequently and faster the effect will become larger. If they are less adaptive, the effect will become smaller.

Interpreting the result qualitatively provides an alarming insight. Schwartz (2019) using an agent-based model shows that judges are likely to facilitate the legislator, i.e. not review unconstitutional laws until the power of the judiciary is established. Combining the results in this chapter with Schwartz's result, one can reach the conclusion that eternity clauses will jeopardize constitutional survival.

For the reasons presented already, looking only at the violations does half the necessary work. Hence, table 3.6 runs the model with a different dependent variable this time. Specifically, I replace the dummy for whether a provision was violated with a dummy for whether the legislator made an attempt to violate a constitutional provision. The structure of the table is the same as the structure of table 3.5.

The effect of strict constraints is robust even when the constraints change type. Contrariwise, the effect of direct democracy disappears. A few sessions with no constraints have a devastating effect on the probability that the legislator will attempt to introduce at least one unconstitutional law. Five sessions of exposure to no constraints are enough to increase the probability that the legislator would want to violate the constitution by approximately 15 percentage points. Interestingly, exposing a society for the first legislative session to a perfectly enforced constitution has the statistically exact opposite effect.

These findings are not counter-intuitive. Acemoglu, Johnson and Robin-

	DV: Dummy whether the legislator tried to introduce an unconstitutional law in policy area d in session s			
	Weak constraints		Strict constraints	
Switched after	started	ended	started	ended
Never	(Reference group)		-0.141^{***} (0.002)	
1rst session	$+0.001$ (0.013)	$+0.001$ (0.003)	$+0.155^{***}$ (0.019)	$+0.138^{***}$ (0.004)
2nd session	$+0.031^{***}$ (0.009)	$-0.006^{(*)}$ (0.003)	-0.056^{***} (0.012)	$+0.143^{***}$ (0.004)
3rd session	$+0.027^{***}$ (0.007)	$+0.006^{(*)}$ (0.003)	-0.108^{***} (0.010)	$+0.129^{***}$ (0.004)
4th session	$+0.015^{(*)}$ (0.006)	$+0.00000002$ (0.003)	-0.113^{***} (0.008)	$+0.134^{***}$ (0.004)
5th session	$+0.008$ (0.006)	-0.015^{***} (0.003)	-0.111^{***} (0.008)	$+0.149^{***}$ (0.004)
6th session	$+0.019^{***}$ (0.005)	-0.011^{***} (0.003)	-0.147^{***} (0.007)	$+0.143^{***}$ (0.004)
7th session	$+0.048^{***}$ (0.005)	$+0.0001$ (0.003)	-0.154^{***} (0.006)	$+0.131^{***}$ (0.004)
8th session	$+0.032^{***}$ (0.005)	-0.013^{***} (0.003)	-0.140^{***} (0.006)	$+0.140^{***}$ (0.004)
9th session	$+0.027^{***}$ (0.004)	-0.013^{***} (0.003)	-0.128^{***} (0.006)	$+0.138^{***}$ (0.004)
	Direct democracy		No constraints	
Switched after	started	ended	started	ended
Never	$-0.005^{(*)}$ (0.002)		-0.002 (0.002)	
1rst session	$+0.031$ (0.019)	-0.004 (0.004)	$+0.02$ (0.019)	-0.146^{***} (0.004)
2nd session	-0.012 (0.013)	$+0.013^{***}$ (0.004)	-0.010 (0.013)	-0.132^{***} (0.004)
3rd session	$+0.034^{**}$ (0.011)	-0.004 (0.004)	$+0.010$ (0.011)	-0.138^{***} (0.004)
4th session	$+0.029^{**}$ (0.009)	-0.002 (0.004)	$+0.020^{(*)}$ (0.009)	-0.128^{***} (0.004)
5th session	$+0.010$ (0.008)	$+0.008^{(*)}$ (0.004)	$+0.017^{(*)}$ (0.008)	-0.103^{***} (0.004)
6th session	$+0.024^{**}$ (0.007)	$+0.0004$ (0.004)	$+0.002$ (0.008)	-0.110^{***} (0.004)
7th session	$+0.010$ (0.007)	-0.001 (0.004)	$-0.015^{(*)}$ (0.007)	-0.113^{***} (0.004)
8th session	-0.003 (0.007)	$+0.006$ (0.004)	-0.008 (0.007)	-0.093^{***} (0.004)
9th session	$+0.016^{(*)}$ (0.006)	$+0.004$ (0.004)	$+0.001$ (0.006)	-0.085^{***} (0.004)
Inclusiveness	-1.76*** (0.002)			

Observations: 1,980,000. Adj. $R^2 = 0.563$. The variable *Inclusiveness* is rescaled to be between zero and one. The coefficients are based on a linear model with society's fixed-effects. Clustered on the society's level robust SE are reported in parentheses. *** p< 0.001, ** p< 0.005**, * p< 0.01, (*) p< 0.05.

Table 3.6: Likelihood of an attempt to violate with a switch

son (2002) show empirically that a society's initial institutions can have a significant effect on how a society will develop. If eternity clauses are at least upheld by the politicians that introduced them -who often end up in office after they draft the constitution- they can be a powerful tool to prevent the emergence of unconstitutional preferences. The effect they can have on whether violations are attempted is also very significant. However, if eternity clauses are used to entrench aspirational elements of the constitution which cannot be directly achieved, they are very likely to fail to facilitate maturation. The citizens learn to ignore the constitutional provisions which are not upheld. Then, when the constitution becomes enforceable they do not care to enforce these provisions.

3.5 Conclusion

In this chapter, I take on the question what determines whether an unamendable constitutional provision becomes self-enforcing. To answer the question, I design an agent-based model. The agents elect legislators who are submitted to different constraints.

In the model, the agents update their ideals based on their society's legal history. Updating influences whether legislators with ideals which are represented in the constitutional core come to power or legislators with ideals outside of the constitutional core win the elections. I expose the same society to different constitutions. I simulate societies where the constitution provides different means for its violation. I also simulate societies where the constraints imposed by the constitution stay constant over time and societies where they change at some point in time.

Simulations show that a constitution which is perfectly enforced gradually becomes self-enforcing, i.e. the legislators stop trying to introduce unconstitutional laws even when they can. In contrast, any means of violating the constitution, independent of whether it is the People, i.e. the agents, or the legislator who decides whether to introduce unconstitutional laws not only increases the probability of violations but also the probability that legislators with unconstitutional ideals will come to power.

Even in countries with a developed legal system, "bad apples" can become a credible threat to the rule of law as recent experiences with populistic parties/leaders in Europe and the USA show. Entrenchment has long been considered as a mean to protect democracy by constraining it. Nevertheless, constitutional entrenchment bears the risk that the constitution will reflect obsolete ideals. I show using simulations that if individuals adapt their ideals based on existing laws this risk is minimal provided that the constitution starts on a good foot, i.e. it is perfectly enforced.

This chapter discusses one variation of the model. One can extend the model to incorporate more complexity and constitutional detail. For example, it would be interesting to see what happens if the constitution is drafted by the agents. Moreover, the model shows that the identity of the first legislator is very important. In the current form of the model, moderate agents are more likely to get elected compared to leftist or rightist agents.

One could expand the model to see how changing the selection process for the first legislator changes the point at which agents converge in the policy space. For example, agents could be allowed to gather information in order to estimate their electability and then decide whether to run for office. Another extension is to implement logrolling in the model, i.e. the agents trading votes to bring the law closer to their ideal, or introduce candidates which do not reveal their true ideals.

Another interesting variation is to allow the legislator to violate the part of the constitution which sets the procedural rules. When an agent is not in power, he wants the legislator to be constrained. The legislator herself only wants to constrain herself as little as necessary. It would be interesting to see how this asymmetry influences the outcomes.

The model is a tool for abduction, i.e. the creation of theory based on observations. Its insights so far were based on simulated data and arbitrary stylized parameters. One could feed the model with survey and experimental data about agent's ideals and preferences and updating and set the parameters according to historical data. This would allow the model to generate country specific benchmarks and predictions which could be used in empirical research. This endeavor is left for future research.

References

Acemoglu, Daron, Simon Johnson and James A Robinson. 2002. Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution. *Quarterly Journal of Economics* 117 (4): 1231–1294.

Albert, Richard. 2013. The Expressive Function of Constitutional Amendment Rules. *McGill Law Journal* 59 (2): 225–281.

———. 2015. How Unwritten Constitutional Norms Change Written Constitutions. *Dublin University Law Journal* 38 (2): 387–417.

———. 2019. *Constitutional Amendments: Making, Breaking, and Changing Constitutions.* 344. New York, NY: Oxford University Press.

Axelrod, Robert. 1984. *The Evolution of Cooperation.* 241. New York, NY: Basic Books.

———. 1986. An evolutionary approach to norms. *American Political Science Review* 80 (4): 1095–1111.

Bjørnskov, Christian, and Stefan Voigt. 2018. The architecture of emergency constitutions. *International Journal of Constitutional Law* 16 (1): 101–127.

Blume, Lorenz, Jens Müller and Stefan Voigt. 2009. The economic effects of direct democracy-a first global assessment. *Public Choice* 140 (3-4): 431–461.

Bonabeau, Eric. 2002. Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences of the United States of America* 99 (SUPPL. 3): 7280–7287.

Carnes, Nicholas. 2018. *The Cash Ceiling: Why only the rich run for office– and What we can do about it.* 322. Princeton, NJ: Princeton University Press.

Chen, Shu-Heng. 2016. *Agent-based computational economics: How the idea originated and where it is going.* 528. London: Routledge.

Chen, Shu-heng, and Ragupathy Venkatachalam. 2017. Agent-based modelling as a foundation for big data. *Journal of Economic Methodology* 24 (4): 362–383.

Contiades, Xenophon, and Alkemene Fotiadou. 2013. Models of constitutional change. In *Engineering Constitutional Change: A Comparative Perspective on Europe, Canada and the USA,* edited by Xenophon Contiades, 417–469. London: Routledge.

Dal Bó, Ernesto, Frederico Finan, Olle Folke, Torsten Persson and Johanna Rickne. 2017. Who Becomes a Politician? *Quarterly Journal of Economics* 132 (4): 1877–1914.

Elkins, Zachary, Tom Ginsburg and James Melton. 2009. *The Endurance of National Constitutions.* 260. Cambridge, MA: Cambridge University Press.

———. 2016. Time and constitutional efficacy. In *Assessing constitutional performance,* edited by Tom Ginsburg and Aziz Z Huq, 233–267. Cambridge, MA: Cambridge University Press.

Elkins, Zackary, Tom Ginsburg and James Melton. 2020. *Chronology of Constitutional Events, Version 1.3.* Accessed 28 April. `https://comparativeconstitutionsproject.org`.

Gavison, Ruth. 2002. What Belongs in a Constitution? *Constitutional Political Economy* 13 (1): 89–105.

Giglioni, Mateo Fabio Nels. 2020. Plutocratic leadership in the electoral arena: three Mitteleuropean cases of personal wealth in politics. *Comparative European Politics* 18:309–329.

Gilbert, Michael D, Mauricio Guim and Michael Weisbuch. 2019. Constitutional Locks. *Virginia Public Law and Legal Theory Research Paper* 2019-50.

Ginsburg, Tom, and James Melton. 2015. Does the Constitutional Amendment Rule Matter at All? Amendment Cultures and the Challenges of Measuring Amendment Difficulty. *International Journal of Constitutional Law* 13 (3): 686–713.

Griffin, John D, Brian Newman and Patrick Buhr. 2020. Class War in the Voting Booth: Bias Against High-Income Congressional Candidates. *Legislative Studies Quarterly* 45 (1): 131–145.

Hein, Michael. 2018. *Constitutional Entrenchment Clauses Dataset.* Göttingen. Accessed 18 May 2020. `http://data.michaelhein.de`.

———. 2019. Impeding constitutional amendments: why are entrenchment clauses codified in contemporary constitutions? *Acta Politica* 54 (2): 196–224.

———. 2020. Do constitutional entrenchment clauses matter? Constitutional review of constitutional amendments in Europe. *International Journal Of Constitutional Law* 18 (1): 78–110.

Holland, John H, and John H Miller. 1991. Artificial Adaptive Agents in Economic Theory. *American Economic Review* 81 (2): 365–370.

Kollman, Ken, John H Miller and Scott E Page. 1992. Adaptive Parties in Spatial Elections. *American Political Science Review* 86 (4): 929–937.

———. 1997. Political Institutions and Sorting in a Tiebout Model. *American Economic Review* 87 (5): 977–992.

Landau, David, and Rosalind Dixon. 2015. Constraining Constitutional Change. *Wake Forest Law Review* 50 (4): 859–890.

Laver, Michael, and Ernest Sergenti. 2012. *Party Competition: An agent-based model.* 293. Princeton, NJ: Princeton University Press.

Lehtinen, Aki, and Jaakko Kuorikoski. 2007. Computing the perfect model: Why do economists shun simulation? *Philosophy of Science* 74 (3): 304–329.

Pilpilidis, Konstantinos. 2018. A Constitution for Eternity: An Economic Theory of Explicit Unamendability. In *An unconstitutional constitution? unamendability in constitutional democracies,* edited by Richard Albert and Oder Bertil Emra, 63–88. Springer.

Pool, Ithiel de Sola, and Allan Kessler. 1965. The Kaiser, the Tsar, and the computer: Information processing in a crisis. *American Behavioral Scientist* 8 (9): 31–38.

Roznai, Yaniv. 2015. Unamendability and the Genetic Code of the Constitution. *European Review of Public Law* 27 (2): 775–825.

———. 2017a. Negotiating the eternal: The paradox of entrenching secularism in constitutions. *Michigan State Law Review* 253:254–331.

———. 2017b. *Unconstitutional Constitutional Amendments:The Limits of Amendment Powers.* 368. Oxford: Oxford University Press.

———. 2018. Necrocracy or Democracy? Assessing Objections to Formal Unamendability. In *An unconstitutional constitution? unamendability in constitutional democracies.* Edited by Richard Albert and Oder Bertil Emrah, 29–62. Springer.

Schelling, Thomas C. 1969. Models of Segregation. *American Economic Review* 59 (2): 488–493.

Schwartz, Alex. 2019. An Agent-Based Model of Judicial Power. *Journal of Law* 9 (1): 21–53.

Tsebelis, George. 2017. The time inconsistency of long constitutions: Evidence from the world. *European Journal of Political Research* 56 (4): 820–845.

Versteeg, Mila, and Emily Zackin. 2016. Constitutions Un-Entrenched: Toward an Alternative Theory of Constitutional Design. *American Political Science Review* 110 (4): 657–674.

Voigt, Stefan. 1999. Implicit Consitutional Change-Changing the Meaning of the Constitution Without Changing the Text of the Document. *European Journal of Law and Economics* 7 (3): 197–224.

———. 2020a. *Constitutional economics: a primer.* 150. Cambridge: Cambridge University Press.

———. 2020b. Mind the Gap – Analyzing the Divergence between Constitutional Text and Constitutional Reality. *ILE Working Paper Series* 32.

Von Neumann, John. 1966. *Theory of self-reproducing automata.* 416. Chicago, IL: University of Illinois Press.

Chapter 4

Even when she wins she loses: Distinguishing gender taste and statistical discrimination

4.1 Introduction

Women CEOs are more likely to get fired than men CEOs and the likelihood of keeping their position does not improve when the company performs good (Gupta et al. 2020).[1] The likelihood of a woman to get tenure in academia is negatively correlated with the number of papers she co-authored with men. Men do not experience a similar malus for publishing with women or men (Sarsons 2017). During an electoral campaign, women candidates stand to lose more votes than men candidates when information is presented that casts a doubt on their competence (Ditonto 2017).

One explanation for the discrepancy in how performance is evaluated and rewarded between men and women is discrimination. Statistical discrimination originates from a belief that women will perform worse in a task than men. Taste discrimination results from a preference to reward or interact with men. Disentangling the two poses several challenges.[2] The decisions to favor a man over a woman due to gender-based beliefs, taste or due to other factors are indistinguishable for an outside observer. Further, it is

1. The chapter uses a binary definition of gender. All the participants in the experiments also defined their gender in a binary way although other options were available.

2. Bertrand et al. (2005) introduce implicit discrimination as a third type of discrimination. Implicit discrimination measures how fast an individual associates an activity with a group (men and work; women and family) and does not require discriminatory actions. I do not consider implicit discrimination because the chapter focuses on behavior and not states of mind.

possible that different factors combined influence how individuals evaluate performance.

I design and run an experiment to disentangle the two types of discrimination. During a pre-experiment, the competitors earn points through a real-effort search-and-find task. Then, they are paired and the competitor with the most points per pair wins a prize. During the experiment, the participants are asked to predict which competitors' gender group got on average more points and which competitor won in each pair. Before they submit their predictions, participants are informed about the competitors' genders, ages and academic degrees. After they submit their predictions, they find out the competitors' number of correct answers and are asked to give a bonus to one competitor per pair.

Most participants predicted that men got on average more correct answers than women. However, participants were more likely to predict that a woman won against a man than that she lost. Although there is no evidence of statistical discrimination against individual women, there is clear evidence of taste discrimination. Men winners were less likely to get the bonus compared to women who lost and women winners were less likely to get the bonus compared to men winners. Further, the group of women which systematically won against men was the group of women least likely to get the bonus. In other words, men were rewarded for winning and women were rewarded for losing against men.

Section 4.2 provides an overview of the existing research. Section 4.3 presents how discrimination is measured and describes the experimental design. Section 4.4 discusses the results and section 4.5 concludes.

4.2 Literature review

4.2.1 Taste versus beliefs

Becker (1957) defines taste discrimination as the differential treatment of a group because of a psychological benefit when interacting with this group. Gender taste discrimination is discrimination which results from a preference to interact with men or women. Akerlof (1985) expands taste discrimination to include cases, where an individual discriminates to comply with discriminatory customs and traditions. For instance, a director might not vote in favor of a female CEO, expecting that the shareholders will not accept her.

Using a double standard is also taste discrimination. Akerlof and Kranton (2010) argue that people reward individuals acting as one's peers and punish them if they do not. They give the example of a very productive woman who was not made partner in her company because she did not act as -how the other partners considered- a *female* should act. Several other examples for double standards can be found in the literature. Investors prefer female entrepreneurs who present a safe investment and male entrepreneurs who present a lucrative investment (Kanze et al. 2018). A woman's attractiveness predicts job market participation but a man's attractiveness does not (Hamermesh and Biddle 1994). Bursztyn, Fujiwara and Pallais (2017) show that female MBA students fear that they will be undesirable as partners when they act in an assertive and ambitious way, i.e. as male MBA students are expected to act.

Aigner and Cain (1977) conceptualize statistical discrimination as follows. When the signals about the agent's productivity potential are noisy, the principal interprets the signals based on his/her beliefs about the distribution of productive agents in an agent's socio-demographic group. Given that two agents send the same signal, principals favor agents from the group which they believe contains more productive agents. Beliefs can be expectations about the probability distribution of an agent's level of ability and that the agent will demonstrate certain behaviors. For example, Becker, Fernandes and Weichselbaumer (2019) provide empirical evidence that companies in German-speaking countries avoid calling back young married women for part-time jobs because of *a particular "risk" of becoming pregnant* and staying at home (Becker, Fernandes and Weichselbaumer 2019, p.149).

Because taste discrimination deters productive investments in human capital, it facilitates statistical discrimination (Becker 1985; Schwab 1986; Neilson and Ying 2016). Guiso et al. (2008) and Nosek et al. (2009) show that women who grew up in cultures which have a different sets of rules for women and men are more likely to perform bad in natural sciences. Carlana (2019) finds that exposure to gender biased math teachers leads to female students getting worse test scores in the mathematics portion of the SAT exams and less female students going to demanding high schools.

Furthermore, discrimination can cause intentional or unintentional under-performance. Psychologists show that discriminated individuals feel pressure to prove the stereotype wrong, which can keep them from performing their best (Steele and Aronson 1995; Spencer, Steele and Quinn 1999; Spencer, Logel and Davies 2016). Bursztyn, Fujiwara and Pallais (2017) show that female MBA students consciously under-perform to conform to female stereotypes. Glover, Pallais and Pariente (2017) observe that minority cashiers in French grocery stores work slower if the store manager is biased against them. Contrary, when the store manager is not biased, they are as productive as the non-minority cashiers.

Because the two types of discrimination feed back into each other, it is not easy to disentangle them. The measurement problem is aggravated by the fact that we lack a reliable measure for the size of total discrimination. For example, both types of discrimination lead to a lower participation rate of women in a labor market. However, the participation rates in a labor market with discrimination can be identical with the participation rates in a market without discrimination, when the provision of men and women workers is not balanced (Klumpp and Su 2013).

