Power, Competition and Voting –
an Experimental Enquiry

Arne Robert Weiß

Dissertation

zur Erlangung des Grades eines Doktors der Wirtschaftswissenschaften

(Dr. rer. pol) der

Universität Erfurt

Staatswissenschaftliche Fakultät

2009

Gutachter: Prof. Dr. Bettina Rockenbach

Prof. Dr. Manfred Königstein

Datum der Disputation: 17.12.2009

Dekan: Prof. Dr. Manfred Königstein

urn:nbn:de:gbv:547-201000028

Zusammenfassung

Macht spielt eine zentrale Rolle in Interaktionen mit Interessenkonflikten. Diese Arbeit studiert experimentell die distributiven und allokativen Konsequenzen verschiedener Machtbasen in wiederholten Spielen und untersucht darauffolgend die Bedeutung von Prozeduren, durch die Macht zugeteilt wird. Im Besonderen werden die Verhaltenseffekte von Wahlprozeduren und ihre Interaktion mit Wettbewerb in Situationen von konfligierenden Interessen analysiert. Diejenigen, die durch die Machtverteilung bevorteilt sind, entweder durch kompetitive, strategische oder allokative Macht, verdienen im Durchschnitt mehr als ihre Gegenparte. Die experimentellen Ergebnisse zeigen zudem, dass die Machtvergabeprozedur eine Rolle spielt. Die bloße Präsenz einer Wahlprozedur kann für eine weniger selbstorientierte Machtausübung sorgen. Dieser Befund ist jedoch abhängig vom Kontext, in den die Wahl eingebettet ist. Die experimentellen Ergebnisse deutet hierbei darauf hin, dass Versprechen, als Basis für Wahlentscheidungen, eine entscheidende Rolle dafür spielen, ob eine Wahl zu weniger oder zu möglicherweise sogar mehr selbstorientierter Machtausübung führt.

Macht

Wahlen

Wettbewerb

Experiment

Abstract

Power plays a central role in interactions with conflicting interests. This thesis provides a behavioural study of the distributive and allocative consequences of different power bases in a repeated play framework and subsequently investigates the role of procedures through which power is granted; in particular, we analyse the behavioural effects of voting and its interaction with competition in situations of conflicting interests. Those who are favoured by the power distribution, either with competitive and strategic power or with allocative power, earn on average more than their counterparts. The experimental results also provide evidence that the procedure matters through which power is granted. The mere procedure of voting can lead to a less self-oriented exercise of power. However, this finding is dependent on the context in which the election is embedded. In particular, the experimental results suggest that promises, as a basis for voting decisions, play a crucial role in determining whether voting leads to a less or, possibly, even to a more self-oriented exercise of power.

Power

Voting

Competition

Experiment

Acknowledgements

First of all, I am greatly indebted to my advisor, Bettina Rockenbach, for her continuous support and for the provision of an inspiring working atmosphere. I would like also thank my co-authors Geng Hong and Gari Walkowitz for their support and for many productive discussions. I also thank Stephan Tontrup for his support in the initial phase of designing the experiments in chapter 3 and 4 and Mu Tao for programming the experiments in chapters 2 and 3. I thank Geng Hong for also running the experiments in chapters 2 and 3 and translating the respective instructions; Fu Hao and Cheng Peng for running the treatments VOT-P and RAN-P in chapter 4 and Mareike Hoffmann for running the German sessions in chapter 4. I would also like to thank Urs Fischbacher for providing us with the z-tree program of Brown, Falk, Fehr (2004), which we adjusted for the experiment in chapter 1.

Furthermore, I would like to express my gratitude to Irenaeus Wolff for his continuous support, to Dennis Dittrich, Fabian Kleine and Manfred Königstein for their advice on econometric modeling and to Özgür Gürerk for always being willing to help and give advice. I also thank my colleagues at the University of Erfurt, the University of Bonn and beyond for their valuable comments and help: Gerhard Wegner, Reinhard Selten, Heike Hennig-Schmidt, Frank Renkewitz, Ringa Raudla, Christiane Pilz, Michel André Marechál, Moritz Mannschreck, Thomas Lauer, Johannes Kaiser, Mareike Hoffmann, Lars Herrmann, Stefan Große, Sebastian Georg, Simon Gächter, Kai Ebenrett and Matthias Altmann as well to CEREB and its members for providing a stimulating, interdisciplinary working environment.

I also thank participants at respective sessions of the ESA-conference 2008 in Lyon, the GfeW-meeting 2008 in Mannheim as well as workshops and colloquiums in Bonn, Cologne, Zurich and, of course, Erfurt.

I am also indebted, not literally though, to the financial support by the *Stiftung der Deutschen Wirtschaft* and by the *Deutsche Forschungsgemeinschaft* (research projects RO 3071/1-2 and HE 2790/2) as well as to funding by the Bonn Laboratory for Experimental Economics. Financial support is also gratefully acknowledged by the Southwest-Jiaotong University in Chengdu, China.

At last, I am very grateful to Severine Bebek as well as to Hans-Jürgen Weiß, Elisabeth Weiß and my extended family for their long-lasting support.

Content

Overview

Power as the "probability that one actor in a social relationship will be in a position to carry out his will despite resistance, regardless of the basis on which this probability rests" (Weber 1947)[1] plays a central role in interactions with conflicting interests. In distribution conflicts, power lies, by Weber's definition, with the actor who is in a better position to push through her interests. Its distributive implications make power a central concept in sociology, with Weber offering a classic and still widely used definition. Power, however, may also have allocative consequences if exerting power means that potential gains from transactions are not fully realised. To start with, power is an uncertain possibility to push through interests; whether this potential is actually sought is still a behavioural question. This thesis provides a behavioural study of the distributive and allocative consequences of different power bases in a repeated play framework and subsequently investigates the role of procedures through which power is granted; in particular, we analyse the behavioural effects of voting and its interaction with competition in situations of conflicting interests. Three bases of power are investigated in this study: a sequential structure of actions with a last-mover advantage when no binding agreements can be made, market conditions and hierarchy.

In the experiments of chapters 1 to 3, the possibility to actually exert power only arises if the transaction partners trust to at least some degree. In the absence of binding agreements, standard economic theory[2] predicts no trust to emerge since power would be fully exploited; as a consequence, the potential to push through interests turns into merely virtual power. This prediction is, however, at odds with an abundant amount of experimental data, which shows gift-exchange and other-regarding behaviour to emerge even among strangers (e.g. Fehr, Kirchsteiger and Riedl 1993; Hoffman, McCabe and Smith 1996; Bolton, Katok and Zwick 1998; Fehr, Kirchsteiger and Riedl 1998; van der Heijden et al. 2001; Charness 2004). Exercisable power may then indeed arise, yet very little is known on how the allocation of power matters for rent-sharing and efficiency. Therefore, experimental data is needed. We study the interaction of different bases of power in a systematic way in chapter 1. Very little is also known on the behavioural relevance of procedures through which power is granted.

[1] The original German quote is: "jede Chance, innerhalb einer sozialen Beziehung den eigenen Willen auch gegen Widerstreben durchzusetzen, gleichviel, worauf diese Chance beruht." (Weber 1985).
[2] Technically speaking, we refer to the subgame-perfect Nash-equilibrium based on common knowledge of rationality.

We pay due attention to this question is chapters 2 to 4. The chapters also provide a methodological ground for an experimental investigation of the research questions.

In chapter 1, we experimentally analyse competitive and strategic power in an incomplete contracts market that allows for relational contracting between buyers and sellers. Buyers pay, while sellers determine quality. Gains from trade, as a measure of efficiency, rise with quality. Competitive power stems from the market environment, while strategic power means "having the last word". As our first research goal, we systematically study the interaction of competitive and strategic power. Our second research goal takes an interesting comparison of two very similar treatments in two different subject pools as a starting point. Results by Wu and Roe (2007a)[3] show less relational contracting on the one hand and lower efficiency and higher inequality on the other hand than results by Brown, Falk and Fehr's (2004) almost identical experiment. In order to test the role of relational contracting for the interaction between competitive and strategic power, the experiment is run in two culturally distinct locations, Germany and China, which were expected to display different tendencies to form relations. This also allows us, as a third research goal, to give a first cross-cultural account on the nature of relational contracting.

We find competitive and strategic power to have a robust effect on rent-sharing across cultures and different degrees of relational contracting. Strategic power has a larger impact on rent-sharing than competitive power, but has an effect on efficiency only in Germany, where subjects rely on relations to a low degree. In this case, efficiency is higher when buyers have strategic power. By contrast, in China, where relations are employed more, strategic power does not affect efficiency. We find relations to be very similar across cultures and conditions of strategic and competitive power. The level of relational contracting is found to depend on its strategic value from the perspective of buyers, to which Chinese buyers seem to react stronger than German buyers. The results show relationships to enhance efficiency and reduce inequality when power is asymmetrically distributed. Yet, they are no universal solution to limit the influence of power and improve cooperation if relations are formed endogenously. The Chinese results also question the general character of earlier results (e.g. by Wu and Roe 2007a and by our German

[3] We refer to the IZA discussion paper of Wu and Roe. They also published a second paper on the same set of experiment (Wu and Roe 2007b). The IZA discussion paper, however, fits more closely to our research question.

treatments) that shows efficiency to improve when those who pay have “the last word”. It rather seems that a high degree of relational contracting and strategic power lying with those who pay are alternative ways to lead to a high, albeit still not optimal, level of efficiency.

In chapter 2, we experimentally investigate the effects of political competition between incumbent power holders and candidates on confiscatory behaviour and investments. Political competition is induced when investors are also voters that elect the holder of power. In the spirit of Grossman and Noh (1994) and Bardhan and Yang (2004), political competition may link the incumbent holder of power´s behaviour more closely to the preferences of voters if incumbents are sufficiently motivated by re-election concerns. If, however, the threat of dismissal by the electorate is as large as to render re-election unrealistic, incumbents may take a short-term view and expropriate all they can while still in office. We embed this experiment in the debate on political regimes and economic performance and argue in favour of controlled experiments as a complementary empirical tool in order to investigate underlying causal relationships. Based on our experimental results, political competition limits confiscatory behaviour as incumbent power-holders are motivated to be re-elected based on a good track record. Investment levels and consequently efficiency as well as investor profits are higher in the presence of political competition. However, rents of holders of power are only somewhat lower with political competition as in the absence of political competition some rulers are surprisingly stuck on the wrong end of the Laffer-curve. Contrary to theoretic predictions, the effects of political competition do not depend on the power holder’s time-horizon. Astonishingly, the effects of political competition even carry over into the last legislative period, i.e. when strategic incentives based on foresight are identical with and without political competition.

Chapter 3 analyses the long and short-run effects of the mere procedure of voting on exertion of power and on trust put into holders of power. Whereas chapter 2 analyses the joint role of political competition and the voting procedure through which political competition is induced, chapter 3 focuses on the procedure of voting itself. Employing the same basic design as chapter 2 but ruling out any reliable individual reputation building, we find elected power holders to use their power in a less self-oriented way. This long-run effect is robust to controlling for player-types, promised back-transfers and the level of received

investments. While we do find some evidence of promise-keeping in both treatments, elected power holders do not keep their word better than randomly drawn power holders. Although the design rules out reliable individual reputation building, differences in dynamic considerations due to instances of subtle identification possibilities may contribute to the effects of voting on the exertion of power over the first four units of the experiment. The small role of identification for electoral prospects as well as evidence from the last unit nevertheless suggest that the procedure of voting itself moderates power; yet, the extent of any possible influence of dynamic incentives remains an open question. We find investors to display higher trust if they can elect the holder of power, independent from any payoff-relevant decisions by the power holders. This effect cannot be explained by treatment-differences in the level or in the role of promises. Investments stay higher in the voting treatment for the first 10 rounds. This effect is driven by differences in the profitability of investing. Hence, an effect of the procedure of voting itself that is independent of the behaviour of power holders can only be observed in the short-run.

Chapter 4 provides a stress-test for the effects found in chapter 3 and a similar voting study by Corazzini, Kube and Maréchal (2007). It introduces the strongest form of power by implementing a clear hierarchy. In the previous chapters the degree of exercisable power is endogenous to decisions by the experimental subjects; by contrast, power in chapter 4 is no longer merely a probability but a certainty as there is no opting out by the subordinates. Furthermore and in contrast to the first three chapters, power comes with no promise by the holder of power to act in the interest of those at the receiving end of power. Instead, voting takes place based on personal descriptions. Furthermore, in light of the unclear role of dynamic considerations for the power moderating effect of voting in chapter 3, the design rules these out entirely: the game is only played once. The experiment was first conducted in Chengdu, China. The results show that voting has no effect on power holders' transfer decisions. In order to test whether the results are different if subjects are more used to formal voting as a mechanism to determine hierarchy, we re-ran the experiment in Erfurt, Germany. The German results show that voting does not have an effect on the power holders' choices either. The results of the stress-testing experiment therefore send a cautionary note as they imply that the power of voting to limit the self-oriented exertion of power shown in previous experiments may be context-specific. In order to test the role of

promises, we ran a new experiment in Chengdu, China, with numerical promises as statements prior to the selection of the power holders. In this case, voting indeed matters: elected holders of power transfer more to their recipients than power holders who are drawn randomly.

The thesis shows that the allocation of power and the procedure through which power is granted matter. Those who are favoured by the power distribution, either with competitive and strategic power (in the first chapter) or with allocative power (in chapters 2 – 4), earn on average more than those who must or may put themselves at the discretion of holders of power. Yet, power matters differently than what standard economic theory predicts. Only in the setting of unconditional power (in chapter 4), standard economic theory predicts the *probability* to carry out one's will to be more than a mere virtual possibility. Yet, power matters in all situations studied in this thesis. The paradoxical reason for this is that power is rarely exploited entirely; thereby, trust may emerge, which is a precondition for power to emerge that may be exerted in the power-holders favour. Economic theories based on heterogeneous agents (e.g. Fehr and Schmidt 1999; Bolton and Ockenfels 2000; Falk and Fischbacher 2006) are emerging to explain cooperation in such situations. They are, however, not generic enough to account for the complex interaction of different bases of power in a repeated game framework. The experimental evidence of this thesis can be fed into their further development.

The experimental results also provide evidence that the procedure matters through which power is granted. The mere procedure of voting can lead to a less self-oriented exercise of power (chapter 3). Voting can therefore be considered more than a means to select among candidates and to sanction bad-performing incumbents. No current theory can provide a comprehensive explanation for this finding. A yet stronger effect on the exertion of power is found when voting induces political competition (chapter 2). The results show that re-election concerns weigh more than the possible shortening of power-holders' horizons. In fact, surprisingly and contrary to theoretic predictions based on heterogeneous agents, the length of incumbent power-holders horizon does not matter at all. The effects of political competition even carry over into the last legislative period. Up to this point, the results seem to paint an optimistic picture of behavioural consequences of voting. However, the stress-testing experiment (chapter 4) sends a cautionary note by emphasising the context-

dependency of the power of voting to limit power. In particular, the results suggest that promises, as a basis for voting decisions, play a crucial role in determining whether voting leads to a less or, possibly, even to a more self-oriented exercise of power.

The thesis provides ample room for future research on power and voting; an outlook is provided in the chapters' concluding sections.

1 When power meets relations – competitive and strategic power in incomplete contracts markets[1]

1.1 Introduction

Standard microeconomic theory and experiments (for an overview see Holt 1995) are surprisingly well-aligned in predicting behaviour in complete contract environments. Little is known, by contrast, about the incentives to cooperate in contracts that are incomplete in the sense that an obligation of at least one party cannot be enforced.[2] Once we depart from the idealised world of zero transaction cost, real-world contracts have to be acknowledged as being necessarily incomplete to at least some degree. Incentives to either defect or to cooperate may be differently strong for the contracting parties. They may either stem from aspects within or outside the contract relationship. In this paper we will study on both internal and external sources of these incentives: *competitive power* and *strategic power*.

Competitive power results from the competitive conditions on the market in which the contract is embedded. The degree of competition may vary if the market is imperfect. The standard case of market imperfections is short-run capacity constraints that make it unprofitable or even technically impossible to quickly build or reduce capacity. A contracting partner has more *competitive power* if she faces less competition than the potential transaction partner. The threat to terminate a relation is, therefore, more credible for the side with competitive power. In case contracts are only partially incomplete, competitive power may already influence the contracting terms to the disadvantage of the side facing more competition.

In a world of incomplete contracts, the contracting parties need not only agree on the content of actions but also on their sequence.[3] Strategic power stems from the resulting

[1] Based on: "When Power Meets Relations – Competitive and Strategic Power in Incomplete Contracts Markets" by Hong Geng, Bettina Rockenbach and Arne Robert Weiß (2009), Working Paper, University of Erfurt. All authors contributed equally.

[2] The contract may not specify an obligation at all, as contract-writing costs increase with the degree of completeness, or it may not be enforceable by a third party due to the absence of a third party with enforcement power or because compliance with the contract is unobservable for a third party.

[3] If a powerful third party is non-existing true simultaneity of actions seizes to be possible. Think of two rivalling gangsters pointing guns at each other who in order to avoid bloodshed agree on dropping their weapons simultaneously. As both know, however, one has to do the decisive step: trust in the other party to follow suit. As we know from many movies, this trust is not always warranted, and one has better kept a gun hidden in the socks. Strategic power not only matters in organised crime but in virtually all commercial transactions. As Arrow's (1972) famous quote says, they always entail an

internal structure of the contract. In cooperation situations with a moral hazard risk the strategic position to still have a powerful option when the other party already made its move can be very valuable. The last mover can then avoid being the sucker or even exploit the strategic position in order to renege in his own favour. We coin this advantage *strategic power*. To put it more general, strategic power lies with the side that still has discretion over whether to abide by its promises once the other side is no longer endowed with a reneging option itself, because the latter either already made its move or because it is forced, by a third party, to comply.

Economic theory gives us little guidance to explore strategic and competitive power if contracts are incomplete and contract compliance is costly. Standard game-theory based on the assumption of common knowledge of rationality only allows for cooperation in equilibrium if the parties have an infinite horizon.[4] Unfortunately, as stipulated by the folk theorem for repeated games, there is typically a multitude of equilibria of which no cooperation is always one. Clear predictions can, therefore, not be derived. If there is a known end to any interaction between the parties, standard game theory fails to predict any cooperation above enforceable levels and hence leaves very little room for effects of competitive and strategic power. Voluntary cooperation, however, is a robust phenomenon in sequential cooperation games in finite play (see Healy 2007). Furthermore, viewing existing experimental evidence in the light of our notions of competitive power and strategic power shows both types of power to matter.[5] Recent economic theories of other-regarding preferences (e.g. Fehr and Schmidt 1999; Bolton and Ockenfels 2000; Falk and Fischbacher 2006) built on experimental evidence do predict cooperation even in finite play and go already some way in explaining consequences of strategic power in static play (see Fehr, Klein and Schmidt 2007). However, they are so far not generic enough to fully understand the complex interaction between strategic and competitive power in repeated play. Hence, empirical data is needed.

element of trust: by the party that moves first. The other side, by contrast, can sit back and observe the transaction partner´s actions.

[4] Technically, cooperative equilibria are also possible in finite games with unknown end in which the players attach a positive probability to the continuation of the finite game in every round.

[5] See the comparison of Brown, Falk and Fehr 2004 and 2008 for an analysis of competitive power and Wu and Roe, henceforth WR, 2007a for an analysis of strategic power.

An ideal data set for this research goal would entail exogenous changes in competitive and strategic power keeping everything else constant. This data set is almost impossible to get in the field as collected data would suffer from endogeneity problems. The allocation of strategic power is likely to also depend on the relative competitive power of the contracting parties. The side with competitive power will try to design the contract in a way that gives itself strategic power as well. Competitive power itself may depend on the institutional determinants of strategic power. A market that suffers for example from enforcement problems for quality is less attractive for buyers to enter than one in which well-functioning enforcement institutions are in place. We circumvent the endogeneity problem by exogenously shifting competitive and strategic power between buyers and sellers using controlled experiments. Previous experiments, albeit not specifically designed to do so, already provide some insight on either competitive or strategic power.

Brown, Falk and Fehr, henceforth BFF, (2008) experimentally analyse the role of unemployment for the emergence of relational contracting and efficiency. They find that wages and, in reciprocal manner, rents react – in our terminology – to changes in competitive power while efficiency remains unaffected. They also show that long-term relations emerge to a larger extent if competitive power sides with the principals. In all their treatments, agents have strategic power; in our terminology principals and agents are analogous to buyers and sellers respectively.

According to Fehr, Klein and Schmidt (2007), efficiency in a one-shot principal-agent game is higher if the side with a lower possible loss from trusting has to move first. Efficiency is higher in a bonus contract which – in our terminology – gives strategic power to the buyers. Buyers also achieved higher profits if they had strategic power. Wu and Roe, henceforth WR, (2007a) support this result in a repeated play framework. They find efficiency to be higher and relational contracting to be reduced if buyers had strategic power. They also find the side with strategic power to earn a larger share of rents. In their experiment, competitive power always stays with the buyers.

To the best of our knowledge, this is the first study which systematically investigates the interaction of competitive and strategic power. This is our first research aim. Our second research goal takes an interesting comparison of two very similar treatments in two different

subject pools as a starting point. WR´s (2007a) results show less relational contracting on the one hand and lower efficiency and higher inequality on the other hand than BFF´s (2004) almost identical experiment (compare section 1.6.2). We want to investigate the interaction of relational contracting with strategic and competitive power. In order to achieve this, we run the experiment with two different subject pools, in Erfurt (Germany) and in Chengdu (China), which we expect to have different tendencies to engage in relational contracting. This also allows us, as a third research goal, to give a first cross-cultural account on the nature of relational contracting.

We find competitive and strategic power both matter. Competitive power influences rent-sharing even if contract enforcement is entirely absent. Strategic power has a larger impact on rent-sharing than competitive power. Competitive power does not affect efficiency. Strategic power only influences efficiency in case buyers react weakly to the strategic value of relational contracting, as they do in Erfurt, Germany. In this case efficiency is raised when buyers have strategic power. If, however, subjects employ relational contracting more as a contract enforcement device, as they do in our sessions in Chengdu, China, neither competitive nor strategic power affects efficiency. Relational contracting is used more when buyers, who initiate offers, do not have strategic power. Despite different levels of relational contracting, the characteristics of relationships are astonishingly similar across cultures as well as market conditions and contract structures. Relationships are always based on a high degree of gift-exchange. We conclude that cross-cultural differences in the behavioural response to the strategic value of relational contracting drive the disparities in the effects of strategic power between our sessions in Germany and China.

The rest of our paper is organised as follows: we introduce the experimental design and derive the game-theoretic predictions in the following section. Behavioural predictions in the absence of common knowledge of money-maximising rationality are discussed in section 1.3.2. Section 1.4 introduces the experimental implementation. In section 1.5, we present the results on competitive and strategic power from our German sessions. We highlight the role of relational contracting and lead over to our second research question, on the robustness of our effects, in section 1.6. In section 1.7, cross-cultural hypotheses are discussed; comparative results of our Chinese sessions are presented in section 1.8. Section 1.9 provides a first account of the cross-cultural nature of relational contracting. The

contributions of this paper are summarised in section 1.10, while the final section 1.11 provides a discussion and an outlook for future research.

1.2 Experimental design

1.2.1 Treatments

We study a repeated market with buyers and sellers which allows for relational contracting. Each buyer may buy one good while each seller may sell one good. Buyers move first by simultaneously offering a price and a desired quality. Each seller may accept exactly one buyer's offer and deliver quality (q) to that buyer. Offered price and desired quality of an accepted offer are called contracted price (p') and contracted quality (q') respectively. The quality (q) delivered by the buyer may differ from the contracted quality (q'). The treatments implement competitive power and strategic power. In the treatments in which buyers have strategic power (spB-teatments), the buyer has an additional move in which she can freely choose the price (p), paid to the seller. The actual price (p) may differ from the contracted price (p'). If the buyer does not have the additional last move, she has to pay the offered price (p'), and the seller is said to have strategic power (spS-treatments).

The allocation of competitive power depends on the relative number of sellers and buyers. If there are an excess number of buyers, sellers are said to have competitive power (cpS-treatments); in case there are an excess number of sellers, buyers are said to have competitive power (cpB-treatments).

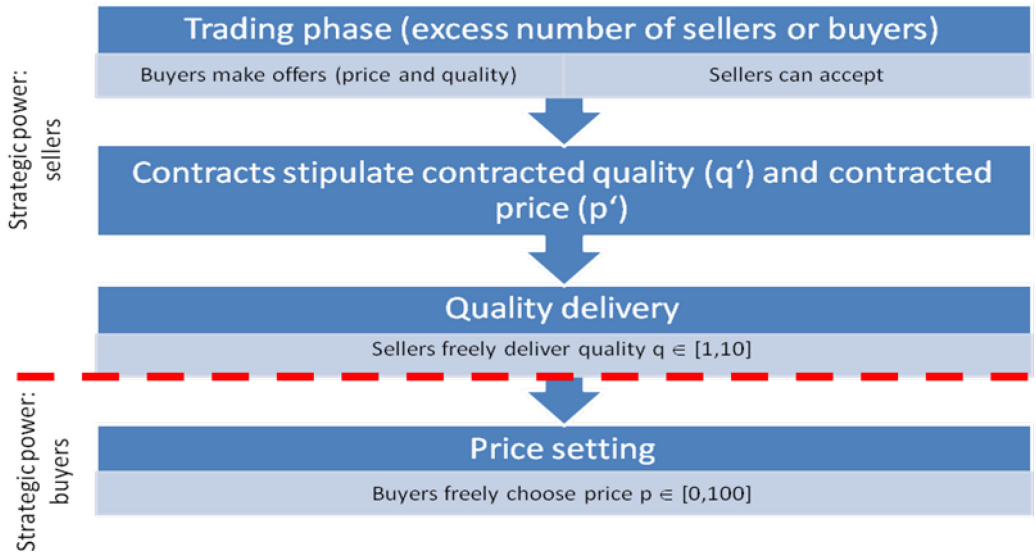


Figure 1-1: Sequence of players´ moves

Figure 1-1 illustrates the sequence of players' moves in our experiment. The combination of the presence and absence of competitive and strategic power constitutes our 2x2 design, as displayed in Table 1-1.

Table 1-1: Experimental treatments

		competitive **p**ower	
		Sellers	**B**uyers
strategic **p**ower	**S**ellers	spS-cpS	spS-cpB
	Buyers	spB-cpS	spB-cpB

We understand the treatments spS-cpB and spB-cpS as mixed-power treatments, as competitive and strategic power are allocated to different sides, whereas we think of spS-cpS and spB-cpB as concentrated-power treatments, as one side has both types of power.

Apart from slight modifications, the treatment spS-cpB corresponds to the treatments ICF by BFF (2004) and IC1 by WR (2007a), while the treatment spS-cpS is equivalent to the IC-treatment in BFF (2008). In a slightly different form spB-cpB was also employed by WR (2007a) as their IC2-treatment. The novel treatment spB-cpS is central to the research question as it allows us to study the relative effects of competitive and strategic power.

1.2.2 Trading environment

Buyers can either make private or public offers. A private offer is only directed to a specific seller and can only be seen and accepted by this seller. A public offer can be seen and accepted by all sellers in the market. In order to make a private offer to a specific seller, a buyer has to type in the identification number of the respective seller and choose "private" as the type of offer. Buyers can make as many offers as they wish until either one of their offers has been accepted or the trading time (3 minutes) has elapsed. Once an offer has been accepted, all outstanding offers of the respective buyer are deleted.[6]

The design allows for relational contracting, i.e. for repeated trading between the same seller and buyer. Technically, repeated trading is possible both by public and private trades. However, the only reliable way for a buyer to form a relation is to make a private offer to the seller of the previous round. By making public offers, buyers are effectively entering a spot-market as they have no control over who accept their offers.

1.2.3 Payoffs

Payoff-functions for buyers (π_B) and sellers (π_S) are given by

$$\pi_B = 10 * q - p$$

$$\pi_S = p - c(q).$$

The cost of providing quality, $c(q)$, rises in quality q according to the following convex schedule

q	1	2	3	4	5	6	7	8	9	10
$c(q)$	0	1	2	4	6	8	10	12	15	18

The outside option that both sellers and buyers earn if they do not conclude a trade is 4. Subtracting the outside option from the payoffs yields rents for buyers and sellers:

[6] Hence, in case more than one seller wants to accept a buyer´s offer, the speed of accepting decides.

$$r_B = 10 * q - p - 4$$

$$r_S = p - c(q) - 4.$$

Buyer-rents are an increasing function of q with a constant marginal rent of 10:

$$\frac{\partial r_B}{\partial q} = 10,$$

while seller-rents are a decreasing function of q with a maximum marginal cost of 3:

$$\frac{\partial r_S}{\partial q} \geq -3.$$

Since the marginal rent for buyers is always larger than the marginal cost for sellers, joint rents increase in quality. The social optimum is, therefore, reached if sellers provide the maximum quality of 10. The maximum achievable joint rent per trade is

$$r_{max} = r_B\,(q = 10) + r_S(q = 10) = 100 - 18 - 8 = 74$$

In this case, trading efficiency is 100%; it is defined as achieved gains from trade relative to maximum possible gains from trade:

$$TE = \frac{r_B(q) + r_S(q)}{r_B(q = 10) + r_S(q = 10)}$$

1.3 Hypotheses

1.3.1 Game-theoretic solution based on common knowledge of rationality

In order to derive unique subgame-perfect Nash-equilibria for the four treatments based on common knowledge of money-maximising rationality, we first assume both transaction partners want to achieve a marginally higher pay-off by concluding a trade (for example due to minimal but positive transaction costs) than by taking their outside option. We secondly abstract from the trading phase and analyse the gift-exchange game that follows for a given pair of seller and buyer. Trying to solve the trading phase game-theoretically would not yield

further insight and is not possible without an extended set of assumptions on trading behaviour as the trading phase cannot be represented in extensive form.[7]

With common knowledge of money-maximising rationality no cooperation above minimum levels of price and quality will take place in any treatment by the logic of backward induction. The last-moving player will not cooperate in the last round of the experiment. The second-to-last-moving player anticipates this and will, by backward induction, not cooperate in any round. In case sellers have strategic power (spS-treatments), sellers will only accept offers at a price of at least 5 and will set minimum quality of 1 for any possible price. Buyers, anticipating this, will offer a price of 5 in order to be marginally better off than their outside option of 4. Since both parties are better off by concluding a trade than by taking their outside option, all trades are predicted to take place and market efficiency (defined as realised number of trades over maximum number of trades) therefore to be 100%.[8]

In case buyers have strategic power (spB-treatments), money-maximising buyers will only pay the minimum price of zero as the contracted price is non-binding. Sellers therefore anticipate being strictly worse off by concluding a trade than by gaining their outside option of 4. Sellers will consequently not accept any offered price. Market efficiency is therefore predicted to be zero.

1.3.1.1 Money-maximising predictions for strategic power

Shifting strategic power to the buyers is predicted to decrease efficiency but not to affect relative rent-sharing. At minimum level of quality trading entails a social gain per trade of 2, which is lost in the spB-treatments. Regardless of strategic power, both parties are therefore predicted to earn the same share of rents.

1.3.1.2 Money-maximising predictions for competitive power

Based on the assumptions made, a price of 5 and a quality of 1 is predicted for both cases of competitive power when sellers have strategic power, while no trades are predicted to take

[7] In the trading phase speed of making offers and accepting offers may play a role. Since speed as a personal characteristic is outside any game-theoretic model, we cannot solve the entire game. We therefore abstract from giving an explicit account of the trading phase and treat the game as if speed played no role. We are then not able to predict which seller or buyer will conclude a trade. Nevertheless, we are able to describe the trades that will occur in equilibrium if players are commonly known to be rational and money-maximising.

[8] We distinguish market efficiency and trading efficiency. Market efficiency refers to the number of undertaken trades as a percentage of all possible trades. Trading efficiency refers to the percentage of achieved gains from trade for the trades undertaken.

place for both cases of competitive power when buyers have strategic power. Competitive power is therefore not predicted to have any effect on the behaviour of the players. Consequently, also efficiency and rent-sharing is predicted not to be affected by competitive power.

Table 1-2 summarises the predictions based on common knowledge of money-maximising rationality:

Table 1-2: Summary of game-theoretic predictions

	spS-cpB	spS-cpS	spB-cpB	spB-cpS
Market efficiency	100%	100%	0%	0%
Price	5	5	-	-
Quality	1	1	-	-
Seller rent	1	1	0	0
Buyer rent	1	1	0	0
Share of buyer profits	50%	50%	50%	50%

1.3.2 Behavioural hypotheses

The hypotheses change substantially if we relax the assumption of common knowledge of money-maximising rationality. As long as subjects expect to earn rents from cooperation at the end of the experiment, due to imperfect knowledge of rationality (for a theoretical analysis see Kreps et al. 1982; for experimental results see Andreoni and Miller 1993), or do not induct the entire game backwards (Rapoport 1997; Weber, Camerer and Knez 2004; Charness and Levin 2005), there is scope for cooperation throughout the experiment. Rents from cooperation even at the end of the game may arise out of other-regarding preferences (Fehr and Schmidt 1999; Bolton and Ockenfels 2000). In order to derive hypotheses in the absence of common knowledge of rationality, we assume the co-existence of two types of players that are hidden information: reciprocal players and money-maximising players. Reciprocal players always reciprocate gifts for intrinsic reasons, i.e. without any expected material gain. We understand gifts to be actions that may lead to positive rents for the last-moving transaction partner. Reciprocal players are assumed to reciprocate in a way that gives the other party a positive and constant share of rents; that is to say, the share of rents is assumed to be independent of the sum of both parties´ rents.[9] Reciprocity then implies a

[9] To be precise, we add a further assumption that follows the models of for instance Fehr and Schmidt (1999) and Bolton and Ockenfels (2000): the reference point of a reciprocal player is solely based on the payoff-comparisons within their own

rising relation between quality and price (when sellers have strategic power) or price and quality (when buyers have strategic power). Money-maximising players, by contrast, always choose the action that maximises their own payoff; money-maximising players consequently only reciprocate if they expect to gain from it. We assume money-maximising players to reciprocate for strategic reasons, but never to a stronger degree than reciprocal players.[10]

In the presence of reciprocal players considerable cooperative gains are available to be distributed between the transaction partners, and incentives to conclude trades are strong. We consequently expect market efficiency not to be significantly lower than 100% in all treatments.

H1: Market efficiency is 100% in all treatments

In order to think about the effects from strategic and competitive power in the presence of both reciprocal and money-maximising players, it is helpful to separate the static and dynamic incentives. Reciprocal players reciprocate gifts in both static and dynamic games. In a one-shot game or at the end of a repeated game, the money-maximising action is unambiguously not to reciprocate. In a repeated game, also money-maximising players may reciprocate if, by reciprocating in one round, they can sufficiently increase the chance to gain rents in the rounds that follow. A player with strategic power may earn rents in future rounds by concluding repeated trades in a high-trust relation. Within relations, we therefore expect money-maximising players to behave as if they were reciprocal players. Outside relations, dynamic incentives are weak so that money-maximising players will reciprocate significantly less than reciprocal players. Building on the findings from BFF (2004, 2008) and WR (2007a) we expect trades to occur both within and outside relations.

To conclude, we expect only part of the trades to be characterised by reciprocal behaviour by the side with strategic power. The side with strategic power may thereby be described to play, as a group, a mixed-strategy of reciprocating and of abrogating on their contractual obligations, each with a positive probability; therefore, the average degree of reciprocity by

transaction with their current partner. A reciprocal second mover is therefore not willing to pay a premium for the risk of the first mover being the sucker.

[10] Money-maximising players may try to mimic the behaviour of reciprocal players in order not to reveal their type (for further discussion see BFF, 2004). However, it never pays to reciprocate stronger than a reciprocal player as a money-maximising player could then increase her payoff without revealing her type.

the side with strategic power is self-serving. This simple framework allows us to think about the treatment effects.

1.3.2.1 Hypotheses for strategic power

In the spB-treatments, buyers are the last movers. For each trade, buyers can always adjust rent-sharing such that they earn at least as much as their trading partner, while sellers are dependent on reciprocating buyers to earn a rent. Given our assumptions, not all buyers will reciprocate. The corresponding analysis holds for the spS-treatments. Buyers are dependent on sellers to reciprocate in order to earn a rent, while sellers can always adjust rent-sharing to their own favour. Again, not all sellers are expected to do so. Consequently, we expect the side with strategic power to earn, ceteris paribus, a larger share of rents.

H2.1: Strategic power increases ceteris paribus the share of rents:

$$\left(\frac{r_B}{r_S+r_B}\right)_{spB-cpB} > \left(\frac{r_B}{r_S+r_B}\right)_{spS-cpB} \Leftrightarrow \left(\frac{r_S}{r_S+r_B}\right)_{spB-cpB} < \left(\frac{r_S}{r_S+r_B}\right)_{spS-cpB}$$

$$\left(\frac{r_B}{r_S+r_B}\right)_{spB-cpS} > \left(\frac{r_B}{r_S+r_B}\right)_{spS-cpS} \Leftrightarrow \left(\frac{r_S}{r_S+r_B}\right)_{spB-cpS} < \left(\frac{r_S}{r_S+r_B}\right)_{spS-cpS}$$

In order to think about the efficiency-consequences of strategic power, it is instructive to first look at the static incentives of the game, i.e. treating the game as if it was played only once, and then analyse how dynamic incentives may change the analysis. When they have strategic power, money-maximising sellers will only provide minimum-quality, leading to low trading efficiency. By contrast, when buyers have strategic power, both reciprocal and money-maximising sellers may provide above minimum quality if, on average, buyers as a group reciprocate sufficiently to render providing above-minimum quality profitable. In this case, non-reciprocal behaviour by money-maximising buyers may have no negative efficiency consequences at all. For any positive degree of rent-sharing, i.e. both sides of the transaction receiving a positive share of the generated rent, the possible loss from unreciprocated trust is higher for buyers than for sellers. Sellers only bear the cost to provide quality while buyers have to additionally pay upfront the share of rent going to the seller in order to induce a reciprocal seller to provide this quality level. We may therefore expect sellers to be more willing to provide a given quality level than buyers are willing to

pay the required price to induce on average the same quality level.[11] In a one-shot game, we therefore expect trading efficiency to be higher when buyers have strategic power. This is also the result found by in a similar one-shot experiment (Fehr, Klein and Schmidt 2007).[12]

Our experiment, of course, is not played once but repeatedly. In a repeated game, also money-maximising players may reciprocate if the dynamic incentives are sufficiently strong. In case sellers have strategic power, efficiency would then be directly raised compared to one-shot play. If money-maximising buyers reciprocate in the spB-treatments, efficiency can also be indirectly affected through the quality choices of sellers. If more buyers reciprocate in repeated compared to one-shot play, the profitability to choose above-minimum quality increases for sellers; sellers may react by increasing their quality choices.

As we argued in the introduction of section 1.3.2, strong dynamic incentives can only be expected within relationships, i.e. through relational contracting. Outside relationships, dynamic incentives are weak. Dynamic incentives are an endogenous outcome of the behaviour of players. They depend on the degree subjects make use of relational contracting as a contract enforcement mechanism, which we cannot predict with the current theories at hand. The relative strength of the incentives for sellers to provide above-minimum quality is therefore a question open to empirical investigation. We therefore stick to existing experimental evidence from WR (2007a) that show efficiency to be higher when buyers have strategic power. The results of WR (2007a) consequently imply that differences in static incentives for sellers to provide above-minimum quality drive efficiency differences. We nevertheless concede that this result may be sensitive to the degree of relational contracting.

H2.2: Quality levels and consequently trading efficiency are higher if buyers have strategic power:

$$q_{spB} > q_{spS}$$

[11] An example helps clarify this point: a buyer has to pay 59 upfront in order to induce a quality level of 10 by a reciprocal player; in this case, both players earn the same rent of 37. If, however, the buyer meets a money-maximising seller who is only willing to provide a quality of 1, the buyer makes a loss of 53. For sellers the equivalent loss is only 22.

[12] Fehr, Klein and Schmidt (2007) formally explain this result by using the Fehr and Schmidt (1999) model of inequity aversion. The main difference in their line of argumentation is that they assume reciprocal players, which they call fair players, not to be willing to display maximum trust because of the fear of suffering from disadvantageous inequality. The qualitative result, however, is robust to this assumption.

1.3.2.2 Hypotheses for competitive power

In cpS-treatments, buyers face strong competition to find a trading partner. Compared to cpB-treatments, this pressure should increase contracted prices.

H3.1: Contracted prices are higher in cpS than in cpB:

$$p'_{cpS} > p'_{cpB}$$

In case sellers have strategic power (spS-treatments), higher contracted prices directly translate into higher actual prices and hence, for a given quality level, higher seller-rents and lower buyer-rents. As the long market side in spS-cpB, sellers have a higher incentive to reciprocate in spS-cpB than in spS-cpS in order to increase their chance of receiving a renewed offer from their current buyer. We therefore expect lower quality choices in spS-cpS than in spS-cpB for any given price level. Both effects from giving competitive power to sellers, higher prices and – ceteris paribus – lower quality, let us predict a higher share of rents for sellers in spS-cpS than in spS-cpB.

In case buyers have strategic power (spB-treatments), contracted prices are non-binding. Nevertheless actual prices may ceteris paribus be higher when sellers have competitive power (in spB-cpS). Buyers have a higher incentive to be re-matched with their current seller if they do not have competitive power. Buyers may be able to increase their chance to enter a relation with their current seller by paying high prices and thereby building the reputation as a reciprocal buyer. For sellers, the incentive to enter a rent-generating relation is higher if buyers have competitive power (in spB-cpB). Sellers may attempt to enter a relation by providing high quality and thereby building a reputation as a trusting seller. To summarise, we expect two effects by shifting competitive power from sellers to buyers when buyers have strategic power: firstly, higher quality given expected prices and secondly, lower prices given quality. Both effects lead us to predict a higher share of rents for buyers in spB-cpB than in spB-cpS and vice-versa for sellers.

H3.2: Competitive power increases ceteris paribus the share of rents:

$$\left(\frac{r_B}{r_S+r_B}\right)_{spB-cpB} > \left(\frac{r_B}{r_S+r_B}\right)_{spB-cpS} \Leftrightarrow \left(\frac{r_S}{r_S+r_B}\right)_{spB-cpB} < \left(\frac{r_S}{r_S+r_B}\right)_{spB-cpS}$$

$$\left(\frac{r_B}{r_S+r_B}\right)_{spS-cpB} > \left(\frac{r_B}{r_S+r_B}\right)_{spS-cpS} \Leftrightarrow \left(\frac{r_S}{r_S+r_B}\right)_{spS-cpB} < \left(\frac{r_S}{r_S+r_B}\right)_{spS-cpS}$$

In order to think about the possible efficiency implications of competitive power, we need to look at the behaviour of sellers as the quality level determines the gains from trade, while the price paid by buyers is only redistributive. Based on the above analysis, no clear prediction on efficiency consequences of competitive power emerge as we predict two counter-running effects for both cases of strategic power: in spS-treatments, shifting competitive power to sellers should lead to higher prices (which may raise quality as a reciprocal response) but, at the same time, to lower quality given prices. In spB-treatments, competitive power lying with sellers instead of buyers should lead to higher quality for a given level of reciprocity by the buyers but at the same time to less reciprocal behaviour by the buyers.

1.3.2.3 Hypotheses for the interaction between strategic and competitive power

One novel aspect of this paper is analysing the interaction between competitive power and strategic power. The first case is comparing concentrated-power and mixed-power treatments.[13] Based on our prediction that both strategic and competitive power matter we predict buyers to earn more than sellers when they have both strategic and competitive power (in spB-cpB) while sellers are predicted to earn more than buyers when they have strategic and competitive power (in spS-cpS).

H4.1: The favoured side will earn higher rents than the unfavoured side in concentrated-power treatments:

$$r_{B\ spB-cpB} > r_{S\ spB-cpB}$$

$$r_{B\ spS-cpS} < r_{S\ spS-cpS}$$

[13] Recall that in concentrated-power treatments one side has both strategic power and competitive power (spB-cpB and spS-cpS), while in mixed-power treatments each sides has either strategic or competitive power (spB-cpS and spS-cpB).

The second case is comparing rent-sharing in the two mixed-power treatments spB-cpS and spS-cpB. Without any more elaborate theory, we cannot predict which side earns a larger share of rents. We do, however, expect rent-sharing to be more pronounced when competitive and strategic power is each allocated to different sides. In order to control for possible differences in quality levels, we predict relative inequality, i.e. equality for given gains from trade, being lower in the mixed-power treatments spS-cpB and spB-cpS than in the concentrated power treatments spS-cpS and spB-cpB.

H4.2: Relative inequality is lower in mixed-power treatments than in concentrated-power treatments:

$$\left(\frac{|r_S - r_B|}{r_S + r_B}\right)_{spB-cpS,spS-cpB} < \left(\frac{|r_S - r_B|}{r_S + r_B}\right)_{spB-cpB,spS-cpS}$$

1.3.2.4 Hypothesis for the interaction of strategic power and competitive power with relational contracting

In light of existing experimental evidence (BFF, 2004), relational contracting is based on a high degree of rent-sharing and trust. We therefore hypothesise that the effects of competitive and strategic power on rent-sharing will be lower within relations than outside relations resulting in a lower degree of relative inequality.

H5: Relative inequality is lower within relations than outside relations:

$$\left(\frac{|r_S - r_B|}{r_S + r_B}\right)_{within\ relations} < \left(\frac{|r_S - r_B|}{r_S + r_B}\right)_{outside\ relations}$$

1.4 Experimental implementation

1.4.1 Specification

The experiment was implemented on the experimental platform of BFF (2004), used also by WR (2007a) and BFF (2008), and adapted to our needs. In each group of all four treatments, 12 subjects interact for 15 rounds in a sequential buyer-seller game in a market environment. In the cpS-treatments, 5 sellers interact with 7 buyers, while in the cpB-treatments 7 sellers face 5 buyers. As each player can only make one transaction per round, at least two subjects of the side without competitive power will therefore not conclude a trade in any round. For

any type of action buyers and sellers may be identified by a unique identification number. Furthermore, subjects are asked to document the history of their trades, including the identification number and the profit of their respective trading partner, on a separate documentation sheet.

In every round, we elicit expectations: In the spB-treatments, the buyer will record a quality expectation and the seller will record a price expectation. In the spS-treatments, only the buyer is asked to provide an expectation of the quality that the seller will set as the buyer cannot deviate from the contracted price.

1.4.2 Procedure

The Erfurt-sessions of the experiment were run in April 2007 at the *Erfurter Laboratorium für experimentelle Wirtschaftsforschung* (elab). The Chengdu-sessions of the experiments were run in April 2008 at the *Herbert A. Simon & Reinhard Selten behavioral decision research lab* of the Southwest Jiaotong University in Chengdu, China. At each location, we collected six independent observations for each treatment. Therefore, in both Erfurt and Chengdu 288 subjects participated.[14] Sessions lasted between 90 and 120 minutes, and subjects earned about 12.5 Euros in Erfurt and about 61 RMB in Chengdu (earnings are equivalent to about USD 20 and USD 8.5 at the time of the experiments in Germany and China respectively). The hourly average payments were set according to local standards and paid in cash after the experiments had been finished. Test questions made sure that subjects were aware of structure of the game, including the number of players on each side, as well as how profits are calculated. The experiment only started when all subjects of a session correctly answered all the test questions. The experiment was programmed in z-tree (Fischbacher 2007).

1.5 Results of the German sessions

Already the rough sketch of the data in Table 1-3 shows that the predictions based on money-maximising players have to be rejected. There is a substantial degree of cooperation in all treatments. Both average prices and average quality choices lie significantly above minimum levels. Furthermore, the prediction that no trade would occur in the spB-treatments has to be strongly rejected. In fact, 99.56% of all possible trades are undertaken

[14] Apart from one subject in Erfurt who attended high-school all subjects were university students.

in spB-cpB, and all possible trades are undertaken in spB-cpS. Consequently, hypothesis H1 is supported by the data. There is also no statistical difference in market efficiency between the four treatments.

Table 1-3: Descriptive summary statistics of the German treatments
(money-maximising predictions in parentheses)

	spS-cpB	spS-cpS	spB-cpB	spB-cpS
Market efficiency	97.56%	99.11%	99.56%	100%
Mean price	34.94 (5)	47.29 (5)	29.47 (-)	33.79 (-)
Mean quality	5.20 (1)	5.83 (1)	6.83 (-)	6.86 (-)
Mean seller rent	23.57 (1)	34.51 (1)	15.01 (0)	19.39 (0)
Mean buyer rent	13.11 (1)	6.98 (1)	34.86 (0)	30.85 (0)
Share of buyer rents	35.72%	16.82% (50%)	69.90% (-)	61.41% (-)
Share of trades within relations	30.24%	33.24%	19.79%	8.89%

1.5.1 The role of strategic power

Recall that strategic power means having the last move in the trading process while the other party is already committed to its decision. For buyers, strategic power means that they can deviate from the contracted price after the seller is already committed to a quality level. If sellers have strategic power, they are free to choose the levels of quality when the buyers are already bound to pay the contracted price. Does strategic power pay off?

Figure 1-2 illustrates the aggregated effects of strategic-power both for rent-sharing and efficiency. The total height of the columns depicts efficiency, and the separate heights of the seller and buyer columns show the share of rents that accrues to sellers and buyers respectively.

Figure 1-2: Effects of strategic power for rent-sharing and trading efficiency

Strategic power strongly influences rent-sharing to the advantage of the strategically favoured side. Buyers gain a significantly larger share of rents in the spB-treatments (on average 65.63%) than in the spS-treatment (on average 25.71%).[15,16] Buyers also gain a significantly larger share of rents than sellers (34.36%) in the spB-treatments, while the opposite is true in the spS-treatments.[17] The impact of strategic power therefore goes beyond the ceteris paribus changes in rent-sharing that were hypothesised in H2.1. The side with strategic power always gains more than the side without strategic power. How does strategic power pay off?

There is a substantial degree of reciprocity in both types of strategic power-treatments as prices and quality-levels are strongly correlated: Spearman's rank correlation coefficients between group-average prices and group-average quality-levels reach 0.846 in spS-treatments and 0.888 in spB-treatments, which are both significant at 0.01 level. The same holds for within-group correlation between prices and quality: Spearman's rank correlation coefficients vary from 0.251 to 0.839 in spS-treatments and from 0.377 to 0.756 in spB-treatments; binomial tests reject the null hypothesis of an equal probability of positive and negative correlation coefficients in both types of treatments.[18]

While investments in above minimum prices or in above minimum quality are profitable on average, the last-moving side, nevertheless, leaves more for itself. Reciprocity is on average not payoff-equalising, but self-serving, which supports our framework on which the hypotheses in section 1.3.2 are built. In the spB-treatments, sellers must provide a higher quality in order to receive a certain price level than sellers are willing to provide in the spS-treatments given this price level. An analogous relation holds for buyers: the price buyers pay in spB-treatments for a certain quality level is lower than the price buyers need to pay in spS-treatments in order to induce sellers to provide, on average, this quality level. Hence, the strategically powerful side profits from being able to leave more for itself; at the same time, it sets, as a group, incentives for the strategically unfavoured side to trust. The result is rent-shifting in favour of the strategically strong side.

[15] All relative rents or relative inequality data in this paper are based on the ratio of group averages (instead of on the group averages of ratios for each trade). Otherwise group level ratios may be biased by individual ratios that can lie considerably below 0 or above 1 because subjects could and did make losses. Nevertheless, the results reported only change marginally if we use group averages of individual ratios.
[16] Mann-Whitney U test, two-sided: $p<0.001$
[17] Wilcoxon signed ranks test, two-sided: $p<0.001$ in both cases
[18] Binomial test, two-sided: $p<0.001$

Result SP1: Rents are higher for the side with strategic power irrespective of competitive power due to self-serving reciprocity.

As hypothesised in H2.2, strategic power influences efficiency (as market efficiency is statistically indistinguishable from 100%, trading efficiency is equivalent to total efficiency).[19] If buyers have strategic power, quality choices and therefore efficiency are significantly higher than if sellers have strategic power: quality reaches an average of 5.52 in spS-treatments but an average of 6.85 in spB-treatments;[20] thereby, nearly 70% of maximum gains from trade are realised when buyers have strategic power, while trading efficiency hovers just above 50% when sellers have strategic power.

In order to understand why efficiency is higher when buyers have strategic power, it again helps to first look at the static incentives sellers face. When buyers have strategic power, providing quality is on average highly profitable for sellers. The average degree of reciprocity turns maximum quality the payoff-maximising choice when buyers have strategic power (see Figure 1-3); therefore, for both money-maximising and reciprocal sellers it pays off to provide above-minimum quality. In the spS-treatments, by contrast, sellers have no static incentive to provide above minimum quality; therefore, only reciprocal sellers would provide above minimum quality in static play. The results seem to reflect the differences in static incentives: there are very few minimum quality choices (5.5% of all quality choices) in the spB-treatments, while the mode is the maximum quality of 10 (23.5% of all quality choices); the mode in the spS-treatments, by contrast, is the minimum quality of 1 (24% of quality choices compared to 21.1% of choices for the maximum quality of 10). Analogous to sellers in the spS-treatments (see Figure 1-12 in the appendix), also many buyers in the spB-treatments behave non-reciprocally so that sellers earn negative rents (see Figure 1-13 in the appendix).[21] In contrast to a seller's behaviour when she has strategic power, non-reciprocal behaviour by a buyer with strategic power has no direct cost for efficiency as buyers´ decisions are only redistributive.

[19] Total efficiency is maximised if all possible gains from trade are realised, i.e. the side with competitive power always finds a transaction partner (market efficiency = 100%) and quality is always set at the maximum of 10 (trading efficiency = 100%). As market efficiency is statistically indifferent from its maximum level of 100% in all treatments total efficiency is statistically the same as trading efficiency.

[20] Mann-Whitney U test, two-sided: p=0.015

[21] Surprisingly, there are even a considerable number of "super-fair" decisions that lead to lower and even negative own rents.

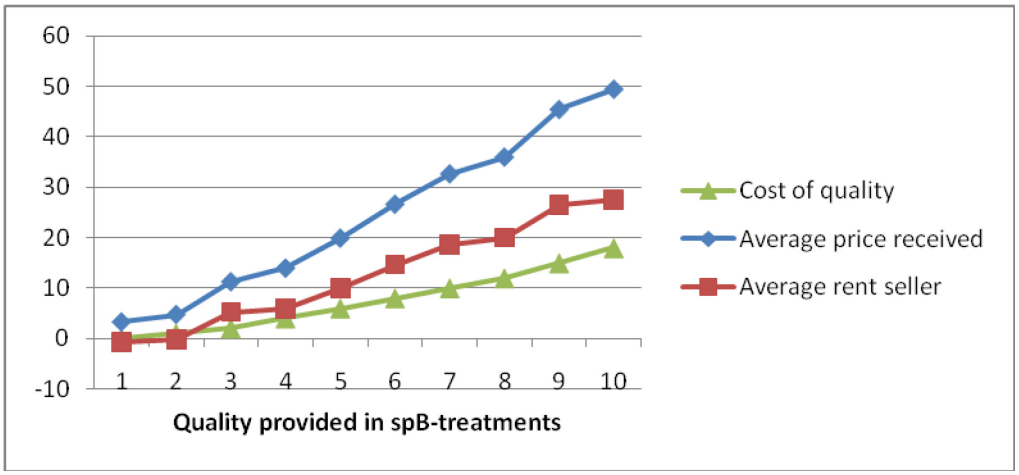


Figure 1-3: Reciprocity of buyers and profitability for sellers to provide quality in spB-treatments

For any degree of rent-sharing, costs to trust are higher for buyers than for sellers (see derivation of hypothesis H2.2). Looking at the data, we see that sellers reciprocate less than buyers in the sense that they favour themselves more than buyers do for given gains from trade. Relative inequality between sellers and buyers is significantly higher when sellers have strategic power.[22] The, on average, less reciprocal behaviour by the sellers further drives up costs to trust for buyers and down the rate of return on trusting. Figure 1-4 shows that, over a large range of prices paid (from about 10 to 40 and above 70), the profitability of trusting is rather low for buyers and much lower than the profitability of trusting for sellers, which can be seen in Figure 1-3. Absolute rents for the strategically unfavoured side are, on average, also lower when sellers have strategic power: sellers in spB-treatments (17.20 on average) earn significantly more than buyers in spS-treatments (10.05 on average).[23] Given the lower profit expectations and the higher possible losses for trusting buyers in the spB-treatments compared to sellers in the spS-treatments, part of the quality differences can be attributed to sellers, as first movers, being more willing to provide a high level of quality than buyers are, in turn, willing to pay up-front to induce, on average, the same quality level by the sellers. This can also be shown in the data: even if all sellers played as reciprocally as to always equalise payoffs (and thereby reciprocated more than buyers in the spB-treatments), efficiency in the spS-treatments would only just about reach the level as efficiency in the

[22] Mann-Whitney U test, two-sided: p=0.020
[23] Mann-Whitney U test, two-sided: p=0.033

spB-treatments.[24] For any self-serving degree of reciprocity, which we have also observed in the spB-treatments, efficiency in the spS-treatments falls further below the level of the spB-treatments.

