Does the threat of overthrow discipline the elites?

Evidence from a laboratory experiment*

Konstantin Chatziathanasiou Svenja Hippel Michael Kurschilgen

July 2021

Abstract. The threat of revolution disciplines the elites, inducing them to strategically share some of their wealth in order to prevent social unrest and preserve their privileges. This behavioral conjecture has been prominently used in the social sciences to explain the extension of the franchise and the expansion of the welfare state. We test this conjecture in a controlled laboratory experiment. We model a society whose members can produce wealth by successfully coordinating conflicting interests. Coordination is facilitated through a pre-game status-ranking. Compliance with the status order yields an efficient yet inequitable payoff distribution, in which a player's wealth is determined by her pre-game status. Between treatments, we vary (a) whether overthrows – which reset the status-ranking via collective disobedience – are possible or not, and (b) whether voluntary redistributive transfers – which high-status players can use to appease the low-status players – are available or not. In contrast to established thinking we find that, on average, the threat of overthrow does not lead to more wealth-sharing. High-status players fail to provide sufficient redistribution to prevent overthrows, and thus lose their privileged position. Groups only stabilize and prosper when an overthrow eventually thrusts intrinsically generous players into high-status positions.

Keywords: redistribution; franchise extension; revolution; elite behavior; coordination; battle of the sexes; experiment.

JEL codes: C72, C92, D74, H23, P48.

---

* K. Chatziathanasiou is affiliated to the University of Münster (kchatzia@uni-muenster.de), S. Hippel to the University of Würzburg (svenja.hippel@uni-wuerzburg.de), M. Kurschilgen to the Technical University of Munich and the Max Planck Institute for Research on Collective Goods (m.kurschilgen@tum.de). We thank an anonymous referee, the editors-in-chief, Samira Akbarian, Stefanie Egidy, Christoph Engel, Jacob Goldin, Oliver Kirchkamp, Julian Nyarko, J.J. Prescott, Cheryl Saunders, Cornelius Schneider, Theresa Wenning, as well as seminar participants at MPI Bonn, U Bonn, U Jena, U Maastricht, U Michigan, the Spring Meeting of Young Economists, the ComplianceNet conference at UC Irvine and the IACL Junior Scholars Forum at the National University of Singapore for very valuable comments and discussion. Financial support by the Max Planck Society is gratefully acknowledged.

1 Introduction

Historically, wealth and power have been distributed unequally (Scheidel 2017). At different times, inequality has fueled social unrest among the disadvantaged, and ultimately revolutions (Muller and Seligson 1987; Goldstone 2016). Facing the threat of being overthrown, elites may try to accommodate demands for redistribution while maintaining the bulk of their privileges. This 'threat of revolution hypothesis' has been prominently used in the social sciences to explain the extension of the franchise and the expansion of the welfare state (Acemoglu and Robinson 2000; Boix 2003; Acemoglu and Robinson 2006; Kim 2007; Burke and Leigh 2010; Brückner and Ciccone 2011; Chaney 2013; Aidt and Jensen 2014; Aidt and Franck 2015; Dasgupta and Ziblatt 2015; Aidt and Leon 2016). Yet, despite being highly intuitive, the threat of revolution hypothesis is not beyond doubt. First, in historical data only real events of social unrest are typically observable, while the latent *threat* of revolution is much harder to observe. Second, even if instances of social unrest are observable, identifying the social unrest as the cause of a change in elites' behavior is notoriously difficult. Potentially, social unrest could be more likely in a political climate already open toward democratic ideas, or with a franchise extension at the horizon. Third, many instances of constitutional reform have happened without revolutionary pressure (Congleton 2010).

This paper complements and contrasts current historical research with a novel laboratory experiment. In particular, we examine key behavioral conjectures underlying the threat of revolution hypothesis: that elites are sensitive to the threat of social unrest, that they react to it by sharing their wealth, and that their observed generosity is motivated by their interest in maintaining their privileges. Our experimental framework models a society in which players can produce wealth by successfully coordinating conflicting interests. Coordination is facilitated through a status order, which ranks each player according to a pre-game lottery. Compliance with the status-ranking yields an efficient yet inequitable distribution of payoffs. To test the effects of the threat of overthrow on elite behavior, we modify this environment in our treatment conditions. Between treatments, we vary (a) whether overthrows – which reset the status-ranking via collective disobedience – are possible or not, and (b) whether voluntary redistributive transfers – which high-status players can use to appease the low-status players – are available or not.

The threat of overthrow is modeled via the possibility to reset the status-ranking through collective disobedience. If, in a given period, the status order is disregarded by a critical mass of society, ranks are allotted anew. After the re-allotment, low-status players may find themselves in the positions of former high-status players, and vice-versa. To trigger an overthrow, players must coordinate on collectively disobeying the prevalent status order, which requires both persistence and the willingness to forfeit monetary payoffs, most likely repeatedly. On our second treatment dimension, we allow high-status players to voluntarily transfer parts of their wealth to low-status players. A credible threat of overthrow should motivate high-status players to be more generous with their transfers in order to prevent the overthrow from happening, thereby generating more voluntary compliance with the prevalent social order. For the reduction of economic disparities, voluntary redistribution can be regarded as a – more efficient – substitute to overthrows. For generating economic surplus, overthrows and redistribution may actually complement one another. In light of a potential overthrow, adequate redistribution might serve to preserve both, elite privileges and overall wealth.

In contrast to established thinking, we find that, on average, the threat of overthrow does not lead to more wealth-sharing. Strikingly, higher ranks' willingness to make redistributive transfers does not increase in the presence of a credible overthrow threat. As a consequence, overthrows happen frequently and act as a highly inefficient equalizer. In most experimental groups, high-status players do not react sufficiently to the threat of overthrow. As overthrows happen, these players lose their high status as ranks are reset, and suffer a substantial – and preventable – drop in earnings. Some groups, however, experience overthrows only in the first half of the experiment, and become perfectly stable in second half. In these stable groups, early overthrows put, eventually, more generous players into high-status positions. By sharing a larger part of their wealth through voluntary redistribution, these players manage to sustain their status. Both low ranks and high ranks earn significantly higher payoffs in the stable groups compared to the unstable groups.

In our experiment, overthrows are not just theoretically possible, they happen repeatedly. In addition, being generous pays off for the elites. And yet, our data suggests that the higher transfers in the stable groups are not the result of successful observational learning from the mistakes of earlier (overthrown) high ranks. Instead, groups stabilize and prosper when an overthrow eventually thrusts intrinsically generous players into high-status positions. The high-status players of the stable groups were already generous *before* they became high-status players. Their generosity correlates with higher levels of trust towards strangers as measured in our post-experimental tests.

In sum, while we do not find evidence supporting the notion that the threat of overthrow leads to more wealth-sharing, we do find that overthrows can function as a selection mechanism. In the absence of coercive means to protect their privilege, only sufficiently generous elites manage to remain in place. Our results challenge the notion that elites engage in strategic wealth sharing, and instead point towards mechanisms of elite exchange, and democratization processes in general, as drivers of wealth sharing.

The remainder of the paper is structured as follows. Section 2 illustrates this paper's contribution to different strands of literature. Section 3 details our experimental design and our modeling choices. Section 4 presents our hypotheses. We present our experimental results in Section 5, and discuss them in Section 6.

2 Related Literature

Threat of Revolution. In the "the long nineteenth century" (Hobsbawm 1987), Europe witnessed a significant democratization process, i.e. the extension of the franchise, as well as the emergence of the welfare-state (Evans 2016). Both developments rest on significant concessions by the elites. What exactly motivated these concessions has been a prominent topic of investigation in the social sciences.

Typically, the historic emergence of the welfare state is understood as a reaction to the popularity of socialist movements (Evans 2016, p. 553). Sociological analysis sees the welfare-state as an instrument to "regulate the poor" (Cloward and Piven 1971). This perspective is taken up in the economic literature, according to which redistributive policies are put in place by the rich to prevent revolutionary action and extralegal appropriation by the poor (Grossman 1995). Acemoglu and Robinson (2000) apply the 'threat of revolution hypothesis' to democratization. Their main example is Britain, which substantially extended the franchise with parliamentary Reform Acts in 1832, 1867, and 1884. This extensions led to reforms in the areas of taxation and education, and to a reduction of inequality through redistribution. As Acemoglu and Robinson (2000, p. 1183) emphasize, each franchise extension was preceded by social unrest including riots. Similar patterns are observable in France, Sweden, and – to a lesser extent – Germany.[1]

Using a panel of twelve countries, Aidt and Jensen (2014) find that the first wave of democratization in Europe, between 1820 and 1938, can be understood as a reaction to the threat of revolution, as there is a relationship between revolutionary events in neighboring countries and franchise reform. The relationship is stronger in cases in which countries are geographically or linguistically close to the center of the revolutionary event. Aidt and Franck (2015) find that the proximity of the so-called Swing Riots induced voters in Britain to elect reform-oriented politicians in 1831, facilitating the adoption of the Great Reform Act in 1832 by the British Parliament. Dasgupta and Ziblatt (2015) study British sovereign bond market responses to the Reform Acts of 1832, 1867 and 1884, as well as to two failed Chartist agitations for reform, finding that perceived political risk increased before the reform episodes, with the risk stemming from social unrest as well as political deadlock. Using a global, country-level dataset starting after World War I, Przeworski (2009, p. 303) reports a positive correlation between franchise extension and observable instances of social unrest. Using historical data on social insurance coverage from 12 European countries from 1880 to 1945, Kim (2007) finds that the expansion of social insurance is associated with democratization which again is influenced by strikes, thus linking the 'threat of revolution hypothesis' to the expansion of the welfare state. In accordance with the 'threat of revolution hypothesis', Burke and Leigh (2010), Brückner and Ciccone

[1] Germany presents itself as a slightly different case. Social unrest and a strong socialist movement provoked welfare reforms, but franchise extension was slow until the Weimar Republik of 1919. Acemoglu and Robinson (2000, p. 1184) argue that a franchise extension was not necessary to make a promise for future redistribution credible, as the threat of revolution was already tangible due to the strong socialist movement.

(2011) and Aidt and Leon (2016) find adverse weather shocks to be triggers of democratic change. As the shocks temporarily lower citizens opportunity cost in contesting power, they open a window for democratic change.

Notwithstanding its prominence, the 'threat of revolution hypothesis' is not beyond doubt. Congleton (2010, p. 15) points out that most constitutional reforms in Europe happened incrementally through "formal and informal bargains over constitutional details" and were not always preceded by large-scale revolts. Lizzeri and Persico (2004) suggest that the elites extended the franchise and thus changed the incentives for politicians in order to create better institutions that were not dominated by clientelism and patronage. Indeed, causal claims about the 'threat of revolution' are notoriously difficult. In historical data, only real events of social unrest are clearly observable, while the *threat* of revolution is much harder to observe. And while instances of social unrest might be observable, it cannot be said with certainty whether they cause franchise extension. Not only could there be different driving forces between unrest and franchise extension, it could even be that social unrest becomes more likely with a franchise extension at the horizon, or in a political climate with more openness to democratic ideas. This paper complements historical research with an experimental study on the behavioral mechanisms underlying the threat of revolution hypothesis.