Researchers used field experiments to control for the under-supply of female workers. Riach and Rich (2002), Azmat and Petrongolo (2014), Lane (2016), Bertrand and Duflo (2017) and Neumark (2018) provide comprehensive surveys of the field experiments on discrimination. Independent of the exact experimental protocol, field experiments involve the introduction of equally qualified fictitious workers with a different gender in the labor market. Because the supply side of labor is kept constant, differences in call-back rates can only reflect gender discrimination. Since field experiments cannot provide sufficient control, researchers cannot use them to distinguish between taste and statistical discrimination or perfectly identify the relative weight of gender compared to other candidate's characteristics (commensurability problem).

To overcome the commensurability problem researchers conducted lab experiments. Lab experiments either asked whether gender determines if someone can be trusted, i.e. will be a reliable agent, or whether participants believe that women are as capable as men. This chapter discusses the latter question. To answer the latter question, vignette experiments have been mainly used in the past. The design involves asking participants to hire one worker from a pair of candidates with different characteristics. These experiments measure how the likelihood of a candidate being hired changes because of gender. In many societies, gender discrimination is condemned and participants want to hide the fact that they discriminate. Hence, one of the benefits of vignette experiments is that they mask the true question from the participants ("are you discriminating or not?").

Lane (2016) provides a meta-analysis of the results from incentivized lab experiments. Balliet, Wu and Dreu (2014) provides a meta-analysis of the results from non-incentivized lab experiments -mainly conducted by psychologists. Both studies estimate that on average gender discrimination reduces the willingness to cooperate or work with a woman by a quarter of a standard deviation. Another insight of the meta-studies is that the prevalence of discrimination is different depending on the domain of life and context. This result is intuitive. If discrimination is the norm in a specific context, both men and women adjust their behavior and expectations, i.e. act according to a discriminating equilibrium.

4.2.2 Similar work to this chapter

The work cited so far acknowledges the distinction between statistical and taste discrimination but does not try to measure the two independently. Coffman, Exley and Niederle (2020) ask whether taste discrimination exists in how people statistically discriminate. In a preliminary study, *workers* answer four quizzes: a pair of a hard and an easy maths quizzes and a pair of a hard and an easy sports quizzes. The authors then divide the *workers* in two groups, according to their gender and birthday: the *female-even month group* and the *male-odd month* group.

During the experiment, the *employers* see in the form of histograms the distribution of correct answers in each group and each worker's sum of correct answers in the easy math and sport quizzes. The authors then present the employers with pairs consisting of one worker from each group and ask them to hire one worker per pair. From the hired workers, one is randomly drawn and the employers get paid for each answer he/she gets correct in a hard quiz. In one treatment, the groups are labeled according to the month of birth and in another according to gender.

Coffman, Exley and Niederle find that when two workers have the same score in the easy quiz the propensity to hire the worker from the better performing group does not depend on the label, i.e. whether the group is labeled as the *male* group or the *odd month* group. To control whether participants believe that women are worse workers but hire them to avoid discriminating, the authors make hiring *women-even month* workers risky. In particular, they do not pay those who hire them with a certain probability. The authors argue that this modification provides a *veil of intentions* to the employers. Even with the veil, the labeling does not influence the likelihood that a worker will be hired.

(Coffman, Exley and Niederle 2020) and this chapter tackle a similar topic in a very different way. In (Coffman, Exley and Niederle 2020), the employers are informed how all workers performed in a related task. Thus, they can compare a worker with the rest of his/her group, the workers in the other group and the other worker. Considering the amount of available relevant information it comes to little surprise that labeling does not change how employers choose who to hire. In this chapter, the focus is not how beliefs are updated based on gender but the gender-based beliefs themselves. Subsequently, the participants do not receive any information about previous performance until they have revealed their beliefs.

Statistical discrimination is the only profit-generating strategy in the experiment reported in this chapter. Even one wrong prediction could cost the participants a significant amount of money. Contrary, Coffman, Exley and

Niederle pay the employers, even if they hire the worse worker. The employers just get paid less. Additionally, Coffman, Exley and Niederle use a knowledge-based task. In such tasks, the participants' performances are to a great extent decided before they start the task. Further, they focus in what they call the *male-typed domain* (sports and math). This allows them to measure gender discrimination, where it is most expected or where it is most likely to be salient. I use the task introduced by Mazar, Amir and Ariely (2008). The task is a gender-neutral effort-based task. Success in the task does not depend on previous skill and the task is designed so that all groups of participants can solve it equally well. Any discrimination in this setting signals a belief that women will put less effort than men or crack under the pressure. Last, I give the participants the opportunity to taste discriminate without risking their pay-offs.

4.3 Experimental design and expected behavior

4.3.1 Design

This subsection describes in detail the experimental design. The participants of the pre-experiment -henceforth competitors- got 15 matrices with twelve three-digit numbers, as seen in Figure 4.1.[3] The competitors had five minutes to find two numbers with a sum of ten in each matrix, henceforth the task.[4] After finishing the task, the competitors were randomly paired and the competitors who solved more matrices than their opponents won a ten Euro prize.

The instructions informed the competitors that a randomly selected participant from the experiment -henceforth the spectator- would decide whether they will get an additional bonus of four Euros. The competitors also knew that spectators would find out their score, i.e. how many matrices they solved, and their socio-demographic characteristics.

The pre-experiment consisted of seven sessions lasting a quarter of an hour each. There were three sessions with only male participants, three sessions

3. I recruited the participants of the pre-experiment and the experiment using hroot from the pool of the economic lab of the University of Hamburg (Bock, Nicklisch and Baetge 2012). The participants of the pre-experiment were excluded from the experiment.

4. Mazar, Amir and Ariely gave participants 20 matrices with a single solution per matrix and four minutes time to solve as many matrices as possible. Because pilots showed that this parameterization made the task impossible to solve, I re-calibrated the task so that a participant who puts constant effort over the entire time could solve all matrices.

Box 1			
4,99	8,88	4,56	4,58
4,98	8,44	5,44	4,95
4,96	0,41	8,43	4,54

Figure 4.1: Example: Matrix task

with only female participants and one mixed session. In the male and female sessions, competitors learned only through the questionnaire which they filled out after finishing the task that they could be matched with a competitor of the other gender. These measures reduce the salience of inter-gender competition and the risk of women adjusting their effort or under-performing.[5] The mixed session is used to see whether women compete differently when they are aware that they might compete against men. The performance of women in this session is not different than performance in the female-only sessions. This is a sign that the participants -at least in the mixed session- did not perceive the task as belonging in the male domain.

The experiment was formatted as a three parts survey, which was run over the internet with the online survey tool LimeSurvey (LimeSurvey GmbH). The survey took on average between eight to nine minutes to fill. The first part explained the task with an example and the conditions under which the competitors solved it (anonymity, time limit, payment for winning etc.). Then, it asked spectators to predict the winners in each competitors' pair. The spectators could predict that either competitor won or that they tied. Before they made their predictions, participants learned the gender, the age group and the academic degree of each competitor.

The second part asked participants to predict which groups of competitors had a higher average performance. The competitors were grouped into: 1) male or female; 2) competitors with at most a bachelor's degree or competitors with a higher degree; 3) competitors who were younger than 29 or older than 28 years old.

The participants were paid for making accurate predictions according to

5. If a competitor asked whether competitors from the other gender existed, they would have gotten a truthful answer. Nevertheless, no one asked.

the following rewards table:

- Most accurate participant: 300 Euro,
- 2nd-4th most accurate participant: 150 Euro,
- 5th-7th most accurate participant: 100 Euro,
- 8th-10th most accurate participant: 50 Euro,
- other participants: 0 Euro

The third part asked the participants to give four Euros (bonus) to one competitor from each pair. The participants were not monetarily remunerated for these decisions. After all participants submitted their decisions, one of them was drawn for each pair and his/her decision was implemented. A spectator could theoretically single-handily decide how the boni are distributed in all pairs (or decide about no bonus). Thus, each decision is equally important and the decisions of the same spectator and between spectators are independent of each other.

In the third part, spectators learned the competitors' scores. To avoid that spectators distribute the bonus based on the joy that they were correct or the disappointment when they weren't, I make it impossible for spectators to match competitors from the first part to competitors from the third part.[6] Specifically, I hide the competitors' age groups. Further, the order with which competitors are presented in the survey is randomized between each part for each spectator. Randomizing the competitors' order also ensures that ordering effects do not play a role.

After finishing the task/filling the survey, participants filled out a questionnaire. Both the spectators' and the competitors' questionnaires asked the participant's age group (<21; 21-24; 25-28; 29-32; >32), gender (man; woman; diverse) and the highest academic degree attained (no degree; Bachelor's degree, diploma or first state exam; Master's degree or second state exam; PhD). The competitors' questionnaires additionally asked whether they had experience with similar tasks or experiments and whether they believed that they would win and afterwards get the bonus. Moreover, the questionnaires invited the competitors to predict how the average male and female competitor scored. The spectators' questionnaires asked additionally to the main questions (gender, age, education) whether the spectator had

6. Only the competitors of two pair can be matched. One is the only pair where a woman competes against a man with a higher degree and the other is the only pair where two men compete and only one of them has a master's degree.

a rural upbringing, migration background and whether they identify themselves as members of a religious community and which. Once the experiment was concluded, payments were made through Amazon vouchers to all participants.

4.3.2 Expected behavior in the experiment

During the first two parts of the experiment, the spectators receive information about the competitors and make predictions which can earn them a prize. Ceteris paribus, spectators who do not correlate gender with performance in the task will predict with an equal probability that a man won and that a woman won and with some probability that the outcome was a tie. Then, they will flip a coin to choose how to answer the question whether women had a higher average than men. A spectator who correlates gender with performance, i.e. statically discriminates, will be more/less likely to predict that a woman won than that a man won. If spectators are more likely to guess that a woman won against a man, then they should also be more likely to guess that women had a higher average score than men.

Hypothesis 4.1a (Pro-male bias) *A spectator who observes a competition between a man and a woman is more likely to predict that the man won rather than that he lost.*

Hypothesis 4.1b (Pro-female bias) *A spectator who observes a competition between a man and a woman is more likely to predict that the man lost rather than that he won.*

During the third part, the spectators can distribute the bonus according to any rule they want. The instructions call the bonus "a money transfer" to avoid priming the spectators to give the bonus to the winner. Spectators who do not care who gets the bonus, i.e. have no or weak other-regarding preferences, will just click-through the third part of the survey. Because LimeSurvey randomized the order of the competitors, such strategies cannot be mistaken for patterns in the data. Spectators who care about who gets the bonus will give the bonus to their favored competitor.

In the latter case, there are several potential criteria with which a spectator can decide which competitor they want to get the bonus. A competitor-based criterion ranks competitors based on their characteristics independent of performance, i.e. relies on taste discrimination. For example, in-group favoritism entails giving the bonus to the competitor, who is more similar to oneself, e.g. has the same gender. Contrary, positive discrimination would have the spectator trying to *nullify the accidents of natural endowment* (Rawls 1971, p.15). In simpler terms, he/she will give the bonus to the competitor who is perceived as less able to complete the task. The identity of this competitor depends on the spectator's beliefs about the probability that the competitor could win against his/her opponent.

A performance-based criterion ranks competitors according to the outcome. A spectator could be rewarding the best/worst score (output-oriented) or compensating for the effort put in the task (input-oriented). Rewarding the winner intuitively appears to be fair. The best performance is often correlated with more effort, more skill in the task and motivates investing effort in future tasks. Although rewarding the worst score seems counter-intuitive, this strategy reduces pay-off inequality. The best performer got the winner's prize. By giving a consolidation prize to the loser, the loser is motivated to participate in future competitions. Additionally, he/she does not leave the competition empty-handed.

Because the actual effort in the task is not observed, an input-oriented decision combines beliefs about productivity and performance in the task. Specifically, the score difference that signals superior effort depends on the beliefs about skill differences. The logic is that a competitor who is believed to be significantly more skilled must have invested less effort to get a marginally better, similar or worse score than the unskilled competitor.

Spectators might use different criteria based on the competitors' genders (double standards). Two men competing might prime equity concerns whereas when a man and a woman compete equality might become more salient. Which mix of criteria is actually used is an empirical question.

Hypothesis 4.2 (Rawlsian preferences) *Spectators are more likely to give the bonus to the competitor who they consider least likely to win than to the competitor they consider more likely to win.*

Hypothesis 4.3 (Same gender favoritism) *Spectators are more likely to give the bonus to the competitor with whom they share a common gender than to a competitor with whom they do not share a common gender.*

Hypothesis 4.4 (No double standard) *Spectators give the bonus to women winners with the same probability as they give it to men winners.*

4.4 Experimental results

4.4.1 Descriptive statistics

This section presents the participants' pool and discusses their decisions during the experiment. Table 4.1 groups the participants in terms of their gender, age group and highest degree. The first number in each cell represents the number of male participants and the second number the number of female participants. Last, the number in the parenthesis is the total number of participants in each group.

Degree/ Age	No degree	Bachelor's or equivalent	Master's or equivalent	PhD	Total
			Competitors		
17-20	1/1 (2)	-/- (0)	-/- (0)	-/- (0)	1/1 (2)
21-24	7/3 (10)	2/3 (5)	-/- (0)	-/- (0)	9/6 (15)
25-28	8/2 (10)	3/3 (6)	1/2 (3)	-/- (0)	12/7 (19)
> 28	2/1 (3)	2/1 (3)	0/4 (4)	-/- (0)	4/6 (10)
Total	18/7 (25)	7/7 (14)	1/6 (7)	-/- (0)	26/20 (46)
			Spectators		
17-20	15/15 (30)	-/2 (2)	-/- (0)	-/- (0)	15/17 (32)
21-24	36/66 (102)	22/31 (53)	3/- (3)	-/- (0)	61/97 (158)
25-28	8/17 (25)	26/46 (72)	12/17 (29)	1/- (1)	47/80 (127)
29-32	5/ 8 (13)	16/11 (27)	4/7 (11)	1/- (1)	26/26 (52)
> 32	3/ 4 (7)	6/9 (15)	2/6 (8)	-/- (0)	11/19 (30)
Total	67/110 (177)	70/99 (169)	21/30 (51)	2/- (2)	160/239 (399)

Table 4.1: Participants' pool

It can be seen that women competitors were not significantly older than men, however they had higher degrees ($\chi^2 = 7.78$, p-value$= 0.021$). Women spectators are neither significantly older nor have a higher degree than men spectators.[7] Further, older spectators tend to have higher degrees ($\chi^2 = 117$, p-value$= 0.000$).

In the questionnaires, almost half of the competitors (21/46) predicted that on average women will get the same score as men. Interestingly, ten men predicted that women would solve on average more matrices than men, whereas only three women made a similar prediction. Five of the top ten

7. One spectator did not disclose her degree. I impute the degree by using the most frequent degree in her age group (Bachelor's), which is also the most common degree for women and women in her age group.

men and two men with below average score stated that they believe that they will get the bonus. Only the fourth best and the eighth worst woman stated that they believed that they will get the bonus. Surprisingly, the top three women did not believe that they will get the bonus.

The division of labor and socialization varies between different countries and communities within the same country. Further, differences in the exposure to successful women can lead to different beliefs about the ability of women. 271 spectators answered that they were born and raised in Germany by German parents. 76 spectators were born and raised in Germany by at least one migrant, i.e. a parent not born and raised in Germany. The rest were born and raised outside of Germany.

144 spectators identified as Christians and 23 identified as Muslims in the questionnaires. 223 did not identify as members of a religious community, were agnostic, atheists or did not disclose this information. The rest identified as members of various other religions (e.g. Judaism, Buddhism or Hinduism). Last, approximately 80% of the spectators grew up in an urban region.

Three men and one woman competitors solved all matrices. Women solved on average less matrices than men but their scores varied less ($\mu_{women} = 7.75$, $SD_{women} = 3.878$ versus $\mu_{men} = 9$, $SD_{men} = 4.699$). The score differences between men and women are not statistically significant (Mann-Whitney U-Test p-value= 0.307). Moreover, a man is not more likely to have a higher score than a woman of the same age or the same degree (Kruskal-Wallis tests p-value= 0.975 and 0.473, respectively).

The competitors were randomly matched in 23 pairs. Table 4.2 reports the competitors' characteristics and their scores in each pair. There never was a tie and the average score difference was approximately 5 matrices (SD=3.175). In no pair the competitors had identical characteristics. As a measure for how different the competitors in a pair were from each other, the table reports the Gower's dissimilarity coefficient for each pair.[8] The table also reports the spectators' predictions for each pair and what portion of the spectators wanted the bonus to go to the winner (BtW).

In most pairs, competitors with a score below median competed against competitors with a below median score or competitors with a score above median competed against competitors with a score above median. Only in seven pairs the winner had an above median score while the loser a below median score. In 16 pairs, women competed against men -henceforth mixed pairs- and in seven pairs the competitors had the same gender -henceforth

8. The Gower's general dissimilarity coefficient allows comparisons, when the characteristics compared are measured in different scales (nominal scale: gender; ordinal scale: age group, degree). Lower numbers reveal more similar pairs. (Gower 1971).

Gender	Age group (Competitor A/Competitor B)	Degree	Score	Gower (%)	Predictions (A/B/Tie)	BtW (%)
Both men	21-24/>32	None/bachelor's	15/14	36.1	238/82/79	60.4
Both men	Both 25-28	None/bachelor's	12/4	11.1	24/200/175	55.4
Both men	Both 25-28	None/master's	9/13	22.2	30/259/110	59.1
Both men	Both 25-28	Bachelor's/none	5/13	11.1	203/31/165	57.4
Both men	>32/21-24	Both none	14/10	25.0	46/289/64	59.9
Both women	17-20/21-24	Both none	9/4	8.3	170/125/104	59.1
Both women	21-24/25-28	Both bachelor's	13/8	8.3	197/63/139	58.4
Woman/man	Both 21-24	Both none	0/2	33.3	133/88/178	59.4
Woman/man	21-24/25-28	Both none	8/5	41.7	94/186/119	59.4
Woman/man	21-24/25-28	Bachelor's/none	10/15	52.8	63/238/98	61.2
Woman/man	21-24/29-32	Both bachelor's	15/8	50.0	94/205/100	57.9
Woman/man	25-28/17-20	Bachelor's/none	10/9	61.1	165/157/77	59.1
Woman/man	25-28/21-24	Both none	6/2	41.7	227/79/93	58.1
Woman/man	25-28/21-24	Both bachelor's	4/15	41.7	207/81/111	58.6
Woman/man	Both 25-28	None/bachelor's	3/7	44.4	207/85/107	54.1
Woman/man	Both 25-28	Master's/none	9/12	55.6	87/247/65	62.9
Woman/man	Both 25-28	Master's/none	7/1	55.6	66/222/111	56.6
Woman/man	>32/21-24	Both none	12/10	58.3	287/49/63	59.1
Woman/man	>32/21-24	Master's/none	10/5	80.6	205/119/75	56.6
Woman/man	>32/21-24	Master's/none	6/7	80.6	206/122/71	61.2
Woman/man	>32/21-24	Master's/bachelor's	2/14	69.4	227/82/90	57.4
Woman/man	>32/25-28	Bachelor's/none	7/12	61.1	189/95/115	58.9
Woman/man	Both >32	Master's/none	12/1	55.6	46/219/134	57.6

Table 4.2: Competitor-pairs and spectators' answers

same gendered pairs.

Women won against men as many times as men won against women. The holder of the lower degree won eight times out of the 14 times the competitors had a different degree. Last, the younger competitors won against older competitors in nine out of 15 pairs in which competitors belonged in a different age group.

In the first part of the survey, 53 spectators never predicted a tie and 16 spectators (9 men and 7 women) always predicted a tie. From those that never predicted a tie, five spectators predicted that the woman and four predicted that the man always won in mixed pairs. Because there is no variation in their predictions, the sixteen spectators who always predicted a tie and the five plus four who always predicted that a competitor with a specific gender won are excluded from the analysis. Spectators were significantly more likely to predict a tie in the same gendered pairs than in the mixed pairs ($\chi^2 = 22$, p-value< 0.000). Overall, women spectators were significantly more likely to predict that the outcome was a tie ($\chi^2 = 22$, p-value$= 0.001$). This is tentative evidence that gender plays a role for spectators.