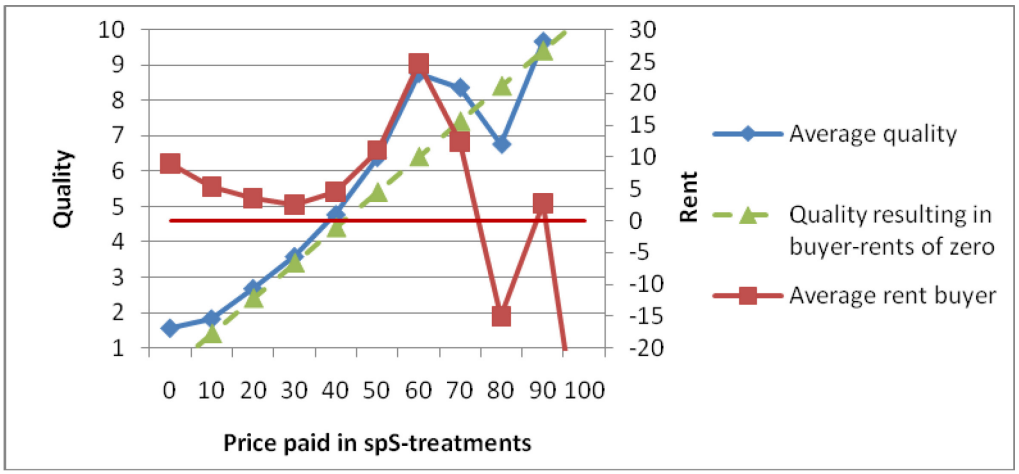


Figure 1-4: Quality-choices of sellers and implied profitability for buyers dependent on prices paid in spS-treatments

As we argued in section 1.3.2, strong dynamic incentives may enter through relational contracting, i.e. repeated trading between the same transaction partners. According to our notion of relational contracting, relationships consist of any type of first round trade plus at least one *private renewal trade*, i.e. a trade that was initiated through a private offer directed at the seller of the previous round. Hence, once a seller accepts a private renewal offer we think of the pair as engaging in relational contracting. Those trades within relations, however, only make up 31.74% of all trades in the spS-treatments and even less, 14.34%, in the spB-treatments. The dynamic incentives for most sellers in the spS-treatments and most buyers in the spB-treatments are therefore not strong. The efficiency effects of strategic power in our data are hence best explained by the differences in static incentives.

Result SP2: Efficiency is higher in the spB-treatments due to strong static incentives for sellers to provide high quality.

[24] We calculate a hypothetical average quality of 6.7 in the spS-treatments if all sellers always chose the payoff-equalising quality level. This would imply a stronger degree of reciprocity than buyers displayed in the spB-treatments. Even then average quality levels would just about reach the level of the spB-treatments (6.85).

1.5.2 The role of competitive power

Let us also recall that being on the short side of the market gives competitive power because each market participant may only conclude a single trade in each round, which, and as we have already seen, generates significant rents on average. Is having competitive power advantageous? The basic competitive power hypothesis H3.2 is supported by the data. Competitive power significantly shifts rent-sharing towards the favoured side. This means that holding strategic power constant, a side earns more if it has competitive power. Buyers earn significantly more in spS-cpB (13.10 on average) than in spS-cpS (6.98 on average) and they earn more in spB-cpB (34.86 on average) than in spB-cpS (30.85 on average),[25] whereas the opposite is true for sellers (see Figure 1-6). There is also a tendency towards rent-sharing being shifted towards the favoured side irrespective of strategic power, as illustrated in Figure 1-5. Overall, however, the impact of competitive power fails to reach a significant level since, as we will analyse in more depth later, the rent-sharing consequences of strategic power dominate.[26]

Result CP1: Holding strategic power constant, competitive power pays off for the favoured side.

Figure 1-5 also reveals that efficiency is unaffected by competitive power. Regardless of which side has competitive power, the transacting parties reach about 60% of maximum gains of trade. How does competitive power pay off?

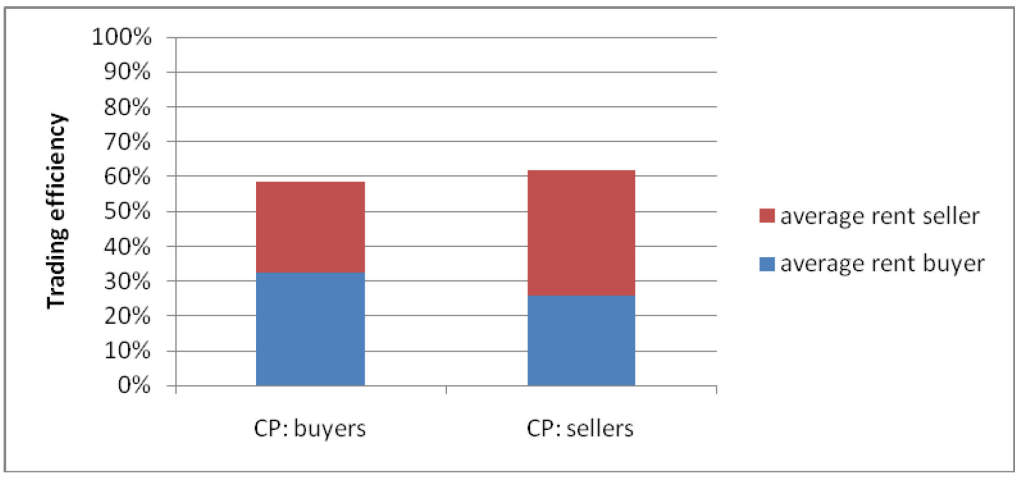


Figure 1-5: Effects of competitive power for rent-sharing and efficiency

[25] Mann-Whitney U test, two-sided: p=0.041 (spS-cpB vs. spS-cpS); p=0.026 (spB-cpB vs. spB-cpS)
[26] Mann-Whitney U test, two-sided: p=0.128 (cpB vs. cpS)

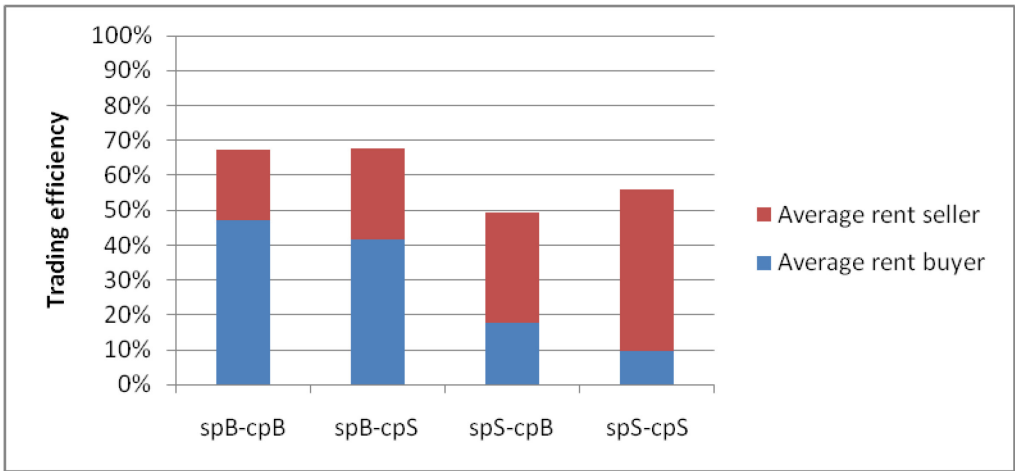


Figure 1-6: Efficiency and rent-sharing comparisons of all treatments

When sellers have competitive power, strong competition among buyers to find a transaction partner leads to significantly higher contracted prices (41.11 on average) compared to when buyers have competitive power (31.63 on average).[27] Hypothesis H3.1 therefore receives support from the data. The different degree of competition among buyers to find a seller is visible in the number of offers buyers make. Each buyer makes, on average, 1.82 offers in spB-cpS and 2.27 in spS-cpS, but only 1.16 offers in spB-cpB and 1.34 in spS-cpB. As in any round two buyers in every group of the cpS-treatment cannot conclude a trade, the number of offers in excess of possible trades is, on average, 116.33 in spB-cpS and 163.17 in spS-cpS, but only 11.83 in spB-cpB and 25.5 in spS-cpB.

Result CP2: Contracted prices are higher in cpS-treatment than in cpB-treatments.

In case sellers have strategic power, higher contracted prices directly shift rents from buyers to sellers if quality levels are unchanged; in fact, quality levels are statistically indistinguishable between the two treatments,[28] despite higher prices in spS-cpS compared to spS-cpB.

Result CP3: When contracted prices are binding (spS-treatments), sellers profit from competitive power as higher contracted prices directly translate into higher seller-rents.

[27] Mann-Whitney U test, two-sided: cpB vs. cpS: p=0.017
[28] Mann-Whitney U test, two-sided: p=0.310

In case buyers have strategic power, contracted prices are not technically linked to actual prices as buyers have full discretionary power to re-set prices; indeed, we do not find significantly higher prices in spB-cpS compared to spB-cpB even though average values are somewhat different (33.79 versus 29.53 respectively).[29] Nevertheless, competitive power also influences rent-sharing when contracted prices are non-binding. The reason is that buyers are more reciprocal and pay a higher price for a given quality level. We use the ratio of actual prices over quality $R_B = p/q$ to measure buyers' reciprocity, which is weakly significantly higher in spB-cpS than in spB-cpB.[30]

Result CP4: When contracted prices are non-binding (spB-treatments), sellers profit from competitive power by receiving higher prices for a given quality level.

1.5.3 The interaction between competitive and strategic power

Result CP3 already hints at an interacting effect between competitive and strategic power, which we now look at in a more general way. The side with strategic power acts in more reciprocal way if it does not have competitive power. In case sellers have strategic power, they provide higher quality for a given price level if they do not have competitive power. Sellers' reciprocity, the ratio of quality over price $R_S = q/p$, is significantly lower in spS-cpS than in spS-cpB.[31] This mirrors result CP3, which was derived for the spB-treatments. The common underlying mechanism is that dynamic incentives change with competitive power. Members of the side without competitive power face strong competition to find a transaction partner. Compared to having competitive power themselves, they have, therefore, a higher incentive to please their current transaction partners in order to increase their chance of being re-matched in the next round.[32] By looking at individual data, we can analyse more deeply how the behaviour of the side with strategic power is affected by competitive power. If the side with strategic power does not have competitive power, more subjects play in a manner that gives their transaction partners at least an equal profit. Let us start with spB-treatments: if buyers have both strategic and competitive power, only 6.7% of

[29] Mann-Whitney U test, two-sided: p=0.310

[30] Mann-Whitney U test, two-sided: p=0.065 based on group level comparisons of $1/n\sum_{i=1}^{n}(p_i/q_i)$, n being the number of trades in a group

[31] Mann-Whitney U test, two-sided: p=0.015 based on group level comparisons of $1/n\sum_{i=1}^{n}(q_i/p_i)$, n being the number of trades in a group

[32] Also recall from section 1.3.2 that a money-maximising player of the strategically favoured side would only pay above-minimum prices or deliver above-minimum quality if he expects this to pay off dynamically.

buyers reciprocate on average as much as to give their transaction partners rents that are at least high as the buyers' own rents. If, by contrast, sellers have competitive power, 26.2% of buyers reciprocate on average as much as for their transaction partners to receive an at least equal payoff. There are similar effects in case sellers have strategic power on the other end of the behavioural spectrum: 43.3% of sellers play in a way that leads on average to negative rents for their transacting buyers when they have both competitive and strategic power; this percentage drops to 28.6% if sellers only have strategic but not competitive power. Hence, if buyers have strategic power, giving competitive power to sellers induces more buyers to play in a payoff-equalising way, whereas if sellers have strategic power, shifting competitive power to the buyers leads to less sellers acting in a non-reciprocal way. Interestingly, the reciprocity-diminishing effects of competitive power happen at different ends of the behavioural spectrum depending on who has strategic power.[33]

Result SP-CP1: Competitive power leads to less reciprocity by the side with strategic power.

The above analysis provides us with the means to explain why efficiency is unaffected by competitive power. In both cases of strategic power, two counter-running effects cancel each other out. For the spB-treatments, endowing buyers also with competitive power leads on the one hand to higher dynamic incentives for sellers to provide above-minimum quality in order for them to increase their chance of receiving an offer in the next round. On the other hand, sellers face a lower profitability of their investments in above-minimum quality as buyers, in turn, have a lower dynamic incentive to please sellers if they have competitive power (see result CP4).[34] In the spS-treatments, sellers act less reciprocally if they are also endowed with competitive power (compare result SP-CP1). At the same time, sellers receive higher prices (see result CP3) to which reciprocal players react with increased quality levels. The resulting net-effect of competitive power on quality is zero in both cases of strategic power.

[33] In case sellers have strategic power and also competitive power, 6.7% of sellers play in a non self-favoured way, whereas this percentage rises to 11.9% if buyers have competitive power. If buyers have strategic power, less than 10% play in a way that leads to negative rents for sellers, regardless of competitive power.

[34] Sellers expect a weakly significantly (Mann-Whitney U test, two-sided: p=0.093) lower degree of reciprocity when buyers have competitive power. This may be interpreted one the one hand as sellers being content with a lower price for a given quality level when they do not have competitive power. On the other hand, it may show that the lower dynamic incentives of buyers to please sellers also entered sellers´ expectations. p-value calculation is based on averages over ratios of expected price over quality.

Result SP-CP2: Efficiency is unaffected by competitive power as two counter-running effects cancel each other out in both cases of strategic power.

The analysis of strategic power already confirmed the dominating effect of strategic power for rent-sharing (see result SP1) as the strategically favoured side always, i.e. regardless of competitive power, gains a larger share of rents (compare Figure 1-6). Hence, in the mixed-power treatments the side with strategic power (sellers in spS-cpB and buyers in spB-cpS) can expect to gain a significantly larger share of rents.[35] Consequently, changes in rent-sharing and profits are larger from spS-treatments to spB-treatments than from cpB-treatments to cpS-treatments. The share of buyer-rents drops rather modestly from 55.42% to 41.25% (or from 23.98 to 19.29 in absolute rents) by shifting competitive power from buyers to sellers. By contrast, the share of buyer-rents decreases considerably, from 65.63% to 25.71% (or from 32.85 to 10.05 in absolute rents), if strategic power shifts from buyers to sellers.

Result SP-CP3: Strategic power dominates rent-sharing.

Furthermore, as hypothesised in H4.2, relative inequality between buyers and sellers is lower in the two mixed-power treatments spS-cpB and spB-cpS than in the two concentrated-power treatments spS-cpS and spB-cpB, in which one side is favoured twice.[36]

Result SP-CP4: Relative inequality is lower in mixed-power than in concentrated-power treatments.

The differences in rent-sharing also translate into absolute differences in rents. The following descending order for rents of buyers emerges: spB-cpB > spB-cpS > spS-treatments. While the order between spB-cpB, spB-cpS and either of the spS-treatments is (weakly) significant,[37] rents for buyers are only somewhat higher in spS-cpB than in spS-cpS.[38] For sellers, we find the following corresponding order of descending rents: spS-cpS > spS-cpB > spB-treatments.[39] We again observe strategic power to dominate rent-sharing so that both sides are better off in absolute terms if they have strategic power. Hypothesis H4.1, which

[35] Wilcoxon signed ranks test, two-sided: p=0.031 in both cases
[36] Mann-Whitney U test, two-sided: p=0.002 for both concentrated-power treatments vs. both mixed-power treatments
[37] Mann-Whitney U test, two-sided: p=0.065 (spB-cpB vs. spB-cpS); p=0.002 (spB-cpB vs. either spS-cpB or spS-cpS)
[38] Mann-Whitney U test, two-sided: p=0.132
[39] Mann-Whitney U test, two-sided: p=0.002 (spS-cpS vs. either spS-cpB, spB-cpS or spB-cpB); p=0.002 (spS-cpS vs. spS-cpB), p=0.053 (spS-cpB vs. cpB-treatments)

predicts the favoured side in concentrated-power treatments to be better off, consequently also receives supported by the data. Interestingly, competitive power only affects absolute rents in the interaction with strategic power. Both buyers and sellers earn more when competitive power is added to strategic power; by contrast, when they do not have strategic power, both sides do not earn significantly more if they are given competitive power.

Result SP-CP5: The side with strategic power is always better off, while competitive power only changes absolute rents when added to strategic power.

1.6 The role of relational contracting

1.6.1 Treatment differences are mainly driven by trades outside relations

Looking at rent-sharing and efficiency in different types of trades (compare Figure 1-7) we see treatment differences to be mainly driven by trades outside relations.[40] There are no treatment-differences in quality levels within relations.[41] Prices within relations are only different between spS-cpB and spS-cpS; in this case, prices are higher if sellers have competitive power.[42]

Result RC1: Within relations, quality levels do not differ between treatments, and prices only differ between spS-cpB and spS-cpS.

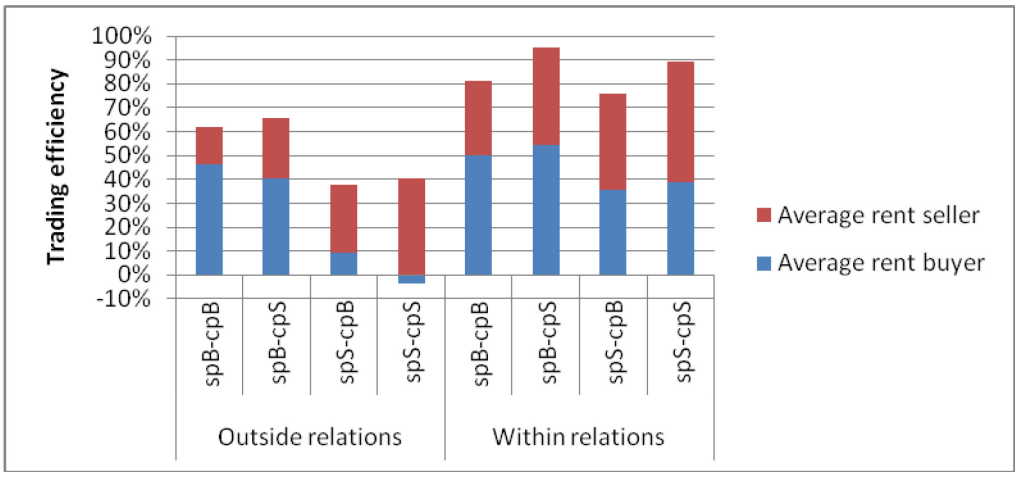


Figure 1-7: Efficiency and rent-sharing within and outside relations

[40] Recall that we consider relational contracting to take place as repeated trading through private offers.
[41] Kruskal-Wallis-Test, p=0.136 for quality; however, a Mann-Whitney U test reveals a weak two-sided significance (p=0.093) for a higher quality level within relations in spS-cpS than in spS-cpB.
[42] Mann-Whitney U two-sided: p=0.009

There are also no treatment differences in relative inequality within relations as rent-sharing is very pronounced within relations.[43] The shares of buyer rents are relatively close to 50% in all treatments (43.43% in spS-cpS, 47.05% in spS-cpB, 57.38% in spB-cpS and 61.24% in spB-cpB). Outside relations, by contrast, relative rents depart considerably more from 50%, as can be seen from Figure 1-7; consequently, relative inequality is significantly higher outside relations than within relations.[44]

Result RC2: Rent-sharing is more pronounced within than outside relations.

Figure 1-7 also impressively shows that within relations transaction partners manage to seize a large part of the available gains from trade. Trading efficiency is above 78% in all treatments, whereas the corresponding trading efficiencies for trades outside relations lie between 37.1% and 65%.[45] Trading efficiency is consequently highly significantly higher within than outside relations, and so are price levels.[46]

Result RC3: Efficiency and price-levels are higher within than outside relations.

Trades within relations are less dependent on market conditions and contract structure than trades outside relations. The fairness norms inherent in relations appear not to take the trading partners' respective outside options on the spot-market fully into account, which are, as our results show, strongly influenced by the allocation of strategic and competitive power; this also implies that the weak side outside relations gains more from being inside a relation. Despite the large cooperative gains from relational contracting, the strong side outside relations gains little (sellers in spS-treatments) or, interestingly, even nothing (buyers in the spB-treatments).[47] In the concentrated-power treatments, the favoured side gains in terms of absolute rents on average as little as 3.85 (sellers in spS-cpS) and 2.19 (buyers in spB-cpB) from relational contracting.

Result RC4: The side without strategic power gains more from relational contracting.

[43] Kruskal Wallis Test for relative inequality across all four treatments, two-sided: p=0.504

[44] Wilcoxon signed ranks test for all 21 groups in which private relations emerged, two-sided: p<0.001

[45] Trading efficiency for trades within and outside relations respectively for each treatment: 89% vs. 62.12% (spB-cpB); 94.5% vs. 65.3% (spB-cpS); 78.2% vs. 37% (spS-cpB); 88.9% vs. 39.6% (spS-cpS)

[46] Wilcoxon signed ranks test for all 21 groups that exhibited relational contracting, two-sided: p<0.001 for both efficiency and prices. Separate tests for every treatment reveal at least weakly significant results (two-sided) on both efficiency and price-levels for all treatments but spB-cpS, in which only 3 groups engaged in relational contracting; nevertheless, the evidence in spB-cpB goes into the same direction as prices and efficiency were also higher within relations wherever relations were formed.

[47] Rents of buyers are not significantly higher within than outside relations.

1.6.2 Are the treatment effects dependent on the subjects´ reliance on relational contracting?

As we already argued in section 1.5.1, the degree of relational contracting is fairly low. Consequently, sellers face rather weak dynamic incentives so that differences in static incentives drive the higher efficiency when buyers, and not sellers, have strategic power. Looking beyond the results of our own experiment supports this view. Comparing the results of our spS-cpB treatment to previous experiments using an almost identical design, our results apparently contain less relational contracting and at the same time lower efficiency than BFF´s (2004) ICF-treatment but similar levels of relational contracting and efficiency as WR's (2007a) IC1-treatment. 45% of all trades in BFF´s (2004) ICF-treatment take place in relationships that lasted more than 5 rounds,[48] and roughly 2/3 of trades occur in relationships of any length. In our spS-cpB treatment, less than 12% of trades occur in relationships of more than 5 rounds, and 42.1% of trades take place within any kind of repeated interactions.[49] WR (2007a) do not report on length of relations but on the share of private trades. The share of private trades may be considered a rough – albeit very imprecise – proxy for the degree of relational contracting as relations are mostly initiated through private offers. They report on 51.4% of private trades, which seems considerable less than the roughly 70%[50] of private trades in BFF-ICF but about the same as the 48.3% of private trades in our spS-cpB treatment. The differences in relational contracting coincide with differences in trading efficiency: BFF's (2004) ICF-treatment boasts an average quality level of 6.9, while WR-IC1 only reaches an average quality of 5.4, which seems roughly the same as the average quality level of 5.2 in our spS-cpB treatment.

The degree of relational contracting in our data may have repercussions on the treatment effects studied. The cross-experimental observation, that BFF´s (2004) data entail both more relational contracting and higher efficiency compared to us and WR (2007a), fits our own inter-experimental observation that efficiency is higher within relations. We have also shown that there are no treatment differences in prices and relative inequality within relations. An

[48] Reported in BFF (2008)

[49] Figure also includes repeated trading by public offers. We think this corresponds to BFF's (2004) definition of relationships which seems to be wider than ours as there is no explicit reference therein to how relations are initiated. The share of trades based on our narrow definition of relationships, which is restricted to explicitly initiated relations (i.e. through private renewal trades) is 30.3%.

[50] Own calculation based on figure 1 in BFF (2004)

obvious follow-up research question therefore is how robust the effects from strategic and competitive power are to the degree of relational contracting.

Possibly, the discrepancies between the results of BFF (2004), WR (2007a) and us are due to a different underlying tendency in the subject pool towards relying on relational contracting as a contract enforcement mechanism.[51] Subjects may have a different understanding of relational contracting as a contract enforcement device, possibly rooted in different abilities to behave strategically in the experiment or in distinct experiences outside the laboratory. These experiences may vary markedly between different subject pools because of unlikeness in formal institutions (such as the judicial system) and informal institutions (such as culture).

If the differences in the level of relational contracting between BFF (2004) on the one hand and us as well as WR (2007a) on the other hand are indeed due to a subject pool effect, we may be able to answer our follow-up research question by replicating the experiment in a subject pool that we expect to be more conducive to relational contracting than our subjects in Erfurt, Germany. For reasons discussed in the following section, we hypothesise to find this subject pool in the People´s Republic of China. The results of the sessions in Chengdu, China, are presented and discussed in section 1.8. A cross-cultural account of the nature of relational contracting is given in section 1.9. Section 1.10 provides a summary, and section 1.11 concludes the paper with a discussion as well as an outlook for future research.

1.7 Cross-cultural hypotheses

Why do we choose China as the control country for our experiment? In contrast to Western countries, relational contracting is said to play a paramount role in business interactions in China (Solinger 1989; Xin and Pearce 1996; Luo and Chen 1997; Li 2002), and (McMillan and Naughton 1996) see contracts in China to be less legal than relational; therefore, the first step to enter a successful business in China is to create a reliable and effective net of relations (Yeung and Tung 1996). The extensive use of informal contract enforcement in China such as self-enforcement through a long-term cooperative solution (see Clark, Murrell and Whiting 2006) may be due to malfunctioning formal contract enforcement institutions (Dixit 2003; Li 2003) or to a cultural emphasis on relations rooted in Confucianism. In China

[51] There are more differences between WR (2007a), BFF (2004) and us such as larger groups in BFF (2004) with a slightly different asymmetry between buyers and sellers, differences in stakes, instructions and so on. While we are not able to rule any of these differences out as explanatory factors, none seems a good candidate to explain the differences in relational contracting.

networks of inter-personal relationship are said to be a widespread socio-cultural phenomenon (King 1991). The Chinese word *guanxi* is becoming a standard term for understanding inter-personal networks (see for example Arias 1998; Farh et al. 1998); it is defined by (Chen and Chen 2004) as "an implicit psychological contract between people to follow social norms such as maintaining long-term relationship". Based on this view, *guanxi* is not restricted to business interactions only.

According to the 5-dimension culture model (Hofstede 2001), Chinese are comparatively very long-term oriented. Out of all 23 investigated countries, China is ranked the country that pays, according to the 5th dimension "long/short term orientation", most attention to long-term relationship whereas Germany is placed at rank 14. The 5th dimension is related to the choice of focus for people's efforts: the future or the present."[52] This behavioural trait should therefore facilitate relational contracting. Translating these findings into our market-experiment, we hypothesise that we will see more relational contracting in Chinese sessions than in German sessions.[53]

H6: Subjects will rely more on relational contracting in the Chinese sessions than in the German sessions.

As we have already seen from the data of the German sessions, trades within relations are characterised by a higher trading efficiency than trades outside relations. If Chinese subjects do indeed rely more on relational contracting, as stipulated in H6, we also expect trading efficiency to reach a higher level in the Chinese than in German sessions.

H7: Trading efficiency is higher in the Chinese sessions than in the German sessions.

The higher share of trades within relations, as hypothesised in H6, means simultaneously a lower level of trades outside relations. In our German experiment, treatment effects on rent-sharing and trading efficiency are driven by trades outside relations. Consequently, we predict these treatment effects to be smaller in Chinese sessions than in German sessions.

[52] In the first version of Hofstede's work (1984), the "long/short term orientation" is not a dimension of the model due to the "Western minds of the designers" (Hofstede 2001, p.351) of the study. The Chinese Value Survey (CVS) found this new dimension, which is added to the second version of Hofstede's work (2001).

[53] For reasons of readability, we speak of German and Chinese data and sessions as well as Germany and China as the respective locations of the experiment. This does, however, not imply that our findings are representative for either Germany or China.

H8: Treatment effects on trading efficiency and rent-sharing are smaller in the Chinese sessions than in the German sessions.

1.8 Results of the Chinese sessions

1.8.1 Relational contracting

Recall from section 1.4 that we define relational contracting as occurring through privately initiated renewed trades between the same pair of transaction partners plus their initial rounds. We consequently use the share of trades within relations to all trades as a measure for the degree of relational contracting in the Chinese and German sessions. The higher is the degree of relational contracting, the stronger subjects rely on relationships in order to enforce incomplete contracts.

Figure 1-8 shows the share of trades within relations to be higher in the Chinese than in the German sessions. Comparison of group averages reveals that the difference is indeed significant. [54] Our hypothesis H6, that Chinese subjects rely more on relational contracting than the German subjects, therefore receives support from the data. In addition to this, the average duration of a relation is 4.75 rounds in China, which is significantly longer than the corresponding average duration of 3.74 rounds in Germany.[55] Hence, if a relationship is established, it survives longer in the Chinese sessions than in the German sessions.

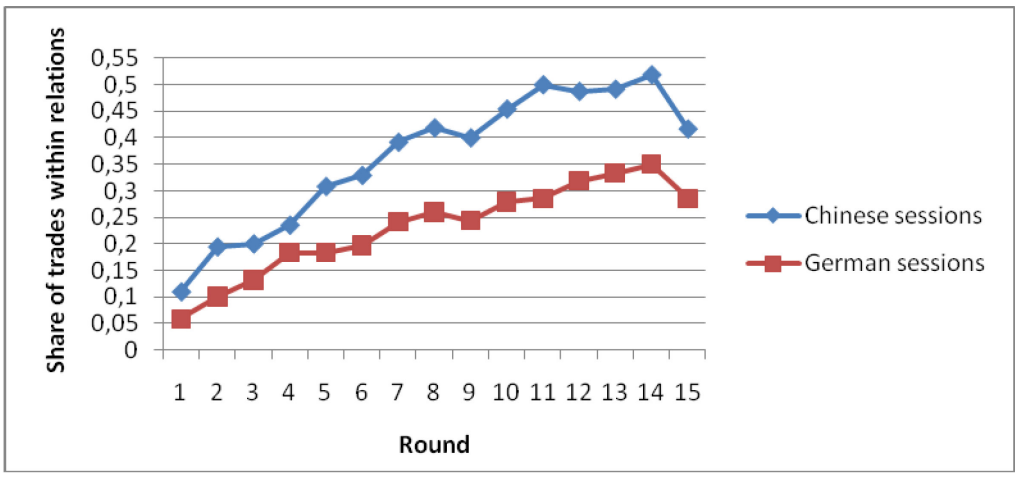


Figure 1-8: Share of trades within relations (private renewal trades plus initial round) in the Chinese and German sessions

[54] Mann-Whitney U test, two-sided: p=0.013
[55] Mann-Whitney U test, two-sided: p=0.017

1.8.2 Cooperation and efficiency

We predicted trading efficiency to be higher in China than in Germany if there is more relational contracting in China. On average, Chinese groups reached a quality level of 7.07 which is significantly higher than the corresponding quality of the German sessions (6.18).[56] Our hypothesis H7 is, therefore, also supported by the data.

Table 1-4: Descriptive summary statistics of the Chinese treatments

	spS-cpB	spS-cpS	spB-cpB	spB-cpS
Market efficiency	99.33%	99.56%	98.22%	98.67%
Mean price	42.79	51.23	30.75	35.78
Mean quality	6.96	7.11	7.11	7.10
Mean seller rent	27.82	35.96	15.41	20.58
Mean buyer rent	22.83	15.85	36.35	31.16
Share of buyer rents	45.07%	30.59%	70.23%	60.22%
Share of trades within relations	51.06%	43.59%	27.67%	23.14%

We have seen that relational contracting is at a higher level in China than in Germany; this allows us to answer our second main research question whether the rent-sharing and efficiency consequences of strategic and competitive power are robust to a higher level of relational contracting.

1.8.3 When relations meet strategic power

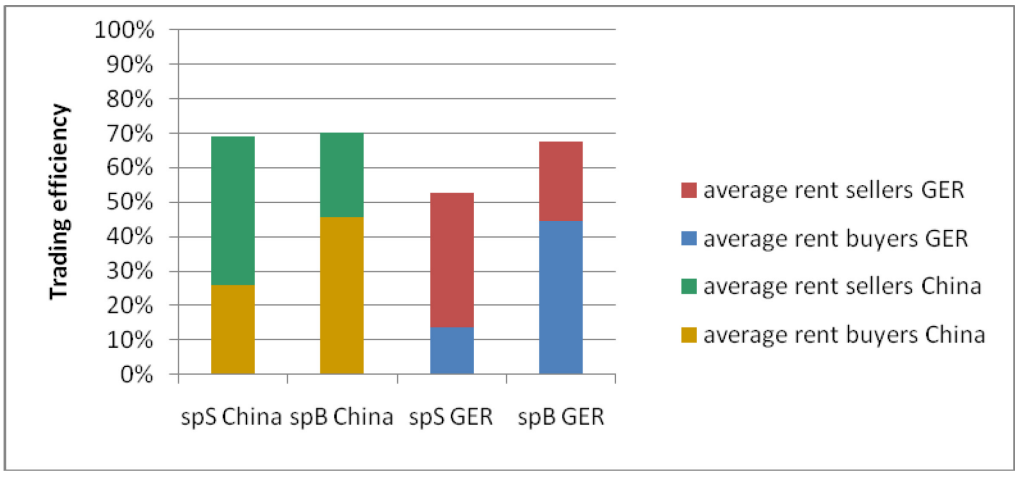


Figure 1-9: Effects of strategic power on rent-sharing and efficiency in China compared to Germany

[56] Mann-Whitney U test, two-sided: p=0.009

The first two columns of Figure 1-9 present the efficiency and rent-sharing consequences of strategic power in the Chinese sessions, whereas the last two columns re-show the respective results from the German sessions (Figure 1-2). It becomes obvious that the efficiency effect of changing strategic power does not exist in the Chinese sessions. The average quality is 7.04 if sellers have strategic power and 7.10 if buyers are strategically favoured, which translates into a trading efficiency of about 70% in both cases.[57] A Mann-Whitney U-test does not detect a different quality level when strategic power is switched from one side to the other.[58] Thus, our hypothesis H2.2, that switching strategic power from seller to buyer leads to a higher efficiency level, has to be rejected for the Chinese data. In the spS-treatments, quality levels are significantly higher in China than in Germany (5.52 on average),[59] while quality-levels in the spB-treatments are almost identical in both countries (on average, 6.85 in Germany).[60]

In order to understand why quality-levels are higher in the Chinese spB-treatments than in the German spB-treatments and efficiency thereby unaffected by strategic power in the Chinese sessions (in contrast to Germany), let us look at the incentives of sellers to provide above-minimum quality. As we discussed in section 1.3.2.1, sellers' incentives to provide quality change with strategic power. In a market where buyers have strategic power, sellers are motivated to provide a high quality if they expect a reciprocal behaviour from buyers. By contrast, in markets where sellers have strategic power, quality choices reflect sellers' reciprocity to buyers' price offers. If buyers deliver a high price offer, it can trigger high quality choices by sellers. Reciprocal sellers would reciprocate both within and outside relations, whereas money-maximising sellers would only reciprocate if they can thereby enter or stay within a rent-generating relation.

The reciprocal behaviour by sellers is also shown by the data: sellers' quality choices are strongly correlated with buyers' prices. Within-group Spearman's rank correlation coefficients vary from 0.492 to 0.888 in spS-treatments; the binomial test rejects the null hypothesis of an equal probability of positive and negative correlation coefficients.[61] This

[57] Recall that trading efficiency is defined as achieved gains from trade relative to maximum gains from trade. Trading efficiency is about 2.7% (1/37) at minimum quality (q=1) and 100% at maximum quality (q=10).
[58] Mann-Whitney U test, two-sided: p=0.831
[59] Mann-Whitney U test, two-sided: p=0.001
[60] Mann-Whitney U test, two-sided: p=0.399
[61] Binomial test, two-sided: p<0.001

finding allows us to find a first explanation why efficiency is higher in the Chinese spS-treatments than in the German spS-treatments.[62] When sellers have strategic power, Chinese buyers choose significantly higher prices than German buyers.[63] As Chinese sellers display, on average, a level of reciprocity at least as high as German sellers, higher prices in Chinese spS-treatments imply a higher quality level.[64] Why do Chinese buyers pay higher prices? We connect this question with our original idea about relational contracting. As expected, we find that the degree of relational contracting is significantly higher in China than in Germany when sellers have strategic power (47.37% in China versus 31.73% in Germany).[65] Re-examining price-differences between German and Chinese data reveals that Chinese buyers in spS-treatments do not pay significantly higher prices within relations (53.22 in China versus 51.39 in Germany) but tend to pay more outside relations (41.05 in China versus 36.1 in Germany);[66] as prices are higher within relations than outside relations,[67] higher average prices in Chinese spS-treatments compared to German spS-treatments are therefore partly driven by a higher level of relational contracting but also to somewhat higher prices outside relations. Outside relations, the somewhat higher prices bring back a significantly higher quality level in China than in Germany (5.59 in China versus 4.20 in Germany).[68] By contrast, quality-levels within relations are indistinguishable between the Chinese and the German spS-treatments (8.65 in China versus 8.29 in Germany) and, in China as in Germany, are higher than outside relations.[69] Therefore, quality is higher in the Chinese spS-treatments partly as a direct consequence of a higher level of relational contracting; it is furthermore a result of a higher quality level outside relations.

While the data do not allow finding a comprehensive explanation for somewhat higher prices and higher quality outside relations in Chinese spS-treatments compared to German spS-treatments, relational contracting may well play a role. More cooperative behaviour outside relations in the Chinese spS-treatments may also be a result of stronger attempts to

[62] Mann-Whitney U test, two-sided: p=0.002
[63] Mann-Whitney U test, two-sided: p=0.060
[64] Both the correlations between prices and quality (Mann-Whitney U test for differences in group-level correlation coefficients, two-sided: p=0.525) and the level of reciprocity, as measured by q/p, (Mann-Whitney U test, two-sided: p=0.319) are indistinguishable between the Chinese and the German spS-treatments, with average level of reciprocity (q/p) being slightly higher in the Chinese than in the German spS-treatments (0.155 in China versus 0.151 in Germany).
[65] Mann-Whitney U test, two-sided: p=0.007
[66] Mann-Whitney U test, two-sided: p=0.713 (within relations); p=0.178 (outside relations)
[67] Wilcoxon signed ranks test, two-sided: p<0.001
[68] Mann-Whitney U test, two-sided: p=0.024
[69] Wilcoxon signed ranks test, two-sided: p<0.001

create relations or to spill-over effects from a higher level of relational contracting. For the Chinese spS-cpB treatment, we do find some evidence for such spill-over effects: quality outside relations is strongly, but only weakly significantly, correlated with the share of relations of all trades in a group.[70] This may be interpreted as the stronger dynamic for sellers to perform arising from a higher level of relational contracting spilling over to trades outside relations. The more sellers and buyers form relations, the harder it is for the remaining sellers, who are on the long side of the market, to enter a relation. Hence, once a seller concludes a trade her incentives to hold onto the buyer rises with the share of trades within relations in the group. Interestingly, in the German spS-cpB treatment, the correlation between share of relations and quality outside relations in the six groups is not strong enough to reach significance.[71] Nevertheless, based on the correlation in the Chinese data a higher level of relational contracting in the Chinese compared to the German spS-cpB treatment may also drive higher quality outside relations in the Chinese compared to the German spS-cpB treatment.[72] For the spS-cpS treatments, such spill-over effects cannot be found, nor can we establish conclusive evidence for differences in attempts to create relations (which may be too subtle to be detected). Therefore, an alternative interpretation is that our Chinese subjects may be more cooperative, for reasons unrelated to the dynamic incentives from relational contracting (we will come to this question in more depth in section 1.9.2).

Interestingly, switching strategic power from sellers to buyers seizes to affect efficiency in China; nor are, in fact, quality levels in the Chinese spB-treatments higher than in the German spB-treatments. At first sight, this seems puzzling; the level of relational contracting is still higher in the Chinese sessions than in German sessions when buyers have strategic power (25.40% in China and 14.34% in Germany) – albeit less so than in the spS-treatments and only at a weakly significant level.[73] Hence, on a group-level, the stronger dynamic incentives arising for sellers to perform from a higher level of relational contracting may not simply be added to stronger static incentives when buyers have strategic power. This is also visible in that we find no positive correlation between the share of relations and the quality level in the Chinese spB-treatments. Looking at individual trades, we nevertheless find, as in

[70] Spearman's rho = 0.771, p=0.072 (two-sided)
[71] Spearman's rho=0.348, p=0.499 (two-sided).
[72] Mann-Whitney U test, two-sided, for differences in the level of relational contracting: p=0.009.
[73] Mann-Whitney U test, two-sided: p=0.091

Germany, trades within relations to be characterised by higher efficiency (9.03 inside relations versus 6.34 outside relations).[74] On a mechanical level, we can therefore say that the level of relational contracting is not sufficiently higher in the Chinese spB-treatments compared to the German spB-treatments to increase efficiency above the level of the corresponding German treatments. Behaviourally, however, we still need an explanation why the dynamic and static incentives do not seem to add up; in fact, they seem to have a substitutive relation: on a group-level, the share of trades within relations is significantly and strongly negatively correlated with quality-levels outside relations.[75] Hence, it seems the higher the quality that sellers provided outside relations, the less did buyers rely on relations in order to enforce contracts.

To conclude, the effect of strategic power on efficiency found in our German data as well as by WR (2007a) is not robust. If subjects rely on relational contracting to a high degree, as in the Chinese sessions, efficiency is raised when sellers have strategic power but not when buyers have strategic power. The reason seems to be Chinese buyers relying less on relational contracting if their profit prospects outside relations are good. The potential efficiency benefits from more relational contracting (in China compared to Germany) is thereby lost when static incentives for sellers to perform are already strong.

> ***Result C-SP1: Efficiency is not affected by strategic power in China. Compared to the corresponding German treatments, efficiency is higher in Chinese spS-treatments but not higher in Chinese spB-treatments.***

From Figure 1-9, we can see that also in China the strategically favoured side gets a larger share of rents, both compared to their trading-partners and to an environment in which they are the strategically unfavoured side. In the spB-treatments, buyers earn a significantly higher share of rents than sellers and than buyers in the spS-treatments.[76] The same holds when sellers have strategic power. Sellers gain a significantly higher share of rents than buyers and compared to their counterparts in the spB-treatments.[77] The finding in Germany

[74] Wilcoxon signed ranks test, two-sided: $p<0.001$
[75] Spearman's rho: -0.706, $p=0.01$ (two-sided)
[76] Wilcoxon signed ranks test, two-sided, for differences in rents between sellers and buyers in spB-treatments: $p<0.001$; Mann-Whitney U test, two-sided, for differences in buyer-rents between spB-treatments and spS-treatments: $p<0.001$
[77] Wilcoxon signed ranks test, two-sided, for differences in rents between sellers and buyers in spS-treatments: $p<0.002$; Mann-Whitney U test, two-sided, for differences in seller-rents between spB-treatments and spS-treatments: $p<0.001$

that strategic power pays off for strategically favoured side is therefore robust to the degree of relational contracting.

Result C-SP2: Strategic power pays off for the strategically favoured side also under more relational contracting.

1.8.4 When relations meet competitive power

The first two columns of Figure 1-10 depict the consequences of competitive power for trading efficiency and rent-sharing in the Chinese sessions, while the last two columns represent the respective results from the German sessions (compare Figure 1-5). As in Germany, quality levels and thus trading efficiency is not affected by competitive power. On average, quality is 7.04 when buyers have competitive power and 7.11 when sellers have competitive power; in both cases, about 70% of maximum gains from trade are reached (see Figure 1-10).

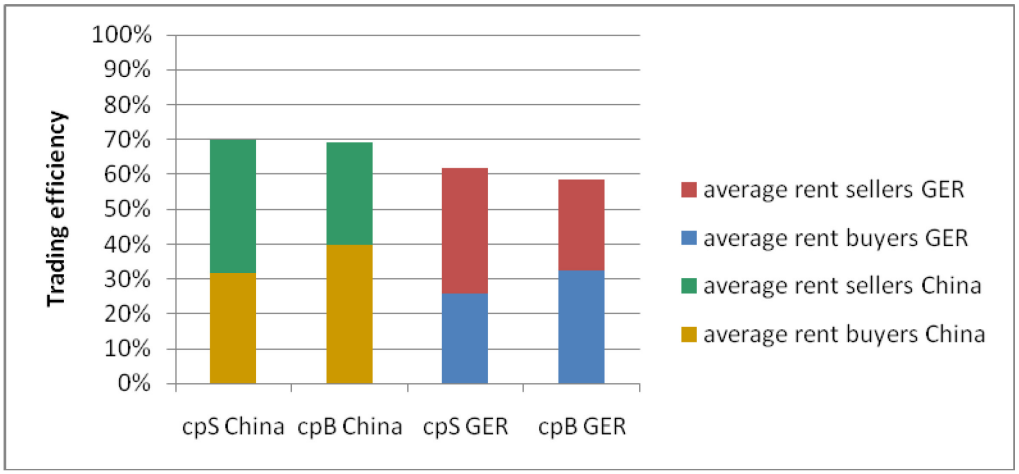


Figure 1-10: Effects of competitive power for trading efficiency and rent-sharing in China compared to Germany

Nevertheless, if we compare the Chinese data with the German data, we find that trading efficiency is higher in China when buyers have competitive power.[78] In the spirit of hypothesis H7, we test whether the higher efficiency level coincides with a higher level of relational contracting. We, indeed, find this to be the case. The share of trades within relations is significantly higher in the Chinese cpB-treatments than in the German cpB-

[78] Mann-Whitney U test, two-sided: p=0.043

treatments.[79] In the cpS-treatments, however, the level of relational contracting is not different between the two countries.

Result C-CP1: Trading efficiency is not affected by competitive power in China, but is higher in the Chinese cpB-treatments than in the German cpB-treatments.

As in Germany, we observe that contracted prices are higher when buyers face competition to find a seller, i.e. when sellers have competitive power.[80] Our hypothesis H3.1 is hence confirmed also in China. Actual trading prices also seem to be higher in the cpS-treatments than in the cpB-treatments, but fail to reach a significance level.[81] When sellers have strategic power, higher contracted prices directly result in higher actual prices as buyers are bound to pay what they offered.[82] When buyers have strategic power, to the contrary, contracted prices are not binding and buyers can arbitrarily adjust their actual prices after sellers made their quality choices. As the data shows, buyers use this advantage so that actual trading prices are not different between spB-cpS and spB-cpB treatments.[83]

Result C-CP2: Contracted prices are higher in cpS-treatments than in cpB-treatments, but actual prices are only significantly affected when contracted prices are binding.

From Figure 1-9, it seems that the competitively favoured side earns relatively more, both in comparison to their trading party and their counterparts that are unfavoured by competitive power. However, differences in rent-sharing do not reach a significant level overall.[84] In absolute terms, however, sellers earn weakly significantly more in cpS treatments than in cpB treatments, while no significant difference can be found for buyer-profits.[85] If we hold strategic power constant, the advantage of competitive power in rent-sharing nevertheless becomes obvious. In treatments where sellers are the last mover (spS-treatments), sellers' shares of rents are significantly higher when they have competitive power.[86] The same is true for treatments in which buyers are the last mover, despite prices being non-binding.

[79] Mann-Whitney U test, two-sided: p=0.034
[80] Mann-Whitney U test, two-sided: p=0.014
[81] Mann-Whitney U test, two-sided: p=0.114
[82] Mann-Whitney U test, two-sided: p=0.041
[83] Mann-Whitney U test, two-sided: p=0.310
[84] Mann-Whitney U test, two-sided, for buyers' share of profits (cpS vs. cpB treatment): p=0.128; Wilcoxon signed ranks test for sellers' profit vs. buyers' profit, two-sided: p=0.424 (cpS treatments); p=0.151 (cpB treatments).
[85] Mann-Whitney U test, two-sided, for sellers' profits: p=0.089 (cpS vs. cpB treatments); Mann-Whitney U test, two-sided, for buyers' profit: p=0.266 (cpS vs. cpB treatments)
[86] Mann-Whitney U test, two-sided: p=0.026

Buyers earn a weakly significantly higher share of rents in spB-cpB than in spB-cpS.[87] The reason is that, as in Germany, buyers reciprocate weakly significantly more if they do not have competitive power.[88]

Result C-CP3: Holding strategic power constant, competitive power pays off for the favoured side.

1.8.5 Relational contracting and the interaction between strategic and competitive power

As stipulated in hypothesis H8, we expect smaller treatment effects given a higher degree of relational contracting. The data provides some support for H8. First, trading efficiency is almost identical across all Chinese treatments, the standard deviation over the treatment averages being as low as 0.07. Neither strategic power nor competitive power consequently has an effect on total gains from trade (see data in table 4), whereas in Germany trading efficiency varies with strategic power. Second, relative inequality is lower in China than in Germany when sellers have strategic power but not different when buyers have strategic power.[89] This is in line with our finding that there is more relational contracting in China than in Germany only when sellers are strategically favoured. Therefore, as we hypothesized in H8, if subjects rely more on relational contracting, as in the Chinese spS-treatments, rent-sharing is more pronounced and, hence, the treatment-effects on rent-sharing are smaller than in Germany. This is also visible in a comparatively smaller change in rents in China when strategic power is shifted between the sides: from spB-treatments to spS-treatments, buyers' share of rents drops from 65.22% to 37.75% (or 14.42 in absolute rents), whilst it drops from 65.63% to 25.71% (or 22.80 in absolute rents) in Germany. The rent-sharing effects of changing competitive power, by contrast, are about the same in both countries: when buyers lose competitive power, their share of rents decreases from 57.79% to 45.40% (or 6.08 in absolute rents) in China and from 55.42% to 41.25% (or 5.06 in absolute rents) in Germany.

As in Germany, we find both powers to affect rent-sharing also when they interact: relative inequality is significantly lower in mixed-power treatments than in concentrated-power

[87] Mann-Whitney U test, two-sided: p=0.065
[88] Mann-Whitney U test, two-sided: p=0.093
[89] Mann-Whitney U test, two-sided: p=0.093 for both spS-cpB and spS-cpS treatments

treatments.[90] Contrary to Germany, however, we do not find the same degree of dominance of strategic power in our Chinese data. In the mixed-power treatments, the side endowed with strategic power does not necessarily earn significantly more than the side with competitive power. Whereas buyers earn more than sellers also in the Chinese spB-cpS treatment,[91] sellers do not earn significantly more than buyers in the Chinese spS-cpB treatment as relative rents of both sides are very close to 50%.[92] The different impact of strategic power may be due to differences in the level of relational contracting: in the Chinese spS-cpB treatment, relational contracting is used to a high degree (51.2%), whereas in both German mixed treatments (30.2% in spS-cpB and in 8.9% spB-cpS) and in the Chinese spB-cpS treatment (23.1%), relations make up a relatively small part of the trades.

> ***Result C-RC1: Strategic power has a larger impact on rent-sharing also when relational contracting is used more. However, the dominance of strategic power found in the German sessions disappears when sellers have strategic power, and a high share of trades take place within relations.***

In all four treatments, Chinese subjects tend to rely more on relational contracting than German subjects (compare data of tables 3 and 4). The tendency seems to be much stronger in the two mixed-power treatments than in the two concentrated-power treatments: the level of relational contracting is significantly higher in the two Chinese mixed-power treatments than in the respective German mixed-power treatments,[93] whilst the cross-cultural differences do not reach statistical significance in the two concentrated-power treatments.[94] Apparently, compared to Germany, power being shared stimulates the formation of relational contracting in China (for a further discussion see section 1.9.2).

> ***Result C-RC2: Chinese subjects rely more on relational contracting than German subjects when strategic power is given to one market side and competitive power is given to the other market side.***

[90] Mann-Whitney U test, two-sided: p=0.003
[91] Wilcoxon signed ranks test, two-sided: p=0.031
[92] Wilcoxon signed ranks test, two-sided: p=0.156
[93] Mann-Whitney U test, two-sided: p=0.009 (spS-cpB); p=0.041 (spB-cpS)
[94] Mann-Whitney U test, two-sided: p=0.195 (spS-cpS), p=0.461 (spB-cpB)

1.9 The nature of relational contracting

1.9.1 The cross-cultural characteristics of relational contracting

As we collected data in two subject pools with two distinct cultures, we can now give a first account of the cross-cultural nature of relational contracting. As in Germany (compare section 1.6.1), we find the rent-sharing effects of strategic and competitive power to be driven more by trades outside than within relations; relative inequality is again significantly lower within than outside relations.[95] We also confirm for the Chinese sessions that relations are characterised by a higher degree of gift exchange than trades outside relations; average prices and quality-levels are both higher within than outside relations.[96] Taken together, in both countries relations are characterised by high efficiency and low inequality.

Since relative inequality is lower within relations, the gains from relational contracting are unevenly distributed in favour of the weak side outside relations. Rent-sharing inside relations is compared to rent-sharing outside relations therefore shifted towards equality. However, treatment effects on rent-sharing are still visible within relations, more so in the Chinese sessions than in the German sessions. In the Chinese sessions, rent-sharing is always affected in the predicted direction,[97] whilst in the German sessions only the effects of strategic power are still statistically significant within relations.[98] In addition, in the concentrated-power treatments spS-cpS and spB-cpB, both in China and Germany, the favoured side earns more than the unfavoured side even within relations.[99] Consequently, equity matters too. In both countries, relations are therefore mostly built on a notion of fairness that entails elements of both equity and equality. This means that relationships are, in both countries, an imperfect buffer to the rent-sharing and efficiency consequences of competitive and strategic power.

[95] Wilcoxon signed ranks test, two-sided: $p<0.001$ for all 24 Chinese groups; this difference is even more visible in the Chinese sessions as we find significantly lower relative inequality within relations even for each separate treatment with only six observations: p=0.031 (spS-cpS, spB-cpB and spB-cpS); p=0.094 (spS-cpB)

[96] Wilcoxon signed ranks test, two-sided: p=0.031 in each treatment

[97] Mann-Whitney U tests for buyer share of rents in the Chinese sessions, two-sided: p=0.015 (spB-cpB > spS-cpB); p=0.002 (spB-cpS > spS-cpS); p=0.026 (spS-cpB > spS-cpS); p=0.009 (spB-cpB > spB-cpS)

[98] A cautionary note: in the German spB-cpS treatment, in only 3 groups private relations were formed. Hence, test statistics for this treatment suffer from limited data; Mann-Whitney U test for buyer share of rents in the German sessions, two-sided: p=0.015 (spB-cpB > spS-cpB); p=0.024 (spB-cpS > spS-cpS); p=0.485 (spS-cpB vs. spS-cpS); p=1.000 (spB-cpB vs. spB-cpS)

[99] Wilcoxon signed ranks test, two sided: p=0.031 (for both German and Chinese spB-cpB and spS-cpS treatments); in the other 4 treatments, no significant differences between seller and buyer rents can be detected within relations

Result R1: Relations are an imperfect buffer to the forces stemming from the market environment and the contract structure.