Democratic Backsliding. Whether elites are sensitive to demands by non-elites to lessen inequality does not only matter for constitutional *reform*. It also matters for the *preservation* of a constitution. Many of today's political conflicts have shifted to questions of so-called democratic backsliding (Levitsky and Ziblatt 2018). One problem in this context is growing socio-economic inequality (Sitaraman 2017; Dixon and Suk 2018; Levitsky and Ziblatt 2018; Khaitan 2019; Pistor 2019) and a lack of elite responsiveness to the preferences of the economically disadvantaged (Gilens 2012). This becomes particularly problematic when a constitution is faced with a transgressor, i.e. a political actor who disregards constitutional provisions. Lacking centralized external enforcement, the enforcement of constitutions poses a problem of collective action (Hardin 1989; Weingast 1997; Bednar 2009, p. 170). Instead, it relies on officials and citizens having a large-enough incentive to act against a constitutional transgression. High levels of inequality, however, might reduce disadvantaged citizens' motivation to protect the constitutional order. Even worse, some citizens might see it in their interest to support a despotic transgressor (Ryvkin and Semykina 2017). Against this background, the behavior of the elites, who benefit more from a stable constitutional order, turns out to be critical for its preservation. A lack of wealth-sharing might lead to a lack of support of institutions.[2]

[2] A social hierarchy supported by the constitution is only one possible case of social order. Our model of a social hierarchy is not specific to constitutionalism but applies to any situation in which widespread adherence to a social order benefits the members of society to different degrees.

Endogenous Institutions. While we are mainly interested in the *disciplining* effect of the threat of overthrow, there could potentially be other effects. A range of experiments shows that an institution's efficacy is increased if it was generated endogenously, typically through some form of voting mechanism which adds a "democratic dividend" (Tyran and Feld 2006; Dal Bó et al. 2010; Sutter et al. 2010; Markussen et al. 2014; Hauser et al. 2014; DeAngelo et al. 2020; Langenbach and Tausch 2019). The overthrow mechanism of our experiment could add such "democratic dividend", as it enables players to affect the social ordering by recurrently reshuffling the ranks. Moreover, the mere option of overthrow might already cast a shadow of legitimacy on the status quo. Analytically, however, such a legitimizing effect of the threat of overthrow is separate from a disciplining effect on elites. We account for this distinction in our research design by testing the threat of overthrow in two distinct environments: (i) In our baseline environment, elites have no means to react to the overthrow threat. Higher compliance with the status order in the presence of an overthrow threat would thus point toward a legitimizing effect. (ii) In our redistributive environment, elites can use voluntary transfers to increase the non-elite's participation in the proceeds from a stable order. Higher compliance due to higher transfers would point toward a disciplining effect of the overthrow threat.

Focality and Competing Interests. Our experimental framework builds on the literature on coordination games and focal points (Schelling 1960). Mehta et al. (1994) and Bardsley et al. (2009) show how focal points are effective when players' interests are perfectly aligned. However, Crawford et al. (2008) point out that the power of focality is considerably reduced when payoffs are minimally asymmetric. In cases of a pronounced payoff asymmetry – as in our experimental implementation – even explicit recommendations to play a specific equilibrium fail, which leads to substantial efficiency losses (Anbarcı et al. 2018). As one would assume, this is due to the resistance of the players who are asked to play their less preferred equilibrium. The problem can be framed in terms of competing saliencies (Crawford et al. 2008, p. 1444). In our game, the salience of higher individual payoffs competes with the salience of the behavioral recommendations derived from the (hierarchical) social order.[3]

Experimental Regime Change. Our work relates to experimental research on regime change. Ryvkin and Semykina (2017) study a game where citizens can choose to replace a democratic regime with an autocrat. In the democratic regime, property rights are secure and redistribution requires a majority vote, while the autocrat promises full redistribution but can potentially expropriate the citizens. Subjects are more likely to voluntarily switch from democracy to autocracy when inequality is high. This resonates with the aforementioned intuition that inequality might have a destabilizing effect benefiting autocracy.

[3] This does not imply that the social order it creates is legitimate. The mere existence of a focal point is analytically separate from its legitimacy (Elster 1984, p. 95; Elkins et al. 2007, p. 1146; McAdams 2015, p. 49).

3 Experimental Design

We first present the general framework of our experimental design (3.1), followed by treatment variations (3.2) and experimental procedures (3.3).

3.1 *General Framework*

We model a society of N strangers. Over numerous periods, members of this society make random bilateral, anonymous encounters, in which they are faced with a simple coordination problem. Coordination is successful and produces wealth if the two players who meet in the stage game agree on who *claims* a coveted resource, and who *concedes* it. The stage game is a *Battle-of-the-Sexes (BoS)* as depicted in Figure 1. Due to $h > l > 0$, the game has two pure-strategy Nash equilibria, with each player preferring the equilibrium in which she claims and the other player concedes.

		player 2	
		claim	*concede*
player 1	*claim*	$0, 0$	h, l
	concede	l, h	$0, 0$

Figure 1: Stage Game

The two players who meet in a stage game can solve their coordination problem by following the recommendation of an exogenous coordination device. This mirrors the efficiency-enhancing effect of societal coordination through rules, e.g. rules over property. There is no means to exogenously enforce compliance with the rules of society; coordination relies on voluntary compliance. The coordination device is provided in the form of a status-ranking, mirroring a legal or social order that ranks players from highest to lowest degree of privilege. Whenever two players meet, they mutually and unambiguously recognize who is of higher status (and thus supposed to *claim* the resource) and who is of lower status (and thus supposed to *concede* that right to the other player). The higher (lower) one's rank, the more often the action recommended by the coordination device is to claim (concede). Players are informed about their counterpart's relative rank (higher or lower than oneself) but not of their counterpart's absolute rank.

In addition to the two pure-strategy equilibria, of which one is always recommended by the coordination device, the game also has a Nash equilibrium in *mixed strategies*. In this *mixed equilibrium*, players ignore the recommendations of the coordination device; each player chooses her preferred action *claim* with the large probability $\frac{h}{h+l}$ and the less attractive action *concede* with the small probability $\frac{l}{h+l}$. In the mixed equilibrium, players only generate economic surplus in the rare occasion that a claiming player meets a conceding player. The mixed equilibrium is both individually and socially inefficient; its expected payoff is below l.[4] This mirrors a situation in which society is stuck in an anarchic state; each player tries to maximize her payoff under the expectation that nobody respects the (potentially available) legal or social institutions.

3.2 Experimental Treatments

We implement 2×2 treatment variations of this general framework. We vary (a) whether overthrows – which reset the status order via collective disobedience – are possible or not, and (b) whether voluntary redistributive transfers – which high-status players can use to appease the low-status players – are available or not. Table 1 summarizes our four experimental treatments. In all treatments, players are fully informed about all the parameters and the choice structure of the experimental environment they participate in. For the experimental instructions, please refer to Appendix B.

treatment name	transfer option	overthrow mechanism
baseline	-	-
overthrow	-	✓
transfer	✓	-
transfer-overthrow	✓	✓

Table 1: Experimental Treatments

Baseline. In the *baseline* treatment, participants are randomly matched into groups of $N = 6$ players. In each period, following a random matching protocol, players within a group are matched in pairs of two and play a BoS as described above. Thus, every period, there are three parallel matches, in which the matched players play a BoS. The actions of the players are color-coded; players claim by choosing *red*, and concede by choosing *blue*. We set $h = 10$ and $l = 1$ to closely mirror inequality in the real world.[5] The parametrization and the actions are depicted in Figure 2.

We implement the status order as follows: Before the beginning of the first period, every player in a group of six gets assigned a letter (a, b, c, d, e, or f) in a pre-game lottery. Every player has the same probability of receiving each letter and every letter gets assigned exactly once. When two players meet, one of the two (pure) stage-game equilibria is highlighted through bold lettering. Which of the two stage-game equilibria is highlighted, depends on whose player's letter comes first in the alphabet. This is commonly known to all players. The exact letter of the other player is never revealed. In any given period, the probability that the advantageous equilibrium will be highlighted

[4] The probabilities of the different outcomes in the mixed equilibrium are: $P(claim, claim) = \frac{h^2}{(h+l)^2}$, $P(claim, concede) = \frac{hl}{(h+l)^2}$, $P(concede, claim) = \frac{hl}{(h+l)^2}$ and $P(concede, concede) = \frac{l^2}{(h+l)^2}$. Players earn zero payoffs whenever the outcome is $(claim, claim)$ or $(concede, concede)$, which happens with probability $\frac{h^2+l^2}{(h+l)^2}$. The mixed equilibrium thus yields an expected payoff of $\frac{hl}{h+l} < l$.

[5] The mean Gini coefficient of the European Union is .31 (Eurostat and World Bank 2020). In our experimental setup, the theoretical scenario in which (i) all players comply perfectly with the social order (i.e. follow the focal point) and (ii) zero transfers are given, yields a Gini of .32. In the US, inequality is more pronounced (Gini coefficient of .39).

		player 2	
		red	***blue***
player 1	***red***	0,0	**10,1**
	blue	1,10	0,0

Figure 2: Stage Game Parametrization

In any given stage game, whichever player's letter comes first in the alphabet (as assigned by lottery) sees the action "red" in bold. The other player sees the action "blue" in bold. This is commonly known to all players.

equals 1 for type a, 4/5 for type b, 3/5 for type c, and so on. As we have a group of six players that are matched randomly and anonymously in each period without any possibility to communicate, we can rule out that players take turns in playing the more beneficial equilibrium. The random allotment of privilege in the status order mimics the role that exogenous attributes like descent, ethnicity, or gender, have played in historically shaping people's access to resources, professions, and other rights and privileges (Elster 1992; Schotter and Sopher 2003; Moulin 2004). To this day, lifetime income (Bowles and Gintis 2002; Kahlenberg 2010; Chetty et al. 2014; Adermon et al. 2018) and political rights (Shachar 2009) are heavily influenced by pre-birth factors.

Transfer. Whereas in the *baseline* high-status players have only a rather limited action space, in the richer, redistributive environment of the *transfer* treatment, high-status players can, additionally, choose to share some of their wealth through a voluntary transfer mechanism. Directly after every stage game, players can make a transfer to the counterpart they just interacted with in that period. The transfer amount is constrained by the payoff generated through successful coordination in that very period, i.e. the player who earned 10 tokens in the stage game, can transfer between 0 and 10 tokens to the other player.[6] This option to make voluntary transfers mirrors the wealth-sharing that takes place in modern societies and is organized through redistribution mechanisms, such as taxes and social benefits.