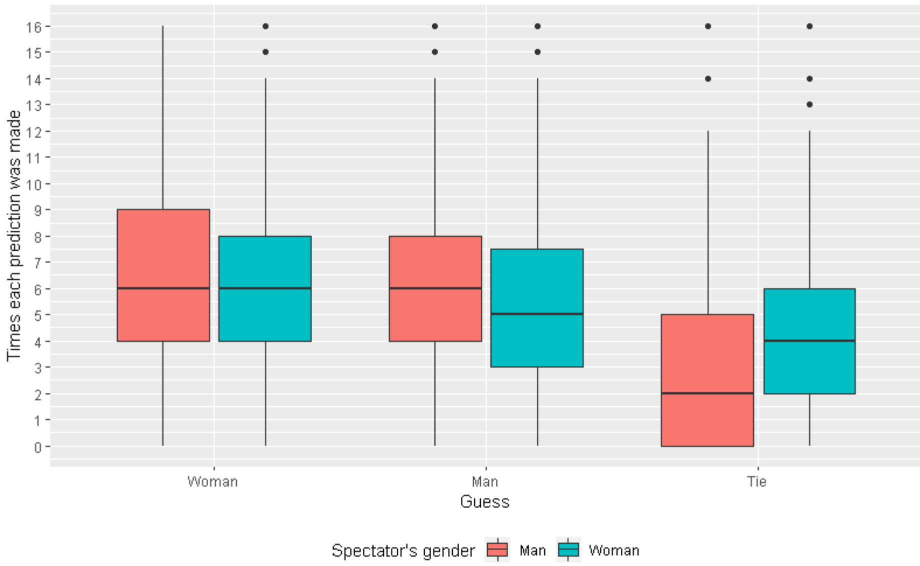


Figure 4.2: Boxplot: Distribution of the predictions for mixed pairs

Figure 4.2 shows the men's and women's distribution of predictions for the mixed pairs. 39.20% of all predictions about the mixed pairs chose the woman as the winner. 35.6% of the predictions said that she will lose and the rest predicted that she will tie. Both men and women were more likely to predict that a woman won rather than that a man won.

In the second part of the survey, 61.4% of the spectators predicted that women solved on average less matrices than men. Less than half the spectators predicted correctly that competitors with a master's degree scored worse that the rest of the competitors (46.9%). Even less spectators (12%) predicted correctly that older participants (>28 years old) managed to solve more matrices compared to younger participants (<29 years old). In total, only two participants predicted all three questions correctly. Interestingly, the predictions in the second part were not consistent with the prediction in the first part. For example, those that predicted that women are on average better than men predicted less often that women won (Mann-Whitney U-Test p-value< 0.000).

In the third part of the survey, one (woman) spectator always gave the bonus to the losing competitor. The rest gave the bonus to the losing competitor in between five and 18 out of the 23 pairs. Spectators gave the bonus to the loser significantly less often in the same gendered pairs ($\chi^2 = 1384$, p-value< 0.000). They also gave the bonus significantly more often to the losing man than the losing woman in the mixed pairs ($\chi^2 = 28$, p-value< 0.000).

161 spectators always gave the bonus to the winner in the same gendered pairs and never in the mixed pairs. 63 spectators always gave the bonus to the winner in the mixed pairs. From them, 62 never gave the bonus to the winner and the remaining one gave the bonus to the winner once in the same gendered pairs. Seven spectators (one man and six women) always gave the bonus to men and there was no spectator who always gave the bonus to the women.

The descriptive statistics can be seen as another indication that gender matters for spectators. Further, they seem to indicate that the rules which apply to mixed gendered competitions are different than the rules for same gendered competitions. In the next subsection, I use regression analysis to ensure that the patterns in the data reflect gender discrimination and are not driven by other competitors' differences or are representative only for specific groups of spectators.

4.4.2 Gender statistical discrimination

The experiment ensures that the probability a spectator makes one of the three possible predictions (A: competitor A won, B: competitor B won, T: they tied) is equal to one, i.e. all spectators are given the same options to choose from and answer all questions. Under these conditions, multinomial regression (ML) can be used to estimate how the odds of predicting that a competitor won over predicting a tie change depending on the competitor's, the opponent's and the spectator's characteristics.

The model estimates the latent propensity to make each prediction by looking at which prediction k from K possible predictions the i-th spectator made for the j-th pair ($y_{ij} = k$). Let a spectator have a latent propensity y^*_{kij} to predict $k \in K$, where:

$$y^*_{ki} = V_{ki} + \epsilon_{ki}$$

with V_{ki} representing the i-th spectator's utility from predicting k and ϵ_{ki} being unobservable factors that influence the spectator's prediction, such as for example the spectator's mood. Because a spectator's utility increases with the reward for being accurate, the prediction's utility increases with the belief that a prediction is correct. The probability to make prediction k can be written as a function of the prediction's utility as:

$$Pr(y_{ij} = k|V_{ki}) = \begin{cases} \dfrac{e^{V_{Ai}}}{1 + e^{V_{Ai}} + e^{V_{Bi}}} & |k = A \\ \dfrac{e^{V_{Bi}}}{1 + e^{V_{Ai}} + e^{V_{Bi}}} & |k = B \\ \dfrac{1}{1 + e^{V_{Ai}} + e^{V_{Bi}}} & |k = T \end{cases}$$

where:

$$V_{ki} = \beta \text{Competitor}_k + \gamma_k \text{Spectator}_i + \eta_k \text{Prediction}_{ki} + \alpha_k$$

α are the prediction-specific intercepts. *Competitor* is a matrix with the predicted winner's characteristics and its values do not vary between spectators. *Spectator* is a matrix with the spectator's characteristics and its values do not vary between pairs. *Prediction* is a matrix with the particular characteristics of each prediction and its values vary between competitors and pairs. All coefficients indicate the change in the odds ratio between predicting a tie, i.e. the baseline prediction, and predicting something else.[9]

The matrix *Competitor* contains one dummy for each of the values of the following three categorical variables:

- *Age* with values: Older, younger, same, not important;
- Degree with values: Higher, lower, same, not important;
- *Gender* with values: Male, female, same, not important.

The *not important* dummies take a value of one when the competitors have a different age, degree or gender and the spectator predicts a tie. The *same* dummies take a value of one, when the competitors have a similar age, degree or the same gender. Otherwise, the appropriate dummy takes the value of one depending on whether the predicted winner is younger than his/her opponent, has a higher degree or is a woman competing against a man.[10]

The matrix *Spectator* contains dummies for the spectator's gender, his/her age group and academic degree. Moreover, it contains dummies for whether

9. For an extensive discussion on how to calculate multinomial models, their mathematical derivation, shortcomings, advantages and extensions see: (Train 2009).

10. I model the βs as correlated random parameters. A fixed effects specification would require estimating $n-1$ coefficient for $K-1$ possible predictions: (3-1)x(399-1) coefficients. Such an endeavor is meaningless because of the incidental parameter problem. Further, due to the size of the data set, it is computationally impossible to estimate a conditional multinomial logit as derived by Chamberlain (1980).

the spectator is a foreigner or a German with migration background, whether he/she was raised in a rural area and whether he/she identifies as a Christian or Muslim. The first set of dummies captures the spectator's sociodemographic characteristics and the latter set of dummies captures the environment in which the spectator grew up and developed his/her beliefs.

The matrix *Prediction* contains six variables: *Other predictions: woman won*, *Other predictions: man won*, *Other predictions: the competitors tied*, *Distance to A*, *Distance to B*, *Distance between A and B*. The spectators can always predict a tie. However, they cannot predict that a woman won when two men compete. The variables *Other predictions* measure the portion of the times a prediction was made given that it was available ($\mu = 0.379$, SD$= 0.211$) and captures the spectator's propensity to chose winners with a certain gender.

There is robust empirical evidence that individuals overestimate the abilities of those they perceive as their peers (Everett, Faber and Crockett 2015). *Distance to A* and *Distance to B* are the Gower distances between the spectator and the competitor who the spectator predicted to be the winner. *Distance between A and B* is the Gower distance between a pair's competitors. This distance is the same between spectator's and the same for each competitor in a pair but changes between pairs.

Table 4.3 reports the results. The reported coefficients are odds ratios and the odds ratio adjusted standard errors (SE) are reported in the parentheses. Column one reports the results of a reduced model with only the competitor's characteristics. The model in column two expands the model in column one by adding the spectator's characteristics. Column three reports the coefficients from the full model. All of the aforementioned models are estimated on the sub-sample of spectators who varied their predictions.

Column four and five report the results from estimating the full model on sub-samples of spectators which could be biased against women. Column four restricts the sample to the spectators who predicted that men have a higher average than women. Column five restricts the sample to men spectators.

The model reported in column six limits the sample to the mixed pairs. In these pairs, the woman was always coded as competitor A and the man as competitor B. To estimate the model on the sub-sample, I have to drop the matrix *Competitor*. Subsequently, the intercepts capture the effect of gender, age and education differences combined.

The first and the second model suggest that the odds a spectator predicts that a woman won are larger than the odds that he/she predicts that no one won. When one controls for the spectator's other predictions the effect of gender disappears. In other words, there are spectators who discriminate systematically in favor of women (e.g. other women) and spec-

Multinomial logistic regression DV: Prediction of i-th spectator for j-th pair						
	(1)	(2)	(3)	(4)	(5)	(6)
Sample	All	All	All	$\hat{\mu_M} > \hat{\mu_F}$	Men	Mixed
Female	1.138**	1.362***	1.110	0.998	1.353	
	(0.072)	(0.105)	(0.103)	(0.115)	(0.250)	
Male	1.099	1.194**	0.894	0.767**	1.182	
	(0.071)	(0.091)	(0.082)	(0.088)	(0.204)	
Younger	1.949***	2.467***	1.872***	1.581***	1.631***	
	(0.124)	(0.188)	(0.137)	(0.145)	(0.212)	
Older	2.721***	3.437***	2.721***	2.405***	2.832***	
	(0.178)	(0.262)	(0.198)	(0.218)	(0.369)	
Lower degree	1.766***	2.285***	1.509***	1.415***	1.729***	
	(0.118)	(0.180)	(0.110)	(0.132)	(0.228)	
Higher degree	1.493***	1.890***	1.246***	1.068	1.171	
	(0.095)	(0.146)	(0.090)	(0.097)	(0.154)	
Prediction: Competitor A won						
Constant		0.916	0.562***	0.727	0.446**	0.244***
		(0.144)	(0.117)	(0.194)	(0.158)	(0.065)
Distance			0.969	0.782	3.345***	7.300***
			(0.134)	(0.134)	(0.939)	(1.653)
Prediction: Competitor B won						
Constant		0.845	0.635**	0.842	1.878*	0.580**
		(0.131)	(0.127)	(0.221)	(0.682)	(0.130)
Distance			2.207***	1.993***	0.148***	52.661***
			(0.385)	(0.440)	(0.052)	(19.130)
Prediction: Competitors tied						
Distance			0.788	0.802	1.956	0.980
			(0.178)	(0.228)	(0.841)	(0.225)
Spectator's control	No	Yes	Yes	Yes	Yes	Yes
Observations	8,602	8,602	8,602	5,428	3,381	5,984
Spectators	374	374	374	236	147	374
LogLikelihood	−8,508	−8,455	−7,693	−4,854	−2,919	−5,295
AIC	17,070	17,020	15,509	-	-	-

The coefficients are the odds ratios and the odds ratio adjusted standard errors are reported in the parentheses. Spectator's controls: the spectator's other predictions, gender, age group and academic degree dummies, whether he/she had a rural upbringing and his/her migration and religious background (Christian and Muslim dummies, German with migration background and foreign dummies). *p<0.1; **p<0.05; ***p<0.01.

Table 4.3: Likelihood of a prediction in part one

tators who discriminate in favor of men. Because the former are more consistent than the latter, they inflate the coefficient for *Female* (Wald test: H_0 $\eta_{Other\ predictions:\ Woman\ won} = \eta_{Other\ predictions:\ Man\ won}$, $\chi^2 = 4$ p-value= 0.049).

In the second model, the coefficient for $Male$ is significantly different than one. Notwithstanding, the result is not robust. Further, the coefficient for $Male$ is significantly smaller than the coefficient for $Female$ in models two and three (Wald test H_0: $\beta_{Female} = \beta_{Male}$, Model 2: $\chi^2 = 4$, p-value= 0.047; Full model: $\chi^2 = 11$, p-value< 0.000). This result indicates that spectators are more likely to predict that a woman won rather that she lost. Considering most spectators predicted that men had a higher average score than women, the result is surprising.

Even the groups of spectators which would be expected to discriminate against women do not. Men are as likely to predict that a woman won against a man as they are to predict that she tied or lost. Also spectators who predict that women have a lower average score are more likely to predict that a woman won against a man rather than that she lost (Wald test H_0: $\beta_{Female} = \beta_{Male}$, $\chi^2 = 10$, p-value= 0.001).

The discrepancy between the predictions in part one and in part two is unexpected but not unexplainable. To compare two competitors, spectators need to consider the entire distribution of scores. For example, if the distribution of women's scores is negatively skewed and the distribution of men's scores is positively skewed, the median woman performs better than the average woman and the median man performs worse than the average man. As a result, women might have a lower average score as a group but still be more likely to win against men.

An alternative explanation is that gender discrimination carries a social stigma. The spectator who were more likely to predict that men won were the spectators most likely to predict that women had a higher average. Although, spectators profit monetarily by maximizing the accuracy of their prediction, their predictions convey information about who they are and their beliefs. Discriminating against women potentially carries a psychological cost (guilt/self-doubt). A feeling of guilt and self-doubt could explain why the spectators refuse to admit that they believe that women are better than men even though they discriminate against them.

There is no clear evidence that spectators favor competitors who are more similar to them. Additionally, there is no evidence that spectators are more likely to choose a tie, when the competitors are equally different from them. The γs, which are not reported in the table, were insignificant. In other words, there was no systematic correlation between the spectator's characteristic and their predictions. The only exception is that women were less likely to choose a winner (competitor A or B) than predict a tie.

When competitors have a different age, spectators are more likely to predict that the older competitor won. The regressions also show that spectators held the competitor with a higher degree as less likely to win. Combining the effects of age and the education difference, the model suggests that the odds a spectator predicts that an older competitor with a higher degree wins are 3.39 (Odds adjusted (OA) SE= 0.013) times higher than the odds of predicting a tie. If the older competitor has a lower degree, being chosen as the winner is 4.11 (OA SE= 0.014) times more like than the spectator predicting that he was as good as his/her opponent. Accordingly, a younger competitor is 2.33 (OA SE= 0.011) and 2.82 (OA SE= 0.012) times more likely to be chosen as the winner when he has a lower and a higher degree respectively. All these differences are significant (p-value$<$ 0.000).

When restricting the sample to mixed pairs, the regression's coefficients suggest that spectators are less likely to predict that competitor A, who was a woman, won compared to predicting that competitor B, who was a man won (Wald test: H_0 $\alpha_A = \alpha_B$: $\chi^2 = 11$, p-value$<$ 0.000). This result seems unexpected considering the results discussed so far. The sixth model does not distinguish between older or younger more or less educated men and women. One could argue that spectators respond to other characteristics of the competitors in mixed pairs and not gender. Such a characteristic could possibly be that only one woman had a lower than her male opponent's degree.

To control for the effect of age and education differences, I estimate a number of regressions on sub-samples of mixed pairs with similar pairs of competitors, i.e. pairs in which the competitors are similarly different. Table 4.4 reports the odds ratio between predicting that a man won and that a woman won based on these regressions. The last column reports the numbers of pairs included in each regression. All regressions are based on the predictions of 374 spectators. I also include the *Other predictions* variable in the model to control for heterogeneity between spectators.

Women's characteristics	Odds ratio	OA SE	p-value	Pairs
Higher degree older	0.288	0.368	0.069	5
Higher degree same age group	18.531	4.908	0.010	3
Higher degree younger	0.023	0.199	0.004	1
Lower degree same age group	4.96	1.778	0.045	1
Same degree older	0.063	0.206	0.001	3
Same degree and age group	1.709	1.304	0.591	1
Same degree younger	0.126	0.369	0.047	2

Table 4.4: Odds ratios estimated with sub-samples of the mixed pairs

The spectators were not significantly more likely to predict that the woman who competed against a man with a similar age and the same academic title won, lost or tied. This is a first sign that gender alone does not make the spectators believe that one competitor is better than another. Every time a woman has a different age than her opponent, the spectators are significantly less likely to predict that she won. However, they do not consistently predict that younger women win more often than older women. This observation goes against the results discussed so far. The full model shows that the spectators consider younger competitors as more likely to lose.

The woman who had a higher degree than her male opponent at a younger age was the second least likely woman to be predicted as a winner. Unexpectedly, the spectators who predicted that her opponent won were correct. Although, she belonged in the top five women she competed against one of the three men who solved all matrices. Otherwise, spectators are more likely to predict that women with a higher degree than their male opponents won rather than predict than women with the same degree as their male opponents won.

In the models, I consider whether the predicted winner was younger/older or had a higher/lower degree than the opponent. I neglect how much the predicted winner is different from the opponent or what is the actual age and degree of the competitor. To account for the impact of the two, I estimate two regressions. One model contains dummies for the competitor's age group and degree, instead of the younger/older and higher/lower degree dummies. The other model contains dummies for how many groups separate the competitor and his/her opponent.[11] How age and education differences enter the model does not significantly influence the coefficients for the competitor's gender dummies.

Including the variable *Other predictions* assumes that the spectators scroll back and forth in the survey and fill the answers in a random order or check that they are consistent before submitting their predictions. This is possible because all questions from the first two parts are shown simultaneously and spectators can fill them in any order. One could argue that spectators are more likely to fill one question after the other.

To account for such an eventuality, I replace the variable *Other predictions* with a variable *Past predictions*. The variable measures how many times the

11. The variables take the value *none* for a tie when the competitors have a different age or degree. When the competitors have the same age or degree, the variables take the value of the competitors' common age group or degree even for a tie. In the second regression, the values for a tie represent how many groups away the competitors are. For example, a 17-20 years old competitor is one group away from a 21-24 years old and two groups away from a 25-28 years old competitor.

spectator made a similar prediction in pairs placed in the survey before the pair for which he/she was currently predicting. The results were robust to how other predictions entered the model.

The *Spectator* matrix ensures that any systematic variation due to the spectator's socialization would not bias the results. Nevertheless, the coefficients for the variables controlling for socialization are all insignificant. As a robustness test, I estimate a model without the controls for the community in which the spectator was socialized. Further, I estimate a model with dummies for each combination of religion, migration background and (rural) upbringing. The coefficients for the variables of interest are not significantly different in these two models compared to the main model.

Although socialization does not appear to correlate with statistical discrimination, it could impact the accuracy of the spectators. To test this intuition, I regress whether the spectator predicted the winner correctly in a mixed pair on the winner's characteristics and the spectator's characteristics. Spectators were more accurate at predicting woman winners rather than man winners. In particular, they overestimate the ability of men to perform better than women. Non-Muslim foreigners were less accurate at predicting the winner when women won. Contrary, foreigners who identified as Muslims were the most accurate spectators at predicting when women won. To sum up, socialization seems to change the distribution of wrong predictions.

Spectators might try to act politically correct. Thus, even spectators who believe that women are worst performers than men might predict a tie to hide their actual beliefs. Assuming that this is true, the odds of predicting that a man won will increase more than the odds of predicting that a woman won if the option Tie is removed, i.e. the independence of irrelevant alternatives (IIA) is violated. How this potential heteroscedasticity is controlled for in the model could bias the results.

In the random parameters model, the results can vary depending on how the random parameters are modeled. As a first step to test the robustness of the results, I estimate models with different correlated and uncorrelated random parameters. The results do not qualitative change with different specifications of the random parameters. I further try to avoid biases because of the random parameters altogether by estimating nested MLs and a heteroscedastic ML as derived by Bhat (1995). The results are not qualitative different between the random parameters, the nested and the heteroscedastic MLs.

To sum up, there is no evidence that the average spectator has a pro-male bias. Thus, the null cannot be rejected for hypothesis 4.1a. However, there is robust evidence that gender plays a role for performance in the eyes of the spectators. Although spectators state that they believe that women are worse than men in the task, they act as if they believe that women are as good and some parts of the spectators as if they believe women are better than men. Subsequently, the null cannot be rejected for hypothesis 4.1b. Interestingly, spectators weigh age and acquired education differently between men and women (double standard). This is a first indication that spectators gender taste discriminate.