Given the similar effects of relations in both countries, how are relations in turn affected by the cultural background of the subjects? In order to answer these questions, we compare average effort, average price, and relative inequality between buyers and sellers within relations in every treatment between China and Germany. Comparing Figure 1-11 and Figure 1-7, we see that efficiency and rent-sharing within relations are very similar in the two countries. In fact, only two instances of (weakly) significant differences can be found when comparing price- and quality-levels as well as relative inequality between China and Germany. In spS-cpB quality levels are weakly significantly higher in China whereas relative inequality is significantly higher in the German spB-cpS treatment than in the corresponding Chinese treatment. In the other 10 cases, differences in the characteristics of relational contracting between the countries are statistically indistinguishable. This result implies that relations not only seem to be a buffer against forces stemming from strategic and competitive power; they are also largely unaffected by the cultural background of the subjects.

Interestingly, also trades outside relations are similar across the subject-pools. In the spB-treatments, no differences emerge in trades outside relations between China and Germany, neither in prices, quality nor relative inequality. On a treatment-basis, cross-cultural differences only reach a weakly significant level in the spS-cpS, in which relative inequality is weakly significantly higher in the German session.[100] However, in the merged Chinese spS-treatments, also prices and quality-levels outside relations tend to be higher than in the merged German spS-treatments; as already shown in section 1.8.3, the differences in quality become significant once we compare both spS-treatments between China and Germany. The same applies to relative inequality, which is also significantly higher in the Chinese spS-treatments than in the German spS-treatments.

Result R2: Relations are very similar across cultures, as are trades outside relations in the spB-treatments; however, some cross-cultural differences exist between trades outside relations in the spS-treatments.

[100] Mann-Whitney U test, two-sided: p=0.093

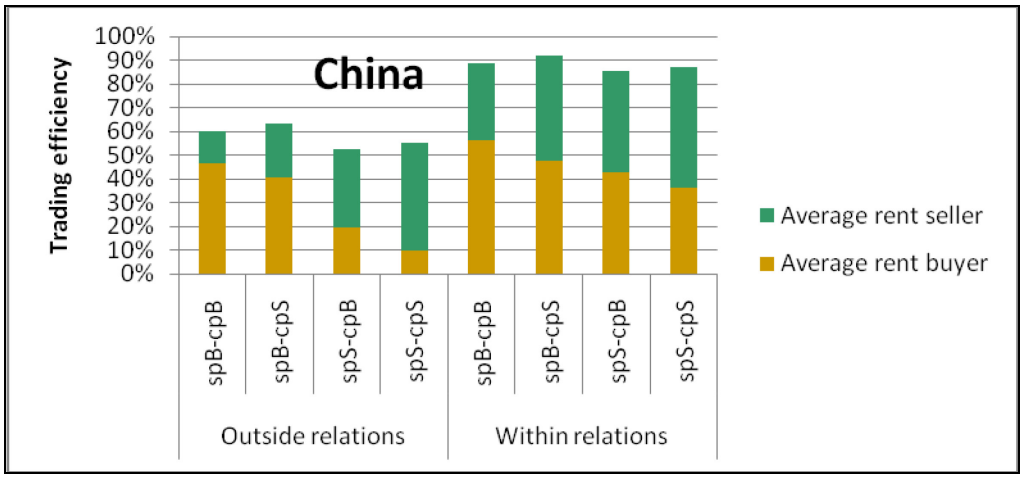


Figure 1-11: Trading efficiency and rent-sharing within and outside relations in Chinese sessions

1.9.2 The determinants of relational contracting

1.9.2.1 Relational contracting – causality or correlation?

As relations arise endogenously in the experiment, a few questions are warranted: first, how can we be assured that differences in relational contracting between the German and the Chinese sessions do indeed provide some explanation for differences in the effects of strategic and competitive power? Secondly, how do we get insight into what drives differences in relational contracting between China and Germany? In this section, we will attempt to find a common answer to both questions that makes us confident that relational contracting in our experiment is not a behavioural black box and indeed provides some insight into cross-cultural differences in the effects of strategic and competitive power.

Imagine our Chinese subjects were more cooperative than our German subjects for reasons that are unrelated to the dynamic incentives within relations, for example rooted in cultural differences. More cooperative behaviour would, within our experiment, result in higher quality and price choices as well as more pronounced rent-sharing. As we have argued, relations are, in turn, based on a high degree of gift exchange and pronounced rent-sharing. Therefore, more relations in China could be a consequence of more cooperative play (or of a greater number of cooperative players who self-select into relations) and not a reason for it. However, for such an explanation to be plausible, we would need an explanation why our Chinese subjects should be more cooperative and, more importantly, we would expect behavioural differences between Chinese and Germans not to systematically vary with treatment-conditions. If the degree of relational contracting and the degree of gift-exchange

were both merely a function of the number of cooperative players in the subject pool, we would also not expect systematic differences in behaviour between the different phases of a relation. Cooperative players should display the same level of cooperation regardless of whether they are in the initial round of a relation or in later phases. In the following paragraphs, we will test the plausibility of a higher level of relational contracting in the Chinese sessions being merely a consequence of more cooperative behaviour by the Chinese subjects.

1.9.2.2 When power meets relations

In fact, as we have already shown, the level of relational contracting reacts to treatment conditions in both countries. We find in both countries relational contracting to be used to a higher degree when sellers have strategic power compared to when buyers have strategic power (see data in Table 1-3 and Table 1-4).[101] This gives a robust support to a similar finding of WR (2007a), who also find less relational contracting in their corresponding spB-cpB treatment compared to their corresponding spS-cpB treatment. In contrast to WR (2007a), we measure this effect by defining relations as happening through private renewal trades rather than looking at the possibly noisy proxy private trades; we also find this effect to be robust to the allocation of competitive power and to distinctly varying the cultural background of the subjects. Competitive power only has a very limited and cross-culturally non-robust effect on the level of relational contracting: albeit starting at a very low level, the level of relational contracting is weakly significantly higher in the German spB-cpB than in the German spB-cpS treatment (see Table 1-3); [102] in the Chinese sessions, competitive power does not affect the level of relational contracting (see Table 1-4).

> ***Result R3: In both countries, the level of relational contracting is higher when sellers have strategic power; by contrast, competitive power only has a limited effect on the level of relational contracting: it is only found in the German sessions.***

When buyers have strategic power, they hold the key both to initiate relations, by making *private renewal offers*, and to maintain a relation, by behaving reciprocally. This situation is

[101] Mann-Whitney U test, two-sided: p=0.001 (spS vs. SpB) for both countries. For comparison between individual treatments: p=0.041 (spS-cpB vs. spB-cpB in Chinese sessions); p=0.026 (spS-cpS vs. spB-cpS in Chinese sessions); in Germany, we find a significant effect of changing strategic power when sellers have competitive power (p=0.011, spS-cpS vs. spB-cpS) but, despite a pronounced treatment effect, only marginally significantly more relations if strategic power is switched from buyers to sellers when buyers have competition power (p=0.119, spS-cpB vs. spB-cpB).

[102] Mann-Whitney U test, two-sided: p=0.063

apparently less conducive for relational contracting to emerge compared to sellers having to maintain a relation.[103] It seems that buyers' incentives to form relations drive the level of relational contracting in both countries.

When sellers have strategic power, buyers are dependent on reciprocating sellers in order to earn a positive rent. By relying on relational contracting, buyers may be able to improve their position in two ways: first, they can potentially select more reciprocal sellers by re-offering them a contract. Second, they can set incentives to money-maximising sellers to provide above-minimum quality by re-offering performing sellers a rent-generating contract. By contrast, when buyers have strategic power, they do not have to fear opportunistic behaviour of sellers as they can always rent-sharing in a way as to earn a positive rent; in fact, sellers have to fear opportunistic behaviour by buyers. As we have shown, buyers as a group set strong incentives for sellers to deliver high quality even outside relations – what sellers duly do. The opposite holds for sellers: sellers would minimise their risks and increase their profit prospects by entering relations when buyers have strategic power; relations are of little use when they themselves have strategic power, as buyers also pay high prices even outside relations. Therefore, the strategic value of relational contracting varies with strategic power. This is also shown by the data: when sellers have strategic power, buyers can considerably improve their profit-prospects by engaging in relational contracting; the average surplus in buyer-rents from trades within relations to trades outside relations is 19.44 (spS-cpB) or 31.80 (spS-cpS) in Germany[104] and 17.55 (spS-cpB) or 19.61 (spS-cpS) in China[105]. By contrast, when they themselves have strategic power, buyers can only expect to make an additional rent of 2.74 (spB-cpB) or 9.70 (spB-cpS) in the German sessions[106] and 7.39 (spB-cpB) or 5.39 (spB-cpS) in the Chinese sessions[107] by entering relations (see also Figure 1-7 and Figure 1-11). Sellers, in turn, profit somewhat more from relations when buyers have strategic power compared to they themselves having strategic power; this

[103] WR (2007a) explain less relational contracting in their corresponding spB-cpB treatment than in their corresponding spS-cpB treatment by arguing that buyers care less about the identity of the sellers when they have the last move. WR (2007a) generalise this finding by linking the prevalence of relational contracting to the degree of discretionary latitude of the contract structure. We argue that this claim is premature as they did not vary the side that initiates relations. Based on our results, we rather think that the incentives to form relations, which vary with strategic power, from the perspective of the initiators of relations, which happen to be the buyers, are crucial for the emergence of relations.

[104] Wilcoxon signed ranks test, two-sided: p=0.063 (spS-cpB); p=0.031 (spS-cpS)

[105] Wilcoxon signed ranks test, two-sided: p=0.031 (spS-cpB); p=0.031 (spS-cpS)

[106] Two-side Wilcoxon signed ranks tests do not detect a significance for either spB-cpB (p=0.438) or spB-cpS (p=0.250). For spB-cpS, it may due to too low observations since only three groups applied relational contracting after all.

[107] Wilcoxon signed ranks test, two-sided: p=0.063 (spB-cpB); p=0.031 (spB-cpS)

difference, however, is much less pronounced than for buyers (see also Figure 1-7 and Figure 1-11). Sellers gain, on average, an additional rent of 11.49 (spB-spB) or 10.11 (spB-spS) in Germany[108] and 13.46 (spB-cpB) or 15.49 (spB-cpS) in China[109] from trades within relations when buyers have strategic power, while they gain, on average, 8.65 (spS-cpB) or 4.93 (spS-cpS) in Germany[110] and 6.81 (spS-cpB) or 3.87 (spS-cpS) in China[111].

The strategic value of relations also varies with competitive power: as the data has just shown, the profitability of concluding trades within relations only changes marginally with competitive power; yet, individual members of the side without competitive power can considerably increase their earnings by engaging in relational contracting. Thereby, they can conclude more, profitable trades than what they can expect by competing on the spot market to find a transaction partner (recall that in any round two members of the side without competitive power have to content themselves with the outside option). Hence, *being* in relationships is also more attractive individually when the other side is endowed with competitive power. However, this may not necessarily hold for trying to *form* relations. Buyers, for instance, may be better off by concluding trades with changing partners rather than by trying, possibly in vain, to re-trade with the seller of the previous round; in line with this, buyers make fewer *private renewal offers* when sellers have competitive power compared to when they have competitive power themselves.[112] Consequently, neither buyer nor seller incentives that stem from the allocation of competitive power seem to explain changes in the level of relational contracting.

To conclude, the buyer's incentives to form relations, rather than to be in relations, seems to drive the level of relational contracting. Buyer-incentives may play a stronger role than incentives of sellers as buyers are the initiators of relations. While seller-incentives may still matter in explaining the absolute level of relational contracting, as relations can only be formed if sellers accept private renewal offers, the incentives of sellers do not provide much ground for explaining why the level of relational contracting changes with the treatment conditions.

[108] Wilcoxon signed ranks test, two-sided: p=0.094 (spB-cpB); p=0.250 (spB-cpS). The insignificance in spB-cpS may be due to too low observations.
[109] Wilcoxon signed ranks test, two-sided: p=0.031 for both spB-cpB and spB-cpS.
[110] Wilcoxon signed ranks test, two-sided: p=0.031 (spS-cpB); p=0.094 (spS-cpS)
[111] Wilcoxon signed ranks test, two-sided: p=0.0031 for both spS-cpB and spS-cpS
[112] Mann-Whitney U test, two-sided: p=0.020 (Chinese data); p=0.014 (German data)

Result R4: In both countries, the strategic value of forming relations, as seen from the buyers' perspective, seems to drive the level of relational contracting.

1.9.2.3 When culture meets relations

The strong reaction of the level of relational contracting to buyers' incentives is not the only piece of evidence that speaks against a higher level of relational contracting being merely driven by cross-cultural differences in cooperativeness; behavioural differences between Chinese and Germans also vary over the four treatments studied. Rent-sharing is not more pronounced in our Chinese sessions when buyers have strategic power, but when sellers have strategic power. This coincides with a high level of relational contracting in both spS-treatments in China, whereas in all other treatments relations make up a relatively small part of the trades. Although there is a tendency towards more relational contracting in all Chinese treatments compared to the German counterparts, differences in relational contracting between China and Germany are considerably unequal between the treatments: we find significantly more relations in China in mixed-power treatments but not in concentrated-power treatments (see result C-RC2); among the concentrated-power treatments, there are a still somewhat more relations when sellers have strategic power (in spS-cpS) but no difference at all when buyers have strategic power (in spB-cpB). As we randomly allocated subjects to the different treatments, subject-differences alone, that is independent from treatment conditions, cannot plausibly explain why we find, compared to the corresponding German treatments, more relational contracting in the mixed-power treatments but not in the concentrated-power treatments; they can also not explain why relative inequality is lower in China than in Germany when sellers have strategic power but not when buyers have strategic power.

We also observe systematic behavioural differences between different phases of a relation: when sellers have strategic power, they provide significantly higher quality in the middle rounds of a relation than in the initial round and again provide less quality in the last round of a relation;[113] when buyers have strategic power, the same pattern holds for prices – albeit

[113] Wilcoxon singed ranks test for middle rounds vs. first round quality-level, two-sided: p=0.004 (German data); p<0.001 (Chinese data). Wilcoxon signed ranks test for last round vs. middle rounds quality-level, two-sided: p=0.002 (German data); p<0.001 (Chinese data)

more clearly in the Chinese sessions, whereas in the German sessions only the drop in prices from the middle rounds to the last round of a relation can be observed.[114]

The evidence, hence, strongly speaks against some unobserved differences in cooperativeness that may explain both different levels of relational contracting and gift-exchange between the Chinese and the German sessions. Therefore, we rather need an explanation for the variations in the level of relational contracting that is related to the interaction between the treatment conditions and the distinct backgrounds of the subjects in our German and Chinese sessions.

Based on our cross-cultural motivation (see section 1.7), we expected Chinese subjects do have more experience with relations as an instrument to enforce incomplete contracts; thereby, they may also have more experience in which contexts relational contracting is most beneficial for them. This may give us a clue in explaining the cross-cultural differences in the level of relational contracting. Taking result R3 as a starting point, we compare the strategic value of relations from the perspective of buyers in all four treatments in both countries with the level of relational contracting. In spB-cpB, neither German nor Chinese buyers have a strong reason to form relations as profitability of relations is low in both countries (difference in rents between trades within and outside relations is 2.74 in Germany and 7.39 in China); we consequently find the same low level of relational contracting in both countries. In spS-cpS, relations are profitable in both countries but yet significantly more so in Germany (rent-differences are 31.80 in Germany versus 19.61 in China); nevertheless, we find still somewhat more relations in China than in Germany. In spS-cpB, incentives to form relations seem to equally strong in both countries (rent-differences are 19.44 in Germany and 17.55 in China), but we find Chinese to rely significantly more on relations. In spB-cpS, incentives to be in relations are high in both countries (while rent-differences are only 9.70 in Germany and 5.39 in China, recall that two buyers will every round be excluded from concluding trades);[115] yet, we find significantly more relations in China than in Germany. At the same time, forming relations may more difficult in spB-cpS than in spB-cpB; as argued above, buyers may rather want to conclude any kind of trade than leaving the trading phase

[114] Wilcoxon singed ranks test for middle rounds vs. first round price-level, two-sided: p=0.219 (German data); p=0.004 (Chinese data). Wilcoxon singed ranks test for last round vs. middle rounds price-level, two-sided: p=0.008 (German data); p=0.001 (Chinese data)

[115] Besides, also incentives for sellers to enter relations are relatively strong in both countries: differences in seller-rents between trades inside and outside relations are 10.12 in Germany and 15.49 in China.

empty-handed because they have unsuccessfully tried to establish a relation. This may explain why there are not more relations in the spB-cpS treatments than in the spB-cpB treatments; in fact, there are significantly less relations in the German spB-cpS than in the German spB-cpB treatment, while there are about the same level of relations in the two corresponding Chinese treatments.

Of course, our data are too limited to give a full account on the cross-cultural determinants of relational contracting; nevertheless, comparing cross-cultural differences in the level of relational contracting with differences in profitability of relations suggests that Chinese buyers, possibly based on their experiences outside the laboratory, react more strongly to the strategic value of relations than German buyers.

> ***Result R5: Cross-cultural differences in the level of relational contracting may be explained by Chinese buyers reacting more strongly to the strategic value of relations than their German counterparts.***

Consequently, the higher is the strategic value of relational contracting for the side who initiates relations, the more relations we expect to be formed. If the strategic value is low, even a subject pool with a presumed good understanding of how to use relations to overcome cooperation problems employs it to a low degree. Consequently, the main cross-cultural differences between China and Germany emerge where relational contracting carries a high strategic value. In terms of average achievable gains from trade, the data seem to imply a trade-off between different sources of incentives for sellers. The same level of efficiency is achieved either when buyers have strategic power or when sellers have strategic power and subjects rely strongly on relational contracting, as in the respective Chinese sessions. When buyers have strategic power, sellers face strong static incentives to perform, whereas a high degree of relational contracting creates strong dynamic incentives for sellers to deliver high quality.

> ***Result R6: The dependence of relational contracting on its strategic value seems to imply a ceiling for achievable gains from trade: high efficiency may either be achieved by sellers having strong static incentives, when buyers have strategic power, or by strong dynamic incentives, arising from a high degree of relational contracting.***

1.10 Summary

In this paper, we analyse two sources of incentives to cooperate in repeated incomplete contracts environments: strategic and competitive power. While strategic power stems from the internal structure of contracts, competitive power results from the competitive market condition in which the contract is embedded. As dynamic incentives prove to be critical for the relative strength of competitive and strategic power when contracts are incomplete, we also investigate the role of relational contracting for the identified effects. The experiments were therefore run in two culturally distinct places that we expected to display different degrees of relational contracting: Erfurt in Germany and Chengdu in China.

We find competitive and strategic power to have a robust effect on rent-sharing across distinct cultures and different degrees of relational contracting. Competitive power influences rent-sharing even if contract enforcement is entirely absent. If the strategically favoured side is not endowed with competitive power, it displays a more reciprocal behaviour. This can be explained by the incentive the side without competitive power has to enter a rent-generating relationship by presenting themselves as a reciprocal transaction partner. Without competitive power, relationship offer the prospect of concluding more trades than on the public spot-market. On average, trades generate significant rents for both transaction partners. Interestingly, competitive power reinforces reciprocity even when the market is predominantly characterised by spot-market trading rather than the bilateral trading islands found by BFF (2004). Competitive power does not affect efficiency.

Strategic power has a larger impact on rent-sharing than competitive power. In Germany, strategic power even dominates competitive power in such a way that the strategically favoured side always gains a larger share of rents regardless of competitive power. In China, by contrast, strategic power only dominates competitive power when buyers have strategic power. Strategic power influences efficiency in case subjects rely weakly on relational contracting, as they do in our German sessions. In this case, efficiency is higher when buyers have strategic power as efficiency is driven by the structure of static incentives for sellers to provide quality. When buyers have strategic power, sellers maximise profits if they provide maximum quality because of buyers´ reciprocal behaviour. By contrast, when sellers have strategic power, providing above-minimum quality may only pay off dynamically, i.e. within

a rent-generating relation. As relational contracting is used to a low degree in our German sessions, efficiency suffers when sellers have strategic power.

This result is not replicated in China: efficiency is unaffected by strategic power in the Chinese sessions. In our Chinese sessions, subjects employ relational contracting more than in our German sessions. This difference is most pronounced when sellers have strategic power; in this case, efficiency is higher in China. When sellers have strategic power, efficiency is higher in China than in Germany because of the stronger reliance of subjects on relational contracting and because of somewhat higher quality levels outside relations. By contrast, efficiency levels are undistinguishable between the Chinese and the German sessions when buyers have strategic power; in this case, differences in relational contracting between our Chinese sessions and our German sessions are less pronounced. More importantly, the higher the quality sellers provide outside relations, the less buyers rely on relations when they have strategic power. As a result, when buyers have strategic power the stronger dynamic incentives in the Chinese sessions do not simply add up to the strong static incentives that we also find in the German sessions. As in Germany, competitive power has no bearing on efficiency in China.

Behaviour within relations is astonishingly similar between China and Germany. This also holds for trades outside relationships when buyers have strategic power. Relational contracting is based on a high degree of trust by the side without strategic power and a high degree of reciprocity by the side with strategic power. Consequently, rent-sharing is more pronounced and efficiency higher within than outside relations. Efficiency-levels within relations are almost undistinguishable across treatments and cultures. Nevertheless, power still matters for rent-sharing also within relations. We therefore conclude that the culturally robust rent-sharing norms inherent in relations are based on a notion of fairness that includes elements of both equality and equity. Relations are therefore a buffer against the forces of competitive and strategic power that works imperfectly in terms of rent-sharing but almost perfectly in terms of efficiency.

In both countries, relational contracting is used more when sellers have strategic power, while competitive power does not influence the level of relational contracting. It therefore seems that the strategic value of forming relations, as seen from the perspective of the

buyers, who are also the initiators of relations, drive the level of relational contracting. In order to explain cross-cultural differences in the level of relational contracting, it seems that the Chinese buyers react more strongly to the strategic value of relations. We conclude that the prevalence of relational contracting depends both on the strategic value of forming relations, which varies mainly with the contractual structure, as well as on the, possibly culturally influenced, strategic understanding of relational contracting as a contract enforcement device.

1.11 Discussion and outlook

In the following, we want to highlight three dimensions to think about the implications of the experimental results as a basis for future research. First, we will assume the perspective of economic actors rather than of economic transactions by hypothetically asking whether players would rather be equipped with competitive or strategic power if they have to choose. Second, we want to ask how robust the experimental results are to the parameterisation and the specific form of strategic and competitive power. Third, we want to indicate possible future research on the nature of relational contracting.

Firstly, we analyse incentives to cooperate when contracts are incomplete; consequently, this paper focuses, in terms of rent-sharing, on transactions rather than on economic actors. We may nevertheless gain further interesting insight into the effects of competitive power and strategic power from switching the perspective by focusing on total rents of actors. By implementing excess demand or excess supply, players without competitive power could only make 5/7 of the number of trades that players with competitive power could conclude. Interestingly, in both countries total rents of buyers and sellers are statistically indistinguishable in both mixed-power treatments, i.e. when competitive and strategic power are allocated to different sides.[116] The higher rents of the side with strategic power in each completed trade is on average eaten up by being restricted to less trades than the side with competitive power. If roles were fixed but players could choose to have either competitive or strategic power, in the German sessions both buyers and sellers would be better off if buyers had strategic and sellers competitive power.[117] In terms of expected

[116] Recall that within each trade the side with strategic power earns a higher rent in all mixed-power treatments apart from the Chinese treatment in which sellers have strategic and buyers competitive power. Therein in which the difference in earnings is not statistically significant.

[117] Mann-Whitney U test, two-sided: p=0.065 (total buyer rents); p=0.093 (total seller rents)

rents, switching powers when sellers have strategic power and buyers have competitive power would therefore be a Pareto-improvement, which is driven by the higher gains from trade when buyers have strategic power. In China, total rents for both buyers and sellers are statistically indistinguishable between the two mixed-power treatments. Hence, looking at total rents, strategic power is no longer an advantage when it has to be traded for competitive power. It is left for future research to find out at which point competitive power even turns into an advantage when excess supply or excess demand are even more pronounced.

Secondly, as we used a specific implementation of strategic and competitive power, we may ask how generic the results of rent-sharing *within* each trade are; more precisely: did we give strategic power a better chance than competitive power to affect rent-sharing given that it implied full, and thereby an extreme form of, discretionary power? It, certainly, seems worth investigating whether a greater asymmetry of the market sides or a less pronounced form of strategic power may lead to a stronger role of competitive power. On the other hand, one may argue that we already implemented an artificially strong form of competitive power in that either demand was strictly greater than supply or vice-versa. Excess demand and excess supply are extreme forms of market imperfections; in the field, we would only expect to observe them in the very short run. We therefore claim that we implemented extreme forms of both strategic and of competitive power. Future research may attempt to investigate forms of strategic and competitive power that are closer to what we may also find more frequently in the field.

Thirdly, still very little is known about the conditions under which relational contracting flourishes and about the role subjects' experience and culture play therein. Exogenous variations in relational contracting may shed further light on a causal relationship between the level of relational contracting and the relative effects of strategic and competitive power. Further cross-cultural evidence is also warranted in order to investigate whether subjects' understanding of the strategic value of relations indeed varies with culture or the institutional environment subjects are accustomed to outside the laboratory. We found relations to be an imperfect buffer against the forces stemming from competitive and strategic power; it is left for future research to find out whether relations, because of

inherent fairness-concerns, may also buffer against *changes* in market conditions and contract structure.

1.12 Appendix

1.12.1 Supporting figures

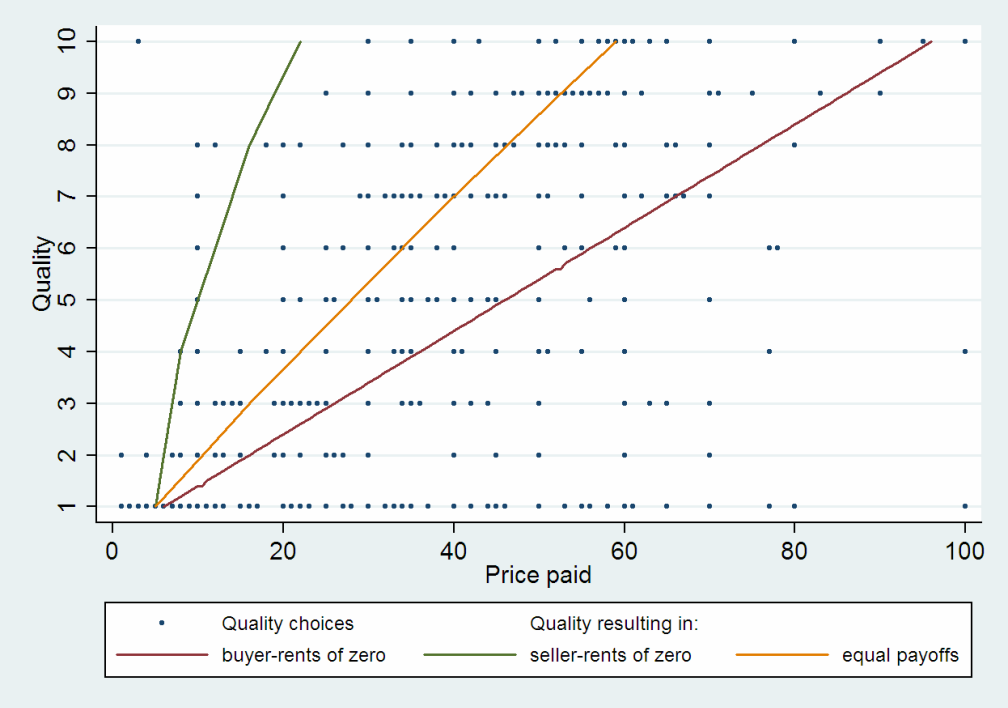


Figure 1-12: Quality-price schedule in German spS-treatments

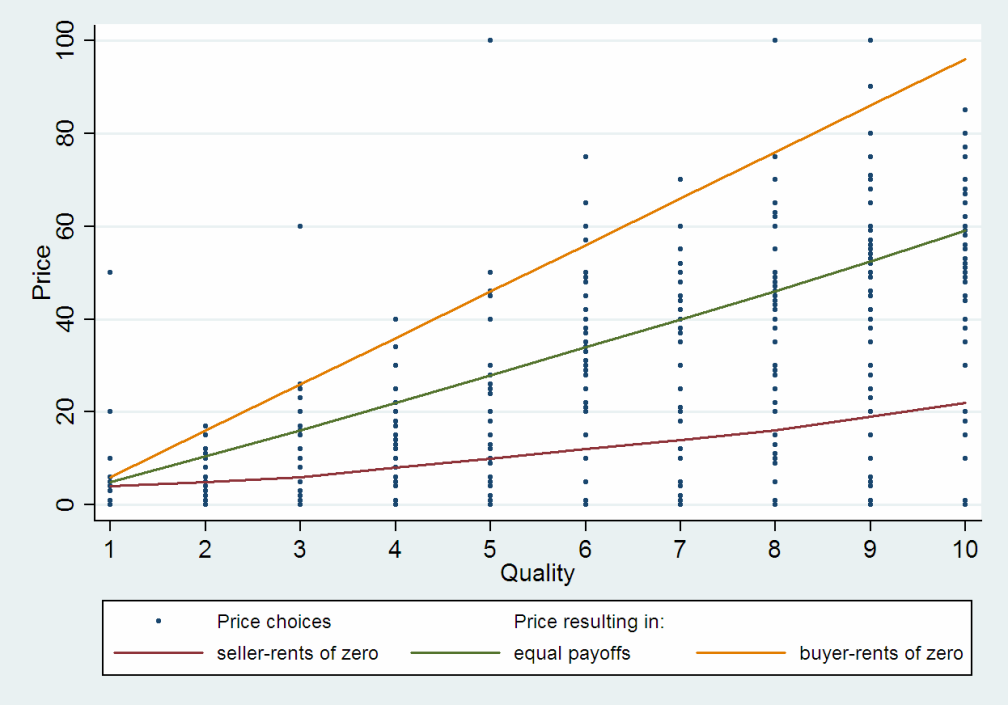


Figure 1-13: Price-quality schedule in German spB-treatments

1.12.2 Appendix: Experimental instructions

[The instruction provided here is the English translation of the original German and Chinese instructions of treatment spB-cpB. The instructions of treatments spB-cpS, spS-cpB and spS-cpS are similar, except for the following parts:

- *Treatment spB-cpS:*
  - *A group consists of 5 sellers and 7 buyers.*
  - *The trading phase will end when all sellers have made a deal or the trading time has elapsed.*
- *Treatment spS-cpB:*
  - *Each round is composed of two phases: the first phase is the trading phase; in the second phase, sellers decide on the quality of the goods.*
- *Treatment spS-cpS:*
  - *Each round is composed of two phases: the first phase is the trading phase; in the second phase, sellers decide on the quality of the goods*
  - *A group consists of 5 sellers and 7 buyers.*
  - *The trading phase will end when: all sellers have made a deal or the trading time has elapsed.*

The original Chinese and German instructions are available from the authors upon request.]

Experimental instructions

General instruction

In today's experiment, you will interact with your group members. A group consists of **7 sellers** and **5 buyers**. You earn points by selling goods as a seller or buying goods as a buyer. The value of the good to a buyer and the cost of the good to a seller depend on the good´s quality. At the beginning of the game, you will be randomly assigned a role either as a buyer or a seller. Your role will not be changed during the game. All participants will also be given a random ID number, which stays the same during the entire experiment. You will be notified about your number at the beginning of the experiment. There are a total number of 15 rounds in the experiment.

In each round, you can earn points. At the end of the experiment, we will exchange the total number of points you earned into RMB/Euro and pay you off. At the beginning of the experiment, **each participant receives 100 points as his endowment**. All information about payoffs in this instruction is expressed in points. The exchange rate is: **8 points = 1 RMB / 35 points = 1 Euro.**

Please note:

Please do not communicate with other participants in any form throughout the experiment. Please keep your mobile phones switched off. In case you have any questions please raise your hand from behind the curtain. All decisions are made anonymously. That is to say no participants know which decisions during the experiment have been taken by which other participants. Payments will be made anonymously and immediately after the experiment.

Procedure:

Each round is composed of the following phases:

1. Trading between sellers and buyers
2. Sellers decide quality of the goods
3. Buyers decide actually price of the goods

Display of payoffs of the current round

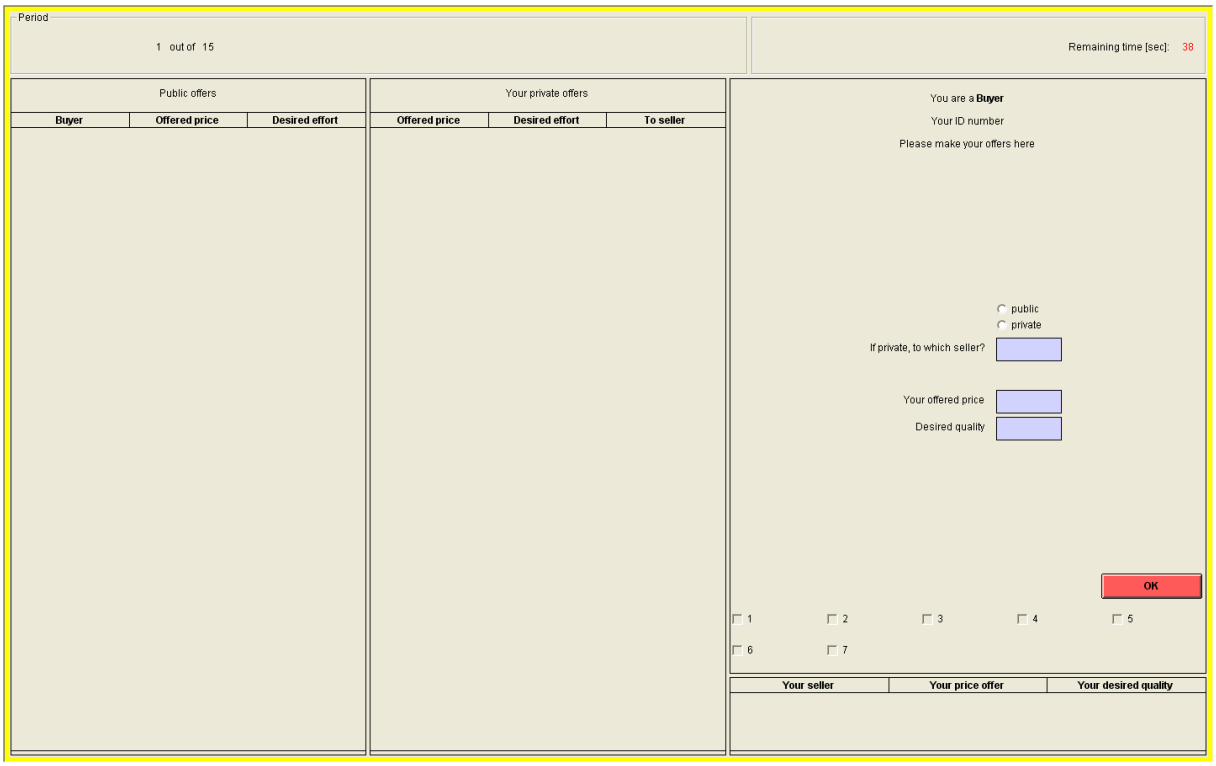
Phase 1: Trading

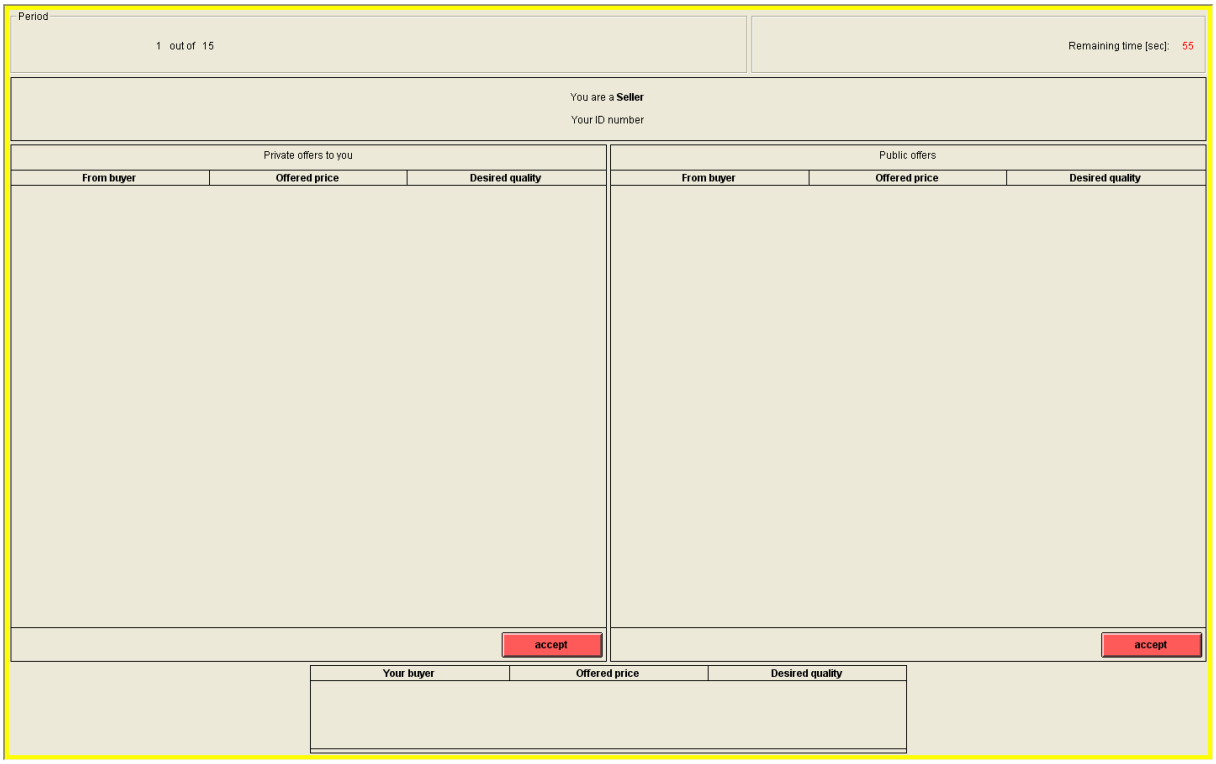
During the trading phase buyers can make offers to sellers.
When making an offer a buyer must deliver the following information:

4. **Price offer (between 0 and 100)**
5. **Desired quality of the good (between 1 and 10)**
6. **Receiver of the offer**

   Buyers can make either a private or a public offer. A private offer will only be sent to one particular seller and can only be viewed and accepted by this seller. In order to make a private offer a buyer must first select "private" and type in the ID number (1, 2, …, 7) of the desired seller in the region underneath. A public offer can be viewed by all sellers and can be accepted by any one seller. To make a public offer a buyer must first select "public".

A buyer can make as many offers as he wants until one of the offers is accepted by a seller. The check box in the lower-right corner of the screen shows to buyers which seller has already accepted an offer. In addition, all buyers and sellers see on screen which trade they made in the current round. **Each buyer and seller can make only one trade in each round.**

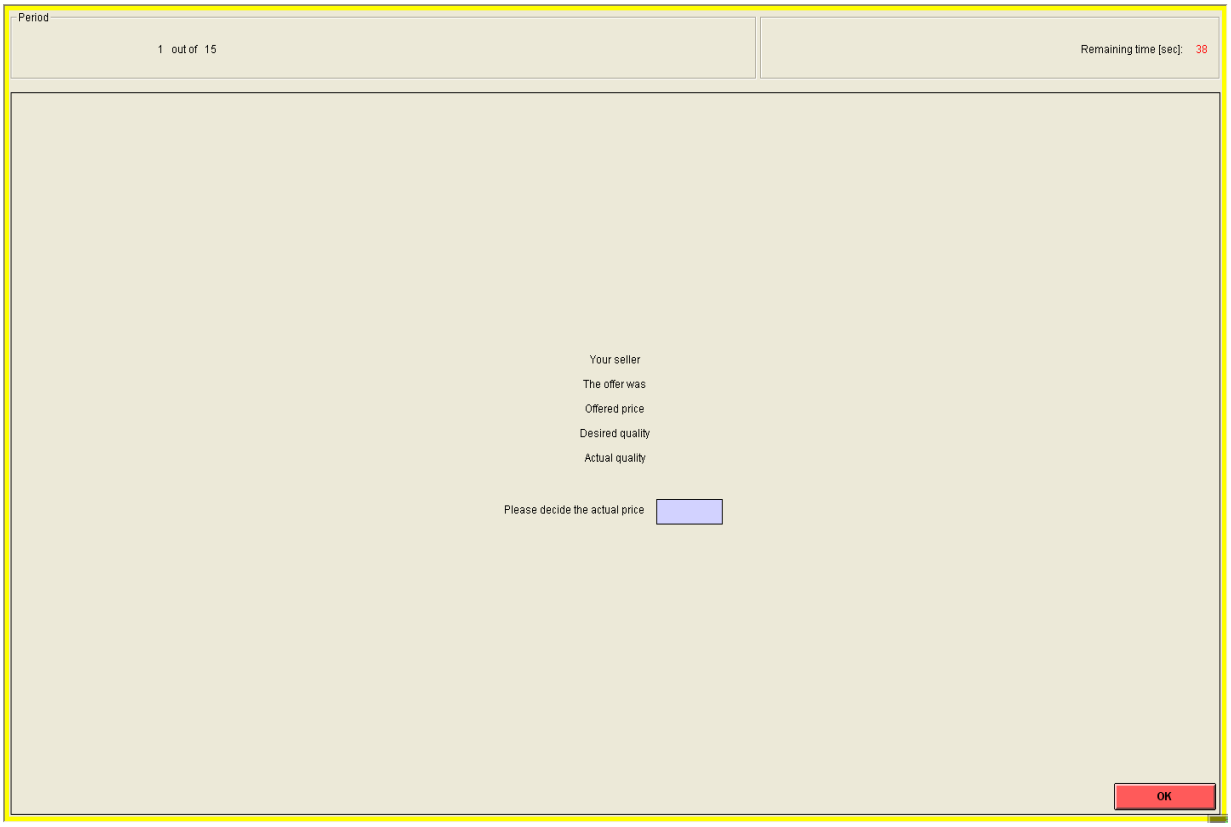




Sellers will see all public offers and those private offers that are sent to him. At any time during the trading phase sellers can accept offers. All received private offers are shown on the left part of screen while all public offers are shown on the right. By clicking on an offer and pressing “accept” a seller can accept one of the listed offers.
No buyer has to make an offer while no seller has to accept an offer. The trading phase will end when all **buyers** have made a deal or the trading time has elapsed. The trading time is 3 minutes (=180 seconds).

Second Phase: deciding quality of the goods

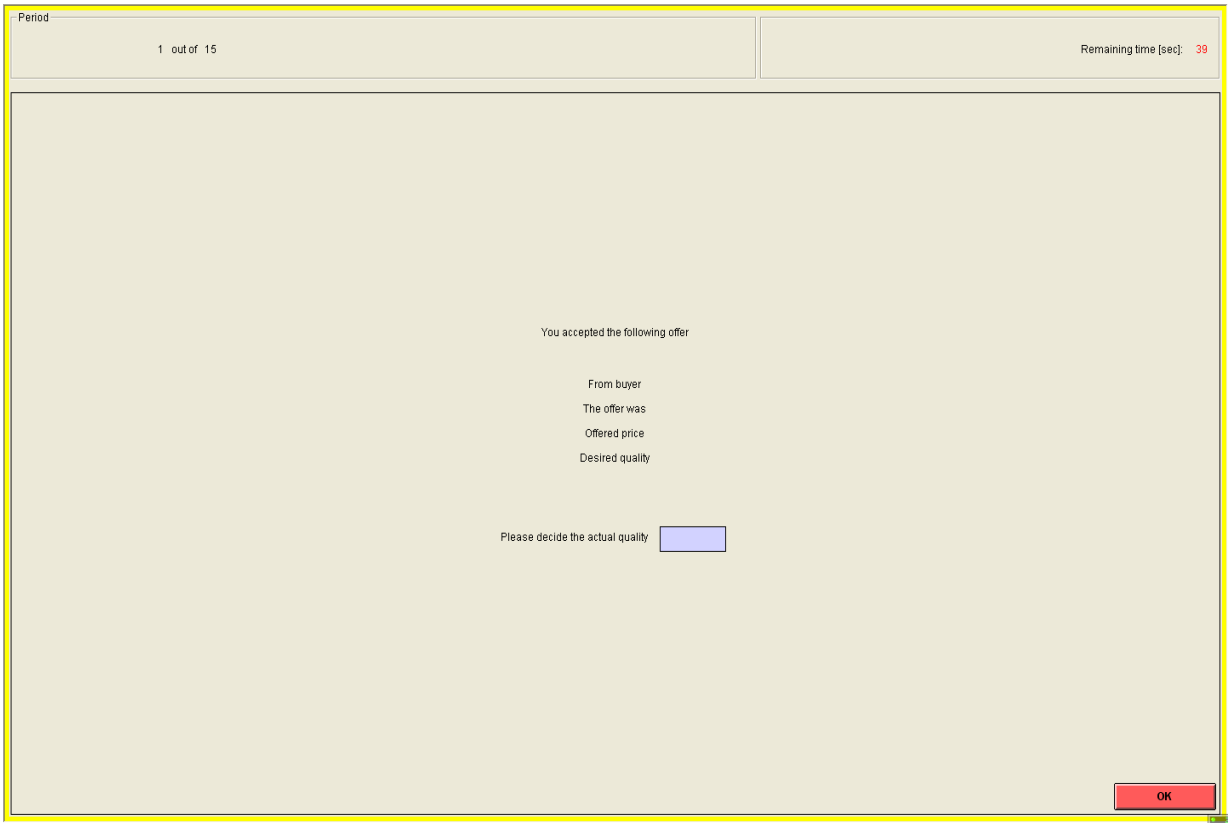
After the trading phase each seller who has accepted an offer will decide the quality of the goods. A seller can select an integer between 1 and 10 as quality of the good. **The seller can deliver a quality equal to, higher or lower than what the buyer desires.**



During the time sellers decide on the quality of the goods each buyer who concluded a trade will estimate what quality of the good he will receive. This estimation will not be seen by any other participants and will not have any influence on the points earned.

Third phase: Deciding actual price

After sellers provide the quality of the goods every buyer who made a deal will decide on the **actual price** of the good. **The buyer can pay a price higher, lower or equal to his offered offer.**



During the time buyers decide on actual prices every seller who concluded a trade estimates the actual price of the goods. This estimation will not be seen by any other participants and will not have any influence on points earned.

Display of payoffs of the current round:

At the end of each round you will be informed of your payoff of this round. If you concluded a trade you will be given the following information: type of offer, ID number of your trading partner, offered price, desired quality of the good, actual quality of the good, actual price of the good, payoff of trading partner and your own payoff of the current round.

Please fill this information into the attached documentation sheet. Any of the 15 rounds ends with payoffs of the current round being displayed.

<u>Calculation of payoff of each round</u>

1. No trade made: payoff of this round = 4
2. Trade was concluded:

Buyer's payoff	**=**	**Value of the good**	**–**	**Actual price**
Seller's payoff	**=**	**Actual price**	**–**	**Cost of good**

3. Cost and value of the good are dependent of its quality, according to the following schedule:

Quality	1	2	3	4	5	6	7	8	9	10
To buyer value of the good is	10	20	30	40	50	60	70	80	90	100
To seller cost of the good is	0	1	2	4	6	8	10	12	15	18

The calculation applies to all buyers and sellers. Hence, all buyers can calculate their sellers' payoff and all sellers can calculate their buyers' payoff.

Advice for making inputs: After each round or to confirm your input please click „OK", „continue" or „accept". Your input will be valid only after you press these buttons.

The experiment will start after all participants correctly answered the test questions. The test questions make sure that all participants have understood the experiment correctly.

Good luck!

2 The economic virtues of voting – how political competition limits confiscatory behaviour[1]

2.1 Introduction

In the enigma of which institutions foster growth, the view that secure property rights are beneficial to economic development is arguably the most widely shared view (see overview in North and Thomas 1973; North 1990; Przeworski and Limongi 1993; Glaeser et al. 2004). The intuition behind the importance of property rights[2] is compelling: the higher the risk to lose one´s returns on investment and effort, the lower the incentives to engage in productive activities. A powerful state may force private parties to respect private property and contracts. At the same time, a state with sufficient power to enforce contracts itself poses a threat to its citizens (Przeworski and Limongi 1993). Notable examples in modern times are the regimes of Suharto, Marcos and Mobuto Sese Seko of Indonesia, Philippines and Zaire (now Democratic Republic of Congo) respectively that are alleged to have embezzled a total sum of 25 to 50 billion USD (Hodess and Transparency 2004). Besides directly seizing assets through expropriation or rent-maximising taxation – on which we focus in this paper – government predation may also come as abrogating on its own contracts with private citizens, defaulting on its debt and spurring inflation by printing money (Olson 1993).

Potential investors may respond to confiscatory behaviour of the state by withholding investments – both into physical and human capital – or by doing business in the informal economy that is by definition not open to the state´s tax administration. Both may lead to foregoing possible economic gains: socially efficient investments may not be undertaken, the efficient size of companies may not be reached as companies need to be as small as to fall under the tax administration´s radar and the set of produced goods may be biased towards goods that are difficult to tax. A lower rate of investments may lead to a lower steady-state of economic development in a Solow-type growth model or even to a lower rate of growth in models of New Growth Theory that do not display diminishing returns to physical or human

[1] Based on: "The Economic Virtues of Voting – How Political Competition Limits Confiscatory Behaviour" by Arne Weiß (2009), Working Paper, University of Erfurt.

[2] We may distinguish de jure and de facto secure property rights. The former points to rule-of-law as a source of guaranteed property rights while the latter focuses on actual choices of possible predators. From the perspective of an investor, which we adopt in this paper, only the expected risk to lose his return matters for his investment decision.

capital (see e.g. Ray 1998). Businesses in the shadow economy are also excluded from using formal means of contract enforcement that the government may provide. This may lead to a further loss of efficiency as investments in transactions, outside trust-based long-lasting relations, will be lower than socially efficient due to the higher risk of moral hazard without formal contract enforcement. Without formal contract enforcement, Mafia structures as private forms of governance may be another, generally unwanted consequence of investors fleeing into the shadow economy. Predatory governments are therefore recognised as a major problem to economic development both by development practitioners (Hodess and Transparency 2004; UNoDC and World Bank UNODC and WorldBank 2007) and academics (North 1990; Olson 1993).

A widely discussed constraint to confiscatory behaviour by the government is democratisation (see Przeworski and Limongi 1993; Glaeser et al. 2004; Pinto and Timmons 2005). In its narrow sense, democracy refers to political competition (Bardhan and Yang 2004). Analogously to economic competition eliminating monopoly or oligopoly rents, political competition may dissipate rents of the political elite in the political market (Becker 1983; Wittman 1989). Democracy may also be a source of credibility for the state, which limits predatory behaviour (North and Weingast 1989). While economists are generally confident about the positive effects of economic competition, much less is known about the effects from political competition (Bardhan and Yang 2004). Furthermore, empirical investigations focusing on the economic effects of political competition are almost completely missing (see Besley, Persson and Sturm 2006).

Political competition may come in several ways. In this paper, we refer to the concept of political accountability.[3] It focuses on the process of political turnover. Political competition in this interpretation is more intense if the public can more easily remove the incumbent ruler from office. Political competition may link the ruler´s behaviour more closely to the voters' preferences if the incumbent rulers are sufficiently motivated by re-election. However, if the threat of dismissal by the public is as large as to render re-election unrealistic, incumbents may take a short-term horizon (Bardhan and Yang 2004). They may then exploit the power they have between elections (Persson, Roland and Tabellini 1997)

[3] A larger strand of literature investigates competition between political jurisdictions (see discussion in Bardhan and Yang 2004)

and expropriate all they can while still in office. In a similar spirit as Bardhan and Yang (2004), Grossman and Noh (1994) formally derive the conditions under which increased political competition leads to less rent-seeking taxation. They show theoretically that in case of sufficient political stability, a higher degree of political competition leads to less confiscatory taxation by the incumbent ruler.[4] In this case, the incentive of incumbents to stay in office (the probability of which they can increase by lowering tax rates) outweighs the shortened horizon (as a higher degree of political competition lowers the probability to stay in office for any given tax rate). As a limit, taxation entails no political rents and welfare is maximised.[5] If, however, political stability is low, the ruler´s time horizon is already short, e.g. because he fears to be overthrown by external forces; in this case, political competition can have perverse effects on taxation: increased political competition further reduces the ruler´s horizon so that only at higher tax rates will the ruler refrain from taking the one-time gain of setting the tax rate at unity.[6] In Grossman and Noh (1994), however, political competition is a black box. The mechanism through which voter´s preferences may enter the political process is left opaque.

The aim of this paper is to partially fill the empirical gap on the possible virtues of political competition. We take the model of Grossman and Noh (1994), henceforth GN (1994), as a starting point and present the first experimental results on the effects from political competition on rent-seeking of rulers. We enrich the model of GN (1994) by filling the black box of political competition with an obvious mechanism: competitive elections. In our experimental model, political competition between incumbents and contestants is either non-existent or endogenously determined through the behaviour of voters. To the best of our knowledge, we employ a hitherto unused method for this field of research: controlled experiments; it circumvents problems of measurement and reverse-causation. At the same time, we make a methodological suggestion to narrow-down research questions within the

[4] Grossman and Noh (1994) speak of a proprietary fiscal authority that is unrestricted in its taxation level. As Grossman and Noh (1994) do not model bureaucratic behaviour but assume the fiscal authority to maximise pecuniary political rents, we feel justified to equate the fiscal authority of Grossman and Noh (1994) with a ruler who is unrestricted in her taxation decision.

[5] Welfare is maximised with arbitrarily large political competition, i.e. when even marginal deviations of the incumbent ruler´s behaviour from the expected equilibrium behaviour of any ruler lead to political turnover, and infinite stability of the political system, i.e. the incumbents´ survival probability being only depended on her behaviour.

[6] Technically, if the survival probability is low, the time-consistency constraint starts binding. Incumbent rulers can only credibly self-commit to a tax rate that is as low as to turn the discounted future losses in tax revenues at least as large as the one-time gain from choosing a tax rate of unity. The discounted future losses increase in the survival probability. Once the time-consistency constraint is binding, increased political competition restricts the set of time-consistent tax rates and thus leads to higher tax rate.

democracy-growth nexus. Our research questions are: Facing political competition, do elected holders of power display less confiscatory behaviour? Do investors respond with higher investments?

Our results show that political competition limits confiscatory behaviour. Political competition motivates incumbents to be re-elected based on a good track-record. Investors react with higher investments in the presence of political competition. In contrast, in the absence of political competition we observe a large variance of investment levels as some non-elected incumbents build a reputation as being benevolent, whereas others even fail to incentivise investments within one legislative period. The effects of political competition do not depend on the time-horizon of rulers, which gives competition a robust role in limiting confiscatory behaviour and increasing efficiency.

The paper is organised as follows: Section 2.2 discusses the limitations of the current empirical literature on the democracy-growth debate and argues for a role of controlled experiments as a complementary tool. Section 2.3 introduces the experimental approach of this paper and puts it into the context of existing experimental research. Section 2.4 generates testable hypotheses, and section 2.5 introduces the experimental procedure. Results are presented in section 2.6. Section 2.7 looks beyond the experimental data and asks what we may, and may not, learn from the experimental results.