Overthrow and Transfer-overthrow. In treatments *overthrow* and *transfer-overthrow*, we introduce an overthrow mechanism to the environments without and with a redistributive option, respectively. The mechanism functions as follows: If, in a given period, all three matched pairs of a given group fail to coordinate on a payoff-yielding equilibrium, then the current status ranking instantly expires, and is replaced by a new – randomly determined – ranking in the subsequent period. The new ranks remain valid until the next overthrow occurs.

Note that triggering the overthrow mechanism is onerous and costly. In expectations, the three lower ranks (d, e, f) would improve their rank after an overthrow while the three higher ranks (a, b, c)

[6] Although we expect mainly the high earners to make use of transfers, we also allow for low earners to transfer (up to 1 token) in order to keep our design symmetric. As expected, the latter case almost never happens in the experiment.

would drop in the status ladder. Consequently, the higher ranks have an interest in preventing an overthrow from happening. Even if the three lower-ranked players deliberately disobeyed the coordination device in order to trigger an overthrow, it still might take several periods until the overthrow is achieved. Due to the random matching of pairings in each period, only in 6 of the 15 possible constellations, each of the three lower ranks is matched with one of the three higher ranks (e.g., a with e, and b with d, and c with f). In the remaining 9 constellations, i.e. in 60% of all periods, there is always one pairing of high-status players that – by complying with the recommendation of the coordination device – can prevent an overthrow from happening (e.g. a with b, or a with c, or b with c). Successfully activating the overthrow mechanism thus requires persistence and the willingness to forfeit monetary payoffs, most likely repeatedly.

The overthrow mechanism mirrors social unrest and the eventual overthrow of the governing elites.[7] In the real world, a single act of non-compliance with the status quo may be perceived as a signal of frustration by non-elites to elites, but it will not lead to overthrow. This changes if non-compliance is collective and persistent. Withdrawal of popular support, typically paired with acts of protest, leads to an exchange of elites (for a classical reference, see de La Boétie 1577, who argued that even tyranny hinges on individual support; for modern examples, see Ackerman and Rodal 2008). How exactly the exchange of elites is achieved, will vary from case to case. There is a great variety in protesting behavior and uprisings (Goldstone 1982; Ackerman and Rodal 2008).[8] It might be that a revolution occurs that violently replaces the elites, it might so be that protests and strikes limit economic activity, or that a social movement pushes for elections or puts pressure on parliament to exchange government. In any case, a heavy investment in practices of resistance is typically needed. Our experimental design accounts for this stylized fact. Triggering the overthrow mechanism is onerous and inherently inefficient, as it requires players to collectively coordinate on disobeying the status quo. Such coordination takes place without explicit communication, but requires nonetheless a persistent group effort.[9]

If the overthrow mechanism is triggered, players again have the same probability to receive a particular letter (a, b, c, d, e, or f) as in the pre-game lottery. In the wake of an overthrow, players can

[7] Thus, the baseline where no overthrow is possible can be understood as modeling a regime so authoritarian that resistance is futile. We thank the editors for pointing this out.

[8] Ackerman and Rodal (2008, p. 111) lay out the possible variance of such protesting non-compliance: "Civilians have used disruptive actions as sanctions, to challenge and delegitimate rulers, mobilize public, constrain authoritarians' power and undermine their sources of support and shift their loyalties. Petitions, marches, walkouts and demonstrations have been used to rouse public support and mobilization. Forms of non-cooperation such as strikes, boycotts, resignations and civil disobedience have served to frustrate the operations of governments. Direct intervention such as blockades, factory occupations and sit-ins have thwarted rulers' ability to subjugate their people. The sequenced, sustained application of these nonviolent operations has engendered historical results: tyrants have capitulated, governments collapsed, occupying armies retreated and political systems that denied human rights been delegitimated and dismantled."

[9] Note that in the real world, it would also be possible to free-ride on the revolutionary efforts of other members of a society. In our experiment, such free-riding is only possible in a weak sense. Players are informed of occuring miscoordinations which could be interpreted as a signal that overthrow is more likely. While it is possible to free-ride on the provision of such a signal, it is not possible to free-ride on the actual overthrow, as the latter requires that all matches miscoordinate.

find themselves in a more privileged position, or in a worse position. We thus mirror not only the investment that is necessary to achieve an overthrow but also the risk connected to its realization. Note that an overthrow (as we model it) does not change the basic *structure* of the social order.[10] After an overthrow, the players may find themselves in different roles, but the roles *per se* remain the same. Keeping the structure of the social order constant, allows us to distill the pure effect of an overthrow threat.

3.3 Procedures

Every subject participated in exactly one experimental treatment (between-subjects design). Each group of six individuals played a total of at least 50 periods. After period 50, the continuation probability dropped to 75%.[11] In each period above 50, there was thus an expected duration of 3 additional periods.[12] All periods were payoff relevant. In the data analysis, we use the first (guaranteed) 50 periods for each individual.

After each period, players received information on the outcomes of their own as well as the other encounters, in the *transfer* and *transfer-overthrow* treatments they then decided on a transfer, and finally received the list of total payoffs. In case the overthrow mechanism was triggered, participants were informed of their new status rank, just before proceeding to a new period.

The instructions were handed out in print to the subjects (see Appendix B) and read aloud, before the experiment started. To ensure full understanding, subjects had to pass extensive control questions. We implemented the experiment with as little contextual information as possible. While contextualization of decisions can sometimes be useful in experiments (Alekseev et al. 2017), decontextualization allows for better experimental control (Smith 1976, p. 278; on limitations, see Engel and Rand 2014). This is of particular importance in relation to experiments on social institutions, where experimental subjects are likely to have rich preconceptions that may affect their choices (McAdams 2015, p. 63).[13]

After the game ended, we elicited participants' (a) other-regarding preferences (Liebrand and McClintock 1988) and (b) risk and trust attitudes as commonly elicited in the German Socioeconomic Panel (SOEP) as well as (c) some socio-demographics (age, gender, number of siblings). For more details on the post-experimental tests, see Appendix C.

The present paper builds on our previous work. For the treatments *baseline* and *transfer*, we report data from Chatziathanasiou et al. (2020). Data for *overthrow* and *transfer-overthrow* is being reported for the first time. All data was collected in one wave in April 2016. The experiment was

[10] This is a deviation from the real world, where social movements (and revolutions) have indeed transformed social order (Ackerman and Rodal 2008; Stephan and Chenoweth 2008; Roberts and Garton Ash 2009; Chenoweth et al. 2011).

[11] This procedure closely follows the protocol of Camera and Casari (2009).

[12] $.75 + .75^2 + .75^3 + .75^4 + .75^5 + ... = 3$.

[13] Preconceptions about law and legitimacy can themselves be a matter of experimental investigation (see, e.g., Engel and Kurschilgen 2013; Chilton 2014). However, in this experiment, for achieving higher levels of control and generality, we deliberately avoid invoking subjects' perceptions of existing institutions.

conducted at the BonnEconLab of the University of Bonn, Germany, and computerized using the software z-Tree (Fischbacher 2007). From a database of more than 5000 people, we recruited 378 subjects[14] using hroot (Bock et al. 2014). Subjects were mainly undergraduate students from a variety of disciplines. Sessions lasted about 90 minutes and subjects earned on average 16.47 € (about 19.38 $), including a show-up fee of 4 €. During the experiment, payoffs where presented in experimental currency units (ECU), with a known exchange rate of ECU 25 = 1 €. Subjects sat in visually completely isolated cubicles.

4 Hypotheses

In order to induce more wealth-sharing from elites to non-elites, the threat of overthrow has to satisfy two conditions. First, the threat needs to be *credible*, i.e. it needs to have a high chance of materializing if it is not actively prevented. Second, the elites need to *adapt* to the threat, i.e. they have to do what is necessary to prevent the overthrow from happening.

Credible Threat. The threat of overthrow can be considered credible if the introduction of the overthrow mechanism comes with a substantial increase in miscoordination, and thus with a high likelihood of an actual overthrow being triggered. The main reason for expecting an increase in miscordination is the changed incentive structure for the low-status players when overthrows are possible. For low-status players, the expected payoff after a reallotment of ranks is (weakly) higher than the expected payoff without reallotment.

For a given period, the expected payoff of every rank depends on the probability to meet a lower-ranked or higher-ranked player, and on that player's propensity to comply with the social order, i.e. to follow the coordination device. For rank a the probability to meet a lower-ranked player and achieve the high payoff of $h = 10$ is 100%. Denoting the probability that the other player follows the coordination device by $\lambda \in [0, 1]$, player a expects a payoff of 10λ tokens in a given period. Similarly, the payoffs of the other ranks are 8.2λ for player b, 6.4λ for player c, 4.6λ for player d, 2.8λ for player e, and 1λ for the lowest-rank player f.[15] The *average rank* has expected payoffs of 5.5λ tokens. If an overthrow occurs and ranks are reset, every player has the same probability to be allotted a certain rank. In expectations, each player becomes an *average rank* after the reset. From the perspective of ranks d, e and f, becoming an *average rank* (who earns 5.5λ in a given period) is more attractive than conserving their current rank. When overthrows are possible, the lower ranks have an

[14] 96 subjects in the treatments *baseline*, *transfer* and *overthrow*, respectively, and 90 subjects in *transfer-overthrow*. Over 50 periods, we thus have a total of 4800 (or 4500) observations per treatment, clustered in 16 (or 15) statistically independent groups per treatment.

[15] The general formula for the expected payoff in one period given coordination according to the device is: $\frac{n-1}{N-1} l\lambda + \frac{N-n}{N-1} h\lambda$ with N being the number of players in the society and n being the numerical rank of a player, i.e. $n = 1$ for rank a, $n = 2$ for rank b, etc.

incentive to trigger an overthrow, by avoiding coordination on the Nash equilibrium recommended by the coordination device.

Potentially, players' risk aversion could mitigate this incentive to miscoordinate. Even for the low-status players there are risks connected to an overthrow. While the low-status players earn at least some small amounts when following the coordination device, these gains might be lost if society slides into an anarchic state. In particular, other players' propensity to follow the coordination device (λ) may drop in the wake of an overthrow, which would reduce the expected payoff after an overthrow and thus its appeal. The erosion of players' willingness to follow the device could eventually lead to the highly unattractive mixed equilibrium, in which all players ignore the recommendations of the coordination device and have expected payoffs of $\frac{10}{11}$ tokens. But as long as risk aversion does not completely nullify the miscoordination incentive (which is unlikely), the possibility of overthrow should decrease lower-ranks' willingness to follow the coordination device.