4.4.3 Gender taste discrimination

I do not find definitive evidence that spectators believe that male competitors are better than female competitors in the task. However, I provide some evidence in the previous subsection that spectators use a double standard to predict how productive a man and a woman are. This subsection relies on conditional logistic regression to answer whether the spectators use gender as a criterium for who gets the bonus, i.e. gender taste discriminate. The full model is:

$$\begin{aligned}\log\frac{Pr(y_{ij}|x)}{1-Pr(y_{ij}|x)} =&\beta_1\ Mixed\ Pair_j + \beta_2\ Woman\ Won_j + \\ &\beta_3(Mixed\ Pair_j \times\ Woman\ Won_j) + \\ &\gamma\ Controls_{ij} + \alpha_i + \epsilon_{ij}\end{aligned}$$

where y_{ij} is a dummy taking the value one when the i-th spectator gives the bonus to the winner in the j-th pair and ϵ_{ij} is the error term. $Mixed\ Pair$ is a dummy for whether the j-th pair is mixed and $Woman\ Won$ is a dummy that takes the value of one, when the winner was a woman. Last, $Controls$ is a matrix with control variables for the competitors' education differences and relative performance, the spectator's past behavior and similarity to the winner and the likelihood that the winner would win.

In the model, α_i are spectator-specific intercepts, i.e. the coefficients for the binary strata indicators. For simplicity, henceforth the β and γ coefficients will be reported as odd ratios. Subsequently, β_1 captures whether the odds that male winners get the bonus change, when they win against women. β_2 captures whether a woman winner competing against a woman is more or less likely to get the bonus compared to a man winning against another man. β_3 shows whether the odds of a woman winner to get the bonus are the same when she competes against a man and when she doesn't.

To remind the reader, spectators see the competitors' genders, degrees and scores before they distribute the boni. Studying for a longer period of time is costly. Indicatively, over a quarter of all students in Hamburg interrupt their studies either to work or because they cannot afford to continue them (Schirmer 2017, p.11). A higher degree not only indicates that a competitor is older but to some extent reflects his/her financial means.

In nine pairs, the winner had the same degree as the loser (reference group). In six pairs, the winner had the previous degree compared to the loser's degree (no degree vs bachelor's or bachelor's vs master's). In four pairs, the winner had a master's degree and the loser no degree and the rest of the pairs were equally split between the winner having no degree and the loser a master's degree and the winner having the next degree compared to the loser. I include one dummy for each of the aforementioned degree differences to control for the effect of a winner being in a different stage of his/her studies.[12]

A competitors' score enters the model in multiple ways. First, I look into how good the competitor and his/her opponent were compared to the median competitor. The intuition is that spectators might be less inclined to reward a winner, who they perceive as lucky to be matched with another "slacking" competitor or more inclined to reward a loser who *puts his/her best* but is matched with a stronger opponent. I divide the pairs into three groups. In one group, both the winner and the loser had a higher score than the median competitor. In another group, the winner and the loser had a lower than the median scores. In the reference group, the winner had a higher score and the loser a lower score than the median.

Second, I include how many more matrices (cubed) the winner solved compared to the loser ($Score\ Difference^3$). Similar scores reflect a comparable effort between the competitors. If spectators value similar effort being compensated similarly, a small score difference is reason to give the bonus to the loser and reduce the pay-off inequality between winner and loser. Contrary, a very big score difference indicates that the loser did not try sufficiently. As a robustness test, I replace the score difference with the rank difference, i.e. how many places apart the competitors were in the ranking. The results are robust to how relative effort is modeled.[13]

The stratification captures any variation due to characteristics of the spectators which are fixed over the different pairs. Thus, a variable like *Other predictions* is not necessary. However, one could argue that spectators decide for each pair as they go and do not spend time scrolling back and forth

---

12. As a robustness test, I estimate the models with higher/lower degree dummies and dummies for the winner-loser degree combinations: both bachelor's, none vs bachelor's, bachelor's vs none, none vs master's, master's vs none, bachelor's vs master's. The coefficients of the main variables remain qualitatively the same.

13. I also estimate the model with the number of matrices, its square, log and inverse. The cubed number provides the best fit. Whether the $Score\ difference$ is transformed and how do not qualitatively change the results for the main variables.

in the survey. This argument is particularly meaningful in the third part of the survey because spectators are not paid for investing time to go back and forth and check their answers. To account for such a behavior, the matrix *Control* contains a variable for how often the spectator gave the bonus to winners in pairs shown before the j-th pair in the survey.

Additional variation due to spectator's characteristics which is not captured by the stratification is whether the winner's gender is the same as the spectator's gender. The dummy *Gender congruency* takes the value of one when the winner and the spectator have the same gender. A significantly positive coefficient for this variable would reveal in-group favoritism.

Last, I control for whether the winner is perceived as more skilled than the loser. As discussed, spectators with Rawlsian preferences will give the bonus to the loser if they believe that he/she did not stand a chance to win. The competitor's objective (true) probability of winning is conditional on two events: how good is a competitor in the task and whether his/her opponent is worse than him/her. Spectator do not know the true probability of winning. Nevertheless, they form priors about the distribution of this conditional probability after they learn the competitors' genders, ages and academic qualifications and the distribution of competitors, i.e. how many women/men etc. are in the sample. Further, the results of phase one give information to the spectator about the true distribution of the conditional probability. Thus, enable them to update their priors.

To estimate the impact of the spectator's beliefs on the odds to give the bonus to the winner, I rely on three measures: the spectator's revealed priors about which competitor is better ($Priors$), the likelihood that a winner from a certain competitors' group would win against an opponent from the opponent's group (*Likelihood*) and the interaction term of the two. Including all three terms in the model accounts for belief updating.

Spectators reveal their priors through their predictions. If a spectator consistently predicts a tie, I assume that the spectator appoints an equal odd to either competitor winning. For the rest of the spectators, I use the model from the previous subsection to estimate the odds ration of the odd that the spectators would predict that the winner won against his/her opponent and the odds that they would predict that the loser won, were they to be asked again to make the same prediction.

I quantify how much more likely is the winner to win against the loser with a three-step process. In the first step, I estimate a mean and a standard deviation for the scores' distribution in six groups of competitors: men/women without a degree/with a bachelor's degree/with a master's degree. For the predictions, I use tobit regression (lower threshold: zero matrices, upper threshold: 15 matrices), where the dependent variable is a count of the matrices solved by a competitor and the independent variables are his/her gender and academic degree. The aim of the first step is to provide an understanding for how the results of phase one could make the spectator update his/her priors and not to estimate the true expected score for each group. Hence, information that could improve the predictions but were not available to the spectators are not included in the model.

In the second step, using the predicted means and standard deviations from the first step I calculate the probability that $x_i - x_{-i} > 0$, where x_i and x_{-i} are the competitor's and the opponent's score, respectively. This probability depends on the group the competitor belongs and the group in which the opponent belongs. Let the probability that the i-th competitor wins against an opponent from the m-th of the aforementioned groups be $p(i|m)$.

The probability that the competitor would win, if phase one is repeated becomes:

$$p = \sum_{m}^{M} p_m p(i|m),$$

where p_m is the probability with which a competitor would be matched with a competitor from the m-th group. Competitors can be matched with any other competitor with an equally probability. Thus, p_m is equal to the number of potential opponents from each group divided by 45. The model predicts the following scores from which the following probabilities of winning are calculated:

- men (no degree) $E(Score) = 8.53$, $(SE = 1.03)$ & $p = 0.588$,
- women with no degree $E(Score) = 6.37$, $(SE = 1.43)$ & $p = 0.219$,
- men (bachelor's degree) $E(Score) = 10.98$, $(SE = 1.48)$ & $p = 0.893$,
- women (bachelor's degree) $E(Score) = 8.82$, (1.46) & $p = 0.627$,
- men (master's degree) $E(Score) = 10.28$, $(SE = 2.19)$ & $p = 0.782$
- women (master's degree) $E(Score) = 8.12$, $(SE = 1.77)$ & $p = 0.504$.

Based on the aforementioned probability, I calculate the odds ratio of the odds that a competitor from the winner's group wins and the odds that a competitor from the loser's group wins. A ratio higher than one indicates that the competitor who won is more likely to win than the loser, if the pre-experiment is repeated. As the ratio increases, the odds that a spectator with Rawlsian preferences gives the bonus to the winner should theoretically decrease.

The variable *Likelihood* compares the groups in which each competitor belongs and does not directly consider the score of a competitor. Moreover, the calculations involved in estimating its value are too complex to do on the fly while solving a survey. I argue that the spectators react intuitively to what estimations quantify. However, the aforementioned arguments indicate that the calculation of the variable might not capture how spectators use the information given.

As a robustness test, I also try a different measure for the likelihood of winning. Specifically, I estimate the models including the odds ratio of the odds that the winner would win and that the loser would win, if I only re-matched the competitors. The odd of winning after a re-match is equal to the number of competitors that had a lower score than the competitor divided by the number of competitors who had an equal or a higher score.[14] The re-match odds-ratio is always larger than one and is unique for each pair (min= 1.219, max= 70, median= 4.594, SD= 19.930). A large odds-ratio indicates that the loser would not be able to win even after re-matching, whereas the winner would win even after re-matching. How I measure the likelihood of winning does not influence the results of the model qualitatively. I report the model with the better fit.

Table 4.5 reports the results. Column one presents the coefficients of a reduced model with only the main variables. In the model reported in column two, the competitors' differences are added as controls. The model of column three includes the controls for how the spectator answered in the previous questions of the survey and whether he had the same gender as the winner. Column four reports the results of the full model. The number of spectators changes between model three and four because the spectators who predicted that the woman or the man always won drop out of the sample. Column five constraints the sample to the decisions for the mixed pairs. Last, column six reports the results of the full model estimated using the sample of male spectators. In this model, the variable *Gender congruency* is omitted because it is perfectly collinear with the variable *Woman won*.

Both male and female winners are more likely to get the bonus when they compete against an opponent of the same gender. The male winners' odds to get the bonus are significantly less reduced compared to the female winners' odds, when they compete against opponents with a different gender, i.e. $\beta_3 < 0$. From the results in the first four columns, it is impossible to say whether women winners are less likely to get the bonus compared to their male opponents who lost. Spectators for which $\alpha + \beta_1 < \beta_3 + \beta_2$ are

14. To avoid having an odd equal to zero for the worst competitor (and an infinite odds ratio), I add one to the numerator and denominator of the odds for all competitors.

	Logistic regression DV: i-th spectator gives the bonus to the winner of the j-th pair					
	(1)	(2)	(3)	(4)	(5)	(6)
Sample	All	All	All	All	Mixed pairs	Men
Mixed pair (β_1)	-0.915^{***}	-0.919^{***}	-1.801^{***}	-1.415^{***}		-1.749^{***}
	(0.045)	(0.058)	(0.051)	(0.052)		(0.083)
Woman won (β_2)	0.102^{*}	0.073	0.670^{***}	0.578^{***}	-0.349^{***}	0.619^{***}
	(0.054)	(0.082)	(0.047)	(0.049)	(0.052)	(0.077)
Mixed pair: woman won (β_3)	-0.286^{***}	-0.185^{**}	-0.310^{***}	-1.269^{***}		-1.286^{***}
	(0.069)	(0.088)	(0.057)	(0.064)		(0.103)
No degree vs Master's		0.234^{***}	0.450^{***}	-0.203^{***}	0.079	-0.296^{***}
		(0.074)	(0.060)	(0.053)	(0.056)	(0.079)
One level lower		0.047	0.112^{***}	-0.234^{***}	0.103^{**}	-0.302^{***}
		(0.056)	(0.035)	(0.033)	(0.046)	(0.051)
One level higher		0.070	-0.280^{**}	1.804^{***}	1.641^{***}	2.109^{***}
		(0.070)	(0.138)	(0.119)	(0.119)	(0.187)
Master's vs no degree		-0.037	-0.245^{***}	-0.106^{***}	-0.016	-0.155^{***}
		(0.055)	(0.027)	(0.027)	(0.036)	(0.042)
Both below median		-0.159^{**}	-0.269^{***}	0.181^{***}	0.136^{***}	0.203^{***}
		(0.066)	(0.042)	(0.043)	(0.038)	(0.068)
Both above median		-0.063	0.160^{***}	0.116^{***}	0.068	0.139^{***}
		(0.057)	(0.033)	(0.035)	(0.043)	(0.054)
Score difference[3]		-0.0001^{**}	0.00002	0.0003^{***}	0.0002^{***}	0.0004^{***}
		(0.0001)	(0.00003)	(0.00004)	(0.00003)	(0.0001)
Portion of previous winners rewarded			4.713^{***}	5.279^{***}	2.296^{***}	5.654^{***}
			(0.096)	(0.098)	(0.117)	(0.155)
Gender congruency			-0.033	-0.031	-0.108^{***}	
			(0.021)	(0.021)	(0.027)	
Likelihood				-0.148^{***}	-0.089^{***}	-0.164^{***}
				(0.008)	(0.007)	(0.012)
Priors				-0.001	0.002	-0.003
				(0.001)	(0.001)	(0.002)
Likelihood: priors				0.001^{*}	0.0003	0.002^{**}
				(0.001)	(0.001)	(0.001)
Observations	8,993	8,993	8,993	8,809	6,128	3,542
Spectators	391	391	391	383	383	154
Log Likelihood	−10,235	−10,224	−7,933	−7,610	−3,604	−2,899
AIC	20,475	20,467	15,891	15,250	-	-
Robust Hausman χ^2	580^{***}	612^{***}	619^{***}	646^{***}	463^{***}	233^{***}

The coefficients represent logits. The robust SE are reported in the parentheses. All models are estimated with spectator's fixed effects. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

Table 4.5: Likelihood of the winner getting the bonus

more likely to give the bonus to women who won and spectators for which $\alpha + \beta_1 > \beta_3 + \beta_2$ are less likely. Constraining the sample to the mixed pairs reveals that indeed women winners are less likely to get the bonus than men winners.

The winner's score difference and education differences do not play a robustly significant role for the decision, who should get the bonus. Contrary, previous decisions play a role. Spectators are consistent in their decision to reward the winners. Whether the spectator and the winner have the same gender does not significantly improve or worsen the odds of the winner to get the bonus. Subsequently, there is no evidence in favor of same gender favoritism and the null cannot be rejected for hypothesis 4.3. As expected, the fact that the winner comes from a group which did well in the matrix task and the loser from a group which did bad reduces the odds that the winner will get the bonus (Rawlsian preferences). Thus, the null can be rejected for hypothesis 4.2

The spectators' priors do not appear to influence the spectator's decision or modulate the effect of experience. However, further tests are necessary. Both the variable *Likelihood* and *Priors* are a function of the competitor's and opponent's genders and education. Thus, multicollinearity could bias the coefficients. One indication that multicollinearity is not significant is that modifications to the model and the sample with which the model is estimated do not significantly change the coefficients. Further, the variance inflation factors are not high. Indicatively, the *Likelihood* variable has the highest $\text{gVIF}^{1/2df} = 4.378$, which is not alarmingly large.

To control whether changing beliefs influences whether spectators taste discriminate, I estimate a model where I interact the winner's gender with the *Likelihood* (squared). Figure 4.3 is based on the results of this model. The detailed regression results along with further regressions with different interaction terms can be found in Appendix A.3.3. To estimate the values of the figure, I assume that the two competitors in a mixed pair have above median scores, the same degree and that the winner solved one matrix more than the loser. Moreover, the deciding spectator believes that both competitors are equally good ($Priors = 1$) and has never given the bonus to the winner in the past.

The red line represents the predicted probability that the female loser will get the bonus and the light blue line the same probability for a female winner. The gray areas represent the 95% predictions' intervals. For easier reading, I convert the *Likelihood* from an odds ratio to likelihoods, assuming that the likelihood that a man and a woman tied is equal to zero. The upper side of the graph is labeled to reflect the likelihood that a man won and the lower side of the graph is labeled to reflect the likelihood that a woman won.

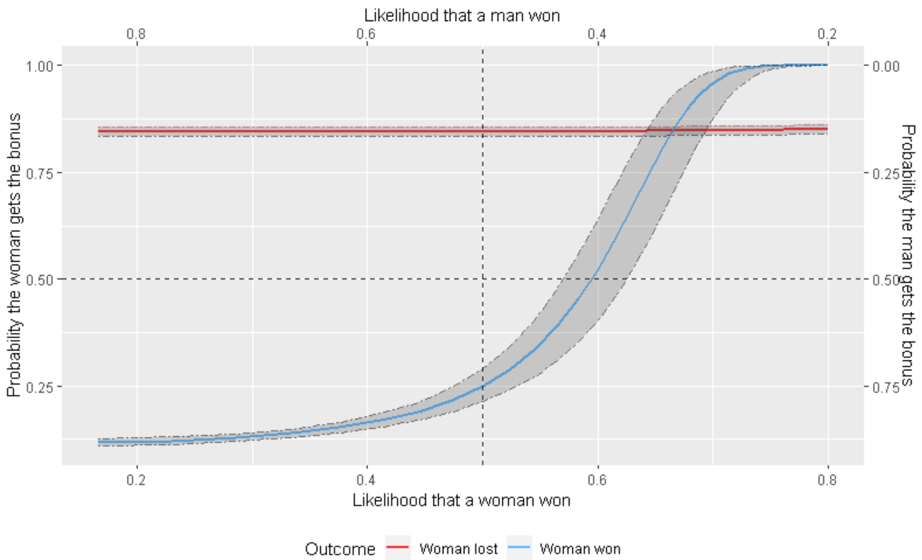


Figure 4.3: Lineplot: Predicted probability of getting the bonus in a mixed pair

Further, on the left side of the y-axis is the probability that a woman gets the bonus and on the right side is the probability that the man gets the bonus.

The dashed lines separate the areas where men are considered more likely to win and where men are considered less likely to win and the areas where men are more likely to get the bonus and less likely to get the bonus. The intersection of the dashed lines is the point where the probability of getting the bonus and the likelihood of winning are equal for men and women.

The figure shows that women are more likely to be rewarded for losing than for winning. Subsequently, the no double standard hypothesis can be rejected (Hypothesis 4.4). Specifically, men winners in mixed pairs are less likely to get the bonus compared to woman losers.

This pattern could be interpreted as a kind of affirmative action. Spectators have not seen women winning and they want to motivate them to participate and become better in the task. However, seeing women doing better than men, i.e. as the likelihood that a woman won increases, does not significantly change the probability that a man who wins will get the bonus.

An alternative explanation could be that spectators indeed gender discriminate, i.e. prefer to give the bonus to a woman rather than a man. Interestingly, women winners are the competitors least likely to get the bonus. Moreover, for spectators to reward women winners more than women losers, they need to have seen women winning almost twice as often as they have seen men winning. Subsequently, also the pro-woman taste conjecture seems to fail.

The most likely explanation based on the data is the following. Spectators enjoy rewarding women that lose more than men that lose, i.e. use a double standard for rewarding men and women, when the competition does not belong in the female domain. The female domain in this sense are the tasks which women have unequivocally proven to be better than men. This result could explain why women avoid competition in the male domain, i.e. in tasks where men are expected to be better. In particular, they expect to get less for their invested effort than their male opponents.

To sum up, this subsection shows that spectators taste discriminate. Particularly, they reward women winners less often than they reward woman losers. The model projects that the premium for losing women will disappear if the task is considered as belonging in the female domain. This subsection connects the spectator's beliefs with their tastes for the first time. The next subsection expands on this relationship.

4.4.4 Mix of discrimination

So far, the paper did not discuss whether some groups of spectators are more or less likely than others to give a certain combination of answers. This subsection fills the gap. To account for the structure of how spectators decide, I use nested multinomial regression.

Nested ML allows groups (nests) of predictions to depend on each other in a random way. Then, the probability that a spectator predicts k becomes the conditional probability that the spectator picks a nest N_k from K nests and the probability of picking prediction k from the nest's predictions. The regression takes as the dependent variable which combination of answers the i-th spectator gave for the j-th mixed pair.

Spectators submit their answers in two waves. First, they make their predictions and then given that either competitor won submit their decision about the bonus. Thus, I define six nests (groups) and allow the elasticities to differ between nests.[15] In the following list, the bullets indicate the choices in each nest.

1. Predicted correctly that the man would win:

15. I further try nesting the choice combinations depending on the spectators' predictions, i.e. define three nests. First, I assume that the spectators do not care for the gender of the winner. Thus, predict the winner and then choose whether to give the bonus to the winner or the loser. Second, I assume that the spectators do not care who won. Thus, after predicting they decide whether to give the bonus to the man or the woman. A likelihood ratio test shows that the six nests specification fits the data better.