2.2 Limitations of existing empirical research

Empirical evidence on the effects of political competition on confiscatory behaviour is mixed. There is only few direct evidence on the economic effects of political competition (see Besley, Persson and Sturm 2006). Most research focuses on the effects of democracy, which is, especially outside economics, mostly understood as more than free and contested elections.[7] As to the effects of democracy, Doucouliagos and Ulubaşoğlu (2008) report in a recent meta-study that two thirds of the included studies find no or a significantly negative effect of democracy on economic growth, whereas one third find a significantly positive effect. This result mirrors earlier and similarly contradictory findings reviewed by Przeworski and Limongi (1993). The inconclusiveness of results already hints at the econometric problems of many cross-country studies that we are going to discuss. Possibly, democracy

[7] The reader is kindly referred to Doucouliagos and Ulubaşoğlu (2008), who discuss commonly employed definitions of democracy.

has both positive and negative effects on growth (for a discussion see Przeworski and Limongi 1993; Vega-Gordillo and Alvarez-Arce 2003). Therefore, separating these effects would be helpful in order to gain a better insight and to provide better policy advice. Many studies, however, implicitly follow a broad research question by analysing the effects of democracy in a wide sense on economic performance. Most critically, many studies analyse the joint effect of political competition and rule-of-law institutions; however, doing so, may not only be theoretically unjustified[8] but may also contribute to the inconclusiveness of the current empirical literature. The more institutional features are included in one democracy measure, the more difficult it is to draw conclusions. Imagine rule-of-law foster economic growth (as it guarantees property rights), but elections do not (as incumbent rulers take a short-term view because of the threat of dismissal from office): a common measure of both elections and rule-of-law will then likely lead to inconclusive results; in case there is a definite result, we can only derive a conclusion on the joint effect of political competition and rule-of-law on economic performance.[9]

The arguably most severe problem in the current empirical literature is the likely endogeneity of democratisation. Possibly, the citizens´ demand for participation in a democratic country also depends on the citizens´ human capital and wealth, both of which are related to economic performance. The possible causal link from democracy to growth has come to be known as the Lipset-hypothesis (Lipset 1960), which received empirical support by Barro (1999), Alvarez et al. (2000) and Glaeser et al. (2004). Ignoring a possible reverse causation from growth to democracy can lead to false conclusions. Glaeser et al. (2004) argue in a fundamental critique on the institutions-cause-growth-literature that using averages of institutional quality over a span of time can falsely attribute economic growth to institutional changes, whereas in reality institutions improved with growth. They also point to a measurement problem of institutions: many measures employed treat deliberate policy

[8] Przeworski and Limongi (1993) for example note that the common employment of democracy as a proxy for secure property rights is unjustified as theoretical analysis is ambiguous.

[9] Of course, a strong case can be made that a stable democratic system that effectively constraints political contestants in respecting the outcomes of fair elections requires some at least rudimentary rule-of-law (Olson 1993; Clague et al. 1996). This needs to include basic economic rights as even the incumbent ruler´s political opponent needs to be able to survive in order for elections to be competitive and fair (Olson 1993). Nevertheless, rule-of-law and political competition do not necessarily come together. While the World Bank governance indicators (Kaufmann, Kraay and Mastruzzi 2008) show a strong correlation between "Voice and Accountability" and "Rule of Law" (0.79, own calculations) there is nevertheless a large continuum of different combinations of both. There are also notable exceptions with a low score on "Rule of Law" but a comparatively high score of "Voice and Accountability" such as Jamaica, Italy, Argentina and Benin. Separating effects would also help to improve theoretic models and set adequate priorities in the international development agenda.

choices of unrestricted rulers, e.g. to hold elections or to respect property rights, as institutions in their *constraining* sense. This may falsely lead to attributing growth to institutions, whereas, in fact, growth may be a consequence of policy choices of growth-friendly dictators.

Methodologically, there is a need for empirical insight into the mechanisms through which democracy may influence economic performance controlling for the direction of causality. One approach to test for causality was introduced by Granger (1969); it relies on changes in one variable preceding changes in another variable.[10] However, we are not aware of any study on political competition employing the Granger methodology. Among the few studies that have tested for Granger causality in the nexus of democracy and economic performance are Farr, Lord and Wolfenbarger (1998), Dawson (2003) and Vega-Gordillo and Alvarez-Arce (2003). Farr et al. (1998) lend support to the Lipset-hypothesis in that they find the level of economic well-being, as measured by GDP, to cause political freedom. Economic freedom, on the other hand, Granger-causes the level of GDP. Somewhat contrary to this, Dawson (2003) finds levels of political freedom and civil liberties as well as economic freedom to cause economic growth while political freedom also causes economic freedom. Vega-Gordillo and Alvarez-Arce (2003) find economic freedom to cause growth, whereas the positive effects of political freedom are less clear. Hence, evidence on the interaction between political freedom and economic performance is mixed, whereas economic freedom causing the level of either growth or economic well-being receives support in all three studies. Besides the inconclusive results on the effects of political freedom, the studies cannot circumvent the problem of mistaking institutional choices for institutional constraints, as discussed by Glaeser et al. (2004).

2.3 An experimental approach

Another approach to investigate causal relationships and at the same time circumvent measurement problems is controlled experiments. Experiments allow for a large degree of control over the data generating process. Problems of measurement and reverse-causation can thereby be circumvented. The ideal data set to answer our research questions would entail exogenous variation of political competition, keeping everything else equal, and then

[10] Rather than causality in the traditional sense, Granger causality can be therefore be thought of as a test of "firstness" (Dawson 2003).

measure the impact on confiscatory behaviour and investments. Properly designed experiments in the laboratory allow the counter-factual analysis that, apart from some rare occurrences of natural experiments, may not be done in the field.[11] By design, we can exogenously implement whether subjects interact in a system with or without political competition and can then observe the role political competition plays for confiscatory behaviour and investment levels. If everything else is controlled for, either by design or by randomisation, causal conclusions can then be drawn. In our case: does political competition *cause* less confiscatory behaviour? There is, however, an obvious catch: while control is experimental economics´ greatest asset, it also is its largest shortcoming. Experimental research is generally confined to rather narrow research questions; and just as with any theoretical model, experimental results must be scrutinised for their external validity, i.e. what they say about the economic systems outside the laboratory. We will come to this in the concluding section 2.7.

2.3.1 Design

We choose a design that incorporates both essential features of the theoretical model of GN (1994) but is also known to be a reliable tool in experimental economics. In GN (1994), the optimal tax rate at which an unrestricted fiscal authority can tax a marketable good is derived. Producers decide how to allocate effort between the production of a non-marketable good and a higher-value marketable good. The non-marketable good is not-taxable and therefore generates a certain profit. The marketable good is prone to confiscation through unrestricted taxation. Efficiency and welfare is maximised if only marketable goods are produced. This, however, requires the producers to expect a tax rate that entails no political rents[12], to which the fiscal authority cannot pre-commit. Reputation may act as an – either perfect or imperfect – substitute for pre-commitment by the ruler.

The effort devoted to the production of the marketable good is, essentially, a risky investment that generates a social surplus. The risk stems from possibly morally hazardous

[11] Counter-factual analysis in the field would attempt to answer questions such as: would Zimbabwe have had less confiscatory taxation had there been fair and competitive elections? Would China have had more investments and faster growth had the communist party been subject to political competition through contested elections?

[12] GN (1994) also derive the optimal level of public good provision. In this case, the welfare maximising tax rate needs to be greater than zero but still entails no political rents as the tax revenue is entirely used to finance the public good. GN (1990) abstract from public good provision in a simplified version of the full model in GN (1990), the latter being the same model as in GN (1994). The qualitative insights of the effects of political competition rent-seeking taxation hold in both versions of the model. Since our research focuses on confiscatory taxation, we abstract from public good provision and thereby follow the simplified model of GN (1990).

behaviour by the fiscal authority. Producers can therefore be understood as investors and, if we abstract from bureaucracy, the fiscal authority as an incumbent ruler who is unrestricted in her taxation of post-investment assets.

We therefore argue that, by adapting an investment game (Berg, Dickhaut and McCabe 1995) to our needs, the basic elements of GN (1994) can be experimentally incorporated. In Berg, McCabe and Dickhaut (1995), a sender can send part of her endowment to an allocator. The amount sent is multiplied by a number greater 1 by the experimenter before reaching the allocator. Sending the entire endowment is therefore the social optimum. After having received the multiplied amount, the allocator decides how much of the amount received he wants to sends back to the investor.

By choosing an investment game as our experimental platform, we are employing a reliable tool that is widely used. We can thereby access a large stock of experimental data, which first shows the investment game to create a cooperation problem between senders and allocators. The cooperation is typically only partly overcome by subjects, which leaves room to analyse the effects of political competition. This applies also to our companion paper (Walkowitz and Weiß 2009, chapter 3 of this thesis), which employs the same basic design.[13] Second, the investment game is generally well understood by subjects. We therefore stayed within the framework of a standard investment game and opted, in particular, against incorporating a direct taxation decision instead of the more standard sending back decision. The relation of the amount sent back and the amount received, nevertheless, unambiguously implies a tax rate. While we could have chosen a design that is yet closer to the problem under study, we opted, for the reasons given, for one that gives us strong confidence in its internal validity.

The investment game incorporates the basic structure of the GN-model: investments generate a social surplus; at the same time, the investors´ post-investment assets are vulnerable to confiscation by the allocator. The latter can be seen as a ruler who is unrestricted in her taxation decision or, in the terminology of GN (1994), as a proprietary fiscal authority. By repeating the investment game for 25 rounds, we allow for reputation building by the allocators. Technically, within the assumptions of standard game theory, a

[13] All subsequent references to Walkowitz and Weiß (2009) also refer to chapter 3 of this thesis.

finite end forbids reputation building by the logic of backward induction.[14] A large amount of experimental evidence on the so-called end game effect, however, shows that cooperation typically only breaks down at the very last rounds of a repeated experiment (this was also found by our companion paper Walkowitz and Weiß 2009). While finite play may sound as a necessary bug – due to the practical restrictions of experimental economics – we actually consider it a feature. It firstly allows us to test whether infinite horizons are indeed necessary for cooperation to emerge and for political competition to have an effect on taxation. It secondly enables us to test GN´s (1994) prediction that the effects from political competition depend on the stability of the political system as the remaining length of the political system is reduced with each round.

In GN (1994), the incumbent ruler's probability to stay in office is linked to the stability of the political system and the differences in utility that producers expect from the incumbent and a fiscal authority displaying equilibrium behaviour. The mechanism through which producers´ assessment of the incumbent ruler affects her survival probability[15] is, however, left opaque. Since we are interested in behaviourally studying the role of political competition, we need to incorporate a transacting mechanism. The most common are elections (Pinto and Timmons 2005), as a core ingredient of democracy. With repeated elections, the voters endogenously determine the degree of political competition, i.e. how sensitive the incumbent ruler's survival probability reacts to her taxation decisions. By experimental design, we can contrast the voting system with a system in which the probability to stay in office is totally insensitive to the tax rate. The latter is analogous to setting the parameter of political competition to zero in the GN-model. We therefore run two treatments: the experimental treatment *PolComp,* which allows for political competition, and the control treatment *NoPolComp* without political competition.

We introduce elections in our design by using five investors, instead of one as in the original investment game, who are also voters in the treatment with political competition (PolComp). Elections take place every five rounds (view Figure 2-1). We thereby give incumbent rulers considerable power between elections, that is, within a legislative period. The survival probability is 1 for four consecutive rounds as the degree of political competition is zero

[14] Compare analysis in section 2.4.1.
[15] Survival probability is the technical term employed by GN (1994), which we use interchangeably with probability to stay in office or in power.

within one legislative period. We thereby deviate from GN (1994), in which the degree of political competition is an exogenous parameter that does not follow electoral patterns. Since we employ a specific mechanism, elections, in order to induce political competition, elections every round would be quite tedious and may seem unnatural for the participating subjects.[16] It also allows both investors and incumbent rulers some time to learn how to interact. Voters elect one allocator (R) out of three candidates (C) by majority rule. In case of a tie between two candidates, a run-off election takes place, again by majority rule. Prior to elections, candidates announce a cheap talk back-transfer strategy to the voters. This is analogous to announcing a tax schedule to which allocators cannot pre-commit. The winning back-transfer promises are displayed to the unelected candidates. Unelected candidates do not participate in the investment game that follow but have to wait until the next election takes place. For every round they wait, they earn the same amount as both investors and the allocator receive as their endowment. Candidates have fixed identification codes for the entire experiment in order to allow for effective reputation building. In the control treatment NoPolComp, the only design difference to PolComp is the random draw of the allocator with equal probability out of the three candidates.[17]

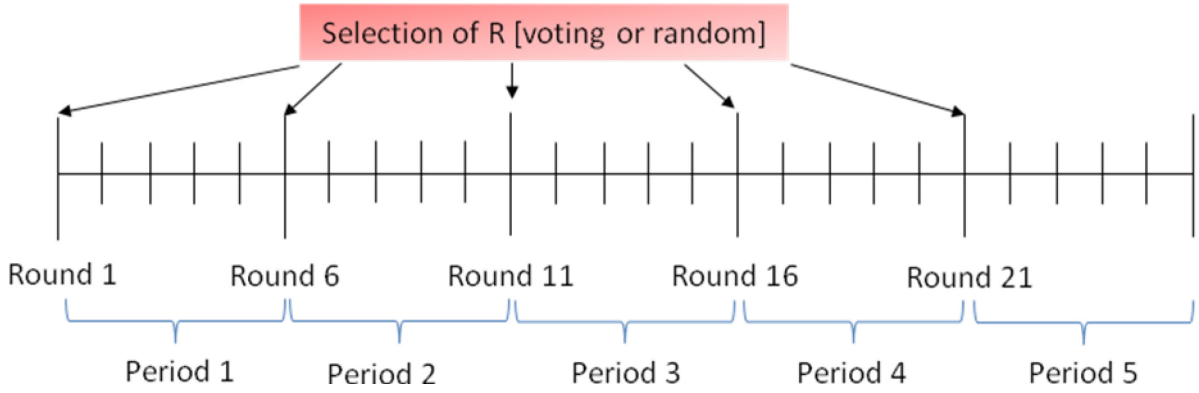


Figure 2-1: Experimental sequence

[16] While we do not pretend to bring real-world democracies into the lab, five rounds nevertheless fit nicely to the length of a legislative period in real-world democratic and socialist countries such as between two presidential elections in France and the length of the planning period in socialist countries or socialist countries in transition such as China. In China, the country´s leaders as well as the National People´s Congress are also elected for five years.

[17] Most notably, candidates in NoPolComp also make cheap-talk promises that are shown to the investors prior to the selection of the allocator. Instructions were worded identically apart from the necessary differences (random draw instead of voting). Instructions are available upon request.

2.3.2 Specification

Figure 2-2 graphically depicts the basic experimental implementation. Every round, 5 investors separately decide which integer between 0 and their endowment of 10 to send to the allocator (R). The allocator is either selected by the voting mechanism introduced above or by a random-draw. The investments are doubled by the experimenter before reaching the allocator. For each amount received, the allocator decides how much she will send back. She can send back any integer between zero and the amount received, i.e. the doubled investment. The allocator only receives information on the invested amounts and not on the corresponding investors. Allocators can hence not discriminate between individual investors in an attempt to satisfy the wishes of only a part of their electorate at the expense of the other.

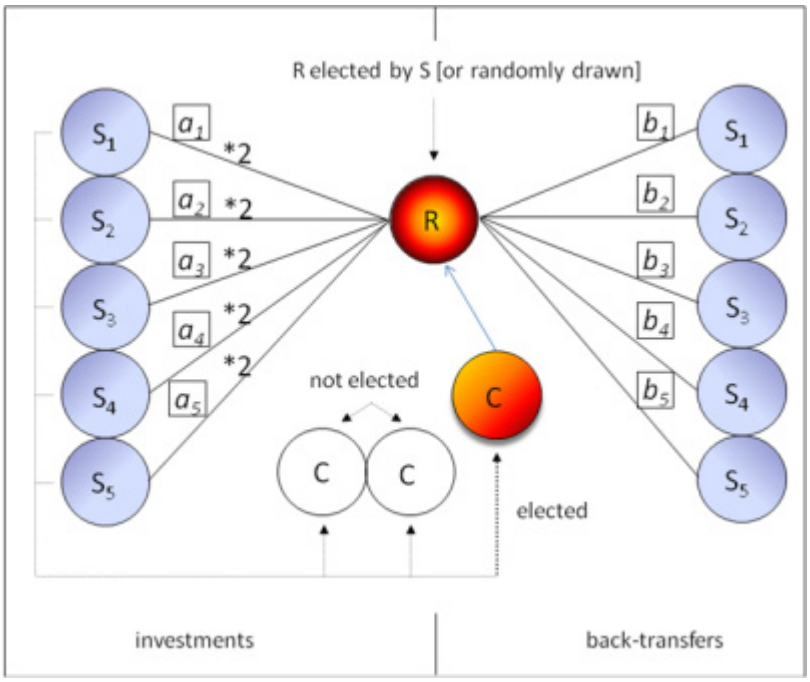


Figure 2-2: Experimental procedure

Payoffs for each round are calculated according to the following schedules:

$$\pi_{investor_i} = 10 - X_i + Y_i$$

$$\pi_{allocator} = 10 + 2 * \sum_{i=1}^{5} X_i - \sum_{i=1}^{5} Y_i$$

$$\pi_{unselected\ candidate} = 10$$

Roles are strictly separated; that is, candidates do not vote and do not make investments themselves. The reason for the strict separation is that we are not interested in the behaviour of candidates as voters and investors and that mixing roles would possibly add noise to the data. Separating candidates (and later allocators) from investors can be understood as few candidates facing a large electorate.[18]

The experiments are run in Chengdu, People's Republic of China. Chinese people have very little experience with voting mechanism as elections only take place in some rural areas. China also scores low on the "Voice and Accountability" indicator of the World Bank Governance indicators (Kaufmann, Kraay and Mastruzzi 2008).[19] We consider the lack of experience with voting an advantage in order to test our hypotheses. Chinese subjects should have less pre-understanding of how an elected holder of power ought or does behave than subjects in countries with more frequent competitive elections. In this way, experimental results may be less confounded by subjects' personal experiences with voting procedures and the behaviour of elected power holders.

We opted for a rather neutral framing that emphasise the hierarchical relation between the allocator and the investors. Investors are called "citizens" and the allocator "president". The decision by investors is framed as "investing". The allocator´s sending back decision is just framed as "sending back". The candidates are called "candidates".[20] The term "citizen" has a political connotation; we nevertheless refrained from explicitly putting subjects into a political context in order to minimise behavioural noise in the data as Chinese subjects may not feel at ease in a political situation in which they vote.

2.3.3 Existing experimental evidence

To the best of our knowledge, political competition has not been experimentally investigated so far. Closest to our design is our companion paper (Walkowitz and Weiß 2009). It analyses the effect of voting on trust and the exercise of power. They employ the same basic design as us, but identification codes are shuffled prior to any selection of allocators; effective

[18] In large electorates, the decisions of candidates *as citizens* are only of marginal importance compared to the aggregate of citizens. In small groups, which we use due to the practical restriction of experimental economics, the behaviour of candidates as citizens would, however, matter.

[19] In 2007, China found itself in the lowest "Voice and Accountability" bracket; its percentile rank was 5.8.

[20] The Chinese terms we employed were: president: 主席(zhǔxí), citizen: 公民(gōngmín), candidate: 候选人(hòuxuǎnrén)

reputation building is thereby ruled-out in order to test whether elected allocators display a reciprocal response to being elected. Their design is consequently not suitable to test the effects of political competition along the lines of Grossman and Noh (1994). Corazzini, Kube and Maréchal (2007) find elected allocators to be willing to send more to the recipients than randomly drawn allocators if their approval rates are higher than what is minimally required to win the election. They attribute this effect to guilt-aversion. In a related treatment, they find some evidence of reputational concerns as well as of motivational crowding-out. In order to improve their chance of one-time re-election, allocators are willing to increase their transfers for low approval rates, but would give somewhat less in the second period compared to the original treatment. Compared to Corazzini et al. (2007), we focus on actual behaviour rather than hypothetical decisions elicited by the strategy method and allow ample room for reputational concerns by playing the basic game 25 times. In contrast to our design, allocators in Corazzini et al. (2007) receive endowments unconditionally and, therefore, do not confiscate investors' assets.

There are several experiments that analyse the effects of competition on cooperation in sequential games. Huck, Lünser and Tyran (2006) show that even with minimal information economic competition is very effective in raising efficiency in a repeated binary trust game[21]. With competition, buyers choose in each period the seller they want to interact with. Bornhorst et al. (2004) show that competition to be selected as an allocator may lead to higher back-transfers and higher efficiency in an investment game. They show that investors send more in a choice treatment than in a random *assignment* treatment while allocators send back more if they are chosen as allocators rather than randomly matched with an investor. They, however, do not report on whether these differences are statistically significant.[22] In contrast to Huck, Lünser and Tyran (2006) and Bornhorst et al. (2004) we focus on the effects of a formal way of choice by a group in a political framework, i.e. democratic voting, instead of individual choice.

[21] In their binary trust game, the trustor decides between "trusting" and "not trusting". In case the trustor trusts, the trustee can decide between "honoring" or "exploiting" trust where "exploiting" trust yields a higher payoff for the trustee. "Trusting" is efficiency-enhancing but only pays off for the trustor if the trustee "honours" trust.

[22] All players were trustors and could potentially be trustees as well. In the *choice* treatment a player could be a trustee for up to four players depending on how many players chose her as a trustee. The experimental set-up is also behaviourally complex as subjects potentially played both roles in the investment game. A trustee in the choice treatment may want to give back a positive amount because of reciprocity concerns, in order to receive further transfers from the same trustor in the future or in order to receive high back-transfers themselves in subsequent rounds from his current trustor as a future trustee.

2.4 Hypotheses

2.4.1 Game-theoretic solution based on common knowledge of rationality

Based on the commonly employed assumption of common knowledge of rationality it is straightforward to derive the game-theoretic solution. In the last round allocators will play their money-maximizing action which is to send zero back regardless of the amount received. Since the finite end of the experiment after 25 rounds is common knowledge investors anticipate this and will consequently not send anything in the first place. Therefore no rent is available in the last round that may induce the allocator to cooperate in earlier rounds. Allocator will therefore not send anything back regardless of the transfers they receive and investors will not to send anything in the first place. The unique subgame perfect Nash-equilibrium for investments and back-transfers in both treatments is zero in any round.[23] Hence, political competition is not predicted to influence relative back-transfers and investments.

2.4.2 Behavioural hypotheses

Based on a vast amount of experimental evidence of cooperation in finite games (see Andreoni and Miller 1993), the game-theoretic solution of section 2.4.1 is not expected to predict the experimental results well. Furthermore, the cited experiments closer to our research question show a considerable level of cooperation in finite play. In particular our companion paper (Walkowitz and Weiß 2009) shows cooperation to emerge in this design even if effective reputation building is ruled out. We therefore expect cooperation to emerge in both treatments and our design therefore to be suitable to test the effects of political competition.

In order to derive qualitative hypotheses for the effects of political competition in our experiment, we need a theoretical framework that allows for reputation building in finite games but still sticks to the assumption that allocators are, on average, driven by self-interest. Models relaxing the assumption of common knowledge of money-maximising

[23] The backward-induction argument goes as follows: With common knowledge of rationality, i.e. everyone knowing that everyone is rational and money-maximising, any allocator will unambiguously confiscate all available assets in the last round. Investors anticipate this and will therefore not send anything in the last round. The allocator therefore has no incentive to send anything back in the penultimate round which rational investors again foresee and so forth. The unique Nash equilibrium that satisfies subgame perfection is therefore zero investments throughout the experiment.
The uncertainty of incumbents over their survival as allocators does not change the rationality prediction. As long there is a definite known end to the game, backward induction along the lines above can be applied as no allocator would send anything back in the last round.

rationality and instead assuming incomplete information about agent´s pro-social preferences and / or rationality allow for reputation building (Kreps et al. 1982). There is scope for cooperation throughout the experiment as long as subjects attach a sufficiently large probability to earning rents from cooperation at the end of the experiment or do not induct the entire game backwards (Rapoport, 1997; Weber, Camerer and Knez, 2004; Charness and Levin, 2005). Investors may expect their allocator to reciprocate investments even in the last round. In this case they may still send positive amounts in the last round. If allocators can strengthen the belief of investors by reciprocating in earlier rounds, it may pay be profitable to build a reputation as a reciprocal allocator. The belief of investors that at least some allocators will reciprocate may also be justified if we depart from the assumption that all players are money-maximisers. Experimental evidence abounds that shows that some subjects do reciprocate at their own cost while, unsurprisingly, the money-maximising homo oeconomicus assumed in most economic models exists, too (compare Andreoni and Miller 1993; Fischbacher, Fehr and Gächter 2001). Money-maximising players will only cooperate if they expect to profit from it. Reciprocal players are willing to cooperate even without material gain but are, based on experimental evidence, self-serving on average.[24] Assuming the co-existence of self-serving reciprocal and money-maximising players when the individual type is private information is a practical starting point to analyse the effects from political competition in a game with finite horizon. It allows for reputation building in finite play but still does not deny the conflicting interests between investors and presidents.

The first hypothesis describes the core ingredient to test the GN-predictions behaviourally: the voting mechanism. We predict voters to condition their votes on the track-record of incumbents. The more incumbents gave back relative to the investments they received, the more votes they are predicted to *receive* from their electorate. This implies

> *H1: The re-election probability increases in relative back-transfers of the previous legislative period*

[24] Fischbacher et al. (2001) find in a public good game that a large part of the players can be classified as either conditional co-operators or money-maximisers. Conditional co-operators are reciprocal in that they cooperate if the counterpart cooperates too. Most conditional co-operators, however, only partially reciprocate and leave more for themselves compared to their counterparts. This finding is also supported in a one-shot investment game played with the same subject pool as this experiment (see Walkowitz and Weiß 2009). Most players therein display self-serving reciprocity in the absence of strategic reasons to cooperate. Reciprocal and also more generally cooperative behaviour can be explained by pro-social preferences that are for example linked to inequity aversion (Fehr and Schmidt 1999; Bolton and Ockenfels 2000).

As political competition is hypothesised to link the survival probability to the level of taxation or, in our design, relative back-transfers, it is instructive to first analyse the effects from political competition until the last election takes place and to treat legislative periods, rather than individual rounds, as analytical units. Based on our assumptions, a money-maximising player reciprocates investments if she believes this to pay off in later legislative periods. Specifically, she will do so if her expected gain in profits in later periods is at least as high as the one-time gain from not reciprocating, i.e. choosing the revenue-maximising level of relative back-transfers. The expected increase in profits can be divided in two parts: first, by building a reputation as a reciprocal allocator, she may receive higher investments in later periods. Second, the possibly higher investments need to be weighted with the expected future time in office. Political competition is hypothesised to affect the second part.

Without political competition, in NoPolComp, the prospective time in office is independent of behaviour and is linked to a constant survival probability of 1/3 from each period to the next. With political competition, in PolComp, the survival probability is hypothesised to increase with the level of relative back-transfers. Thus, on the one hand, incumbents in PolComp may ceteris paribus give back higher amounts than incumbents in NoPolComp in order to boost their re-election chance. On the other hand, if incumbents believe the survival probability to be too insensitive to their back-transfer decisions they would be materially better off to choose the revenue-maximising level of relative back-transfers. Incumbents in PolComp may also believe increasing their re-election probability above 1/3 to be too costly to be profitable. In this case, they have a shorter prospective time in office than incumbents in NoPolComp. A shorter horizon would ceteris paribus lead to lower back-transfers in PolComp than in NoPolComp. Hence, the effects from political competition on relative back-transfers, and consequently on taxation, depend largely on the relation between the survival probability and relative back-transfers. The relation is an empirical question as it depends on the voting behaviour of the investors for which we do not have a theory at hand. The inconclusiveness also mirrors the non-linear effect of political competition on the behaviour of the leader in the model of Grossman and Noh (1994) and its dependence on the survival probability of the political system. We can, nevertheless, say that if voters reward incumbents who give back more than the revenue maximising level, by voting for them in subsequent elections, the following re-election hypothesis emerges:

H2: With political competition allocators send back more for a given level of investments than without political competition as long as there is still the prospect of re-election (i.e. for the first four legislative periods)

H2 is the principal hypothesis as it is directly linked to the research question whether political competition can lower rent-seeking taxation by increasing relative back-transfers in the experiment.

For any individual round, the revenue-maximising level of back-transfers is zero regardless of investments. However, as investments receivers serve a minimum length in office of five rounds, back-transfers of zero will not be revenue-maximising over the course of a legislative period as investors will lose their trust in their allocator. Instead, allocators may try to build a reputation within each legislative period by setting relative back-transfer higher than 1, to which trusting investors would respond with positive investments.

In the last round of any non-last legislative period (i.e. rounds 5, 10, 15 and 20, which we also call intermediate last rounds), the survival probability drops from 1 to 1/3 in the absence of political competition. We therefore predict tax rates in intermediate last rounds to be higher than in the first four rounds of any legislative period in the absence of political competition.

H3: Without political competition relative back-transfers will drop at intermediate last rounds

Based on our framework, investors only care about payoffs; nevertheless, their previous payoffs may not be the only relevant information in order for them to form expectations about future payoffs. When taking their investment and voting decisions, investors may evaluate the credibility of candidates' promises and may also view current trustworthiness, i.e. how closely allocators stick to their promises, as a signal of their future trustworthiness. Hence, candidates may carefully choose the promises they make. In both treatments, allocators are interested in triggering high investments; this should, at least before investors have made experience with actual back-transfers behaviour of the allocator, require a promised rate of return that is both positive and credible. Candidates in PolComp have an additional incentive to promise more than their opponents as long as their promises are still credible; otherwise, voters may avoid candidates promising to be very generous. Hence, we

can expect promised relative back-transfers to be at least as high in PolComp as in NoPolComp. However, it seems too far-stretched to predict promised relative back-transfers to be even higher due to political competition. The reason is that the parsimonious framework we employ does not allow for predictions on credibility.

H4: Promised relative back-transfers are not lower in PolComp than in NoPolComp

If promises are not higher in the presence of political competition, we would, based on H2, expect allocators to be more trustworthy in PolComp than in NoPolComp as long as re-election still motivates incumbent allocators. If, however, promises are higher in PolComp, treatment effects on how well allocators stick to their promises are unclear. On the one hand, higher relative back-transfers would mean less cheating by allocators in PolComp. On the other hand, higher promised relative back-transfers would mean more cheating in PolComp.

In case H2 is correct, investors face higher returns on their investments. Nevertheless, all incumbents, independent of political competition, have an incentive to receive high investments. For any level of relative back-transfers, the higher are investments the higher will be profits of allocators. Therefore, without any more elaborate theory on investor behaviour, which may predict the elasticity of investments to their profitability, we can only predict that investment will, throughout the experiment, not be lower in the presence of political competition.

H5: Investments will not be lower in PolComp than in NoPolComp

Taking hypotheses 2 and 5 together leads to the following prediction:

H6: Investor profits are higher in PolComp than in NoPolComp for the first four legislative periods

Predicting allocator profits based on H2 and H5 is less straightforward as allocators are predicted to give back relatively more in PolComp than in NoPolComp but are also predicted to receive more. However, by the logic underlying H2 we predict political competition to drive relative back-transfers above the revenue-maximising level. Consequently, the hypothesised lower relative back-transfers in NoPolComp will be closer to the revenue-maximising level. We hence predict

H7: Political competition restricts rent-seeking behaviour by presidents and therefore leads to lower profits of presidents in PolComp than in NoPolComp for the first four legislative periods

By the logic underlying H2, the prospect of re-election should motivate incumbents more the longer their horizon is. In the first legislative period, incumbents can be re-elected four times, whereas in the last period no re-election awaits anymore. Consequently, the differences in strategic incentives for incumbent allocators between the treatments decrease over time. As investors' anticipation of allocator behaviour also drives investment levels, we hypothesise:

H8: the effects of political competition on relative back-transfers and investments levels decrease over time

In the last legislative period, the strategic incentives for the allocator are the same in PolComp as in NoPolComp. Hence, based on strategic reasoning we do not expect any effects of political competition.

H9: in the last legislative period, relative back-transfers and investment-levels do not differ between PolComp and NoPolComp

2.5 Procedure

We collected 9 independent observations for both treatments. In total 144 subjects participated in the computerised laboratory-experiment run in October 2007 at the Herbert A. Simon & Reinhard Selten Behavioral Decision Research Lab, Southwest Jiaotong University, Chengdu, China. The experiment was programmed in ASP.NET. Separate cubicles with view-proof curtains made sure that interaction between the subjects was completely anonymous. It was common knowledge that the experiment only started after all subjects correctly answered test-questions about the experimental set-up, including how earnings are calculated. Prior to the main treatments, a bilateral one-shot investment game was played using the strategy method (Selten 1967), without feedback and with constant roles between the one-shot game and the treatments PolComp and NoPolComp. The one-shot game was played in order to allow for comparability between PolComp and NoPolComp and the treatments VOT and RAN of our companion paper (Walkowitz and Weiß 2009).[25] Subjects´

[25] In Walkowitz and Weiß (2009), the procedure of the one-shot game is described in more detail (see section 3.2 of this thesis as well the instructions provided in the appendix of this chapter).

earnings were set at local standards. Sessions lasted on average about 3 hours[26], in which subjects earned an average of 92.1 RMB (about USD 12.3 at the time of the experiment), paid out in cash individually at the end of the experiment.

2.6 Results

2.6.1 Hypotheses testing

In contrast to the theoretical model of GN (1994), we fill the black box of political competition with a voting mechanism. Therein, voters endogenously determine how incumbents' chance to stay in power is linked to their implicit taxation decisions. Within the framework of our design, H1 predicts that the survival probability of incumbents is increasing in relative back-transfers. The data reveal that this is indeed the case. Average relative back-transfers of the previous period are 1.78 for re-elected allocators, while average relative back-transfers of the previous period are only 1.43 for non re-elected allocators. In all groups in which voters both re-elected and ousted incumbents, relative back-transfers of the previous period are higher for re-elected allocators.[27] The voting hypothesis H1 therefore receives strong support by the data. Figure 2-3 also reveals graphically that the ex-post re-election probability is strongly increasing in relative back-transfers.

Result 1: The re-election probability increases in relative back-transfers.

For relative back-transfers approaching maximum relative back-transfers (above 1.8) the ex-post re-election probability is 1. In these instances, which occur in 22.22% of cases, the decisions by allocators entail little rent-seeking as the implicit tax rate is less than 10%. In this region (more precisely for relative back-transfers of 11/6 (about 1.83), the allocator receives the same profit as the investors. While it seems difficult to argue that 11/6 sticks out as a possible fairness point, it is nevertheless interesting to note that back-transfers which give investors the same average profits as the allocator virtually assures re-election.[28]

[26] Besides the actual experimental play, the duration includes giving instructions; clarifying open questions; subjects answering test questions; the one-shot investment game; a risk measure experiment and pen & paper hand-written survey questions after the experiment as well as individual payments.

[27] p = 0.063 (Wilcoxon paired two-sided); in 2 out of the 9 groups allocators were never re-elected while in 2 groups the allocator never changed.

[28] It requires some – albeit simple – algebra to calculate the relative back-transfers that equalises average investor payoffs with the allocators´ payoff. It seems hard to imagine that investors explicitly calculated the corresponding back-transfers for each of their investments.

Note on the other hand that no allocator was re-elected who chose, on average, a lower relative back-transfer than 1.49; this also happened in only 25% of all cases. Allocators were obviously aware that they could significantly increase their re-election chance by giving back more than the equal split.

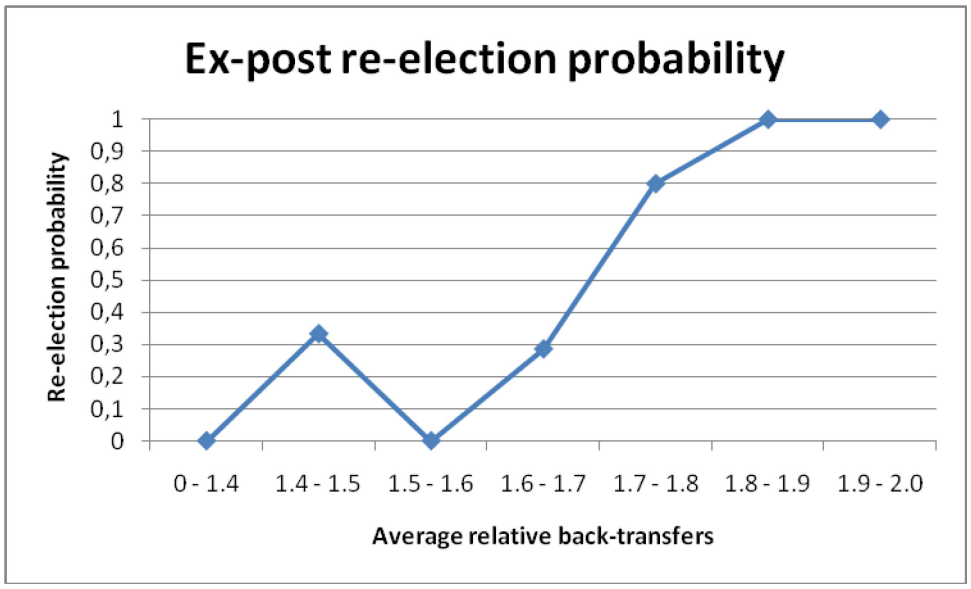


Figure 2-3: Re-election probability dependent on relative back-transfers

By giving back more than the equal split of the returns from investing, incumbents can increase their re-election probability above the probability to be randomly drawn in the control treatment NoPolComp. On average, the ex-post re-election probability is 52.8% in PolComp, which, at first sight, seems to be higher than the probability of a random draw with probability 1/3, as in NoPolComp; however, this is only true for 6 of the 9 groups. In 3 groups, incumbent allocators' ex-post re-election probability is less than 1/3. In those groups, voters were, at least initially, very demanding. The ex-post probability to be elected is therefore not significantly higher than either the realised probability to be chosen in NoPolComp[29] or the ex-ante probability of 1/3 to be selected as in the random mechanism in NoPolComp.[30]

Result 1.1: Highly reciprocal incumbents are rewarded with a high re-election probability.

[29] Mann-Whitney U test PolComp versus NoPolComp: p=0.218 (two-sided)
[30] Binomial-test for whether significantly more groups have an ex-post re-election probability above 1/3 than below 1/3: p=0.508 (two-sided)

The principal political competition hypothesis H2 predicts that in this case political competition will lead to higher relative back-transfers as long as allocators are still motivated by re-election, i.e. for the first 20 rounds. Treatment comparison shows that allocators in PolComp indeed choose higher relative back-transfers than their counterparts in the control treatment NoPolComp (see Figure 2-4). While elected allocators send back an average of 61.37% more than investments received, the corresponding figure for randomly drawn allocators is only 27.66%. In the first round, relative back-transfers are higher in PolComp but not significantly so but increase afterwards highly significantly above the level in NoPolComp ($p<0.001$ Mann-Whitney U, two-sided). On average, relative back-transfers in PolComp are always incentive-compatible for investors.

> *Result 2: Political competition leads to higher relative back-transfers as long as re-election is still possible.*

Our experiment therefore supports the prediction of GN (1994) that political competition may lead to more benevolent behaviour by the ruler. Interestingly, however, a high re-election probability is no prerequisite for high relative back-transfers. Even in those 3 groups that had a high turn-over of allocators, relative back-transfers are higher than 1.45, which still is higher than the average relative back-transfers in NoPolComp, and higher than 1 even at intermediate last rounds. The evidence, therefore, suggests that the concerns of incumbents taking a short-term view if their survival probability is low, as formulated by Grossman and Noh (1994) and Bardhan and Yang (2004), are largely unfounded. It seems that the mere pressure of competition already increases relative back-transfers as allocators are strongly motivated to stay in office.

The strategic incentives between PolComp and NoPolComp are most divergent at the end of each legislative period. Over the course of one legislative period, both with and without political competition, incumbents may try to build a reputation as a reciprocal allocator in order to incentivise investments. At the end of a legislative period, however, only elected allocators can be motivated by re-election; furthermore, they can, as we have shown, push the re-election probability above 1/3. The survival probability for incumbents in NoPolComp, by contrast, drops to 1/3 irrespective of their behaviour. H3 therefore predicts relative back-transfers to drop in NoPolComp at the end of legislative periods in the absence of political

competition. We indeed find weakly significantly lower relative back-transfers (p=0.098 Wilcoxon paired, two-sided) in the last rounds of the first four legislative periods (1.13) than in the corresponding preceding four rounds (1.30). By contrast, with political competition relative back-transfers do not drop significantly prior to the selection of allocators. Even in intermediate last rounds, elected allocators give, on average, back more than 50% of the gains from investment (55.6% in intermediate last rounds versus 62.9% in the first four rounds).

> *Result 3: Without political competition relative back-transfers drop prior to selections of allocators but are still incentive-compatible on average. By contrast, with political competition, allocators do not give back relatively less in the last round of their reigns than in the first four rounds.*

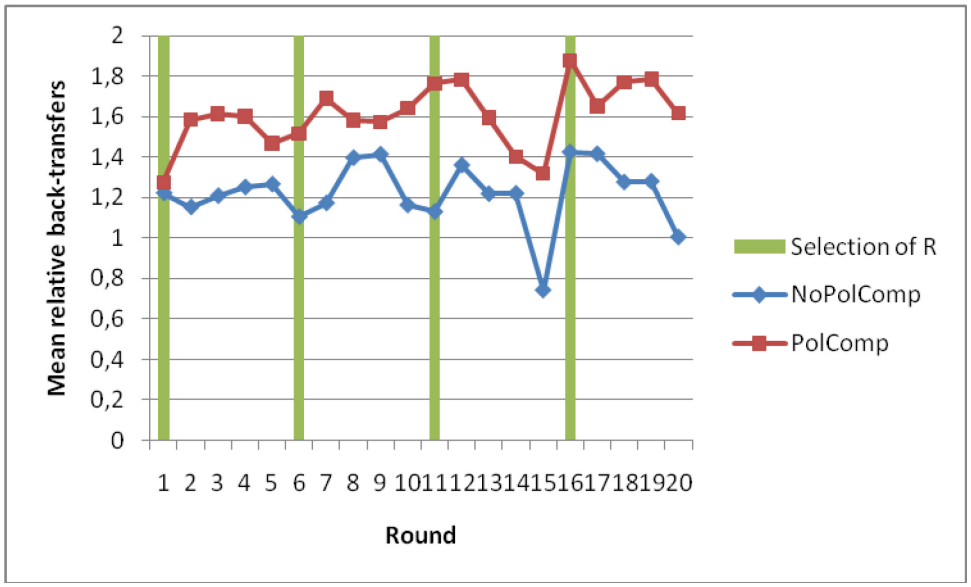


Figure 2-4: Relative back-transfers of allocators (R) with and without political competition

Note that already the re-chosen probability of 1/3 in the control treatment NoPolComp seems to be sufficient for incumbents to send back incentive-compatible back-transfers (that turn investing profitable) with the exception of the third legislative period.

Do the higher relative back-transfers in the presence of political competition mean allocators in PolComp stick more closely to what they promised to the investors? As it turns out, they do not: allocators in PolComp give back, on average, 1.59 points less than what they promised, compared to 1.15 points in NoPolComp; this difference is, however, not

significant[31]. The obvious reason for allocators in PolComp not cheating less on their investors than allocators in NoPolComp is that they promise investors significantly[32] higher returns on their investments. In the presence of political competition, allocators promise (for those investment-levels chosen by the investors) an average return of 84.4%, compared to only 47.9% in the absence of political competition.

> *Result 4: Political competition induces higher promised relative back-transfers. As a result, allocators do not cheat less in PolComp than in NoPolComp.*

In both treatments, investing is profitable on average. Average relative back-transfers lie above 1 with and without political competition[33] as most allocators try to build a reputation as a reciprocal receiver of investments. Because of the incentives of reputation building for allocators in both treatments, H5 only went as far as predicting investments not to be lower with political competition. The results, however, go beyond H5 as investments are significantly higher with political competition than without political competition (p=0.011, Mann-Whitney U, two-sided). Based on the aggregate comparison of profitability, this very clear effect of political competition on investment levels may seem surprising. A closer look at the group and the individual level, however, reveals that aggregated values do not tell the whole story. At first sight, even then, the differences between the two treatments do not seem to be high: the share of investments that lead to positive rents of investors is 90.7% with political competition and 82.2% without political competition. A group-level comparison shows this difference to be of marginal significance (p=0.1095, Mann-Whitney two-sided). However, digging deeper we find this comparison to underestimate the true extent of differences between the treatments as it neglects the behavioural response of investors to negative rents from investing. In 3 of 9 groups in the control treatment the mode of investments and in 2 groups even the median of investments is zero, mainly because investors seize to send positive investments after having received a negative return on their investments. Looking at the share of profitable investments, 83.02% of all investment decisions lead to positive rents in the presence of political competition, but only 57.16% in the absence of political competition. This difference is significant (p=0.011, Mann-

[31] Mann-Whitney U, two-sided: p=0.605
[32] Mann-Whitney U, two-sided: p<0.001
[33] Binomial-test for average relative back-transfers lying with equal probability above and below 1: p=0.004 (PolComp) and p=0.039 (NoPolComp)

Whitney U, two-sided). Interestingly, political competition still leads to weakly significantly higher investments even if we take out the three idiosyncratic groups that entail close to or even above 50% of investments of zero (p-value = 0.088, Mann-Whitney U, two sided). As efficiency, defined as realised total payoffs over maximum total payoffs, is strictly increasing in the level of investments, political competition also raises efficiency in our experiment. On average, 41.11% more surplus is generated in PolComp than in NoPolComp. In NoPolComp, average investments are 55.92% of endowment, whereas this figure reaches 78.90% in PolComp.

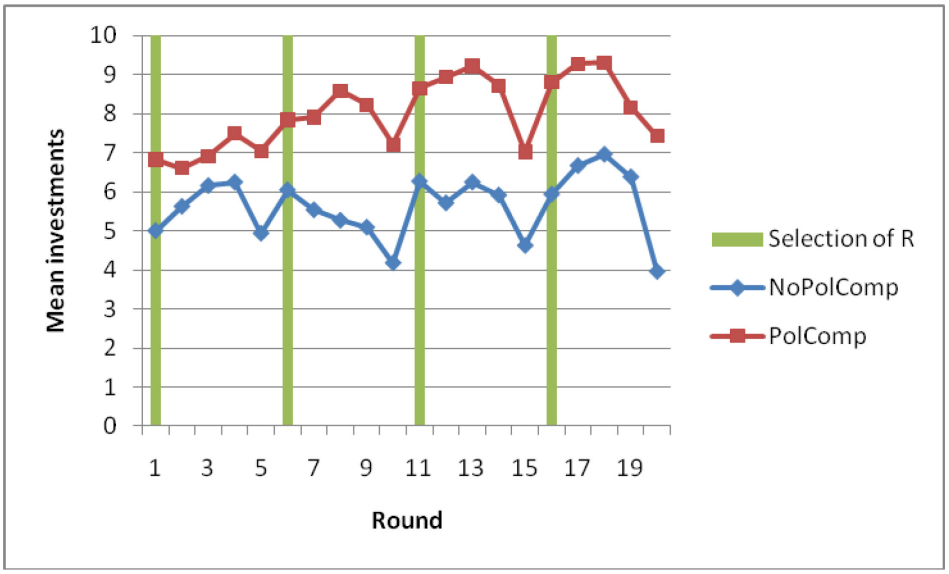


Figure 2-5: average investments with and without political competition

Result 5: Political competition leads to higher investments and, consequently, higher efficiency as investors respond to the higher profitability of investing in PolComp

We find political competition not only to increase investment levels but to also to lead to a more stable level of investments. Between-group variance of investment levels is only 1.192 in PolComp but 5.803 in NoPolComp, which is a significant difference (p=0.023, two-sided, two-sample randomization test for differences in variance). The same applies to within-group variance, which is significantly higher (p=0.050, Mann-Whitney U, two-sided) in NoPolComp (12.377) than in PolComp (8.463). The main reason seems to be the more varying behaviour of allocators without political competition: some incumbents in NoPolComp fail to incentivise investments (which happens in only one period in PolComp),

while others build a good reputation. Investment-levels consequently vary more without political competition.

Result 5.1: Political competition leads to more stable investment levels.

The higher relative back-transfers, which are, on average, also larger than 1, imply that investing is more profitable in the presence of political competition. Since investments are also higher with political competition, investor profits are necessarily higher too. On average, investors in PolComp earn every round a rent of 5.32 (or a payoff of 15.32), whereas investors in NoPolComp have to be content with a rent of 2.09 (or a payoff of 12.09) per round. As this difference is highly significant ($p < 0.01$, Mann-Whitney U two-sided), hypothesis H6 receives strong support from the data.

Result 6: Political competition leads to higher investor profits.

In deriving H7, we predicted allocators to be closer to the revenue-maximising level of relative back-transfers in the absence of political competition. We therefore hypothesised allocators to earn more without political competition. Average values are indeed somewhat different (27.61 in NoPolComp versus 24.06 in PolComp). Surprisingly, however, the difference between the treatments is only of borderline significance ($p=0.1359$, Mann-Whitney U, two-sided). Does this result mean that political competition does not come at a cost for incumbents? One may argue that political competition may act as a commitment device through which allocators become more credible in choosing higher relative back-transfers and therefore receive higher investments. While this reasoning sounds intuitively appealing, we think it only partly survives closer scrutiny of the data. Recall that three out of the nine groups in the control treatment, i.e. without political competition, are characterised by a large number of investments of zero. Closer inspection of the reasons for low investments reveals that in all of these three groups allocators are drawn who fail to incentivise investments; they do not build a reputation as a reciprocal allocator.[34] Hence, in these three groups, allocators choose relative back-transfers clearly lower than the revenue-maximising level. Once allocators are stuck at the wrong end of the Laffer-curve of allocator

[34] In the first rounds of their reigns, investors are initially trusting and send positive investments. The respective allocators, however, send nothing back and do not even change their strategy when they are offered smaller investments in later rounds. In one group, efficiency suffers additionally from a morally hazardous decision of the former allocator in her last round. Nevertheless, 3 out of the 5 investors are initially willing to put some trust in the new allocator on which he, however, fails to build.

revenues depending on implicit tax levels, they are virtually excluded from future revenues within the legislative period. Re-running significance tests reveals that the treatment comparison is fairly sensitive to the inclusion of allocators stuck at the wrong end of the Laffer-curve. Taking out of the data set the observations of either of the two allocators who were re-chosen two times and therefore had a large negative impact on allocator´s profits leads to significantly lower profits of allocators with political competition.[35] Taking out all observations of either of the two corresponding groups also yields significantly lower profits of allocators with political competition.[36] To conclude, H7 at first sight receives only weak support by the data. A closer inspection of the data, however, lets us support a restricted version of H7: political competition limits rent-seeking behaviour by allocators. If we just compare cases in which allocators do not fail to incentivise investments, political competition also leads to lower profits of allocators.

> *Result 7: Political competition has a limiting effect on rent-seeking of allocators. This leads to somewhat lower profits of allocators. Allocator profits are significantly lower with political competition if we do not consider allocators in NoPolComp stuck at the wrong end of the Laffer-curve.*

By the logic of inter-temporal incentives for allocators, we expect, as stipulated in H8, that political competition will have a larger effect on relative back-transfers and consequently on investment levels in the earlier rounds than in later rounds of the experiment. Surprisingly, neither treatment differences in relative back-transfers nor in investment levels decrease over time. Spearman´s rho is insignificantly positive (instead of negative) in both cases (p=0.838 for relative back-transfer differences and p=0.325 for investment differences), which contradicts H8.

> *Result 8: The effects of political competition on relative back-transfers and investments do not decrease over time.*

In the last legislative period, re-election cannot motivate incumbent allocators any more. H9 consequently predicts no effects of political competition on the behaviour of allocators and investors in the last period. Surprisingly, relative back-transfers are still considerably higher

[35] p=0.0745 and p=0.0464 Mann-Whitney U two-sided
[36] p=0.0315 and p=0.0503 Mann-Whitney U two-sided

in PolComp than in NoPolComp as investors get an average rate of return 47.63% with political competition but only 15.16% in the control treatment (see also Figure 2-6 and Figure 2-7). This difference, is, however, weakly significant only if we test differences with a Fisher-Pitman Permutation Test for Two Independent Samples (FPPI), and not with the previously employed Mann-Whitney U Test.[37] The same applies to investments: they are considerably higher in PolComp (7.46) than in NoPolComp (5.42), but weakly significantly so only by testing with a FPPI.[38] Profit of citizens are weakly significantly higher (p=0.0625) with political competition. The last-period treatment differences seem puzzling and contradictory to the intuition behind H9 but follow along the lines of result 8, which shows effects of political competition not to decrease over time.

To further explore the effects of political competition in the last legislative period, it is insightful to separate the latter into its first four rounds, that is for as long as reputation building is still possible, and the experiment´s very last round. For the first four rounds, relative back-transfers continue to be significantly higher with political competition (p=0.04, Mann-Whitney U two-sided). While average values again suggest investments to be higher with political competition for rounds 21-24 (8.3 versus 6.07) significance tests reveal only a marginal significance (p=0.114, Mann-Whitney U two-sided and p=0.090, FPPI two-sided). Investor profits are significantly higher with political competition (p=0.024 Mann-Whitney U two-sided), whereas allocator profits are unchanged by political competition because some allocators, as also observed in previous periods, fail to incentivise investments. In the last round of the experiment, political competition seizes to effect relative back-transfers, investments and profits of either investors or allocators. In both treatments, there is a substantial end-game effect. Yet, if there is any treatment difference at all, elected allocators tend to give back relatively less than their randomly drawn counterparts.[39] This leads to somewhat lower average profits of investors (p=0.138 Mann-Whitney U, two-sided) in

[37] In contrast to the standard Mann-Whitney U test, which is only based on rank-comparison of two distributions, the Fisher-Permutation Test for Independent Samples (FPPI) takes the actual values of the observations into account. Despite the considerably lower average values in NoPolComp, some allocators without political competition still provide high rates of return. Consequently, the Mann-Whitney U leads to an insignificant result, whereas treatment differences are weakly significant according to the FPPI: FPPI (two-sided): p=0.092; Mann-Whitney U (two-sided): p=0.258.

[38] FPPI: p=0.097 (two-sided); Mann-Whitney U: p=0.161 (two-sided)

[39] Without political competition, 3 presidents choose average relative back-transfers of less than 1 while 3 turn investing profitable by giving back more than they received. The other 3 were not trusted by the investors and received zero investments. With political competition, 7 presidents send back less than or equal to what they received (of which 5 send back zero) while 2 send back more than what they received. The differences in relative back-transfers (0.47 with political competition versus 0.86 without political competition) are, however, not significant (p=0.243 Mann-Whitney U two-sided).

PolComp. The endgame-effect is also pronounced in investor behaviour. Nevertheless, investments still tend to be higher in PolComp (4.11) than in NoPolComp (2.8) – albeit not at a statistically significant level (p=0.188 Mann-Whitney U, two-sided).

> *Result 9: The effects of political competition on relative back-transfers and investor profits carry over into the last legislative period until the penultimate round.*

The results from the last legislative period are puzzling. On the one hand, political competition limits rent-seeking in the first four rounds but leads to, if there is any difference at all, more rent-seeking in the very last round. From an economic point of view, the differences in the first four rounds are astonishing given that strategic incentives are identical with and without political competition. Because of this, no economic rational based on pecuniary outcomes alone, as in standard economic theory or in the theoretical framework of section 2.4.2, can explain these differences. It goes without saying that human interaction is more complex than our parsimonious framework assumes; more importantly, the predictive power of our framework is poor in the first four rounds of the last legislative period, whereas it is doing a fairly good job in the first four periods. This invites us to think about treatment effects that have hitherto not been considered and that may help us explain last period behaviour and that can possibly enrich our understanding of the treatment effects.