Hypothesis 1. *In the absence of transfer opportunities, the possibility of overthrow will lead to an increase in miscoordination.*

Adaptation of Elites. Historical research across the social sciences has suggested that elites engage in strategic wealth-sharing in order to preserve their privileges.[16] Underlying this 'threat of revolution hypothesis' is the behavioral conjecture that elites adapt to the threat of being overthrown. In view of a credible overthrow threat, the elites in our experiment have more to lose. In the presence of transfer opportunities, they could adapt to this threat by displaying a higher willingness to redistribute economic surplus to lower ranks. Voluntary transfers have a soothing function, as they reduce the gap between the payoffs of low-status players and high-status players (Chatziathanasiou et al. 2020). From a theoretical perspective, voluntary redistribution can be understood as a way of honoring a tacit social contract (Binmore 1998). In exchange for appropriate compensation, low-status players accept the status quo and concede privilege to whomever has higher status. If, however, high-status players do not make sufficient transfers, low-status players now have an indirect punishment tool at their disposal as they can now overthrow the current high-status players. This should discipline high-status players and motivate them to make higher transfers. Such behavior would be in line with 'threat of revolution hypothesis'.

Hypothesis 2. *In the presence of transfer opportunities, the possibility of overthrow will lead to an increase in redistributive transfers.*

[16] See Sections 1 and 2.

5 Results

We first report treatment differences on the incidence of overthrow events, and then look into transfer behavior and payoff implications. Subsequently, we zoom into the *transfer-overthrow* treatment and examine the behavior of high-status players in the face of potential overthrow in order to understand why certain groups manage to stabilize while others do not.[17]

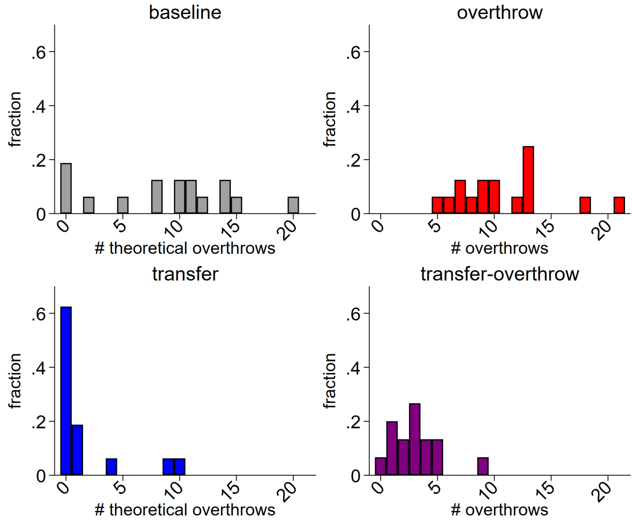


Figure 3: Incidence of Overthrow Events
Distribution of the number of overthrow events, by treatment. An overthrow event is a situation in which all matched pairs fail to coordinate on a payoff-yielding equilibrium in a given period. In treatments *overthrow* and *transfer-overthrow* the miscoordination of all three pairings triggered a reshuffling of the status order. In *baseline* and *transfer* the overthrow event is purely theoretical.

Overthrow Events. Figure 3 shows the number and distribution of overthrow events across treatments. We denote as *overthrow event* a situation in which all matched pairs fail to coordinate on a payoff-yielding equilibrium in a given period. As noted in Section 3, the miscoordination of all three pairings triggered an overthrow of the prevailing status ranking in treatments *overthrow* and *transfer-overthrow* whereas for *baseline* and *transfer* the overthrow event is purely theoretical.

In the absence of transfer opportunities, overthrow events are rather frequent. In *baseline*, only 19% of all groups never experience an overthrow event. The mere possibility of overthrow makes overthrow events, i.e. full miscoordination in a given period, even more likely (p = 0.073). In the *overthrow* treatment there is not a single group that never experiences an overthrow. Put differently, in the absence of transfers, the probability that a group experiences at least one overthrow, and thus has its status order reset, equals 100%. The threat of overthrow is credible and the first condition for a disciplining effect is fulfilled, yielding support for Hypothesis 1:

[17] Unless stated differently, for comparisons between treatments we report two-sided Wilcoxon rank-sum tests over group-level means. Summary statistics for all figures can be found in the Appendix A.

Result 1. *In the absence of redistribution, overthrows occur in all groups.*

In principle, appropriate redistributive transfers could prevent overthrows from happening. And indeed, over the course of 50 periods the average group experiences 8.8 (theoretical) overthrow events in *baseline* but only 1.6 (theoretical) overthrow events in *transfer* ($p < 0.001$). Similarly, the average number of (real) overthrows drops from 10.9 in *overthrow* to 3.1 in *transfer-overthrow* ($p < 0.001$).

However, also in the presence of transfer opportunities we do find overthrow events to become more likely when the threat of overthrow is real. Whereas in *transfer*, the share of groups that never experience a single overthrow event is 63%, in *transfer-overthrow* this share drops to merely 7% ($p < 0.001$). Put differently, in 93% of groups in *transfer-overthrow* the initial high ranks were unable to sustain the status order and consequently lost their high status.

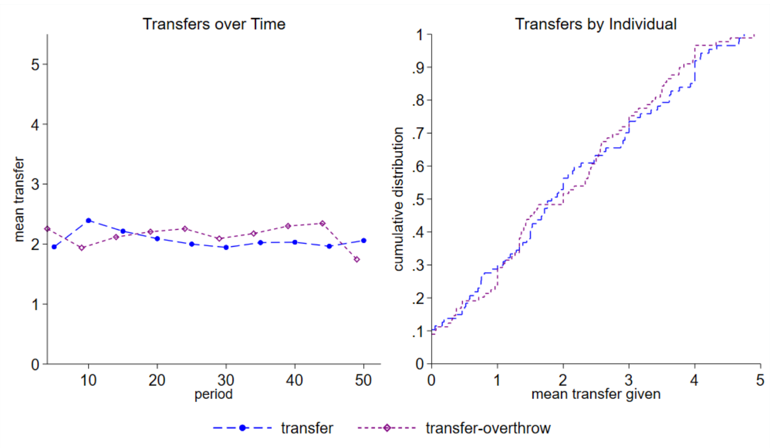


Figure 4: Transfers

Mean transfers over time and cumulative distribution of mean transfers given by each individual.

Transfers and Payoff Implications. As stated in Section 4, the second condition for a disciplining effect requires high status players to adapt to the threat of being overthrown. Figure 4 shows mean transfers, over time (left panel) and by individual (right panel).

According to the threat of revolution hypothesis, high status players should react to the possibility of overthrow by increasing their transfer levels in order to appease the low ranks and prevent overthrows from happening. But the level of transfers in *transfer* (1.99) and *transfer-overthrow* (2.06) is virtually identical ($p = 0.843$). Thus, we find no support for Hypothesis 2 and note:

Result 2. *The threat of overthrow does not increase high-status players' willingness to make redistributive transfers.*

Figure 5 compares the payoff development over time in all treatments.[18] We distinguish between players who were allotted a high rank (a, b, c) at the beginning of the game, and those who

[18] Average payoffs over all ranks do not follow a noticeable time trend, see Figure A1.

were initially allotted a low rank (d, e, f). In the absence of transfer opportunities, the overthrow mechanism acts as an equalizer, leveling the differences in payoffs among the players. Initial lower ranks earn significantly more in *overthrow* than in *baseline* (p = 0.006) whereas initial upper ranks earn significantly less when overthrows are possible (p = 0.004). The Gini coefficient drops significantly from .30 in *baseline* to .18 in *overthrow* (p < 0.001).

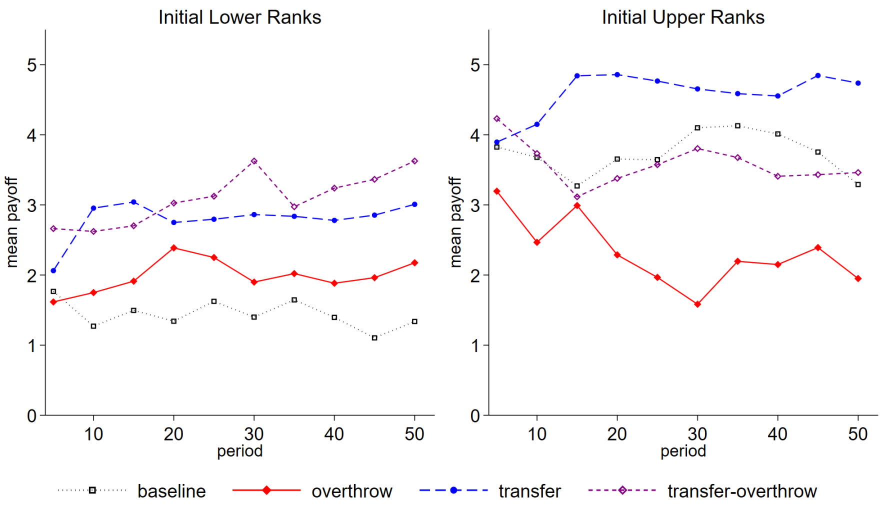


Figure 5: Payoffs over Time by Initial Rank

Mean payoffs (after transfers) over time. We distinguish between players who were allotted a high rank (a, b, c) at the beginning of the game, and those who were initially allotted a low rank (d, e, f).

The overthrow mechanism still acts as an equalizer in the presence of transfer opportunities. Despite the fact that, in principle, voluntary transfers would be a more efficient substitute for that purpose. The Gini coefficient drops from .18 in *transfer* to .15 in *transfer-overthrow* (p = 0.058). Underscoring the inefficiency, the substantial payoff drop of the initial upper ranks from 4.59 in *transfer* to 3.58 in *transfer-overthrow* (p = 0.010) is not matched by a payoff increase of the initial lower ranks (p = 0.441). As illustrated in the right-hand panel of Figure 5, the payoffs of the initial upper ranks start off at a similar level (around 4.0) in both transfer treatments but then diverge over time. While they increase to about 4.9 in *transfer* as coordination improves, they decrease to about 3.4 in *transfer-overthrow* as overthrows occur and initial high ranks lose their privileged position.

Stable vs. Unstable Groups. When overthrows occur, high status players lose their status as ranks are reset, and experience a substantial drop in earnings. Almost all groups in *overthrow* and *transfer-overthrow* experience (one or more) overthrows in the first half of the game (see Figure A2). In the absence of transfers, all groups continue to experience overthrow events throughout the game. But in the presence of transfer opportunities, some groups become perfectly stable in the second half of the game. Figure 6 shows the period in which an overthrow occurs *for the last time* in a given group. In the following, we explore the differences between the 6 groups from the *transfer-overthrow* treatment that do not experience an overthrow after period 25 (*stable groups*) and the 9 groups that experience one or more overthrows after period 25 (*unstable groups*).[19]

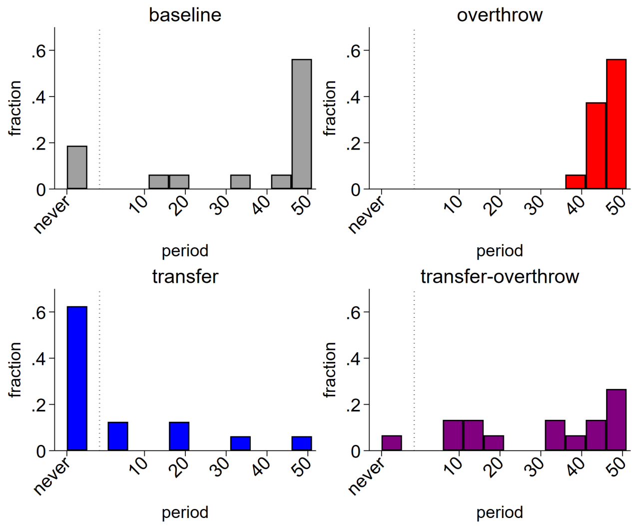


Figure 6: Period of Last Overthrow Event

Distribution of the period of the last overthrow event in a given group, by treatment. An overthrow event is a situation in which all matched pairs fail to coordinate on a payoff-yielding equilibrium in a given period. In treatments *overthrow* and *transfer-overthrow* the miscoordination of all three pairings triggered a reshuffling of the status order. In *baseline* and *transfer* the overthrow event is purely theoretical.