- The man will win and the winning man should get the bonus
- The man will win and the losing woman should get the bonus

2. Predicted wrongly that the man would win:
   - The man will win and the winning woman should get the bonus
   - The man will win and the losing man should get the bonus
3. Predicted wrongly that the woman would win:
   - The woman will win and the winning man should get the bonus
   - The woman will win and the losing woman should get the bonus
4. Predicted correctly that the woman would win:
   - The woman will win and the winning woman should get the bonus
   - The woman will win and the losing man should get the bonus
5. Predicted that no one would win and the man won:
   - They will tie and the winning man should get the bonus
   - They will tie and the losing woman should get the bonus
6. Predicted that no one would win and the woman won:
   - They will tie and the winning woman should get the bonus
   - They will tie and the losing man should get the bonus

The model contains dummies for the spectator's characteristics and whether he/she said that women are better in the task than men. Specifically, there is a dummy for whether a spectator is a woman, a dummy whether he/she participates in a Bachelor's or Master's program or has a Master's already (*Basic Studies*) and age group dummies. Last, I include one dummy for whether the competitors have the same degree.

Table 4.6 reports the results of the model. The sample is limited to the mixed pairs, in which it is impossible for the spectator to match the competitors of part one to the competitors of part three and thus the spectators do not know whether they predicted correctly. Further, I drop the participants who at either part of the survey did not vary their answers. The last column of the table reports the portion of the 5,505 choice combinations of the model which corresponds to each choice combination. The reference choice combination is when the spectator predicts that no one won and then gives the bonus to the winning woman.

Nested multinomial logistic regression DV: Combination of choices of the i-th spectator for the j-th pair

Who gets the bonus?	Constant	Woman	Basic studies	21-24	25-28	>28	Same degree	$\hat{\mu_M} < \hat{\mu_W}$	Freq.
Predicts that the man won									
Man wins									
Man	17.638**	3.699	-5.301**	3.471**	4.253**	3.024	1.091	-2.24**	0.056
	(7.212)	(2.299)	(2.198)	(1.703)	(2.125)	(1.894)	(1.213)	(0.995)	
Woman	13.257**	4.832**	-3.501*	3.107*	4.54**	4.256**	-1.326	-1.287	0.108
	(6.669)	(2.284)	(2.083)	(1.618)	(2.059)	(1.801)	(0.826)	(0.933)	
Woman wins									
Man	11.293	4.338*	-4.736**	4.149**	4.309**	4.121**	-0.014	-1.755*	0.162
	(6.985)	(2.281)	(2.195)	(1.965)	(2.117)	(1.958)	(0.763)	(0.934)	
Woman	2.141	4.624**	-1.153	0.17	4.188*	2.047	-0.192	-1.475	0.054
	(12.95)	(2.33)	(3.548)	(3.56)	(2.348)	(2.652)	(0.974)	(1.086)	
Predicts that the woman won									
Man wins									
Man	3.460	5.480**	-4.047*	2.094	3.514	2.271	-7.684	-0.949	0.055
	(8.564)	(2.31)	(2.165)	(1.774)	(2.186)	(1.968)	(4.855)	(1.038)	
Woman	17.167**	5.263**	-4.075*	3.206*	4.389**	3.94**	4.159**	-0.963	0.138
	(7.186)	(2.292)	(2.114)	(1.678)	(2.112)	(1.871)	(1.906)	(1.023)	
Woman wins									
Man	8.592	4.055*	-3.813*	3.659**	4.662**	3.901**	3.583**	-3.019***	0.151
	(7.075)	(2.28)	(2.088)	(1.64)	(2.078)	(1.814)	(1.688)	(0.934)	
Woman	43.455**	2.279	-3.121	3.043	3.924	1.363	-8.311*	-6.476***	0.053
	(17.011)	(2.447)	(2.405)	(2.133)	(2.526)	(2.413)	(4.406)	(1.515)	
Predicts that no one won									
Man wins									
Man	4082.384	152.081	39.298	-14.016	-36.197	-8.907	9.394	11.418	0.033
	(27448.743)	(529.671)	(140.655)	(61.138)	(134.796)	(48.24)	(30.268)	(46.041)	
Woman	1349.143	-60.42	-22.598	11.032	22.638	9.569	-3.19	-7.95	0.075
	(8716.14)	(232.87)	(61.692)	(26.89)	(59.144)	(21.31)	(13.337)	(20.179)	
Woman wins									
Man	5.525	7.304**	-4.91*	4.957**	7.067**	5.956**	0.291	-2.83**	0.081
	(4.73)	(3.253)	(2.635)	(2.489)	(3.052)	(2.698)	(1.005)	(1.265)	

The results are based on 5,505 obervations, i.e. the answers of the spectators who varied both their predictions and whom they gave the bonus ($n = 367$) and the mixed pairs, where women had the same or a higher degree than their opponent ($m = 15$). The combinations were nested according to the spectator's predictions and for each prediction depending whether the woman won or lost in the j-th pair (pseudo-$R^2 = 0.028$). The coefficients are logits. The reference category are the predictions, where the spectator predicts a tie, a woman wins and the spectator gives her the bonus. The standard errors are reported in the parentheses. *p<0.1; **p<0.05; ***p<0.01.

Table 4.6: Likelihood of a combination of answers

When spectators predict that the woman won, they are more likely to give the bonus to the woman than the man, independent of who won. Moreover, they are more likely to give the bonus to the winning woman rather than the losing woman. Interestingly, when women have the same degree as their opponents, spectators are more likely to give her the bonus, if she loses rather than if she wins.

Women are more likely to predict a tie. Additionally, the regression shows that when they do, they are more likely to give the bonus to the losing man rather than the winning woman. Contrary, when the woman loses, they are as likely to give the bonus to the pair's woman or man.

Spectators in their basic studies are as likely to give the bonus to the winning man or the losing winner. However, they are more likely to give the bonus to a winning woman rather than a losing man. Older spectators tend to favor losing men rather than winning women, independent of what they predict. Last, those that predict that women had a higher average are less likely to predict that the woman won and when they do, they are more likely to give the bonus to the man, if she wins. In other words, they both taste and statistically discriminate against women.

To sum up, the results suggest that the spectators who believe that women can win are reluctant to reward them when they do. Although, women spectators are the most likely to predict that a woman won, the biggest chunk of this group gives the bonus to the man independent of the outcome.

4.5 Conclusion

Stopping gender discrimination and paving the way to gender equality has been on the list of things to achieve for many countries the last years. However, we know very little about how different sources of discrimination interact. This chapter reported the results of an experiment designed to disentangle taste and statistical discrimination.

In an online survey, participants were incentivized to identify the set of winners in competitions. Then, once informed about the outcomes of the competitions the same spectators had to give boni to the competitors. The experiment's participants do not appear to statistically discriminate. However, they taste discriminate. Specifically, they weigh women's academic achievements differently compared to men's academic achievements. Further, they reward men for winning and women for losing against men. These finding are alarming. If the incentives described for women, i.e. to do less in order to get more, are not taken care of, gender discrimination cannot be eradicated.

One could protest that the results of the experiment are not representative. I would agree. Students in Germany are indoctrinated to avoid gender discrimination and the perils of gender discrimination are known to them. An increasing number of universities fights for an equal distribution of male and female students and teaching staff. Gender studies chairs arise and funds and scholarships are created to close the gap -if any exists in a field- between men and women. Running the experiment on the general population in a society where women are less integrated in the social, political and economic life would probably show that the results discussed underestimate the threat double standards pose for gender equality.

This paper does its best to dampen discrimination. I use a gender-neutral task to ensure that the experimental design will not produce discrimination. I run the experiment on a sub-group of the population which arguably would be expected to have transcended traditional gender stereotypes and sexism. Nevertheless, I still find significant discrimination.

References

Aigner, Dennis J, and Glen G Cain. 1977. Statistical Theories of Discrimination in Labor Markets. *Industrial and Labor Relations Review* 30 (2): 175–187.

Akerlof, George A. 1985. Discriminatory, Status-based Wages among Tradition-oriented, Stochastically Trading Coconut Producers. *Journal of Political Economy* 93 (2): 265–276.

Akerlof, George A, and Rachel E Kranton. 2010. *Identity Economics.* 184. Princeton, NJ: Princeton University Press.

Azmat, Ghazala, and Barbara Petrongolo. 2014. Gender and the labor market: What have we learned from field and lab experiments? *Labour Economics* 30:32–40.

Balliet, Daniel, Junhui Wu and Carsten K W De Dreu. 2014. Ingroup Favoritism in Cooperation: A Meta-Analysis. *Psychological Bulletin,* no. 6: 1556–1581.

Becker, Gary S. 1957. *The Economics of Discrimination.* 1rst. 167. Chicago, IL: University of Chicago Press.

———. 1985. Human Capital, Effort, and Sexual Division of Labor. *Journal of Labor Economics* 3 (1): 33–58.

Becker, Sascha O, Ana Fernandes and Doris Weichselbaumer. 2019. Discrimination in hiring based on potential and realized fertility: Evidence from a large-scale field experiment. *Labour Economics* 59:139–152.

Bertrand, Marianne, Dolly Chugh, Sendhil Mullainathan, Dan Black, David Neumark, Shelly Lundberg and Kerwin Charles. 2005. Implicit discrimination. *American Economic Review* 95 (2): 94–98.

Bertrand, Marianne, and Esther Duflo. 2017. Field Experiments on Discrimination. Chap. 8 in *Handbook of field experiments,* edited by Vinayak Banerjee Abhijit and Esther Duflo, 309–393. Amsterdam: North Holland.

Bhat, Chandra R. 1995. A Heteroscedastic Extreme Value Model of Intercity Mode Choice. *Transportation Research Part B: Methodological* 29 (6): 471–483.

Bock, Olaf, Andreas Nicklisch and Ingmar Baetge. 2012. *hroot: Hamburg registration and organization online tool.* WiSo-HH.

Bursztyn, Leonardo, Thomas Fujiwara and Amanda Pallais. 2017. Acting wife: Marriage market incentives and labor market investments. *American Economic Review* 107 (11): 3288–3319.

Carlana, Michela. 2019. Implicit Stereotypes: Evidence from Teachers' Gender Bias. *Quarterly Journal of Economics* 134 (3): 1163–1224.

Chamberlain, Gary. 1980. Analysis of Covariance with Qualitative Data. *The Review of Economic Studies* 47 (1): 225.

Coffman, Katherine B, Christine L Exley and Muriel Niederle. 2020. The role of beliefs in driving gender discrimination. *Harvard Business School Working Paper* 18-054:60.

Ditonto, Tessa. 2017. A High Bar or a Double Standard? Gender, Competence, and Information in Political Campaigns. *Political Behavior* 39 (2): 301–325.

Everett, Jim AC, Nadira S Faber and Molly Crockett. 2015. Preferences and beliefs in ingroup favoritism. *Frontiers in Behavioral Neuroscience* 9 (FEB): 1–21.

Glover, Dylan, Amanda Pallais and William Pariente. 2017. Discrimination as a Self-Fullfilling Prophecy: Evidence from French Grocery Stores. *Quarterly Journal of Economics* 132 (3): 1219–1260.

Gower, John C. 1971. A General Coefficient of Similarity and Some of Its Properties. *Biometrics* 27 (4): 857–871.

Guiso, Luigi, Ferdinando Monte, Paola Sapienza and Luigi Zingales. 2008. Culture, Gender, and Math. *Science* 320 (May): 1164–1165.

Gupta, Vishal K, Sandra C Mortal, Sabatino Silveri, Minxing Sun and Daniel B Turban. 2020. You're Fired! Gender Disparities in CEO Dismissal. *Journal of Management* 46 (4): 560–582.

Hamermesh, Daniel S, and Jeff E Biddle. 1994. Beauty and the Labor Market. *American Economic Review* 84 (5): 1174–1194.

Kanze, Hana, Laura Huange, Mark A Conley and Tory A Higgins. 2018. We ask men to win and women not to lose: Closing the gender gap in startup funding. *Academy of Management Journal* 61 (2): 586–614.

Klumpp, Tilman, and Xuejuan Su. 2013. A theory of perceived discrimination. *Economic Theory* 53 (1): 153–180.

Lane, Tom. 2016. Discrimination in the laboratory: A meta-analysis of economics experiments. *European Economic Review* 90:375–402.

LimeSurvey GmbH. *Limesurvey: An Open Source survey tool.* Hamburg, Germany. `http://www.limesurvey.org`.

Mazar, Nina, On Amir and Dan Ariely. 2008. The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research* 45 (6): 633–644.

Neilson, William, and Shanshan Ying. 2016. From taste-based to statistical discrimination. *Journal of Economic Behavior and Organization* 129:116–128.

Neumark, David. 2018. Experimental research on labor market discrimination. *Journal of Economic Literature* 56 (3): 799–866.

Nosek, Brian A, Frederick L Smyth, N Sriram, Nicole M Lindner, Thierry Devos, Alfonso Ayala, Yoav Bar-Anan et al. 2009. National differences in gender–science stereotypes predict national sex differences in science and math achievement. *Proceedings of the National Academy of Sciences* 106 (26): 10593–10597.

Rawls, John. 1971. *A Theory of Justice.* Revised. Cambridge, MA: Harvard University Press.

Riach, Peter A., and Judith Rich. 2002. Field experiments of discrimination in the market place. *Economic Journal* 112 (483): F480–F518.

Sarsons, Heather. 2017. Recognition for group work: Gender differences in academia. *American Economic Review* 107 (5): 141–145.

Schirmer, Hendrik. 2017. *So leben Studierende in Hamburg. Zur wirtschaftlichen und sozialen Lage der Studierenden in Hamburg 2016. Online-Befragung an Hamburger Hochschulen.* Technical report. Hamburg, Germany: Studierendenwerk Hamburg.

Schwab, Stewart. 1986. Is Statistical Discrimination Efficient? *American Economic Review* 76 (1): 228–234.

Spencer, Steven J, Christine Logel and Paul G Davies. 2016. Stereotype threat. *Annual Review of Psychology* 67:415–437.

Spencer, Steven J, Claude M Steele and Diane M Quinn. 1999. Stereotype Threat and Women's Math Performance. *Journal of Experimental Social Psychology* 35 (1): 4–28.

Steele, Claude M, and Joshua Aronson. 1995. Stereotype Threat and the Intellectual Test Performance of African American. *Journal of personality and social psychology* 69 (5): 797–811.

Train, Kenneth E. 2009. *Discrete Choice Methods with Simulation.* 388. New York, NY: Cambridge University Press.

List of cited books

Acemoglu, Daron, and James A Robinson. 2012. *Why Nations Fail: The Origins of Power, Prosperity, and Poverty.* 2nd. London: Profile Books LTD.

Akerlof, George A, and Rachel E Kranton. 2010. *Identity Economics.* 184. Princeton, NJ: Princeton University Press.

Albert, Richard. 2019. *Constitutional Amendments: Making, Breaking, and Changing Constitutions.* 344. New York, NY: Oxford University Press.

Axelrod, Robert. 1984. *The Evolution of Cooperation.* 241. New York, NY: Basic Books.

Beard, Charles A. 1913. *An Economic Interpretation of the Constitution of the United States.* New York, NY: Macmillan.

Becker, Gary S. 1957. *The Economics of Discrimination.* 1rst. 167. Chicago, IL: University of Chicago Press.

Bicchieri, Cristina. 2006. *The grammar of society: the nature and dynamics of social norms.* 278. Cambridge: Cambridge University Press.

Carnes, Nicholas. 2018. *The Cash Ceiling: Why only the rich run for office–and What we can do about it.* 322. Princeton, NJ: Princeton University Press.

Chen, Shu-Heng. 2016. *Agent-based computational economics: How the idea originated and where it is going.* 528. London: Routledge.

Elkins, Zachary, Tom Ginsburg and James Melton. 2009. *The Endurance of National Constitutions.* 260. Cambridge, MA: Cambridge University Press.

Harsanyi, John C, and Reinhard Selten. 1988. *A general theory of equilibrium selection in games.* Cambridge, MA: MIT Press.

Kahneman, Daniel. 2011. *Thinking Fast and Slow.* 499. New York, NY: Farrar, Straus / Giroux.

Laver, Michael, and Ernest Sergenti. 2012. *Party Competition: An agent-based model.* 293. Princeton, NJ: Princeton University Press.

North, Douglass C. 1990. *Institutions, Institutional Change and Economic Performance.* 164. Cambridge, MA: Cambridge University Press.

Rawls, John. 1971. *A Theory of Justice.* Revised. Cambridge, MA: Harvard University Press.

Roznai, Yaniv. 2017b. *Unconstitutional Constitutional Amendments:The Limits of Amendment Powers.* 368. Oxford: Oxford University Press.

Schelling, Thomas C. 1960. *The Strategy of Conflict.* 322. Cambridge MA: Harvard University Press.

Tarrow, Sidney G. 1994. *Power in Movement.* 328. New York, NY: Cambridge University Press.

Train, Kenneth E. 2009. *Discrete Choice Methods with Simulation.* 388. New York, NY: Cambridge University Press.

Voigt, Stefan. 2019. *Institutional Economics An introduction.* 293. Cambridge: Cambridge University Press.

———. 2020a. *Constitutional economics: a primer.* 150. Cambridge: Cambridge University Press.

Von Neumann, John. 1966. *Theory of self-reproducing automata.* 416. Chicago, IL: University of Illinois Press.

Wintrobe, Ronald. 2000. *The Political Economy of Dictatorship.* Cambridge MA: Cambridge University Press.

Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data.* 752. Cambridge, MA: MIT press.

List of cited articles and unpublished papers

Acemoglu, Daron, Simon Johnson and James A Robinson. 2002. Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution. *Quarterly Journal of Economics* 117 (4): 1231–1294.

Acemoglu, Daron, and James A Robinson. 2006. De Facto Political Power and Institutional Persistence. *American Economic Review* 96 (2): 325–330.

Aidt, Toke S, and Raphaël Franck. 2013. How to get the snowball rolling and extend the franchise: Voting on the Great Reform Act of 1832. *Public Choice* 155 (3-4): 229–250.

———. 2015. Democratization Under the Threat of Revolution: Evidence From the Great Reform Act of 1832. *Econometrica* 83 (2): 505–547.

Aidt, Toke S, and Peter S Jensen. 2014. Workers of the world, unite! Franchise extensions and the threat of revolution in Europe, 1820–1938. *European Economic Review* 72:52–75.

Aidt, Toke S, and Gabriel Leon. 2016. The Democratic Window of Opportunity: Evidence from Riots in Sub-Saharan Africa. *Journal of Conflict Resolution* 60 (4): 694–717.

Aigner, Dennis J, and Glen G Cain. 1977. Statistical Theories of Discrimination in Labor Markets. *Industrial and Labor Relations Review* 30 (2): 175–187.

Akerlof, George A. 1985. Discriminatory, Status-based Wages among Tradition-oriented, Stochastically Trading Coconut Producers. *Journal of Political Economy* 93 (2): 265–276.

Albert, Richard. 2013. The Expressive Function of Constitutional Amendment Rules. *McGill Law Journal* 59 (2): 225–281.

———. 2015. How Unwritten Constitutional Norms Change Written Constitutions. *Dublin University Law Journal* 38 (2): 387–417.

Alós-Ferrer, Carlos, Michele Garagnani and Sabine Hügelschäfer. 2016. Cognitive reflection, decision biases, and response times. *Frontiers in Psychology* 7 (1402): 1–21.

Arellano, Manuel. 1993. On the testing of correlated effects with panel data. *Journal of Econometrics* 59 (1-2): 87–97.

Aumann, Robert J. 1976. Agreeing to Disagree. *Annals of Statistics* 4 (6): 1236–1239.

Axelrod, Robert. 1986. An evolutionary approach to norms. *American Political Science Review* 80 (4): 1095–1111.

Azmat, Ghazala, and Barbara Petrongolo. 2014. Gender and the labor market: What have we learned from field and lab experiments? *Labour Economics* 30:32–40.

Balliet, Daniel, Junhui Wu and Carsten K W De Dreu. 2014. Ingroup Favoritism in Cooperation: A Meta-Analysis. *Psychological Bulletin,* no. 6: 1556–1581.

Bardsley, Nicholas. 2000. Control without deception: Individual behaviour in free-riding experiments revisited. *Experimental Economics* 3 (2000): 215–240.

Becker, Gary S. 1985. Human Capital, Effort, and Sexual Division of Labor. *Journal of Labor Economics* 3 (1): 33–58.

Becker, Sascha O, Ana Fernandes and Doris Weichselbaumer. 2019. Discrimination in hiring based on potential and realized fertility: Evidence from a large-scale field experiment. *Labour Economics* 59:139–152.

Bertrand, Marianne, Dolly Chugh, Sendhil Mullainathan, Dan Black, David Neumark, Shelly Lundberg and Kerwin Charles. 2005. Implicit discrimination. *American Economic Review* 95 (2): 94–98.

Bhat, Chandra R. 1995. A Heteroscedastic Extreme Value Model of Intercity Mode Choice. *Transportation Research Part B: Methodological* 29 (6): 471–483.