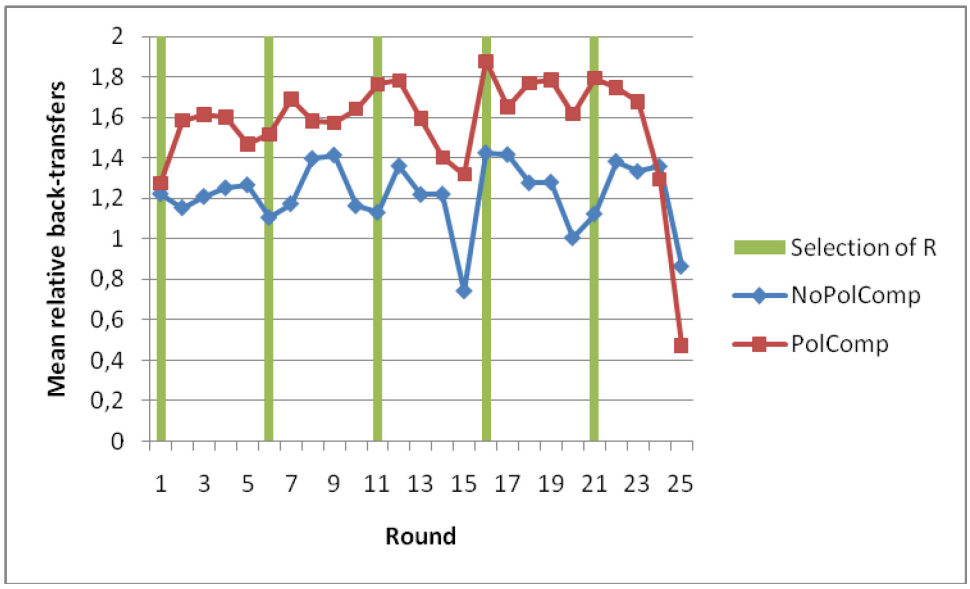


Figure 2-6: mean relative back-transfers over all 25 rounds

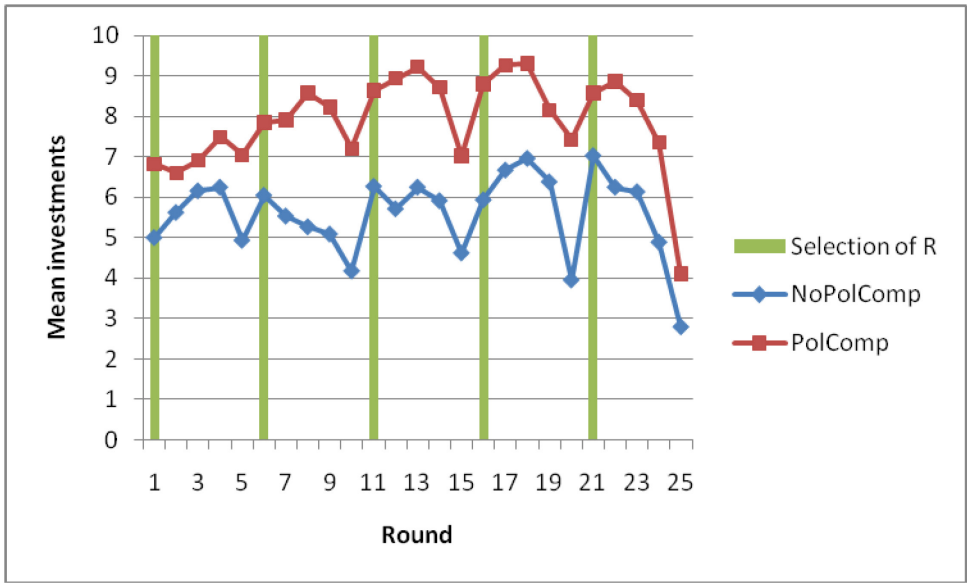


Figure 2-7: Mean investments over all 25 rounds

2.6.2 The last-period puzzle and alternative explanations for treatment-differences

The voting mechanism we employed to induce political competition may itself have behavioural effects. In Walkowitz and Weiß (2009), even in the absence of effective reputation, elected allocators give back relatively more than their randomly drawn counterparts. Independent from allocators' sending back decisions, investors are also more trusting initially. The authors conclude that the voting procedure by itself make a difference to allocator and investor behaviour. A crucial element for elections leading to more group-oriented behaviour by the allocator seems to be whether or not voting is based on explicit promises (Geng and Weiß 2009, chapter 4 of this thesis).[40] In case candidates make promises, elected candidates are likely to attribute their electoral success to the promises they gave and not to them being the better candidate in some intrinsic sense. Consequently, elected dictators transfer more to their group if voting is based on promises, but not if voting is based on personal characteristics (Geng and Weiß 2009). Corazzini et al. (2007) also show the relevance of guilt-aversion in explaining behavioural effects of voting. Compared to a random draw, elections may trigger higher second-order beliefs of allocators (i.e. beliefs of allocators about recipients' expectations), which would lead to more group-oriented behaviour by elected guilt-averse allocators. Being elected may also be considered as a gift, both in material terms, when being elected is potentially profitable, as well as in non-material terms, when allocators appreciate the transfer of power for its own sake (for

[40] All following references to Geng and Weiß (2009) also refer to chapter 4 of this thesis.

further discussion see Geng and Weiß 2009; Walkowitz and Weiß 2009). This gift may be reciprocated by a more group-oriented behaviour by the elected allocator. This reciprocity-motive receives support by Geng and Weiß (2009) as the more group-oriented behaviour by elected dictators can only be partly explained by either guilt-aversion or commitment to promises.

While we have no direct measure of second-order-beliefs, it is plausible to assume that second-order-beliefs increase with the promised relative back-transfers and with relative back-transfers of previous periods. Carry-over effects from previous periods were also found by Bohnet and Huck (2004) who showed second movers in a modified trust game to be more trustworthy after having been exposed to an environment of repeated interaction with the same first mover.[41] If there are no treatment differences in the consistency of second-order-beliefs of allocators, differences in first-order-beliefs of investors, which we did measure, would also guide to differences in second-order-beliefs. Promised relative back-transfers are indeed still considerably higher in PolComp than in NoPolComp in the last period: allocators in PolComp promise, on average, to provide a return of 89.70% on investments, while allocators in NoPolComp present themselves less generous and only promise investors an average 37.39% return on their investments.[42] Investors expect a higher rate of return in PolComp than NoPolComp, already in the first round of the last legislative period. Based on both the history of back-transfer behaviour as well on the promises made, second-order-beliefs of allocators are, as first-order-beliefs of investors, expected to be higher in PolComp than in NoPolComp.

Previous evidence as well as likely differences in second-order-beliefs of allocators turn guilt-aversion and a reciprocal response to being elected into candidates to explain higher relative back-transfers in PolComp than in NoPolComp in the last period. These two explanations, however, are at odds with last round behaviour, for which we would have to add a post-hoc explanation or concede that the last round is behaviourally different from the rest of the experiment.[43] Furthermore, in the last period, allocators cheat more on their promises in PolComp than in NoPolComp; significantly more over all five rounds (on average, allocators

[41] They employ a binary trust game similar to the design later used by Huck et al. (2006).
[42] Mann-Whitney U, two-sided: $p<0.001$
[43] One may speculate that the very last round is salient for allocators and that thus the possibility to exploit investors' trust dominates allocators' behaviour.

give back 3.10 less than what they promised in PolComp, while average cheating is only 1.01 in NoPolComp), but still somewhat more even in the first four rounds.[44] Back-transfer decisions by allocators are also not significantly correlated with promised relative back-transfers: in PolComp, Spearman's rho is even negative (-0.1), albeit far from significant.[45] Obviously, allocators in PolComp aren´t too worried about keeping their promises; they rather seem to make promises strategically in order to win the election. Therefore, if guilt-aversion matters in the last period, its influence seems to be very limited. A third – albeit tentative – explanation for treatment differences in the last period is the establishment of different norms of cooperation between the treatments. Differences in social norms may be due to history effects, i.e. past experiences of investors, or higher expectations on allocators' cooperative behaviour if allocators have been elected based on a promise they made. Possibly, in order to make the same level of investments, investors in PolComp demand higher relative back-transfers than investors in NoPolComp. Different expectations of investors would create different incentives for allocators for the first four rounds of the last period but not for the last round. In the last round, in the absence of strategic incentives, the personal type of the allocator, grossly divided into money-maximisers or reciprocators, determines the sending back decision. It is not straightforward and beyond the scope of this paper to measure the establishment of norms; it seems, however, an intriguing avenue for future research.

Of course, guilt-aversion and reciprocity for being elected are also potential candidates for explaining treatment differences in the first four periods. While we cannot rule out that guilt-aversion matters, as we do not have any direct measure of second-order-beliefs of allocators, a strong role of guilt-aversion seems, for a number of reasons, implausible: first, mirroring the result for the last legislative period, actual relative back-transfers are not positively correlated with promised relative back-transfers in PolComp; this is not even the case in the first round of any legislative period, when promises should be most salient. Interestingly, these correlations are positive and significantly different from zero in NoPolComp: Spearman's rho for correlating average relative back-transfers with corresponding average relative back-transfer promises for all rounds in the first four

[44] For the entire last legislative period: Mann-Whitney U, two-sided: p=0.04; for rounds 21-24: average cheating is 1.97 in PolComp and 0.78 in NoPolComp, p=0.127 (Mann-Whitney U, two-sided) for treatment differences

[45] PolComp: Spearman's rho = -0.100, p=0.798 (9 independent observations); NoPolComp: Spearman's rho = 0.350, p=0.356 (9 independent observations)

legislative periods is 0.04 (p=0.987, two-sided) in PolComp, but 0.473 (p=0.020, two-sided) in NoPolComp; only correlating decisions in the first rounds of the first four periods (i.e. rounds 1, 6, 11 and 16) leads to a Spearman's rho of 0.156 (p=0.500, two-sided) in PolComp, but a Spearman's rho of 0.409 (p=0.047, two-sided) in NoPolComp. Secondly, comparing PolComp to the treatment VOT from our companion paper (Walkowitz and Weiß 2009), we find differences in actual relative back-transfers to be more pronounced than differences in promised relative back-transfers: in the first four periods, the promised return on investments by allocators in PolComp is, on average, 17.76% higher than the promised return in VOT; the difference in actual rates of return is even higher with 24.38% in favour of PolComp. Both treatment differences are significant.[46] The significant and markedly higher rate of return in PolComp compared to VOT also speaks against a reciprocal response to being elected driving the differences between PolComp and NoPolComp. Nevertheless, guilt-aversion and reciprocity for being elected may also contribute to political competition that is induced by a voting procedure leading to considerably less confiscatory behaviour. Possibly, the surprising independence of the treatment effects on the remaining time-horizon (see result 8) can be explained by such motives. It is left for future research to further investigate the behavioural channels through which voting and political competition influence confiscatory behaviour.

The descriptive statistics of Table 2-1 summarises the differences between the treatments:

[46] p=0.001 (for relative back-transfer promises) and p=0.014 (for relative back-transfers), Mann-Whitney U, two-sided

Table 2-1: Summary statistics with and without political competition

Treatment	**PolComp**	**NoPolComp**
Re-chosen probability	Empirical (ex-post): 0.53	Theoretical (ex-ante): 1/3 Empirical (ex-post): 0.31
Mean back-transfers	12.97	7.66
Mean return on investments	59.05%	26.89%
Mean investments	7.89	5.59
Mean profits investors	15.07	12.07
Mean profits allocators	24.06	27.61
Mean promised relative back-transfers (for actual investment decisions)	1.85	1.47

The summary statistics of Table 2-1 also clearly show that, as hypothesised, the game-theoretic solution based on common knowledge of rationality does not describe the data well. Relative back-transfers lie significantly above zero in both treatments and so do, as a prerequisite for positive back-transfers, investments.[47] This vindicates our experimental finite-play approach in testing the effects of political competition.

2.6.3 Summary of results

Our results provide causal evidence for the mechanism of political competition to limit confiscatory behaviour. Political competition motivates incumbents to be re-elected based on a good track-record. Voters reward incumbents for high back-transfers, which incumbents anticipate. The re-election probability of incumbents is therefore higher than the corresponding probability to be re-chosen in the absence of political competition. Investors react to the higher return on investments with political competition in that investment levels and consequently efficiency are higher. Investors are, hence, better off with political competition and earn higher profits. Political competition also has a limiting effect on political rents. However, as some allocators without political competition are

[47] $p<0.001$ (one-sample t-test one-sided) for mean relative back-transfers and mean investments in both treatments as well as for median investments in PolComp. $p=0.01$ (one-sample t-test one-sided) for median investments in NoPolComp.

interestingly stuck on the wrong-end of the Laffer-curve, political rents are only somewhat lower with political competition. Contrary to what economic theory would suggest, the effects of political competition do not depend on the time-horizon of allocators as treatment differences do not decrease as the experiment progresses. Astonishingly, the effects of political competition on confiscatory behaviour and investor profits carry even over into the last legislative period, i.e. when incumbents are not motivated by re-election anymore. Yet more puzzling, behaviour in the very last round entails, if there is any non-accidental treatment difference at all, more rent-seeking by elected allocators. A tentative explanation, the deeper investigation of which would go beyond the scope of this paper, is the establishment of different social norms between the treatments on how the allocator ought to behave. This difference in norms may affect incentives for allocators to reciprocate investments until the penultimate round, but no longer in the very last round.

2.7 What do we learn from the experimental results?

As we argued in section 2.3, the great advantage of properly designed – that is, internally valid – experiments over happenstance data analysis is control over the data-generating process. We exogenously implemented the regime type, either a system with political competition or a system without political competition, in which randomly allocated subjects made their decisions. This allows us to draw causal inference from the experimental results: political competition *causes,* with the reservation of the reported likelihood of a type I error, allocators to confiscate post-investment assets less – in our experiment. Obviously, we would like to draw conclusions beyond the confines of the experiment.

First, we argued in section 2.3 that we incorporated the essential features of the model of Grossman and Noh (1994) in order to test whether political competition can lead to less confiscatory taxation by rulers or – in the terminology of Grossman and Noh (GN) – by a proprietary fiscal authority. Tax levels are implicitly set by the back-transfer decisions of the allocator, who we consider an appropriate experimental incorporation of an authority unrestricted in its taxation decisions. Relative back-transfers of zero are equivalent to a tax rate of unity while maximum back-transfers, i.e. relative back-transfers of 2, imply a tax-rate of zero. Translating relative back-transfers into tax levels, political competition almost halves confiscatory taxation to a level of 21.9% from a level of 39.1% without political competition. By itself, limiting rent-seeking is not necessarily welfare enhancing. Investing full endowment

pays off in our design as long as the implicit tax rate is lower than 50%. Nevertheless, investment levels are different between the treatments and so is, consequently, efficiency. Political competition therefore has positive effects on welfare in our experiment. Transferring the limiting effects of political competition on rent-seeking into the context of the GN-model should lead to an even stronger impact on efficiency. In GN (1994), it is only profitable for producers to contribute their entire effort to the production of the higher-value marketable good if the tax level entails no political rents.

The results also show that elections can induce the welfare-enhancing effects from political competition. First, voters reward incumbents for low tax rates so that their re-election probability increases significantly above the corresponding probability to be re-chosen in the absence of political competition. For average tax rates of 10% or lower over the course of a legislative period, the ex-post survival probability is unity. Second, the competition induced by elections works even with little information. Voters have no information on the possible behaviour of un-elected candidates, who themselves receive no information on the behaviour of either investors or the elected allocator. Yet, newly elected allocators almost never start off with low relative back-transfers. Allocators therefore do not seem to need a learning phase neither on investor nor on voting behaviour. Either already the mere possibility not to get re-elected or the fact that they were newly elected, and therefore investors were likely not content with the previous allocator, are sufficient to guarantee high relative back-transfers. This robust effect of political competition resonates the findings of Huck, Lünser and Tyran (2006) in similarly structured economic contexts. They showed minimal information to be sufficient for competition to increase efficiency in experience goods markets. The results also show that a model on political competition does not need to have assumptions as restrictive as GN (1994) in the following two respects: Firstly, an infinite horizon is not necessary in order for political competition to affect rent-seeking taxation. Secondly, agents may not necessarily be assumed to be money-maximising. We find some allocators to behave cooperative even in the last round and in the absence of political competition; nevertheless, political competition limits confiscatory behaviour. We have therefore shown that political competition curbs confiscatory behaviour with very little information, with a finite horizon and with real-world people that are neither purely money-maximising nor perfectly rational. This result therefore turns the model of GN (1994) rather

more applicable to real-world economic and political systems. As a more general note, the results also vindicate the largely confident view of economists on the positive effects of competition, which especially the Chicago school of thought considers applicable to political markets as well.

While the experimental results can therefore be fruitfully exchanged with future theoretical work, the external validity of the results is still open to debate. How well can the experimental results be translated into real-world contexts? As any theoretic model, an experiment can only focus on a specific part of real-world politico-economic systems. The essential question to assess its external validity is not whether the experiment is a good abstraction of the world outside the laboratory; it goes without saying that real-world democracies or autocracies are in many respects different from the laboratory set-up (and very heterogeneous too). Therefore, the experimental results cannot provide an answer to whether democracies limit confiscatory behaviour compared to autocracies. The appropriate question is rather whether the experiment sheds light on a relevant nexus in the outside world. Firstly, does the experimental setup implement an appropriate abstraction of a structure that may also be found to be relevant in the field? And secondly, given that the structure of the problem is indeed relevant, do we expect the tested *mechanism* to work qualitatively the same in the field too?

We argue that the basic structure of an investment game indeed incorporates a good abstraction of the moral-hazard problem between the state and investors in the absence of rule-of-law. The closest real-world representation may be found in the decision problem faced by small-scale enterprises and workers whether to undertake socially productive investments into physical or human capital. The investor may thereby tap part of the social gain from investing but at the same make herself more vulnerable to confiscatory behaviour by the state as such investments reduce liquidity or may mean moving from the informal into the formal sector (as the formal economy offers better paid jobs or because investments raises the company's size above the tax-administration´s radar).[48]

Assessing the external validity of the experimental mechanism is basically to turn the question on its toes. What aspects are missing in the tested mechanism that may lead to

[48] The reduction of liquidity through investment applies in particular to human capital as investments in human capital cannot be sold as a tangible asset.

different conclusions in the field? In the remainder of the paper, we will consider three dimensions of omitted aspects discussed in the literature: first, informational asymmetries between the ruler and the constituency; second, voter heterogeneity and third, the instability of regimes.

First, in the experiment, political competition works well in aligning the incumbent ruler's incentives with the voters' interests. Incumbents are motivated by re-election, the probability of which they can increase by running on a good track-record. In the field, the disciplinary function of political competition may be limited if voters strongly discount past experiences (Frey 1994) or suffer from informational disadvantages compared to the ruler (as e.g. in the model of Besley and Kudamatsu 2007). The first case opens leeway for incumbents to confiscate assets at the beginning of electoral periods, as long as they satisfy their voters' wishes close to the next election. The second case introduces imperfect information of the voters on the ruler's choices. In this experiment, deviations from promises by the ruler and low profits for the investors can be unambiguously attributed to the ruler's choices; she is always able to meet her promises. In reality, information is incomplete and, as a result, possibly imperfect, too. Changes in economic conditions outside the control of the ruler may render fulfilling her electoral promises difficult or even impossible. In case of asymmetric information about exogenous shocks[49] to the economy, the ruler may be able to hide confiscatory behaviour by blaming changed circumstances. This may lead to a less effective firing-policy by the voters. Furthermore, we abstracted from group decision processes within a ruling group of people and also from principal-agent problems between the ruler and the executing bureaucracy. Both may lead to the relevant actors responding less strongly to the degree of political competition. On the other hand, the experiment also omits the selecting function of elections, which should ceteris paribus lead to more voter-oriented behaviour of the ruler: through often lengthy election campaigns that screen candidates' credentials and personal history, voters may be able to select among "good" and "bad" candidates (for related models see Carrillo and Mariotti 2001; Cooter 2003; Maskin and Tirole 2004; Besley 2005). While incorporating the missing aspects into the analysis may or may not lead to less stark effects of political competition, they do not provide a ground for arguing for political competition leading to *less* alignment of interests.

[49] Shocks are exogenous if they are outside the model considered here; hence, shocks may also arise from within the politico-economic system if they lie outside the control of the ruler.

Second, voter's interests may not necessarily lie at a better protection of property and may therefore not necessarily be conducive to stronger incentives to invest. In this experiment, voters are homogenous in their monetary interests; furthermore, re-distribution between voters was ruled out. A large debate on the economic virtues and vices of democracy, however, focuses on voter heterogeneity and the possibility of the (poor working class) majority expropriating the (rich capitalist) minority (see the discussion in Przeworski and Limongi 1993). Theoretic analysis based on money-maximising rationality and homogeneity in income sources suggests that a majority will re-distribute, by means of non-discriminatory taxation, less from a minority than a revenue-maximising dictator would appropriate himself (Olson 1993; McGuire and Olson 1996). However, this may not necessarily hold if voters differ in their endowment with assets and we allow for direct re-distribution between groups of voters. Incorporating voter heterogeneity therefore seems a fruitful direction for future experimental research.

Third, we did not vary the stability of the political system, which is an important parameter in GN (1994) and also in the empirical investigations of Clague et al. (1996). Based on GN's (1994) analysis, if the survival probability of the political system is sufficiently low, an increase in the degree of political competition will lead to more confiscation, and not less, as the time-consistency constraint is binding. Our experimental results suggest that time-consistency is less of a problem empirically than theoretically. Nevertheless, it may be interesting to test the effects of political competition when the regimes themselves are unstable.

Of course, experiments are only one way of empirical analysis and are complementary to, inter alia, cross-country regressions and analytic narratives. Comparing the results of different methods may be considered a process of academic checks and balances. In order to be confident of both internal and external validity, empirical findings by different methods should tell a similar story. One tentative step in this direction is abstracting from the static nature of the investment game employed in the experiment and thinking of investments as an important determinant of a transitory or even balanced growth rate as in models of Endogenous Growth Theory (see e.g. the overview in Ray 1998). The system with political competition would then generate both higher and more stable growth rates. Tentatively interpreted in this way, the experimental results resemble real-world observations of highly

varying economic performance of countries without political competition in contrast to more stable and on average higher growth rates of democracies (Besley and Kudamatsu 2007). Our experimental results suggest that it may be worthwhile to further investigate political competition as a possibly important reason for (and not only a consequence of) the differences in economic performance between democratic and non-democratic countries.

2.8 Appendix: Experimental instructions

2.8.1 Instructions of the OS-game

Welcome to join us in an economic experiment.

Please read the instructions of the experiment carefully. Do not communicate with other participants in any form throughout the experiment. If you have any questions, read the instruction again, or raise your hand if it is still puzzling. We will reach you and answer your question.

You will be paid cash for participating in the experiment. The money you earn will first be calculated in points. The amount of points you earn depends on your and other participants' decisions in the experiment. You will receive RMB from the points you made in the experiment.

Please make your decision independently and do not communicate with other participants in any form. If you are found to do so, you will be dismissed from the experiment.

All data and answers will be analyzed anonymously. You have already drawn a code, which ensures anonymity. Please make your decision on the computer in the cubicle numbered by that code. We will only know what decision is made by which code, instead of whom.

Today's experiment consists of two independent parts. Now we run the first part. When the first experiment is finished, you will get the experimental instructions for the second experiment. Attention: your decision in the first experiment has no effect on your decisions in the second experiment!

After the first experiment, you will not know about your payoff from the first experiment. The experiment directly goes to the second experiment. Your payoff from the first experiment will be added to your payoff from the second experiment and will be paid to you in cash after the second experiment.

Experimental instruction of the first experiment

You will receive RMB from the points you earned by the following exchange rate:

10 points = 10 RMB

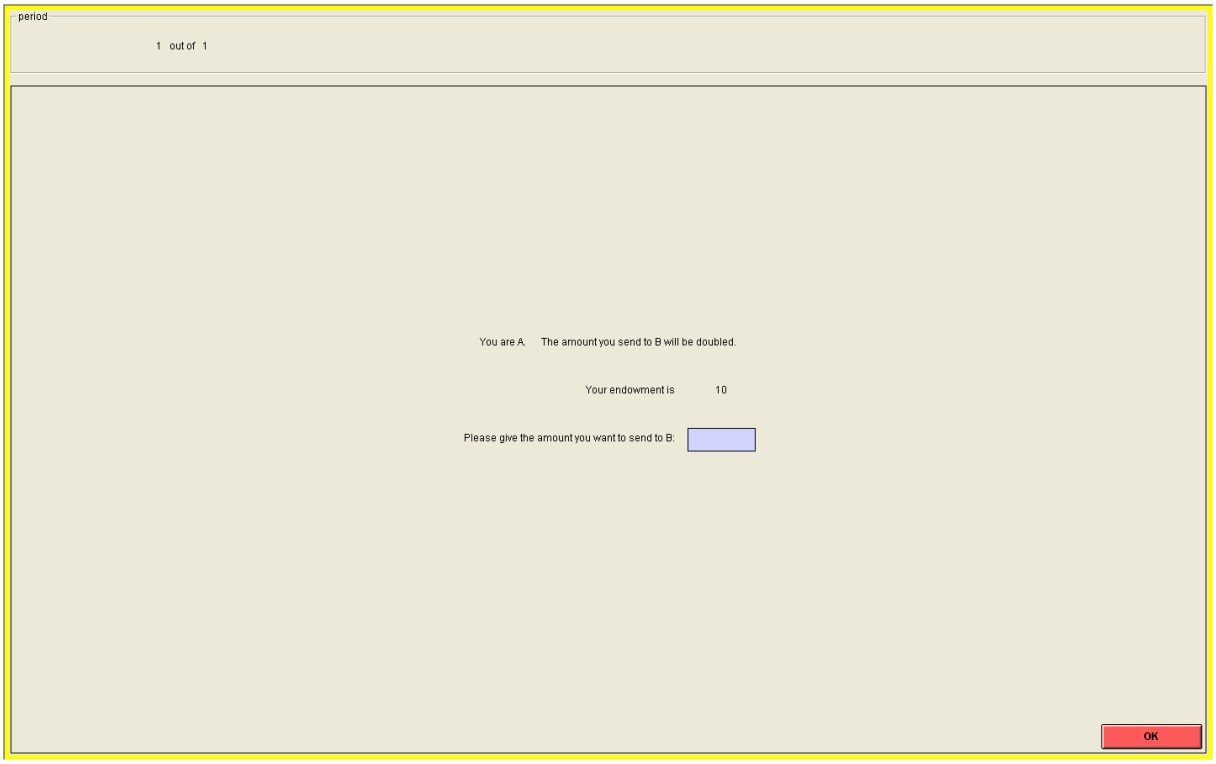
In this experiment, all participants will be assigned either of two types: 15 as type A and 9 as type B. Your type will be randomly assigned in the beginning of the experiment.

The experiment has only one decision round! A and B decide simultaneously. A makes one transfer decision, B makes multiple back-transfer decisions.

At the beginning of the experiment, A and B receive each 10 points as initial endowments. B puts this endowment aside and will get it paid after the experiment.

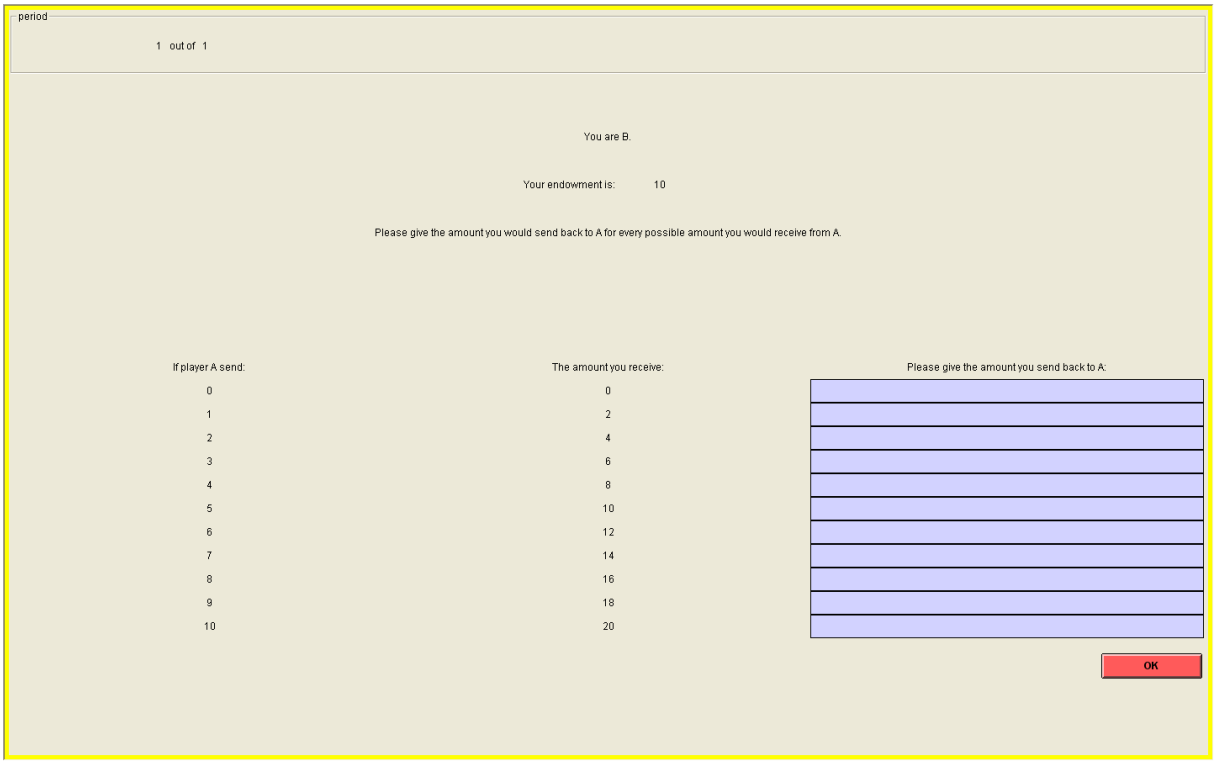
1. A's decision: the transfer decision

A can send any integer part of her endowment to B. That is to say, A can choose any integer number from the set {0; 1; 2; 3; 4; 5; 6; 7; 8; 9; 10} to send to B. The amount sent by A will be doubled by the experimenter. Thus, B will not get A's initial transfer, but twice of the transfer. Here is a screen shot of what A sees on her computer:



2. B's decision: the back-transfer decision

B receives the doubled amount sent by A. B can transfer any integer part of this doubled amount back to A. That means, B can choose an integer number from the interval [0; 2 x A's transfer] to back-transfer to A. However, B does not know the actual amount which was sent by A and doubled by the experimenter. Thus, B must decide a back-transfer amount for each possible transfer of A. For both decision-makers, the decision relevant for the payoffs is that back-transfer amount which B chooses according to the actual transfer made by A. The amount back-transferred from B to A will not be doubled. Here is the screen shot of what B sees on his computer.



How to calculate the income from the first experiment?

A's income:

Every one of the 15 As will be randomly matched to one of the Bs. The payoff is calculated according to the actual decisions of the matched A and B.

A's income depends on the points she sent to B and the points back-transferred from B to A. A's income is calculated as follows:

A's income = A's endowment (10 points) – points sent from A to B + points back-transferred from B to A

You can see from the formula, the more B back-transfers to A, the higher is A's income. The less A sends to B, the higher is A's income.

B's income:

Every one of the 9 Bs will be randomly matched to one of the As. The payoff is calculated according to the actual decisions of the matched A and B.

B's income depends on the points A sent to him and the points back-transferred from him to A. B's income is calculated as follows:

B's income = B's endowment (10 points) + 2 x points sent from A to B – points back-transferred from B to A

You can see from the formula, the more A sends to B, the higher is B's income. The less B back-transfers to A, the higher is B's income.

Payoffs of all As and Bs are calculated by the formulas mentioned above. Every A can calculate the B's income whom she is matched with; every B can calculate the A's income whom he is matched with.

2.8.2 Instructions for treatment PolComp

Now the second experiment starts. Please read the experimental instructions carefully. You will be again paid cash for participating in the experiment. The money you earn will first be represented in points. The amount of points you earn is depending on your and other participants' decisions in the experiment.

You will receive RMB from the points you earned by the following exchange rate:

10 points = 2 RMB

In this experiment, you will interact with other 7 participants. At any time, you will not know who they are. Similarly, other participants will not know who you are.

Experimental instruction of the second experiment

There are 24 participants in the experiment. At the beginning, you will be randomly assigned to a group of 8 members. Then you will be assigned a role in the experiment. All participants will be assigned either of the two types: 5 will be citizens and 3 will be candidates. Your current type depends on your type in the first part of the experiment. That is to say, if you were type A in the first experiment, you are citizen now; if you were type B in the first experiment, you are candidate now. Every candidate will receive a unique code. Attention: the composition of each 8-member group, the role of every participant and the code of each candidate remain unchanged throughout the experiment.

There will be 25 decision rounds in total, with every 5 rounds as a unit. The decision environments of each unit are identical, which means that the experiment process of the first unit (rounds 1-5) is the same as that of the second unit (rounds 6-10); of the third unit (rounds 11-15); of the fourth unit (rounds 16-20) and of the fifth unit (rounds 21-25).

At the beginning of each round, all group members receive 10 points as initial endowment. The 5 citizens will decide on how many points to send. At the beginning of each unit (round 1, 6, 11, 16 and 21), one of the 3 candidates will be selected as the president for the following 5 rounds of each unit (described below). In each round of the unit, all citizens can decide to send any integer point to the president from their initial 10 points. That means, citizens can send any integer they wish to the president from the set {0; 1; 2; 3; 4; 5; 6; 7; 8; 9; 10}.

The president will receive double points as sent by the 5 citizens. That means, each amount of points sent by citizens will be doubled and sent to the president. Thus, the president will get twice the amount of points sent by citizens, instead of the same amount. The amount of points that citizens sent to the president will be shown to the president in a descending order. That means, the president will see how many points he gets from the citizens, but not which citizen sent him how much. The president must decide, for each amount of points, how much to back-transfer. The president can back-transfer any integer from the doubled points he received. That means, the president can back-transfer any integer from the interval [0; 2 x sent amount], to citizens. The amount of points the president back-transfers to citizens will not be doubled.

How to select the president?

At the beginning of each unit (1st, 6th, 11th, 16th and 21st round), the 3 candidates can announce a back-transfer suggestion, as if they were selected to be president. The back-transfer suggestion is how much each candidate, as president, will back-transfer given citizens' transfers (0, 1, 2, 3, 4, 5 , 6 ,7 ,8 ,9 ,10) after he receives points sent from citizens (0, 2, 4, 6, 8, 10, 12, 14, 16, 18, 20). These suggestions serve as candidates' declaration. However, the president is not restricted to his announcement. That means the president is not bound to his announcement. After reading candidates' declaration, citizens will elect a president from the 3 candidates for the following 5 rounds of this unit. Each citizen has 1 vote. According to the majority principle, the candidate who receives at least 3 votes will be elected as president. If two candidates both receive 2 votes and the third candidate gets 1 vote, the third candidate is eliminated. The 2 candidates who get 2 votes have to re-announce their declaration. Citizens will vote again to elect a president. The eliminated candidates see the declaration of the president and wait for the start of the next unit.

After a unit is finished, all 3 candidates will again propose a back-transfer suggestion as their declaration. Citizens will see the new declaration from all candidates and vote one of them to be the president for the next unit. Attention: in each unit, the candidates' codes remain the same. That is to say, citizens do not know which candidate has been elected as president in previous units, and which candidate has not been elected as president in previous units.

Steps of experiment in detail

1. At the beginning of the experiment, each participant will be randomly assigned to an 8-member-group. The three candidates receive their codes.

2. The following only happens at rounds 1, 6, 11, 16 and 21:

    a. The 3 candidates propose the back-transfer suggestions, as if they were elected to be the president. The back-transfer suggestion is how much he will back-transfer given citizens' transfer (0, 1, 2, 3, 4, 5, 6 ,7 ,8 ,9 ,10) after he receives from citizens (0, 2, 4, 6, 8, 10, 12, 14, 16, 18, 20). These suggestions will be served as candidates' declaration.

    b. All citizens will read the declaration from all candidates and elect one of the candidates to be president for the following 5 rounds of the current unit.

    c. If 2 candidates both receive 2 votes and the third candidate gets 1 vote, the 2 candidates who get 2 votes will have to re-announce their declaration. Citizens will vote again and elect 1 of these candidates to be president.

    d. The eliminated candidates see the president's declaration. After one minute the program automatically enters a waiting screen for the start of next unit.

    e. At the end of each unit, the 3 candidates make a new declaration for the next unit. Back to step 2a).

3. In every decision round:

    a. Citizens make a decision how many points to send to the president. Citizens can send any integer they wish to the president from the set {0; 1; 2; 3; 4; 5; 6; 7; 8; 9; 10}.

    b. President sees the amount of points he gets from the 5 citizens, without knowing exactly who sent how much. The president will then decide upon the amount of points to be back-transferred to each citizen. The president can back-transfer any integer from the set [0; 2 x sent amount] to citizens. The amount of points the president back-transfers to citizens will not be doubled.

c. Every participant receives information about his own payoff from the current round.

4. (Only after round 25). Every participant receives information about her/his total payoff and receives payment.

How to calculate your income from the experiment?

Your total income from the experiment is the sum of income you receive from each decision round. The income of each round is calculated as follows:

Income of citizens in each round = Citizen's initial endowment (10 points) – the points sent to president + the points back-transferred by president

Income of candidates who are not elected as president in each round = candidate's initial endowment (10 points)

Income of president from each round = President's initial endowment (10 points) + the points sent by the 5 citizens x 2 – the points back-transferred to the 5 citizens

2.8.3 Instructions for treatment NoPolComp

Now the second experiment starts. Please read the experimental instructions carefully. You will be again paid cash for participating in the experiment. The money you earn will first be represented in points. The amount of points you earn is depending on your and other participants' decisions in the experiment.

You will receive RMB from the points you earned by the following exchange rate:

10 points = 2 RMB

In this experiment, you will interact with other 7 participants. At any time, you will not know who they are. Similarly, other participants will not know who you are.

Experimental instruction of the second experiment

There are 24 participants in the experiment. At the beginning, you will be randomly assigned to a group of 8 members. Then you will be assigned a role in the experiment. All participants will be assigned either of the two types: 5 will be citizens and 3 will be candidates. Your current type depends on your type in the first part of the experiment. That is to say, if you were type A in the first experiment, you are citizen now; if you were type B in the first experiment, you are candidate now. Every candidate will receive a unique code. Attention: the composition of each 8-member group, the role of every participant and the code of each candidate remain unchanged throughout the experiment.

There will be 25 decision rounds in total, with every 5 rounds as a unit. The decision environments of each unit are identical, which means that the experiment process of the first unit (rounds 1-5) is the same as that of the second unit (rounds 6-10); of the third unit (rounds 11-15); of the fourth unit (rounds 16-20) and of the fifth unit (rounds 21-25).

At the beginning of each round, all group members receive 10 points as initial endowment. The 5 citizens will decide on how many points to send. At the beginning of each unit (round 1, 6, 11, 16 and 21), one of the 3 candidates will be randomly selected as the president for the following 5 rounds of each unit (described below). In each round of the unit, all citizens can decide to send any integer point to the president from their initial 10 points. That means,

citizens can send any integer they wish to the president from the set {0; 1; 2; 3; 4; 5; 6; 7; 8; 9; 10}.

The president will receive double points as sent by the 5 citizens. That means, each amount of points sent by citizens will be doubled and sent to the president. Thus, the president will get twice the amount of points sent by citizens, instead of the same amount. The amount of points that citizens sent to the president will be shown to the president in a descending order. That means, the president will see how many points he gets from the citizens, but not which citizen sent him how much. The president must decide, for each amount of points, how much to back-transfer. The president can back-transfer any integer from the doubled points he received. That means, the president can back-transfer any integer from the interval [0; 2 x sent amount], to citizens. The amount of points the president back-transfers to citizens will not be doubled.

How to select the president?

At the beginning of each unit (1^{st}, 6^{th}, 11^{th}, 16^{th} and 21^{st} round), the 3 candidates can announce a back-transfer suggestion, as if they were selected to be president. The back-transfer suggestion is how much each candidate, as president, will back-transfer given citizens' transfers (0, 1, 2, 3, 4, 5 , 6 ,7 ,8 ,9 ,10) after he receives points sent from citizens (0, 2, 4, 6, 8, 10, 12, 14, 16, 18, 20). These suggestions serve as candidates' declaration. However, the president is not restricted to his announcement. That means the president is not bound to his announcement. One of the three candidates will be randomly selected as president. All citizens will see the declarations of all candidates and know which candidate has been randomly selected as president. The candidates who are not randomly selected as president see the declaration of the president and wait for the start of the next unit.

After a unit is finished, all 3 candidates will again propose a back-transfer suggestion as their declaration. Computer will randomly select one candidate to be the president of next unit. Citizens will see the new declaration from all candidates and know which one of them to be the president for the next unit. Attention: in each unit, the candidates' codes remain the same. That is to say, citizens know which candidate has been randomly selected as president in previous units, and which candidate has not been randomly selected as president in previous units.

Steps of experiment in detail

1. At the beginning of the experiment, each participant will be randomly assigned to an 8-member-group. The three candidates receive their codes.

2. The following only happens at rounds 1, 6, 11, 16 and 21:

    a. The 3 candidates propose the back-transfer suggestions, as if they were elected to be the president. The back-transfer suggestion is how much he will back-transfer given citizens' transfer (0, 1, 2, 3, 4, 5 ,6 ,7 ,8 ,9 ,10) after he receives from citizens (0, 2, 4, 6, 8, 10, 12, 14, 16, 18, 20). These suggestions will be served as candidates' declaration.

    b. One candidate is randomly selected as president.

    c. All citizens will read the declaration from all candidates and know which candidate to be president for the following 5 rounds of the current unit.

    d. The candidates who are not randomly selected as president see the president's declaration. After one minute the program automatically enters a waiting screen for the start of next unit.

    e. At the end of each unit, the 3 candidates make a new declaration for the next unit. Back to step 2a).

3. In every decision round:

    a. Citizens make a decision how many points to send to the president. Citizens can send any integer they wish to the president from the set {0; 1; 2; 3; 4; 5; 6; 7; 8; 9; 10}.

    b. President sees the amount of points he gets from the 5 citizens, without knowing exactly who sent how much. The president will then decide upon the amount of points to be back-transferred to each citizen. The president can back-transfer any integer from the set [0; 2 x sent amount] to citizens. The amount of

points the president back-transfers to citizens will not be doubled.

c. Every participant receives information about his own payoff from the current round.

4. (Only after round 25). Every participant receives information about her/his total payoff and receives payment.

How to calculate your income from the experiment?

Your total income from the experiment is the sum of income you receive from each decision round. The income of each round is calculated as follows:

Income of citizens in each round = Citizen's initial endowment (10 points) – the points sent to president + the points back-transferred by president

Income of candidates who are not randomly selected as president in each round = Candidate's initial endowment (10 points)

Income of president from each round = President's initial endowment (10 points) + the points sent by the 5 citizens x 2 – the points back-transferred to the 5 citizens

3 Does voting moderate power and foster trust? A behavioural view on democracy[1]

3.1 Introduction

Power often comes with conflicts of interest between the holders of power and their electorates: public leaders may embezzle funds; executives of companies may opt for short-term window-dressing instead of long-term profit maximisation; both may spend money on pet-projects instead on good-value-for-money investments. A widely-used mechanism of collective choice to allocate power and determine hierarchy is voting. Given the widespread use of elections, do formal voting procedures themselves make a difference to the exercise of power in a situation of conflicting interests? Do those at the receiving end of power trust the holders of power more if they can elect them? In case voting matters for trust and power, do effects wear off over time?

We experimentally implement our research questions by using a modified investment game repeated over five elections units, in which reliable reputation building across units is ruled out. Investors can trust by sending investments to a responder who is either elected by the investors or randomly determined and has power to seize post-investment assets. The higher is the amount that responders give back to the investors, the less do they exercise their power in their own favour. The higher are investments, the higher is the trust investors put into the holder of power. Measuring player-types prior to the main treatments improves our statistical power by allowing us to control for a potentially crucial covariate.

Our research questions have so far received only little academic attention despite considerable amount of evidence in social psychology and experimental economics that participatory procedures are behaviourally relevant. The main focus of the economics literature is the potentially disciplinary role that elections have on incumbent leaders by threatening dismissal from office (see e.g. Bardhan and Yang 2004; for experimental

[1] Based on: "Does Voting Moderate Power and Foster Trust? A Behavioural View on Democracy" by Gari Walkowitz and Arne Weiß (2009), Working Paper, University of Erfurt. All authors contributed equally.

evidence see Weiß 2009, chapter 2 of this thesis).[2] A second, albeit less developed, strand of literature analyses how elections may serve to select holders of power. Voters may want to give power to those whose preferences are supposedly most in line with their own (for a related model see Maskin and Tirole 2004), who is up to the task in terms of her ability (Carrillo and Mariotti 2001) or who may even feel a public service motivation (Cooter 2003); this motivation is intrinsic in nature (see e.g. Frey 1997).[3] This approach already departs considerably from the assumption of homogenous candidates who are solely motivated by pecuniary or non-pecuniary benefits of power; yet, this approach still considers only the outcomes of elections to matter: winning candidates may either be 'good' or 'bad', but are otherwise unaffected by the procedure through which they assume power.

This, however, may not be the whole story. Possibly, being elected by her constituency triggers the leader to feel a sense of commitment or reciprocity to act in her voters' interests, leading to power being exercised in a less self-oriented way. Voting may also foster trust as participatory procedures have been shown to turn people more confident about personal success in uncertain situations (Langer 1975). Indeed, a vast amount of evidence on procedural justice in social psychology shows procedures to matter. However, the literature almost exclusively concentrates on the procedural determinants of subordinates' willingness to abide by the decisions of an authority (see e.g. Thibaut and Walker 1975; Tyler 1989; Tyler and Lind 1992; Tyler and Blader 2003; Tyler 2006), rather than on either trust or on the effects of procedures for the behaviour of those who hold power. In psychological economics and related fields in political philosophy, the utility consequences of participatory procedures are investigated. According to research on procedural utility (e.g. Frey, Benz and Stutzer 2004), which draws on earlier work by Sen (1995), people may have a preference for participation independent of outcomes and yield substantially higher life satisfaction from participatory rights.

A growing strand of literature analyses the effects of voting procedures on cooperation. The largest body of evidence points to ballot voting (Ostrom, Walker and Gardner 1992; Maier-

[2] All subsequent references to Weiß (2009) also refer to chapter 2 of this thesis.
[3] Related to the idea of being motivated to serve the public is also the concept of civic duty (for a discussion and a related model see Besley 2005).

Rigaud and Apesteguia 2003; Tyran and Feld 2006; Dal Bó, Foster and Putterman 2008) or foot voting (Gürerk, Irlenbusch and Rockenbach 2006) enhancing cooperation within groups. It has also been shown that merely choosing one´s transaction partner can mitigate the moral hazard problem: agents behave more reciprocally if they have accepted an offer by a principal that was only targeted at them instead of an offer that was open to all agents in a group (Brown, Falk and Fehr 2004). To our best knowledge, no experimental evidence exists on the effects of voting on trust put into holders of power and on the long-term effects of voting on the exercise of power.

Some evidence exists on the joint role of elections and approval rates for the behaviour of holders of powers in the short-run: Corazzini, Kube, Maréchal (2007) find elected allocators to send more to the recipients than randomly drawn allocators if their approval rates are higher than what is minimally required to win the election. In contrast to us, they do not study the long-term effects of voting on power and only analyse the effects of voting contingent on approval rates.[4] Furthermore, their voting procedure triggers considerably higher promises compared to their control treatment, in which the computer votes; promises are more than cheap-talk for dictators who have a preference for either promise keeping (Vanberg 2008) or consistency (Ellingsen and Johannesson 2004) or who exhibit guilt-aversion and whose second-order beliefs rise with the promises they give. The less self-oriented behaviour of elected allocators in Corazzini, Kube and Marechál (2007) compared to their randomly drawn counterparts may therefore be explained by a joint effect of the procedure of voting and the higher promises induced by competition for votes. In our design, responders have an incentive to strategically choose their promises also in the control treatment so that considerable treatment-differences in promises may not arise.

The ideal dataset to study the effect of voting on the exercise of power and the trust put into holders of power would entail exogenous changes of the mechanism to select the holder of power keeping everything else constant. For a number of reasons, this dataset is almost impossible to come by outside controlled experimental conditions. First, in field settings, the selection function of elections may not be controlled for. It is virtually impossible to measure

[4] Corazzini, Kube, Maréchal (2007) repeated-play treatment only consists of two elections (and two rounds of allocation decisions) in total, which seems insufficient to test for long-run effects.

whether voters may have been successful in electing candidates that are, intrinsically, less corrupted by power. Second, in real-life hierarchies, further attributes of the candidates, such as her ability, may also be important in determining the trust put into her. Third, the length of a player´s horizon, which crucially influences his strategic incentives, is difficult to control for and may well depend on whether a voting mechanism is employed (for a discussion see Weiß 2009). Fourth, whether voting is used in the field may depend on the behaviour or the personal types of the holders of power. Endogeneity problems may therefore arise that are difficult to control for. With a properly designed experiment, we can not only control for differences in environments and socio economic backgrounds that are unrelated to voting but also test whether a voting procedure can causally lead to short- and long-run differences in trust and in the exertion of power.

We find elected power holders to use their power in a less self-oriented way. This long-run effect is robust to controlling for player-types, promised back-transfers and the level of received investments. While we do find some evidence of promise-keeping in both treatments, elected power holders do not keep their word better than randomly drawn power holders. We find investors to display higher trust if they can elect the holder of power, independently from any payoff-relevant decisions by the power holders. This effect cannot be explained by treatment-differences in the level or in the role of promises. Investments stay higher in the voting treatment for the first 10 rounds. This effect is driven by differences in profitability of investing. Hence, an effect of the procedure of voting itself that is independent of the behaviour of power holders can only be observed in the short-run.

The paper is organised as follows: Section 3.2 introduces the experimental implementation of the research questions. Hypotheses are derived in section 3.3. Results are presented in sections 3.4 to 3.6. Section 3.7 concludes and provides an outlook for future research.

3.2 Experimental implementation

3.2.1 Experimental design

In order to answer our research questions our experimental design is based on a game that both entails hierarchy and power and is able to measure trust put into holders of power. In the investment game as introduced by Berg, Dickhaut, and McCabe (1995), a sender and an

anonymous responder are randomly paired and endowed with a fixed amount *X* each. The sender can send any non-negative part *a* of his endowment to the responder. The responder pockets her endowment. While being sent, *a* is multiplied by *k* >1 by the experimenter. The responder then decides to back-transfer any non-negative amount *b* of the received amount *k*a* to the sender. The responder therefore has full discretionary power to decide on the post-investment resource-allocation and is always able to at least receive the same payoff as the sender. The responder effectively plays a dictator game, the pie being the received amount *k*a*. The amount *a* chosen by the sender is interpreted as sender's degree of trust towards the responder. The fraction *b*/(*k*a*) that is transferred back by the responder delivers a measure of reciprocity towards the sender. Within the context of our research question, the responder´s reciprocity negatively corresponds to the degree that the responder exploits her power. The higher the responder's reciprocity, the higher is her short-term loss of profit, and hence, the less does she exercise her power in her own favour.

In our experimental treatment (VOT), participants are randomly selected into anonymous groups of eight players consisting of five senders (S) and three responder candidates (C). Senders and candidates are randomly determined. All players initially receive an endowment of 10 points. To introduce voting into the basic investment framework, we implement an initial election stage before the group investment game starts. First, all candidates have to announce a non-binding back-transfer strategy, that is the vector *b* they would transfer back for any possible amount received from an individual investor in case of their election. Candidates are assigned a unique identification code so that each candidate can be unambiguously linked to a promised back-transfer strategy. This code, which consists of a 3 randomly generated letters, is valid for one unit (i.e. five decision rounds). Based on this information, the five senders are asked to elect one responder (R) out of the three candidates by majority rule.[5] Through the election procedure, one candidate is determined as responder. Unelected candidates do not participate in the investment game that follows.[6] Finally, the five senders receive information about the outcome of the election; the corresponding identification code of the winning candidate is shown to the senders. In

[5] In case of a tie between two candidates, a second voting takes place again by majority voting.
[6] They will neither act as senders nor as responders. The reason for this strict separation is to avoid noise in the data caused by mixed roles.

addition, the back-transfer strategy of the elected responder is shown to the non-chosen candidates in order to set a focal point on the winning back-transfer strategy. In a next step, all senders (S_i) individually choose an investment level a_i ϵ {0,1,...,10}. Investments are doubled (k=2) by the experimenter and transferred to the elected responder. The responder then separately decides for every investment a_i the amount b_i ϵ {0;2∗a_i} to be transferred back. Responders receive no information on the identity of the senders that correspond to the invested amounts. In the back-transfer decision the responder is not bound to the initially announced back-transfer strategy.

The payoff of a sender is given by $\pi_{S_i} = 10 - a_i + b_i$. Accordingly, a responder's payoff is calculated by $\pi_{S_i} = 10 + 2 * \sum_{i=1}^{5} a_i - \sum_{i=1}^{5} b_i$. The two unelected candidates keep their endowment of 10 points.

The control treatment (RAN) is identical for the only difference that responders are randomly drawn with equal probability of 1/3 out of the three candidates. Figure 3-1 graphically shows the basic experimental design.

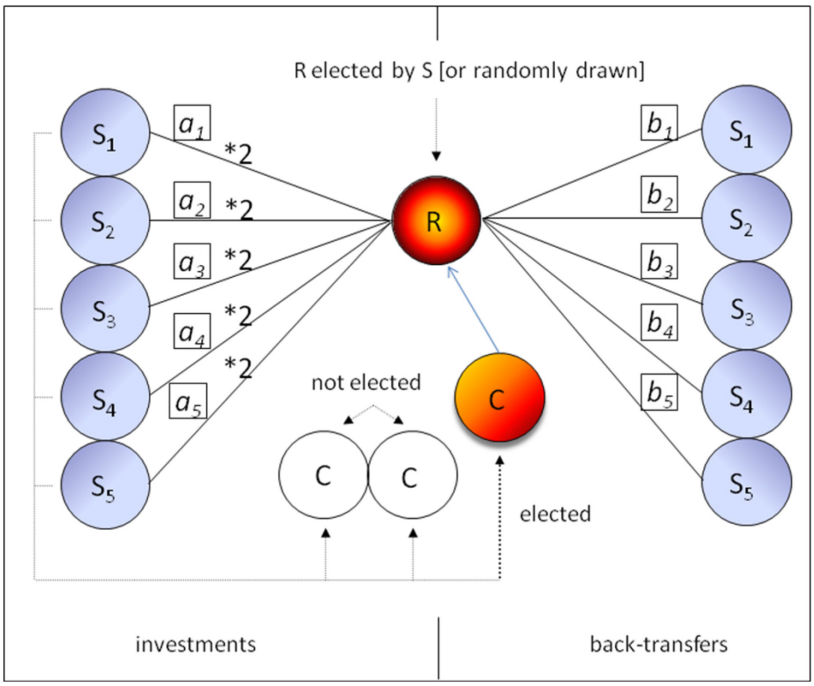


Figure 3-1: Experimental design

In order to analyse the long-term effects of voting we use a repeated game that consists of 25 rounds in total. Roles and groups are kept constant. Every five rounds a new responder is chosen (see Figure 3-2). Prior to any responder selection, candidates receive new identification codes. Furthermore, the position of candidates' promises on the computer screen is shuffled for each responder selection. In order to set a strong focal point on the promises of the chosen responder after each responder selection, the promises of the chosen responder are shown to the non-chosen candidates. These three measures render reliable identification across elections impossible. The exact voting outcome, i.e. how many votes each candidate receives, is not displayed in order to focus on the procedure of voting, rather than the role of approval rates.[7]

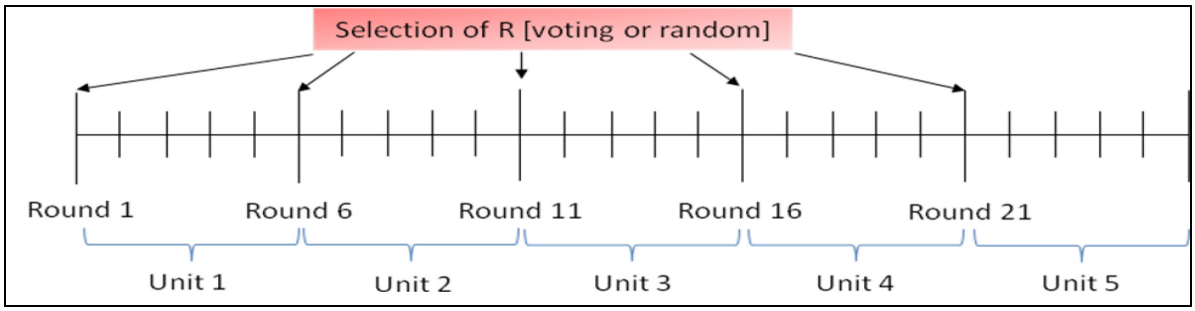


Figure 3-2: Experimental sequence

We ran the experiments in Chengdu, People's Republic of China. Chinese people have very little experience with voting mechanism as elections only take place in some rural areas. China also scores low on the "Voice and Accountability" indicator of the World Bank Governance indicators (Kaufmann, Kraay and Mastruzzi 2008).[8] We consider the lack of experience with voting procedures an advantage in order to test our hypotheses. Chinese subjects should have less pre-understanding of how an elected holder of power ought or does behave than subjects in countries with more frequent elections. In this way, experimental results may be less confounded by subjects' personal experiences with voting procedures and the behaviour of elected power holders.