What distinguishes stable groups from unstable groups in the *transfer-overthrow* treatment? Two plausible mechanisms come to mind. On the one hand, stable groups could simply be the result of successful observational learning from the mistakes of earlier high-ranks. Initial low ranks may observe that transfers of the initial high ranks are too low to prevent overthrows. When eventually put into a high-status position in the aftermath of an overthrow, those players start making more generous transfers for *strategic* reasons, i.e. in order to preserve their status. On the other hand, the stable groups may just benefit from the fact that overthrows put, eventually, more *intrinsically* generous players into high-status positions.

To shed light into the mechanism, we take advantage of the fact that (a) the group formation, (b) the allotment (and reallotment) of ranks within a group, and (c) the pairing of players in a given period were exogenous. In addition, our design allows us to not only observe the generosity of high ranks (a, b, c) but also of low ranks (d, e, f). Whenever a player earned 10 tokens in the stage game, she could transfer up to 10 tokens to the other player. The lower a player's rank, the less frequently she earned 10 (because she was usually matched with higher-ranked players), but from time to time lower ranks did earn 10, and could thus display their generosity.

Figure 7 shows average transfers over time, distinguishing between groups that were stable (i.e. experienced no overthrow in periods 26-50) or unstable in the second half of the game, and

[19] For comparisons between stable and unstable groups we report two-sided Wilcoxon rank-sum tests over individual-level means.

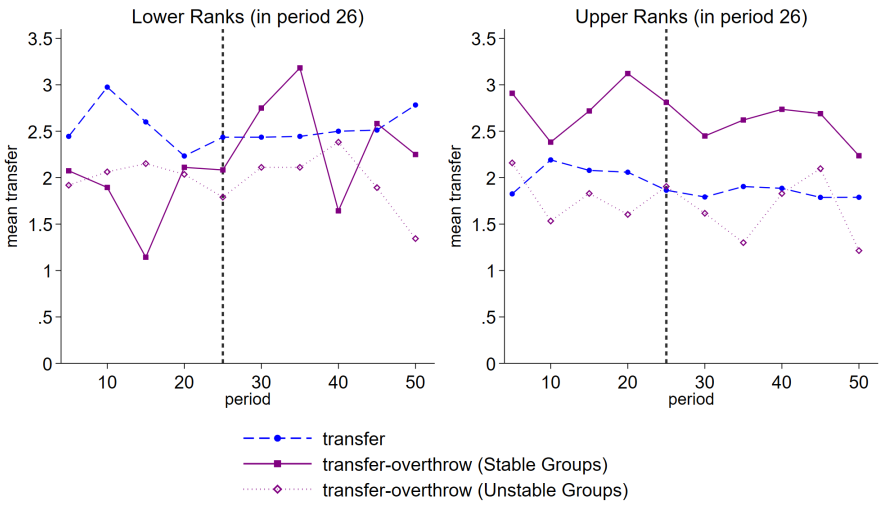


Figure 7: Stability and Transfers

Mean transfers. Of the 15 groups in *transfer-overthrow* 6 are stable (i.e. no overthrow after period 25), and 9 are unstable (i.e. at least one overthrow after period 25).

between players who were high ranks or low ranks at the beginning of the second half (i.e. in period 26). For the lower ranks, there are no systematic differences. But for the upper ranks, there are two important differences. First, in the second half of the game, the high ranks of the stable groups are significantly ($p = 0.022$) more generous (transferring on average 2.45 tokens) than the high ranks of unstable groups (1.38 tokens). Stability thus emerges in groups with systematically higher levels of redistribution. Second, the high-status players of the stable groups did not learn to become more generous by observing the failure of the initial high ranks. Rather, they were already more generous *before* they became high-status players. In the first half of the game, the later-to-be high-status players of the stable groups transfer on average 2.66 tokens whereas the later-to-be high-status players of the unstable groups transfer on average 1.80 tokens ($p = 0.039$).

These findings are confirmed by regression analysis. Column 1 of Table 2 shows that the probability of a group becoming stable in the second half of the game increases by 10.3 percentage points ($p = 0.030$) when a player who is of high-status in period 26 gave on average 1 token more in the first half of the game. Column 2 shows that a player's generosity early in the game predicts her generosity late in the game. A player who gives 1 token more in the first half gives, on average, .755 tokens more in the second half ($p < 0.001$). Column 3 shows that generosity is mainly associated with higher levels of individual trust towards strangers ($p = 0.058$) measured in the post-experimental tests. Indeed, trust has been shown to be a critical ingredient of economic development (Algan and Cahuc 2010; Bigoni et al. 2016) and the stability of institutions (Alesina and Giuliano 2015).

Result 3. *The new elites do not learn from the mistakes of the old elites. Stability emerges when intrinsically generous players are (accidentally) allotted high-status positions.*

Table 2: Determinants of Stable Groups

	(1) DV: *stable group*	(2) DV: *second-half transfers*	(3) DV: *first-half transfers*
first-half transfers	0.103** (0.05)	0.755*** (0.13)	
ringdegree			0.589 (0.46)
trust			1.511* (0.73)
risk			-0.336 (0.36)
gender			-0.267 (0.26)
siblings			0.186 (0.17)
age			0.0129 (0.04)
Constant		0.184 (0.22)	2.431*** (0.43)
Observations	45	45	45
Clusters	15	15	15
R^2		0.577	0.132

Column (1) shows marginal effects of probit regressions. Columns (2) and (3) show coefficients of OLS regressions. Cluster-robust standard errors in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Who benefits from stability? Figure 8 shows average payoffs over time, distinguishing between stable and unstable groups, and between players who were high ranks and players who were low ranks at the beginning of the second half (i.e. in period 26). We find that, unsurprisingly, being in a stable group is beneficial for the lower ranks as they receive higher transfers. After period 25, low-status players in the stable groups have significantly higher earnings (unstable 2.66; stable 3.35; p = 0.046). Strikingly, however, being generous pays off for the high-status players, too. By sharing a larger part of their wealth through voluntary redistribution, the high ranks of stable groups not only manage to sustain their status, they actually end up earning significantly more than the high rank players in the unstable groups. After period 25, the mean payoff for high-status players in a stable group is 5.01, as compared to 3.31 in an unstable group (p < 0.001).

Result 4. *High status players' generosity leads to stability and higher payoffs both for low-status and for high-status players.*

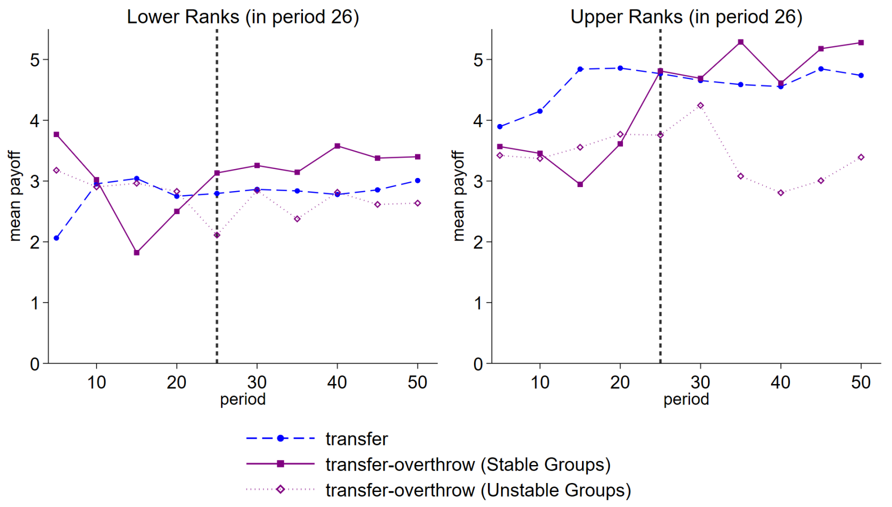


Figure 8: Stability and Payoffs

Mean payoffs after transfers. Of the 15 groups in *transfer-overthrow* 6 are stable (i.e. no overthrow after period 25), and 9 are unstable (i.e. at least one overthrow after period 25).

6 Discussion and Outlook

Despite being credible, the threat of overthrow did, on average, *not* have a disciplining effect on the elites. In most experimental groups, high-status players did not adapt their transfer behavior to the threat of overthrow. As a result, overthrows happened and the initial elites lost their privileged position. However, in the second half of the game *some* groups stabilized and prospered. In these groups, the overthrow mechanism brought – eventually, through the reshuffling of ranks – players into high-status positions who were intrinsically generous. The price of stability was not too high for those generous high ranks. To the contrary, they earned significantly more than the (less generous) high ranks of the unstable groups.

Our results suggest that elites are more myopic than the threat of revolution hypothesis assumes them to be. This resonates with the historical observation that ruling elites, particularly in authoritarian regimes, tend to be overconfident (Goldstone 2001, p. 162). Of course, stylized laboratory experiments do not fully map the rich context of the real world. They provide a complement to studies with observational data by allowing to model social phenomena and test behavioral theories (Schmidt 2009; Falk and Heckman 2009; in the context of institutions, see Camerer and Talley 2007; Engel 2013). While concerns over external validity (Voigt 2011, p. 247) call for a discussion of experimental results, these concerns pose no general argument against experiments. A solid body of research has documented that experimental choices have predictive power for the real world (Rustagi et al. 2010; Sausgruber and Tyran 2011; Esarey et al. 2012; Durante et al. 2014; Kosfeld and Rustagi 2015). Revealed preferences for redistribution elicited in the laboratory are largely in line with observational field data (Fong 2001; Alesina and Angeletos 2005; Alesina and Giuliano 2011).

This holds not only for the general population but also for the economic elite (Cohn et al. 2019). In a recent study in Switzerland, laboratory behavior has been shown to predict individuals' support for redistributive policies in national plebiscites (Epper et al. 2020). Therefore, we are confident that our results, while not disproving the threat of revolution hypothesis historically, have predictive power for the real world and challenge the notion that elites engage in strategic wealth-sharing.