Bjørnskov, Christian, and Stefan Voigt. 2018. The architecture of emergency constitutions. *International Journal of Constitutional Law* 16 (1): 101–127.

Blume, Lorenz, Jens Müller and Stefan Voigt. 2009. The economic effects of direct democracy-a first global assessment. *Public Choice* 140 (3-4): 431–461.

Bolton, Gary E, and Axel Ockenfels. 2000. ERC : A Theory of Equity, Reciprocity, and Competition. *American Economic Review* 90 (1): 166–193.

Bonabeau, Eric. 2002. Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences of the United States of America* 99 (SUPPL. 3): 7280–7287.

Bursztyn, Leonardo, Thomas Fujiwara and Amanda Pallais. 2017. Acting wife: Marriage market incentives and labor market investments. *American Economic Review* 107 (11): 3288–3319.

Caplin, Andrew, and Daniel Martin. 2016. The dual-process drift diffusion model: Evidence from response times. *Economic Inquiry* 54 (2): 1274–1282.

Carlana, Michela. 2019. Implicit Stereotypes: Evidence from Teachers' Gender Bias. *Quarterly Journal of Economics* 134 (3): 1163–1224.

Cason, Timothy N, and Vai Lam Mui. 2007. Communication and coordination in the laboratory collective resistance game. *Experimental Economics* 10 (3): 251–267.

———. 2013. Coordinating resistance through communication and repeated interaction. *Economic Journal* 124 (574): 226–256.

———. 2015a. Individual versus group play in the repeated coordinated resistance game. *Journal of Experimental Political Science* 2 (1): 94–106.

———. 2015b. Rich communication, social motivations, and coordinated resistance against divide-and-conquer: A laboratory investigation. *European Journal of Political Economy* 37:146–159.

Chamberlain, Gary. 1980. Analysis of Covariance with Qualitative Data. *The Review of Economic Studies* 47 (1): 225.

Chen, Fadong, and Urs Fischbacher. 2020. Cognitive processes underlying distributional preferences: A response time study. *Experimental Economics* 23:421–446.

Chen, Shu-heng, and Ragupathy Venkatachalam. 2017. Agent-based modelling as a foundation for big data. *Journal of Economic Methodology* 24 (4): 362–383.

Chong, Dennis, Mingxing Liu and Qi Zhang. 2016. Collective resistance under authoritarianism: Elite–mass strategies in an experimental Game. *Political Behavior* 38 (4): 951–976.

Coffman, Katherine B, Christine L Exley and Muriel Niederle. 2020. The role of beliefs in driving gender discrimination. *Harvard Business School Working Paper* 18-054:60.

Cornelißen, Thomas, and Katja Sonderhof. 2009. Partial effects in probit and logit models with a triple dummy-variable interaction term. *Stata Journal* 9 (4): 571–583.

Dal Bó, Ernesto, Frederico Finan, Olle Folke, Torsten Persson and Johanna Rickne. 2017. Who Becomes a Politician? *Quarterly Journal of Economics* 132 (4): 1877–1914.

Dal Bó, Pedro, Andrew Foster and Louis Putterman. 2010. Institutions and behavior: Experimental evidence on the effects of democracy. *American Economic Review* 100 (5): 2205–2229.

Denes, Matthew R., Jonathan M. Karpoff and Victoria B. McWilliams. 2017. Thirty years of shareholder activism: A survey of empirical research. *Journal of Corporate Finance* 44:405–424.

Ditonto, Tessa. 2017. A High Bar or a Double Standard? Gender, Competence, and Information in Political Campaigns. *Political Behavior* 39 (2): 301–325.

Elster, Jon. 1995. Forces and mechanisms in the constitution-making process. *Duke Law Journal* 45 (2): 364–396.

Engelmann, Dirk, and Martin Strobel. 2004. Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments. *American Economic Review* 94 (4): 857–869.

Everett, Jim AC, Nadira S Faber and Molly Crockett. 2015. Preferences and beliefs in ingroup favoritism. *Frontiers in Behavioral Neuroscience* 9 (FEB): 1–21.

Fehr, Ernst, and Klaus M Schmidt. 1999. A theory of fairness, competition and cooperation. *Quarterly Journal of Economics* 114 (3): 817–868.

Gavison, Ruth. 2002. What Belongs in a Constitution? *Constitutional Political Economy* 13 (1): 89–105.

Giglioni, Mateo Fabio Nels. 2020. Plutocratic leadership in the electoral arena: three Mitteleuropean cases of personal wealth in politics. *Comparative European Politics* 18:309–329.

Gilbert, Michael D, Mauricio Guim and Michael Weisbuch. 2019. Constitutional Locks. *Virginia Public Law and Legal Theory Research Paper* 2019-50.

Ginsburg, Tom, Zachary Elkins and Justin Blount. 2009. Does the Process of Constitution-Making Matter? *Annual Review of Law and Social Science* 5:201–223.

Ginsburg, Tom, and James Melton. 2015. Does the Constitutional Amendment Rule Matter at All? Amendment Cultures and the Challenges of Measuring Amendment Difficulty. *International Journal of Constitutional Law* 13 (3): 686–713.

Glover, Dylan, Amanda Pallais and William Pariente. 2017. Discrimination as a Self-Fullfilling Prophecy: Evidence from French Grocery Stores. *Quarterly Journal of Economics* 132 (3): 1219–1260.

Gower, John C. 1971. A General Coefficient of Similarity and Some of Its Properties. *Biometrics* 27 (4): 857–871.

Greene, William. 2010. Testing hypotheses about interaction terms in nonlinear models. *Economics Letters* 107 (2): 291–296.

Griffin, John D, Brian Newman and Patrick Buhr. 2020. Class War in the Voting Booth: Bias Against High-Income Congressional Candidates. *Legislative Studies Quarterly* 45 (1): 131–145.

Grossman, Guy, and Delia Baldassarri. 2012. The Impact of Elections on Cooperation: Evidence from a Lab-in-the-Field Experiment in Uganda. *American Journal of Political Science* 56 (4): 964–985.

Guiso, Luigi, Ferdinando Monte, Paola Sapienza and Luigi Zingales. 2008. Culture, Gender, and Math. *Science* 320 (May): 1164–1165.

Gupta, Vishal K, Sandra C Mortal, Sabatino Silveri, Minxing Sun and Daniel B Turban. 2020. You're Fired! Gender Disparities in CEO Dismissal. *Journal of Management* 46 (4): 560–582.

Hamermesh, Daniel S, and Jeff E Biddle. 1994. Beauty and the Labor Market. *American Economic Review* 84 (5): 1174–1194.

Hein, Michael. 2019. Impeding constitutional amendments: why are entrenchment clauses codified in contemporary constitutions? *Acta Politica* 54 (2): 196–224.

———. 2020. Do constitutional entrenchment clauses matter? Constitutional review of constitutional amendments in Europe. *International Journal Of Constitutional Law* 18 (1): 78–110.

Helton, William S, and Paul N Russell. 2015. Rest is best: The role of rest and task interruptions on vigilance. *Cognition* 134:165–173.

Holland, John H, and John H Miller. 1991. Artificial Adaptive Agents in Economic Theory. *American Economic Review* 81 (2): 365–370.

Kahneman, Daniel. 2003. Maps of bounded rationality: Psychology for behavioral economics. *American Economic Review* 93 (5): 1449–1475.

Kanze, Hana, Laura Huange, Mark A Conley and Tory A Higgins. 2018. We ask men to win and women not to lose: Closing the gender gap in startup funding. *Academy of Management Journal* 61 (2): 586–614.

Klumpp, Tilman, and Xuejuan Su. 2013. A theory of perceived discrimination. *Economic Theory* 53 (1): 153–180.

Kollman, Ken, John H Miller and Scott E Page. 1992. Adaptive Parties in Spatial Elections. *American Political Science Review* 86 (4): 929–937.

———. 1997. Political Institutions and Sorting in a Tiebout Model. *American Economic Review* 87 (5): 977–992.

Kool, Wouter, Joseph T McGuire, Zev B Rosen and Matthew M Botvinick. 2010. Decision making and the avoidance of cogntive demand. *Journal of Experimental Psychology: General* 139 (4): 665–682.

Landau, David, and Rosalind Dixon. 2015. Constraining Constitutional Change. *Wake Forest Law Review* 50 (4): 859–890.

Lane, Tom. 2016. Discrimination in the laboratory: A meta-analysis of economics experiments. *European Economic Review* 90:375–402.

Lehtinen, Aki, and Jaakko Kuorikoski. 2007. Computing the perfect model: Why do economists shun simulation? *Philosophy of Science* 74 (3): 304–329.

Mazar, Nina, On Amir and Dan Ariely. 2008. The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research* 45 (6): 633–644.

McGuire, Robert A, and Robert L Ohsfeldt. 1989. Self-interest, Agency Theory, and Political Voting Behavior: The Ratification of the United States Constitution. *American Economic Review* 79 (1): 219–234.

Mullen, Elizabeth, and Benoît Monin. 2016. Consistency versus licensing effects of past moral behavior. *Annual Review of Psychology* 67 (1): 363–385.

Neilson, William, and Shanshan Ying. 2016. From taste-based to statistical discrimination. *Journal of Economic Behavior and Organization* 129:116–128.

Neumark, David. 2018. Experimental research on labor market discrimination. *Journal of Economic Literature* 56 (3): 799–866.

North, Douglass C, and Barry R Weingast. 1989. Constitutions and commitment: The evolution of institutions governing public choice in seventeenth-century England. *Journal of Economic History* 49 (4): 803–832.

Norton, Edward C, Hua Wang and Chunrong Ai. 2004. Computing interaction effects and standard errors in logit and probit models. *The Stata Journal: Promoting communications on statistics and Stata* 4 (2): 154–167.

Nosek, Brian A, Frederick L Smyth, N Sriram, Nicole M Lindner, Thierry Devos, Alfonso Ayala, Yoav Bar-Anan et al. 2009. National differences in gender–science stereotypes predict national sex differences in science and math achievement. *Proceedings of the National Academy of Sciences* 106 (26): 10593–10597.

Ostrom, Elinor. 1986. An agenda for the study of institutions. *Public Choice* 48 (1): 3–25.

Pool, Ithiel de Sola, and Allan Kessler. 1965. The Kaiser, the Tsar, and the computer: Information processing in a crisis. *American Behavioral Scientist* 8 (9): 31–38.

Posner, Eric A, Kathryn E Spier and Adrian Vermeule. 2010. Divide and Conquer. *Journal of Legal Analysis* 2 (2): 417–471.

Rand, David G. 2016. Cooperation, fast and slow: Meta-analytic evidence for a theory of social heuristics and self-interested deliberation. *Psychological Science* 27 (9): 1192–1206.

Rand, David G, Joshua D Greene and Martin A Nowak. 2012. Spontaneous giving and calculated greed. *Nature* 489 (7416): 427–430.

Riach, Peter A., and Judith Rich. 2002. Field experiments of discrimination in the market place. *Economic Journal* 112 (483): F480–F518.

Roznai, Yaniv. 2015. Unamendability and the Genetic Code of the Constitution. *European Review of Public Law* 27 (2): 775–825.

———. 2017a. Negotiating the eternal: The paradox of entrenching secularism in constitutions. *Michigan State Law Review* 253:254–331.

Sarsons, Heather. 2017. Recognition for group work: Gender differences in academia. *American Economic Review* 107 (5): 141–145.

Schelling, Thomas C. 1969. Models of Segregation. *American Economic Review* 59 (2): 488–493.

Schwab, Stewart. 1986. Is Statistical Discrimination Efficient? *American Economic Review* 76 (1): 228–234.

Schwartz, Alex. 2019. An Agent-Based Model of Judicial Power. *Journal of Law* 9 (1): 21–53.

Spencer, Steven J, Christine Logel and Paul G Davies. 2016. Stereotype threat. *Annual Review of Psychology* 67:415–437.

Spencer, Steven J, Claude M Steele and Diane M Quinn. 1999. Stereotype Threat and Women's Math Performance. *Journal of Experimental Social Psychology* 35 (1): 4–28.

Stahl, Dale O, and Paul W Wilson. 1995. On player's models of other players: Theory and experimental evidence. *Games and Economic Behavior* 10 (1): 218–254.

Steele, Claude M, and Joshua Aronson. 1995. Stereotype Threat and the Intellectual Test Performance of African American. *Journal of personality and social psychology* 69 (5): 797–811.

Tsebelis, George. 2017. The time inconsistency of long constitutions: Evidence from the world. *European Journal of Political Research* 56 (4): 820–845.

Versteeg, Mila, and Emily Zackin. 2016. Constitutions Un-Entrenched: Toward an Alternative Theory of Constitutional Design. *American Political Science Review* 110 (4): 657–674.

Voigt, Stefan. 1999. Implicit Consitutional Change-Changing the Meaning of the Constitution Without Changing the Text of the Document. *European Journal of Law and Economics* 7 (3): 197–224.

———. 2018. How to measure informal institutions. *Journal of Institutional Economics* 14 (1): 1–22.

———. 2020b. Mind the Gap – Analyzing the Divergence between Constitutional Text and Constitutional Reality. *ILE Working Paper Series* 32.

Weingast, Barry R. 1995. The economic role of political institutions: Market-preserving federalism and economic development. *Journal of Law, Economics, and Organization* 11 (1): 1–31.

———. 1997. The Political Foundations of Democracy and the Rule of Law. *American Political Science Review* 91 (2): 245–263.

Woolcock, Michael. 1998. *Theory and Society* 27 (2): 151–208.

Wooldridge, Jeffrey M. 2019. Correlated random effects models with unbalanced panels. *Journal of Econometrics* 211 (1): 137–150.

List of software, databases and reports used

Bock, Olaf, Andreas Nicklisch and Ingmar Baetge. 2012. *hroot: Hamburg registration and organization online tool.* WiSo-HH.

Elkins, Zackary, Tom Ginsburg and James Melton. 2020. *Chronology of Constitutional Events, Version 1.3.* Accessed 28 April. https://comparativeconstitutionsproject.org.

Fischbacher, Urs. 2007. *Z-Tree: Zurich toolbox for ready-made economic experiments.*

Hein, Michael. 2018. *Constitutional Entrenchment Clauses Dataset.* Göttingen. Accessed 18 May 2020. http://data.michaelhein.de.

LimeSurvey GmbH. *Limesurvey: An Open Source survey tool.* Hamburg, Germany. http://www.limesurvey.org.

Appendix A

Appendices

A.1 Chapter 2

A.1.1 Instructions for the participants (translation)

General instructions for the participants

Welcome to the Experimental Laboratory, you are now taking part in an economic experiment. The duration of this experiment is about 90 minutes. In this experiment you can earn - depending on your decisions - a not-inconsiderable amount of money. It is therefore important that you read these instructions carefully.

For the duration of the experiment, we ask you to comply to certain rules: From now on, there is an absolute ban on communication. If you have questions, please ask us alone. To do this, notify us and we will come to your cubicle. Failure to observe this rule will result in immediate exclusion from the experiment and all payments. We also ask you to put in silent mode and pack away your cell phones and all other technical equipment. We do not want you or other participants to be disturbed or distracted by them. We look forward to your participation and hope for your cooperation.

The experiment consists of two parts. We will distribute the instructions for each part and read them out loud. Thereby we ensure that all participants are informed at least once about the structure of the experiment. If you have questions, notify us from your cubicle. Questions do not mean that you are slow of comprehension, but that we have formulated something incomprehensibly. Please do no hesitate to ask.

Anonymity The experiment is anonymous. You will not be asked for your name or any information that will allow the lab team to identify you. In the experiment, we also ask about your socio-demographic characteristics (age, gender), but not for your name.

Afterwards the pay-offs are made. We will call you individually based on your location numbers for the pay-offs. Just one of the experimenters and you will learn what you earned. Therefore, all decisions in the experiment and the pay-offs at the end remain anonymous.

Pay-offs The size of your pay-offs is calculated based on your decisions, the decision of the other participants and chance. In the experiment we are not talking about Euro, but about ECU (Experimental Currency Unit). Your earnings from the experiment are calculated in ECU. The earnings are then converted from ECU to the Euro at the end of the experiment, where:

- 10 (ten) ECUs equals 1 (one) Euro

The first part of the experiment

Before the second part begins, you vote whether you are in favor of Option A or Option B.

- Option A: The following rule is introduced.
- Option B: The rule is not introduced.

Rule: All payouts for Action Y are increased by 2 ECU.

Whether option A or option B is implemented is determined as follows:

1. You vote.
2. The computer decides whether the voting outcome is taken into account.
   (a) If the election outcome is taken into account, the option, which received the simple majority of votes is implementing
   (b) If the election outcome is ignored, option A is implemented, independent of the voting outcome.

The second part of the experiment

The second part of the experiment consists of 38 independent rounds. In each round, a so-called situation is randomly drawn. A situation consists of 4 income boxes. There are 9 different situations altogether.

- In addition to the instructions, you have received in a separate document (one for option A and one for option B) the description of the 9 situations.

Division of roles One of the participants is randomly assigned the role of actor E. All other participants are randomly assigned the role of actor A or actor B. Each actor A forms a group with a randomly determined actor B.

- You remain in your role throughout the experiment (actor E, A or B)

Rounds (total: 38)

- In one of the rounds you interact with two other participants. You only interact with real participants in this one round. This round is also the only round that is ultimately relevant for your earnings.
  - In this round you will be informed about the actual decisions of actor E and the other actor A or B respectively.
  - Your earnings from this round, depend on actual decisions.
  - In which of the 38 rounds you interact with participants is randomly determined by the computer. You will not know what round this is until the end of the experiment.
- You interact with a computer for the remaining 37 rounds. These rounds are not relevant to your earnings.
  - Nonetheless, the terms actor E, actor A, and actor B continue to be used (although you interact exclusively with a computer).
  - Because you interact with the computer in these 37 rounds, all the information you receive about other player decisions is generated by the computer.
  - The computer randomly chooses between actions X and Y with equal probability.

Please note that the round which is relevant for the pay-offs and you interact with real participants is determined randomly. Therefore, you do not know by the end of the experiment in which of the 38 rounds you interact with other participants and when with the computer. Therefore, no participant can know if a particular round is relevant for the pay-offs.

Course of a round

- **Step 1:** Decision of the actor E
  1. Actor E learns which situation is randomly drawn for the respective round.

2. He choose one of the four possible income boxes: I, II, III or IV.
   - This decision determines the income box in which all other participants interact.

- **Step 2:** Decisions of Actors A and B

  1. Actors A and B learn which situation was drawn and which income box was selected for each round.
  2. Each actor (A and B) chooses at the same time, without knowing the other actors (A or B) decision, between act X and act Y. (See: Figure 1a and 1b)

  Please note that the income box displayed is optimized for easier reading. Your action will always be on the left side and your pay-off will always be displayed first. [Figures 1a and 1b]

- **Step 3: Information**

  - Actors A and B are informed about the decision of the other actors and their earnings in each round after both decisions have been entered. Please note that the information you receive in the 37 (out of 38) rounds that do not result in a pay-off is generated by the computer. As a result, you can not draw conclusions about the actual behavior of the other actors.
  - Actor E does not receive information about the decisions of actors A and B, until the end of the experiment. Only at the end of the experiment is actor E informed as to what actors A and B have decided.

Earnings actor E

Since only one of the participants is assigned the role of actor E and actor E chooses the income box for all groups, each group generates earnings for actor E. Actor E receives the average payout resulting from the decision of all groups.

Example:

If option A applies: You are Actor A. In Situation 1, Actor E selects Income Box I. Then Actor B and you yourself decide on Action X.

- In this case you will receive 90 ECU, Actor B will also receive 90 ECU and Actor E will receive 210 ECU from your group.

If actor B and you choose action Y instead, under the same circumstances

- you will receive ECU 140, player B will also receive ECU 140 and player E will receive ECU 0 from your group.

If in the same situation one actor chooses X and the other Y

- Actor E receives ECU 210 from your group , the actor who has chosen Action Y receives ECU 80 and the actor who has chosen Action X receives ECU 90.

If option B applies: You are Actor A. In Situation 1, Actor E selects Income Box I. Then Actor B and you yourself decide on Action X.

- In this case you will receive 90 ECU, Actor B will also receive 90 ECU and Actor E will receive 210 ECU from your group.

If actor B and you choose action Y instead, under the same circumstances

- you will receive ECU 138, player B will also receive ECU 138 and player E will receive ECU 0 from your group.

If in the same situation one actor chooses X and the other Y

- Actor E receives ECU 210 from your group , the actor who has chosen Action Y receives ECU 78 and the actor who has chosen Action X receives ECU 90.