[7] For the role of approval rates in a voting setting, the reader is kindly referred to Corazzini et al. (2007).
[8] With a percentile rank of 5.8, China found itself in the lowest "Voice and Accountability" bracket in 2007.

Instructions onscreen and on paper presented the experiment in rather neutral terms that emphasise the hierarchical relation between the responder and the investors.[9] Investors are called "citizens" and the responder "president". The decision by investors is framed as "investing". Allocators´ back-transfer decision is just framed as "sending back". The candidates are called "candidates".[10] The term "citizen" has a political connotation; we nevertheless refrained from explicitly putting subjects into a political context as Chinese subjects may not feel at ease in a political situation in which they vote.

3.2.2 Measuring player-types through prior one-shot game

In order to measure trust and reciprocity dispositions of later senders and candidates (and responders once candidates have been selected) and to gain information about different player types, we conducted a classical one-shot investment game (henceforth OS) before the experimental treatments RAN and VOT were launched. In this game, we applied the same strategy space for investors (X=10), an identical multiplier k=2, and the strategy vector method (Selten 1967) for responders. Thus, responders had to decide for every possible amount sent by the investor how much to back-transfer. Our experimental design is aimed at preventing the election to have a selective function, i.e. to select "good" instead of "bad" candidates. By measuring player-types, we can test whether our design is indeed successful in this regard. Furthermore, we can more precisely measure the effect of voting on trust and reciprocity by controlling for their initial propensity to trust and to reciprocate, which is potentially an important covariate in determining players' behaviour in VOT and RAN.

Roles between the two experiments were kept constant to compare choices across treatments but were framed differently. Senders were called player *Type A* and candidates player *Type B*. In this way, we measure a baseline trust level for later senders and a baseline-strategy of all later responders for all possible investments that they might receive in VOT and RAN. Subjects received no feedback on the OS-choices of the matched interaction partner; thereby, subjects could not update beliefs and made a fresh start in the repeated game.

[9] Please refer to Appendix for instructions

[10] Chinese terms we employed were: president: 主席 (zhǔxí), citizen: 公民 (gōngmín), candidate: 候选人 (hòuxuǎnrén)

3.2.3 Experimental procedure

We collected 9 independent observations for both treatments. In total 144 subjects participated in the computerised laboratory-experiment run in October 2007 at the Herbert A. Simon & Reinhard Selten Behavioral Decision Research Lab, Southwest Jiaotong University, Chengdu, China. The experiment was programmed in ASP.NET.[11] Separate cubicles with view-proof curtains made sure that interaction between the subjects was completely anonymous. It was common knowledge that the experiment only started after all subjects correctly answered test-question about the experimental set-up, including how earnings are calculated. Subjects´ earnings were set at local standards. Sessions lasted on average 3.25 hours, in which subjects earned an average of 91 RMB (about USD 12 at the time of the experiment), paid out in cash individually at the end of the experiment.[12]

3.3 Hypotheses

3.3.1 Game-theoretic solution based on common knowledge of rationality

Based on the commonly employed assumption of common knowledge of rationality, it is straightforward to derive the subgame-perfect Nash-equilibrium of our game. In the last round (t = 25) responders will play their money-maximizing action, which is to send back zero regardless of the amount received ($b_{i,t}$ = 0, i = 1,...,5 and t = 25 for $a_i \in$ {0,1,...,10}). Since the finite end of the experiment after 25 rounds is common knowledge, senders anticipate this and will consequently not send anything in the first place ($a_{i,\ t=25}$ = 0). Therefore no rent is available in the last round that may induce responders to cooperate in earlier rounds. Responders will therefore not send anything back regardless of the transfers they receive and senders will not to send anything in the first place ($a_{i,t}$ = 0, i = 1,...,5 and t = 1,...,25). The unique subgame-perfect Nash-equilibrium path for investments and back-transfers in both treatments is zero:[13]

[11] For instructions please refer to the appendix.

[12] Besides the actual experimental play, the duration includes giving comprehensive instructions; clarifying open questions; subjects answering test questions; the one-shot investment game; a risk measure experiment and pen & paper hand-written survey questions after the experiment as well as individual payments. Further data is available upon request.

[13] The backward-induction argument runs the following: With common knowledge of rationality, i.e. everyone knowing that everyone is rational and money-maximising, any responder will unambiguously confiscate all available assets in the last round. Senders anticipate this and will therefore not send anything in the last round. Responders therefore have no incentive to send anything back in the penultimate round which rational senders again foresee and so forth. The unique subgame-perfect Nash-equilibrium path therefore is zero investments throughout the experiment. The uncertainty of

$b_{i,t} = 0$, $i = 1,...,5$ and $t = 1,...,25$ for $a_i \in \{0,1,...,10\}$

$a_{i,t} = 0$, $i = 1,...,5$ and $t = 1,...,25$

Hence, based on common knowledge of money-maximising rationality, voting is predicted not to influence investments and back-transfers and thereby not to matter for trust and the exercise of power.

3.3.2 Behavioural hypotheses

If we relax the assumptions of money-maximising rationality, a broader set of behavioural predictions can be made. By electing responders**,** voters transfer at the same time considerable power and an economic rent to responders**;** responders can never fall below the earning of an unelected candidate and will only earn less than senders if they voluntarily send back more than the equal split. The elected responders may consider the transfer of power as an act of trust and the implied economic rent as a gift. Previous research has shown that distrust can be self-fulfilling in that it lowers trustworthiness (Fehr and Rockenbach 2003; Falk and Kosfeld 2006). By the same idea, trust may be considered a kind act to be positively reciprocated (Falk and Fischbacher 2006). Gift-exchange, i.e. the reciprocation of a material gift with another gift, has been found in many experimental studies (e.g. Fehr, Kirchsteiger and Riedl 1993; Fehr, Kirchsteiger and Riedl 1998; van der Heijden et al. 2001; Charness 2004), that may be explained by inequity aversion (Fehr and Schmidt 1999; Bolton and Ockenfels 2000) or by reciprocity utility (Falk and Fischbacher 2006). In RAN, by contrast, responders have no reason to consider their selection as a gift from their citizens. Hence, we expect responders to behave in a less self-oriented way if they have been elected rather than randomly drawn. Translated into our design, we expect responders' reciprocity to be enhanced by the voting procedure. We consequently predict

incumbent responders over whether they will hold onto their power does not change the rationality prediction. As long there is a definite known end to the game, backward induction along the lines above can be used as no responder would send anything back in the last round.

H1: For given investments elected responders in VOT will send back higher amounts b than selected responders in RAN. Hence, returns on investment are predicted to be higher in VOT than in RAN.

Technically, the promises of responders are cheap-talk. Nevertheless, they may matter to senders, especially prior to any belief updating on the elected responder's back-transfer behaviour. When senders take their voting decision and invest in the first round of any unit, they do not have any specific knowledge about any of the three candidates (as new identification codes are allocated at any new unit). Senders may also believe that promises matter to responders. Indeed, recent experimental evidence shows that individuals may have a preference for promise keeping (Vanberg 2008) or consistency (Ellingsen and Johannesson 2004). Nevertheless, rather than naively trusting any promise, senders may evaluate the credibility of candidates' promises and may also view current trustworthiness, i.e. how closely responders stick to their promises, as a signal of their future trustworthiness. Hence, candidates may carefully choose the promises they make. In both treatments, responders are interested in triggering high investments; this should, at least before senders have made experience with actual back-transfers behaviour of responders, require a promised rate of return that is both positive and credible. In VOT, candidates have to additionally think of how to beat the other two candidates. While senders will certainly value high back-transfer, it is nevertheless unclear whether this competition leads to higher promises in VOT compared to RAN. Promises that are higher than those promises that trigger high investments in RAN may not be deemed to be credible anymore. Large differences in promised returns on investment (henceforth RoI) therefore seem implausible. Hence, the competition that candidates face may or may not lead to higher promises in VOT than in RAN, but not to lower promises than in RAN.

H2: Promised RoI are not lower in VOT than in RAN.

If promises are not higher in VOT, we would, based on H1, expect responders to be more trustworthy in VOT than in RAN. Furthermore, promises may carry greater weight if responders have been elected based on these promises. If, however, promises are higher in VOT, treatment effects on how well responders stick to their promises are unclear. On the

one hand, higher RoI would mean less cheating by responders in VOT. On the other hand, higher promised RoI would mean more cheating in VOT. Therefore, no prediction can be derived on treatment effects on the trustworthiness of responders.

In order to derive hypotheses on the behaviour of senders**,** we may first think about their behaviour prior to belief updating, i.e. in the first round of the experiment. First, based on the group engagement model (Tyler and Blader 2003) in social psychology, cooperative behaviour in groups is strengthened when procedures lead to more engagement in the group. As the voting procedure in VOT engages senders more than the random draw in RAN, senders may be more cooperative in VOT than in RAN. This corroborates with findings of cooperation games in which participatory procedures lead to more cooperative play (see introduction). As joint payoffs increase in investments *a*, investing may be perceived as cooperating from sender's perspective. Second, the engagement of voters in VOT may lead to sense of agency (Sen 1995) and control, fostering trust. Research in social psychology on the "illusion of control" (Langer 1975) has shown that participatory procedures may turn people more confident about personal success in uncertain situations, even when the objective probabilities of success have not changed. Third, senders may expect, in line with H1, a higher profitability of investing in VOT than in RAN. If senders expect negative returns on investment, only those who are either risk-loving, altruistic or have a preference for efficiency would invest. Whereas risk-neutral and money-maximising investors would send their full endowment if they expect a positive return on investment, risk or inequality-averse investors would increase their investments with positive expected RoI. Hence, if senders expect a positive and higher return on investments in VOT than in RAN, they may be willing to invest more. Therefore, we predict

H3.1: First round investments are higher in VOT than in RAN.

After the first round, senders can update their beliefs on the behaviour of their responders based on past back-transfers. Over the course of the experiment, we expect actual behaviour of responders to be become salient in determining beliefs of senders. Based on H1, we expect RoI to be higher in VOT than in RAN. We therefore expect investments to be consistently higher in VOT than in RAN.

H3.2: As long as returns on investment are not lower in VOT than in RAN, investments are higher in VOT than in RAN.

Combining H1 and H3.2, we also predict

H4: Senders' profits are higher in VOT than in RAN.

Deriving a hypothesis for responders' profits is not so straightforward, however. On the one hand, we predict higher investments in VOT than in RAN, which, for any given RoI < 1 would imply higher profits of responders. On the other hand, we predict higher RoI in VOT than in RAN. Nevertheless, the basis of H1 assumes responders in VOT to exercise their power in a less self-oriented way; by contrast, we do not predict responders in VOT to be better in maximising profits. If higher RoI in VOT compared to RAN would imply higher profits for responders, there is little reason why RoI should not be raised in either treatment; raising RoI is a Pareto-improvement from which both responders that only care for their own profits and those who have pro-social preferences would profit.[14] Consequently, we predict

H5: responders' profits are lower in VOT than in RAN.

3.4 Results I: Does voting moderate power?

We start with looking at the behaviour of elected responders. We find the average returns on investment (RoI)[15] per group – as our measure for the responders' reciprocity – to be weakly significantly higher in VOT (34.88%) compared to RAN (16.63%) (p=.084, Fisher-Pitman Permutation Test for two independent samples,[16] henceforth FPPI). The same holds for the accumulated back-transfer volume over all 25 rounds: back transfers in VOT are, on average, 37.26 % higher than in RAN (p=0.129, FPPI). When we separately analyse the first unit of play – consisting of the first five decision rounds where no re-election has occurred so far – we find that elected responders in VOT reward trust higher (35.9%) compared to randomly chosen responders in RAN (16.8%). This difference, however, does not reach

[14] One caveat to this line of argumentation, however: responders may not expect raising RoI to be profitable or may be risk averse (i.e. value certain payoffs in the current round higher than higher, but uncertain, profits in the next round). In this case, it is still rational for responders to set below profit-maximising RoI in RAN. Ironically, responders in VOT would then profit, on average, from the effect of voting on power that we hypothesise.

[15] The RoI is given by RoI = (b_i/a_i)-1. For ease of reading, we drop the term "relative" in the remaining text.

[16] We apply this test because it does not require a distribution assumption and its asymptotic efficiency equals 1. All statistical tests without further reference are carried out two-sided.

statistical significance (p=0.176, FPPI). If we calculate the moving average RoI over all decision rounds, we find a clearly pronounced difference between the two treatments and less variance in VOT compared to RAN (see Figure 3-3).

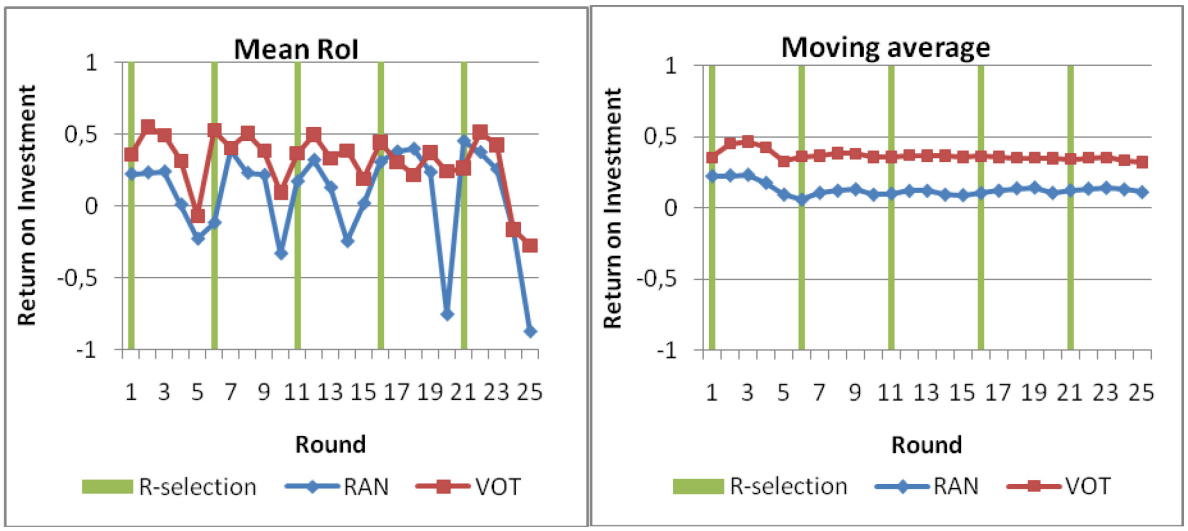


Figure 3-3: Treatment comparison of mean and moving average of return on investment over time.

In the last round of any unit, there are no obvious strategic reasons to transfer back positive amounts as the game between senders and the selected responders ends with probability of 2/3 due to the allocation of new identification codes at the subsequent round. When we investigate responders' behaviour *before* selection in each unit and in the very last round of the game, hence, in rounds 5, 10, 15, 20, and 25, we find end-game RoIs to be higher and still non-negative in VOT (0.02) compared to RAN (-0.46) (p=0.088, FPPI, see also Figure 3-3). Our findings on RoI support our first hypothesis.

> *Observation 1: Returns on investments are higher in VOT compared to RAN. Whereas significant treatment differences cannot be detected in the very first round, RoI in units' last rounds are significantly higher in VOT.*

3.4.1 What drives higher RoI?

In the next step of our investigation we systematically investigate drivers behind the observed treatment effects on power holders' back-transfer behaviour in order to analyse whether the procedure of voting itself matters.

3.4.1.1 Population and selection effects

By controlling for players' dispositions, we are able to measure the treatment effect more precisely, as a deviation from players' behaviour exposed in OS. In a first step, we will compare OS-behaviour of responders across both treatments. We then control whether our design indeed inhibits selection of more reciprocal types through a successful signalling of subsequent back-transfer behaviour in voting. We will therefore test for differences in reciprocity levels of players selected in VOT compared to randomly determined responders in RAN and compared to those candidates who were not chosen in VOT.

VOT-players on average generated an OS-RoI of -22.8%, which is very similar to the OS-RoI of -20.5% in RAN (p=0.884, FPPI). Moreover, elected responders in VOT (-31.3%) have a somewhat higher baseline reciprocity-level than selected responders in RAN (-16.4%), albeit not significantly so (p=0.409, FPPI). As the baseline-reciprocity of elected responders in VOT (-31.3%) is lower than the average baseline-reciprocity in VOT (-22.8%), selected responders in VOT were not more reciprocal in OS compared to unelected candidates in VOT, rather the opposite. Consequently, our design prevented reliable signalling of candidate types.

Based on their OS-behaviour, we find that the treatment effect even *overcompensated* for underlying average OS-differences: If we take into account to what extent responders deviated from their baseline strategy as revealed in the one-shot game, the treatment difference increases considerably.[17] Responders in VOT deviate on average more (62.0%) from their OS RoI-level compared to RAN (28.3%) (p=0.018, FPPI, see also Figure 3-4, which shows responders' average deviation from their OS-strategy over time). Consequently, voting in fact fosters player's reciprocity.

[17] The deviation is calculated as the average difference between the RoI provided by responders in the repeated game and the RoI that the respective responder would have provided in OS, given his OS-strategy and the amounts invested by senders in the repeated game. An example might clarify the calculation: assume that a = 5, b = 10 and the back-transfer based on the OS-strategy of the selected responder for an investment of 5 is 5, then the difference in RoI is 2 – 1 = 1.

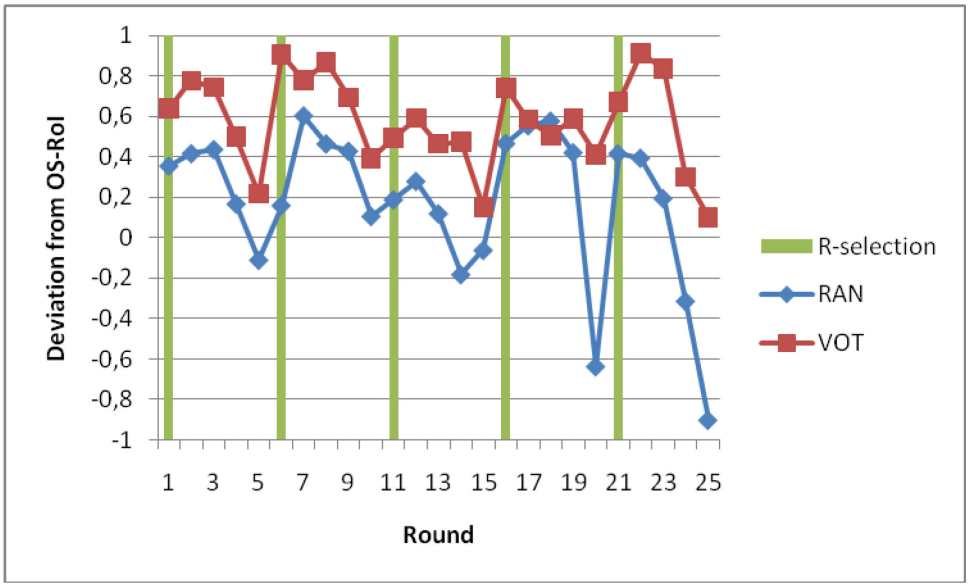


Figure 3-4: Treatment comparison of mean deviation of responders from their respective OS-strategy expressed in RoI.

Observation 2: Higher RoI in VOT are not due to a population effect. Players deviate (positively) much more from their OS strategy than in RAN. There is no selection effect.

3.4.1.2 Differences in promises and commitment toward them?

Next we investigate whether promised RoI constitute one source for the observed effects in RoI. That is, we first investigate whether elected responders in VOT promise to give back more compared to RAN**.** Then we try to answer the question whether promised RoI and actual RoI are correlated, i.e. whether, and if so, how strongly responders feel obliged to act in line with their prior statements. This procedure should reveal whether promises and related actions might partly explain our treatment differences.

For VOT (66.1%) we find that responders promised on average a higher RoI compared to RAN (52.1%).[18] This difference is not statistically significant, however (p=0.244, FPPI).[19] Only in RAN**,** RoI are weakly correlated to promises (Spearman's ρ = 0.346, p=0.088).[20] In VOT the

[18] This refers to the promises for those investments actually undertaken by senders in VOT.
[19] Nevertheless, variance in promises is substantially lower in VOT (11.7) than in RAN (32.3) (p=0.02, Exact Sample Permutation Test for Differences in Variances).
[20] The correlations are calculated for each elected responder's average promise and associated average RoI.

correlation between promises and RoI is close to zero (Spearman's ρ = 0.090, p=0.683).[21] Comparing responders' deviation in actual RoI from promised RoI, we find no significant difference between RAN (0.42) and VOT (0.33) (p=0.426, FPPI). Consequently, our treatment effects are neither visually driven by differences in given promises nor by clear differences in the commitment toward them. As a further result, responders in VOT are, in the promise-keeping sense, not more trustworthy than responders in RAN.

> *Observation 3: Promises in VOT and RAN do not differ significantly. In VOT power holders show no (higher) commitment to promises.*

3.4.1.3 More rewarded trust or a reward of more trust?

Do responders provide higher RoI in VOT because of differences in investment-levels between the treatments (which we will analyse in more depth in section 3.5) or because responders reward investments in general more. In other words: responders may be moving from RAN to VOT along the basically same function of back-transfers dependent on investments, or the back-transfer function itself may change. We first look at RoI in the most prominent case when responders receive an investment of 10 points, which represents the mode investment in both treatments (see Figure 3-5) and, simultaneously, full investment. We find that in VOT (32.70%) power holders reward higher investments significantly more than in RAN (16.10%) (p=0.010, FPPI).[22]

[21] This is reasonable given the fact that in VOT promises are a strong strategic means to attract voters which on the other hand might make them less meaningful.

[22] For this analysis we need to omit one independent observation (group) in VOT (RAN) with only 1 (3) observation(s) respectively. The low number of investments of 10 in these two groups would inhibit meaningful statistical testing.

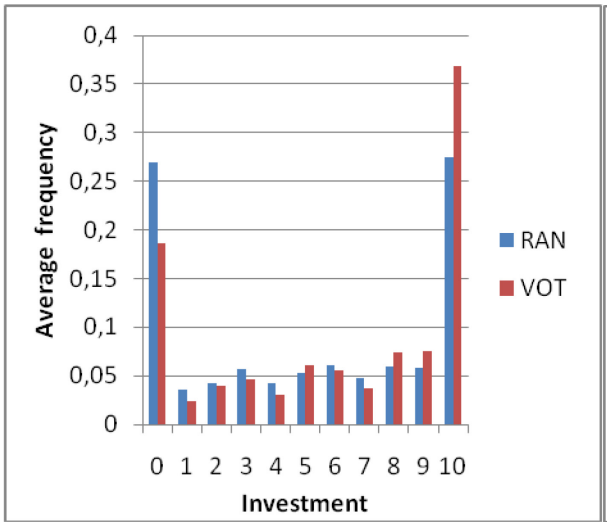


Figure 3-5: Distribution of individual investments by treatment.

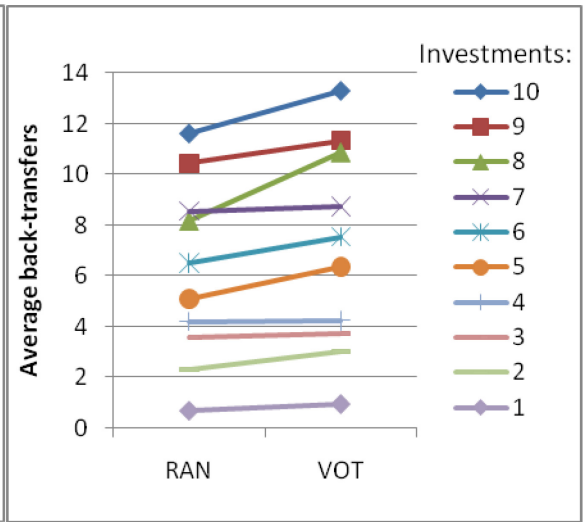


Figure 3-6: Average back-transfers for all positive investment levels in both treatments.

Hence, responders do not only move upward along their back-transfer functions; the function itself shifts from RAN to VOT (see also Figure 3-6 for other investments).

Observation 4: Trust is rewarded more in VOT.

3.4.2 An integrated model for back-transfers

So far we have examined the impact of single drivers responsible for observed RoI differences by applying non-parametrical statistics. In the next step we summarize our investigation on back-transfer behaviour by simultaneously controlling for treatment condition, types elicited through the preceding OS-game, promises, investments and the rounds of play (within each unit). Conducting an OLS-regression with robust standard errors adjusted for cluster over all decision rounds and groups, we find the treatment dummy, which is 1 for VOT and 0 for RAN, to positively influence RoI (see Table 3-1, first column).[23]

The coefficient of the dummy *Treatment* is about 0.2, which translated into RoI being 20% higher (or 0.2) in VOT than in RAN given everything else is kept constant. Hence, in VOT**,** trust is about 20% more rewarded compared to RAN. Furthermore, RoI do significantly and positively depend on responders' back-transfer strategy in the preceding OS-game: an

[23] As only few observations lay at either the lower bound or the upper bound of possible RoI (less than 5% for the lower bound and less than 10% for the upper bound) we refrain from explicitly modeling potential censoring, in order to facilitate the interpretation of estimated coefficients.

increase of OS-RoI by 1 (or 100%) would yield a nearly 25% higher RoI in the repeated game.[24] We can furthermore see that in the integrated model using all individual data at hand promises do have a positive impact on RoI; responders in both treatments feel to some extent positively committed to their promises: precisely, an increase in promised RoI by 1 (or 100%) leads, on average, to an increase in actual RoI by about 0.2.[25] However, as we have also shown previously, promises do not significantly differ across treatments. Therefore, the treatment effect is not driven by differences in promises.[26] Our model also shows that higher investments lead to higher RoI in both treatments. As a last result, Table 3-1 also suggests that RoI decline over the course of each unit.

An interesting question is how consistent our findings are dependent on the progress of the game. Does behaviour converge to a common level in both treatments; does the treatment effect therefore disappear in later periods? When separately analyzing the last unit (rounds 21 to 25) of our experiment (see Table 3-1, column 2), we find correlations to be quite consistent with our findings related to all periods of the game. The treatment dummy still has weakly significant explanatory power. Interestingly, whereas investments do not explain RoI anymore, OS-game behaviour still has highly significant explanatory power for the last five decisions rounds' RoI.

> *Observation 5: RoI in VOT are significantly higher compared to RAN. RoI depend on responders' one-shot game behaviour, on their promises, and on received investments. Differences in reciprocity are robust over the course of the game.*

[24] The explanatory variable is the RoI that the responder would provide for the investment received by senders in the repeated game given her back-transfer vector in the OS-game. For example if senders send 5, OS-RoI is the RoI that results from what responders would send back in the OS-game based on her entry for 5 in her back-transfer vector. If this entry is 5, then the OS-RoI is 0.

[25] This finding is different from what we have analysed with simple rank-correlations and non-parametric testing for differences in commitment toward promises based on group averages.

[26] This also becomes obvious by regression analysis in that the slope dummy *Promised RoI * Treatment* turns out not to significantly influence RoI (p=0.181).

Table 3-1: OLS-regression analysis on responder's RoI with robust standard errors adjusted for clusters (group-level).

Dependent variable: Return on Investments: (b_i/a_i)-1	**All units**	**Last unit**
Explanatory Variable	**Coefficient (robust standard errors adjusted for cluster (group-level) in parentheses)**	
OS-RoI	0.248*** (0.033)	0.347*** (0.104)
Promised RoI[27]	0.198** (0.075)	0.367** (0.138)
Investment	0.026*** (0.009)	0.010 (0.015)
Unit round (1,2,3,4,5)	-0.069*** (0.0143)	-0.149*** (0.050)
Treatment (1 for VOT, 0 for RAN)	0.198*** (0.048)	0.205* (0.115)
Constant	-0.105 (0.098)	-0.112 (0.231)
Number of observations	1737	309
Number of clusters	18	18
R^2	0.191	0.230

*** - p < 0.01 (2-sided); ** - p < 0.05 (2-sided); * - p < 0.1 (2-sided)

3.4.3 Differences in reputation building incentives?

In the course of our analysis on the origins of the positive effect on responders' reciprocity caused by an introduction of a voting institution, we have previously identified observable interdependencies in a controlled statistical model. Now we want to analyze potential other, more subtle drivers behind higher RoI. One subsequent question is whether differences in dynamic considerations contributed to differences in responders' behaviour across the two treatments. As we have already shown, end-unit effects are less pronounced in VOT than in RAN; one potential explanation that is not linked to the procedure of voting itself is responders in VOT trying more strongly to build a reputation across units than responders in RAN. This explanation, of course, is not consistent with our experimental design that rules

[27] Promised RoI are the implicit RoI that results if a responder sends back what she promised in her back-transfer strategy for the investment made by senders: for example, if sender sends 5 and the responder promises 10 for an investment of 5, then promised RoI would be 1.

out reliable reputation building and turns the election-winning promises strategies a focal point.[28] If our design indeed succeeded in ruling out individual identification across units, we would expect the voting mechanism to be essentially random. To start with, we investigate whether ex-post re-election probabilities differed across treatments. Despite a somewhat higher ex-post re-election probability for direct re-election in consecutive units in VOT (47.2%) compared to RAN (33.3%), no significant difference emerges (p=0.272, FPPI, for observed ex-post re-election probabilities).[29] A Pearson's chi-square test for the observed ex-post distribution of the attributes "direct re-election" and "no direct re-election" confirms this finding (p=0.230).[30] The same finding holds when direct re-elections and re-elections in later units are considered. Hence, at first sight, the voting mechanism seems to be random.

However, the incumbent has, interestingly, a significant bonus in elections. Based on a Poisson-regression to explain the number of votes that each candidates receives (prior to any run-off election between the two leading candidates), the incumbent receives significantly more votes, even when we control for the rank of the average promised RoI: the coefficient of *Incumbent* is positive and significantly different from 0 (see Table 3-2, Model 1). As expected, *rank of promises*, which is 1 for the highest average promised RoI and 3 for the lowest, has a highly significantly negative influence on the number of votes received. We employ a measure of the entire promised back-transfer vector as this has most explanatory power.[31] The highly significant and large constant nevertheless shows that the unexplained part of voting behaviour still makes up the largest part of the votes; the incumbent enjoys, on average, a expected bonus of 0.591 votes, whereas the constant is equivalent to 2.404 votes.[32]

[28] Recall that we implemented a randomised identification switch after each unit and displayed the winning back-transfer strategy to all candidates.

[29] A binomial test to check whether the ex-post re-election probability for a consecutive re-election is different from 1/2 shows no significant deviations in RAN (p=0.508) and VOT (p=0.508) from this theoretical benchmark.

[30] In both treatments 23 different subjects where in fact elected.

[31] Aggregating the entire vector produces most variance among the promises of the three candidates. Employing specific entries of the vector, e.g. for possibly focal investments such as for the mode of 10 or the average investments of the previous unit, often results in ties among two candidates and ignores possibly important differences in other parts of the vector. Using ranks of several entries of the vector runs in addition into multicollinearity-problems (as ranks for different investments are strongly correlated) and leads to a loss of degrees of freedom. Furthermore, different ways on how to enter candidates' promises into a regression framework has no significant effect on the incumbent-bonus in Model 1.

[32] All coefficients are expressed in expected logs or in the change of expected logs. The constant can be transformed into count data by taking e to the power of the constant's coefficient. Marginal effects are calculated with STATA.

As the significant *Incumbent*-dummy in Model 1 is surprising, we analyse in the next step what may explain the incumbent bonus in elections; in particular, we want to investigate whether back-transfer behaviour may influence electoral prospects. If, contrary to our expectations, identification was somehow possible and back-transfers mattered for voting behaviour, we would expect the votes of unelected candidates to negatively depend on the average RoI of the previous unit; by contrast, we would expect the incumbent to be awarded more votes if she provided a high RoI in the previous unit. We would also expect the incumbent to receive more votes, the closer she sticks to her previous back-transfer promise strategy (as identification codes are changed from one unit to the next, identification is only possible through the promised back-transfer strategy). If identification is possible, rank of promises may furthermore matter less for the electoral prospects of incumbents than of non-incumbents as the former can then, in contrast to the latter, be judged on their back-transfer behaviour.

In Table 3-2, Model 2, we therefore run a separated regression for incumbents and non-incumbents on the following variables: *rank of promises*, *lagged RoI* (RoI of the previous unit) as well *absolute deviation of promised RoI* (the deviation in promised RoI from the previous unit to the current unit as a proxy for possible identification from one unit to the next). Model 2 reveals that the incumbent-dummy is no longer significant. It also reveals that *rank of promises* do not have differing impacts on electoral prospects for incumbents and non-incumbents: both coefficients are almost equal in size and not significantly different.[33]

In a next step, we therefore separate the regression for incumbents and non-incumbents only for *lagged RoI* and *absolute deviation from promises.* Model 3 shows that non-incumbents indeed receive fewer votes if the incumbent provided higher RoI in the previous unit. While the coefficients of *absolute deviation from promises* seem to be different for incumbents and non-incumbents, neither coefficients is significantly different from zero, nor does a parameter-test reveal a significant difference[34] between them. In a final step, we therefore run a regression simply on *rank of promises* as well as *lagged RoI* separated for incumbents and non-incumbents. Model 4 in this case indeed shows (weakly) significant

[33] A Chi-squared test for equal parameters shows the two coefficients not to be significantly different: p=0.673.
[34] Chi-squared test for equal parameters: p=0.190.

influences of *lagged RoI* on electoral prospects that go into opposite directions for incumbents and non-incumbents. However, both coefficients are rather small. A 100% higher RoI in the previous unit only leads to 0.204 more expected votes for the incumbent and 0.231 fewer expected votes for the two other candidates. In all Models, the constant has, by some distance, the largest coefficient.

Table 3-2: Poisson-regression for number of votes received (prior to any run-off election) with robust standard errors adjusted for clusters (group-level).

Dependent variable: Number of votes received (prior to run-off)	**Model 1**	**Model 2**	**Model 3**	**Model 4**
Explanatory variables:	**Coefficient (robust standard errors adjusted for clusters (group-level) in parentheses)**			
Rank of promises	-0.258*** (0.085)		-0.278*** (0.092)	-0.264*** (0.083)
Incumbent (1 for incumbent, 0 for non-incumbents)	0.345** (0.144)	0.099 (0.452)		
Constant	0.877*** (0.177)	1.188*** (0.318)	1.222*** (0.268)	1.072*** (0.187)
Rank of promises * Incumbent		-0.324*** (0.075)		
Rank of promises * Non-incumbent		-0.248 (0.152)		
Lagged RoI* Incumbent		0.110 (0.257)	0.091 (0.075)	0.127* (0.074)
Lagged RoI * Non-incumbent		-0.225 (0.167)	-0.215** (0.102)	-0.144** (0.068)
Absolute Deviation of Promised RoI * Incumbent		-2.154 (1.596)	-2.020 (1.480)	
Absolute Deviation of Promised RoI * Incumbent		-0.126 (0.274)	-0.112 (0.267)	
Number of observations	108	108	108	108
Number of clusters	9	9	9	9
Prob > chi^2	0.0001	0.0000	0.0000	0.0000
Log pseudolikelihood	-159.793	-158.792	-158.880	-159.357
	*** - p < 0.01 (2-sided); ** - p < 0.05 (2-sided); * - p < 0.1 (2-sided)			

To conclude, the regression analysis provides a mixed picture: the significant bonus of incumbents in elections (see Model 1) seems to be explainable by responders' back-transfer behaviour, which has opposite effects on the electoral prospects of the incumbent and the two non-incumbents.[35] However, how closely the incumbent sticks to her promises of the previous unit does not provide significant explanatory power (see Model 2). The size of the influence of responders' back-transfer behaviour also remains small and is found, for both incumbents and non-incumbents, only in Model 4. Also note that in all models, *rank of promises* has considerable explanatory power and has the same effect on electoral prospects for incumbents and non-incumbents. If incumbents could be identified clearly, we would expect past behaviour to be more important for voters than cheap-talk promises. The evidence is therefore certainly not strong enough to convincingly show that incumbents indeed had incentives to behave reciprocally in order to improve their electoral prospects; nor do the results suggest that incumbents could be clearly identified. Nevertheless, we also have to acknowledge the surprising possibility that incumbents could, at least in some instances, be identified by their voters. This may be considered similar to findings reported from spectrum auctions, in which identification of different bidders, even though unreliable and costly, was still observed (Cramton and Schwartz 2000). We also do not know whether responders believed at all that they could increase their electoral prospects in the next unit by providing high RoI in the current unit. If they did, responders in VOT may have been more sensitive to dynamic considerations across units. This may contribute to VOT causing higher RoI than RAN.

While we cannot rule out the possibility of differences in dynamic considerations in the first four periods, we can do so for the last unit. Therein, strategic incentives based on foresight are identical between the treatments as no further election takes place; this was also common knowledge among subjects. As we have shown in Table 3-1, column 2, a weakly significant treatment effect is still detectable in the last unit. Furthermore, responders in VOT still deviate weakly significantly more from their OS-disposition compared to responders in RAN (p=0.064, FPPI); this seems a meaningful measure as OS-game behaviour can still predict responders' back-transfer behaviour for the five last rounds (see Table 3-1,

[35] The incumbent-dummy would not turn significant either if we added it to Models 3 and 4.

column 2). Last-unit results therefore suggest that the voting procedure has an effect on back-transfer behaviour that is independent from any possible differences in dynamic considerations.

> *Observation 6: There is no sufficient evidence that the observed treatment effect is entirely driven by differences in dynamic considerations between the treatments. However, incumbents receive more votes compared to candidates that were not elected in the previous unit. Treatment differences in the last unit suggest that the voting procedure itself nevertheless matters; voting by itself seems to moderate power.*

To sum up our observations on power holders' reciprocity, we conclude our main findings:

> ***Summary result 1**: We find power-holders to exercise their power less in a self-oriented way if they have been elected compared to being randomly drawn. Power holders' promises and the commitment toward them do not substantially differ across treatments. This also implies that in a promise-keeping sense, trustworthiness is not enhanced by the voting procedure. Unexpected reputation building opportunities across units may have arisen, but last-unit results suggest that these are not crucial to explaining treatment-differences. Rather, the voting procedure we employed may by itself have a power-moderating effect.*

3.5 Results II: Does voting foster trust?

We will now concentrate on the effect of voting on senders' behaviour. In order to look at the effects of the voting procedure on investors' trust, it is insightful to first look at behaviour in the first round. Therein, senders have not yet been exposed to any payoff-relevant decisions by the responders; hence, any differences in first round investments would be clear evidence for a treatment effect independent of responders' back-transfer behaviour. After that, we will look at senders' behaviour over time in order to investigate the stability of any treatment effects on trust.

3.5.1 First round treatment effects on trust

Investments start off considerably higher in VOT (6.69) than in RAN (5.47). This difference reaches weak significance if we compare all 45 individual investment decisions (p=0.092, FPPI), but stays insignificant if we compare strictly independent group-averages (p=0.148, FPPI).[36] Looking more closely at the data, we find a decisive difference in first-round behaviour between the treatments to be the number of trusting players. While all 45 investors send a positive amount in VOT, 8 investors refrain from investing at all in RAN. If we compare only those investors that have at least invested 1 point, average investments in RAN rise to 6.58.

> *Observation 7: Initial investments in VOT are higher than in RAN; in VOT, all senders invest, whereas in RAN, some investors refrain from investing at all.*

As for RoI, we will now explore potential explanatory variables for higher investments in VOT in the first round. Apart from the voting procedure itself, differences in first-round investments may be driven by differences in baseline investment levels, responders' promises, or expectations on responders' behaviour.

3.5.1.1 Population differences?

In fact, differences in the baseline trust-level, as displayed in the OS-game, run again against the observed treatment effect: average OS-investments by senders in VOT (3.24) are, on average, lower compared to RAN (4.16) – albeit not significant (p=0.140, FPPI). Taking the baseline trust-level into account, the treatment effect on trust rises to 2.13, which is highly significant (p=0.003, FPPI). The treatment effect on investment levels consequently over-compensates for somewhat underlying population differences. We therefore find strong evidence in support of H3.1.

> *Observation 8: Voting fosters trust; senders deviate significantly and positively more from their baseline trust level in VOT than in RAN.*

[36] While investment decisions within a group are, strictly speaking, not independent in VOT, the only interaction between investors in round one is through electing responders based on back-transfer promises. In RAN, no voting takes place but candidates' back-transfer strategies may have a common influence on investments. Average promised RoI, at which we will look in more depth later, do not significantly differ between the treatments (p=0.338, FPPI), nor do they correlate with investments (see next page). Consequently, it may be worth wile to also look at individual observations.

Interestingly, first-round investment behaviour is correlated with OS-behaviour in RAN (Spearman´s ρ=0.565, *p*=0.000, correlating individual observations, Spearman's ρ=0.519, *p*=0.152, correlating group-averages), but not in VOT (Spearman´s ρ=0.056, *p*=0.716, correlating individual observations; Spearman's ρ = 0.135, *p*=0.730, correlating group-averages). While the baseline investment-level that investors displayed in OS seems to act as an anchor in RAN, the voting procedure initially appears to induce higher investments independent of senders' baseline investment-level. Hence, the voting procedure seems to put senders initially into a separate (and trust-breading) context compared to the OS-game; this does not happen in RAN.

Observation 9: The baseline trust-level serves as an anchor for first-round investments in RAN, but not in VOT.

3.5.1.2 Higher promises?

If promised RoI were higher in VOT than in RAN, investors may be lured into investing more in VOT. However, average promised RoI in the first round are neither significantly higher in VOT (*p*=0.338, FPPI), nor do investment decisions significantly correlate with promises made by responders in either treatment – neither when correlating group-averages (Spearman's ρ=0.167, *p*=0.668 in VOT and Spearman's ρ=0.219, *p*=0.572 in RAN), nor when looking at all 45 individual investment decision in each treatment (Spearman's ρ = 0.001, p=0.993 in VOT and Spearman's ρ = 0.144 p=0.346 in RAN). Hence, the voting procedure itself seems to trigger higher trust, independently from responders' behaviour. H2 is supported by the data as promised RoI are not lower in VOT than in RAN.

Observation 10: First-round promises by winning candidates do neither significantly differ across treatments nor are they significantly correlated with actual investments.

3.5.1.3 Higher expectations?

For those senders who actually invest, expectations on responders' behaviour are somewhat but not significantly higher in VOT compared to RAN**:** senders expect to receive, on average, a RoI of 48.61% in VOT and of 38.37% in RAN (*p*=0.367, FPPI). As we do not have data on the expectation of senders who refrain from investing, we cannot directly test expectation differences for all 45 senders in each treatment. However, since investment levels are

positively correlated with expected RoI,[37] it seems, nevertheless, plausible to assume that the 8 senders in RAN expected investing not to pay off and therefore refrained from investing in the first place. Based on this assumption, i.e. setting expected RoI of the non-investors at 0, the (hypothetical) back-transfer expectations of senders may be seen as somewhat more pessimistic in RAN compared to VOT.[38]

> *Observation 11: Observed profitability expectations are not significantly higher in VOT than in RAN in the first round; however, it cannot be ruled out that senders nevertheless expect lower RoI in RAN than in VOT as the non-trusting senders in RAN may refrain from investing in RAN because they expect to make losses.*

3.5.2 Integrated analysis of first-round investments

So far, we have examined the impact of potential single drivers responsible for observed initial differences in senders' behaviour by applying non-parametrical statistics. As with responders' behaviour, in the next step we will summarise our investigation on senders' behaviour by simultaneously controlling for the treatment, types elicited in the preceding OS-game and RoI promised by responders. Given the data structure, we conduct an Ordered-Probit regression with cluster-robust standard errors (group level).[39] The analysis shows a significant, positive treatment effect: the treatment dummy (*Treatment*), which is 1 for VOT and 0 for RAN, is weakly significantly different from 0 (see Table 3-3, Model 1). Hence, controlling for the level of OS-investments and the promised RoI, first-round investments are weakly significantly higher in VOT than in RAN.

In line with the non-parametric statistics, we observe a markedly different role of player-types in the two treatments by comparing the coefficients of OS-RAN and OS-VOT (see Table

[37] Based on expectations and investment decision of each investing sender: VOT: Spearman's ρ=0.367, *p*=0.013; RAN: Spearman's ρ=0.520, *p*=0.001; based on group-averages, a significant correlation is observed if we jointly analyse both treatments: Spearman's ρ=0.400, *p*=0.099.

[38] Naturally, all assumptions on their expectations are arbitrary and potentially wrong; nevertheless, assuming these investors expected to receive at most a return of zero might provide us with further insight in the role of expectations. In this case, profitability expectations are weakly significantly higher in VOT than in RAN if we base the treatment comparison on all individual decisions: p=0.040, FPPI. Comparing the strictly independent group-averages yields no significant difference: p=0.130, FPPI.

[39] An Ordered-Probit is an appropriate regression model for a discrete distribution of values with a low number of categories; it is also more robust in the sense that it is less restrictive in on the data-generating process. While investments of 0 are less frequent in the first round than over the entire experiment, maximum investments are still the modes in both treatments.

3-3, Model 2).[40] Whereas, players' baseline trust level, as measured by the investments made in the preceding OS-game, has highly significantly explanatory power in RAN, no significance influence can be found in VOT.[41] Once we allow for treatment-differences in the influence of player-types, we observe, somewhat contrary to the non-parametric correlations, *average promised RoI* to weakly significantly predict investments. At the same time, the coefficient of the treatment dummy increases considerably and becomes highly significant. Hence, the voting treatment is found to have a robust, significant effect on first-round investments; therefore, H3.1 is also supported by means of regression analysis that simultaneously controls for the influence of promised RoI and baseline trust-level.

[40] *OS-investment X RAN* is 0 for all observations in VOT and equal to the OS-investments for all observations in RAN. *OS-investment X VOT* is 0 for all observations in RAN and takes up the value of the OS-investments for all observations in VOT.
[41] A Chi-squared test for equal parameters confirms the two coefficients to be significantly different: p=0.012. By contrast, no significant difference can be detected in the influence of promised RoI (p=0.922) so that respective slope dummies are omitted.

Table 3-3: Ordered-Probit regression models for first round investments with robust standard errors adjusted for clusters (group-level).

Dependent variable: **First round investments**	**Model 1**	**Model 2**
Explanatory Variables:	**Coefficient (robust standard errors adjusted for clusters (group-level) in parentheses)**	
OS-investment	0.161*** (0.060)	
Average promised RoI	0.689 (0.434)	0.754* (0.426)
Treatment (1 for VOT, 0 for RAN)	0.400* (0.222)	1.302*** (0.444)
OS-investment X VOT (0 for RAN)		0.041 (0.071)
OS-investment X RAN (0 for VOT)		0.291*** (0.084)
cut1	-0.391 (0.418)	0.106 (0.382)
cut2	-0.178 (0.364)	0.329 (0.353)
cut3	0.047 (0.332)	0.559 (0.317)
cut4	0.53 (0.297)	1.054 (0.284)
cut5	0.643 (0.327)	1.172 (0.300)
cut6	0.955 (0.351)	1.496 (0.320)
cut7	1.215 (0.365)	1.766 (0.322)
cut8	1.438 (0.383)	1.997 (0.353)
cut9	1.725 (0.338)	2.303 (0.351)
cut10	1.825 (0.343)	2.412 (0.364)
Number of observations	90	90
Number of clusters	18	18
Pseudo R2	0.048	0.071
Pseudolikelihood	-186.755	-182.294
*** - p < 0.01 (2-sided); ** - p < 0.05 (2-sided); * - p < 0.1 (2-sided)		

Observation 12: The voting treatment fosters trust as it has a significant positive effect on investment levels. This effect is independent from promised RoI and senders' baseline-trust level.

3.5.3 Stability of treatment effects on trust

To investigate whether the treatment also affects trust once investors get exposed to back-transfer behaviour by the responders, we now look at investments aggregated over all decision rounds. Despite an accumulated investment volume that is 20.46% higher in VOT compared to RAN, we find no significant difference in investment levels between VOT (6.23) and RAN (5.17) (p=0.240, FPPI). Looking at the evolution of treatment differences in investment-levels, we find investments to be higher for roughly the first half of the experiment (see also Figure 3-7): For the first 10 rounds (and also for both units separately)[42], investment levels are weakly significantly higher in VOT (6.90) than in RAN (5.12) (p=0.073, FPPI). From unit 3 onwards, investment levels are no longer significantly higher in VOT than in RAN;[43] differences in average values of investment levels between the treatments also narrow over time and are basically inexistent in the last unit.[44] Hence, the treatment effect on investments peters out over time. The moving average of investments (see Figure 3-7) shows how differences in investments between the treatments approach over time. Mirroring a similar finding on responders' behaviour, end-game investments tend to be higher in VOT (3.66) compared to RAN (2.32), albeit not at a statistically significant level (p=0.165, FPPI, see also Figure 3-7), when we look at investments in each unit's last round.

> *Observation 13: Investments are higher in VOT compared to RAN for the first two units of the experiment. Treatment differences peter out in later units.*

[42] Unit 1: 6.76 (VOT) compared to 5.21 (RAN**)** (p=0.060, FPPI); unit 2: 7.04 (VOT) versus 5.04 (RAN)

[43] Comparing average investments from round 1 to round 15, differences in investment levels between VOT (5.11) and RAN (6.61) nevertheless almost reach weak significance (p=0.120, FPPI).

[44] Unit 3: 6.05 (VOT) versus 5.07 (RAN); unit 4: 6 (VOT) versus 5.48 (RAN); unit 5: 5.29 (VOT) versus 5.05 (RAN).

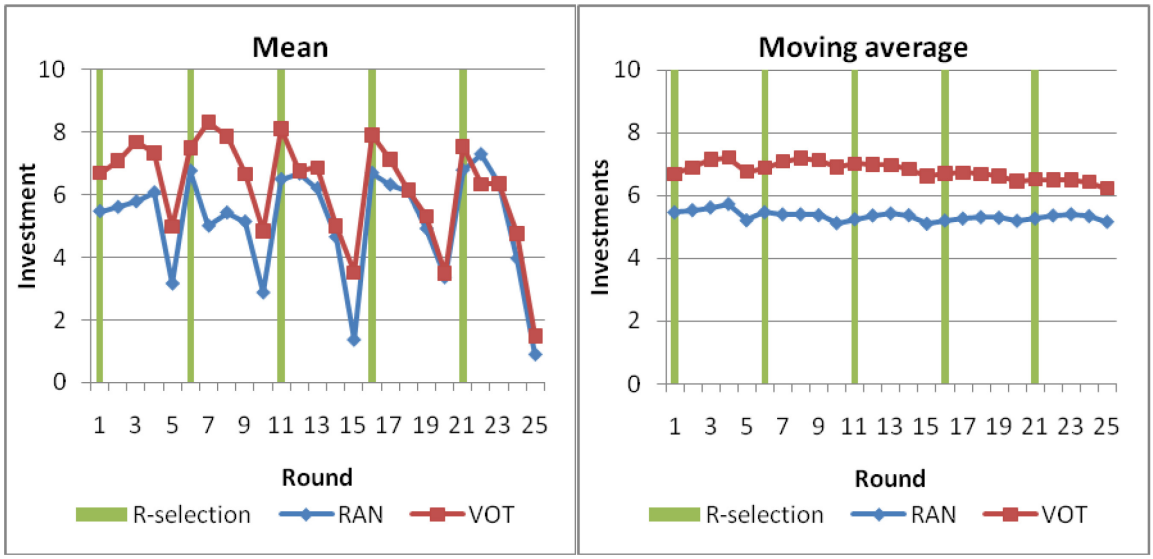


Figure 3-7: Mean investments and moving average of investments over time.

Let us now look at the effects of the treatment on trust by taking the players' baseline trust level into account. As we have already seen, senders in RAN have a somewhat higher baseline trust-level, according to their OS-behaviour, than senders in VOT. Controlling for this, treatment differences are pronounced over the entire 25 rounds: in VOT, senders send on average 2.98 more than in the preceding OS-game, whereas senders in RAN only deviate by 1.01 (see also Figure 3-8). This treatment difference (1.97) is weakly significant (p=0.056, FPPI). Furthermore, investments in RAN are not significantly higher than investments in the preceding OS-game (p=0.219, FPPI), whereas they are (and highly significantly so) in VOT (p=0.01, FPPI). Hence, trust is fostered by voting, whereas RAN does not lead to a lasting, positive effect on trust. We therefore find mixed-support for H3.2: investments are only higher for the first 2 units, but the voting treatment nevertheless fosters trust consistently.

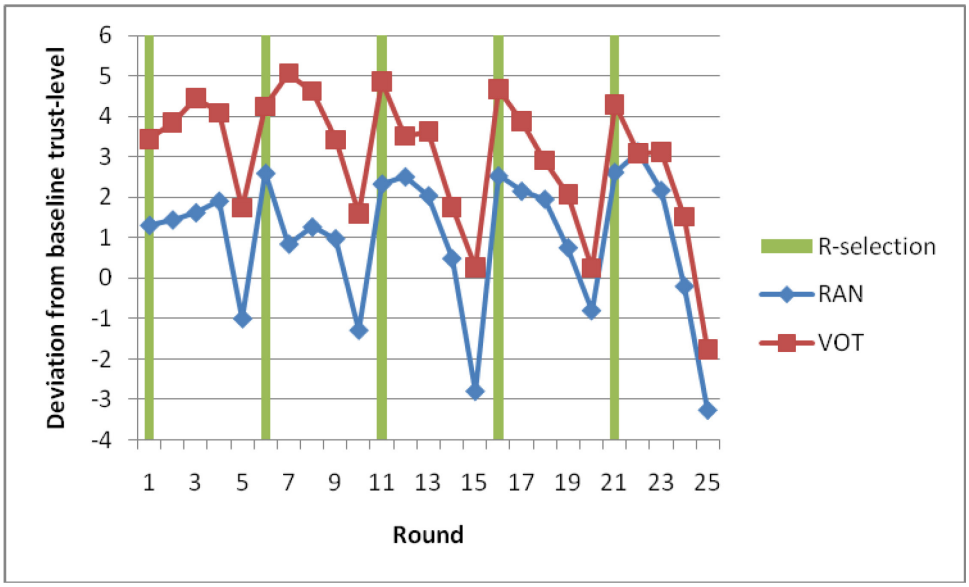


Figure 3-8: Mean deviation from baseline trust level over time: investment (round t) - OS-investment.

Observation 14: Senders positively deviate more from their baseline trust level in VOT than in RAN. Furthermore, trust is over the entire experiment only fostered in VOT, but not in RAN.

3.5.4 The roles of promises and responder-behaviour for explaining higher investments in the first two units

In order to explain treatment differences in investment levels over the first two units of the experiment, we will now refine our analysis and look at the role of promises and, most of all, the explanatory power of responders' back-transfer behaviour. As for the first round, responders do not give significantly higher promises in VOT than in RAN in the first two units either: responders promise to deliver, on average, a RoI of 60.81% in VOT and of 49.80% in RAN (p=0.375, FPPI). Group-level investments also do not correlate with promised RoI, neither in VOT nor in RAN: Spearman's ρ is -0.167 (p=0.668) in VOT and 0.151 in RAN (p=0.699).[45] Note that the correlation coefficient for VOT is even negative, albeit clearly not significantly so. Hence, promises do not drive treatment differences in investment levels.