Further, while we did not find that the threat of overthrow disciplines the elites, we found that overthrows can function as a selection mechanism. In the experimental groups that ended up being stable and prosperous, elites where exchanged via the overthrow mechanism until eventually, generous players were thrust into high-status positions. In the real world, old elites can resort to authoritative means to prevent such change. In our experiment, this was excluded by design. On the one hand, this only underlines the myopia of the old elites. In a setting with no coercive means, elites should be even more sensitive to the threat of overthrow. On the other hand, the finding that – absent coercion – overthrow functions as a screening device resonates with real world experiences. Tunesia is the only case among the countries of the Arab Spring that so far has been considered a success (Willis 2016; Roberts 2016). Here, recurring protests continued even after the ousting of former President Ben Ali and subsequent governments agreed to peaceful transitions of power.[20] Arguably, the recurring protests functioned as a screening device that influenced the composition of subsequent governments, until a reform-oriented constellation was found.[21]

Our finding of elite myopia also has implications for the debate on inequality and constitutionalism. Our results suggest that it might not be enough to rely on elite's goodwill to form equitable and thus robust institutions. Given the lack of political responsiveness to the demands of the economically disadvantaged (Gilens 2012), the establishment of social rights has been proposed as a means to restrain elite behavior more strongly (Young 2012; Peters 2017; Moyn 2018). If effective (on doubts, see Chilton and Versteeg 2017), such guarantees might lead to more robust institutions by incentivizing citizens to protect the constitution against transgressions.

Finally, further research should look into constitutional provisions that help or hinder coordination on collective disobedience. Rights to resistance, for instance, have been understood to help stabilize a constitution by reminding the people of their collective power to constrain the rulers (Ginsburg et al. 2012; Pilpilidis 2015; Gutmann and Voigt 2021). Their inclusion in a constitutional document functions as a precommitment device which is supposed to discipline the governing elites (Ginsburg et al. 2012, p. 1184). However, this understanding rests on the conjecture that the threat of overthrow does indeed have a disciplining effect – a conjecture that did not hold in our experimental

[20] In 2015, the reform-oriented negotiation parties were even awarded the Nobel Peace Prize (The Norwegian Nobel Committee 2015).

[21] Of course, the case of Tunesia can again be qualified as special. A nuanced description of the case points to an "attachment to values of procedure, process, and forms of legality [...] that marks out Tunisia and explains the success of forms of civil resistance" (Willis 2016, p. 52).

setting. Yet, rights to resistance might facilitate the behavior of the 'protesting' non-elites, allowing them to put protesting behavior and overthrow threats to socially beneficial use.

References

Acemoglu, D. and J. A. Robinson (2000). Why did the west extend the franchise? Democracy, inequality, and growth in historical perspective. *Quarterly Journal of Economics 115*(4), 1167–1199.

Acemoglu, D. and J. A. Robinson (2006). *Economic Origins of Dictatorship and Democracy*. Cambridge University Press.

Ackerman, P. and B. Rodal (2008). The strategic dimensions of civil resistance. *Survival. Global Politics and Strategy 50*(3), 111–126.

Adermon, A., M. Lindahl, and D. Waldenström (2018). Intergenerational wealth mobility and the role of inheritance: Evidence from multiple generations. *Economic Journal 128*(612), F482–F513.

Aidt, T. S. and R. Franck (2015). Democratization under the threat of revolution: Evidence from the Great Reform Act of 1832. *Econometrica 83*(2), 505–547.

Aidt, T. S. and P. 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, T. S. and G. Leon (2016). The democratic window of opportunity: Evidence from riots in Sub-Saharan Africa. *Journal of Conflict Resolution 60*(4), 694–717.

Alekseev, A., G. Charness, and U. Gneezy (2017). Experimental methods: When and why contextual instructions are important. *Journal of Economic Behavior & Organization 134*, 48–59.

Alesina, A. and G.-M. Angeletos (2005). Fairness and redistribution. *American Economic Review 95*(4), 960–980.

Alesina, A. and P. Giuliano (2011). Preferences for redistribution. In *Handbook of Social Economics*, Volume 1, pp. 93–131. Elsevier.

Alesina, A. and P. Giuliano (2015). Culture and institutions. *Journal of Economic Literature 53*(4), 898–944.

Algan, Y. and P. Cahuc (2010). Inherited trust and growth. *American Economic Review 100*(5), 2060–92.

Anbarcı, N., N. Feltovich, and M. Y. Gürdal (2018). Payoff inequity reduces the effectiveness of correlated-equilibrium recommendations. *European Economic Review 108*, 172–190.

Bardsley, N., J. Mehta, C. Starmer, and R. Sugden (2009). Explaining focal points: Cognitive hierarchy theory versus team reasoning. *Economic Journal 120*(543), 40–79.

Bednar, J. (2009). *The Robust Federation. Principles of Design*. Cambridge University Press.

Bigoni, M., S. Bortolotti, M. Casari, D. Gambetta, and F. Pancotto (2016). Amoral familism, social capital, or trust? The behavioural foundations of the Italian North–South divide. *Economic Journal 126*(594), 1318–1341.

Binmore, K. (1998). *Game Theory and the Social Contract: Just Playing (Vol. 2)*. MIT Press.

Bock, O., I. Baetge, and A. Nicklisch (2014). hroot: Hamburg registration and organization online tool. *European Economic Review 71*, 117–120.

Boix, C. (2003). *Democracy and Redistribution*. Cambridge University Press.

Bowles, S. and H. Gintis (2002). The inheritance of inequality. *Journal of Economic Perspectives 16*(3), 3–30.

Brückner, M. and A. Ciccone (2011). Rain and the democratic window of opportunity. *Econometrica 79*(3), 923–947.

Burke, P. J. and A. Leigh (2010). Do output contractions trigger democratic change? *American Economic Journal: Macroeconomics 2*(4), 124–57.

Camera, G. and M. Casari (2009). Cooperation among strangers under the shadow of the future. *American Economic Review 99*(3), 979–1005.

Camerer, C. and E. Talley (2007). Experimental study of law. In A. M. Polinsky and S. Shavell (Eds.), *Handbook of Law and Economics, Vol. 2*, pp. 1619–1650. Elsevier.

Chaney, E. (2013). Revolt on the nile: Economic shocks, religion, and political power. *Econometrica 81*(5), 2033–2053.

Chatziathanasiou, K., S. Hippel, and M. Kurschilgen (2020). Property, redistribution, and the status quo: A laboratory study. *Experimental Economics*, 1–33.

Chenoweth, E., M. J. Stephan, and M. Stephan (2011). *Why Civil Resistance Works: The Strategic Logic of Nonviolent Conflict*. Columbia University Press.

Chetty, R., N. Hendren, P. Kline, and E. Saez (2014). Where is the land of opportunity? The geography of intergenerational mobility in the United States. *Quarterly Journal of Economics 129*(4), 1553–1623.

Chilton, A. S. (2014). The influence of international human rights agreements on public opinion: An experimental study. *Chicago Journal of International Law 15*, 110–137.

Chilton, A. S. and M. Versteeg (2017). Rights without resources: The impact of constitutional social rights on social spending. *Journal of Law and Economics 60*(4), 713–748.

Cloward, R. A. and F. F. Piven (1971). *Regulating the Poor: The Functions of Public Welfare*. New York: Pantheon Books.

Cohn, A., L. Jessen, M. Klasnja, and P. Smeets (2019). Why do the rich oppose redistribution? An experiment with America's top 5%. *mimeo*.

Congleton, R. D. (2010). *Perfecting Parliament: Constitutional Reform, Liberalism, and the Rise of Western Democracy*. Cambridge University Press.

Crawford, V., U. Gneezy, and Y. Rottenstreich (2008). The power of focal points is limited: Even minute payoff asymmetry may yield large coordination failures. *American Economic Review*, 1443–1458.

Dal Bó, P., A. Foster, and L. Putterman (2010). Institutions and behavior: Experimental evidence on the effects of democracy. *American Economic Review 100*(5), 2205–2230.

Dasgupta, A. and D. Ziblatt (2015). How did britain democratize? Views from the sovereign bond market. *Journal of Economic History*, 1–29.

de La Boétie, É. (1577). *Discours de la Servitude Volontaire* (The Politics of Obedience: The Discourse of Voluntary Servitude. 1975, translated by Harry Kurz ed.). Free Life Editions.

DeAngelo, G. J., D. Dubois, and R. Romaniuc (2020). The perils of democracy. *Journal of Economic Behavior & Organization 175*, 328–340.

Dixon, R. and J. Suk (2018). Liberal constitutionalism and economic inequality. *University of Chicago Law Review 85*, 369–401.

Durante, R., L. Putterman, and J. van der Weele (2014). Preferences for redistribution and rerception of fairness: An experimental study. *Journal of the European Economic Association 12*(4), 1059–1086.

Elkins, Z., T. Ginsburg, and J. Melton (2007). Baghdad, Tokyo, Kabul...: Constitution making in occupied states. *William & Mary Law Review 49*, 1139–1178.

Elster, J. (1984). *Ulysses and the Sirens. Studies in Rationality and Irrationality*. Cambridge University Press.

Elster, J. (1992). *Local Justice: How Institutions Allocate Scarce Goods and Necessary Burdens*. Russell Sage Foundation.

Engel, C. (2013). *Legal Experiments: Mission Impossible?* Eleven International Publishing.

Engel, C. and M. Kurschilgen (2013). The coevolution of behavior and normative expectations: An experiment. *American Law and Economics Review 15*(2), 578–609.

Engel, C. and D. G. Rand (2014). What does "clean" really mean? The implicit framing of decontextualized experiments. *Economics Letters 122*(3), 386–389.

Epper, T., E. Fehr, and J. Senn (2020). Other-regarding preferences and redistributive politics. *mimeo*.

Esarey, J., T. C. Salmon, and C. Barrilleaux (2012). What motivates politicalpreferences? Self-interest, ideology, and fairness in a laboratory democracy. *Economic Inquiry 50*(3), 604–624.

Eurostat and World Bank (2020, December). *Europäische Union: Ranking der Mitgliedsländer und Beitrittskandidaten nach Ungleichheit bei der Einkommensverteilung auf Basis des Gini-Index im Jahr 2019, Chart 2*. Statista. Available at https://de.statista.com/statistik/daten/studie/942729/umfrage/ranking-der-eu-laender-nach-einkommensungleichheit-im-gini-index/. Accessed on July 25, 2021.

Evans, R. J. (2016). *The Pursuit of Power: Europe 1815-1914*. Penguin.

Falk, A. and J. J. Heckman (2009). Lab experiments are a major source of knowledge in the social sciences. *Science 326*(5952), 535–538.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics 10*(2), 171–178.

Fong, C. (2001). Social preferences, self-interest, and the demand for redistribution. *Journal of Public Economics 82*(2), 225–246.

Gilens, M. (2012). *Affluence and Influence*. Princeton University Press.