You have received a questionnaire with comprehension questions to make sure that all participants have understood how to read the income boxes and that other important elements of the instructions have been understood. Please fill in the questionnaire now. Lift your arm out of the cabin as soon as you are done. We will then come to you and check your answers. Please note that the experiment will not start until all participants have completed the questionnaire correctly.

A.1.2 Instructions for the participants (original)

Allgemeine Instruktionen für die Teilnehmende

Herzlich Willkommen im Experimentallabor, Sie nehmen nun an einem wirtschaftswissenschaftlichen Experiment teil. Die Dauer dieses Experiments beträgt etwa 90 Minuten. In diesem Experiment können Sie – je nach Ihren Entscheidungen – einen nicht unbeträchtlichen Geldbetrag verdienen. Es ist daher wichtig, dass Sie diese Instruktionen genau durchlesen.

Für die Dauer des Experiments bitten wir Sie ein paar Regeln zu beachten: Von nun an herrscht ein absolutes Kommunikationsverbot. Wenn Sie Fragen haben, dann richten Sie diese bitte an uns. Dazu melden Sie sich und wir kommen dann zu ihrer Kabine. Die Nicht-Beachtung dieser Regel führt zum sofortigen Ausschluss aus dem Experiment und allen Zahlungen. Auch bitten wir Sie ihre Handys und sonstigen technischen Geräte aus, oder zumindest auf lautlos zu stellen und wegzupacken. Wir möchten nicht, dass Sie oder andere Teilnehmende dadurch gestört oder abgelenkt werden. Wir freuen uns über Ihre Teilnahme und hoffen auf Ihre Kooperation.

Das Experiment gliedert sich in zwei Teile. Wir werden Ihnen zu jedem Teil die Instruktionen austeilen und diese auch laut vorlesen. Dadurch stellen wir sicher, dass alle Teilnehmenden den Aufbau des Experiments einmal vollständig zur Kenntnis genommen haben. Falls Sie hierbei (oder ganz generell) Fragen oder Probleme haben, melden Sie sich bitte aus ihrer Kabine. Fragen bedeutet nicht, dass sie begriffsstutzig sind, sondern meist, dass wir etwas missverständlich formuliert haben. Also zögern Sie bitte nicht zu fragen.

Anonymität: Das Experiment ist anonym. Sie werden an keiner Stelle nach Ihrem Namen oder nach Informationen gefragt, die es dem Laborteam ermöglichen, Sie zu identifizieren. Im Experiment stellen wir Ihnen noch einige Fragen, die Sie nach einigen soziodemographischen Informationen fragen (Alter, Geschlecht), nicht aber nach Ihrem Namen.

Danach erfolgt ihre Auszahlung. Zur Auszahlung werden wir Sie einzeln anhand ihrer Platznummern aufrufen. Nur einer der Experimentatoren und Sie erfahren, was Sie verdient haben. Alle Entscheidungen in diesem Experiment sowie die Auszahlungen am Ende bleiben also anonym.

Auszahlung: Die Höhe Ihrer Einnahmen ergibt sich aus Ihren Entscheidungen, den Entscheidungen der anderen Teilnehmenden und dem Zufall. In dem gesamten Experiment sprechen wir nicht von Euro, sondern von ECU (Experimental Currency Unit). Ihre Einnahmen aus dem Experiment werden zunächst in ECU berechnet. Der von Ihnen während des Experiments verdiente Betrag wird am Ende von ECU in Euro umgerechnet, wobei gilt:

10 (Zehn) ECU entsprechen 1 (Ein) Euro.

Der erste Teil des Experiments

Bevor der zweite Teil beginnt, geben Sie ein, ob Sie für Option A oder Option B sind.

Option A: **Die folgende Regel wird eingeführt.**

Option B: **Die Regel wird nicht eingeführt.**

Regel: Alle Auszahlungen für Handlung Y sind um 2 ECU erhöht.

Ob Option A oder Option B implementiert wird, wird wie folgt bestimmt:

1. Sie stimmen ab.
2. Der Computer entscheidet, ob der Wahlausgang berücksichtigt wird.
    a. Wenn der Wahlausgang berücksichtigt wird, wird die Option implementiert, die die einfache Mehrheit der Stimmen erhalten hat.
    b. Wenn der Wahlausgang nicht berücksichtigt wird, wird die Option A unabhängig vom Wahlausgang implementiert.

Der zweite Teil des Experiments

Der zweite Teil des Experiments besteht aus 38 voneinander unabhängigen **Runden**. In jeder Runde wird eine sogenannte **Situation** zufällig ausgelost. Eine Situation besteht aus 4 **Einkommensboxen**. Insgesamt gibt es 9 verschiedene Situationen.

- Sie haben zusätzlich zu den Instruktionen in einem **Dokument** die Beschreibung der 9 Situationen (jeweils für Option A und Option B) erhalten.

Rollenaufteilung

Einem(r) der Teilnehmenden wird zufällig die Rolle **Akteur E** zugeteilt. Allen anderen Teilnehmenden wird zufällig die Rolle **Akteur A** oder **Akteur B** zugeteilt. Jeder Akteur A bildet mit einem zufällig bestimmten Akteur B eine Gruppe.

- Sie verbleiben das gesamte Experiment über in Ihrer Rolle (Akteur E, A oder B)

Runden (insgesamt: 38)

- *In einer der Runden interagieren Sie mit zwei anderen Teilnehmenden.* Sie interagieren nur in dieser einen Runde ausschließlich mit echten Teilnehmenden. *Diese Runde ist auch die einzige Runde, die am Ende für Ihre Auszahlung relevant ist.*
    - In dieser Runde werden Sie über die tatsächlichen Entscheidungen des Akteurs E und des jeweils anderen Akteurs A oder B informiert.
    - Ihre Einnahmen hängen von diesen tatsächlichen Entscheidungen ab.
    - In welcher der 38 Runden Sie mit Menschen interagieren, wird zufällig vom Computer bestimmt. Sie werden bis zum Ende des Experiments nicht erfahren, welche Runde dies ist.
- *In den verbleibenden 37 Runden interagieren Sie mit einem Computer. Diese Runden sind nicht für Ihre Auszahlung relevant.*
    - *Trotzdem werden weiterhin die Bezeichnungen Akteur E, Akteur A und Akteur B verwendet (obwohl Sie ausschließlich mit einem Computer interagieren).*
    - Da Sie in diesen 37 Runden mit dem Computer interagieren, sind alle Informationen, die Sie über Entscheidungen der anderen Akteure erhalten, vom Computer generiert.
    - Der Computer wählt zufällig mit gleicher Wahrscheinlichkeit zwischen den Handlungen X und Y.

Bitte beachten Sie, dass die für die Auszahlungen relevante Runde, in der Sie mit echten Teilnehmenden interagieren, per Zufall bestimmt wird, d.h. Sie wissen bis zum Ende des Experiments nicht, in welcher der 38 Runden Sie mit anderen Teilnehmenden interagieren und wann es sich ausschließlich um den Computer handelt. Daher kann kein(e) Teilnehmende(r) wissen, ob eine bestimmte Runde für die Auszahlungen relevant ist.

Ablauf einer Runde

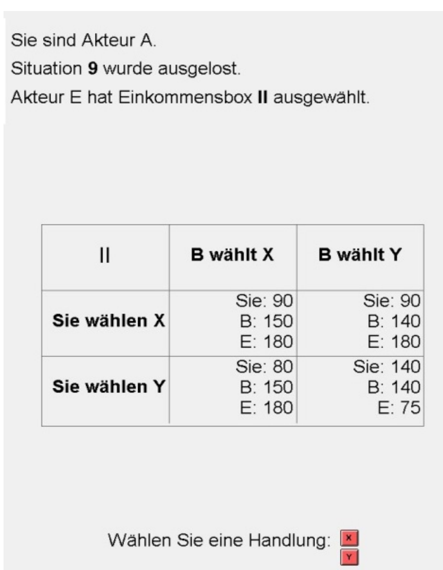
Schritt 1: Entscheidung des Akteurs E

1. Akteur E erfährt, welche Situation für die jeweilige Runde zufällig ausgelost ist.
2. Er wählt eine der vier möglichen Einkommensboxen: I, II, III oder IV.
    - Diese Entscheidung bestimmt die Einkommensbox, in der alle anderen Teilnehmenden interagieren.

Schritt 2: Entscheidungen der Akteuren A und B

1. Akteure A und B erfahren, welche Situation ausgelost wurde und welche Einkommensbox für die jeweilige Runde ausgewählt wurde.
2. Jeder Akteur (A und B) wählen zeitgleich, also ohne die Entscheidung des jeweils anderen zu kennen, zwischen Handlung X und Handlung Y. (Siehe: Abbildung 1a und 1b)

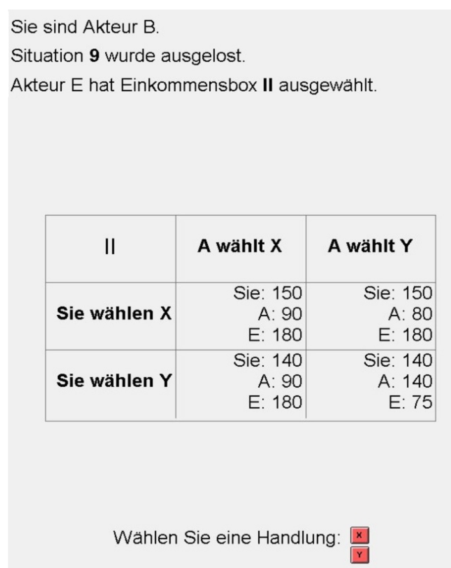
Bitte beachten Sie, dass die angezeigte Einkommensbox für einfacheres Lesen optimiert ist. Ihre Aktion wird immer auf der linken Seite und Ihre Auszahlung wird immer als erste angezeigt.


Sie sind Akteur A.
Situation **9** wurde ausgelost.
Akteur E hat Einkommensbox **II** ausgewählt.

II	**B wählt X**	**B wählt Y**
Sie wählen X	Sie: 90 B: 150 E: 180	Sie: 90 B: 140 E: 180
Sie wählen Y	Sie: 80 B: 150 E: 180	Sie: 140 B: 140 E: 75

Wählen Sie eine Handlung: X Y

Abbildung 1a: Entscheidungsbildshirm Akteur A


Sie sind Akteur B.
Situation **9** wurde ausgelost.
Akteur E hat Einkommensbox **II** ausgewählt.

II	**A wählt X**	**A wählt Y**
Sie wählen X	Sie: 150 A: 90 E: 180	Sie: 150 A: 80 E: 180
Sie wählen Y	Sie: 140 A: 90 E: 180	Sie: 140 A: 140 E: 75

Wählen Sie eine Handlung: X Y

Abbildung 1b: Entscheidungsbildschirm Akteur B

Schritt 3: Informationen

- Die Akteure A und B werden über die Entscheidung der anderen Akteure und ihre Einnahmen in jeder Runde informiert, nachdem beide Entscheidungen eingegeben wurden.

Bitte beachten Sie, **dass die Informationen die Sie in den 37 (von 38) Runden erhalten, welche keine Auszahlung zur Folge haben, vom Computer generiert sind. Demzufolge können Sie keine Rückschlüsse auf das tatsächliche Verhalten der anderen Akteure ziehen.**

- Akteur E erhält bis zum Ende des Experiments keine Informationen über die Entscheidungen der Akteuren A und B. Erst am Ende des Experiments wird Akteur E informiert, wie die Akteure A und B sich entschieden haben.

Einnahmen Akteur E

Da nur einem(r) der Teilnehmenden die Rolle Akteur E zugeteilt wird und Akteur E sich die Einkommensbox für alle Gruppen aussucht, werden durch jede Gruppe Einnahmen für Akteur E generiert. Akteur E erhält die durchschnittliche Auszahlung aus den Interaktionen aller Gruppen.

Beispiel:

Sofern Option A gilt:

Sie sind Akteur A. In Situation 1 wählt Akteur E die Einkommensbox I. Danach entscheiden Akteur B und Sie selbst sich für die Handlung X.

- In diesem Fall erhalten Sie 90 ECU, Akteur B erhält ebenfalls 90 ECU und Akteur E erhält von Ihrer Gruppe 210 ECU.

Wenn Akteur B und Sie sich unter den gleichen Umständen stattdessen für Handlung Y entscheiden

- erhalten Sie 140 ECU, Akteur B erhält ebenfalls 140 ECU und Akteur E bekommt von Ihrer Gruppe 0 ECU.

Falls sich in der gleichen Situation ein Akteur für X und der andere für Y entscheidet

- bekommt Akteur E von Ihrer Gruppe 210 ECU, der Akteur, der sich für Handlung Y entschieden hat, erhält 80 ECU und der Akteur, der sich für Handlung X entschieden hat, erhält 90 ECU.

Sofern Option B gilt:

Sie sind Akteur A. In Situation 1 wählt Akteur E die Einkommensbox I. Danach entscheiden Akteur B und Sie selbst sich für die Handlung X.

- In diesem Fall erhalten Sie 90 ECU, Akteur B erhält ebenfalls 90 ECU und Akteur E erhält von Ihrer Gruppe 210 ECU. (**Keine Veränderung zu Option A**)

Wenn Akteur B und Sie sich unter den gleichen Umständen stattdessen für Handlung Y entscheiden

- erhalten Sie **138 ECU**, Akteur B erhält ebenfalls **138 ECU** und Akteur E bekommt von Ihrer Gruppe 0 ECU.

Falls sich in der gleichen Situation ein Akteur für X und der andere für Y entscheidet

- bekommt Akteur E von Ihrer Gruppe 210 ECU, der Akteur, der sich für Handlung Y entschieden hat, erhält **78 ECU** und der Akteur, der sich für Handlung X entschieden hat, erhält 90 ECU.

Wir haben Ihnen einen Fragenbogen mit Verständnisfragen austeilen, um sicherzustellen, dass alle Teilnehmenden verstanden haben, wie man die Auszahlungsboxen liest und dass auch andere wichtige Hinweise in den Instruktionen verstanden wurden. Bitte füllen Sie den Fragebogen nun aus. Heben Sie Ihren Arm aus der Kabine, sobald Sie fertig sind. Wir werden dann zu Ihnen kommen und Ihre Antworten überprüfen.

Bitte beachten Sie, dass das Experiment erst beginnt, wenn alle Teilnehmenden den Fragebogen richtig ausgefüllt haben.

Ende der Instruktionen

A.1.3 Random information feed

Period	Sub-game	Var.	Response	Period	Sub-game	Var.	Response
1	DAC_B	F_M	Challenge	20	DAC_A	R_M	Challenge
2	DAC_A	α_M	Acquiescence	21	DAC_A	C_L	Challenge
3	BP	C_H	Challenge	22	DAC_B	R_M	Acquiescence
4	BP	α_H	Challenge	23	NC	F_L	Acquiescence
5	BP	α_M	Challenge	24	NC	α_H	Acquiescence
6	BP	F_M	Acquiescence	25	DAC_B	C_L	Challenge
7	NC	C_L	Challenge	26	DAC_A	α_H	Challenge
8	NC	α_M	Acquiescence	27	DAC_B	F_L	Acquiescence
9	DAC_B	R_H	Acquiescence	28	DAC_A	R_M	Acquiescence
10	NC	$Base$	Challenge	29	DAC_B	α_M	Acquiescence
11	BP	F_L	Acquiescence	30	NC	F_M	Challenge
12	BP	C_L	Acquiescence	31	DAC_B	C_H	Acquiescence
13	NC	R_H	Challenge	32	DAC_A	F_L	Challenge
14	NC	R_M	Acquiescence	33	DAC_B	α_H	Challenge
15	DAC_A	F_M	Acquiescence	34	DAC_B	$Base$	Challenge
16	BP	R_H	Challenge	35	DAC_A	R_H	Challenge
17	NC	C_H	Acquiescence	36	BP	$Base$	Acquiescence
18	BP	R_M	Challenge	37	NC	$Base$	Challenge
19	DAC_A	C_H	Acquiescence	38	DAC_B	C_L	Acquiescence

A.1.4 Population-averaged probit model

DV: i-th responders decision in t-th period								
Main explanatory variables								
			DAC_v		DAC_w		BP	
			1.856***	(0.243)	1.281***	(0.217)	3.254***	(0.286)
C_L	0.213	(0.245)	+0.001	(0.319)	−0.448	(0.293)	−0.530(*)	(0.316)
C_H	−0.157	(0.257)	+0.300	(0.288)	−0.364	(0.295)	+0.119	(0.369)
R_M	−0.345	(0.324)	+1.065**	(0.386)	−0.029	(0.419)	+0.340	(0.335)
R_H	−0.386	(0.307)	+1.015**	(0.366)	+0.071	(0.366)	+0.487	(0.363)
α_M	−0.570	(0.383)	+1.028*	(0.396)	+0.113	(0.408)	+0.623	(0.418)
α_H	−0.383	(0.294)	+0.730*	(0.309)	−0.009	(0.379)	+0.580	(0.392)
F_L	−0.294	(0.209)	+1.007**	(0.320)	+0.113	(0.261)	+0.619(*)	(0.369)
F_M	−0.389	(0.274)	+1.043**	(0.366)	+0.187	(0.395)	+0.764*	(0.362)
Control variables								
Experience with sub-game					−0.034	(0.025)		
Experience with variation					0.047	(0.062)		
Lagged partner's response					0.149(*)	(0.083)		
Victim of BP at t-1					−0.036	(0.168)		
Victim of DAC at t-1					0.019	(0.173)		
Witness of DAC at t-1					−0.305*	(0.138)		
Rest time in seconds					0.003(*)	(0.001)		
System-I thinking (<15s)					0.211*	(0.107)		

Obs= 4,218 (111 participants). The model did not converge. Clustered on the participant's level semirobust SE are in parentheses. *** p<0.001, ** p<0.01, * p<0.05, (*) p<0.1.

A.2 Chapter 3

A.2.1 Model parameterization

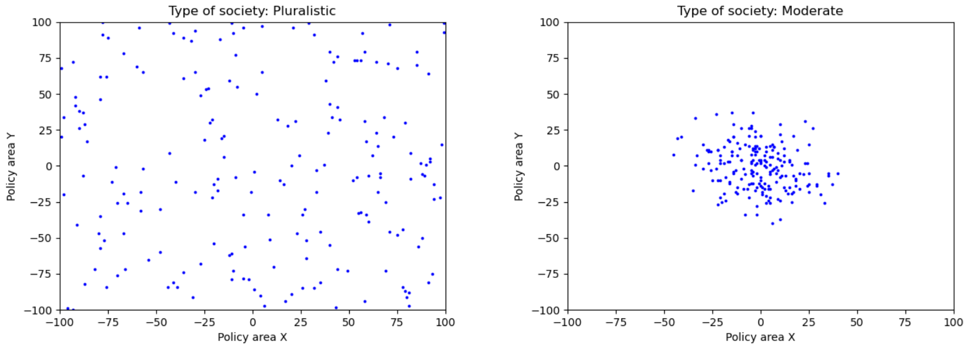
This appendix presents the simulations' parameters. All simulation experiments are programmed and run using Python (version 3.7.0). The pseudo random number generation algorithms used are taken from the package *random* (version 3.8.5).

All simulations have the following common parameters:

1. The society consists of 200 agents ($N = 200$).
2. The game lasts 30 legislative sessions ($S = 30$).
3. The policy space is ten-dimensional ($D = 10$).
4. The laws in each policy area can take 201 integer values between $d_{min} = -100$ and $d_{max} = 100$.
5. In $s = 1$, the constitutional core allows exactly half of the agent's ideal laws in each policy area.
6. Agents specialize high with a 0.25 probability and specialize low with a 0.75 probability.
7. $S^s_{nd} = 1 - \frac{a^s_{nd}}{d_{max} - d_{min}}$.
8. $C^s_{nd} = 1 - \frac{|d^s_d - d^s_{nd}|}{d_{max} - d_{min}}$.
9. All agents are equally likely to be nominated as candidates in $s = 1$.[1]
10. In $s = 1$, before the first legislator legislates the law in force is equal to zero (free market).
11. After the legislator legislates, agents have four time units to comply to laws ($t = 4$).

Depending on how the initial agents' ideal laws are generated, the societies can be divided in five types. In *pluralistic* societies, the computer draws for each agent a point from the policy space with each point being equally likely to be drawn. The point's coordinates become the agent's ideal laws. The

1. This parameter is chosen because it ensures that the results of the experiments will not be biased by the model's parameters. Any other probability distribution would influence which path the society takes.