[45] Correlations would not even become "significant" if we looked at correlations between promises and each sender's average investments (which, of course, are not statistically independent): Spearman's ρ=-0.52, p=0.736 (VOT) and Spearman's ρ=0.175, p=0.251 (RAN). Over the entire 25 rounds, the results are qualitatively the same: Spearman's ρ is -0.300 (p=0.433) in VOT and 0.075 in RAN (p=0.847). Again this insignificant role of promises would not change even if we treated all individual investments as independent observations: Spearman's ρ = 0.079, p=0.605 (VOT) and Spearman's ρ = -0.219, p=0.148 (RAN).

After the initial round, investors can update their beliefs on the behaviour of responders. Therefore, we now want to analyse whether once investors are exposed to responders' behaviour profitability considerations drive differences in investment levels. If investors expect a negative RoI, only those who are either risk-loving, altruistic or have a preference for efficiency would still invest. Indeed, in both treatments we find only very few instances in which investors are willing to invest despite expecting a negative RoI (4.18% of all investment decisions in VOT and 5.69% of all investment decisions in RAN). Whereas risk-neutral and money-maximising investors would send their full endowment if they expect a positive RoI, risk or inequality averse investors would still respond to changes in RoI above zero.

As we have shown, RoI are higher in VOT than in RAN; interestingly, however, this only applies partly to the first two units. Despite a considerable difference in average values, statistical tests only reveal a marginal significance: average RoI for the first two units are 1.36 in VOT and 1.13 in RAN (p=0.119, FPPI). Only comparing the share of investments that led to positive profits, we nevertheless find weakly significant treatment differences: 86.50% of all non-zero investments are profitable in VOT, but only 66.70% in RAN (p=0.078, FPPI). Likely, past experiences determine expectations of the investors on the profitability of their investments. We therefore expect past experiences with profitability of investing to also influence investment behaviour. Indeed, both RoI as well as the share of profitable investments are highly correlated with investments in both treatments: Spearman's ρ lie between 0.77 and 0.94 and are but one highly significant.[46] As profitability of investing is higher in VOT than in RAN, differences in responders' behaviour seem to be an important driver of higher investments in VOT compared to RAN in the first two units of the experiment.

> *Observation 15: While promised RoI do not visibly provide any explanatory power, differences in RoI are an important driver of treatment differences in the first two units.*

[46] Correlating RoI and investments: Spearman's ρ=0.940, p<0.001 (VOT); Spearman's ρ=0.917, p<0.001 (RAN); correlating share of profitable investments and investments: Spearman's ρ=0.917, p<0.001 (VOT); Spearman's ρ=0.767, p=0.016 (RAN).

By integrating the above analysis into an econometric framework, we can first simultaneously control for the role of promised RoI and baseline trust-level and, in a next step, test whether the procedure of voting still matters in the first two periods once we control for treatment-differences in profitability of investing.

3.5.5 An integrated model for trust-formation

Regression Model 1 in Table 3-4 confirms the non-parametric analysis on the role of promises, baseline trust-level and patterns within each unit: in contrast to *Average Promised RoI*, which do not significantly affect investments, *OS-investment* has a highly significant impact on investments. Both treatments display a pattern of decreasing investments over the course of units 1 and 2 each. We do not find the effects of promised RoI, baseline trust-level and the unit structure to differ between the treatments; therefore, a pooled regression is conducted.[47] As the treatment dummy is positive and significantly different from 0, neither promised RoI nor baseline trust-levels drive treatment differences.

[47] Chi-squared tests for equal parameters: *Unit Round RAN* = *Unit Round VOT*: p=0.821; *Average promised RoI RAN* = *Average promised RoI VOT*: p=0.429; *OS-investments RAN* = *OS-investments VOT*: p=0.322.

Table 3-4: Ordered-Probit regression models for explaining investments in the first 2 units with robust standard errors adjusted for clusters (group-level)

Dependent variable: Investments in rounds[48] 2 – 10	**Model 1**	**Model 2**
Explanatory Variables:	**Coefficient (robust standard errors adjusted for clusters (group-level) in parentheses)**	
OS-investment	0.099*** (0.198)	0.090*** (0.024)
Average promised RoI	0.291 (0.396)	-0.354 (0.288)
Treatment (1 for VOT, 0 for RAN)	0.536* (0.276)	0.132 (0.137)
Unit Round (1,2,3,4,5)	-0.213*** (0.030)	-0.307*** (0.037)
Mean of lagged RoI		1.581*** (0.115)
Lagged positive RoI dummy (1 if investor made a profit in previous round, 0 otherwise)		0.721*** (0.086)
cut1	-0.891 (0.301)	0.486 (0.206)
cut2	-0.751 (0.307)	0.662 (0.208)
cut3	-0.618 (0.315)	0.84 (0.2)
cut4	-0.435 (0.315)	1.081 (0.188)
cut5	-0.293 (0.324)	1.269 (0.187)
cut6	-0.129 (0.331)	1.482 (0.181)
cut7	0.083 (0.324)	1.766 (0.177)
cut8	0.193 (0.324)	1.906 (0.179)
cut9	0.375 (0.33)	2.14 (0.195)
cut10	0.596 (0.327)	2.415 (0.221)
Number of observations	810	787
Number of clusters	18	18
Pseudo R^2	0.041	0.128
Pseudolikelihood	-1610.211	-1420.766

*** - p < 0.01 (2-sided); ** - p < 0.05 (2-sided); * - p < 0.1 (2-sided)

[48] For comparative reasons, we restricted Model 1 to the same rounds as Model 2, in which round 1 is automatically omitted as no lagged data exists therein.

The comparison of Model 1 and 2 of Table 3-4 shows the driver of higher investments in rounds 2 to 10 to be differences in responders' behaviour. Once we include measures of senders' profitability experiences, the treatment dummy seizes to significantly influence investments. Both *mean of lagged RoI*[49] as well as the dummy *last round positive RoI,* which is 1 if senders made a positive profit in the previous round and 0 if she did not, highly significantly influence investment levels.[50]

There are no significant treatment differences in the effects of *average past RoI* and *last round positive RoI dummy*, which is why we again run a pooled regression for both treatments.[51] As responders turn investing more profitable in VOT than RAN, these differences drive the higher investments in VOT compared to RAN in rounds 2 to 10.[52] Hence, the initial effect of the voting procedure on trust does not last. This may be due to two reasons: First, profitability considerations may be so dominant for senders that initial, more subtle psychological effects of the procedure of voting no longer matter sufficiently. Second, already the initial effects of voting may be due to the expectation of higher profitability of investing in VOT compared to RAN. Possibly, if we had a measure for the expectations of those senders that did not invest, we would find differences in expectations, rather than the procedure itself, to drive first-round differences. Interestingly, once we control for past actual RoIs (in Model 2), higher promised RoI may even negatively influence investments (albeit not at a statistically significant level).

To conclude the analysis on trust, Table 3-5 reports the drivers of trust over the entire experiment. Table 3-5 is, by and large, in line with the analysis of rounds 2 – 10 (see Table 3-4). Baseline trust-level as well as measures for past profitability experiences are significantly and positively correlated with investments over the entire experiment. Unsurprisingly and as obvious from Figure 3-7, we find investments to have a clearly pattern in each unit by decreasing as the unit progresses. The coefficient of the treatment dummy is

[49] *Mean of lagged RoIs* in round T and for player S_i equals $\sum_{t=1}^{T-1}(b_i/a_i - 1)_t/(T-1)$ for all $a_i > 0$

[50] Consequently, the dummy-variable *last round positive RoI* is also zero if the investor has not invested.

[51] Chi-squared tests for equal parameters when running a full set of slope dummies: *Mean of lagged RoI RAN* = *Mean of lagged RoI VOT*: *p*=0.820; *Lagged positive RoI dummy RAN* = *Lagged positive RoI dummy VOT*: *p*=0.267. The same holds again for the other explanatory variables: *Unit Round RAN* = *Unit Round VOT*: *p*=0.827; *Average promised RoI RAN* = Average *promised RoI VOT*: *p*=0.913; *OS RAN* = *OS VOT*: *p*=0.755.

[52] Round 1 is automatically excluded as no lagged data exists therein.

now negative; however, this may be purely coincidental as it is very close to 0. We therefore do not reject the Null-hypothesis of the procedure of voting not to influence investments over the entire experiment.

Table 3-5: Ordered-Probit regression models for explaining investments in rounds 2 – 25 with robust standard errors adjusted for clusters (group-level).

Dependent variable: **Investments** **Explanatory Variables:**	**Coefficient** **(robust standard errors adjusted for clusters (group-level) in parentheses)**
OS-investment	0.068*** (0.023)
Average promised RoI	-0.315 (0.231)
Treatment (1 for VOT, 0 for RAN)	-0.185 (0.119)
Unit round (1,2,3,4,5)	-0.364*** (0.026)
Mean of lagged RoIs	1.751*** (0.183)
Lagged positive RoI dummy (1 if investor made a profit in previous round, 0 otherwise)	0.665*** (0.083)
cut1	0.717 (0.209)
cut2	0.845 (0.215)
cut3	1.013 (0.211)
cut4	1.199 (0.202)
cut5	1.333 (0.203)
cut6	1.527 (0.212)
cut7	1.723 (0.208)
cut8	1.863 (0.211)
cut9	2.085 (0.215)
cut10	2.319 (0.247)
Number of observations	2137
Number of clusters	18
Pseudo R2	0.114
Pseudolikelihood	-3832.734
*** - p < 0.01 (2-sided); ** - p < 0.05 (2-sided); * - p < 0.1 (2-sided)	

Observation 16: The integrated analysis over the entire experiment shows the treatment not to have an effect on investment that is independent from profitability experiences.

To conclude the subsection on the role of voting on trust, we sum up our main findings:

***Summary result 2**: We find the voting procedure to foster trust independently of responders' back-transfer behaviour, i.e. in the first round. Once investors get exposed to power-holders' back-transfer behaviour, differences in profitability of investing drive treatment differences. Long-run effects of the procedure on investments cannot be observed.*

3.6 Result III: Does voting affect the payoff-distribution?

In the last subsection we take a look at senders' and responders' profits in order to evaluate players' monetary outcomes in both treatment conditions. As we have shown, both profitability and average values of investments are higher in VOT. As a result, senders earn weakly significantly more in VOT (12.59 points) compared to RAN (11.25) (p=0.060, FPPI). This result supports hypothesis H4.

Interestingly, however, H5 is not supported by the data. Despite a difference in average values along the lines of H5, responders in RAN (29.58) do not make significantly higher profits than responders in VOT (28.20) (p=0.439, FPPI). Nevertheless, the higher RoI in VOT certainly did not pay off for responders. Responders in RAN were therefore not irrational in the sense that they could have achieved higher profits for themselves (and for the group) by providing higher RoI as in VOT. In this sense, we have to qualify the answer to our research question. In any stage game, responders in VOT do exert their power less to their own advantage; over time, however, part of this sacrifice for the benefit of the group is compensated by somewhat higher investments in VOT compared to RAN.

As efficiency, measured as the sum of incomes of senders and responders, increases with investments, both treatments' efficiency levels are similar (82.85% in VOT and 78.04%,

p=0.242, FPPI).[53] However, the distribution of income slightly differs across treatments. In RAN**,** senders earn on average 18.3 points less than responders. In VOT, this difference (15.6) is smaller than in RAN at a marginal level of significance (p=0.103, FPPI).

Our next result concludes the main findings on the role of voting for profits:

> ***Summary result 3****: Whereas voting leads to a more equal distribution of incomes and higher profits for investors, over the course of the game, holders of power do not make significantly lower profits under the voting mechanism; part of power-holders sacrifice for the benefit of the group is compensated by somewhat higher investments triggered by the voting procedure****.***

3.7 Conclusion and outlook for future research

In this paper, we analyse the short- and long-run effects of voting on the exercise of power and trust put into holders of power. Investors can trust by sending investments to a responder who has power to seize post-investment assets. The return on investments that responders provide is a measure of her reciprocity. The higher the return on investment, the higher is the immediate loss of profit of the power holder, and hence, the less they exercise their power to their own advantage. The responder is either elected by the investors, in the voting treatment, or selected by a random draw, in the random treatment. As candidates also make specific back-transfer promises before election, we also measure the effect of voting on promised-based trustworthiness. Prior to the repeated game, a one-shot bilateral investment game is played, which provides us with a baseline measure of trust for all investors and of reciprocity for all later holders of power. We find three main results:

First, voting has a long-run effect on the exercise of power, but no significant effect on trustworthiness; by providing higher returns on investment, elected holders of power exercise their power less to their own advantage than randomly drawn counterparts. This effect is robust to controlling for player-types, promised returns on investments and the level of investments. In fact, the voting treatment over-compensates for underlying differences in players' reciprocity. While we do find some evidence of power holders'

[53] Thereby, we do not take the incomes of un-selected candidate into account as they do not participate in the actual game. Doing so would not change the treatment comparison as un-selected candidates always earn the same in both treatments.

commitment to promises in both treatments, elected candidates do not keep their word better than randomly drawn holders of power. In the promise-keeping sense, voting fails to increase trustworthiness. Although the design rules out reliable individual reputation building, differences in dynamic considerations due to instances of subtle identification possibilities may contribute to the effects of voting on the exercise of power over the first four units of the experiment. The small role of identification for electoral prospects as well as evidence from the last unit nevertheless suggests that the procedure of voting itself moderates power.

Second, voting has a long-run effect on trust, but not on investment-levels. The voting procedure induces investors to deviate positively from their initial trust-level, as displayed by the investments made in the preceding one-shot game. This is not found when power-holders are randomly drawn. Investment levels, by contrast, are only higher under the voting procedure compared to the random procedure for the first 10 rounds of the experiment. Long-run effects of voting on investments can therefore not be observed. The effects of voting on investments are robust for controlling for promised return on investments. Voting also has an effect on investments and trust that is independent of power holders' back-transfer behaviour, i.e. in the first round of the experiment. Once investors get exposed to power-holders' back-transfer behaviour, however, differences in profitability of investing drive treatment differences. Hence, an effect of the procedure of voting itself can only be observed in the short-run.

Third, we find the implementation of a voting institution to raise investors' incomes. Investors are considerably better off under the voting treatment than under the random treatment. Whereas in each round, power-holders sacrifice profits for the benefits of the group under the voting procedure, they do not make significantly lower profits under the voting procedure compared to the random mechanism; somewhat (albeit not significantly) higher investments when power holders are elected compared to randomly drawn compensate for the loss of profits in each round.

Our experiment opens ample room for further research. Future experiments may shed more light on the relative strengths of dynamic considerations versus the reciprocity-based effects

of voting that were hypothesised to matter. Further research is also warranted in finding out more on the role expectations play for voting fostering trust. Is the trust-fostering effect of voting only due to investors expecting higher reciprocity or is there an independent effect that is linked to the procedure itself? It also seems interesting to investigate more closely the long-run effects of voting on trust. Why does the voting effect on investments wear off so quickly, despite still somewhat higher profitability of investing even in later units? A promising avenue for future research on the effects of voting and trust and power may also be to change the cultural background of subjects. Based on the World Bank governance indicators (see section 3.3), our subjects come from a background of very little exposure to democratic procedures in situations of conflicting interests. Possibly, subjects having more experience with elections would attach a greater weight to the voting procedure we employed. In this case, voting may be more powerful in fostering trust and moderating power.

3.8 Appendix: Experimental instructions

3.8.1 Instructions of the OS-game

Welcome to join us in an economic experiment.

Please read the instructions of the experiment carefully. Do not communicate with other participants in any form throughout the experiment. If you have any questions, read the instruction again, or raise your hand if it is still puzzling. We will reach you and answer your question.

You will be paid cash for participating in the experiment. The money you earn will first be calculated in points. The amount of points you earn depends on your and other participants' decisions in the experiment. You will receive RMB from the points you made in the experiment.

Please make your decision independently and do not communicate with other participants in any form. If you are found to do so, you will be dismissed from the experiment.

All data and answers will be analyzed anonymously. You have already drawn a code, which ensures anonymity. Please make your decision on the computer in the cubicle numbered by that code. We will only know what decision is made by which code, instead of whom.

Today's experiment consists of two independent parts. Now we run the first part. When the first experiment is finished, you will get the experimental instructions for the second experiment. Attention: your decision in the first experiment has no effect on your decisions in the second experiment!

After the first experiment, you will not know about your payoff from the first experiment. The experiment directly goes to the second experiment. Your payoff from the first experiment will be added to your payoff from the second experiment and will be paid to you in cash after the second experiment.

Experimental instruction of the first experiment

You will receive RMB from the points you earned by the following exchange rate:

10 points = 10 RMB

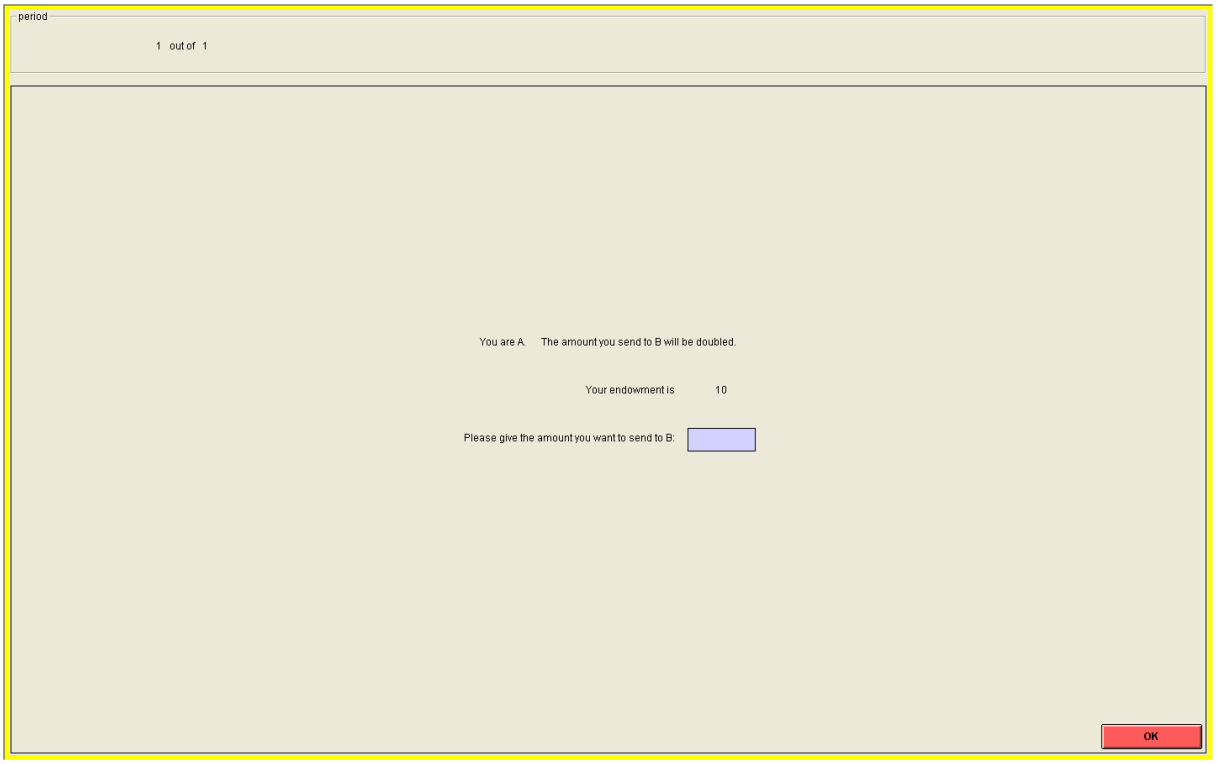
In this experiment, all participants will be assigned either of two types: 15 as type A and 9 as type B. Your type will be randomly assigned in the beginning of the experiment.

The experiment has only one decision round! A and B decide simultaneously. A makes one transfer decision, B makes multiple back-transfer decisions.

At the beginning of the experiment, A and B receive each 10 points as initial endowments. B puts this endowment aside and will get it paid after the experiment.

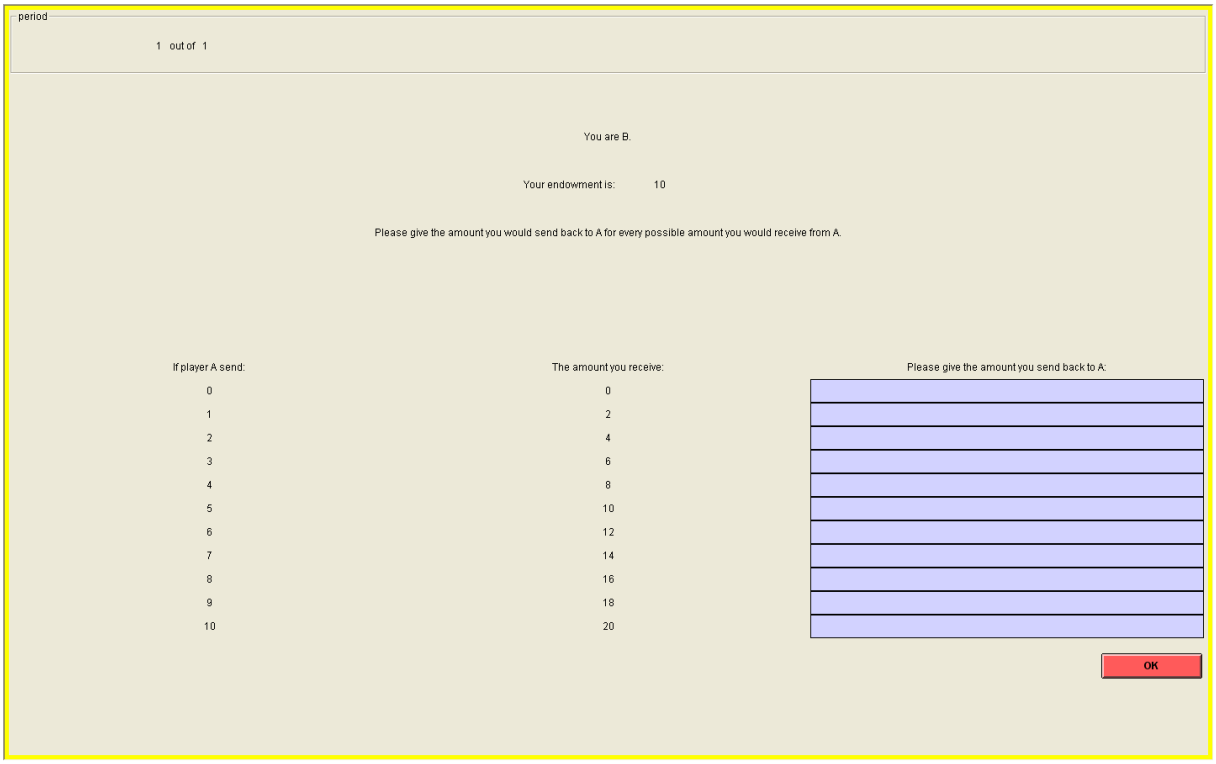
1. A's decision: the transfer decision

A can send any integer part of her endowment to B. That is to say, A can choose any integer number from the set {0; 1; 2; 3; 4; 5; 6; 7; 8; 9; 10} to send to B. The amount sent by A will be doubled by the experimenter. Thus, B will not get A's initial transfer, but twice of the transfer. Here is a screen shot of what A sees on her computer:



2. B's decision: the back-transfer decision

B receives the doubled amount sent by A. B can transfer any integer part of this doubled amount back to A. That means, B can choose an integer number from the interval [0; 2 x A's transfer] to back-transfer to A. However, B does not know the actual amount which was sent by A and doubled by the experimenter. Thus, B must decide a back-transfer amount for each possible transfer of A. For both decision-makers, the decision relevant for the payoffs is that back-transfer amount which B chooses according to the actual transfer made by A. The amount back-transferred from B to A will not be doubled. Here is the screen shot of what B sees on his computer.



How to calculate the income from the first experiment?

A's income:

Every one of the 15 As will be randomly matched to one of the Bs. The payoff is calculated according to the actual decisions of the matched A and B.

A's income depends on the points she sent to B and the points back-transferred from B to A. A's income is calculated as follows:

A's income = A's endowment (10 points) – points sent from A to B + points back-transferred from B to A

You can see from the formula, the more B back-transfers to A, the higher is A's income. The less A sends to B, the higher is A's income.

B's income:

Every one of the 9 Bs will be randomly matched to one of the As. The payoff is calculated according to the actual decisions of the matched A and B.

B's income depends on the points A sent to him and the points back-transferred from him to A. B's income is calculated as follows:

B's income = B's endowment (10 points) + 2 x points sent from A to B – points back-transferred from B to A

You can see from the formula, the more A sends to B, the higher is B's income. The less B back-transfers to A, the higher is B's income.

Payoffs of all As and Bs are calculated by the formulas mentioned above. Every A can calculate the B's income whom she is matched with; every B can calculate the A's income whom he is matched with.

3.8.2 Instructions of the treatment VOT

Now the second experiment starts. Please read the experimental instructions carefully. You will be again paid cash for participating in the experiment. The money you earn will first be represented in points. The amount of points you earn is depending on your and other participants' decisions in the experiment.

You will receive RMB from the points you earned by the following exchange rate:

10 points = 2 RMB

In this experiment, you will interact with other 7 participants. At any time, you will not know who they are. Similarly, other participants will not know who you are.

Experimental instruction of the second experiment

There are 24 participants in the experiment. At the beginning, you will be randomly assigned to a group of 8 members. Then you will be assigned a role in the experiment. All participants will be assigned either of the two types: 5 will be citizens and 3 will be candidates. Your current type depends on your type in the first part of the experiment. That is to say, if you were type A in the first experiment, you are citizen now; if you were type B in the first experiment, you are candidate now. Attention: the composition of each 8-member group and the role of every participant remain unchanged throughout the experiment.

There will be 25 decision rounds in total, with every 5 rounds as a unit. The decision environments of each unit are identical, which means that the experiment process of the first unit (rounds 1-5) is the same as that of the second unit (rounds 6-10); of the third unit (rounds 11-15); of the fourth unit (rounds 16-20) and of the fifth unit (rounds 21-25).

At the beginning of each round, all group members receive 10 points as initial endowment. The 5 citizens will decide on how many points to send. At the beginning of each unit (round 1, 6, 11, 16 and 21), one of the 3 candidates will be selected as the president for the following 5 rounds of each unit (described below). In each round of the unit, all citizens can decide to

send any integer point to the president from their initial 10 points. That means, citizens can send any integer they wish to the president from the set {0; 1; 2; 3; 4; 5; 6; 7; 8; 9; 10}.

The president will receive double points as sent by the 5 citizens. That means, each amount of points sent by citizens will be doubled and sent to the president. Thus, the president will get twice the amount of points sent by citizens, instead of the same amount. The amount of points that citizens sent to the president will be shown to the president in a descending order. That means, the president will see how many points he gets from the citizens, but not which citizen sent him how much. The president must decide, for each amount of points, how much to back-transfer. The president can back-transfer any integer from the doubled points he received. That means, the president can back-transfer any integer from the interval [0; 2 x sent amount], to citizens. The amount of points the president back-transfers to citizens will not be doubled.

How to select the president?

At the beginning of each unit (1^{st}, 6^{th}, 11^{th}, 16^{th} and 21^{st} round), the 3 candidates can announce a back-transfer suggestion, as if they were selected to be president. The back-transfer suggestion is how much each candidate, as president, will back-transfer given citizens' transfers (0, 1, 2, 3, 4, 5 , 6 ,7 ,8 ,9 ,10) after he receives points sent from citizens (0, 2, 4, 6, 8, 10, 12, 14, 16, 18, 20). These suggestions serve as candidates' declaration. However, the president is not restricted to his announcement. That means the president is not bound to his announcement. After reading candidates' declaration, citizens will elect a president from the 3 candidates for the following 5 rounds of this unit. Each citizen has 1 vote. According to the majority principle, the candidate who receives at least 3 votes will be elected as president. If two candidates both receive 2 votes and the third candidate gets 1 vote, the third candidate is eliminated. The 2 candidates who get 2 votes have to re-announce their declaration. Citizens will vote again to elect a president. The eliminated candidates see the declaration of the president and wait for the start of the next unit.

After a unit is finished, all 3 candidates will be assigned a new code. They will again propose a back-transfer suggestion as their declaration. Citizens will see the new declaration from all candidates and vote one of them to be the president for the next unit. Attention: candidates

will be assigned a new code in each unit. That is to say, citizens do not know which candidate has been elected as president in previous units, and which candidate has not been elected as president in previous units.

Steps of the experiment in detail

1. At the beginning of the experiment, each participant will be randomly assigned to an 8-member-group. The three candidates receive their codes.

2. The following only happens at rounds 1, 6, 11, 16 and 21:

    a. The 3 candidates propose the back-transfer suggestions, as if they were elected to be the president. The back-transfer suggestion is how much he will back-transfer given citizens' transfer (0, 1, 2, 3, 4, 5, 6 ,7 ,8 ,9 ,10) after he receives from citizens (0, 2, 4, 6, 8, 10, 12, 14, 16, 18, 20). These suggestions will be served as candidates' declaration.

    b. All citizens will read the declaration from all candidates and elect one of the candidates to be president for the following 5 rounds of the current unit.

    c. If 2 candidates both receive 2 votes and the third candidate gets 1 vote, the 2 candidates who get 2 votes will have to re-announce their declaration. Citizens will vote again and elect 1 of these candidates to be president.

    d. The eliminated candidates see the president's declaration. After one minute the program automatically enters a waiting screen for the start of next unit.

    e. At the end of each unit, the 3 candidates make a new declaration for the next unit. Back to step 2a).

3. In every decision round:

    a. Citizens make a decision how many points to send to the president. Citizens can send any integer they wish to the president from the set {0; 1; 2; 3; 4; 5; 6; 7; 8; 9; 10}.

b. President sees the amount of points he gets from the 5 citizens, without knowing exactly who sent how much. The president will then decide upon the amount of points to be back-transferred to each citizen. The president can back-transfer any integer from the set [0; 2 x sent amount] to citizens. The amount of points the president back-transfers to citizens will not be doubled.

c. Every participant receives information about his own payoff from the current round.

4. (Only at the end of rounds 5, 10, 15 and 20). All 3 candidates are assigned a new code, back to step 2.

5. (Only after round 25). Every participant receives information about her/his total payoff and receives payment.

How to calculate your income from the experiment?

Your total income from the experiment is the sum of income you receive from each decision round. The income of each round is calculated as follows:

Income of citizens in each round = Citizen's initial endowment (10 points) – the points sent to president + the points back-transferred by president

Income of candidates who are not elected as president in each round = candidate's initial endowment (10 points)

Income of president from each round = President's initial endowment (10 points) + the points sent by the 5 citizens x 2 – the points back-transferred to the 5 citizens

3.8.3 Instructions for treatment RAN

Now the second experiment starts. Please read the experimental instructions carefully. You will be again paid cash for participating in the experiment. The money you earn will first be represented in points. The amount of points you earn is depending on your and other participants' decisions in the experiment.

You will receive RMB from the points you earned by the following exchange rate:

10 points = 2 RMB

In this experiment, you will interact with other 7 participants. At any time, you will not know who they are. Similarly, other participants will not know who you are.

Experimental instruction of the second experiment

There are 24 participants in the experiment. At the beginning, you will be randomly assigned to a group of 8 members. Then you will be assigned a role in the experiment. All participants will be assigned either of the two types: 5 will be citizens and 3 will be candidates. Your current type depends on your type in the first part of the experiment. That is to say, if you were type A in the first experiment, you are citizen now; if you were type B in the first experiment, you are candidate now. Attention: the composition of each 8-member group and the role of every participant remain unchanged throughout the experiment.

There will be 25 decision rounds in total, with every 5 rounds as a unit. The decision environments of each unit are identical, which means that the experiment process of the first unit (rounds 1-5) is the same as that of the second unit (rounds 6-10); of the third unit (rounds 11-15); of the fourth unit (rounds 16-20) and of the fifth unit (rounds 21-25).

At the beginning of each round, all group members receive 10 points as initial endowment. The 5 citizens will decide on how many points to send. At the beginning of each unit (round 1, 6, 11, 16 and 21), one of the 3 candidates will be randomly selected as the president for the following 5 rounds of each unit (described below). In each round of the unit, all citizens can decide to send any integer point to the president from their initial 10 points. That means,

citizens can send any integer they wish to the president from the set {0; 1; 2; 3; 4; 5; 6; 7; 8; 9; 10}.

The president will receive double points as sent by the 5 citizens. That means, each amount of points sent by citizens will be doubled and sent to the president. Thus, the president will get twice the amount of points sent by citizens, instead of the same amount. The amount of points that citizens sent to the president will be shown to the president in a descending order. That means, the president will see how many points he gets from the citizens, but not which citizen sent him how much. The president must decide, for each amount of points, how much to back-transfer. The president can back-transfer any integer from the doubled points he received. That means, the president can back-transfer any integer from the interval [0; 2 x sent amount], to citizens. The amount of points the president back-transfers to citizens will not be doubled.

How to select the president?

At the beginning of each unit (1^{st}, 6^{th}, 11^{th}, 16^{th} and 21^{st} round), the 3 candidates can announce a back-transfer suggestion, as if they were selected to be president. The back-transfer suggestion is how much each candidate, as president, will back-transfer given citizens' transfers (0, 1, 2, 3, 4, 5 , 6 ,7 ,8 ,9 ,10) after he receives points sent from citizens (0, 2, 4, 6, 8, 10, 12, 14, 16, 18, 20). These suggestions serve as candidates' declaration. However, the president is not restricted to his announcement. That means the president is not bound to his announcement. One of the three candidates will be randomly selected as president. All citizens will see the declarations of all candidates and know which candidate has been randomly selected as president. The candidates who are not randomly selected as president see the declaration of the president and wait for the start of the next unit.

After a unit is finished, all 3 candidates will be assigned a new code. They will again propose a back-transfer suggestion as their declaration. Computer will randomly select one candidate to be the president of next unit. Citizens will see the new declaration from all candidates and know which one of them to be the president for the next unit. Attention: candidates will be assigned a new code in each unit. That is to say, citizens do not know which candidate has

been randomly selected as president in previous units, and which candidate has not been randomly selected as president in previous units.

Steps of the experiment in detail

1. At the beginning of the experiment, each participant will be randomly assigned to an 8-member-group. The three candidates receive their codes.

2. The following only happens at rounds 1, 6, 11, 16 and 21:

    a. The 3 candidates propose the back-transfer suggestions, as if they were elected to be the president. The back-transfer suggestion is how much he will back-transfer given citizens' transfer (0, 1, 2, 3, 4, 5 ,6 ,7 ,8 ,9 ,10) after he receives from citizens (0, 2, 4, 6, 8, 10, 12, 14, 16, 18, 20). These suggestions will be served as candidates' declaration.

    b. One candidate is randomly selected as president.

    c. All citizens will read the declaration from all candidates and know which candidate to be president for the following 5 rounds of the current unit.

    d. The candidates who are not randomly selected as president see the president's declaration. After one minute the program automatically enters a waiting screen for the start of next unit.

    e. At the end of each unit, the 3 candidates make a new declaration for the next unit. Back to step 2a).

3. In every decision round:

    a. Citizens make a decision how many points to send to the president. Citizens can send any integer they wish to the president from the set {0; 1; 2; 3; 4; 5; 6; 7; 8; 9; 10}.

    b. President sees the amount of points he gets from the 5 citizens, without knowing exactly who sent how much. The president will then decide upon the

amount of points to be back-transferred to each citizen. The president can back-transfer any integer from the set [0; 2 x sent amount] to citizens. The amount of points the president back-transfers to citizens will not be doubled.

c. Every participant receives information about his own payoff from the current round.

4. (Only at the end of rounds 5, 10, 15 and 20). All 3 candidates are assigned a new code, back to step 2.

5. (Only after round 25). Every participant receives information about her/his total payoff and receives payment.

How to calculate your income from the experiment?

Your total income from the experiment is the sum of income you receive from each decision round. The income of each round is calculated as follows:

Income of citizens in each round = Citizen's initial endowment (10 points) – the points sent to president + the points back-transferred by president

Income of candidates who are not randomly selected as president in each round = Candidate's initial endowment (10 points)

Income of president from each round = President's initial endowment (10 points) + the points sent by the 5 citizens x 2 – the points back-transferred to the 5 citizens

4 The Limited Power of Voting to Limit Power – a Stress-test in China and Germany[1]

4.1 Introduction

Power is a double-edged sword: it is an essential instrument to resolve collective-action problems; at the same time, it may be abused to the power holder's own advantage.[2] In many cases of organised human interaction, power is transferred through the means of elections. Elections may serve a disciplinary device by threatening dismissal from office (see e.g. Bardhan and Yang 2004) and for experimental evidence Weiß 2009, chapter 2 of this thesis).[3] Elections may also act as a means of selection, possibly leading to a choice of leader who is most capable (see for example Carrillo and Mariotti 2001), whose preferences are most in line with the constituency's own preferences (for a related model see Maskin and Tirole 2004) or who may even be motivated to serve the public (Cooter 2003; for a related model see Besley 2005). Elections, however, may matter beyond either reputational incentives or the choice of leader. Possibly, being elected by her constituency triggers the leader to feel a stronger sense of duty to act in her voters' interests. In this paper, we analyse whether the mere procedure of voting matters for the exercise of power.

Recent experimental evidence indeed shows holders of power to act in a less self-oriented way if they have been elected by their constituencies compared to being appointed by a random mechanism.[4] Corazzini, Kube and Marechál (2007) find elected allocators to send more to recipients than randomly drawn allocators if their approval rates are higher than what is minimally required to win the election. Weiß (2009) shows elected allocators to send back considerably more than randomly drawn allocators; this effect is found even in the last election period when re-election cannot motivate incumbents anymore. Walkowitz and Weiß (2009, chapter 3 of this thesis)[5] find a less pronounced but qualitatively same effect

[1] Based on: "The Limited Power of Voting to Limit Power – a Stress-test in China and Germany" by Hong Geng and Arne Robert Weiß (2009), Working Paper, University of Erfurt.
[2] This fundamental dilemma was already noted, forcefully, by John Locke (Locke and Laslett 1988).
[3] All subsequent references to Weiß (2009) also refer to chapter 2 of this thesis.
[4] In their experimental instructions, the holder of power is formulated either as "the winner" (Corazzini, Kube and Maréchal 2007) or as "the president" (in Weiß 2009 and in Walkowitz and Weiß 2009).
[5] All subsequent references to Walkowitz and Weiß (2009) also refer to chapter 3 of this thesis.

even if reliable reputation building is ruled out. In the existing experiments, not only the procedure of voting may explain the effects of voting, but also commitment to promises (Ellingsen and Johannesson 2004; Vanberg 2008), dynamic considerations or an interaction with investments. Therefore, they cannot provide a conclusive picture on whether the mere procedure of voting matters. In the first series of experiments reported in this paper, we rule these factors out and thereby provide a *stress-test* for the power of voting to limit the self-oriented exercise of power.

In our design, recipients elect one candidate to become – effectively – a dictator based on positively connoted, personal descriptions that candidates choose prior to the election. The elected dictator decides how to split an endowment between herself and the voters. In the control treatment, the dictator is selected by a random draw. The experiment is only played once, which is known to all subjects. Apart from voting, recipients do not make any further decisions.

Compared to Corazzini, Kube and Marechál (2007) as well as Walkowitz and Weiß (2009), we rule out any interaction with competition for votes and commitment to promises. In both Corazzini, Kube and Marechál (2007) as well as Walkowitz and Weiß (2009), candidates announce promises on how they will behave in case they win the election. Winning an election based on a promise sends a message from the voters of what they expect from the winning candidate, which matters to allocators who exhibit guilt-aversion (Charness and Dufwenberg 2006; Battigalli and Dufwenberg 2007). Furthermore, the voting procedure in Corazzini, Kube and Marechál (2007) triggers higher promises compared to their control treatment, in which the computer votes;[6] promises are more than cheap-talk for dictators who have a preference for either promise keeping (Vanberg 2008) or consistency (Ellingsen and Johannesson 2004) or who exhibit guilt-aversion and whose second-order beliefs rise with the promises they give. The less self-oriented behaviour of elected allocators in Corazzini, Kube and Marechál (2007) compared to their randomly drawn counterparts may therefore be explained by a joint effect of the procedure of voting and the higher promises induced by competition for votes. In our design, election takes place based on personal

[6] Compare tables 2 and table 3 in their appendix.

descriptions unrelated to the later task of the winning candidate; therefore, competition for votes cannot lead to differences in promises confounding the results.

Compared to Walkowitz and Weiß (2009), which is our experimental starting point, we also rule out any sort of dynamic considerations. Their experiment is based on a repeated investment game with five election periods. Reliable reputation building is ruled out in order to test for voting effects that are unrelated to strategic incentives; nevertheless, it remains unclear how strong the role of dynamic incentives is in explaining the voting effect on allocator behaviour in penultimate units. Even if reputational building is effectively ruled out, incumbents may still believe that their re-election probability is not independent of their transfer behaviour (in contrast to the control treatment in which it clearly is). As a consequence, they may strategically choose higher transfers in earlier rounds. As the one-shot play is common knowledge among the subjects, we rule out any sort of dynamic considerations and, consequently, any possible interaction between dynamic considerations and voting. The experimental design also rules out monetary reciprocity, which may play a role in Walkowitz and Weiß (2009). Put differently, if an elected dictator wants to reward her recipients in our design, she will do so because she is elected by the recipients but not because recipients invested in the first place.

The experiment was first conducted in Chengdu, China. The results show that voting has no effect on dictators' transfer decisions. In order to test whether the results are different if subjects are more used to formal voting as a mechanism to determine hierarchy, we re-ran the experiment in Erfurt, Germany. The German results show that voting does not have an effect on the dictators' choices either. The results of the stress-testing experiment therefore send a cautionary note as they imply that the power of voting to limit the self-oriented exercise of power shown in previous experiments may be context-specific and dependent on, possibly, an interaction with specific promises, monetary reciprocity or dynamic considerations. In order to test the role of promises, we ran a new experiment in Chengdu, China, with subjects recruited from the same subject pool as in the first experiment. We altered the treatments by implementing, as statements prior to the selection of the dictators, numerical promises on dictators' transfers instead of personal characteristics. In this case,

voting indeed matters: elected dictators transfer more to their recipients than dictators who are randomly drawn.

The paper is organized as follows: In section 4.2 we describe experimental design and procedure. Section 4.3 provides a game theoretic solution and our intra- and cross-cultural hypotheses. In section 4.4, we first present intra-cultural results and subsequently cross-cultural results. We discuss and put the experimental results into a wider context in section 4.5. Section 4.6 provides an outlook and first experimental results on a follow-up research question that is immediately derived from the discussion in section 4.5. Section 4.7 concludes our findings and provides an outlook for future research.

4.2 Experimental design and procedure

4.2.1 Basic experimental design

Each experimental group consists of five players. Two of them are candidates (CA) and the other three are recipients (R). In the selection stage, the candidates first choose a personal description that can be seen by all three recipients. In the experimental voting treatment (VOT), each R casts one vote for one candidate to become D. The candidate who receives at least two votes becomes D. In the subsequent allocation stage, D decides how to distribute 100 points among herself and the three R. D can choose any amount $s \in \{0,1,...,100\}$ to allocate to the recipients. The remaining of the 100 points (100 – s) is D's payoff. As payoff, each R gets exactly the amount s transferred by D.[7] The experiment is only played once. In the control treatment (RAN), the only difference arises at the selection stage, in which D is selected by a random draw. At the point in which R vote in VOT, the computer draws one CA to be D in RAN. **Fehler! Verweisquelle konnte nicht gefunden werden.** illustrates the experimental process. An experimental instruction is provided in the appendix.

[7] As each recipient receives s but the dictator pays only once, our design implies a multiplicator of 3 for the amount s transferred by the dictator. Efficiency, as measured by total payoffs of R and D, therefore increases with transfers s. We employ a multiplicator of 3 in order to make life cognitively easier for the dictators; they only have to think about the amount they are willing to allocate to the recipients and not also about how much each recipient will thereby receive.

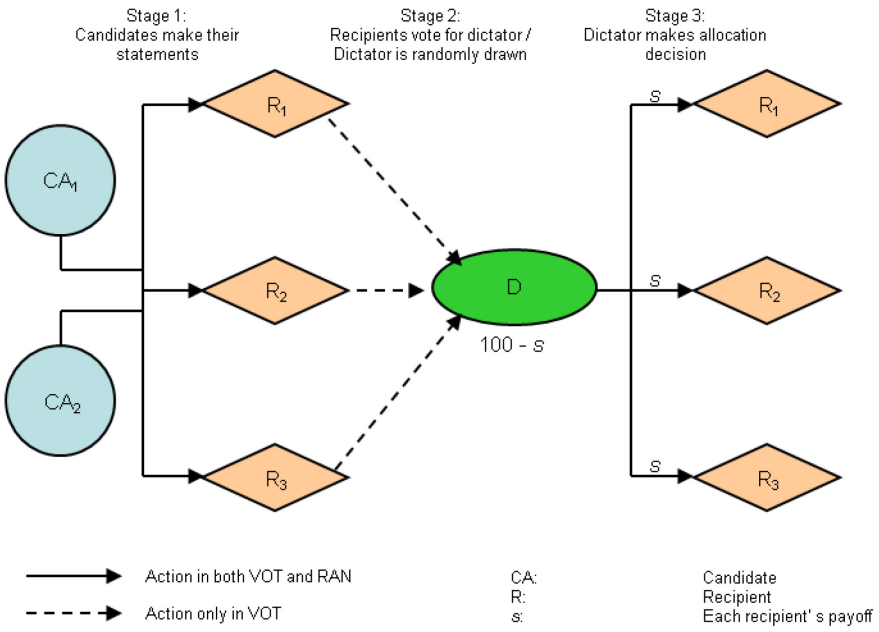


Figure 4-1: Experimental procedure

In the allocation stage, D effectively plays a dictator game between herself and the group of recipients as she has full discretionary power on how to split the endowment. Therefore, we speak of D as being a dictator.

4.2.2 Framed text and separated roles

As the power asymmetry is important for the research question, players' roles are framed with terms that create power distance: the dictator is framed as "president" and recipients as "citizens". Candidates are just called "candidates".[8] The unelected candidate does not turn into a recipient of the president because we want to concentrate on the relationship between the elected president and the voters (citizens): how will the president treat his constituency?

4.2.3 Statements of candidates

The challenge we face in the experimental design is to create a meaningful election but at the same time control for unwanted differences between the treatments. In order to test for the effects of voting, the voting procedure needs to be perceived differently than the

[8] In local language, we employed the following terms in the experiment (Chinese, German): president: 主席(zhǔxí), Bürgervorsitzender; citizen: 公民(gōngmín), Bürger; candidate: 候选人 (hòuxuǎnrén), Kandidat.

random draw by the dictators. At the same time, we want to rule out any possible selection of more group-oriented candidates as well as an interaction between competition for votes and commitment to promises. We therefore use statements that are both non-strategic, and purely positively connoted. By employing only positively connoted terms, we effectively rule out any meaningful differences in personal descriptions of the selected dictators between the treatments that may have unwanted, and immeasurable, repercussions for the dictator's allocation decision.[9]

Precisely, candidates are asked to choose from the following eight positively connoted terms: "optimistic", "erudite", "creative", "musical", "sportive", "lively", "diligent", and "fond of travelling".[10] Candidates select three of these eight terms and are asked to rank them according to a decreasing order of conformity to their personality. These ranked terms are candidates' statement prior to either the voting procedure or the random draw.

4.2.4 Payoff for unselected candidate

The payoff of the unselected candidate is chosen in a way as to minimise confounding effects on candidates' attitude towards the election or on dictators' decision. If the payoff for the unselected candidate is too high, candidates may not want to be elected as dictator. If the payoff for unselected candidate is a fixed number, the dictator may take it as a focal point for his transfer decision. Therefore the unselected candidate receives a random payoff from the same interval as the dictator's decision range {0,1,...,100}. The expected payoff for an unselected candidate is 50, which is also the equal-split solution, i.e. when the dictator sends half of the 100 points to his recipients.

4.2.5 Beliefs and hypothetical decisions

In order to gain further insight into the relevance of the voting procedure, we also collect hypothetical transfers from the unselected candidates as well as recipients' expectations on the dictator's transfer. In order to elicit hypothetical decisions as closely to actual decisions

[9] In RAN the selection of candidates is independent of their statements, while in VOT recipients are likely to vote for candidates they consider being trustworthy. If statements were meaningful for the later allocation decision, dictators in D may be described by those terms that imply high transfers to the recipients, whereas no such selection effect can arise in RAN.

[10] Small-scale surveys among Chinese and Germans, including subjects in a first trial session, were conducted in order to be confident that these terms do indeed have a positive connotation in both countries.

as possible but without actually employing the strategy method (Selten 1967), we let the unselected candidates state directly before announcing the election result how much they will transfer in case they have been elected as dictator.[11] The wording of the question is chosen in a way as to blur its hypothetical nature.[12] For the same reason, the timing for the elicitation of recipients' beliefs is after announcing the winner of the election or the random draw. As soon as they are informed about the result of the selection, recipients are asked to estimate how many points they will get from the selected dictator. In the German sessions, recipients afterwards have to make a second estimation on how many points the unselected candidate would have transferred if he had been selected as dictator.

4.2.6 Procedure

The Chinese sessions of the experiment were run in November 2007 at the Herbert A. Simon & Reinhard Selten behavioral decision research lab of the Southwest Jiaotong University in Chengdu, China. The German sessions of the experiment were run in January 2008 at the Laboratory for Experimental Economics (eLab) at the University of Erfurt. For each treatment, we collected 15 independent observations in China and 10 in Germany. Therefore, 150 students participated in the Chinese sessions and 100 students in the German sessions. Test questions made sure that subjects are aware of the one-shot play, the power asymmetry in their group as well as of how profits are calculated. The experiment started only when all subjects in the session correctly answered all test questions. Each experimental session lasted about one hour including instructions and payments. On average, Chinese students earned about 42.5 RMB (approximately 4 Euro), and German students earned about 11.5 Euros. The exchange rates between points and cash / local currency were set according to local standards.

Experimenter effects are a sensitive issue in running cross-cultural experiment (see for example (Roth et al. 1991). In order not to let subjects wonder about cross-cultural research

[11] We opted against employing the strategy method for eliciting decisions from both the selected and the un-selected candidate as we want to stay as closely as possible to a real-world voting context. Besides, as possible influences of the strategy method on the effects of voting procedures are untested, we want to refrain from discussing whether either observed or unobserved treatment differences may be an artefact of having employed the strategy method.

[12] The text unselected candidates see on their screen is "It will be announced soon which candidate has been elected / randomly drawn as president. Please insert how many points you will transfer to the citizens if you have been elected / drawn as president".

questions (if they faced a foreign experimenter), we kept the nationality of the experimenter constant to the subject pool. We also kept the gender of the experimenter constant across experimental sessions. One of the authors (Weiß) was present both in China and in Germany. The other author (Geng) was the experimenter in China. Another female German experimenter ran the German sessions. The experimental instructions were originally written in Chinese. The translation of instructions and computer screens into German was done applying the back-translation method.[13] The experiment was programmed in z-Tree (Fischbacher 2007).

4.3 Theoretic solutions and behavioural hypotheses

4.3.1 Payoffs and game-theoretic solution

Payoff functions for dictator (π_D), recipients (π_R) and the unselected candidate (π_L):

$$\pi_D = 100 - s$$

$$\pi_R = s$$

$$\pi_L = random\ number \in \{0,1,\ldots,100\} \text{ with } E(\pi_L) = 50$$

The unique game theoretic prediction based on money-maximising rationality is the dictator, regardless of whether he is elected in VOT or randomly drawn in RAN, keeping the entire 100 points for himself, i.e. s = 0. No recipients will therefore receive anything from the dictator.

4.3.2 Behavioural hypotheses

4.3.2.1 Intra-cultural hypotheses

In our design, only a direct behavioural response to being elected, rather than being randomly drawn, may constitute an effect of voting on power. By electing a dictator, considerable power is transferred. The dictator is completely unrestricted to decide on her own payoff as well on the payoff of the three recipients. If the dictator decides to send less

[13] See Brislin (1970) as well as Eco and McEwen (2000) for the method of back-translation.

than 50% of the 100 points, the transfer of power also comes with an expected positive material gift (as the un-elected candidate has an expected payoff of 50).

The elected dictator may consider the transfer of power as an act of trust and the implied potentially profitable position as a material gift. Previous research has shown that distrust can be self-fulfilling in that it lowers trustworthiness (Fehr and Rockenbach 2003; Falk and Kosfeld 2006). By the same idea, trust may be considered a kind act to be positively reciprocated (Falk and Fischbacher 2006). Gift-exchange, i.e. the reciprocation of a material gift with another gift, has been found in many experimental studies (see e.g. Fehr, Kirchsteiger and Riedl 1993; Fehr, Kirchsteiger and Riedl 1998; van der Heijden et al. 2001; Charness 2004), that may be explained by inequity aversion (Fehr and Schmidt 1999; Bolton and Ockenfels 2000) or by reciprocity utility (Falk and Fischbacher 2006). Hence, based on previous research on trust and gift-exchange, we expect dictators to show a reciprocal response to being elected. Within our design, reciprocity by a dictator means a transfer s that is higher than what the dictator would send in the absence of reciprocity considerations. As dictators have no reason to consider their selection in RAN as a gift from their recipients, we consequently predict:

H1: Transfers are higher in VOT than in RAN

The voting procedure may also affect recipients' expectations of the behaviour of the chosen dictator. Firstly, voters may expect elected dictators to behave reciprocal and group-oriented because of the voting procedure. Secondly, as the recipients in VOT decide who will be elected as the dictator, they may feel less vulnerable towards the dictator than recipients who have to be content with the outcome of a random draw. Research in social psychology on the "illusion of control" (Langer 1975) has shown that participatory procedures may turn people more confident about personal success in uncertain situations even when the objective probabilities of success have not changed. Translated into our design, recipients may trust more in the group-oriented behaviour of the dictator in VOT than in RAN merely because they are able to choose. We therefore predict:

H2: Recipients expect higher transfers in VOT than in RAN

4.3.2.2 Cross-cultural hypothesis

Democratic procedures are considerably more common in Germany than in China. In Germany, already school kids use ballot voting in order to determine hierarchy and to take decisions of collective importance. From about the age of 10, school kids vote for their class presidents and may use voting to decide on issues such as in which colour their class room may be painted. Later in their school life, pupils may vote directly for the pupil´s president, and their representatives are increasingly involved to vote on issues of collective importance. The typical mechanism is voting with a simple majority rule and universal suffrage. The outcomes of voting procedures are generally well respected. German recipients are typically awarded the right to vote in political elections at the age of 18 and in some local elections even at the age of 16. The electoral system in Germany is a mixture of direct and proportional elements. All German participants have likely participated in political elections besides elections in school and at university. In the World Bank´s Governance Indicators, Germany is consequently ranked in the top bracket in terms of "Voice and Accountability" with a percentile rank of 94.7 (Kaufmann, Kraay and Mastruzzi 2008). Furthermore, in the democracy index of World Audit, an international non-for-profit company providing global geopolitical perspectives, Germany is ranked in the top division. Countries in the top division score highest on "political rights" and "civil liberties" according to their rating by Freedom House.[14] In China, by contrast, direct elections only occur for village councils in designated rural areas and for the local people's congress. Young people rarely get a chance to participate in formal voting procedures. From school to university, class representatives are usually assigned by teachers. In few cases voting is used, but the candidates are nominated by teachers. Chinese pupils have therefore no experience with free and contested elections. According to the World Bank´s Governance Indicators, China consequently scores low on "Voice and Accountability": China is ranked in the bottom bracket with a percentile rank of 5.8 (Kaufmann, Kraay and Mastruzzi 2008). In the democracy index of World Audit, China is assigned to the bottom division. Although voting procedures are rarely used in political and organisational contexts, ballot voting is nevertheless known even to city dwellers; it is

[14] The ranking in the democracy index is foremost based on the rating on "political rights" and "civil liberties" by Freedom House. See www.worldaudit.org/democracy.htm for tables and http://www.worldaudit.org/methodology.htm for their methodology.

increasingly used for example in popular TV-shows. Here, however, the content of voting is questions of taste, rather than issues with conflicting interests, which are typical of political contexts. As the Chinese media also reports on elections taking place outside China, our Chinese subjects should be familiar with the voting mechanism we employ.