Ginsburg, T., D. Lansberg-Rodriguez, and M. Versteeg (2012). When to overthrow your government: The right to resist in the world's constitutions. *UCLA Law Review 60*, 1184–1260.

Goldstone, J. A. (1982). The comparative and historical study of revolutions. *Annual Review of Sociology 8*(1), 187–207.

Goldstone, J. A. (2001). Toward a fourth generation of revolutionary theory. *Annual Review of Political Science 4*(1), 139–187.

Goldstone, J. A. (2016). *Revolution and Rebellion in the Early Modern World: Population Change and State Breakdown in England, France, Turkey, and China, 1600-1850*. Routledge.

Grossman, H. I. (1995). Robin Hood and the redistribution of property income. *European Journal of Political Economy 11*(3), 399–410.

Gutmann, J. and S. Voigt (2021). Militant constitutionalism: A promising concept to make constitutional backsliding less likely? *Public Choice*, 1–28.

Hardin, R. (1989). Why a constitution? In B. Grofman and D. Wittman (Eds.), *The Federalist Papers and the New Institutionalism*, pp. 100–120. Agathon Press.

Hauser, O., D. Rand, A. Peysakhovich, and M. Nowak (2014). Cooperating with the future. *Nature 511*(7508), 220–223.

Hobsbawm, E. (1987). *Age of Empire: 1875-1914*. Weidenfeld & Nicolson.

Kahlenberg, R. D. (2010). *Affirmative Action for the Rich: Legacy Preferences in College Admissions*. Century Foundation Press New York, NY.

Khaitan, T. (2019). Political insurance for the (relative) poor: How liberal constitutionalism could resist plutocracy. *Global Constitutionalism 8*(3), 536–570.

Kim, W. (2007). Social insurance expansion and political regime dynamics in Europe, 1880-1945. *Social Science Quarterly 88*(2), 494–514.

Kosfeld, M. and D. Rustagi (2015). Leader punishment and cooperation in groups: Experimental field evidence from commons management in Ethiopia. *American Economic Review 105*(2), 747–83.

Langenbach, P. and F. Tausch (2019). Inherited institutions: Cooperation in the light of democratic legitimacy. *Journal of Law, Economics, and Organization 35*(2), 364–393.

Levitsky, S. and D. Ziblatt (2018). *How Democracies Die*. Crown.

Liebrand, W. and C. McClintock (1988). The ring measure of social values: A computerized procedure for assessing individual differences in information processing and social value orientation. *European Journal of Personality 2*(3), 217–230.

Lizzeri, A. and N. Persico (2004, 05). Why did the elites extend the suffrage? Democracy and the scope of government, with an application to Britain's "Age of Reform". *Quarterly Journal of Economics 119*(2), 707–765.

Markussen, T., L. Putterman, and J.-R. Tyran (2014). Self-organization for collective action: An experimental study of voting on sanction regimes. *Review of Economic Studies 81*(1), 301–324.

McAdams, R. (2015). *The Expressive Powers of Law: Theories and Limits*. Harvard University Press.

Mehta, J., C. Starmer, and R. Sugden (1994). The nature of salience: An experimental investigation of pure coordination games. *American Economic Review 84*(3), 658–673.

Moulin, H. (2004). *Fair Division and Collective Welfare*. MIT Press.

Moyn, S. (2018). *Not Enough: Human Rights in an Unequal World*. Harvard University Press.

Muller, E. N. and M. A. Seligson (1987). Inequality and insurgency. *American Political Science Review*, 425–451.

Peters, A. (2017). Global constitutionalism: The social dimension. *Global Constitutionalism from European and East Asian Perspectives (Cambridge: CUP 2018), Max Planck Institute for Comparative Public Law & International Law (MPIL) Research Paper* (2017-25).

Pilpilidis, K. (2015). When rock trumps paper: A theory of constitutional resistance. Available at SSRN 2624459.

Pistor, K. (2019). *The Code of Capital: How the Law Creates Wealth and Inequality*. Princeton University Press.

Przeworski, A. (2009). Conquered or granted? A history of suffrage extensions. *British Journal of Political Science*, 291–321.

Roberts, A. (2016). Civil resistance and the fate of the arab spring. In A. Roberts, M. J. Willis, R. McCarthy, and T. G. Ash (Eds.), *Civil Resistance in the Arab Spring: Triumphs and Disasters*, pp. 271–325. Oxford University Press.

Roberts, A. and T. Garton Ash (2009). *Civil Resistance and Power Politics: The Experience of Non-violent Action from Gandhi to the Present*. Oxford University Press.

Rustagi, D., S. Engel, and M. Kosfeld (2010). Conditional cooperation and costly monitoring explain success in forest commons management. *Science 330*(6006), 961–965.

Ryvkin, D. and A. Semykina (2017). An experimental study of democracy breakdown, income and inequality. *Experimental Economics 20*(2), 420–447.

Sausgruber, R. and J.-R. Tyran (2011). Are we taxing ourselves? *Journal of Public Economics 95*(1-2), 164–176.

Scheidel, W. (2017). *The great leveler: Violence and the history of inequality from the stone age to the twenty-first century*. Princeton University Press.

Schelling, T. (1960). *The Strategy of Conflict*. Harvard University Press.

Schmidt, K. M. (2009). The role of experiments for the development of economic theories. *Perspektiven der Wirtschaftspolitik 10*, 14–30.

Schotter, A. and B. Sopher (2003). Social learning and coordination conventions in intergenerational games: An experimental study. *Journal of Political Economy 111*(3), 498–529.

Shachar, A. (2009). *The Birthright Lottery: Citizenship and Global Inequality*. Harvard University Press.

Sitaraman, G. (2017). *The Crisis of the Middle-Class Constitution. Why Economic Inequality Threatens our Republic*. Knopf.

Smith, V. L. (1976). Experimental economics: Induced value theory. *American Economic Review 66*(2), 274–279.

Stephan, M. J. and E. Chenoweth (2008). Why civil resistance works: The strategic logic of nonviolent conflict. *International Security 33*(1), 7–44.

Sutter, M., S. Haigner, and M. Kocher (2010). Choosing the carrot or the stick? Endogenous institutional choice in social dilemma situations. *Review of Economic Studies 77*(4), 1540–1566.

The Norwegian Nobel Committee (October 10, 2015). The Nobel Peace Prize for 2015. Available at https://www.nobelprize.org/prizes/peace/2015/press-release/. Accessed on July 25, 2021.

Tyran, J.-R. and L. Feld (2006). Achieving compliance when legal sanctions are non-deterrent. *Scandinavian Journal of Economics 108*(1), 135–156.

Voigt, S. (2011). Positive constitutional economics II? A survey of recent developments. *Public Choice 146*(1), 205–257.

Weingast, B. (1997). The political foundations of democracy and the rule of the law. *American Political Science Review 91*(2), 245–263.

Willis, M. J. (2016). Revolt for dignity, Tunisia's revolution and civil resistance. In A. Roberts, M. J. Willis, R. McCarthy, and T. G. Ash (Eds.), *Civil Resistance in the Arab Spring: Triumphs and Disasters*, pp. 30–52. Oxford University Press.

Young, K. (2012). *Constituting Economic and Social Rights*. Oxford University Press.

A Additional Results and Summary Statistics

Figure A1 depicts the means of overall payoffs over time. They do not follow a notable time trend.

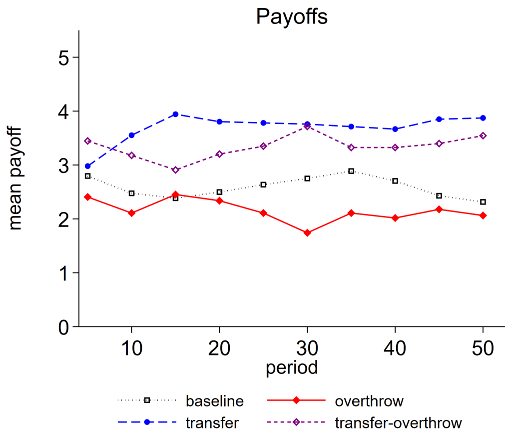


Figure A1: Payoffs over Time

Mean payoffs (after transfers).

Figure A2 shows when an overthrow occurs for the first time. It thus shows how long the initial social order holds up. Almost all groups experience overthrows in the first half of the game.

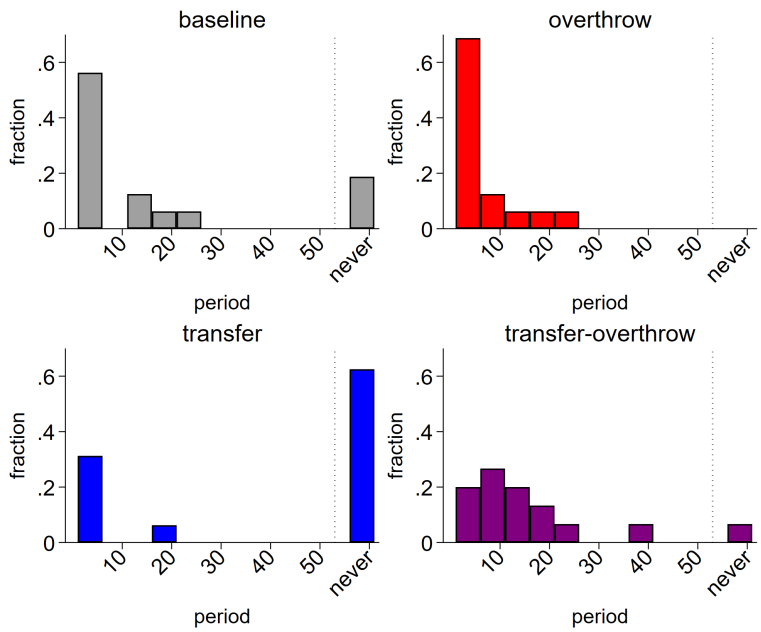


Figure A2: Period of First Overthrow Event

Distribution of the period of the first overthrow event in all groups of a treatment. An overthrow event is a situation in which all matched pairs fail to coordinate on a payoff-yielding equilibrium in a given period.

	baseline	*overthrow*	*transfer*	*transfer-overthrow*
number of overthrow events	8.75	10.88	1.63	3.07
	(5.97)	(4.30)	(3.24)	(2.22)
period of first overthrow event	15.75	6.31	34	14.6
	(18.59)	(6.23)	(22.91)	(13.55)
period of last overthrow event	43.75	46.44	39.25	32.8
	(12.42)	(3.20)	(18.81)	(16.38)
transfers	-	-	1.99	2.06
	-	-	(.75)	(.9)
payoffs (after transfers)	2.59	2.15	3.69	3.34
	(.93)	(.36)	(.9)	(.5)
initial lower ranks	1.44	1.99	2.79	3.1
	(.46)	(.5)	(.74)	(.46)
initial upper ranks	3.74	2.32	4.59	3.58
	(1.59)	(.69)	(1.3)	(.78)
Gini coefficients	.30	.18	.18	.15
	(.09)	(.08)	(.06)	(.06)