(a) Pluralistic society

(b) Moderate society

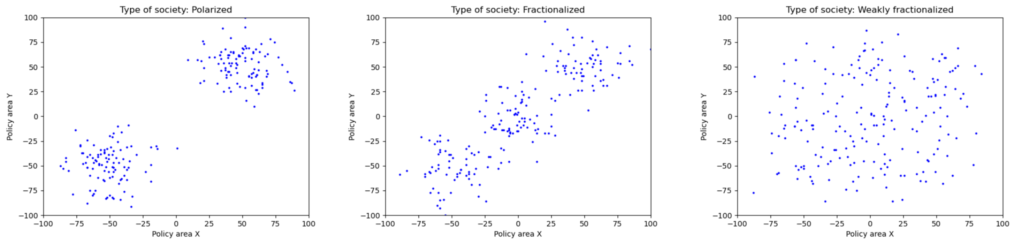
Figure A.1: Types of simulated societies (1)

characteristic of pluralistic societies is that the ideal laws are uncorrelated between policy areas and agents. In other words, the probability that an agent is in favor of laws right of zero in one policy area and laws left of zero in another area or the other way around is high (≈ 0.5). Figure A.1a shows an example generated using a two-dimensional policy space of a pluralistic society.

In *moderate societies*, the computer draws ideal laws from a normal distribution with a mean of zero and a standard deviation of $d_{max}/6$. As a result, the agents' ideal legal orders form a hyper-spherical cloud around the center of the policy space. Ideal laws generated according to this distribution will be referred to as moderate ideal laws and agents who have moderate ideal laws will be referred to as liberal agents. In moderate societies, the probability that an agent will perceive a law as optimal increases as the value of the law approaches zero. Figure A.1b shows an example of a moderate society.

In *polarized* societies, the agents are divided into two groups according to the mean of the distribution from which their ideals are drawn. For one of the groups, the computer draws the values of the ideal laws randomly following a truncated at d_{min} normal distribution with a mean of -50. Agents whose ideal laws follow this distribution will be referred to as leftist agents. For the other group, henceforth the rightist agents, the computer draws ideal laws following a truncated at d_{max} normal distribution with a mean of 50. Figure A.2a shows an example of a polarized society.

Fractionalized societies combine moderate and polarized societies. Specifically, the agents in this society type are equally divided between liberal, leftist and rightist agents. As a result, the agents' ideal legal orders form three clouds, one in the center of the policy space and two on two opposite quadrants of the policy space. Figure A.2b shows an example of a fractionalized society.



(a) (b) (c)

Figure A.2: Types of simulated societies (2)

Last, *weakly fractionalized* societies are fractionalized societies where the agents' ideal laws are drawn from more than one distributions. An ideal law is drawn with a probability equal to 0.5 from moderate laws, from rightist laws or from leftist laws for one third of the society respectively. Those agents whose ideal laws are predominantly drawn from either leftist or rightist laws get with a 0.16 probability ideal laws drawn from rightist or leftist laws, respectively. Last, they can get ideal laws drawn from moderate laws with a 0.34 probability. The agents whose ideals are predominantly drawn from moderate laws can get ideal laws drawn from leftist laws or rightist laws with a 0.25 probability each. Figure A.2c gives an example for a weakly fractionalized society.

In weakly fractionalized societies, the ideal legal orders are spread more in the policy space compared to the ideals in fractionalized and polarized societies. However, the ideals are less spread compared to pluralistic societies. Although the ideal laws are correlated, it is still possible for agents to favor different types of laws in different areas. Hence, one could say that weakly fractionalized societies are a mixture of all the other types of societies.

Societies can also be divided according to how specialized agents begin the game. In $s = 1$ during all simulations, generalists can within the given time adjust to a switch from zero to d_{min} or d_{max} ($a = 25$). Experts can adjust to changes half the size the changes that generalists can adjust to rounded up ($a = 13$).

In *homogeneous* societies, all agents begin as generalists. In *heterogeneous* societies, there are experts and generalists. In these societies, the experts are the minority in the population, i.e. an agent is an expert with a 0.25 probability and a generalist with a 0.75 probability. Experts are specialized in all policy areas. Respectively, generalists are not specialized in any policy area. Last, in *diversified* societies agents can be experts in one area and generalist in another area. Combining the two criteria, the fifteen type of societies are generated.

The constitutional core is generated by a pseudo random number generation algorithm. The computer uses the algorithm to select two numbers between d_{min} and d_{max} to represent d_l and d_r. If d_l is smaller (d_r is bigger) that the value for the ideal law of the most leftist (rightist) agent in the society the computer increases d_l (reduces d_r) to the value of the most leftist (rightist) agent's ideal law.

This step of the procedure is designed to reflect our current understanding of constitution-making. Specifically, we have evidence that drafters act in accordance with their personal interests (Beard 1913; McGuire and Ohsfeldt 1989; Elster 1995, 2012; Ginsburg, Elkins and Blount 2009). Had the agents been allowed to draft the constitution- as long as there is no drafter, interest group or coalition of drafters which would profit by fighting for laws further to the right or left (away from his ideal law or outside of the coalition core), these laws would also not be allowed by the constitution.

To ensure that all constitutions begin by allowing the same portion of the agent's ideal laws, the computer is programmed to check whether between d_l and d_r lie the ideal laws of exactly 100 agents. If the two numbers are not sufficiently apart, the computer expands the core. If the constitutional core allows more that half of the agents' ideal laws, the computer is programmed to shrink the core.

As a result, the probability that a law is allowed by the constitution increases as the law's value approaches zero. By extension, not regulating a policy area, i.e. the free market has the highest probability to be allowed by the constitution. Last, the ideal law of the median agent in each policy area is always constitutional.

Figure A.3 shows the distribution of coverage according to the society's type. Within the violin plots are boxplots. The vertical red line shows the median coverage and the dots represent the outliers in the sample. The provision with the smallest coverage allowed for 9% of all possible laws. The provision with the largest coverage allowed 61% of all possible laws and the average coverage was 39.81% (SD=13.72). The ideals in moderate societies were the most concentrated. Thus, it comes to no surprise that the provisions in these societies were on average the most restrictive.

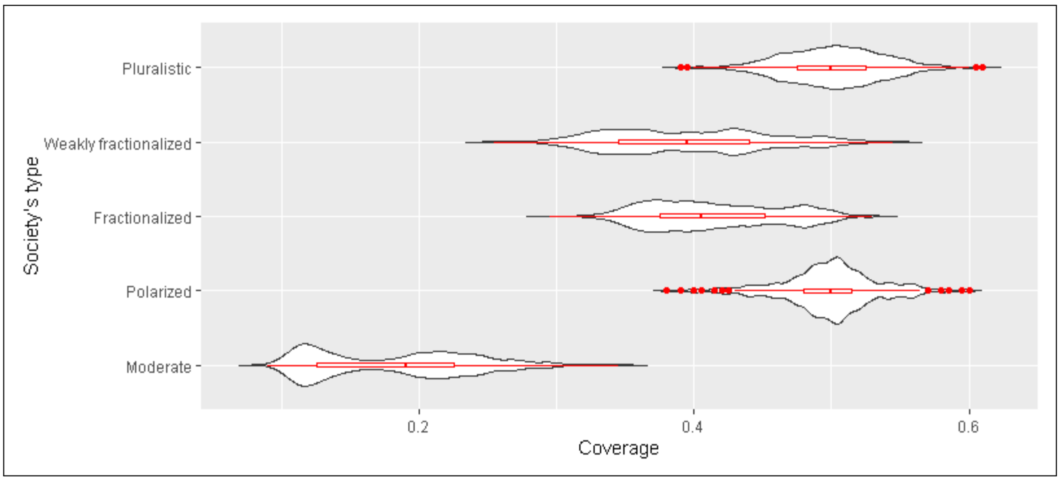


Figure A.3: Violin plots: Coverage by society's type

References

Beard, Charles A. 1913. *An Economic Interpretation of the Constitution of the United States.* New York, NY: Macmillan.

Elster, Jon. 1995. Forces and mechanisms in the constitution-making process. *Duke Law Journal* 45 (2): 364–396.

———. 2012. The Optimal Design of a Constituent Assembly. In *Collective wisdom: principles and mechanisms,* edited by Hélène Landemore and Jon Elster, 148–172. Cambridge, MA: Cambridge University Press.

Ginsburg, Tom, Zachary Elkins and Justin Blount. 2009. Does the Process of Constitution-Making Matter? *Annual Review of Law and Social Science* 5:201–223.

McGuire, Robert A, and Robert L Ohsfeldt. 1989. Self-interest, Agency Theory, and Political Voting Behavior: The Ratification of the United States Constitution. *American Economic Review* 79 (1): 219–234.

A.3 Chapter 4

A.3.1 Instructions for the participants: phase one (translation)

General instructions

Welcome! We look forward to your participation and hope for your cooperation! You are now taking part in an economic experiment. The duration of this experiment is approximately 20 minutes. The participation is anonymous. An identification number has already been given to you. All your decisions are traced back to your identification number only. Neither the experimenters nor other participants have the opportunity to match your identification number to you. At no point will you be asked for your name or information that will allow the experimenters to identify you. If you receive a reward, the Labor team will inform you of the email address you used to register in hroot. For the duration of the experiment, we ask you to observe a few rules: From now on, please stop communicating with other participants in the experiment and put your cell phones and other technical devices away. We do not want you or other participants to be disturbed or distracted by this. Otherwise we will ask you to leave the experiment and we will also exclude you from the remuneration.

The experiment consists of two phases. You only take part in the first phase. You will shortly be given a task. Please read the instructions for the task carefully. If you have any questions or problems, please raise your hand and we will come to you. Asking does not mean that you are confused, but mostly that we have formulated something misleading. So please do not hesitate to ask.

The task

Your task is to find and mark a single combination of pairs of numbers in each box, the sum of which is exactly ten (10.00). You will be given five (5) minutes to complete the task.

Hints:

- There are 15 boxes, each with 12 numbers between 0.01 and 9.99 per box (see examples).
- The boxes are identical for all participants.
- There is at least one correct pair of numbers per box!

- The type of marking (e.g. circling, ticking, etc.) does not matter as long as it is clear.
- If you have marked more than two (2) numbers (e.g. by mistake), you can enter the correct number combination at the edge of the box.
  - If the marked and the entered numbers do not match, the entered numbers will be used as your answer.
  - If you have marked or entered more or less than two numbers, the answer will be counted as wrong.

-Example figures-

At the end of the 5 minutes, i.e. after the task:

1. Stow the task boxes in the envelope.
2. If you edit the task boxes after the five minutes, you will be excluded from all payments.
3. Complete the survey and also stow it in the envelope.
4. Wait until your seat number is called and throw the envelope into the box in front of the exit.

Remuneration

Your remuneration is determined based on your performance in the task and a bonus from the second phase. All payments are made via Amazon vouchers up to four weeks after the end of the experiment (probably at the end of August).

Main remuneration (10 euros): All participants from today's sessions are divided into groups of two, i.e. You will be assigned to a partner. Your number of correct answers forms your score. The number of correct answers from your partner forms his score.

- If you have a higher score than your partner, you will receive ten (10) euros. Your partner receives zero (0) euros.
- If you have a lower score than your partner, you will receive zero (0) euros. Your partner receives ten (10) euros.
- If the scores are the same, a random draw will be made to determine which of the two receives the full remuneration.

Bonus (4 euros): You could also get a four (4) euro bonus. In the second part of the experiment, randomly chosen independent observers decide whether you will receive the bonus. The bonus is paid out when all observers have cast their votes. Within the next month, 400 external observers will independently decide whether you or your partner will receive the bonus. Only one of these decisions will be relevant for you and your partner. Before the observers make a decision, they get the same information as you. In addition, they will find out your score and that of your partner, as well as information from the answer sheets. The scores are not assigned to you personally or your identification number. The distribution of the bonus leads to no further earnings for the observers.

A.3.2 Instructions for the participants: phase one (original)

Allgemeine Instruktionen für die Teilnehmenden

Herzlich willkommen! Wir freuen uns über Ihre Teilnahme und hoffen auf Ihre Kooperation!

Sie nehmen nun an einem wirtschaftswissenschaftlichen Experiment teil. Die Dauer dieses Experiments beträgt etwa 20 Minuten. Das Experiment ist anonym. Eine Identifikationsnummer ist Ihnen bereits ausgehändigt worden. Alle Ihre Entscheidungen werden ausschließlich auf Ihre Identifikationsnummer zurückgeführt. Weder die Experimentatoren noch andere Teilnehmende haben die Möglichkeit, Ihre Identifikationsnummer Ihrer Person zuzuordnen. Sie werden an keiner Stelle nach Ihrem Namen oder nach Informationen gefragt, die es den Experimentatoren ermöglichen, Sie zu identifizieren. Sofern Sie eine Entlohnung erhalten, wird das Labour Team Sie über die E-Mail-Adresse informieren, mit der Sie in hroot registriert sind.

Für die Dauer des Experiments bitten wir Sie, ein paar Regeln zu beachten: Stellen Sie bitte von jetzt an die Kommunikation mit anderen Teilnehmenden des Experiments ein und legen Sie Ihre Handys und sonstigen technischen Geräte weg. Wir möchten nicht, dass Sie oder andere Teilnehmende dadurch gestört oder abgelenkt werden. Wir werden Sie andernfalls bitten, das Experiment zu verlassen und werden Sie außerdem von der Entlohnung ausschließen.

Das Experiment besteht aus zwei Phasen. Sie nehmen nur in der ersten Phase teil. Sie bekommen gleich eine Aufgabe. Bitte lesen Sie die Instruktionen für die Aufgabe sorgfältig durch. Falls Sie hierbei Fragen oder Probleme haben, heben Sie bitte die Hand und wir kommen zu Ihnen. Fragen bedeutet nicht, dass Sie begriffsstutzig sind, sondern meist, dass wir etwas missverständlich formuliert haben. Also zögern Sie bitte nicht, zu fragen.

Aufgabe

Ihre Aufgabe ist es, in jeder Box eine einzige Kombination von Zahlenpaaren, deren Summe genau zehn (10,00) beträgt, zu suchen und zu markieren. Sie bekommen fünf (5) Minuten Zeit für die Aufgabe.

Hinweise:

- Es gibt 15 Boxen mit jeweils 12 Zahlen zwischen 0,01 und 9,99 pro Box (siehe Beispiele).
- Die Boxen sind identisch für alle Teilnehmenden.
- Pro Box existiert **mindestens** **ein** richtiges Zahlenpaar!

- Die Art der Markierung (z.B. einkreisen, ankreuzen usw.) spielt keine Rolle, solange sie eindeutig ist.
- Falls Sie (z.B. aus Versehen) mehr als zwei (2) Zahlen markiert haben, können Sie am Rand der Box die korrekte Zahlenkombination eintragen.
    - Falls die markierten und die eingetragenen Zahlen nicht übereinstimmen, werden die eingetragenen Zahlen als Ihre Antwort benutzt.
    - Falls Sie mehr oder weniger als zwei Zahlen markiert oder eingetragen haben, wird die Antwort als falsch angerechnet.

Korrekte Antwort

Beispiel 0			
9,31	~~2,86~~	9,33	2,85
1,43	8,58	7,09	2,89
6,88	2,88	~~7,14~~	6,87

Korrekte Antwort

Beispiel 2			
9,31	2,86	9,33	2,85
1,43	8,58	7,09	2,89
6,88	~~2,88~~	~~7,14~~	6,87

2,86 und 7,14

Korrekte Antwort

Beispiel 1			
9,31	~~2,86~~	9,33	2,85
1,43	8,58	7,09	2,89
6,88	~~2,88~~	~~7,14~~	6,87

2,86 und 7,14

Falsche Antwort

Beispiel 3			
9,31	~~2,86~~	9,33	2,85
1,43	8,58	7,09	2,89
6,88	2,88	~~7,14~~	6,87

2,88 und 7,14

Am Ende der 5 Minuten, also nach der Aufgabe:

1. Verstauen Sie die Aufgabeboxen in dem Umschlag.
2. Bearbeiten Sie die Aufgabenboxen über die Zeit hinaus, werden Sie von allen Zahlungen ausgeschlossen.
3. Füllen Sie die Umfrage aus und verstauen Sie auch diese in dem Umschlag.
4. Warten Sie, bis Ihre Platznummer aufgerufen wird und werfen Sie den Umschlag in die Kiste vor dem Ausgang.

Entlohnung

Ihre Entlohnung wird anhand Ihrer Leistung in der Aufgabe und einem Bonus aus der zweiten Phase bestimmt. Alle Auszahlungen erfolgen über Amazon Gutscheine bis zu vier Wochen nach Experimentsende (voraussichtlich Ende August).

Hauptentlohnung (10 Euro):

Alle Teilnehmenden aus den heutigen Sessions werden in Zweiergruppen aufgeteilt, d.h. Sie werden einem Partner zugeteilt. Ihre Anzahl richtiger Antworten bildet Ihren Score. Die Anzahl richtiger Antworten ihres Partners bildet seinen Score.

- Falls Sie einen höheren Score als Ihr Partner haben, erhalten Sie **zehn (10) Euro**. Ihr Partner erhält null (0) Euro.
- Falls Sie einen niedrigeren Score als Ihr Partner haben, erhalten Sie **null (0) Euro**. Ihr Partner erhält zehn (10) Euro.
- Falls die Scores gleich hoch sind, wird zufällig gelost, wer von beiden die komplette Entlohnung erhält.

Bonus (4 Euro):

Sie könnten zusätzlich einen Bonus von **vier (4) Euro** erhalten. Ob Sie den Bonus erhalten, wird von zufällig gewählten unabhängigen Beobachtenden im zweiten Teil des Experiments entschieden.

Die Auszahlung des Bonus erfolgt, wenn alle Beobachtenden ihre Stimmen abgegeben haben. Innerhalb des nächsten Monats werden 400 externe Beobachtende unabhängig voneinander entscheiden, ob Sie oder Ihr Partner den Bonus erhalten. Nur eine dieser Entscheidungen wird für Sie und Ihren Partner relevant sein.

Bevor die Beobachtenden eine Entscheidung treffen, bekommen sie die gleichen Informationen wie Sie. Zusätzlich erfahren sie Ihren Score und den Ihres Partners, sowie Informationen aus den Antwortbögen. Die Scores sind nicht Ihrer Person oder Ihrer Identifikationsnummer zugeordnet. Die Verteilung des Bonus führt zu keinem weiteren Verdienst für die Beobachtenden.

Ende der Instruktionen

Haben Sie noch Fragen? Bitte melden Sie sich und wir kommen zu Ihnen. Nach Beginn der Zeit dürfen Sie keine weiteren Fragen stellen.

A.3.3 Robustness tests

	Logistic regression: i-th spectator gives the bonus to the winner of the j-th pair								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Woman won	0.566***	0.667***	−0.482***	0.669***	0.562***	0.562***	0.563***	−0.473***	−0.457***
	(0.047)	(0.048)	(0.134)	(0.048)	(0.048)	(0.048)	(0.048)	(0.137)	(0.138)
Mixed pair	−1.679***	−1.784***	−1.723***	−1.794***	−1.657***	−1.656***	−1.664***	−1.700***	−1.707***
	(0.050)	(0.052)	(0.051)	(0.052)	(0.051)	(0.051)	(0.051)	(0.052)	(0.052)
Mixed pair:	−0.803***	−0.309***	0.094	−0.285***	−0.805***	−0.805***	−0.786***	0.082	0.091
woman won	(0.052)	(0.058)	(0.116)	(0.059)	(0.052)	(0.052)	(0.053)	(0.118)	(0.119)
Likelihood^2	−0.004***		−0.003***		−0.004***	−0.004***	−0.004***	−0.003***	−0.003***
	(0.0002)		(0.0003)		(0.0002)	(0.0002)	(0.0002)	(0.0003)	(0.0003)
Priors OR		0.001		0.004**	0.0005	0.0003	0.003	0.0003	0.003
		(0.001)		(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)
Likelihood^2:						0.00004	0.00004	0.00005	0.00004
Priors OR						(0.00003)	(0.00003)	(0.00003)	(0.00003)
Woman won:			0.970***					0.959***	0.944***
Likelihood^2			(0.114)					(0.116)	(0.118)
Woman won:				−0.006**			−0.005*		−0.006*
Priors OR				(0.003)			(0.003)		(0.003)
Woman won: Priors:									0.002
Likelihood^2									(0.005)
Observations	8,993	8,809	8,993	8,809	8,809	8,809	8,809	8,809	8,809
Log Likelihood	−7,747	−7,794	−7,740	−7,793	−7,609	−7,608	−7,607	−7,601	−7,600

The coefficients represent logits. All models contain all the variable in the matrix *Control*. The robust standard errors are reported in the parentheses. All models are estimated with spectator's fixed effects. *p<0.1; **p<0.05; ***p<0.01.