Because of large differences in personal experiences with democratic procedures we, nevertheless, expect Germans subjects to be more influenced by the voting procedure than Chinese subjects. We therefore predict

> H3: The difference in transfers between VOT and RAN (sVOT – sRAN) is larger in the German sessions than in the Chinese sessions

4.4 Results

4.4.1 Results of the Chinese sessions

Table 4-1: Descriptive summary statistics of Chinese sessions
(standard deviation in parentheses)

	C-VOT	C-RAN
Mean transfer of dictators	29.93	34.80
	(17.248)	(17.877)
Mean hypothetical transfer of unselected candidates	33.73	36.80
	(23.864)	(21.301)
Mean expected transfer of recipients	41.20	39.62
	(7.578)	(14.796)

Table 4-1 shows the summary statistics of VOT and RAN in the Chinese sessions. Surprisingly, both average values of transfers (29.93) and hypothetical transfers (33.73) are lower in VOT than in RAN (34.80 and 36.80 respectively). This already provides strong evidence against hypothesis H1, which predicts transfers to be higher in VOT than in RAN. Significance tests deliver evidence that actual transfers are indeed not significantly different between the two treatments.[15] The same holds true when we compare the hypothetical transfers of un-elected candidates in VOT to either the hypothetical transfers of unselected candidates in RAN or to the actual transfers in RAN: differences are statistically insignificant.[16] Even when

[15] Mann-Whitney U test, two-sided: p=0.624
[16] Mann-Whitney U test, two-sided: p=0.775 in the former case and p=0.766 in the latter case.

we merge dictators' actual transfers and unselected candidates' hypothetical transfers, still no significant difference emerges between VOT and RAN.[17] As significance tests do not reveal a significant difference in either actual or hypothetical transfers and as average values are lower in VOT than in RAN, hypothesis H1 needs to be strongly rejected for our Chinese sessions. The average beliefs of recipients are very similar in VOT (41.20) and RAN (39.62). A significance test cannot deliver any support of hypothesis H2, which predicts recipients to expect higher transfers in VOT than in RAN.[18] Hence, hypothesis H2 does not hold for the Chinese data either.

Result C-1: In the Chinese sessions, voting has no effect either on dictators' transfers or on recipients' expectations.

4.4.2 Results of the German sessions

Table 4-2: Descriptive summary statistics of German sessions
(standard deviation in parentheses)

	G-VOT	G-RAN
Mean transfer of dictators	38.30	49.70
	(13.873)	(24.221)
Mean hypothetical transfer of unselected candidates	40.40	36.80
	(12.518)	(17.479)
Mean expected transfer of recipients	35.03	40.53
	(8.972)	(4.11)
Mean of recipients' expected transfer of unselected candidates[19]	34.27	42.27
	(7.694)	(3.042)

Table 4-2 shows the summary statistics of VOT and RAN in the German sessions. The average transfer in RAN (49.79) is very close to the equal split of 50, while the average transfers in

[17] Mann-Whitney U test, two-sided: p=0.547. If by merging actual transfers and hypothetical transfers we still do not find significant differences, the evidence to reject H1 is even stronger. While the purist may not view actual and hypothetical transfers as statistically independent, we may still treat them as independent observations – in order to get further insight into the results – as there was no direct informational exchange between the candidates. Only the winning candidate in VOT is aware of being favoured by the voters but still has no information on the other candidate. Hence, while actual transfer decisions may, for a number of reasons, be different from hypothetical transfer decisions, candidates within one group should not influence each other. In fact, statistical tests reveal no difference between actual and hypothetical transfers, according to both two-sided Mann-Whitney U tests, p=0.845 for VOT and p=0.976 for RAN and two-sample Kolmogorow-Smirnow tests, p=0.999 for both VOT and RAN; there is also no significant difference in variances according to the Two-Sample Randomization Test for Differences in Variances: p=0.591 (two-sided) for VOT and p=0.731 (two-sided) for RAN.

[18] Mann-Whitney U test, two-sided: p=0.992

[19] Recall from section 2 that recipients' expectations on unselected candidates' hypothetical transfers were only elicited in the German sessions.

VOT (38.30) are 11.40 points lower. Significance tests, however, do not show a difference.[20] The hypothetical transfers of unselected candidates in VOT (40.40 on average) also do not differ both compared to the actual (49.70 on average) and the hypothetical transfers (36.80 on average) in RAN.[21] The merged data of actual and hypothetical transfers do not show a difference between VOT and RAN either.[22] We thus reject hypothesis H1 for our German sessions, too.

Result G-1: In the German sessions, voting has no effect on dictators' transfers.

Surprisingly and contrary to our hypothesis H2, German recipients' expectation on their dictators' transfer choices is weakly significantly higher in RAN (40.53 on average) than in VOT (35.03 on average).[23] Furthermore, recipients' expectations on the hypothetical transfers of the unselected candidates are also significantly higher in RAN (42.27 on average) than in VOT (34.27 on average).[24]

Result G-2: In the German sessions, recipients expect to get higher transfers in RAN than in VOT.

4.4.3 Cross-cultural results

Although the German transfers (see Table 4-2) are on average higher than the Chinese transfers (see Table 4-1) in both VOT and RAN, neither difference reaches a significant level.[25] The same holds true for the hypothetical transfers of the unselected candidates. The German unselected candidates would not behave differently than the Chinese unselected candidates, neither in VOT nor in RAN.[26] However, when we compare the merged data of actual and hypothetical transfers, we find that they are (weakly) significantly higher in German VOT than in Chinese VOT, whereas there is no significant difference between the

[20] Mann-Whitney U test, two-sided: p=0.424
[21] Mann-Whitney U test, two-sided: p=0.424 in the former case and p=0.565 in the latter case
[22] Mann-Whitney U test, two-sided: p=0.877. As in China (see footnote 10 on merging actual and hypothetical transfers), both types of transfers are not statistically different according to both two-sided Mann-Whitney U tests, p=0.811 for G-VOT and p=0.159 for G-RAN, and two-sample Kolmogorov-Smirnov tests, p=0.976 (G-VOT) and p=0.294 (G-RAN); there is also no significant difference in variances according to the Two-Sample Randomization Test for Differences in Variances: p=0.702 (two-sided) for VOT and p=0.461 (two-sided) for RAN.
[23] Mann-Whitney U test, two-sided: p=.069
[24] Mann-Whitney U test, two-sided: p=.007
[25] Mann-Whitney U test, two-sided: p=0.229 (VOT); p=0.115 (RAN)
[26] Mann-Whitney U test, two-sided: p=0.180 (VOT); p=0.946 (RAN)

two RAN treatments.[27] This finding suggests the German dictators may care more about the welfare of their group than the Chinese dictators if they are put into a voting environment. The reason for not directly observing cross-cultural differences in actual transfers, neither in VOT nor in RAN, may be due to the small number of observations (10 German observations vs. 15 Chinese observations).

> *Result CG-1: In VOT, German candidates tend to transfer more than Chinese candidates. In RAN, the transfers are not different between German and Chinese sessions.*

The average difference between the transfers in VOT and RAN is larger in the German sessions (11.40) than in the Chinese sessions (4.87), yet in another direction than stipulated in hypothesis H3: in both countries, transfers are higher in RAN than in VOT. Hence, dictators' transfers are not raised more due to the voting procedure in Germany than in China and hypothesis H3 needs to be rejected. Nevertheless, German dictators may still be more affected by the voting procedure, albeit in a different direction than hypothesised. In order to test whether the treatment effect between VOT and RAN is different in the two countries, we use the Monte-Carlo approximation of a two-sided permutation test with 50.000 draws. The result fails to reach a significant level.[28]

> *Result CG-2: The difference in transfers between VOT and RAN does not differ between the German and the Chinese sessions*

4.5 Discussion of the stress-testing experiment

Hence, based on our experimental setup, voting fails to limit the self-oriented exercise of power. Or put differently: the voting effect identified by Walkowitz and Weiß (2009) and (Corazzini, Kube and Maréchal 2007) do not survive the stress-test of our experiment. Based on our results, merely being elected instead of randomly drawn does not lead to power being used in a less self-oriented way. The implications of our results are therefore not in line with Walkowitz and Weiß (2009) and (Corazzini, Kube and Maréchal 2007). The tendency of our results even goes contrary to what has been previously observed. If voting

[27] Mann-Whitney U test, two-sided: p=0.072 (VOT); p=0.238 (RAN)
[28] p=0.560, two-sided

matters at all, elected dictators send less than randomly drawn dictators.[29] The power of voting to limit the self-oriented exercise of power therefore seems to be highly context-specific. And voting may even have perverse effects on group-oriented behaviour. Voting leads to considerable less morally hazardous behaviour by a powerful allocator in an investment context when voting is based on promises and reputation building is possible (Weiß 2009) and even when reliable reputation building is ruled out – albeit to a lesser extent (Walkowitz and Weiß 2009). Voting also leads to higher transfers when elections are based on promises and approval rates are higher than what is minimally required to win the election (Corazzini, Kube and Maréchal 2007). By contrast, when promises are absent and power is unconditional (as in our design) voting seems to have no capacity to limit power. Our experimental results therefore send a cautionary note not to put too much trust into voting or participatory mechanisms *by themselves* limiting power.

The question is: How pessimistic should we be about the power of voting to limit power? How may we reconcile the results from this experiment on the one hand with the results by (Corazzini, Kube and Maréchal 2007) as well as Walkowitz and Weiß (2009) on the other hand? In order to think about both questions and indicate directions for future research, we will, in the remainder of this section, evaluate features of our design and compare it to previous voting experiments. As our results do not yield significant differences between the treatments but a tendency towards elected dictators sending less than randomly drawn dictators, we will think in two directions. First, sticking to the basis of our hypotheses in section 4.3.2, we will consider aspects of our stress-test that may lead to elected dictators *not* behaving less self-oriented. We will then discuss whether a hitherto unconsidered behavioural mechanism, namely the creation of entitlements, may lead to voting causing power to be exercised in a *more* self-oriented way.

[29] In order to gain more insight in which direction the voting procedure we employed may affect behaviour of the dictators, we compare the merged transfers from China and Germany and test for treatment differences using both a Mann-Whitney U test and a less conservative t-test. We are aware of the limitations of this approach and will therefore treat the results with a grain-of-salt. Nevertheless, the results may give us a better picture of the experimental data. Both results are clearly insignificant: Mann-Whitney U, p=0.412, two-sided; t-test, p=0.172, two-sided and therefore can clearly not refute the Null-hypothesis of no difference between the treatments; nevertheless, employed as a research-compass, the results of the t-test seem to turn looking for reasons why elected dictators may behave in a *more* self-oriented way a worthwhile direction for a discussion of the experimental results and for future research.

In deriving the hypotheses of section 4.3.2 the candidates' perceptions of the voting procedure are critical. For the dictator to respond to the transfer of power and the potential material gift of being selected differently by being elected than by being randomly drawn, she has to view the voting procedure as something different than a random draw. If dictators consider being elected as a deliberate act of trust by the voters, we would expect at least some dictators to reciprocate this trust by sending more than a randomly drawn dictator. If, however, the dictator sees the voting mechanism as merely a substitute for some random mechanism, we would not expect him to behave any different by being elected than by being randomly drawn. We will separate this question into two elements: the perception of the voting mechanism itself and the role of intentionality.

4.5.1 Is the voting mechanism perceived as meaningful?

Candidates' statements, which voters base their decisions on, are unrelated to the later task of the elected candidate and leave little room for differences in taste. By contrast, in Walkowitz and Weiß (2009) candidates state a non-binding back-transfer strategy that is directly related to the decisions elected candidates would later take. We purposefully refrained from using such kind of information in order to avoid unwanted differences between the statements of elected and randomly drawn candidates.[30] Possibly, however, the differences between the treatments thereby become so subtle that elected dictators consider their electoral success as just as much an outcome of chance as their randomly drawn counterparts.

In order to investigate the significance of the method we employed let us look first at the selection of terms. Evidence for candidates perceiving the term selection procedure as meaningful would be candidates deliberately choosing instead of randomly picking terms to describe themselves. Note that the distribution of selected terms should be uniform if subjects choose them randomly. Looking at Figure 4-2 and Figure 4-3, the distributions of

[30] Recall that the voting procedure in Corazzini, Kube, Maréchal (2007) triggered higher promises compared to the control treatment. If some dictators are either guilt-averse or have a preference for promise keeping or consistency, higher promises in the voting treatment would lead to higher transfers. This effect would be unrelated to the voting procedure itself.

selected terms in VOT are clearly not uniformly distributed.[31] Even in the Chinese random treatment, where the chosen terms are irrelevant for the selection of the dictator, candidates did not randomly pick terms.[32] Only in G-RAN the distribution of terms is not distinguishable from being randomly selected.[33]

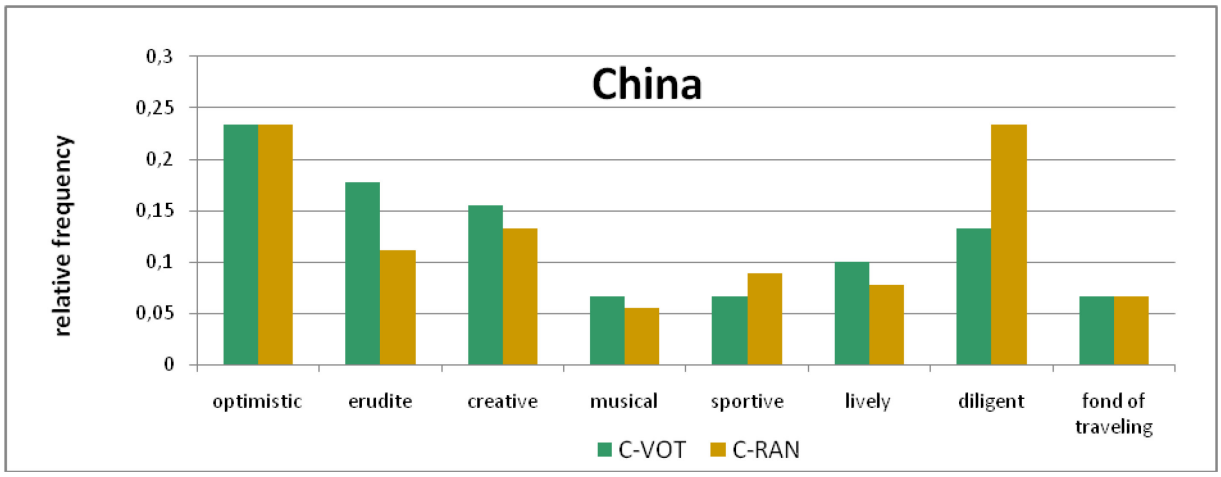


Figure 4-2: Distribution of chosen terms in Chinese sessions

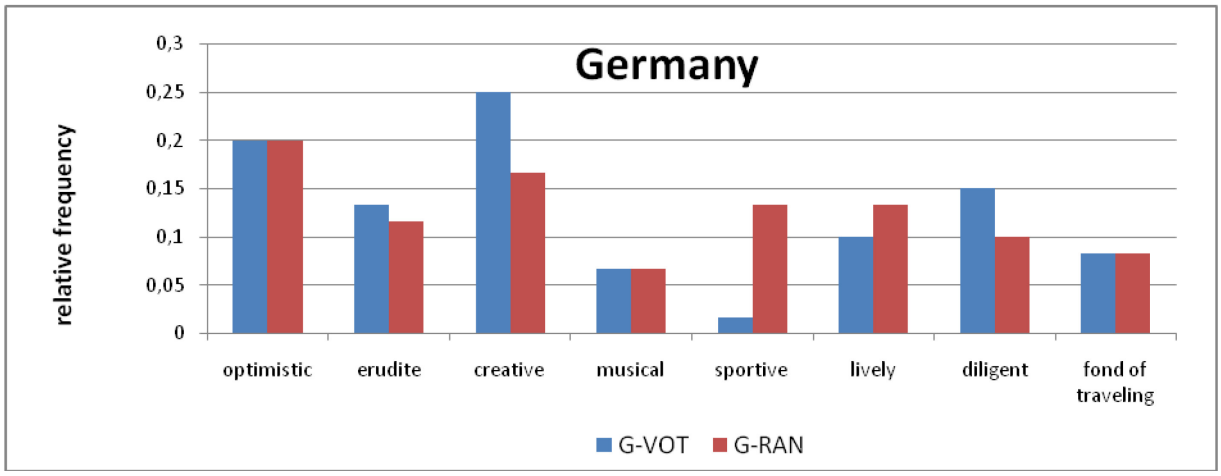


Figure 4-3: Distribution of chosen terms in German sessions

Dictators may perceive the voting procedure as meaningful if they attribute their electoral success as a result of a deliberate act by the voters based on the chosen terms of the candidates. While we have no direct access to dictators' perceptions, we can nevertheless

[31] We used two-sided Chi square Goodness-of-Fit test to test whether the distribution of chosen terms may be drawn from a uniform distribution of terms: p=0.008 (C-VOT), p=0.0084 (G-VOT)
[32] Chi square Goodness-of-Fit test, two-sided: p=0.0006
[33] Chi square Goodness-of-Fit test, two-sided: p=0.4939. Note that the test can nevertheless not show that candidates did not care about which terms to choose. The distribution of self-described characteristics may also be uniform.

look at the voting procedure from the perspective of the recipients – assuming that dictators' expectations of voters' behaviour are not systematically false.

First, did recipients have a real choice? In only one case (out of 50) the two candidates in one group chose the same three terms. But even in this case the ordering of terms was not the same. Recipients therefore always had a choice in the sense that the two options were not identical. Second, did recipients perceive the candidates´ statements as important for their voting decision? 87% of all German recipients affirmed this question when being asked after they took their voting decision.[34] As we have shown in section 4.5, recipients' voting decisions were in line with the transfers they expected from the respective candidates. Hence, recipients expected the personal statements of the candidates to be informative.

To conclude, the voting procedure we employed in VOT is certainly different from the random device in RAN. This is also reflected in the views of the subjects: significantly more subjects favoured the voting mechanism to a random mechanism in all treatments in both countries based on a hypothetical question in a questionnaire after subjects received feedback on their dictator's decision.[35] As recipients were neither better off in VOT than in RAN nor expected to receive more in VOT than in RAN, this seems to be in line with the notion of procedural utility (see e.g. Frey, Benz and Stutzer 2004): recipients seem to have a taste for having a say. Hence, the voting procedure worked differently and was perceived in a different way than the random mechanism; nevertheless, we cannot rule out that the differences between the mechanisms in VOT and RAN were *insufficient* to lead to less self- -oriented behaviour of dictators in VOT than in RAN.

4.5.2 Is intentionality critical?

Elected dictators, even if they feel to be elected purposefully, may not view the transfer of power as an intentional act of trust. Intentions have received considerable attention in the study of reciprocity (see e.g. Falk and Fischbacher 2006). A common view seems to emerge that sees intentions as at least enhancing reciprocity (see e.g. McCabe, Rigdon and Smith 2003; Charness 2004). The basic experimental structure we employed is based on the

[34] Chinese citizens were not asked about their perception of the candidates' statements.
[35] Chi square Goodness-of-Fit test, two-sided: $p<0.0001$ (for C-VOT, C-RAN and G-VOT), $p=0.0339$ (G-RAN)

power-distribution of a dictator game. The advantage of a dictator game for our research question may at the same time limit the power of voting: a dictator's power is unconditional. In our design, the recipients have no choice but to transfer power to one of the two candidates. Even if they deliberately choose one candidate over the other, they have no choice but to put their material well-being in the hands of the elected candidate. In other words: the transfer of power itself is not an act of choice. By contrast, in Walkowitz and Weiß (2009) voters made two choices: they transferred power by giving a candidate their votes and they decided through their investments to which degree they would trust the elected candidate. Possibly, for voting to be reciprocated, the intentional element of trust needs to be more visible.

4.5.3 Does voting lead to greater entitlements for elected dictators?

It has been shown that giving-behaviour in dictator game reacts sensitively to how entitled the dictator feels to keeping the amount to be distributed between herself and the recipient. Dictators give considerably less if they feel they deserve to have received the endowment, e.g. by first passing a test (Cherry, Frykblom and Shogren 2002). Even subtle forms of entitlements seem to matter: in Bolle and Vogel (2007) provision allocations (by the experimenter) between the dictator and the recipient influence the final allocation of the dictator as long as the provisional allocation is privately and socially acceptable. Possibly, also the voting procedure we employed may create entitlements.

Candidates in VOT may consider the voting procedures as a contest between the candidates. Winning this contest may be attributed to being the "better" candidate in two possible dimensions: dictators who described themselves truthfully may consider their personality more appreciated by the voters while those dictators choosing terms strategically may consider themselves better in guessing voters' preferences.[36] As either the more appreciated or smarter candidate, the dictator in VOT may feel entitled to a monetary prize. In RAN, by contrast, the selection of the dictator is unrelated to their personal characteristics and only a matter of chance. Dictators in RAN may therefore feel less entitled to a reward. Consequently, the stronger sense of entitlement in VOT compared to RAN would lead to

[36] The game played by strategic candidates may be seen as a sort of beauty contest.

dictators, contrary to our hypothesis H1, sending less if they are elected than randomly drawn.

Entitlements have, to the best of our knowledge, hitherto not been explicitly considered in voting experiments. Entitlements may indeed matter more in our design than in previous voting studies. In Walkowitz and Weiß (2009) and Corazzini, Kube, Marechál (2007) allocators are elected based on specific promises on their behaviour in the position of the allocator. No personal information is given about the candidates. Hence, elected allocators are likely to be aware that their electoral success was due to their promises and not to them being the "better" candidate for intrinsic reasons. Furthermore, in Corazzini, Kube, Marechál (2007), elected allocators know before taking their transfer decisions that first and second-order beliefs on the transferred amounts would be elicited.[37] Thereby, the transfer expectations of the voters may become salient both in attributing electoral success and in triggering guilt-aversion or commitment to promises. In Walkowitz and Weiß (2009), the separate and repeated investment decisions may also make clear that trust is conditional. Electoral success would therefore not justify exploiting the position of the allocator.

4.6 Testing the role of promises

In order to test the role of promises in explaining differences between our study on the one hand and Walkowitz and Weiß (2009) and Corazzini, Kube, Marechál (2007) on the other hand, we ran two new treatments. These were run in January 2009 also at the Herbert A. Simon & Reinhard Selten behavioral decision research lab of the Southwest Jiaotong University in Chengdu, China.[38] We altered the treatments VOT and RAN by implementing numerical promises instead of personal characteristics as statements prior to the selection of the dictators. 20 independent observations were collected for both treatments, which we call VOT-P and RAN-P. We also elicited second order beliefs of the selected dictators in order to test for guilt-aversion as a possible driver of treatment differences.[39] We elicited second order beliefs after dictators chose how many points to transfer to their group; in this way,

[37] Compare the instruction in Corazzini et al. (2007)

[38] We thank Peng Cheng for running the experiment for us. The experiment will be presented and analysed in more depth in his diploma thesis, supervised by Prof. Dr. Armin Falk of the University of Bonn.

[39] The text dictators see on their screen is "The three citizens in your group are estimating how many points you will allocate to them. Please estimate the average points of the estimation of the three citizens."

numerical promises instead of personal characteristics are the only difference between the new treatments VOT-P and RAN-P and the original treatments VOT and RAN.

Table 4-3: Descriptive summary statistics of treatments with numerical promises (standard deviation in parentheses)

	VOT-P	RAN-P
Mean transfer of dictators	34.00	25.30
	(17.592)	(20.303)
Mean hypothetical transfer of unselected candidates	37.55	29.35
	(21.132)	(18.004)
Mean promises of dictators	52.25	49.75
	(7.887)	(13.396)
Mean expected transfers of recipients	40.99	34.13
	(9.954)	(10.095)
Mean second order belief of dictators	40.75	36.70
	(11.616)	(16.658)

Table 4-3 summarizes the results of the new treatments. We find actual transfers to be weakly significantly higher in VOT-P (34.00 on average) than in RAN-P (25.30 on average);[40] a joint comparison of actual and hypothetical transfers even imply significantly higher transfer decisions in VOT-P compared to RAN-P.[41]

Result CP-1: Dictators' transfers are weakly significantly higher in VOT-P than in RAN-P.

The average expected transfers of recipients are, on average, higher in VOT-P (40.99) than in RAN-P (34.13). The statistical test rejects the null hypothesis that recipients have the same expectations in the two treatments at a weakly significant level.[42]

Result CP-2: Recipients' expectations on the transfers are weakly significantly higher in VOT-P than in RAN-P.

[40] Mann-Whitney U test, two-sided: p=0.098

[41] Mann-Whitney U test, two-sided: p=0.050. As in previous sections (see footnote 10 and 16 on merging actual and hypothetical transfers), both two-sided Mann-Whitney U tests (p=0.738 for VOT-P and p=0.341 for RAN-P) and Kolmogorov-Smirnov tests (p=0.933 for VOT-P and p=0.699 for RAN-P) do not reveal significant difference between actual and hypothetical transfer decisions in both VOT-P and RAN-P treatments. There is also no significant difference in variances according to the two-sample Randomization test for differences in variances: p=0.682 (two-sided) for VOT-P and p=0.776 (two-sided) for RAN-P.

[42] Mann-Whitney U test, two-sided: p=0.081

Second order beliefs of dictators and promises are strongly and highly significantly correlated in VOT-P but only marginally significantly in RAN-P.[43] Transfers are weakly significantly correlated with second order beliefs of dictators in VOT-P but again only at marginal significance in RAN-P.[44] Surprisingly, elected dictators did not give significantly higher promises than randomly selected dictators (on average 52.25 in VOT-P vs. 49.75 in RAN-P);[45] second order beliefs of dictators are only somewhat higher in VOT-P than in RAN-P (on average 40.75 in VOT-P vs. 36.70 in RAN-P).[46] Interestingly, transfers are not positively correlated with promises in RAN-P, while they are, at a weakly significant level, in VOT-P.[47] Although average values of commitment to promises are considerably different (on average, actual transfers fall short of promises by 18.25 in VOT-P and 24.45 in RAN-P), tests reveal no significant difference between the treatments.[48]

4.7 Conclusions and outlook

The experimental results of the stress-testing experiment using personal descriptions as statements not only fail to support our basic hypotheses of voting limiting the self-oriented exercise of power; they even provide astonishingly strong evidence to reject them: in both countries, China and Germany, average transfers tend to be higher in the random treatment than in voting treatment – albeit not at any reasonable significance level. Hence, if there is any non-accidental treatment-difference in transfers at all, elected dictators exploit their power in a *more* self-oriented way than randomly drawn dictators. Furthermore, in our German sessions, recipients expect to receive more from randomly drawn dictators than from their elected counterparts. Albeit not significant, the tendency of differences in expected transfers goes into the same direction in our Chinese sessions.

In a follow-up experiment using explicit promises as statements prior to voting, we do find elected dictators transferring more to their group than randomly drawn dictators. Hence, the promises seem critical. The significant differences in transfers also show that an intentional transfer of power is not necessary for voting to affect behaviour. The results are

[43] VOT-P: Spearman's rho = 0.662, p=0.001 (two-tailed); RAN-P: Spearman's rho = 0.372, p=0.106 (two-tailed)
[44] VOT-P: Spearman's rho = 0.408, p=0.074 (two-tailed); RAN-P: Spearman's rho = 0.374, p=0.104 (two-tailed)
[45] Mann-Whitney U, two-sided: p=0.297
[46] Mann-Whitney U, two-sided: p=0.158
[47] VOT-P: Spearman's rho = 0.404, p=0.078 (two-tailed); RAN-P: Spearman's rho = -0.072, p=0.763 (two-tailed)
[48] Mann-Whitney U, two-sided: p=0.427

partly consistent with guilt-aversion: dictators in both treatments using promises seem to be motivated by guilt-aversion, more clearly so in the voting treatment than in the random treatment. As Corazzini, Kube, Marechál (2007) also argue, promise-based voting sends a clear message about voter expectations; consequently, second order beliefs of dictators are considerably stronger correlated with promises in the voting than in the random treatment, while in the latter transfers are not even positively correlated with promises. Nevertheless, guilt-aversion cannot be the whole story: the treatment-difference is higher in transfers than in second order beliefs of dictators, the latter difference not even reaching a significant level. Direct commitment to promises, as a mechanism independent of guilt-aversion, does not provide any explanatory power in order to account for the effects of voting; elected dictators are not more committed to their promises than their randomly drawn counterparts. Rather, at least the basis for our behavioural hypothesis seems to be vindicated: motives other than commitment to promises or guilt-aversion, such as reciprocity, may lend a role for voting procedures limiting the self-oriented exercise of power. More research is nevertheless needed to clearly attribute the effects of voting to behavioural mechanisms. More experimental data are also warranted to find out more about whether voting procedures create entitlements. Compared to the random draw, the voting procedure based on personal descriptions may create a stronger feeling of being entitled to power, counter-running the hypothesised reciprocity-motive. Possibly, depending on procedural details, voting can therefore lead to an either more or less self-oriented exercise of power.

4.8 Appendix: Experimental instructions

General instruction

Welcome to our experiment!

Please read through the instructions carefully. You are not allowed to communicate with other participants by any means during the experiment.

If you are not clear about the experiment, please read through the instructions once again. For any further questions please raise your hand and we will come to answer your questions individually.

Your payoff will be expressed in points. The amount of the points depends on the decisions made by you and the other participants. After the experiment, we will exchange the points into RMB/Euro according to the following exchange rate:

100 Points = 75 RMB / 18.75 Euro

Besides, each participant will receive 10 RMB / 4 Euro for participating in the experiment.

During the whole experiment, please make the decision on your own. Be sure not to communicate with other participants in any way, or else you have to be ruled out of the experiment.

All the data and answers will be analyzed anonymously. To ensure anonymity, you have been instructed to choose a code number. Please find your seat in the cabinet with the corresponding number and make your own decision during the experiment. We can match decisions only to code numbers, but not to persons.

4.8.1 Instruction for the treatment VOT

There are 25 participants in the experiment.

This Experiment has only one round!

At the beginning of the experiment, each participant is allocated randomly into a five person group. The decisions within each group are independent of the other groups, that is to say, your decision only influences your own group members. There are two types of players in each group: 3 participants are citizens and the other 2 are candidates. One of the 2 candidates will be elected by the 3 citizens to be the president.

The elected president decides on how to distribute 100 points among the citizens in his group and himself. He can arbitrarily distribute the 100 points between him and the citizens of his group. The president decides how many of the 100 points he will transfer to **each** citizen in his group. The payoff for the president is the difference between the 100 points and the amount he transferred to the citizens. The amount transferred to each citizen is identical, that is, each citizen receives the same amount from the president.

The citizens vote for the president in the following way:

The two candidates choose among 8 descriptive adjectives those 3 that best represent his personality and rank them according to how well the adjectives describe his personality. (The first one is the adjective that best explains his personality, the second one is the second suitable adjective for his personality and the third one represents the third adjective matching his personality.) These 3 ranked adjectives are the personality statements of the candidates.

The citizens see the personality statements of the 2 candidates and elect one of them to become the president. Each citizen has only one vote and the voting result is determined by

majority rule, that is, the candidate with at least 2 votes becomes the president. The elected president then makes the allocation decision as described above.

The Steps of the Experiment:

1. At the beginning of the experiment, each participant is randomly allocated to a five person group and receives the role either as a citizen or as a candidate.
2. Each candidate chooses 3 descriptive adjectives that best represent his personality among the 8 descriptive adjectives and rank them according to the conformity of the adjectives with his personality.
3. Citizens see the personality statement of the candidates and elect one of them to become the president.
4. The elected president makes the transfer decision.
5. Each group member is informed on his own payoff.
6. The experiment ends.

How to calculate your payoff in the experiment

1. Citizen: Citizen's payoff = amount transferred by the president.

   The payoff of each citizen is equal to the amount transferred by the president. The higher the amount is, the higher will be his payoff; the lower this amount is, the lower will be his payoff.

2. President: President's payoff = 100 – amount transferred to the citizens

   The payoff of the president is equal to the difference between 100 points and the amount he transferred to the citizens. The higher this amount is, the lower will be his payoff; the lower the amount is, the higher will be his payoff.

3. Payoff of the unelected candidate: The computer draws one number from the interval {0,1,…,100}. This number is the payoff of the unelected candidate.

4.8.2 Instruction for the treatment RAN

There are 25 participants in the experiment.

This Experiment has only one round!

At the beginning of the experiment, each participant is allocated randomly into a five person group. The decisions within each group are independent of the other groups, that is to say, your decision only influences your own group members. There are two types of players in each group: 3 participants are citizens and the other 2 are candidates. One of the 2 candidates will be randomly selected to be the president.

The randomly selected president decides on how to distribute 100 points among the citizens in his group and himself. He can arbitrarily distribute the 100 points between him and the citizens of his group. The president decides how many of the 100 points he will transfer to **each** citizen in his group. The payoff for the president is the difference between the 100 points and the amount he transferred to the citizens. The amount transferred to each citizen is identical, that is, each citizen receives the same amount from the president.

How to randomly select the president

Computer program select randomly one of the two candidates to be president. The two candidates choose among 8 descriptive adjectives those 3 that best represent his personality and rank them according to how well the adjectives describe his personality. (The first one is the adjective that best explains his personality, the second one is the second suitable adjective for his personality and the third one represents the third adjective matching his personality.) These 3 ranked adjectives are the personality statements of the candidates.

The citizens see the personality statements of the 2 candidates and know which one of them is randomly selected to be the president.

The Steps of the Experiment:

1. At the beginning of the experiment, each participant is randomly allocated to a five person group and receives the role either as a citizen or as a candidate.
2. Each candidate chooses 3 descriptive adjectives that best represent his personality among the 8 descriptive adjectives and rank them according to the conformity of the adjectives with his personality.
3. Citizens see the personality statement of the candidates and know which one of them is randomly selected to be the president.
4. The randomly selected president makes the transfer decision.
5. Each group member is informed on his own payoff.
6. The experiment ends.

How to calculate your payoff in the experiment

1. Citizen: Citizen's payoff = amount transferred by the president.

   The payoff of each citizen is equal to the amount transferred by the president. The higher the amount is, the higher will be his payoff; the lower this amount is, the lower will be his payoff.

2. President: President's payoff = 100 – amount transferred to the citizens

   The payoff of the president is equal to the difference between 100 points and the amount he transferred to the citizens. The higher this amount is, the lower will be his payoff; the lower the amount is, the higher will be his payoff.

3. Payoff of the unselected candidate: The computer draws one number from the interval {0,1,…,100}. This number is the payoff of the unselected candidate.

4.8.3 Instruction for the treatment VOT-P

There are 25 participants in the experiment.

This experiment has only one round!

At the beginning of the experiment, each participant is allocated randomly into a five person group. The decisions within each group are independent of the other groups, that is to say, your decision only influences your own group members. There are two types of players in each group: 3 participants are citizens and the other 2 are candidates. One of the 2 candidates will be elected by the citizens to be president.

The elected president decides on how to distribute 100 points among the citizens in his group and himself. He can arbitrarily distribute the 100 points between him and the citizens of his group. The president decides how many of the 100 points he will transfer to **each** citizen in his group. The payoff for the president is the difference between the 100 points and the amount he transferred to the citizens. The amount transferred to each citizen is identical, that is, each citizen receives the same amount from the president.

The citizens vote for the president in the following way:

The two candidates make promises about how many points they are going to distribute to the citizens if they win the election.

The citizens see the promises of the 2 candidates and elect one of them to become the president. Each citizen has only one vote and the voting result is determined by majority rule, that is, the candidate with at least 2 votes becomes the president. The elected president then makes the allocation decision as described above.

The steps of the experiment:

1. At the beginning of the experiment, each participant is allocated to a five person group at random and receives the role either as a citizen or as a candidate.
2. Each candidate makes a promise about how many points he will allocate to the citizens if he wins the election.
3. Each citizen sees the promises of the candidates and votes for one of them for presidency.
4. The elected president makes the transfer decision.
5. Each group member is informed of his own payoff.
6. The experiment ends.

How to calculate your payoff in the experiment?

1. Citizen: Citizen's payoff = amount transferred by the president.

The payoff of each citizen is equal to the amount transferred by the president. The higher the amount is, the higher will be his payoff; the lower this amount is, the lower will be his payoff.

2. President: President's payoff = 100 – amount transferred to the citizens

The payoff of the president is equal to the difference between 100 points and the amount he transferred to the citizens. The higher this amount is, the lower will be his payoff; the lower the amount is, the higher will be his payoff.

3. Payoff of the unelected candidate: The computer draws one number within the interval {0,1,…,100}. This number is the payoff of the unelected candidate.

4.8.4 Instruction for the treatment RAN-P

There are 25 participants in the experiment.

This experiment has only one round!

At the beginning of the experiment, each participant is allocated randomly into a five person group. The decisions within each group are independent of the other groups, that is to say, your decision only influences your own group members. There are two types of players in each group: 3 participants are citizens and the other 2 are candidates. One of the 2 candidates will be selected to be president randomly by computer.

The randomly selected president decides on how to distribute 100 points among the citizens in his group and himself. He can arbitrarily distribute the 100 points between him and the citizens of his group. The president decides how many of the 100 points he will transfer to **each** citizen in his group. The payoff for the president is the difference between the 100 points and the amount he transferred to the citizens. The amount transferred to each citizen is identical, that is, each citizen receives the same amount from the president.

How to randomly select the president

The two candidates make promises about how many points they are going to distribute to the citizens if they win the election. The citizens see the promises of the 2 candidates. Computer program selects randomly one of the two candidates to be president. The selected president then makes the allocation decision as described above.

The steps of the experiment:

1. At the beginning of the experiment, each participant is allocated to a five person group at random and receives the role either as a citizen or as a candidate.

2. Each candidate makes a promise about how many points he will allocate to the citizens if he becomes the president.
3. Each citizen sees the promises of the candidates and knows which one of them is randomly selected to be the president.
4. The selected president makes the transfer decision.
5. Each group member is informed of his own payoff.
6. The experiment ends.

How to calculate your payoff in the experiment?

1. Citizen: Citizen's payoff = amount transferred by the president.

The payoff of each citizen is equal to the amount transferred by the president. The higher the amount is, the higher will be his payoff; the lower this amount is, the lower will be his payoff.

2. President's payoff = 100 – amount transferred to the citizens

The payoff of the president is equal to the difference between 100 points and the amount he transferred to the citizens. The higher this amount is, the lower will be his payoff; the lower the amount is, the higher will be his payoff.

3. Payoff of the unselected candidate: The computer draws one number within the interval {0,1,…,100}. This number is the payoff of the unselected candidate.

References

Alvarez, M., J. A. Cheibub, F. Limongi and A. Przeworski (2000). Democracy and Development: Political Institutions and Material Well-being in the World, 1950-1990, Cambridge University Press.

Andreoni, J. and J. H. Miller (1993). "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma: Experimental Evidence." Economic Journal **103**(418): 570-85.

Arias, J. T. G. (1998). "A relationship marketing approach to guanxi." European Journal of Marketing **32**: 145-156.

Arrow, K. (1972). "Gifts and Exchanges." Philosophy and Public Affairs **1**(4): 343-362.

Bardhan, P. K. and T. T. Yang (2004). Political Competition in Economic Perspective, Berkeley. Dept. of Economics, University of California.

Barro, R. J. (1999). "Determinants of Democracy." Journal of Political Economy **107**(6): 158-183.

Battigalli, P. and M. Dufwenberg (2007). "Guilt in Games." American Economic Review **97**(2): 170-176.

Becker, G. S. (1983). "A Theory of Competition among Pressure Groups for Political Influence." Quarterly Journal of Economics **98**(3): 371-400.

Berg, J., J. Dickhaut and K. McCabe (1995). "Trust, Reciprocity, and Social History." Games and Economic Behavior **10**(1): 122-142.

Besley, T. (2005). "Political Selection." Journal of Economic Perspectives **19**(3): 43-60.

Besley, T. J. and M. Kudamatsu (2007). Making Autocracy Work, CEPR Discussion Paper.

Besley, T. J., T. Persson and D. M. Sturm (2006). Political Competition and Economic Performance: Theory and Evidence from the United States, CEPR Discussion Paper.

Bohnet, I. and S. Huck (2004). "Repetition and Reputation: Implications for Trust and Trustworthiness When Institutions Change." American Economic Review **94**(2): p362 - 366.

Bolle, F. and C. Vogel (2007). "Even Weak Entitlements Matter in the Dictator Experiment." Working Paper, University of Frankfurt (Oder).

Bolton, G. E., E. Katok and R. Zwick (1998). "Dictator game giving: Rules of fairness versus acts of kindness." International Journal of Game Theory **27**(2): 269-299.

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

Bornhorst, F., A. Ichino, O. Kirchkamp, K. Schlag and E. Winter (2004). How do People Play a Repeated Trust Game? Experimental Evidence. SFB 504 Discussion Paper No. 04-43

Brislin, R. W. (1970). "Back-Translation for Cross-Cultural Research." Journal of Cross-Cultural Psychology **1**(3): 185.

Brown, M., A. Falk and E. Fehr (2004). "Relational Contracts and the Nature of Market Interactions." Econometrica **72**(3): 747-780.

Brown, M., A. Falk and E. Fehr (2008). "Competition and Relational Contracts: The Role of Unemployment as a Disciplinary Device." Working Paper, Institute for the Study of Labor.

Carrillo, J. D. and T. Mariotti (2001). "Electoral competition and politician turnover." European Economic Review **45**(1): 1-25.

Charness, G. (2004). "Attribution and Reciprocity in an Experimental Labor Market." Journal of Labor Economics **22**(3): 665-688.

Charness, G. and M. Dufwenberg (2006). "Promises and Partnership." Econometrica **74**(6): 1579-1601.

Charness, G. and D. A. N. Levin (2005). "The Origin of the Winner's Curse: A Laboratory Study." Working Paper, University of California.

Chen, X. P. and C. C. Chen (2004). "On the Intricacies of the Chinese Guanxi: A Process Model of Guanxi Development." Asia Pacific Journal of Management **21**(3): 305-324.

Cherry, T. L., P. Frykblom and J. F. Shogren (2002). "Hardnose the Dictator." American Economic Review **92**(4): 1218-1221.

Clague, C., P. Keefer, S. Knack and M. Olson (1996). "Property and Contract Rights in Autocracies and Democracies." Journal of Economic Growth **1**(2): 243-276.

Clark, C. C., P. Murrell and S. A. Whiting (2006). "The Role of Law in China's Economic Development." Working Paper, George Washington University Law School.

Cooter, R. (2003). "Who Gets on Top in Democracy-Elections as Filters." Supreme Court Economic Review **10**: 127.

Corazzini, L., S. Kube and M. A. Maréchal (2007). "Towards Behavioral Public Choice: Guilt Aversion and Accountability in the Lab." Working Paper, Max Planck Institute for Research on Collective Goods.

Cramton, P. and J. A. Schwartz (2000). "Collusive bidding: Lessons from the FCC spectrum auctions." Journal of Regulatory Economics **17**(3): 229-252.

Dal Bó, P., A. Foster and L. Putterman (2008). Institutions and Behavior: Experimental Evidence on the Effects of Democracy, NBER Working Paper No. W13999.

Dawson, J. W. (2003). "Causality in the Freedom–Growth Relationship." European Journal of Political Economy **19**(3): 479-495.

Dixit, A. (2003). "Trade Expansion and Contract Enforcement." Journal of Political Economy **111**(6): 1293-1317.

Doucouliagos, H. and M. A. Ulubasoglu (2008). "Democracy and Economic Growth: A Meta-Analysis." American Journal of Political Science **52**(1): 61-83.

Eco, U. and A. McEwen (2000). Experiences in Translation, University of Toronto Press.

Ellingsen, T. and M. Johannesson (2004). "Promises, threats and fairness." The Economic Journal **114**: 397-420.

Falk, A. and U. Fischbacher (2006). "A theory of reciprocity." Games and Economic Behavior **54**(2): 293-315.

Falk, A. and M. Kosfeld (2006). "The Hidden Costs of Control." American Economic Review **96**(5): 1611-1630.

Farh, J., A. S. Tsui, K. R. Xin and B. Cheng (1998). "The Influence of Relational Demography and Guanxi: The Chinese Case." Organization Science **9**(4): 471-488.

Farr, W. K., R. A. Lord and J. L. Wolfenbarger (1998). "Economic Freedom, Political Freedom, and Economic Well-Being: A Causality Analysis." Cato Journal **18**(2): 247-262.

Fehr, E., G. Kirchsteiger and A. Riedl (1993). "Does Fairness Prevent Market Clearing? An Experimental Investigation." Quarterly Journal of Economics **108**: 437-437.

Fehr, E., G. Kirchsteiger and A. Riedl (1998). "Gift exchange and reciprocity in competitive experimental markets." European Economic Review **42**(1): 1-34.

Fehr, E., A. Klein and K. M. Schmidt (2007). "Fairness and Contract Design." Econometrica **75**(1): 121-154.

Fehr, E. and B. Rockenbach (2003). "Detrimental effects of sanctions on human altruism." Nature **422**: 137-140.

Fehr, E. and K. M. Schmidt (1999). "A Theory of Fairness, Competition, and Cooperation." Quarterly Journal of Economics **114**(3): 817-868.

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

Fischbacher, U., E. Fehr and S. Gächter (2001). "Are People Conditionally Cooperative? Evidence from a Public Goods Experiment." Economic Letters **71(3)**: 397-404.

Frey, B. S. (1994). "Direct Democracy: Politico-Economic Lessons from Swiss Experience." American Economic Review **84**: 338-338.

Frey, B. S. (1997). Not Just for the Money: An Economic Theory of Personal Motivation, Edward Elgar Publishing.

Frey, B. S., M. Benz and A. Stutzer (2004). "Introducing Procedural Utility: Not Only What, but Also How Matters." Journal of Institutional and Theoretical Economics **160**: 377-401.

Geng, H. and A. R. Weiß (2009). "The Limited Power of Voting to Limit Power – a Stress-Test in China and Germany." Working Paper, University of Erfurt.

Glaeser, E. L., R. La Porta, F. Lopez-de-Silanes and A. Shleifer (2004). "Do Institutions Cause Growth?" Journal of Economic Growth **9**(3): 271-303.

Granger, C. W. J. (1969). "Investigating Causal Relationships by Econometric Models and Cross-Spectral Methods." Econometrica **37**(3): 424-438.

Grossman, H. I. and S. J. Noh (1990). "A Theory of Kleptocracy with Probabilistic Survival and Reputation." Economics & Politics **2**(2): 157-171.

Grossman, H. I. and S. J. Noh (1994). "Proprietary Public Finance and Economic Welfare." Journal of Public Economics **53**(2): 187-204.

Gürerk, O., B. Irlenbusch and B. Rockenbach (2006). "The Competitive Advantage of Sanctioning Institutions." Science **312**(5770): 108-111.

Healy, P. J. (2007). "Group Reputations, Stereotypes, and Cooperation in a Repeated Labor Market." American Economic Review **97**(5): 1751-1773.

Hodess, R. and I. Transparency (2004). Global Corruption Report 2004: [Special Focus: Political Corruption]. London; Sterling, VA, Pluto Press/Transparency International.

Hoffman, E., K. McCabe and V. L. Smith (1996). "Social Distance and Other-Regarding Behavior in Dictator Games." American Economic Review **86**: 653-660.

Hofstede, G. (1984). Culture's Consequences: International Differences in Work-Related Values, Sage Publications.

Hofstede, G. (2001). Culture's Consequences: Comparing Values, Behaviors, Institutions, and Organizations Across Nations, Sage Publications.

Holt, C. A. (1995). "Industrial Organization: A Survey of Laboratory Research." The Handbook of Experimental Economics: 349-443.

Huck, S., G. K. Lünser and J.-R. Tyran (2006). Competition Fosters Trust, C.E.P.R. Discussion Papers.

Kaufmann, D., A. Kraay and M. Mastruzzi (2008). "Governance Matters Vii: Aggregate and Individual Governance Indicators 1996-2007." Working Paper, World Bank Policy Research.

King, A. Y. (1991). "Kuan-hsi and Network Building: A Sociological Interpretation." Daedalus: 63-84.

Kreps, D., P. Milgrom, J. Roberts and R. Wilson (1982). "Rational Cooperation in the Finitely Repeated Prisoners' Dilemma." Journal of Economic Theory **27**(2): 245-52.

Langer, E. J. (1975). "The illusion of control." Journal of Personality and Social Psychology **32**(2): 311-328.

Li, J. S. (2003). "Relation-based versus Rule-based Governance: an Explanation of the East Asian Miracle and Asian Crisis." Review of International Economics **11**(4): 651-673.

Li, S. (2002). "Does east love guanxi more than west? The evolution of relation-based governance: Contemporary and historical evidences." Global Economic Review **31**(1): 1-11.

Lipset, S. M. (1960). Political Man, Heinemann.

Locke, J. and P. Laslett (1988). Two treatises of government, Cambridge Univ Press.

Luo, Y. and M. Chen (1997). "Does Guanxi Influence Firm Performance?" Asia Pacific Journal of Management **14**(1): 1-16.

Maier-Rigaud, F. P. and J. Apesteguia (2003). "The Role of Choice in Social Dilemma Experiments." Working Paper, Max Planck Institute for Research on Collective Goods.

Maskin, E. and J. Tirole (2004). "The Politician and the Judge: Accountability in Government." American Economic Review **94**(4): 1034-1054.

McCabe, K. A., M. L. Rigdon and V. L. Smith (2003). "Positive reciprocity and intentions in trust games." Journal of Economic Behavior and Organization **52**(2): 267-275.

McGuire, M. C. and M. Olson (1996). "The Economics of Autocracy and Majority Rule." Journal of Economic Literature **34**: 72-96.

McMillan, J. and B. Naughton (1996). Reforming Asian Socialism, University of Michigan Press Ann Arbor.

North, D. and B. Weingast (1989). "Constitutions and Commitment: The Evolution of Institutions Governing Public Choice in 17th Century England." Journal of Economic History **49**(4): 803-832.

North, D. C. (1990). Institutions, Institutional Change, and Economic Performance, Cambridge University Press.

North, D. C. and R. P. Thomas (1973). The Rise of the Western World: A New Economic history, Cambridge University Press.

Olson, M. (1993). "Democracy, Dictatorship, and Development." American Political Science Review **87**(3): 567-75.

Ostrom, E., J. Walker and R. Gardner (1992). "Covenants With and Without A Sword: Self-Governance is Possible." American Political Science Review **86**(2): 404-417.

Persson, T., G. Roland and G. Tabellini (1997). "Separation of Powers and Political Accountability*." Quarterly Journal of Economics **112**(4): 1163-1202.

Pinto, P. M. and J. F. Timmons (2005). "The Political Determinants of Economic Performance: Political Competition and the Sources of Growth." Comparative Political Studies **38**(1): 26.

Przeworski, A. and F. Limongi (1993). "Political Regimes and Economic Growth." Journal of Economic Perspectives **7**(3): 51-69.

Rapoport, A. (1997). "Order of play in strategically equivalent games in extensive form." International Journal of Game Theory **26**(1): 113-136.

Ray, D. (1998). Development Economics, Princeton University Press.

Roth, A. E., V. Prasnikar, M. Okuno-Fujiwara and S. Zamir (1991). "Bargaining and Market Behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An Experimental Study." American Economic Review **81**(5): 1068-1095.

Selten, R. (1967). "Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eines Oligopolexperiments", in Sauermann. H. "Beiträge zur exprimentellen Wirtschaftsforschung", JCB Mohr (Paul Siebeck), Tübingen: 136-168.

Sen, A. (1995). "Rationality and Social Choice." American Economic Review **85**(1): 1-24.

Solinger, D. J. (1989). "Urban reform and relational contracting in post-Mao China: An interpretation of the transition from plan to market." Studies in Comparative Communism **22**(2/3): 171-185.

Thibaut, J. W. and L. Walker (1975). Procedural justice, L. Erlbaum Associates.

Tyler, T. R. (1989). "The psychology of procedural justice: a test of the group-value model." Journal of Personality and Social Psychology **57**(5): 830-838.

Tyler, T. R. (2006). Why People Obey the Law, Princeton University Press.

Tyler, T. R. and S. L. Blader (2003). "The Group Engagement Model: Procedural Justice, Social Identity, and Cooperative Behavior." Personality and Social Psychology Review **7**(4): 349.

Tyler, T. R. and E. A. Lind (1992). "A relational model of authority in groups." Advances in Experimental Social Psychology **25**: 115-191.

Tyran, J. R. and L. P. Feld (2006). "Achieving Compliance when Legal Sanctions are Non-deterrent." The Scandinavian Journal of Economics **108**(1): 135-156.

UNODC and WorldBank (2007). Stolen Asset Recovery (StAR) Initiative: Challenges, Opportunities, and Action Plan, United Nations Office on Drugs and Crime, The World Bank.

van der Heijden, E. C. M., J. H. M. Nelissen, J. J. M. Potters and H. A. A. Verbon (2001). "Simple and complex gift exchange in the laboratory." Economic Inquiry **39**(2): 280-297.

Vanberg, C. (2008). "Why Do People Keep Their Promises? An Experimental Test of Two Explanations." Econometrica **76**(6): 1467-1480.

Vega-Gordillo, M. and J. L. Alvarez-Arce (2003). "Economic Growth and Freedom: A Causality Study." The Cato Journal **23**(2): 199-216.

Walkowitz, G. and A. R. Weiß (2009). "Does Voting Moderate Power and Foster Trust? A Behavioural View on Democracy." Working Paper, University of Erfurt.

Weber, M. (1947). "The theory of economic and social organization." Trans. AM Henderson and Talcott Parsons, Oxford University Press.

Weber, M. (1985). "Wirtschaft und Gesellschaft (1922)." Fünfte Auflage, Tübingen.

Weber, R. A., C. F. Camerer and M. Knez (2004). "Timing and Virtual Observability in Ultimatum Bargaining and "Weak Link" Coordination Games." Experimental Economics **7**(1): 25-48.

Weiß, A. R. (2009). "The Economic Virtues of Voting - How Political Competition Limits Confiscatory Behaviour." Working Paper, University of Erfurt.

Wittman, D. (1989). "Why Democracies Produce Efficient Results." The Journal of Political Economy **97**(6): 1395.

Wu, S. Y. and B. Roe (2007a). "Discretionary Latitude and Relational Contracting." IZA Discussion Paper No. 2879.

Wu, S. Y. and B. Roe (2007b). "Contract Enforcement, Social Efficiency, and Distribution: Some Experimental Evidence." American Journal of Agricultural Economics **89**(2): 243-258.

Xin, K. R. and J. L. Pearce (1996). "Guanxi: Connections as Substitutes for Formal Institutional Support." Academy of Management Journal **39**: 1641-1658.

Yeung, I. Y. M. and R. L. Tung (1996). "Achieving business success in Confucian societies: The importance of guanxi (connections)." Organizational Dynamics **25**(2): 54-65.

Curriculum Vitae

Arne Robert Weiß

Address
University of Erfurt
Microeconomics
Nordhäuser Str. 63
99089 Erfurt
Deutschland
Fax: +49 (0) 361 737 4529
e-mail: arne.weiss@uni-erfurt.de

Current Position

Since 10/2010	Research Assistant at the University of Erfurt

Education & Previous Positions

06/2009 – 08/2010	Research Associate of the Harvard Business School / consultant with the National Bureau of Economic Research, working as a research director of a field experiment in Zambia
01/2007-05/2009	Research Assistant at the University of Erfurt
10/2002-04/2005	Economics at the University of St. Gallen (“Lic. Oec. HSG”) Diploma-Thesis: „O-Ring Theorie und Effizienzlöhne“ (O-Ring Theory and Efficiency Wages), Supervisors: Prof. Dr. Simon Gächter, Prof. Dr. Jean-Robert Tyran
11/2003 – 10/2004	Assistant of Prof. Dr. Jörg Baumberger, University of St. Gallen
09/2000 – 06/2002	Economics at the London School of Economics (“Diploma in Economics”)
10/1998 – 08/2000	Economics and Business Administration at University Witten/Herdecke