Table A1: Summary Statistics

Group-level means. For average *period of first overthrow event* and *period of last overthrow event*, we assume groups that *never* experience an overthrow to experience their first overthrow in period 51. Standard deviations in parentheses.

	transfer	*transfer-overthrow*	
		unstable groups	*stable groups*
transfers	2.02	1.78	2.30
	(1.42)	(1.24)	(1.42)
lower ranks (in period 26)	2.28	1.94	2.05
	(1.43)	(1.39)	(1.33)
upper ranks (in period 26)	1.81	1.63	2.54
	(1.39)	(1.07)	(1.5)
first half transfers	2.22	1.88	2.3
	(1.38)	(1.43)	(1.41)
lower ranks (in period 26)	2.54	1.96	1.88
	(1.29)	(1.6)	(1.23)
upper ranks (in period 26)	1.98	1.80	2.66
	(1.41)	(1.27)	(1.48)
second half transfers	1.89	1.69	2.33
	(1.68)	(1.26)	(1.56)
lower ranks (in period 26)	2.18	2.02	2.16
	(1.91)	(1.35)	(1.61)
upper ranks (in period 26)	1.68	1.38	2.45
	(1.47)	(1.12)	(1.56)

Table A2: Summary Statistics for Stability and Transfers

Mean transfers at individual-level. First half of the game are periods 1–25, second half are periods 26–50. Standard deviations in parentheses. 6 of 15 groups in *transfer-overthrow* are stable (i.e. no overthrow after period 25).

	transfer	*transfer-overthrow*	
		unstable groups	*stable groups*
payoffs (after transfers)	3.69	3.08	3.72
	(1.55)	(.84)	(1.3)
lower ranks (in period 26)	2.79	2.73	3.1
	(.97)	(.78)	(1.13)
upper ranks (in period 26)	4.59	3.44	4.34
	(1.50)	(.75)	(1.17)
first-half payoffs (after transfers)	3.61	3.19	3.26
	(1.66)	(1.17)	(1.22)
lower ranks (in period 26)	2.72	2.8	2.85
	(1.16)	(1.07)	(.95)
upper ranks (in period 26)	4.50	3.57	3.68
	(1.62)	(1.16)	(1.34)
second-half payoffs (after transfers)	3.77	2.98	4.18
	(1.65)	(1.08)	(1.64)
lower ranks (in period 26)	2.87	2.66	3.35
	(1.08)	(.96)	(1.45)
upper ranks (in period 26)	4.68	3.31	5.01
	(1.63)	(1.12)	(1.40)

Table A3: Summary Statistics for Stability and Payoffs

Mean payoffs (after transfer) at individual-level. First half of the game are periods 1–25, second half are periods 26–50.

Standard deviations in parentheses. 6 of 15 groups in *transfer-overthrow* are stable (i.e. no overthrow after period 25).

B Instructions

Note: The text below shows the instructions for the ***baseline*** *treatment, on which all other treatments build. The additional text in blue was included in the instructions for the* ***overthrow*** *as well as the* ***overthrow-transfer*** *treatment. The additional text in red was included in the instructions for the* ***transfer*** *as well as the* ***overthrow-transfer*** *treatment. Instructions displayed here are a translation into English.*[22] *Original instructions where in German and are available from the authors upon request.*

Welcome to our experiment!

If you read the following instructions carefully, you can earn a substantial sum of money, depending on your decisions. It is therefore very important that you read these instructions carefully.

Absolutely no communication with the other participants is allowed during the experiment. Anyone disobeying this rule will be excluded from the experiment and all payments. Should you have any questions, please raise your hand. We will then come to you.

During the experiment, we will speak not of Euros, but of points. Your entire income will therefore initially be calculated in points. The total number of points accumulated by you during the experiment will be paid out to you in Euros at the end, at a rate of:

25 points = 1 Euro.

At the end of the experiment, you will be paid, in cash, the number of points you will have earned during the experiment. In addition to this sum, you will receive payment of 4 Euro for showing up at this experiment.

The experiment consists of at least 50 periods.

After period 50, a draw will decide in each period whether there shall be a further period. With a probability of 75%, there will be a period 51. Should there be a period 51, there will be a period 52 as well, once again with a probability of 75%, etc.

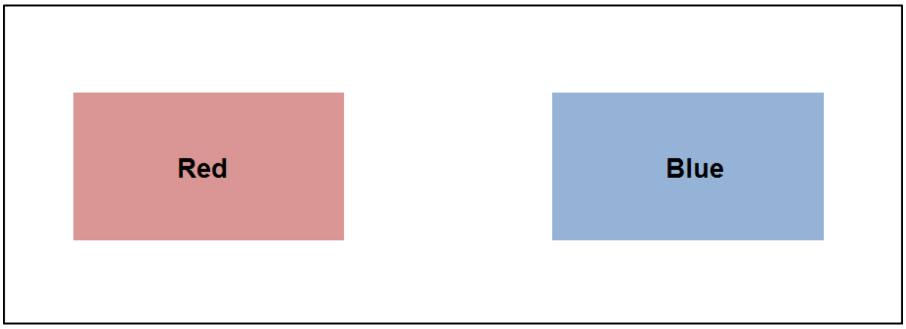
At the beginning of the experiment, participants will be randomly divided into groups of six. Apart from you, your group will therefore be made up of another 5 members. The constellation of your group of six will remain unchanged throughout the entire experiment.

[22] We thank Brian Cooper (MPI Bonn) for the translation.

Also at the beginning of the experiment, the computer will name the participants of each group of six, assigning to each a randomly drawn letter (a, b, c, d, e, or f). Each participant in the group is equally likely to receive a particular letter (a, b, c, d, e, f). Each letter is distributed once in each group of six.

In each period, you will interact with exactly one of the other participants from your group. The computer will randomly determine at the beginning of each period who that other player is. The other 4 participants in your group of six will each also be randomly matched with another participant from the group. In total, there will hence be three parallel encounters in your group in each period.

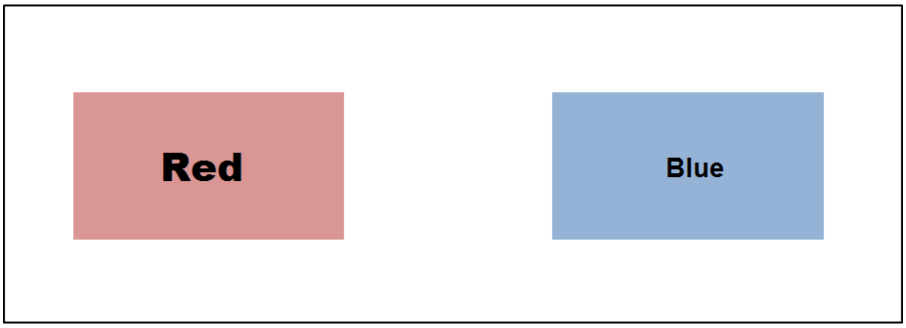
In each period, your task is to choose one of two decision fields:



How many points you earn in a period depends on your decision as well as on the decision of the participant with whom you are interacting.

- If you choose "Red" and the other participant chooses "Blue", you will earn 10 points, and the other participant will earn 1 point.
- If you choose "Blue" and the other participant chooses "Red", you will earn 1 point, and the other participant will earn 10 points.
- If both participants choose "Red", you will both earn 0 points.
- If both participants choose "Blue", you will both earn 0 points.

In each period, one of these fields will be in bold:



Whenever you see the field “Red” in bold, the other participant sees the field “Blue” in bold, and vice versa. You are free to decide whether you wish to follow the marking or not.

The computer decides on the basis of your letter which field is in bold. Whichever participant’s letter comes first in the alphabet sees the field “Red” in bold. If, for example, the computer assigned you the letter c at the beginning of the experiment, and you interact with participant d, e, or f, you will see “Red” in bold. If you interact with participant a or b, however, you will see “Blue” in bold.

For example, if you were assigned the letter a at the start, “Red” will be in bold in all periods. If you are participant f, “Blue” will always be in bold, etc.

Only once both participants have made their decisions will you find out what the other participant has chosen.

At the end of each period, your computer screen will give you an overview of:

- which field you have opted for;
- which field the other participant has chosen;
- the income you and the other participant have each earned in this period;
- how the participants of the other encounters have chosen.

Further, you have the chance to transfer to the other participant any part of your income from the current period. To do this, enter on your screen the number of points you wish to transfer to the other participant, and confirm your entry by clicking “Continue”. You are free to decide whether or not you wish to transfer points and, if you do, how many points you wish to donate.

In addition, your computer screen will give you an overview of:

- the number of points transferred by you to the other participant;

- the number of points the other participant has transferred to you;
- the income you have earned in this period, after transfers;
- the total income you have made up to now;
- the total income each of the other participants in your group of six has made so far.

At the end of each period, the letters that were randomly allotted at the beginning of the experiment (a, b, c, d, e, f) can be allotted anew by the computer. As in the beginning, every participant has the same probability of receiving a specific letter (a, b, c, d, e, f). This probability is independent from the letter that he had so far.

This reallotment takes place, if in all three matches within your group of six both participants choose the same color. For instance, if in all three matches both participants choose "Red". Or, for instance, if in two matches both choose "Red" and in the other match both choose "Blue".

In the next period, you will once again be randomly matched with one other participant from your group of six, with whom you will then interact.

Do you have any questions? If yes, please raise your hand. We will come to you.

C Post-Experimental Tests

Other-regarding preferences. Other-regarding preferences were measured using the social value orientation (SVO) ring measure (Liebrand and McClintock 1988). The ring measure consists of 15 modified dictator games in which players allocate money between themselves and another player. Players are characterized on a continuous type space ranging from competitiveness and individualism to prosociality and altruism. To make the *Social Value Orientation (SVO)* comparable to the other scales, we divide the ring degree by 45, such that 0 describes a perfectly selfish (0 degree) individual, and 1 a perfectly pro-social (45 degree) individual.

Risk. The risk elicitation consisted of 1 general risk question and 6 domain-specific questions. The general question read: "Are you, generally speaking, a person willing to take risks or do you rather try to avoid risks?" The domain-specific questions read: "How would you rate your willingness to take risks in the following domains...(i) when driving a car, (ii) in financial investments, (iii) in leisure and sports, (iv) in your career, (v) concerning your health, (vi) when trusting unfamiliar people?". There are 10 answer options ranging from *not at all willing to take risks* to *very willing to take risks*. We use the arithmetic mean over the 7 questions as our measure of an individual's risk attitude.

Trust. The trust questions read: "Please rate the following three statements: (i) Generally, people can be trusted. (ii) Nowadays you cannot trust anybody. (iii) When dealing with strangers it's better to be careful before trusting them." The answer options are: *fully agree*, *rather agree*, *rather disagree*, *fully disagree*. The composite trust measure is the arithmetic mean of the three question whereby the first is coded negatively and the other two positively.